Proceedings of

The International Conference on Artificial Intelligence Management and Trends

ICAIMT

May 21-22, 2025

Abu Dhabi School of Management
Abu Dhabi
United Arab Emirates

Proceedings of

The International Conference on Artificial Intelligence Management and Trends

May 21-22, 2025

Abu Dhabi School of Management, Abu Dhabi, UAE

Organizing Committee:

Dr. Neda Abdelhamid (Chair)
Chair of MSBA Program
Associate Professor of Business Analytics

Dr. Evi Indriasari Mansor
Associate Professor of Business Analytics
Abu Dhabi School of Management

Dr. Constance Van Horne
Director of International Affairs and Executive Development & Associate Professor of Management
Abu Dhabi School of Management

Editorial Board

The editorial board members for this conference are:

Preface

The International Conference on Artificial Intelligence Management and Trends (ICAIMT 2025) was held on May 21–22, 2025, at the Abu Dhabi School of Management in the capital of the United Arab Emirates. As a leading international forum, the conference brought together academics, researchers, and industry professionals to exchange knowledge and discuss the rapidly evolving role of artificial intelligence (AI) in transforming organizational strategies and operational models.

Held under the theme "Shaping the Future: Strategic AI Integration Across Industries," ICAIMT 2025 welcomed more than 100 participants from over 18 countries including the UAE, UK, Qatar, Oman, Jordan, Italy, Lebanon, Bahrain, Kazakhstan, and Russia. The event featured a diverse program of keynote speeches by academic and industry leaders, research paper presentations, expert panel discussions, technology showcases, and hands-on workshops. These sessions explored how AI is being used to support innovation and informed decision-making across critical sectors such as healthcare, finance, education, logistics, and management.

Key topics addressed included AI- automation, ethical and responsible AI, data governance, intelligent systems in education, and strategic AI deployment. Discussions emphasized the importance of aligning AI technologies with institutional objectives, regulatory compliance, and broader societal values. ICAIMT 2025 also encouraged interdisciplinary collaboration by bridging gaps between theoretical research and applied AI solutions. Attendees engaged with real-world use cases, frameworks, and tools that demonstrated AI's potential to enhance performance, optimize resources, and unlock new growth avenues for organizations.

The ICAIMT 2025 received a large number of submissions. Each paper underwent a rigorous peer-review process by at least two reviewers, evaluating criteria such as originality, thematic relevance, technical soundness, and contribution to the field. Accepted papers were presented during the conference sessions, facilitating scholarly exchange and enabling potential research collaborations. All accepted papers are compiled in this proceedings volume, published by ADSM with an assigned ISSN and individual DOIs to ensure global visibility and citation accessibility.

We hope that the work presented in these proceedings contributes meaningfully to the academic and professional AI community. We also look forward to seeing continued research, strategic initiatives, and cross-sector innovations inspired by the discussions and partnerships formed during ICAIMT 2025.

Dr. Neda Abdelhamid (ICAIMT Chair)

Table of Contents

Title **Page**

Incorporation Of Generative AI Into The Assessment Of A Practical Advertising Production University Course – A Case Study

Philip Dennett
London Southbank University – UAE
Ajman, United Arab Emirates
philip.dennett@lsbuac.ae

Abstract— Students are increasingly using generative AI (artificial intelligence) tools in their university assignments, often running afoul of academic integrity rules. The purpose of this case study was to reduce the risk of academic integrity breaches and integrate AI into elements of an advertising communication course. The changes made included practice and guidance on generative AI models, AI elements incorporated into a major assignment, and assessable personal reflections from students. The methods used were the development and presentation of an assessment brief and associated rubric incorporating the use of AI tools. Preliminary results suggest students have a high level of comfort in the use of generative AI tools and a clear understanding of its place in academic assessments.

Keywords—generative AI, academic integrity, advertising course.

I. INTRODUCTION

Following the COVID-19 pandemic there has been several changes in education and learning, including widespread use of online learning and the use of generative AI tools by both academics and students[1]. The Australian government's Tertiary Education Quality and Standards Agency (TEQSA) warns that the use of generative AI is generally not detectable and that it can produce responses for course-specific assessments. This has resulted in higher incidences of academic integrity breaches[2].

Given both the widespread use and difficulty of detecting AI content, this research seeks to identify a way in which the use of AI tools can be incorporated into a written assessment to overcome these issues. The scope of the research is to redesign an existing assessment in an undergraduate advertising course to incorporate the use of AI tools and revise the outputs to clearly identify the intellectual efforts of students. This will be of value to other academics who teach courses where a creative output is required.

II. BACKGROUND

Despite a growing body of literature, significant research gaps exist in understanding the impact of AI in university settings[3]. Generative AI is a deep learning model (DLM) designed to create original content, including written and visual material, in response to a user's request[4]. The availability of DLM's such as ChatGPT have given rise to questions such as the future role of teachers and the design of assessments[5]. Ghanimi and colleagues suggest that rather than focus on the negatives we should separate data gathering tasks (which AI can assist with) and critical thinking and creative tasks (which prove the students understanding of the course material. By working in tandem in this way assessments can become programmatic rather than summative allowing students multiple opportunities to demonstrate their competence[6].

Lodge, and colleagues[7] advise that assessments in the age of AI should emphasise appropriate engagement with AI tools, integrated into the learning process. For example, framing the use of AI tools as a learning competency, demonstrated through student reflections[8]. The goal should be to integrate AI into the learning process in a way that reinforces academic integrity and cultivates essential digital literacy skills[9].

Taking this advice, I decided to review the principal assessment in a course called Production: Creative Advertising that is aimed at second- and third-year undergraduates studying communications.

III. METHODOLOGY

This research uses a case study methodology to assess the design and implementation of an assessment that reduces the vulnerability of it to academic misconduct, and secondly, to incorporate the use of AI as a learning competency into the assessment rubric.

The existing assessment, worth 30% of the overall grade, asked students (in groups) to produce a creative rationale and associated execution based on a client brief. Marks were awarded for the appropriateness of the rationale (to the brief) and the quality of the creative execution. The assessment submission consisted of a completed Creative Rationale template and an example creative execution.

Three problems were identified: There was no assessment of individual input into the finished work; the course asks students to produce a range of creative works, and the existing assessment asks for only one creative example and leaves it up to the students to choose the type and media; there was no mention of generative AI and its acceptable use.

The tools used in this research were an updated assessment brief and an associated rubric. Both these tools were presented in a tutorial environment allowing for student feedback and questions.

The assessment brief asked students to show evidence of research justifying their proposals as well as any working papers or meeting notes showing the development process for the campaign. Additionally, each student in the group also had to separately submit a 250-word reflection discussing their role in the assignment and the use of AI tools in the process of campaign development. This changed the emphasis from the end-product to the process.

Each group was also asked to include the following creative executions:

An A4 sized print advertisement
Three Instagram posts
A storyboard for a 30 second video ad (suitable to be shown on social media).

Students were advised that all elements should be produced to a professional standard and that any images used should be AI generated with prompts used being incorporated in the rationale or appendix.

From an academic integrity viewpoint students were told that AI tools could be used to assist with tasks like brainstorming, structuring, and editing, as long as the final submission is students' own work, and any AI assistance is acknowledged.

The new assessment rubric included an additional section (with a 20% weighting) on AI Use & Critical Engagement. The other criteria were: creative concept; strategic thinking; visual and written execution; and presentation and structure. All the criteria were weighted at 20% each.

Previously the course itself did not have any instruction relating to generative AI, so the following components were added to the course outline:

In week 9, the tutorial introduced two generative AI tools: Chat GPT and Rendernet, plus an activity where students were asked to create a virtual influencer. This task gave students experience in writing prompts and creating images. Both these skills were relevant to the redesigned assessment. Students were also asked to review and reflect on the TEQSA advice to students on the use of artificial intelligence.

The data generated from student reflections in this study are not publicly available due to confidentiality agreements and ethical considerations. Informed consent was obtained to share the example used in figure 1.

IV. RESULTS

At the time of writing, we have just completed the week 9 tutorial and briefed the students on the assessment. Insights from this were overall positive. Students expressed surprise that they were to be allowed to use AI tools and appreciated ssthe ethical restrictions imposed. After the session on creating an AI influencer they felt quite comfortable in using AI to support their creativity and based on the AI influencers they created, I felt they now had the technical skill necessary to undertake the assessment tasks.

Once the assessments have been presented and graded, I will use the 250-word reflections by each student to measure their level of engagement with the assessment and their understanding of the use of AI tools in the university context.

Figure 1 below, shows a sample of a student generated image. The prompt this student used was: "Create a realistic image of a beautiful young Australian woman in a university classroom. She has medium length light brown hair and a pleasant expression, wearing a white lace top. There is a whiteboard in the background and there is light coming through the windows from the side. Make it look like it was taken on a film camera, and have the main focus be on the woman with the background slightly out of focus."

Figure 1

V. CONCLUSION

The assessment design incorporates the use of generative AI tools and provides students with experience in the ethical use of them. The new emphasis on process mitigates the risk of academic conduct breaches and the addition of the individual reflection encourages students to participate in the process. Reviewing these personal reflections will enable me to determine whether the aims of this redesign have been met.

This research will provide a guide to academics who wish to incorporate AI tools into their course assessments. The final results will demonstrate if students are able to responsibly and ethically use AI as an integral part of their assessment writing process. The research also provides a potential model for the teaching of the use AI tools responsibly, including prompting, evaluating outputs, and citing assistance.

This preliminary research can be expanded to include longitudinal studies that measure the effects over successive courses of the students' growing competency in AI use and management. While this study was situated in the Advertising discipline similar studies could be carried out in other creative areas to determine the more widespread applicability of this model.

REFERENCES

[0] Baidoo-anu, D., & Owusu Ansah, L. (2023). Education in the Era of Generative Artificial Intelligence (AI): Understanding the Potential Benefits of ChatGPT in Promoting Teaching and Learning. Journal of AI, 7(1), 52-62. https://doi.org/10.61969/jai.1337500

[1] https://www.teqsa.gov.au/guides-resources/higher-education-good-practice-hub/artificial-intelligence

[2] Russell Butson, R., & Rachel Spronken-Smith, R. (2024) AI and its implications for research in higher education: a critical dialogue, Higher Education Research & Development, (43)3, 563-577, DOI: 10.1080/07294360.2023.2280200

[3] Gui, J., Sun, Z., Wen, Y., Tao, D., & Ye, J. (2021). A review on generative adversarial networks: Algorithms, theory, and applications. IEEE Transactions on Knowledge and Data Engineering. Doi: 10.1109/TKDE.2021.3130191.

[4] Ghanimi, R., Ghanimi, I., Al Karkouri, A., Essahb, H., El Janati, B., Ghanimi, F. (2024). Generative artificial intelligence (AI) on the field of education and learning: threats, limitations and future directions, Journal of Innovation and Digital Health 1(2), 79-84.

[5] van der Vleuten, C., Lindemann, I., & Schmidt, L. (2018). Programmatic assessment: The process, rationale and evidence for modern evaluation approaches in medical education, Medical Journal of Australia, 209(9), 386–388.

[6] Lodge, J. M., Howard, S., Bearman, M., Dawson, P, & Associates (2023). Assessment reform for the age of Artificial Intelligence. Tertiary Education Quality and Standards Agency.

[7] Chiu, T., Ahmad, Z., Ismailov, M., Sanusi, I. (2024). What are artificial intelligence literacy and competency? A comprehensive framework to support them. Computers and Education Open, Vol 6, June. https://doi.org/10.1016/j.caeo.2024.100171

[8] Corbin, T., Dawson, P., Kelli Nicola-Richmond, K., Partridge, H. (2025). 'Where's the line? It's an absurd line': towards a framework for acceptable uses of AI in assess ment, Assessment & Evaluation in Higher Education, DOI: 10.1080/02602938.2025.2456207

Leveraging Emerging Technologies for Entrepreneurship Education in the United Arab Emirates

Prof. Naveed Yasin
Abu Dhabi School of Management
Dubai, United Arab Emirates
0000-0003-3927-3762

Dina Samhouri
London South Bank University
UAE Centre
Dubai, United Arab Emirates
dina.samhouri@gmail.com

Dr. Aiden Salamzadeh
University of Tehran
Tehran, Iran
salamzadehaidin@gmail.com

Dr. Marc Poulin
Abu Dhabi School of Management
Abu Dhabi, UAE
m.poulin@adsm.ac.ae

Dr. Lizana Oberholzer
University of Wolverhampton
Wolverhampton, UK
lizana.oberholzer@buckingham.ac.uk

Abstract— The adoption of emerging technologies is reshaping entrepreneurship education, fostering innovation and adaptability. This paper explores how artificial intelligence (AI), blockchain, and virtual reality (VR) can enhance entrepreneurial learning in the United Arab Emirates (UAE). While technological advancements offer significant opportunities, challenges such as digital literacy gaps, regulatory concerns, and infrastructure limitations persist. By analyzing the UAE's educational and entrepreneurial ecosystem, this study highlights strategies for integrating technology into curricula, emphasizing experiential learning and university-industry collaboration. Findings suggest that embedding technology modules and fostering industry collaboration are key to effective implementation. Addressing barriers through public-private partnerships and faculty training can ensure an inclusive, resilient ecosystem that empowers future entrepreneurs.

Keywords— Entrepreneurship Education, Emerging Technologies, AI, Blockchain, VR, Higher Education, UAE.

I. INTRODUCTION

Entrepreneurship education is evolving in response to rapid technological advancements, shifting from conventional teaching to technology-enhanced experiential learning [1]. The UAE's strategic vision—encompassing the National Innovation Strategy and Vision 2031—prioritises digital transformation in both education and economic diversification [2]. However, integrating emerging technologies into entrepreneurship education presents ongoing challenges, including uneven access, digital literacy disparities, and cultural resistance to risk-taking [3].

This paper addresses the pressing need to understand how emerging technologies such as artificial intelligence (AI), blockchain, virtual reality (VR), and machine learning (ML) [19] can be effectively embedded in entrepreneurship education within the UAE. The study is guided by three central research questions:

- How can these technologies enhance entrepreneurship education in the UAE?
- What barriers and enablers influence their implementation?
- How can experiential learning and university–industry collaboration support this transformation?

The aim is to explore how these technologies can enhance entrepreneurial learning, while the objectives are to analyse the current educational landscape, identify challenges and opportunities, and offer strategies for effective integration. The paper contributes to the field by bridging the gap between national digital ambitions and practical implementation, providing insights relevant to policymakers, educators, and industry stakeholders.

The paper is structured as follows: the next section reviews existing literature on technology and entrepreneurship education. This is followed by an outline of the methodology, a discussion of key findings, and concluding recommendations for building a more inclusive, innovation-driven learning ecosystem.

II. METHODOLOGY

This study employs a qualitative, exploratory approach, drawing on secondary sources, policy documents, and case studies of entrepreneurship-focused initiatives within UAE-based universities. Notable examples include the Hult Prize programme at the American University in Dubai, the in5 Innovation Centre affiliated with the Dubai Knowledge Park, and the Startup Bootcamp hosted by the University of Sharjah. These institutions were selected based on their active incorporation of emerging technologies, links to government innovation strategies, and engagement with industry partners.

The analysis identifies how digital tools, specifically AI, blockchain, and VR, are being used to enhance experiential learning and bridge the gap between theory and practice. A thematic analysis was conducted to extract recurring patterns related to curriculum design, institutional support, public–private collaboration, and student engagement. Global best practices are also examined to contextualise the UAE's advancements and ongoing challenges in integrating technology into entrepreneurship education [4].

PRELIMINARY FINDINGS AND DISCUSSION

A. Role of Emerging Technologies in Entrepreneurship Education

TABLE I. ROLE OF EMERGING TECHNOLOGIES IN ENTREPRENEURSHIP EDUCATION

Technology	**Key Contributions**	**Challenges**	**Implementation Strategies**
AI	• Adaptive learning and personalised feedback • Market-driven simulations	• Bias in algorithms • Need for data privacy safeguards	• Embed AI tools in curricula • Faculty training in AI applications
Blockchain	• Secure credentialing • Transparent start-up funding and IP tracking	• Regulatory uncertainty • Limited awareness of utility	• Collaboration with fintech incubators • Regulatory alignment
VR	• Immersive experiential learning • Real-world scenario simulations	• High infrastructure costs • Limited technical expertise	• Shared VR labs • Curriculum-linked simulations
Sources: Chen et al. (2024)[5]; Yasin et al. (2021, 2023)[6][7]; Pan (2022)[8]; Hyams-Ssekasi & Yasin (2022)[1]; Dennett (2025)[5]			

These technologies demonstrate strong potential in enhancing learner engagement, real-time decision-making, and scenario-based training. When effectively applied, they create immersive environments where students can experiment, fail, and refine entrepreneurial competencies with reduced real-world risk.

B. Challenges in Technology Integration

Despite this promise, implementation is constrained by multiple factors:

a.) Digital skills gap: Uneven digital literacy among students and educators poses a barrier to adoption [1].

b.) Regulatory and ethical concerns: Data privacy and AI biases require governance frameworks to protect academic integrity and innovation [9].

c.) Infrastructure and resource limitations: Smaller institutions struggle to afford advanced digital tools, creating disparities in access to technology-driven entrepreneurship education [10, 17].

C. Strategies for Effective Implementation

To address these barriers and promote sustainable adoption, the following strategies have proven effective:

a.) Curriculum integration: Embedding AI, blockchain, and VR modules into entrepreneurship programs ensures practical, future-oriented learning across diverse contexts [11, 15, 16, 18].

b.) University-industry collaboration: Strengthening public-private partnerships fosters knowledge exchange, funding opportunities, and access to cutting-edge technology [7, 12, 13, 14].

c.) Faculty development programs: Training educators in emerging technologies enhances pedagogical effectiveness and ensures up-to-date teaching methodologies [5].

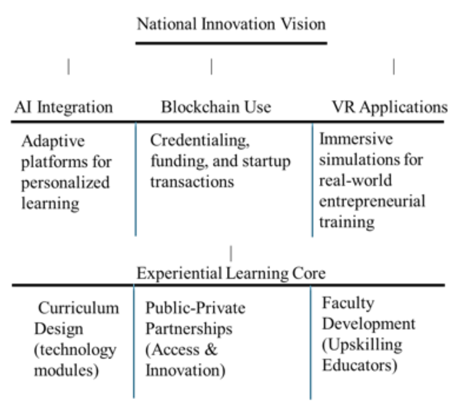


Fig 1 : National Innovation Vision

This framework in Figure 1 illustrates the interaction between emerging technologies (AI, blockchain, and VR) and their functional integration into entrepreneurship education in the UAE. It shows how national policy underpins digital transformation, supported by aligned strategies in curriculum design, industry collaboration, and educator training—all centred on fostering experiential learning.

III. CONCLUSION AND FUTURE RESEARCH

The UAE's entrepreneurship education sector is poised for transformation through the integration of AI, blockchain, and VR. However, overcoming digital literacy gaps, infrastructure constraints, and regulatory challenges remains essential to effective implementation. For policymakers, this includes supporting public-private partnerships, investing in digital infrastructure, and developing regulatory frameworks that facilitate responsible technology use. For educators, professional development and curriculum redesign are key to embedding emerging technologies in meaningful and accessible ways.

This study contributes to the UAE's Vision 2031 by supporting national efforts to build a knowledge-based, innovation-driven economy through education reform. It also offers practical insights for strengthening entrepreneurship ecosystems more broadly, particularly in emerging markets with similar ambitions.

Future research should examine scalable models for technology adoption in resource-limited institutions, explore cultural perceptions of tech-enabled learning, and assess the long-term impact of digital pedagogy on entrepreneurial outcomes. By fostering an inclusive, technology-driven educational framework, the UAE can further establish itself as a global leader in entrepreneurship education.

REFERENCES

[1] Hyams-Ssekasi and N. Yasin, "The Future of Enterprise and Entrepreneurship Education in Relation to Technology," in Technology and Entrepreneurship Education, D. Hyams-Ssekasi and N. Yasin, Eds., Cham: Springer International Publishing, 2022, pp. 251–259. doi: 10.1007/978-3-030-84292-5_11.

[2] UAE Government, "The National Agenda for Entrepreneurship and SMEs | The Official Portal of the UAE Government." Accessed: Mar. 05, 2025. [Online]. Available: https://u.ae/en/about-the-uae/strategies-initiatives-and-awards/strategies-plans-and-visions/finance-and-economy/the-national-agenda-for-entrepreneurship-and-smes

[3] A. Salamzadeh, M. Tajpour, and E. Hosseini, "Measuring the Impact of Simulation-Based Teaching on Entrepreneurial Skills of the MBA/DBA Students," in Technology and Entrepreneurship Education: Adopting Creative Digital Approaches to Learning and Teaching, 2022, pp. 77–104. doi: 10.1007/978-3-030-84292-5_4.

[4] A. Al-Gindy, N. Yasin, M. Aerabe, and A. Omar, "Integrating Digital Technology in Enterprise and Entrepreneurship Education," 2022, pp. 53–75. doi: 10.1007/978-3-030-84292-5_3.

[5] P. Dennett, "Training Needs for Entrepreneurship Education in the Creative Industries in the UAE," in Entrepreneurship in the Creative Industries, Routledge, 2025.

[6] N. Yasin, Z. Khansari, and K. Tirmizi, "Exploring the challenges for entrepreneurship business incubator hubs in the United Arab Emirates," Int. J. Glob. Small Bus., vol. 12, p. 190, Jan. 2021, doi: 10.1504/IJGSB.2021.114575.

[7] N. Yasin, S. A. M. Gilani, G. Nair, G. M. Abaido, and S. Askri, "Establishing a nexus for effective university-industry collaborations in the MENA region: A multi-country comparative study," Ind. High. Educ., vol. 37, no. 6, pp. 838–859, Dec. 2023, doi: 10.1177/09504222231175862.

[8] Y. Pan, "Designing Smart Space Services by Virtual Reality-Interactive Learning Model on College Entrepreneurship Education," Front. Psychol., vol. 13, p. 913277, Jul. 2022, doi: 10.3389/fpsyg.2022.913277.

[9] Rasul, T., Nair, S., Kalendra, D., Balaji, M. S., Santini, F. de O., Ladeira, W. J., Rather, R. A., Yasin, N., Rodriguez, R. V., Kokkalis, P., Murad, M. W., & Hossain, M. U. (2024). Enhancing academic integrity among students in GenAI era: A holistic framework. The International Journal of Management Education, 22(3), 101041. doi.org/10.1016/j.ijme.2024.101041

[10] A. Balawi, "Entrepreneurship ecosystem in the United Arab Emirates: An empirical comparison with Qatar and Saudi Arabia," Int. Entrep. Rev., vol. 7, no. 2, pp. 55–66, 2021, doi: 10.15678/IER.2021.0702.05.

[11] N. Yasin and K. Hafeez, "Enterprise Simulation Gaming: Effective Practices for Assessing Student Learning with SimVenture Classic and VentureBlocks," in Experiential Learning for Entrepreneurship: Theoretical and Practical Perspectives on Enterprise Education, 2018, pp. 51–69. doi: 10.1007/978-3-319-90005-6_3.

[12] Yasin, N., & Gilani, S.A.M (2022). Assessing the current state of university-based business incubators in Canada. Industry and Higher Education, 0(0). https://doi.org/10.1177/09504222221124749

[13] Yasin, N., Gilani, S. A. M., Nair, G., Abaido, G. M., & Askri, S. (2023). Establishing a nexus for effective university-industry collaborations in the MENA region: A multi-country comparative study. Industry and Higher Education, 0(0). https://doi.org/10.1177/09504222231175862

[14] Yasin, N., & Gilani, S. A. M. (2022). 'Imitate or Incubate?' Evaluating the Current State of University-Based Business Incubators in the United Arab Emirates. FIIB Business Review, 0(0). https://doi.org/10.1177/23197145221112744

[15] Yasin, N. and Khansari, Z. (2021) Evaluating the impact of social enterprise education on students' enterprising characteristics in the United Arab Emirates. Education+ Training 63(6) pp.827-905

[16] Salamzadeh, A., Rezaei, H., Hadizadeh, M. Ansari, G., & Yasin, N., (2023) The Application of Strategic Foresight in Women's Entrepreneurship Development. Journal of Women's Entrepreneurship Education (Special Issue) 15(1-2) pp.16-36

[17] Ladeira, W. J., Santini, F. de O., Rasul, T., Cheah, I., Elhajjar, S., Yasin, N., & Akhtar, S. (2024). Big data analytics and the use of artificial intelligence in the services industry: a meta-analysis. The Service Industries Journal, 44(15–16), 1117–1144. https://doi.org/10.1080/02642069.2024.2374990

[18] Yasin, N., Hafeez, K. and Salamzadeh, A. (2025), A Cross-country Comparative Ethnographic Analysis of Immigrant Enclave Entrepreneurship, Qualitative Market Research: An International Journal, 28(1), pp. 1-37. https://doi.org/10.1108/QMR-11-2023-0164

[19] Gilani S.A.M., Copiaco A, Gernal L, Yasin N., Nair G, & Anwar I. (2023) Savior or Distraction for Survival: Examining the Applicability of Machine Learning for Rural Family Farms in the United Arab Emirates. Sustainability (MDPI). 2023; 15(4):3720. https://doi.org/10.3390/su15043720

Sentiment Analysis for Government E-Services: A CRISP-DM Machine Learning Approach

Saeed AlRashdi
Department of Government Enablment (GovDigital)
Abu Dhabi , U.A.E.
Saeed.AlRashdi@dge.gov.ae

Abstract—The increasing adoption of digital platforms by government organizations necessitates a deeper understanding of user sentiment to enhance public services. This paper presents an AI-driven sentiment analysis framework focused on Arabic-language Twitter data related to Abu Dhabi's TAMM e-services platform. Utilizing the CRISP-DM methodology, this study collected and preprocessed over 3,000 Arabic tweets, applied machine learning algorithms including Logistic Regression, Naïve Bayes, SVM, Random Forest, and MLP, and evaluated their performance in classifying sentiment. The results reveal that Multinomial Naïve Bayes achieved the best balance of accuracy (79.4%) and training efficiency (0.68s). Comparative evaluations using Cohen's Kappa and Matthews Correlation Coefficient validate inter-model agreement. This work demonstrates the feasibility and value of leveraging Arabic sentiment analysis for real-time feedback, aligning with national strategies such as the UAE Happiness Index [1], [2].

Keywords— Sentiment Analysis, Arabic NLP, CRISP-DM, TAMM, Machine Learning, Public Services, UAE Happiness Index

I. INTRODUCTION

As governments embrace digital transformation, e-services like TAMM in Abu Dhabi aim to streamline service delivery and improve citizen satisfaction. Understanding public sentiment through social media provides actionable insights that support policymaking and service enhancement. Sentiment analysis, particularly for Arabic text, poses unique challenges due to linguistic complexity and dialectal variation [3], [4].

Previous research has explored various machine learning methods for sentiment analysis [3], with specific focus on Arabic sentiment modeling [4]. However, limited studies have applied these techniques to the domain of government e-services, especially in the Arabic-speaking context. This study bridges that gap by applying a structured data mining methodology (CRISP-DM) and comparative machine learning evaluation to Arabic tweets related to TAMM.

II. METHODOLOGY

This study followed the CRISP-DM methodology [5], a structured process that includes six major phases, applied specifically to sentiment analysis:

- **Business Understanding:** The primary objective was to extract and analyze public sentiment about the TAMM government e-service platform in Abu Dhabi to help enhance decision-making and improve user satisfaction.
- **Data Understanding:** We collected Arabic-language tweets related to TAMM using Tweet Flash [6]. This data was derived from both labeled sentiment datasets and live user-generated content.
- **Data Preparation:** Preprocessing was a critical step to ensure data quality and accuracy. The steps included:
    - Removing non-Arabic words, emojis, URLs, and repeated characters.
    - Normalizing Arabic text by unifying variations in spelling (e.g., replacing أ/آ/إ with ا).
    - Tokenizing and applying TF-IDF vectorization to transform the raw text into numerical feature representations [7], [8].
    - Merging normalized labeled data (Arabic Sentiment Corpus) with the preprocessed TAMM tweets to form a unified dataset.

The full flow of handling the data is illustrated in **Figure 1**, showcasing the end-to-end pipeline from tweet collection to model training and evaluation.

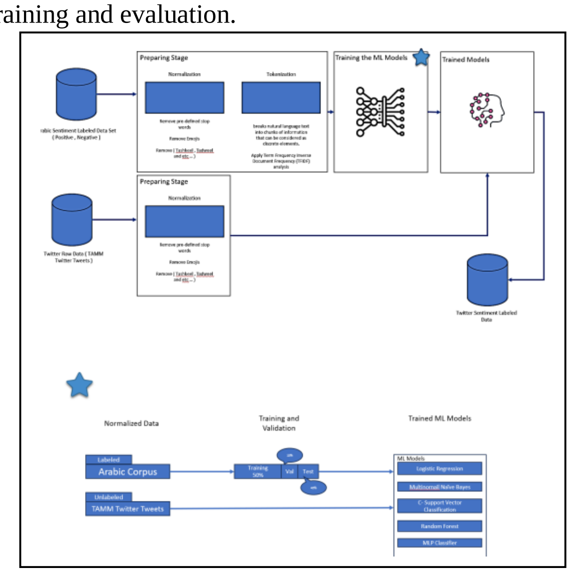


Figure 2: Workflow for Arabic Sentiment Analysis (from preprocessing to model evaluation)

- **Modeling:** Five classifiers were selected to evaluate their suitability for real-time sentiment analysis in Arabic: Logistic Regression, Multinomial Naïve Bayes, SVM (C-SVC), Random Forest, and MLP. Naïve Bayes was selected for its simplicity and performance with sparse data; SVM and MLP for their ability to capture nonlinear patterns; and Random Forest for robustness against overfitting.

- **Evaluation:** The dataset was split into 50% training, 10% validation, and 40% testing. Metrics included accuracy, training time, Cohen's Kappa, and MCC [9], [10].

III. RESULTS

The performance of each model is summarized in Table 1. Figures 2 and 3 illustrate comparative accuracy and training time.

TABLE 1: Model Performance Comparison

Algorithm	**Accuracy (%)**	**Training Time (s)**	**Notes**
Logistic Regression	78.9	7.28	Balanced accuracy with fast training
Multinomial Naïve Bayes	**79.4**	**0.68**	Best balance between accuracy and efficiency
SVM (C-SVC)	79.1	3694.44	High test score but slowest training
Random Forest	78.9	966.59	High test score; potential overfitting
MLPClassifier	77.4	866.90	Complex model, risk of overfitting

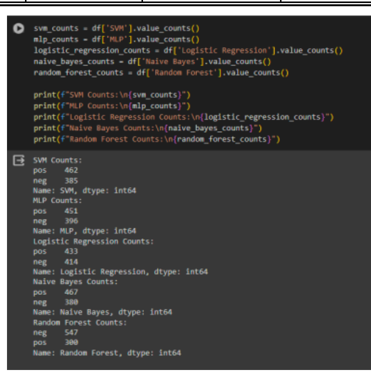


Fig 4 : Training Time for Each Model

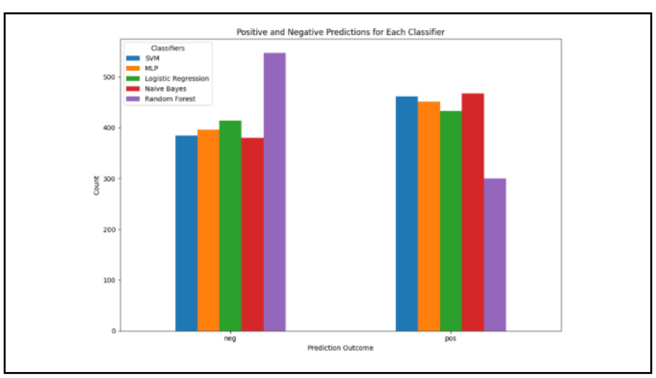


Fig 3: Accuracy of Sentiment Models

Agreement metrics:

- **Cohen's Kappa**: Strong agreement between SVM and Logistic Regression (0.80).

- **MCC**: Validated model consistency across predictions.

These results indicate that while complex models such as Random Forest and MLP offer high performance, their training cost and overfitting potential make them less ideal for real-time applications. Multinomial Naïve Bayes, on the other hand, combines strong accuracy with minimal computational overhead, making it a suitable candidate for integration into live systems.

IV. DISCUSSION

Arabic sentiment analysis requires rigorous preprocessing and model tuning [4], [11]. The Multinomial Naïve Bayes model proved most efficient, showing both high accuracy and the fastest training time—making it ideal for government applications where responsiveness is ke.

A qualitative complement was added through sentiment word clouds. These visualizations (Figure 4) reveal key user themes. Positive sentiments included words like "سهل" (easy), "ممتاز" (excellent), and "سريع" (fast), indicating satisfaction with usability. Negative terms like "تأخير" (delay), "صعب" (difficult), and "غير واضح" (unclear) reveal areas for service refinement.

Fig 4 : Word Clouds for Positive and Negative Sentiment Classification

These visuals provide intuitive support for identifying service strengths and weak points. Importantly, they align directly with real-world citizen experiences, informing data-driven improvements in TAMM.

V. CONCLUSION

This study presents a CRISP-DM-based framework to apply sentiment analysis to Arabic-language Twitter data regarding TAMM e-services. It highlights the Multinomial Naïve Bayes model as a practical, efficient classifier for deployment in real-time feedback systems.

The findings have direct implications for service development teams: user feedback can be rapidly interpreted and categorized to inform dynamic improvements, aligning with goals like the UAE Happiness Index.

Limitations: The analysis focused on Modern Standard Arabic and did not account for dialectal diversity. Also, social media users may not represent the full demographic range of TAMM users.

Future Work: Research can expand into dialect detection, apply deep learning models like AraBERT, and integrate with real-time dashboards used by government support teams.

REFERENCES

[1] The Official Portal of the UAE Government, “Happiness and National Agenda”, 2023. [Online]. Available: https://u.ae/en/about-the-uae/the-uae-government/government-of-future/happiness

[2] [M. Wankhade, A. C. S. Rao, and C. Kulkarni, "A survey on sentiment analysis methods, applications, and challenges," Artif. Intell. Rev., vol. 55, no. 7, pp. 5731–5780, 2022.

[3] S. Mohammad, "Sentiment Analysis in Arabic," 2021. [Online]. Available: https://saifmohammad.com/WebPages/ArabicSA.html

[4] C. Schröer, F. Kruse, and J. M. Gómez, "A systematic literature review on applying CRISP-DM process model," Procedia Comput. Sci., vol. 181, pp. 526–534, 2021.

[5] Shane, “Tweet Flash”, Apify. [Online]. Available: https://apify.com/shanes/tweet-flash

[6] M. I. Alfarizi, L. Syafaah, and M. Lestandy, “Emotional Text Classification Using TF-IDF And LSTM,” JUITA, vol. 10, no. 2, pp. 225–232, 2022.

[7] H. T. Sueno, B. D. Gerardo, and R. P. Medina, "Multi-class document classification using SVM based on improved Naïve Bayes vectorization," Int. J. Adv. Trends Comput. Sci. Eng., vol. 9, no. 3, 2020.

[8] G. Rau and Y. S. Shih, “Evaluation of Cohen's kappa and other measures of inter-rater agreement,” J. Engl. Acad. Purp., vol. 53, 2021.

[9] J. Yao and M. Shepperd, “Assessing software defection prediction performance: Why using the Matthews correlation coefficient matters,” in Proc. Eval. Assess. Softw. Eng. Conf. (EASE), 2020.

[10] S. H. Janjua et al., “Multi-level aspect based sentiment classification of Twitter data using hybrid approach in deep learning,” PeerJ Comput. Sci., vol. 7, p. e433, 2021.

AI Early Detection System for Autism Screening

Aiyda Salem Alseiari
Master of Science in Business Analytics
Abu Dhabi School of Management (ADSM)
adsm-215295@adsm.ac.ae

Abstract—This research aims to develop an AI-based framework for early detection of autism spectrum disorder (ASD) in children using eye-tracking data. This approach uses convolutional neural networks (CNNs) introduced by [13} combined with hybrid models to automatically classify the distinctive gaze recognition patterns of individuals with ASD, providing non-invasive detection. This framework was designed after reviewing existing AI models, structuring relevant components, and obtaining expert feedback. A case study analysis indicated that AI models, particularly adaptive CNN models, outperformed traditional screening methods in terms of accuracy and efficiency [16]. The goal is to develop a practical, interpretable, and scalable ASD screening framework to enhance early clinical decision-making.

Keywords—**Autism Spectrum Disorder, Convolutional Neural Networks, Screening, Deep Learning, Deep Neural Networks.**

I. INTRODUCTION

The global prevalence of autism spectrum disorder (ASD) has increased dramatically. The Autism Spectrum Disorder Monitoring (ASDM) network states that in the United States, the prevalence among 8-year-olds increased from 1 in 150 in 2000 to nearly 1 in 36 in 2020. The ASDM network suggests that this increase can be attributed to improved screening procedures, increased awareness, and changes in screening criteria [8]. Early screening for ASD, particularly before age 3, is critical. Interventions implemented during this critical developmental stage will significantly enhance children's social, cognitive, and communication abilities [17]. However, the average screening age in developed countries remains around four years, and even higher in low-resource countries, highlighting the urgent need for rapid and effective early screening systems.

The methodology applied in this research study follows a structured, five-step approach to developing an AI-based framework for early detection of ASD using eye-tracking data. In the first step, the researcher systematically examines the problem by examining gaps in current ASD screening methods and identifying how AI and eye-tracking technology can be used for improvement. In the second step, a detailed literature review is conducted to evaluate current AI screening frameworks, with a particular focus on those that incorporate CNNs and DNNs as their most prominent features, and to identify gaps and best practices.

In the third step, based on the literature review, a preliminary framework is proposed that combines the features of AI and eye-tracking technology for early ASD screening. In the fourth step, a survey method is implemented, gathering responses from healthcare professionals, AI specialists, and educators to evaluate the framework's effectiveness, practicality, and feasibility. In the final step, the framework is enhanced, taking into account the feedback received to ensure accurate screening, treatment, and integration with the healthcare and education systems in the UAE.

The proposed AI-driven framework for early ASD screening has clear benefits for children, parents, healthcare professionals, and even policymakers. For children, early and accurate screening allows for interventions that are more likely to improve cognitive, social, and communication skills[17]. Parents and caregivers benefit from the availability of simple, non-invasive tools that reduce the emotional and logistical burden associated with traditional assessments [4].

Healthcare professionals use a more objective, data-driven approach, increasing the accuracy of the method and reducing overreliance on behavioral observations to record subjective measurements as noted by [7]. Integrating this framework into healthcare and education systems fosters collaboration between clinicians and educators, enhancing the use of AI-based screening tools to better identify ASD [14]. This framework also enables policymakers to study and integrate it into existing healthcare and education systems in the UAE. Furthermore, the framework encourages collaboration across disciplines and supports research-based strategies to improve screening for ASD [5; 18].

II. METHODOLOGY

Step 1: Understanding the research problem: Identifying existing challenges in early screening for ASD. This includes examining the limitations of current screening tools and the potential for combining AI and eye-tracking data to achieve accurate and non-invasive screening. This basic understanding supports the development of a practical AI-based screening framework.

Step 2: Conduct a comprehensive literature review: To analyse existing research on AI-driven ASD screening frameworks, particularly focusing on Convolutional Neural Networks (CNNs) and Deep Neural Networks (DNNs), and hybrid models. This step highlights the methodological strengths and gaps in current practices, contributing to a solid theoretical foundation for framework design.

Step 3: Framework Design: Develop a preliminary framework that features effective AI models, data preprocessing strategies, and feature extraction methods identified in previous literature revies. This includes defining model components, data flow, and evaluation metrics. The framework also addresses data quality variability through normalization, handling missing values, and noise reduction. Techniques such as SMOTE and

GANs are conceptually integrated to manage imbalanced datasets, while evaluation across diverse datasets supports demographic generalization and fairness.

Step 4: Expert Consultation and Feedback: This step focuses on collecting primary data through expert surveys. Healthcare professionals, AI researchers, and educators will provide feedback through structured questionnaires, providing insights into the framework's feasibility and practical effectiveness. This step ensures that the framework supports the needs of professionals directly involved in ASD screening and interventions. Incorporating interpretation techniques, such as SHAP and Grad-CAM, improves clinical confidence and model interpretability, making it more transparent in clinical settings.

Step 5: Refining the Framework: This involves refining the framework based on expert feedback. The reviews will address any identified gaps, with a focus on improving screening accuracy and enhancing practical application. The final framework will be refined for practical application in both the healthcare and education settings in the UAE.

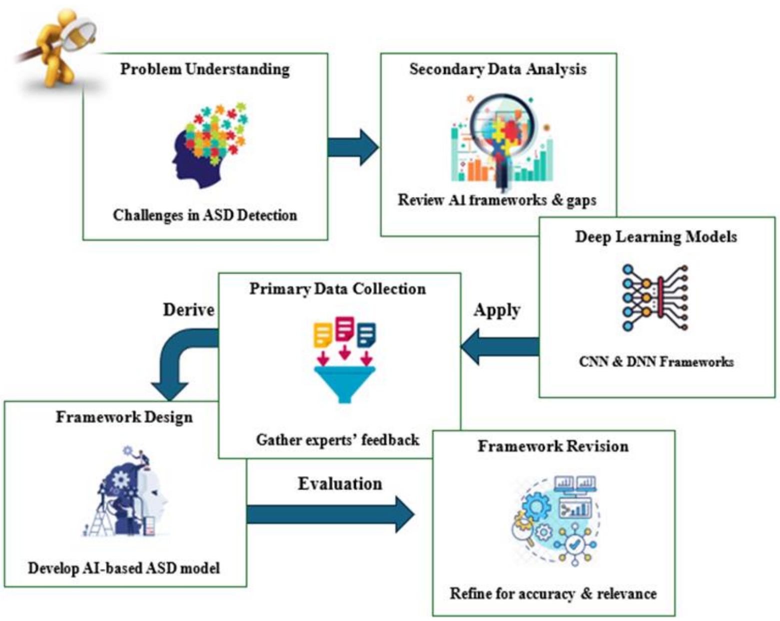


Fig 1. Flowchart of the Research Approach

III. LITERATURE REVIEW

Recent advances in AI have revolutionized the screening of ASD by integrating eye-tracking data and DL models. Multiple studies have revealed the effectiveness of CNNs, hybrid models, and other AI techniques in identifying atypical gaze behaviors associated with ASD [2]. This literature review highlights the most important findings from recent studies that contribute to the development of AI-based screening framework.

The study by [16] developed an Adaptive Deep Convolutional Neural Network (ADCNN) framework to enhance early screening of ASD in children. The study used two datasets: the Eye-Tracking Scan Path (ETSP) dataset containing 547 images from 59 children (219 ASD, 328 TD), and the ABIDE dataset, which provided additional personal information. The methodology combined Discrete Wavelet Transform (DWT) for feature extraction, Kernel PCA for dimensionality reduction, and Grey Wolf Optimization (GWO) for feature selection. In total, 131 features were analyzed. The ADCNN model, built four convolutional layers and achieved an AUC of 94.51%, outperforming Decision Tree by 30%, KNN by 28%, and conventional CNNs by 8%. The study highlights the potential of AI-integrated, non-invasive methods for accurate and early ASD screening.

In another review [6] proposed a new T-CNN-ASD model that applies to CNN to classify children with ASD based on their eye-tracking scan paths. The dataset included a total of 547 images of 59 children with 29 screened with ASD and 30 with TD. The T-CNN-ASD model architecture comprised of two hidden layers with 300 and 150 neurons, respectively. The model was trained using 10-fold cross-validation and a 20 percent dropout rate to reduce overfitting. It also outperformed traditional machine learning algorithms such as RF, DT, KNN, and MLP models, achieving an accuracy of 95.59%, the results highpoint the model's powerful ability to classify children with ASD versus their non-ASD peers based on eye-tracking behavior and enhance the potential of deep learning techniques in early screening for ASD [6].

Similarly, [10] proposed the Involution Fused ConvNet model, which uses convolutional and involutional layers to classify eye-tracking patterns in children with ASD. The dataset consisted of two eye-tracking datasets, one containing 547 images (219 ASD, 328 non-ASD) and the other containing 300 images from 28 children. Their hybrid model consisted of three involution layers followed by three optimally convolution layers, designed to enhance spatial feature extraction. The model achieved an accuracy of 99.43% on one dataset and 96.78% on another dataset. These results highlight the effectiveness of incorporating location- -sensitive spatial processing through involution layers in improving the accuracy of ASD detection. [10].

In another study conducted by [7] developed an intelligent eye-tracking system using DL to support early screening of ASD. The study integrated the use of MobileNet, VGG19, DenseNet169, and a hybrid MobileNet-VGG19 on a dataset of 547 images of children from 219 with ASD and 328 TD. Among these, MobileNet had outstanding accuracy of 100% while VGG19 had 92%, DenseNet169 achieved 78%, and the hybrid model scored 91%. DL-based autonomous eye tracking systems have proven their value in non-invasive screening solutions for ASD.

A systematic literature review by [19] included 130 studies exploring applications of deep learning techniques in image classification from MRI scans, eye tracking data, and face recognition. The review concluded that CNNs had classification accuracy between 75% and 95%. The highest screening accuracies were achieved with multimodal approaches using eye-tracking and MRI scans. The study emphasized the need for standardized datasets and culturally inclusive models to improve the accuracy and real-world applicability of AI-based ASD screening systems.

Combining Deep Learning Models with eye-tracking technology improves early detection of ASD in children. As for the dataset of 59 children, included 29 screened with ASD and 30 TD. Different models were tested including LSTM, CNN-LSTM, GRU, and BiLSTM. Following data scaling and missing value treatment, the LSTM model resulted in an accuracy of 98.33%, with CNN-LSTM close behind at 97.94%. GRU and BiLSTM obtained 97.49% and 96.44%, respectively. The review by [2] indicates that the accuracy of early detection of ASD significantly increases by combining DL and eye-tracking data.

In another investigation, [1] designed a CNN-based model to predict ASD using scan paths of eye-tracking. This research trained CNN on scanning path images of children with ASD and TD by using deep feature extraction for classification. The model showed excellent results, having an accuracy of 98%, proving how powerful CNNs are in detecting unusual gaze patterns. This study emphasized the combination of eye-tracking in addition to deep learning to objectively and quickly screen ASD [1].

This study used eye-tracking scan paths as biomarkers and analyzed data from 547 participants (219 ASD and 328 TD children) to evaluate ML models. The research tested several classifiers, Boosted Decision Tree (BDT), Deep Support Vector Machine (DSVM), Decision Jungle (DJ), and a Deep Neural Network (DNN). The DNN outperformed by achieving an AUC of 97%, sensitivity of 93.28%, and specificity of 91.38%. These results highlight the potential of machine learning algorithms to provide reliable non-invasive screening techniques for the early detection of ASD, as noted by [11].

A survey review of 35 selected studies focused on the application of AI tools in initial screening for ASD between 2011 and 2021. The review focused on the application of ML and DL techniques on eye movements, facial gestures, motor activities, and Electronic Health Records (EHR). According to the results, screening accuracy rates ranged from 70% to 95%, with eye tracking and facial recognition achieving the highest accuracy. As [12] noted, the provision of standardized data suitable for use in clinical contexts remains insufficient.

A computer-assisted screening model was developed and designed based on visual attention patterns derived from eye-tracking technology. The study included 106 participants, including 76 children with ASD and 30 TD. Classification was performed using artificial neural networks (ANNs) and support vector mechanisms (SVMs). A genetic algorithm was able to extract 15 relevant variables out of 28 potential variables. The accuracy of the ANN model was 90%, while the achieved sensitivity and specificity were 69% and 93%, respectively. [15] claim that these results provide further evidence of specialized gaze patterns in children with ASD, confirming the importance of eye-tracking data.

TABLE II. METHODS USED IN ASD DETECTION

Methods Used	No. of Features	Results	Findings	Citation
Adaptive Deep CNN (ADCNN)	Yes (5-Fold Cross Validation)	AUC: 94.51%	ADCNN outperformed DT by 30%, KNN by 28%, and conventional CNNs by 8% in ASD detection.	Palanichamy et al. (2025) [16]
CNN-ASD (CNN)	Yes, (10-fold cross-validation, Dropout 20%)	Accuracy: 95.59%	The CNN model showed strong performance in distinguishing between children with ASD and TD.	Alsaidi et al. (2024) [6]
Involution Fused ConvNet	Yes, (Spatial Feature Extraction)	99.43% (Dataset 1), 96.78% (Dataset 2)	The highest accuracy among all models; spatial feature extraction enhances classification.	Islam et al. (2024) [10]
MobileNet, VGG19, DenseNet169, Hybrid MobileNet-VGG19	Yes, (Hybrid Model with Transfer Learning)	MobileNet: 100% accuracy; VGG19: 92% accuracy; DenseNet169: 78% accuracy; Hybrid: 91% accuracy	The study showed that MobileNet achieved the highest accuracy in classifying ASD from eye-tracking data, indicating its potential for effective ASD screening.	Alsharif, R., & Ghulam, A. (2024)[7]
Systematic Review of Multiple Deep Learning Models (e.g., CNNs)	Not Applicable	CNNs: Classification accuracies ranging from 75% to 95%; Multimodal approaches combining eye-tracking and MRI data led to even higher screening accuracies	The systematic review highlights the effectiveness of DL models, particularly CNNs, in ASD screening. It also emphasizes that incorporating multimodal data sources, such as eye-tracking	Uddin, M. I., & Rahman, M. M. (2024) [19]

STM, CNN-LSTM, GRU, BiLSTM	Yes, (Scaling, Handling Missing Values)	LSTM: 98.33%, CNN-LSTM: 97.94%, GRU: 97.49%, BiLSTM: 96.44%	LSTM accomplished the highest level of accuracy, outperforming deep learning models over traditional ML.	Ahmed et al. (2023) [2]
ANN, FFNN, CNN (GoogleNet, ResNet-18), Hybrid CNN-Support Vector Machine (SVM)	Yes, (Scaling, Feature Extraction, Handling Missing Values)	ANN: 99.8% accuracy; ResNet-18: 97.6% accuracy; Hybrid models: 95.5% and 94.5% accuracy	The study reveals that combining AI models with eye-tracking data can effectively help in the early screening of ASD, with ANN achieving the highest accuracy.	Ahmed, M., & Jadhav, A. (2022) [1]
BDT, DSVM, DJ& DNN	Yes, Features extracted from visualized eye-tracking scanpaths; image augmentation applied	DNN: AUC 97%, sensitivity 93.28%, specificity 91.38%	The study shows that visual representations of eye-tracking data, when analyzed using DL models, can effectively serve as biomarkers for early ASD screening.	Kanhirakadavath, M., & Chandran, S. (2022) [11]
Various ML/DL Methods	Not Applicable	Detection accuracies ranging from 70% to 95%.	Eye-tracking and facial recognition technologies offer the highest screening accuracy for early ASD detection.	Kohli, D., & Sharma, P. (2022) [12]
Artificial Neural Networks (ANN), Support Vector Machines (SVM)	Yes, (Feature Selection via Genetic Algorithm)	ANN: 90% precision, 69% sensitivity, 93% specificity	The study examined a computational method combines Visual Attention Models and AI techniques, showing the ability of eye-tracking data in ASD screening.	Oliveira, G., & Silva, H. (2021) [15]

IV. RESEARCH ETHICS

The development of this framework ensures compliance with data protection regulations, protects the identity of expert participants, and ensures the responsible use of AI technologies. Consent will be obtained from experts providing feedback, ensuring that participation is not forced [9]. This research adheres to the laws of the GDPR and the UAE Data Law, ensuring data is protected and handled anonymously [18]. Ethical sensitivity is maintained by designing the framework within the boundaries of ethical AI and regional boundaries [5]. Since this research does not involve any human interventions or clinical trials, it poses minimal ethical risks [20]. No medical procedures will be collected along with personal health information.

V. CONCLUSION

ASD remains one of the most concerning conditions across the globe, making early detection essential to improve developmental outcomes. This chapter examined the limitations of traditional approaches to screening ASD, which rely primarily on behavioral and clinical assessment. Such approaches often delay screening and intervention. The introduction of eye-tracking technology, which analyses gaze fixation and visual attention movements, is an innovative non-invasive approach for ASD screening. Several studies were examined that demonstrated high accuracy classifiers for ASD based on ML and DL approaches with CNNs being the most effective. The primary models of CNNs, known as ResNet, MobileNet, and VGG19, surpassed traditional ML techniques by achieving over 90% classification accuracy. Nevertheless, dataset and standardization limitations combined with low model generalisability remain imperative issues. The chapter also discussed combining multimodal data and eye-tracing data with other behavioral and communication data to improve screening accuracy. There is no doubt that AI-enhanced screening tools for ASD have the potential to transform early screening, but a significant amount of work is needed in terms of model review, ethical issues, and practical clinical considerations.

REFERENCES

[1] Ahmed, Z. A., & Jadhav, M. E. (2020). Convolutional neural network for prediction of autism based on eye-tracking scanpaths. International Journal of Psychosocial Rehabilitation, 24(05).

[2] Ahmed, Z. A., Albalawi, E., Aldhyani, T. H., Jadhav, M. E., Janrao, P., & Obeidat, M. R. M. (2023). Applying eye tracking with deep learning techniques for early-stage detection of autism spectrum disorders. Data, 8(11), 168.

[3] Ahmed, Z. A., Aldhyani, T. H., Alzahrani, E. M., Albalawi, E., Algarni, M. H., Jadhav, M. E., ... & Mehdi, A. (2024). Effectiveness of Histogram Equalization and Ensemble Deep Learning Techniques for Detecting Autism Using Eye-Tracking.

[4] Alarifi, H., Aldhalaan, H., Hadjikhani, N., Johnels, J. Å., Alarifi, J., Ascenso, G., & Alabdulaziz, R. (2023). Machine learning for distinguishing saudi children with and without autism via eye-tracking data. Child and Adolescent Psychiatry and Mental Health, 17(1), 112.

[5] AlQahtani, O., & Efstratopoulou, M. (2023). The UAE and Gulf countries' cultural characteristics and their influence on autism. Review Journal of Autism and Developmental Disorders, 1-5.

[6] Alsaidi, M., Obeid, N., Al-Madi, N., Hiary, H., & Aljarah, I. (2024). A convolutional deep neural network approach to predict autism spectrum disorder based on eye-tracking scan paths. Information, 15(3), 133.

[7] Alsharif, N., Al-Adhaileh, M. H., Al-Yaari, M., Farhah, N., & Khan, Z. I. (2024). Utilizing deep learning models in an intelligent eye-tracking

system for autism spectrum disorder diagnosis. Frontiers in Medicine, 11, 1436646.

[8] Baio, J. (2018). Prevalence of autism spectrum disorder among children aged 8 years—autism and developmental disabilities monitoring network, 11 sites, United States, 2014. MMWR. Surveillance Summaries, 67.

[9] Durango, I., Gallud, J. A., & Penichet, V. M. (2024). Human-Data Interaction Framework: A Comprehensive Model for a Future Driven by Data and Humans. arXiv preprint arXiv:2407.21010.

[10] Islam, M. F., Manab, M. A., Mondal, J. J., Zabeen, S., Rahman, F. B., Hasan, M. Z., ... & Noor, J. (2024). Involution Fused ConvNet for Classifying Eye-Tracking Patterns of Children with Autism Spectrum Disorder. arXiv preprint arXiv:2401.03575.

[11] Kanhirakadavath, M. R., & Chandran, M. S. M. (2022). Investigation of eye-tracking scan path as a biomarker for autism screening using machine learning algorithms. Diagnostics, 12(2), 518.

[12] Kohli, D., & Sharma, P. (2022). Various datasets for intelligent technologies in ASD detection [Data sets].

[13] LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. Proceedings of the IEEE, 86(11), 2278-2324.

[14] Mansour, T., & Bick, M. (2024). How can physicians adopt AI-based applications in the United Arab Emirates to improve patient outcomes?. Digital Health, 10, 20552076241284936.

[15] Oliveira, G., & Silva, H. (2021). Eye-tracking data for ASD diagnosis research [Data set].

[16] Palanichamy, N., Haw, S. C., & Ng, K. W. (2025). Adaptive Deep Convolution Neural Network for Early Diagnosis of Autism through Combining Personal Characteristic with Eye Tracking Path Imaging. JOIV: International Journal on Informatics Visualization, 9(1)

[17] Ridout, S., & Eldevik, S. (2024). Measures used to assess treatment outcomes in children with autism receiving early and intensive behavioral interventions: A review. Review Journal of Autism and Developmental Disorders, 11(3), 607-619.

[18] The UAE National Policy for Empowering People of Determination.

[19] Uddin, M. Z., Shahriar, M. A., Mahamood, M. N., Alnajjar, F., Pramanik, M. I., & Ahad, M. A. R. (2024). Deep learning with image-based autism spectrum disorder analysis: A systematic review. Engineering Applications of Artificial Intelligence, 127, 107185.

[20] World Medical Association. (2013). World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. Jama, 310(20), 2191-2194.

A Preliminary Comparative Analysis of National AI Strategies in the KSA and the UAE

Muhammd Tahir
College of Computer Science and Engineering
University of Jeddah
Jeddah, 21577, Saudi Arabia
mtyousaf@uj.edu.sa

Abstract—This paper analyzes the national AI strategies of the Kingdom of Saudi Arabia (KSA) and the United Arab Emirates (UAE) through document analysis. It compares their key official documents and describes key components, priority sectors, similarities, differences, and national visions. The results show that both countries: 1) prioritize AI in resources, energy, and healthcare, among other sectors, but differ in economic diversity, data governance, and investment targets; 2) are well-equipped to become regional and global AI leaders; and 3) lack in the areas of skills gaps, regulatory frameworks, and global competitiveness. This work concludes with recommendations for policymakers, researchers, and industry stakeholders.

Keywords—document analysis, AI strategies, KSA, UAE

I. INTRODUCTION

Recent trends in Artificial Intelligence (AI) are rapidly changing industries, economies, and societies worldwide. The research [7] shows "global GDP could be up to 14% higher in 2030 as a result of AI". This shows how important AI is for shaping future prosperity. Furthermore, nations are harnessing the potential of AI and trying to develop and implement national AI strategies. In this context, the UAE and the KSA each have launched ambitious national strategies. Both countries want to position themselves as global hubs for AI innovation and implementation.

The UAE's National AI Strategy 2031 [9] and KSA's National Strategy for Data & AI [8] are the important components of their broader national visions—UAE Centennial 2071 [10] and Vision 2030 [12], respectively.

These strategies aim to use AI for economic growth, tech progress, and better quality of life. For instance, UAE's National AI Strategy states that "AI has the potential to generate up to AED 335 billion in the UAE economy" [9] by 2031. Similarly, KSA targets "attracting 75 billion SAR in AI-related investments by 2030" [8]. These targets highlight the transformative potential of AI and by analyzing their strategies it is evident that KSA and UAE are positioning themselves to compete on the global stage. In this context, this work tries to answer the following general research question(s):

> What are the key components of the national AI strategies in the KSA and the UAE, their priority sectors, and how they compare in terms of similarities and differences?

By performing the document analysis of the UAE and KSA's AI strategies, this work; 1) contributes to the growing body of literature on national AI strategies, 2) provides a broad understanding of the key components of these national strategies and highlights the priority sectors, and 3) provides insights for policymakers, researchers, and industry stakeholders.

II. LITERATURE REVIEW

AI is proving to be a transformative force in the global economy. References [5] states that "Generative AI's impact on productivity could add trillions of dollars in value to the global economy." This immense economic potential has prompted governments worldwide to develop national AI strategies. The Organisation for Economic Co-operation and Development [6] reports that "many countries have announced national AI strategies and policy initiatives, which commonly aim to ensure a leadership position in AI."

In the Middle East, the UAE and KSA have emerged as regional leaders in AI adoption, having announced their national AI strategies [9][8], respectively. These strategies demonstrate a dedication to using AI as a driving force for the economy, technology, and humanity. Existing studies on national AI strategies often focus on developed economies such as the US, China, and the European Union, where AI adoption is more advanced. Reference [4] noted that China, Canada, and Finland were the first countries to announce their national AI strategies (section 7.3 of the report).

The US National AI Initiative Act [11] emphasizes on forming a national Artificial Intelligence advisory committee, directing the National Academies to study AI's impact on the workforce and China's Next Generation Artificial Intelligence Development Plan [13] aims to leverage China's strengths in AI to boost national competitiveness and improve public services. Similarly, the European Union's AI Strategy focuses on ethical AI and human-centric approaches [2].

Despite increasing research on national AI strategies, there is a limited comparative analysis of AI strategies in the Middle East region, especially regarding KSA and UAE. This highlights the need for further examination of their AI initiatives. A recent online report has examined national AI strategies (NASs) in the Arab region and identified six key themes, including AI's role in economic growth, job creation, regional leadership, and potential risks such as job displacement and loss of control over AI [3]. Their work differs in scope and methodology from our work. References [3] describe a generic regional perspective, while our work presents a detailed but concise comparative examination of the UAE and KSA's strategic AI documents. Our work also focuses on identifying the key components of national strategies, describing both similarities and differences. It also shows how these AI strategies align with the broader national visions (of both countries), including the emphasis on the priority sectors.

III. METHODOLOGY

Our work used a document analysis methodology [1] to systematically examine and interpret national AI strategy documents and to draw conclusions. This method includes reading and interpreting documents to extract meaningful data and also to identify emerging patterns.

A. Document Analysis Process

Document Selection: The two primary documents selected were: 1) The UAE National AI Strategy 2031 [9], and 2) the KSA National Strategy for Data & AI [8]. These were chosen based on their relevance as official, strategic-level documents of national AI priorities.

Initial Review: Documents were read to understand their structure and main focus areas.

Theme Identification and Coding: Recurring themes and strategic components were identified and manually coded. Key themes (components) identified included: vision, objectives, talent development, research and Innovation, data governance, economic impact, priority sectors among others.

Comparative Analysis: The analysis involved reviewing identified patterns and extracting relevant information on key themes or components (Table I). Furthermore, the comparative analysis highlighted similarities and differences based on key components.

Insights & Interpretation: This step provided insights into how each country is using AI to achieve national goals and address significant regional challenges.

B. Methodology Flowchart:

Below is a simple flowchart (Figure 1) outlining the steps followed in the document analysis.

Figure 1: Flowchart of methodology

IV. RESULTS

The UAE National AI Strategy 2031 describes a clear vision to position the country as a global leader in AI by 2031 [9] (page 2). The strategy focuses on eight strategic objectives, with key priorities including building a reputation as an AI destination, increasing competitive assets, and ensuring strong governance. The strategy also emphasizes the importance of talent development, public AI training, upskilling students, and government training programs (objective 5) for a future-ready workforce.

It is also a priority for the UAE to focus on research and innovation. This is planned to be achieved by establishing a National Virtual AI Institute. Another important objective is to attract global AI researchers to build world-leading research capabilities. Data governance is another key focus area, aiming to create a secure data infrastructure (objective 7) and promote ethical and responsible AI practices (objective 1). The UAE estimates that AI could contribute AED 335 billion to its economy by 2031, with major focuses on different priority sectors (Figure 2). Finally, the strategy also highlights the importance of strong governance, regulations, collaboration, and ensuring responsible AI development (objective 8).

On the other hand, the KSA National Strategy for Data & AI [8] emphasizes on positioning the Kingdom as a regional and global hub for Data & AI by 2030. Their key emphasis is on open data, talent development, and investment. Their strategy is aligned with KSA's Vision 2030 [12] for economic diversification and considers "data as the new oil" [8] (Narrative, Section 3). An important feature of KSA's AI strategy is its commitment to open data which they are planning to roll out across government entities after 2025 (Section 4.1). This may drive innovation and attract businesses especially those which rely on data for AI development.

Talent development (Section 4.1, and Objectives 2 & 3) is another key priority with the integration of AI into the education system (Section 4, Objective 2). This will further help to create specialized tracks, and offer professional training to build a talented local workforce that will be skilled in Data & AI (Section 4, Objective 2). Their AI strategy also emphasizes the need for a regulatory environment by keeping in view the ethical AI development and data protection. The strategy also highlights the importance of establishing innovation hubs and test-beds in smart cities like NEOM and promoting partnerships with global research institutions (Section 1.0). The KSA is aiming to attract 75 billion SAR in AI-related investments by 2030, which will be supported by targeted funds and investor support programs (Section 5.0).

TABLE III. COMPARATIVE DOCUMENT ANALYSIS OF UAE AND KSA NATIONAL AI STRATEGIES **(KEY COMPONENTS)**

Key Components	**UAE National AI Strategy 2031**	**KSA National Strategy for Data & AI**
Vision	Become a global leader in AI by 2031.	Become a global hub for Data & AI by 2030.
Alignment with National Vision	UAE Centennial 2071.	KSA Vision 2030.
Key Objectives	1. Build a reputation as an AI destination. 2. Increase competitive assets. 3. Ensure strong governance.	1. Position KSA as a worldwide hub for Data & AI. 2. The development of local talent. 3. Promote open data.
Talent Development	Focus on public AI training, upskilling students, and government training programs.	Integrate AI into education, create specialized tracks, and offer professional training.
Research & Innovation	Establish a National Virtual AI Institute and attract global researchers.	Create innovation hubs and test-beds in smart cities like NEOM.
Data Governance	Create a secure data infrastructure and promote ethical AI practices.	Roll out open data by default across government entities by 2025.
Economic Impact	AI could contribute AED 335 billion to the economy by 2031.	Target attracting 75 billion SAR in AI-related investments by 2030.
Priority Sectors	Energy, healthcare, tourism, logistics, and cybersecurity.	Energy, healthcare, mobility, education, and government services.
Governance & Regulation	National governance and international collaboration on AI ethics.	Develop regulatory frameworks and focus on ethical AI development.
Key Similarities	Both strategies focus on the talent development, research and innovation, and governance. They also highlight the economic venues i.e., the relevant financial contributions in AI. They are both aligned with their broader national visions i.e., UAE Centennial 2071 and KSA Vision 2030. Both of these strategies also focus on Energy and Healthcare as shown in Figure 2.	
Key Differences	The focus of the UAE is to build a global reputation for AI by improving its many sectors like tourism, logistics, and cybersecurity. While, the emphasis of the KSA is on open data, investment, and innovation, economic diversification, and reducing the dependence on oil revenues.	

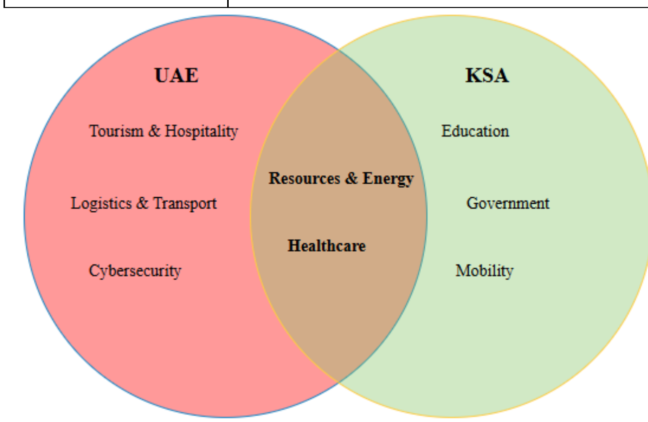


Figure 2: AI Priority Sectors: UAE vs KSA

V. DISCUSSION AND CONCLUSION

The analysis of the strategies of both countries show that they are committed to becoming global AI leaders, focusing on economic growth, research and innovation, and social development. This is crucial to build sustainable AI ecosystems. The focus of UAE's strategy is on building a global reputation for AI with its emphasis on multiple sectors like tourism and logistics. Similarly, the emphasis of KSA's AI strategy is on open data, investment and innovation which describes its goal of becoming a hub for Data & AI by 2030.

However, both countries face significant challenges that should be addressed in order to fully realize the potential of their AI strategies. For example, the UAE's reliance on global (expatriate) talent and KSA's skill gaps are critical barriers that requires to take initiatives for local future-ready workforces. Regulatory uncertainties and the need for policy frameworks are also other important challenges. The intense global competition from established AI leaders such as the US, China, and the EU further requires extra struggle for the path to global leadership in AI for both countries.

To overcome these challenges both countries must advance international collaboration with focus on local talent development, invest in research and innovation, and continuously update their AI strategies with respect to latest technological advancements.

The analysis of key documents also provides important insights for policymakers, researchers, and industry stakeholders to effectively apply AI in the UAE and KSA (Table II).

TABLE IV. KEY INSIGHTS FOR POLICY MAKERS, RESEARCHERS AND INDUSTRY STAKEHOLDERS

Stakeholder	Key Insights
Policymakers	1) Increase AI investments through incentives and local and global partnerships. 2) Create regulatory frameworks for ethical and responsible AI adaptation. 3) Develop AI talent through education and international collaboration.
Researchers	1) Analyze the effectiveness of AI strategy. 2) Explore sector-specific AI adoption, especially research collaboration in priority sectors. 3) Investigate open data initiatives and ethical concerns.
Industry Stakeholders	1) Leverage AI strategies for business and investment opportunities in priority sectors. 2) Align AI initiatives with national priorities and ensure regulatory compliance. 3) Invest in AI workforce development to address talent shortages.

By focusing on these insights, policymakers can refine AI policies, researchers can bridge knowledge gaps, and businesses can capitalize on emerging AI-driven opportunities, collectively providing a strong and sustainable AI ecosystem in both countries.

While no formal framework such as SWOT or PESTLE was applied in this initial analysis, the emergent themes offer a strong foundation for applying such frameworks in future research. For example, SWOT analysis could help evaluate the internal strengths and weaknesses of each strategy, alongside external opportunities and threats. Also, PESTLE framework could guide a more macro-level analysis focusing on Political, Economic, Social, Technological, Legal, and Environmental factors influencing AI strategy development.

REFERENCES

[1] Bowen, G. A. (2009). Document analysis as a qualitative research method. Qualitative Research Journal, 9(2), 27–40. https://doi.org/10.3316/QRJ0902027

[2] European Commission. (2020). White paper on artificial intelligence: A European approach to excellence and trust. https://ec.europa.eu

[3] Hendawy, M., & Kumar, N. (2024). AI in the national AI strategies of the Arab region (Critical Policy Analysis 2021–2025, July 2024). Arab Reform Initiative. https://s3.eu-central-1.amazonaws.com/storage.arab-reform.net/ari/2024/11/13112245/2024-08-09-EN-Menna.pdf

[4] Maslej, N., Zhang, B., Barron, B., & Etchemendy, J. (2024). The AI index 2024 annual report. Stanford University, Institute for Human-Centered AI. https://hai.stanford.edu/ai-index/2024-ai-index-report

[5] McKinsey & Company. (2021). The economic potential of generative AI: The next productivity frontier. https://www.mckinsey.com

[6] OECD. (2019). Artificial intelligence in society. OECD Publishing. https://doi.org/10.1787/eedfee77-en

[7] PwC. (2021). Sizing the prize: What's the real value of AI for your business and how can you capitalise? https://www.pwc.com

[8] Saudi Data and Artificial Intelligence Authority (SDAIA). (2020). National strategy for data & AI: Strategy narrative (October 2020). https://www.carringtonmalin.com/wp-content/uploads/2020/08/NDAIS-Strategy-Narrative-V2-19Oct20.pdf

[9] U.A.E. Government. (2018). UAE national strategy for artificial intelligence 2031. https://ai.gov.ae/wp-content/uploads/2021/07/UAE-National-Strategy-for-Artificial-Intelligence-2031.pdf

[10] U.A.E. Government. (2023). UAE Centennial 2071. https://u.ae/en/about-the-uae/strategies-initiatives-and-awards/strategies-plans-and-visions/innovation-and-future-shaping/uae-centennial-2071

[11] United States Congress. (2020). National Artificial Intelligence Initiative Act of 2020. https://www.congress.gov/bill/116th-congress/house-bill/6216/text

[12] Vision 2030 Kingdom of Saudi Arabia. (2016). KSA Vision 2030. https://www.vision2030.gov.sa/media/rc0b5oy1/saudi_vision203.pdf

[13] Xinhua News Agency. (2017). China's next generation artificial intelligence development plan. http://fi.china-embassy.gov.cn/eng/kxjs/201710/P020210628714286134479.pdf

Energy Efficiency of Residential Building using Decision Tree Classification

Fatmah Mohame Alkaabi
Business Analytics Programme
Abu Dhabi School of Management
Abu Dhabi, United Arab Emirates:
adsm-215057@adsm.ac.ae

Abstract— Building cooling demand is rising due to urbanization and climate change, especially in hot regions like Abu Dhabi[1]. Traditional HVAC systems are energy-intensive[2], increasing the need for innovative, efficient solutions. This study presents a decision tree–based model utilizing C4.5, CART, and Decision Stump algorithms to predict residential cooling load based on thermal and environmental data. Among these, CART achieved the highest predictive accuracy. The proposed model is transparent, interpretable, and cost-effective, supporting energy efficiency efforts. It can help reduce operational costs and guide decision-making for sustainable residential energy planning in rapidly urbanizing and climate-vulnerable regions [3] [4]

Keywords— Artificial Intelligence; CART; Decision Tree; Decision Stump; Cooling Load

I. BACKGROUND

Energy consumption in buildings constitutes a significant portion of global energy use, particularly in hot climates where cooling demands are intensified by climate change and urbanization. Traditional estimation methods often face challenges in capturing dynamic factors such as weather variability, building geometry, and occupant behaviour.

II. MOTIVATION

With increasing urbanization and climate change, there is a pressing need for accurate and efficient cooling load predictions in residential settings. Prior research demonstrates that machine learning can significantly improve energy efficiency and operational cost savings, especially in commercial sectors[5]. This study aims to extend those benefits to residential applications by leveraging data-driven approaches

III. PROBLEM STATEMENT

Despite advances in predictive modeling, current cooling load estimation techniques often lack transparency and adaptability. Furthermore, explainable artificial intelligence (XAI) remains underutilized in this domain, making it difficult for stakeholders to trust and implement AI-based solutions.

IV. RESEARCH QUESTIONS

How accurately can decision tree algorithms (C4.5, CART, and Decision Stump) predict cooling loads in residential buildings?
Can explainable models improve stakeholder trust and usability in energy forecasting?

V. OBJECTIVES

To develop a predictive, explainable decision tree–based model for cooling load estimation using thermal and environmental data.
To evaluate and compare the performance of C4.5, CART, and Decision Stump algorithms.
To provide interpretable results that support both residential users and policy decision-makers.

VI. CONTRIBUTIONS

This research contributes an interpretable and efficient modeling approach that enables accurate cooling load forecasting while enhancing transparency through XAI techniques. It supports energy efficiency strategies, reduces operational costs, and aligns with sustainability goals.

VII. SCOPE AND VALUE

The study focuses solely on decision tree (DT) algorithms and their ability to forecast cooling loads in residential buildings. Other machine learning models and broader energy systems are outside the scope. The proposed framework provides actionable insights for homeowners, energy providers, and policymakers to establish more innovative, AI-integrated energy strategies. Future research will address the implementation and evaluation of the developed system.

VIII. METHODOLOGY AND RESULTS

This study compares three decision tree (DT) algorithms—C4.5, CART, and Decision Stump—to predict cooling loads in residential buildings. While existing literature highlights the effectiveness of decision tree models in energy forecasting, most studies refer to general DT approaches or ensemble methods (such as Random Forest) without directly comparing these specific standalone classifiers. By focusing on interpretable DT algorithms, this study aims to evaluate which model offers the best trade-off between prediction accuracy, computational efficiency, and explainability.

Prior research supports this direction: Moon et al. (2024) demonstrated that ensemble-based DT models like Random Forest and Gradient Boosting yield high predictive accuracy for residential electricity forecasting, supported by SHAP for interpretability[6]. Similarly, decision tree–based models achieved strong performance ($R^2 > 0.95$) in predicting heating and cooling loads [4] [7]. However, comparative benchmarking of standalone DT algorithms such as C4.5, CART, and Decision Stump remains limited in energy-related applications,

highlighting the need for this study.

TABLE I. Key Aspects of Literature Review Summary

Author/Year	Purpose/Objective	Methodology	Findings
Moon et al. (2024)	Residential building electricity consumption forecasting using explainable AI.	Random Forest, Gradient Boosting, Decision Tree Bagging, SHAP for interpretability.	Random Forest and Gradient Boosting outperformed regression models, particularly in cooling load prediction.
Khorrami et al. (2024)	Comparative study on heating and cooling loads forecasting.	Decision Tree, Linear Regression, Neural Networks.	The decision Tree model achieved 98.96% accuracy for the heating load and 93.24% for the cooling load.
Moradzadeh et al. (2020)	Performance evaluation of machine learning for heating and cooling loads.	Decision Trees, Multilayer Perceptron (MLP), Support Vector Regression (SVR).	MLP achieved $R^2 > 0.95$, outperforming SVR in heating and cooling load predictions.

The methodology is guided by the Cross-Industry Standard Process for Data Mining (CRISP-DM) framework[8]. which is widely adopted in practical data science workflows for its structured and iterative development process[9].

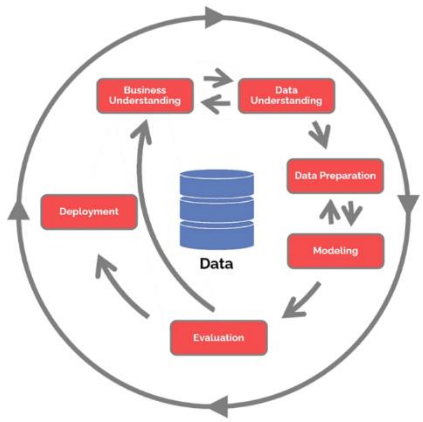


Fig. 1. CRISP-DM Process Model (Adapted from Satyasai, 2024).

- Data Source: This study utilizes the energy efficiency dataset retrieved from Kaggle, which includes 768 instances and eight building-related input features such as wall area, roof area, glazing area, and orientation[10].

- Tools: Model development was performed using standard machine-learning environments that support classification, feature selection, and evaluation.

- Preprocessing: Data cleaning, normalization, and feature selection were applied to improve model robustness.

- Modeling: The three DT algorithms—C4.5, CART, and Decision Stump—were trained and tested using a 70/30 split. Performance was assessed using accuracy, precision, recall, and F1-score.

This methodology answers:

- How do C4.5, CART, and Decision Stump compare in predicting residential cooling loads?

- Which model best balances performance and interpretability for real-world energy applications?

This short paper proposed a machine learning–based framework for predicting cooling load in residential buildings using three interpretable decision tree (DT) algorithms: C4.5, CART, and Decision Stump. The approach is grounded in the CRISP-DM methodology and leverages publicly available thermal and environmental building data. Although prior literature supports the utility of DT-based forecasting in energy modeling, this study contributes by comparatively analyzing specific DT variants that have not been widely benchmarked together. The proposed methodology supports energy efficiency efforts by enabling transparent and replicable decision-making that can inform residential energy planning and regulatory strategies.

To extend this research, future studies may consider:

- **Validation**: Validate predictive reliability by applying and testing the trained models on real-world residential energy datasets across various climate zones and building types.

- **Model Enhancement:** Explore hybrid models by integrating DT algorithms with ensemble learning or neural networks to improve performance and scalability.

- **Broader Applications**: Extend the methodology beyond residential buildings to other sectors—such as commercial real estate and infrastructure—where accurate energy forecasting supports sustainability, cost reduction, and operational planning.

This work serves as a foundation for developing practical tools to support long-term energy sustainability strategies.

To extend this research, future studies may consider:

- **Validation**: Validate predictive reliability by applying and testing the trained models on real-world residential energy datasets across various climate zones and building types.

- **Model Enhancement:** Explore hybrid models by integrating DT algorithms with ensemble learning or neural networks to improve performance and scalability.

- **Broader Applications**: Extend the methodology beyond residential buildings to other sectors—such as commercial real estate and infrastructure—where accurate energy forecasting supports sustainability, cost reduction, and operational planning.

This work serves as a foundation for developing practical tools to support long-term energy sustainability strategies.

REFERENCES

[1] Department of Energy Abu Dhabi. (2023). Sustainability Report 2023. Abu Dhabi: DOE.

[2] Ionescu, C., Baracu, T., Vlad, G. E., Necula, H., & Badea, A. (2015). The historical evolution of energy-efficient buildings. Renewable and Sustainable Energy Reviews, 49, 243–253. https://doi.org/10.1016/j.rser.2015.04.062

[3] Bekdaş, G., Aydın, Y., Işıkdağ, Ü., Sadeghifam, A. N., Kim, S., & Geem, Z. W. (2023). Prediction of the cooling load of tropical buildings with machine learning. Sustainability, 15(9061).

[4] Khorrami, B. M., Soleimani, A., Pinnarelli, A., Brusco, G., & Vizza, P. (2024). Forecasting heating and cooling loads in residential buildings using machine learning: A comparative study of techniques and influential indicators. Asian Journal of Civil Engineering, 25, 1163–1177.

[5] Zingre, K. T., Wan, M. P., & Ng, B. F. (2021). Prediction of the cooling load of tropical buildings with machine learning. Energy and Buildings, 252, 111403.

[6] Moon, J., Maqsood, M., So, D., Baik, S. W., Rho, S., & Nam, Y. (2024). Advancing ensemble learning techniques for residential building electricity consumption forecasting: Insight from explainable artificial intelligence. PLOS ONE, 19(11), e0307654.

[7] Moradzadeh, A., Mansour-Saatloo, A., Mohammadi-Ivatloo, B., & Anvari-Moghaddam, A. (2020). Performance evaluation of two machine learning techniques in heating and cooling loads forecasting of residential buildings. Applied Sciences, 10(11), 3829.

[8] Satyasai, U. S. (2024). What is CRISPR-DM? Retrieved April 28, 2025, from https://medium.com/@udaysrisatyasai/what-is-crisp-dm-cbdae4dfcc84

[9] Li, X., Zhang, C., & Zhuang, J. (2021). A systematic literature review on applying CRISP-DM in data mining. Proceedings of the 2021 International Conference on Intelligent Computing (ICIC), 245–251.

[10] Chowdhury, U. (2022). Energy Efficiency Data Set [Data set]. Kaggle. https://www.kaggle.com/datasets/ujjwalchowdhury/energy-efficiency-data-set

Predictive Analysis in Personalized Medicine - Research Directions

Raha AlAssaf
Abu Dhabi School of Management
Abu Dhabi, United Arab Emirates
raha_assaf@yahoo.com

Ishtiaq Rasool Khan
Abu Dhabi School of Management
Abu Dhabi, United Arab Emirates
i.khan@adsm.ac.ae

Abstract— This paper reviews how Artificial intelligence (AI) is used in personalized medicine through predictive analysis of genomic data. A thematic literature review of ten academic studies was conducted to examine the use of AI models in disease risk prediction, ethical challenges and the limitations of existing models trained primarily on Western genomic data. The findings highlight the lack of UAE specific genomic research and raise concerns about fairness and accuracy of models in local clinical settings. This paper identifies the research gaps, proposes future directions to improve the utilization of AI models in personalized medicine by leveraging UAE Genome Program.

Keywords—Artificial Intelligence, Predictive Analysis, Personalized Medicine, Genomic Data, Deep Learning, Machine Learning.

I. INTRODUCTION

Predictive analytics and Artificial are transforming genetic testing in healthcare. Machine learning, especially deep learning models, can analyse massive amounts of genomic data to improve disease detection, risk assessment and personalized treatment. However, the effectiveness of these models is heavily relied on the representation and quality of the training data, which often excludes non-Western populations.

Background and Motivation

UAE is one of the leading countries investing in genomic research to address the growing number of chronic diseases and rare genetic conditions. The UAE genome program aims to improve genomic healthcare by mapping the genetic makeup of the local population. This initiative creates opportunities for using AI to support personalized medicine. However, as genomic data grows, the needs to develop AI models tailored to Emirati population becomes urgent for effective healthcare service delivery.

Problem Statement

The real-world application of AI and predictive analytics in genomic medicine in the UAE remains limited. This is mainly because current AI models are developed using non-representative datasets, primarily from Western populations. These models may produce biased or inaccurate results for the UAE's diverse population. Challenges around the explainability of AI outputs, ethical governance, and local clinical integration limit the adoption of AI-driven personalized medicine in the region.

Aims and Objectives

This study aims to:

- Review the current use of AI and predictive analysis in personalized medicine with a focus on genomic data.
- Highlight research gaps in the applicability of AI models trained on Western specific datasets to the UAE context.
- Propose future directions to enhance model accuracy, transparency, and ethical governance within genomic healthcare systems.

Research Questions

RQ1: How are AI and predictive analytics currently being used in personalized genomic medicine?

RQ2: What are the limitations of applying Western-trained AI models in the UAE?

RQ3: What ethical and regulatory challenges must be addressed to support local implementation?

Contributions

This paper contributes by conducting a thematic literature review of AI applications in personalized genomic medicine, identifying the critical gaps related to region-specific data and ethical considerations, and recommending pathways for developing more inclusive and effective AI models for the UAE healthcare system.

Structure of the paper

The paper presents the methodology, summarizes key themes from literature, outlines results and research gaps, proposes future research directions, and concludes with recommendations for advancing AI in personalized medicine in the UAE.

II. METHODOLOGY

This study follows a qualitative thematic literature review approach to explore the application of Artificial Intelligence and predictive analysis in personalized genomic medicine, with a specific focus on the UAE context. A total of ten peer-reviewed articles published between 2022 and 2024 were selected from reputable journals using academic databases such as PubMed, Scopus, and Google Scholar.

The papers selection criteria included studies that applied AI or predictive analysis to genomic data, studies relevant to disease prediction, personalized medicine or ethical concerns, and English-language articles. Studies that focused on technical model development without healthcare application or unrelated to genomics, were excluded from the review.

Selected papers were thematically coded and grouped into four key themes: Predictive Analytics and AI in Genomic Healthcare, Deep Learning and Machine Learning Models, Ethical and Regulatory Challenges in Genomic AI, and Genomic Research including UAE Genome Program.

Table 1 summarizes the key characteristics of the selected studies, including study type, findings, limitations, and their relevance to personalized medicine.

TABLE V. SUMMARY OF REVIEWED STUDIES

Study	**Summary**		
	Study Type	**Findings /limitations**	**Relevance to Personalized Medicine**
Alzoubi et al. [1]	Experimental	High accuracy but black-box limitations	Demonstrates deep learning's potential in predicting complex genetic diseases, supporting personalized risk assessment.
Atieh et al. [2]	Review	Highlights need for regional policies	Emphasizes policy development necessary to support ethical AI use in regional personalized medicine initiatives.
Choon et al. [3]	Applied Research	Accurate NGS support, lacks UAE dataset	Applies AI to enhance rare disease diagnosis through NGS, contributing to personalized treatment pathways.
Hassan et al. [4]	Review/ Applied Analysis	Strong models, lacks clinic integration	Explores AI integration of genomic and clinical data, supporting individualized healthcare strategies.
Khan, Mohsen & Shah [5]	Scoping Review	Overfitting risk from small datasets	Highlights how biomarkers and AI can predict diabetes risk, a core aspect of personalized preventive medicine.
Quazi [6]	Review Article	Deep learning outperforms ML, lacks explainability	Compares model performance for analyzing genomic sequences, aiding in mutation detection for tailored care.
Rahma et al. [7]	Cross-sectional study	Shows lack of trust in AI tools	Assesses public perception of genomic medicine and genetic testing in the UAE, addressing a key aspect of societal readiness for the implementation of personalized medicine.
Vilhekar & Rawekar [8]	General Review	Need for explainable AI, general focus	Reviews AI's use for analyzing genetic data to enable more accurate personalized diagnosis, individual disease risk prediction, and tailored treatment strategies.
Zhang & Imoto [9]	Review/ Applied Research	Uses image conversion, lacks regional data	Applies image-based AI to classify genome variants, which is fundamental step in personalized medicine and can support diagnostic methods.
Zhang et al. [10]	Review/Applied Analysis	Biased data, strong sequencing tools	Analyzes big data applications in genomics, relevant to refining individual-level predictions in medicine.

III. LITERATURE REVIEW

A. Predictive Analytics and AI in Genomic Healthcare

The Predictive analytics and AI create opportunities, in the area of genomics, to enhance prediction of disease risk and personalized medications. But there are challenges that surround the use of these tools, caused by the variations in methodologies and sources of data. One challenge is the generalizability of AI models, in addition to limited clinical settings to apply these models.

Zhang et al. [10] and Hassan et al. [4] investigate how predictive analysis and AI models enhance disease risk assessment using genetic data, but in different ways. Zhang et al. [10] concentrate on analysis using big data and show that next generation sequencing (NGS) has an important role in enhancing rare disease predictions; whereas, Hassan et al. [4] focus on improving personalized medicine by the integration of biomedicals data, transcriptomics and clinical data.

The two studies highlight the important role of predictive analysis and AI in disease risk assessment, but their shortcomings show critical research gaps. Zhang et al. [10] face bias in the dataset and the models used in the study depend on Western-centric genomic data, which limits the generalizability to non-Western populations. On the other hand, Hassan et al. [4] underscore the absence of frameworks that regulate the integration of genetics AI models into clinics, which creates data privacy concerns.

A third study presented by Khan, Mohsen & Shah [5] relates closely to the work of Zhang et al. [10] and Hassan et al. [4] through a systematic review that focuses on genetics biomarkers for predicting diabetes. The review findings show that AI improves diabetes risk assessment and that multimodal diabetes prediction models perform higher than unimodal models. There were limitations of overfitting caused by small sample sizes as highlighted by Khan, Mohsen & Shah [5].

A. Deep Learning and Machine Learning Models

Multiple studies examine the role of deep learning in genomic medicine. Each study participates in the findings through focusing on different applications. Quazi [6] observes how CNNS, GANS and traditional machine learning models can improve the analysis of genomic data by conducting a comparative analysis of these models. The results prove that deep learning models perform better than traditional ML models in detecting mutation and analysing genome sequence.

Alzoubi et al. [1] propose a deep learning framework that can be utilized to predict disease risks. The study applies deep learning models on genome-wide association study (GWAS) datasets and proves that these models perform better than the traditional statistical techniques by achieving 94% accuracy in identifying complex genetic disorders. Both Quazi [6] and Alzoubi et al. [1] prove that deep learning models have high performance, however the researchers underscore that these models are black box functions that lack explainability.

Vilhekar & Rawekar [8] examine the application of deep and machine learning in different areas of drug repurposing and disease detection. The researchers underlines the strength of AI models in predicting genetic disorder diagnosis and personalized treatment, yet these models face shortcomings of transparency and validation. Vilhekar & Rawekar [8] do not focus on specific AI tool or specific disease and follow more general approach than Alzoubi et al. [1] and Quazi [6]. However, all of these studies conclude that AI need to be explainable and to be trained on specific region data for more reliability.

Choon et al. [3] explores databases powered by AI and deep learning models that help the Next Generation Sequencing (NGS) to diagnose rare diseases. The study shows that these models have the capabilities to accurately classify variations in genetics and improve the accuracy of disease diagnosis. Zhang & Imoto [9] employ deep learning in image processing which is used to transform genetic sequences to visual datasets used for variant classification. The potential of deep learning and machine learning in analysing genomics is emphasized by both Choon et al. [3] and Zhang & Imoto [9], however their studies lack training dataset from UAE genome program.

B. Ethical and Regulatory Challenges in Genomic AI

Many ethical challenges limit the implementation of predictive analytics and AI models in genomics despite their potential to enhance the genomic diagnosis and personalized medicine. Privacy concerns, public trust, biased datasets, models and lack of regulations are ones of these challenges. Rahma et. Al [7] and Ateia et al. [2] investigate the ethical concerns of applying AI in genomics, but each study looks into the problem from different motivations.

Rahma et al. [7] focus on the AI genetic testing in the UAE from public perception using mixed-method approach. A cross-sectional study design using questionnaires and validation using descriptive statistics. The findings of the study brief that lack of public trust in AI applications and concerns of possible misuse of private information hinder the adoption of AI and predictive analytics tools in genome medicine.

Ateia et al. [2] concentrate on regulatory challenges and policies for data sharing in regional genome programs. The study highlights the necessity to have a governance framework that regulates how AI use the genomics information and establishes politics for sharing data. The researchers underline that the Emirati Genome Program (EGP) aligns with highest ethical and governance standards and make sure that the data of all participants including the date of blood donations kept anonymous.

C. Genomic Research including UAE Genome Program

Western-centric datasets are used by most genetics AI models for training and validation, which limits the ability to use these models for non-western populations. To address this challenge, the UAE Genome Program develops AI models that use UAE-specific genetic data for training. Unfortunately, there is a lack of research in this field.

Atieh et al. [2] highlights the UAE Genome Program objectives to improve predictive medicine using AI. Rahma et al. [7] explore the attitudes of UAE people to the use of genomic AI. Both studies underline the necessity to develop AI models that are trained on UAE genomic data to improve predictions accuracy, public trust and unbiased results.

Results

The review of ten recent studies revealed the following key insights. “Fig. 1” illustrates the main themes identified across the reviewed studies.

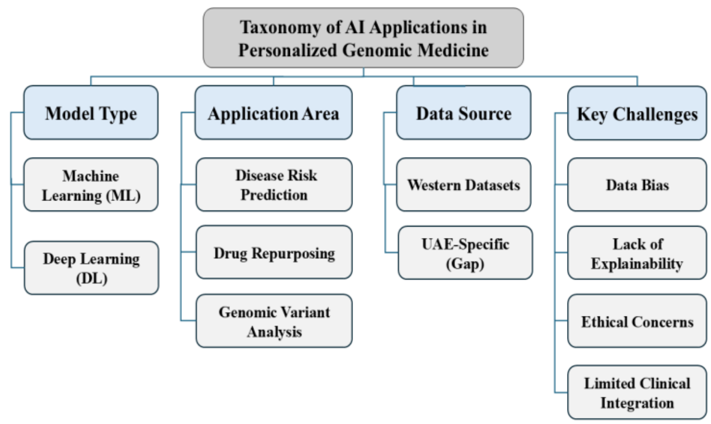


Fig. 1. Taxonomy of AI applications in personalized genomic medicine.

- AI models, especially those using deep learning, are effective in disease prediction using genetic data.
- Deep learning models perform better than traditional ML models in finding patterns in genomic data, but are hard to interpret in clinical settings.
- Several studies point out that people worry about privacy and the misuse of their genetic data. Policies and regulations are still under development.
- None of the studies applied data from UAE Genome Program.

IV. RESEARCH GAPS

The literature review found multiple gaps in the studies within the UAE context.

- Most AI models are trained on Western-centric datasets. These models may not fit into UAE population data, which limits their ability to be applied at local clinics and reduces their efficiency to UAE-specific population datasets.

- Deep learning models are unexplainable black box functions, therefore cannot be integrated into clinical processes. Clinicians are not able to interpret the risk assessments resulted from AI models, which makes their adoption in personalized medicine more difficult.
- Lack of policies and regulatory frameworks for using AI in genomic healthcare and personalized medicine, which leads to ethical concerns and people mistrust in such models regarding the misuse of their data.

Research Directions

Based on the research gaps presented in the literature, several research directions are proposed to focus future efforts to build reliable and effective AI-based genomic healthcare in UAE.

First, future research should train AI models using Emirati genomic data to improve accuracy. This will improve prediction accuracy and enhance the relevance of AI models in local healthcare settings.

Second, integrating Explainable AI (XAI) into predictive models to improve transparency and build trust among clinicians and patients. When healthcare professionals understand how these models generate the results, they are more likely to use them in decision-making.

Third, to develop a framework to govern the use of AI in genomic medicine in UAE, this should address data privacy, patient consent and authorization of AI models to ensure ethical implementation and build public trust in the genome program.

Finally, to close the gap between research and clinical practice. To ensure that AI models are applicable in real-world settings, pilot studies and clinical trials should be conducted in collaboration with hospitals to validate the reliability and effectiveness of AI applications in actual healthcare environment.

V. CONCLUSION

Artificial intelligence and predictive analysis are transforming personalized medicine by enhancing how health risks are predicted to be more accurate through the use of genomic data. However, most AI models are built on datasets that are not specific to population of the UAE. Key challenges identified in the literature include data bias, ethical concerns, limited model explainability, and lack of integration into clinical practice. There is an urgent need for UAE-trained AI models, explainable methods, and clear regulations. Future research should focus on utilizing local datasets, conducting clinical validation through pilot studies, and addressing public trust through transparent and ethical implementation. By bridging these gaps, researchers can support the development of more accurate and trusted AI models for personalized medicine and genomic healthcare system in the UAE.

REFERENCES

[1] Alzoubi, H., Alzubi, R., & Ramzan, N. (2023). Deep Learning Framework for Complex Disease Risk Prediction Using Genomic Variations. Sensors, 23(9), 4439. https://doi.org/10.3390/s23094439

[2] Ateia, M., Ogrodzki, P., Wilson, A., Ganesan, S., Halwani, R., Koshy, S., & Zaher, S. (2023). Population Genome Programs across the Middle East: Landscape, Challenges, and Opportunities. Biomed Hub, 8, 60–71. https://doi.org/10.1159/000530619

[3] Choon, Y. W., Choon, Y. F., Nasarudin, N. A., Al Jasmi, F., Remli, M. A., Alkayali, M. H., & Mohamad, M. S. (2024). Artificial intelligence and database for NGS-based diagnosis in rare disease. Frontiers in Genetics, 14, 1258083. https://doi.org/10.3389/fgene.2023.1258083

[4] Hassan, M., Awan, F. M., Naz, A., deAndrés-Galiana, E. J., Alvarez, O., Cernea, A., Fernández-Brillet, L., Fernández-Martínez, J. L., & Kloczkowski, A. (2022). Innovations in Genomics and Big Data Analytics for Personalized Medicine and Health Care: A Review. International Journal of Molecular Sciences, 23(9), 4645. https://doi.org/10.3390/ijms23094645

[5] Khan, S., Mohsen, F., & Shah, Z. (2024). Genetic biomarkers and machine learning techniques for predicting the onset and progression of diabetes mellitus: a scoping review. Cardiovascular Diabetology, 23(1), 80. https://doi.org/10.1186/s12933-024-02140-8

[6] Quazi, S. (2022). Artificial intelligence and machine learning in precision and genomic medicine. Medical Oncology, 39(8), 120. https://doi.org/10.1007/s12032-022-01711-1

[7] Rahma, A. T., Elhassan, I., Al Anouti, F., Ahmed, S. A., Ali, A., El Gamal, M., AbdelWareth, L. O., Bawazir, A., Elamin, M. A., Mahboub, B., & Tayoun, A. (2023). Knowledge, attitudes, and perceptions of genomic medicine among the general population in the United Arab Emirates. Human Genomics, 17(1), 63. https://doi.org/10.1186/s40246-023-00509-0

[8] Vilhekar, R. S., & Rawekar, A. (2024). Artificial Intelligence in Genetics. Cureus, 16(1), e52035. https://doi.org/10.7759/cureus.52035

[9] Zhang, Y. Z., & Imoto, S. (2024). Genome analysis through image processing with deep learning models. Journal of Human Genetics, 69(10), 519-525. https://doi.org/10.1038/s10038-024-01275-0

[10] Zhang, Y., Yu, J., Xie, X., Jiang, F., & Wu, C. (2024). Application of Genomic Data in Translational Medicine During the Big Data Era. Frontiers in Bioscience (Landmark Edition), 29(1), 7. https://doi.org/10.31083/j.fbl2901007

Artificial Intelligence in Oil & Gas Drilling Operations: Managing Key Performance Indicators through Machine Learning Predictive Model

Assad Sheraz
M.Sc. in Business Analytics
Abu Dhabi School of Management
Abu Dhabi, UAE
0009-0003-5272-4158

Shahjahan Khan
Drilling and Completion Department
MOL Pakistan Oil and Gas B.V
Islamabad, Pakistan
0009-0006-6260-6560

Adnan Manzoor
Drilling and Completion Department
MOL Pakistan Oil and Gas B.V
Islamabad, Pakistan

Abstract—This paper presents the applications of Artificial Intelligence (AI) and Machine Learning (ML) in optimizing oil and gas drilling operations and then presents how macro-level drilling Key Performance Indicators (KPIs) are managed through machine learning predictive model. Drilling operations are challenging in nature whereas drilling in Pakistan's northern region is more challenging due to complex geological and tectonic conditions. CRISP-DM methodology is adopted to develop ML models for the prediction of macro-level drilling KPIs such as Dry Hole Drilling Days (DHDD), Dry Hole Drilling Cost (DHDC), and Clean Time (CT). Historical data of the Wells located in Pakistan's northern region are used to train, test, and evaluate ten different ML algorithms. Two top performing models for each KPI are then used to develop ML based predictive calculator in Google Colab. Deployment results from unseen data of six Wells show that the predictions are either better or complimenting the traditional methods. These deployment results show the effectiveness of ML methods and potential of AI & ML in enhancing drilling efficiency, reducing costs, and overcoming challenges of drilling oil & gas Wells.

Keywords—Artificial Intelligence, Machine Learning, Drilling, Oil & Gas Wells, Key Performance Indicators, Dry Hole Drilling Days, Dry Hole Drilling Cost, Clean Time

I. INTRODUCTION

This paper intends to summarize the ongoing applications of Artificial Intelligence (AI) and Machine Learning (ML) in oil & gas drilling operations, followed by a real case example of the deployment of machine learning predictive models for managing the macro-level drilling Key Performance Indicators (KPIs). Previous applications of AI and ML in the field of oil & gas drilling mainly focused on micro-level KPIs like drilling parameters, downhole vibration and Rate of Penetration (ROP), whereas this study focuses on the prediction of macro-level drilling KPIs.

Oil & gas industry play an important role in global economy and serve as a backbone for the world's development and growth. In oil & gas industry, drilling is the Capital Expenditure (CAPEX) intensive and one of the most important domain as it enables the extraction of oil & gas present deep below the subsurface. Oil and gas drilling operations are complex in nature due to their reliance on subsurface data. Drilling in Pakistan's northern region is even more challenging since the geology in this region is complex and tectonically active [1]. Even the Wells located close to each other are entirely different in operational difficulties and hence simple prediction methods produce inaccurate results [2]. In order to tackle these drilling challenges, proactive approach is needed and hence accurate prediction and optimization of Key Performance Indicators (KPIs) is very important. There are two main categories of drilling KPIs, which are micro-level KPIs and macro-level KPIs [3].

RELATED WORK

Artificial intelligence (AI) was applied to predict gamma-ray (GR) logs in real time during drilling operations, addressing delays caused by logging-while-drilling (LWD) tools measuring already-penetrated formations. Using surface drilling parameters and 4609 data entries from three different Wells, Support Vector Machine (SVM) and Random Forests (RF) models were developed and validated. SVM was better than RF, having correlation coefficient (R) of 0.98 and an average absolute percentage error of 1.42%. These models helped to predict real-time GR and identify formation lithology more efficiently [4]. AI-based workflow was created to enhance drilling performance by controlling and reducing the drill-string vibrations. Support Vector Regression (SVR), Multi-Layer Perceptron (MLP), and Decision Tree Regression (DTR) were utilized to predict & optimize key drilling parameters like Bit RPM, Rate of Penetration (ROP), and torque. SVR was the best model and helped to achieve 43% increase in ROP and reduced the drilling time from 66 to 31 hours [5].

Mechanical Specific Energy (MSE) optimization was carried out through machine learning for improving the drilling performance in one of the offshore field. Random Forest (RF) proved to be the best model for ROP. It also performed better than traditional methods and helped to make real-time adjustments that saved time and cost [6]. In order to detect downhole vibrations (axial, torsional, and lateral) using real-time Rig surface data, machine learning models were developed.

5750 drilling data points were used to develop Radial Basis Function (RBF), Support Vector Machine (SVM) and Adaptive Neuro-Fuzzy Inference System (ANFIS). This resulted in achieving high accuracy with R values ranging from 0.91 to 0.98 and average absolute percentage error from 1.1% to 7.3%. In addition to this, validation was also done through the unseen data. This system provided real-time vibration monitoring and helped to overcome the failures and reduce costs [7].

For extracting data from Rig reports, a conversational AI chatbot using generative AI was designed. It was designed through large language models (LLMs) and machine learning models (ML) to answer the queries related to drilling. In addition to this, it also performed diagnostic analysis and provided operational recommendations. This work showed that how generative AI can change energy sector operations by improving decision-making [8]. In order to perform directional drilling tool selection like Rotary Steerable Systems (RSS) or Positive Displacement Motors (PDM) a machine learning model using offset Wells data was developed. Multiple input variables like Rig hydraulics, inclination-ROP relationships, and tripping quality were used. XGBoost proved to be the best model for this task [9]. Machine learning was applied on data from four Wells to predict & optimize drilling parameters like Weight on Bit (WOB), Flowrate (FR), and Rotation per Minute (RPM). This helped to improve drilling efficiency with the increase in Rate of Penetration (ROP) [10].

Later part of this paper presents the deployment results of the machine learning predictive model developed for macro-level drilling KPIs.

II. METHODOLOGY

A. Machine Learning Predictive Model For Macro-Level Drilling Kpis

A machine learning model is developed in order to predict macro-level drilling KPIs like Dry Hole Drilling Days (DHDD), Dry Hole Drilling Cost (DHDC), and Clean Time (CT). CRISP-DM (Cross Industry Standard Process for Data Mining) methodology is applied, utilizing historical data from Wells drilled in Pakistan’s complex geological region. CRISP-DM (Cross Industry Standard Process for Data Mining) is shown in Figure 1. Multiple input variables are used in order to develop these machine learning models. Ten machine learning algorithms are trained, tested, and evaluated where Support Vector Machine (SVM) & Random Forest (RF) remained the best models for DHDD, Random Forest (RF) & Grid Search CV XGBoost remained best for DHDC and Stacking (SVM Base & LR Meta) & Grid Search CV XGBoost remained best for CT. Single algorithm could not best predict all the KPIs and hence ML based predictive calculator is developed based on two top performing algorithm for each KPI. Best and second best algorithm can vary based on the data used for training, testing and evaluating the machine learning model [11].

B. DEPLOYMENT

In order to manage the Key Performance Indicators (KPIs) through the developed machine learning predictive model, a machine learning model based calculator is developed in Google colab. In order to make the calculator user friendly, unseen input data for carrying out the prediction is also asked to be inserted in the form of csv prediction file. As shown in Figure 2, DataF.csv is the file used to input the data for training, testing and evaluating the machine learning algorithms, whereas Prediction.csv is the file used to input unseen data for prediction. Snapshot of the prediction file showing the required input variables is shown in Figure 3. Once the model is run, it train, test and evaluate the model based on DataF.csv file uploaded in it and presents the two best machine learning models for each of the KPI. Thereafter, using the Prediction.csv file it produces the prediction results for each KPI separately from the two best algorithms.

III. RESULTS & DISCUSSION

A. Best Performing Ml Models

Data used in deployment phase for training, testing and evaluating the machine learning model, resulted in Support Vector Machine (SVM) & Random Forest (RF) as the best performing models for DHDD with the error percentages of 14.2% & 14.9% respectively, Random Forest (RF) and Grid Search CV XGBoost as the best performing models for DHDC with the error percentages of 11.2% & 14.5% respectively and Stacking (SVM Base & LR Meta) & Grid Search CV XGBoost as the best performing models for CT with the error percentages of 10.56% & 10.64% respectively. Summary of these results is shown in Figure 4.

B. Deployment Results

During the deployment phase, six Wells are selected, which are not part of the data that is used to train, test and evaluate the machine learning models. Well 1 to 4 are already drilled, so their planned as well as the actual results are available for comparison, whereas Well 5 & 6 are not drilled completely, so only planned values; from traditional method; are available. ML based calculator made prediction for each of the three KPIs using two best performing models for that specific KPI as mentioned in the previous section. Prediction results for all three KPIs are shown in Table 1. Whereas, Table 2 is showing the planned and actual values for these three KPIs for the same six Wells. Planned values mentioned in this table are from the traditional methods currently in practice, whereas actual values for these KPIs are from the actual drilling of these Wells.

Figure 5 presents the comparison of traditional method & ML model results for DHDD. It can be observed that actual values are exceeding the planned values for all the four Wells which are already drilled, which shows the need of improvement in traditional method adopted for calculating the planned values of DHDD. Predicted values from best performing machine learning model SVM are matching with planned values in Well 1, 2, 5 and 6, whereas they are matching with the actual values in Well 3 and 4. Figure 6 presents the comparison of traditional method & ML model results for DHDC. It can be observed that actual DHDC on two instances is less than the planned DHDC, despite of the fact that actual DHDD are more than the planned DHDD for all the Wells here. Where Actual vs Planned DHDC trend is matching with the Actual vs Planned DHDD trend, predicted values from best performing model RF is matching more closely with the actual DHDC values. Figure 7 presents the comparison of traditional method & ML model results for

CT. Since the error percentages for both the models are almost same, the predicted values of both the models Stacking (SVM Base & LR Meta) & Grid Search CV XGBoost are matching closely with the actual CT values.

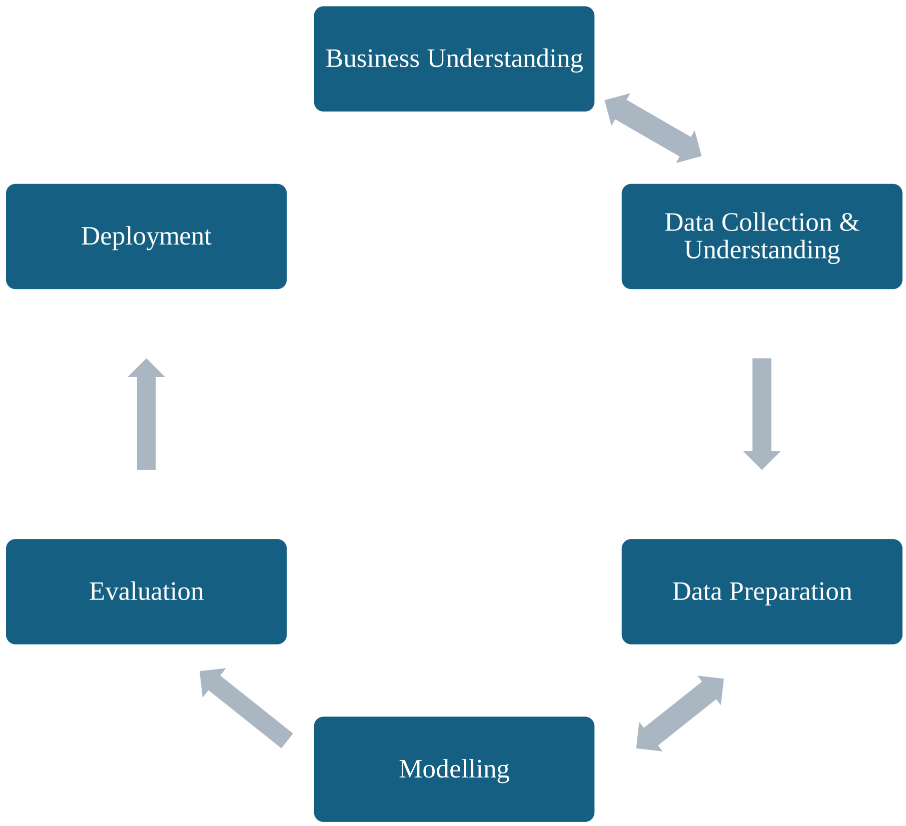


Fig. 1. CRISP-DM Framework [11]

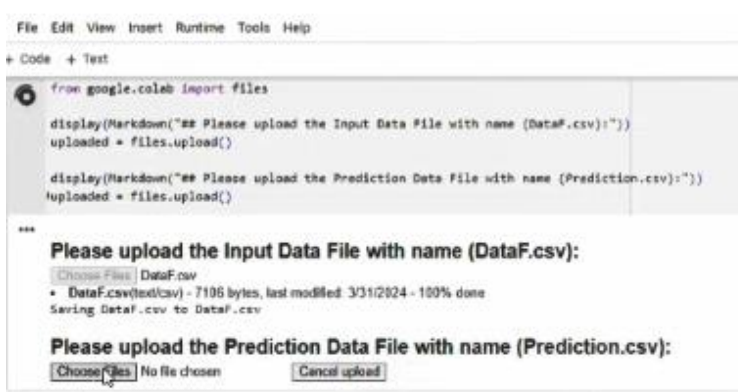


Fig. 2. Snapshot from ML Based KPI Prediction Calculator

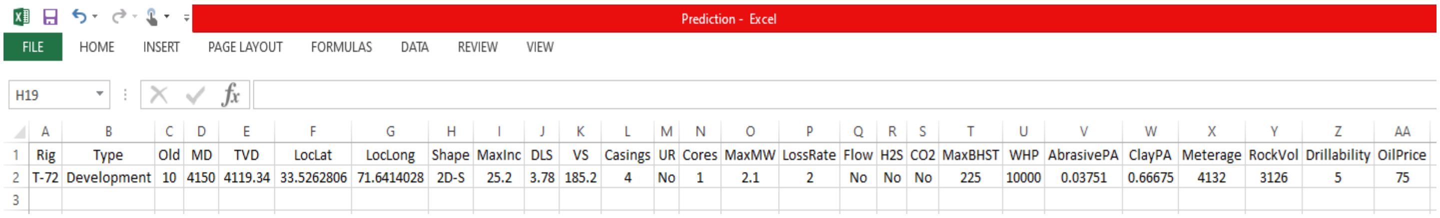


Fig. 3. Snapshot of the Prediction.csv File

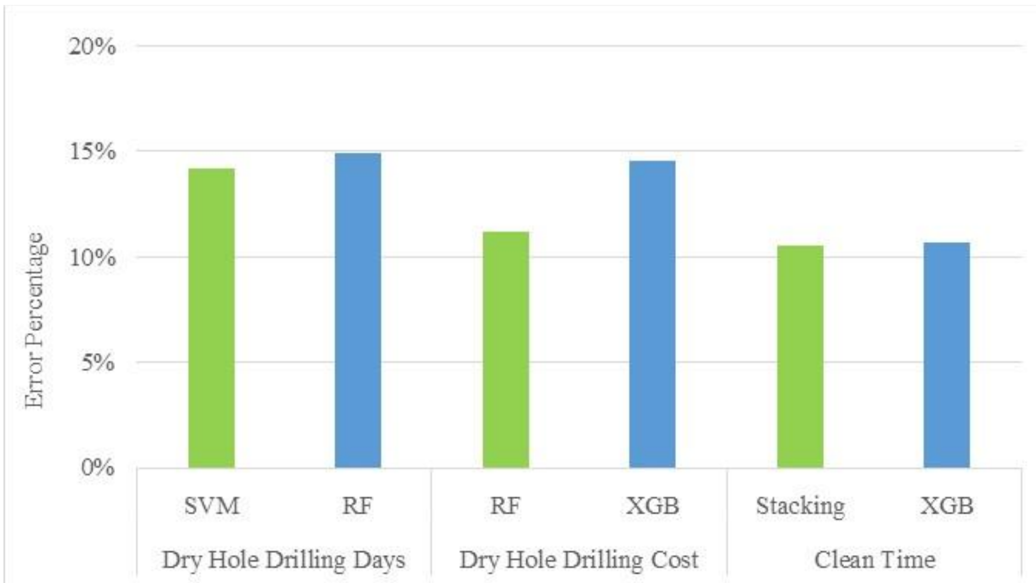


Fig. 4. Two Best Performing Machine Learning Models for each KPI

TABLE VI. PREDICTION RESULTS OF TWO BEST PERFORMING ML MODELS

Well Name	**DHDD (days)**		**DHDC (m$)**		**CT (days)**	
	SVM	**RF**	**RF**	**XGB**	**Stacking**	**XGB**
Well 1	135	194	15.4	7.2	115	101
Well 2	218	266	21.7	18.2	162	189
Well 3	199	280	22.7	16.1	214	174
Well 4	233	272	21.7	19.9	191	205
Well 5	108	203	16.0	13.1	121	112
Well 6	116	113	10.1	11.1	88	100

TABLE VII. PLANNED & ACTUAL VALUES OF THE KPIS

Well Name	**DHDD (days)**		**DHDC (m$)**		**CT (days)**	
	Planned	**Actual**	**Planned**	**Actual**	**Planned**	**Actual**
Well 1	133	162	12.5	13.6	Not Applicable	101
Well 2	225	235	17.4	15.5		179

Well 3	160	183	15.6	12.7		175
Well 4	208	249	17.6	17.8		181
Well 5	129	Not Available	13.8	Not Available		Not Available
Well 6	110	Not Available	10.5	Not Available		Not Available

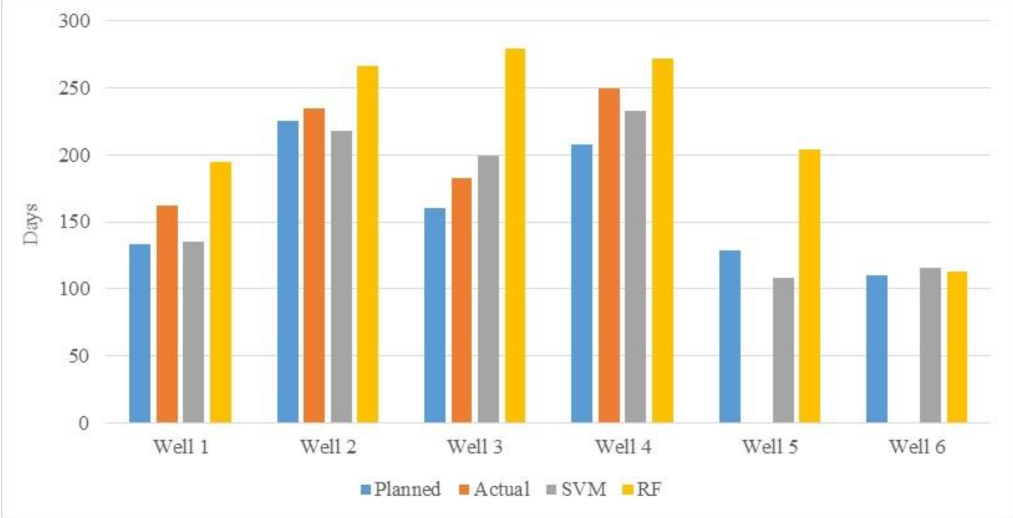


Fig. 5. Comparison of Traditional Method & ML Model Results for DHDD

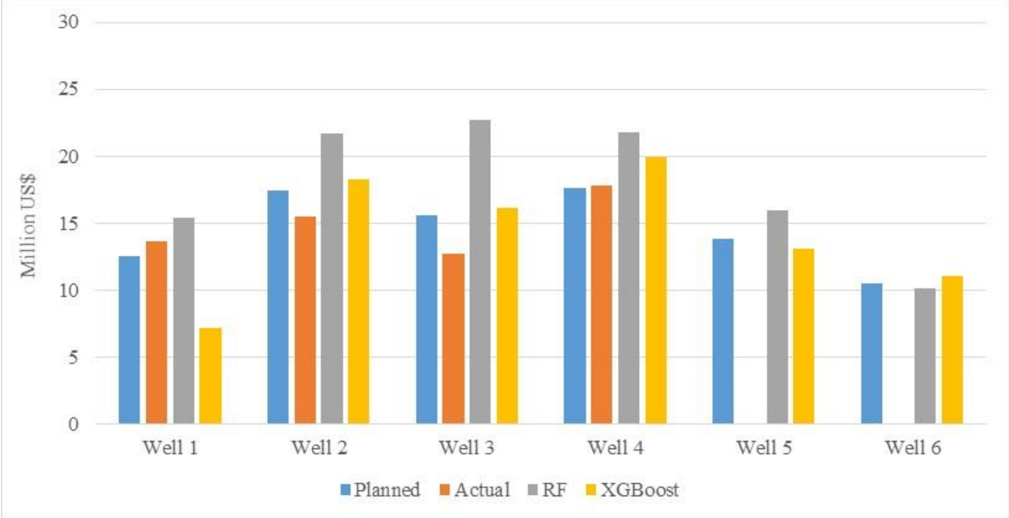


Fig. 6. Comparison of Traditional Method & ML Model Results for DHDC

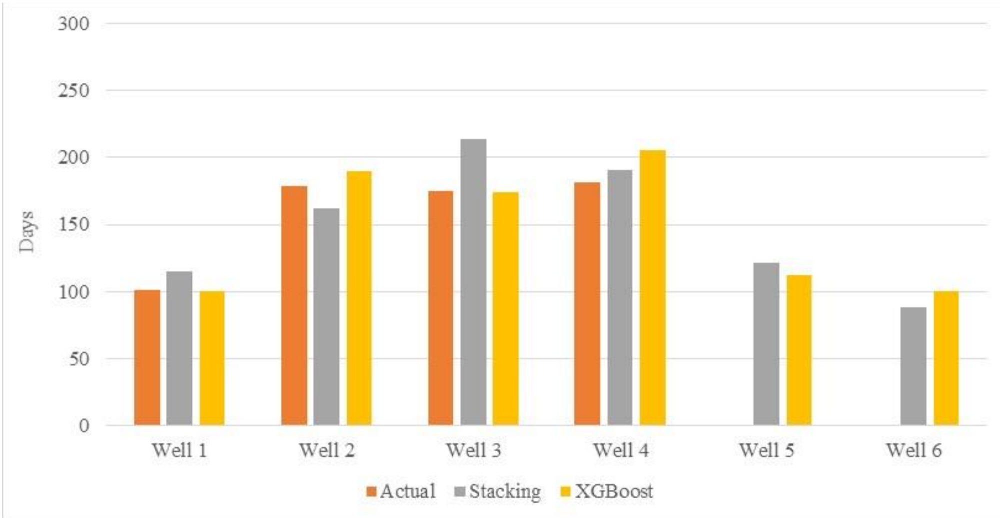


Fig. 7. Comparison of Traditional Method & ML Model Results for CT

IV. CONCLUSION

This paper shows the application of AI & ML in optimizing oil and gas drilling operations, particularly in complex geological regions like Pakistan's northern region. Machine learning predictive models are developed to predict macro-level drilling KPIs such as DHDD, DHDC, and CT and accordingly then deployed by predicting the KPIs of the unseen data. Development of a user friendly ML based predictive calculator in Google Colab assists in the deployment phase. The results from six Wells show that the predictions from these models either complement or perform better than the traditional methods, highlighting the potential of AI & ML in enhancing drilling efficiency and reducing the cost. This paper focuses on macro-level drilling KPIs and hence contributes to AI & ML growing applications in the field of oil & gas drilling which has previously focused on micro-level drilling KPIs.

Since this study is based on the data availability from one of the fields from Pakistan's northern region. Future study can focus on training, testing, evaluating and deploying the same ML based predictive calculator on other fields and geological settings. This study will help to expand this application in other drilling regions. Moreover, future study can also be conducted to reduce variance between the planned / actual DHDC and ML predicted DHDC value by including some additional cost related input variables. This will help to refine DHDC prediction.

ACKNOWLEDGMENT

Authors express their heartfelt gratitude to MOL Pakistan for their support in publishing this research paper and for sharing the actual Wells data, which played a pivotal role in this research study and develop machine learning models. Special thanks to Abu Dhabi School of Management (ADSM) for providing the platform to conduct this research study.

REFERENCES

[1] S. Khan, A. N. Awan, Z. Sarfraz, and I. Muhammad, "Geomechanics Application in the North of Pakistan: Unlocking Drilling Excellence-A Comprehensive Analysis of 10 Years," in SPE/PAPG Pakistan Section Annual Technical Conference, pp. SPE-219509, 2023.

[2] J. Haneef and A. Sheraz, "Development of well complexity calculator and its integration into standard well engineering management system/well delivery system," Journal of Petroleum Exploration and Production Technology, vol. 12, no. 6, pp. 1727–1757, 2021.

[3] J. Haneef and A. Sheraz, "A Comparative Analysis of Well Key Performance Indicators (KPIs) with Well Complexities Using Well Complexity Calculator," Arabian Journal for Science and Engineering, pp. 1–18, 2022.

[4] F. Ibrahim and S. Elkatatny, "Real-time GR logs estimation while drilling using surface drilling data; AI application," Arabian Journal for Science and Engineering, vol. 47, no. 9, pp. 11187–11196, 2022.

[5] F. S. Boukredera, M. R. Youcefi, A. Hadjadj, C. P. Ezenkwu, V. Vaziri, and S. S. Aphale, "Enhancing the drilling efficiency through the application of machine learning and optimization algorithm," Engineering Applications of Artificial Intelligence, vol. 126, p. 107035, 2023.

[6] Nautiyal and A. K. Mishra, "Machine learning application in enhancing drilling performance," Procedia Computer Science, vol. 218, pp. 877–886, 2023.

[7] R. Saadeldin, H. Gamal, and S. Elkatatny, "Detecting downhole vibrations through drilling horizontal sections: Machine learning study," Scientific Reports, vol. 13, no. 1, p. 6204, 2023.

[8] Singh, T. Jia, and V. Nalagatla, "Generative AI enabled conversational Chatbot for drilling and production analytics," in Abu Dhabi International Petroleum Exhibition and Conference, p. D021S065R002, 2023.

[9] M. Nour, S. K. Elsayed, and O. Mahmoud, "A supervised machine learning model to select a cost-effective directional drilling tool," Scientific Reports, vol. 14, no. 1, p. 26624, 2024.

[10] Elahifar, "Real-Time Artificial Intelligence-Enhanced Machine Learning Technique for Accurate Drilling Parameter Prediction and Optimization," in SPE Annual Technical Conference and Exhibition, p. D021S017R001, 2024.

[11] Sheraz, S. Khan, N. Abdelhamid, and A. Manzoor, "Application of Machine Learning for Comprehensive Predictive Modelling of Drilling Key Performance Indicators Using Historical Drilling Data - Collaborative Case Study Between Academia and Industry," in Oman Petroleum & Energy Show. SPE, 2025.

Artificial Intelligence in Russia: Evaluating Yandex GPT-5 and GigaChat 2.0 MAX in Global Context

Arsenii Moshninov
Graduate School of Business
HSE University
Moscow, Russia
akmoshninov@edu.hse.ru

Mikhail Komarov
Graduate School of Business
HSE University
Moscow, Russia
mmkomarov@hse.ru

Abstract— Study evaluates Russian large language models Yandex GPT-5 and GigaChat 2.0 MAX against global leaders—ChatGPT 4o, Grok 3, and DeepSeek V3—in business-relevant tasks. This study uses a mixed-method approach, combining custom-designed tasks, LM Arena data, and standardized benchmarks to evaluate logic, mathematics, programming, content creation, and image generation. Russian models match global peers in logic and mathematics, lag slightly in programming and creativity due to resource constraints, and excel in native language tasks, partially outperforming Western models. Their performance aligns closely with international standards, with potential to lead in Russian tasks, reflecting Russia's strategic AI.

Keywords—AI development, Russian AI, artificial intelligence, Yandex GPT-5, GigaChat 2.0 MAX

I. INTRODUCTION

The global artificial intelligence (AI) landscape has entered an era of geopolitical stratification, wherein national AI systems increasingly embody regional technological priorities and cultural contexts. Russia's pursuit of technological sovereignty has caused the rapid advancement of domestic large language models (LLMs), notably Yandex GPT-5 and GigaChat 2.0 MAX, developed by Russia's leading IT companies, Yandex and Sber, respectively. Yandex GPT- 5, created by the search engine giant Yandex, is optimized for Russian language processing and integrates with the Alice voice assistant [1]. GigaChat 2.0 MAX, developed by the fintech leader Sber, excels in Russian language contexts [2]. These advancements occur amidst international sanctions and an increased emphasis on import substitution. They foster a unique AI innovative ecosystem that balances limited resources with strategic investments in technology and linguistics [3,4,5].

Recent analyses by CNA Russia Studies indicate that 78% of Russian AI research funding is channeled through state-aligned entities, emphasizing applications that ensure information control and linguistic autonomy [3]. This state-driven approach contrasts sharply with the commercial paradigms prevalent in the West, yet it delivers significant outcomes. For instance, GigaChat 2.0 MAX achieves an 80% accuracy rate on the Massive Multitask Language Understanding (MMLU) benchmark within Russian contexts, surpassing certain Chinese models in regional assessments. Similarly, Yandex's integration of the Alice voice assistant with GPT-5 Pro highlights Russia's tailored approach to human-computer interaction, with a focus on optimizing Cyrillic script processing and Slavic linguistic structures [6,7].

Western advancements show scale and versatility, while Russian LLMs like Yandex GPT-5 prioritize efficiency, achieving 92% of GPT-4's coding accuracy with 40% fewer computational resources—a critical advantage given Russia's reliance on sanctioned hardware alternatives [8,9,10].

Russia's strategic AI positioning blends technological resilience with cultural specificity. The 2024 CSET analysis shows only 12% of Russian studies address cross-cultural evaluation, compared to 34% in U.S. research [4,5].

Comprehensive comparisons between Russian and global LLMs in business applications are scarce. Existing research on Russian AI, like Yandex GPT-5 and GigaChat 2.0 MAX, is outdated due to rapid innovation, lacking rigorous scrutiny. This study addresses this gap by systematically comparing their strengths and limitations for organizational performance

This research contributes to AI in business management by introducing a novel evaluation framework accounting for cultural and linguistic variations, comparing Russian and Western LLMs to describe global AI dynamics, and providing actionable insights for enterprises and policymakers in cross-cultural contexts [11,12,13].

II. METHODOLOGY

This study adopts a mixed-methods framework to assess the performance of five LLMs in business-relevant tasks—ChatGPT 4o by OpenAI [14], Grok 3 by xAI, a company associated with Elon Musk [15], DeepSeek V3, a non-commercial model from a Chinese startup [16], GigaChat 2.0 MAX by Sber, a fintech leader excelling in Russian-language contexts [2], and Yandex GPT-5 by Yandex, a search engine giant optimized for Russian language processing [1]. The approach combines custom-designed tasks, LM Arena data, and standardized benchmarks for a robust evaluation, as illustrated in the methodology flowchart (Figure 1) [17].

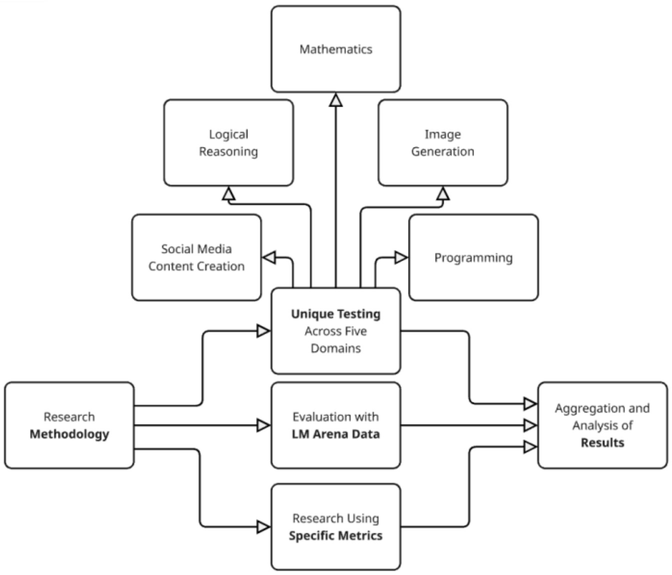


Fig. 1. Flowchart with research methodology

We designed tasks across five categories: logical reasoning (e.g., letter counting, scored 0 or 1) [18], mathematics (e.g., equation solving, partial credit), programming (JavaScript 'Snake' game, 0–2 for functionality), social media (Instagram post, 1–5 for creativity), and image generation (futuristic city, assessed by three for accuracy, averaged).

The authors acknowledge uncertainty regarding whether the results may have been affected by the service's concurrent load, potentially reducing response accuracy and task-solving efficacy, thus constituting a study limitation.

LM Arena comparison utilizes the platform's Elo rating system to rank ChatGPT 4o, Grok 3, and DeepSeek V3 based on user interactions across various Russian, English tasks. For models not included in LM Arena—GigaChat 2 MAX and YandexGPT 5—simulated pairwise comparisons or analogous methods will estimate their performance relative to listed models [17].

Standardized benchmarks provide objective metrics: MMLU (Russian and English) for language comprehension, GSM8K and MATH for mathematical reasoning, HumanEval for coding, and IFEVAL (Russian and English) for instruction-following [17,19,20,21,22].

Evaluation and comparison integrate results from all components. Custom task scores (binary or scaled), LM Arena rankings, and benchmark outcomes are aggregated to compare model performance holistically, identifying strengths and limitations for business uses like customer service and content creation.

III. RESULTS

All models demonstrated proficiency in logical reasoning and mathematical tasks, achieving maximum scores across the board (e.g., 5/5 points each for ChatGPT 4o, Grok 3, DeepSeekV3, Yandex GPT, and GigaChat 2 MAX in Table I), which indicates a strong baseline capability in structured problem-solving based on accuracy and efficiency. In programming tasks, ChatGPT 4o, Grok 3, and DeepSeekV3 distinguished themselves with fully functional solutions that met all specified requirements, showcasing robust code generation abilities and earning top marks (e.g., 4/4 points each) for executable and efficient code. Conversely, Yandex GPT's website design appeared outdated, lacking modern aesthetic and functional standards, scoring lower (e.g., 2/4 points) due to its less effective output, while GigaChat 2 MAX failed to produce operational game code, yielding non-executable results and receiving minimal points (e.g., 1/4). For social media content creation, DeepSeekV3 exhibited exceptional creativity, crafting a post that excelled in engagement, relevance, and narrative depth, outperforming its peers in this qualitative assessment and securing the highest score (e.g., 5/5 points), while others scored based on comparative quality (e.g., 3–4/5). In the image generation task, ChatGPT 4o accurately integrated text and imagery as per the prompt, earning a high score (e.g., 5/5 points) for precision and coherence. Yandex GPT and GigaChat 2 MAX achieved moderate success, though their results were marred by inaccuracies in text rendering and visual coherence, resulting in moderate scores (e.g., 3–4/5 points). Notably, DeepSeekV3 lacks image generation functionality entirely, a limitation inherent to its current design, rendering it unable to compete in this domain and thus scoring 0/5. Despite these variations, the performance gap between Russian models (GigaChat 2 MAX and Yandex GPT) and global models was not substantial, suggesting comparable overall competence as reflected in Table I's aggregate scores. However, ChatGPT 4o and Grok 3 slightly led the evaluation, due to their consistent performance, balancing technical precision with creative adaptability, achieving higher total scores (e.g., 18–19/20 vs. 14–16/20). Detailed results and scores for each model across the evaluated tasks are presented in Appendix A, a brief version in Table I.

TABLE VIII. BRIEF VERSION AI TEST

Task Type	**ChatGPT-4o**	**Grok-3**	**DeepSeekV3**	**GigaChat 2 MAX**	**Yandex GPT**
Logical reasoning tasks	5	5	5	5	5
Mathematical tasks	3	3	3	3	3
Programming	4	4	4	2	2
SMM	4	4	4	5	3
Image generation	5	4	3	4	4
Total score	**21**	**20**	**17**	**16**	**17**

Fig. 2. Brief version AI test, Table I

Chatbot Arena data (no Russian LLMs) shows ChatGPT-4o ranking 2nd overall (1466), Grok-3 4th (1404), and DeepSeek-V3 5th (1370). In Russian tasks, ChatGPT-4o and Grok-3 tie for 1st (1430, 1426), DeepSeek-V3 ranks 3rd (1376). Grok-3 excels in creativity (2nd, 1406), ChatGPT-4o in coding (1st, 1425), while DeepSeek-V3 lags in math (6th, 1341), indicating ChatGPT-4o's versatility across tasks [17] (Appendix B).

Benchmark analysis reveals GigaChat 2 MAX's competitive edge in Russian-language tasks, scoring 80.46 on MMLU (RU), surpassing GPT-4o (80.00) and DeepSeek-V3 (73.74). In English tasks, GPT-4o leads with 88.70 on MMLU (EN), followed by GigaChat 2 MAX (86.00) and DeepSeek-V3 (85.24). GigaChat 2 MAX excels in instruction-following for Russian (IFEVAL RU: 83.62), while DeepSeek-V3 dominates in coding (HumanEval: 91.46) and English instruction-following (IFEVAL EN: 92.21). These results highlight the strength of Russian LLMs in native language contexts, positioning GigaChat 2 MAX as a formidable contender globally [23] (Appendix C).

IV. CONCLUSION

This study has demonstrated that the latest Russian LLMs, Yandex GPT-5 and GigaChat 2.0 MAX, exhibit performance levels closely aligned with global models such as DeepSeek V3, ChatGPT 4o, and Grok 3. The research reveals that Russian LLMs achieve competitive outcomes across diverse business-relevant tasks. Notably, their proficiency in logical reasoning and mathematics matches that of global counterparts, though they lag slightly in programming and creative domains, reflecting resource constraints rather than inherent limitations.

In Russian-language tasks, global models showcased varying strengths based on LM Arena data (excluding Russian models). ChatGPT 4o and Grok 3 tied for the highest performance, achieving Elo ratings of 1430 and 1426, respectively, while DeepSeek V3 trailed at 1376. This indicates that ChatGPT 4o and Grok 3 excel in versatility and adaptability across linguistic contexts, positioning them as leaders among global models for Russian tasks [13].

Comparatively, Russian models partially outperform Western counterparts in native language benchmarks. GigaChat 2.0 MAX, for instance, scored 80.46 on MMLU (RU), surpassing ChatGPT 4o (80.00) and DeepSeek V3 (73.74), and excelled in instruction-following (IFEVAL RU: 83.62). These results underscore the tailored efficacy of Russian models in their linguistic domain, driven by a focus on cultural and functional specificity.

The unique geopolitical and geotechnical motivations of Russian developers—emphasizing technological sovereignty and linguistic autonomy—suggest a trajectory where Russian models may soon outpace global analogs in Russian-language tasks. With state-driven investments and progress evident in current benchmarks, models like GigaChat 2.0 MAX are steadily advancing toward this potential, leveraging regional priorities to bridge performance gaps.

Globally, Russian LLMs stand roughly equivalent to their international peers, mirroring the developmental pace of Western AI systems. This parity highlights Russia's strategic resilience in AI innovation despite sanctions and limited resources. This study does not address the data infrastructure challenges of LLMs, which remain beyond its scope. Future research should explore longitudinal comparisons and the broader impact of national strategies on AI ecosystems, offering deeper insights into their implications for global business management.

REFERENCES

[1] Yandex. (n.d.). Alice voice assistant. Retrieved April 22, 2025, from https://alice.yandex.ru/

[2] Sber. (n.d.). GigaChat. Retrieved April 22, 2025, from https://giga.chat/

[3] Center for Naval Analyses. (2024). Artificial intelligence and autonomy in Russia. Retrieved April 4, 2025, from https://www.cna.org/centers-and-divisions/cna/sppp/russia-studies/artificial-intelligence-and-autonomy-in-russia

[4] Geneva Internet Platform. (2024, February 11). Russia struggles to catch up in global AI race. Retrieved April 4, 2025, from https://dig.watch/updates/russia-struggles-to-catch-up-in-global-ai-race

[5] Popkova, E. G., & Stefanovic, M. (2024). [TRENDS OF THE AI ECONOMY IN RUSSIA]. Journal of Artificial Intelligence, 1(1), 1–10. https://jai.aspur.rs/archive/v1/n1/1.pdf

[6] Evrim Ağacı. (2025, February 25). YandexGPT-5 Pro revolutionizes AI with Alice integration. Retrieved April 4, 2025, from https://evrimagaci.org/tpg/yandexgpt-5-pro-revolutionizes-ai-with-alice-integration-227608

[7] The Hans India. (2025, March 13). Sber presents new neural network GigaChat 2.0. Retrieved April 4, 2025, from https://www.thehansindia.com/business/sber-presents-new-neural-network-gigachat-20-953634

[8] Konaev, M., & Gilli, A. (2020). Russian AI research 2010 to 2018. Center for Security and Emerging Technology. https://cset.georgetown.edu/wp-content/uploads/CSET-Russian-AI-Research-2010-to-2018-2.pdf

[9] Future Skills Academy. (2025, February 25). GPT-5 vs GPT-4. Retrieved April 4, 2025, from https://futureskillsacademy.com/blog/gpt-5-vs-gpt-4/

[10] Fello AI. (2024, August). Claude AI: Everything you need to know. Retrieved April 4, 2025, from https://felloai.com/2024/08/claude-ai-everything-you-need-to-know/

[11] Chen, D., Esperança, J. P., & Wang, S. (2022). The impact of artificial intelligence on firm performance: An application of the resource-based view to e-commerce firms. Frontiers in Psychology, 13, Article 884830. https://doi.org/10.3389/fpsyg.2022.884830

[12] DataCube Research. (2024, June). Russia generative AI market: Analysis 2019-2032 (Report AI4212). Niche Industry Monitor.

[13] Business Insider. (2024). US, China compete for AI dominance while Russia's model lags behind. Retrieved April 4, 2025, from https://www.businessinsider.com/us-china-compete-ai-dominance-while-russia-model-lags-behind-2025-2

[14] OpenAI. (n.d.). ChatGPT. Retrieved April 22, 2025, from https://chatgpt.com/

[15] xAI. (n.d.). Grok. Retrieved April 22, 2025, from https://grok.com/

[16] DeepSeek. (n.d.). DeepSeek chat. Retrieved April 22, 2025, from https://chat.deepseek.com

[17] Chatbot Arena. (2024). LM Arena Leaderboard. Retrieved April 4, 2025, from https://arena.lmsys.org/

[18] Hendrycks, D., Burns, C., Basart, S., Zou, A., Mazeika, M., Song, D., & Steinhardt, J. (2020). Measuring massive multitask language understanding. arXiv. https://arxiv.org/abs/2009.03300

[19] Cobbe, K., Kosaraju, V., Bavarian, M., Chen, M., Jun, H., Kaiser, L., Plappert, M., Tworek, J., Hilton, J., Nakano, R., Hesse, C., & Schulman, J. (2021). Training verifiers to solve math word problems. arXiv. https://arxiv.org/abs/2110.14168

[20] Jagtap, A. D., Shin, Y., Kawaguchi, K., & Karniadakis, G. E. (2022). Deep Kronecker neural networks: A general framework for neural networks with adaptive activation functions. Neurocomputing, 468, 165–180.

[21] Chen, M., Tworek, J., Jun, H., Yuan, Q., Pinto, H. P. de O., Kaplan, J., Edwards, H., Burda, Y., Joseph, N., Brockman, G., Ray, A., Puri, R., Krueger, G., Petrov, M., Khlaaf, H., Sastry, G., Mishkin, P., Chan, B., Gray, S., … Zaremba, W. (2021). Evaluating large language models trained on code. arXiv. https://arxiv.org/abs/2107.03374

[22] Ouyang, L., Wu, J., Jiang, X., Almeida, D., Wainwright, C., Mishkin, P., Zhang, C., Agarwal, S., Slama, K., Ray, A., Schulman, J., Hilton, J., Kelton, F., Miller, L., Simens, M., Askell, A., Welinder, P., Christiano, P. F., Leike, J., & Lowe, R. (2022). Training language models to follow instructions with human feedback. Advances in Neural Information Processing Systems, 35, 27730–27744. https://proceedings.neurips.cc/paper_files/paper/2022/hash/b1efde53be364a73914f58805a001731-Abstract-Conference.html

[23] T-J.ru. (2025, March 13). Sber GigaChat 2. Retrieved April 4, 2025, from https://t-j.ru/news/sber-gigachat-2/

Appendix A

Type	Task	ChatGPT 4o	Grok 3	DeepSeekV3	GigaChat 2 MAX	Yandex GPT
Logical reasoning tasks	Letter count in word	1	1	1	1	1
	Next number in sequence	1	1	1	1	1
	Logical deduction about tails	1	1	1	1	1
	Calculate average speed	1	1	1	1	1
	Anagram verification	1	1	1	1	1
Total Logical		**5**	**5**	**5**	**5**	**5**
Mathematical tasks	Solve linear equation	1	1	1	1	1
	Find function derivative	1	1	1	1	1
	Prove irrationality	1	1	1	1	1
Total mathematical		**3**	**3**	**3**	**3**	**3**
Programming	Create course website HTML	2	2	2	2	1
	Develop JavaScript snake game	2	2	2	0	1
Total programming		**4**	**4**	**4**	**2**	**2**
SMM	Write eco-product Instagram post	**4**	**4**	**5**	**3**	**3**
Image generation	Generate futuristic city image	**5**	**4**	**0**	**3**	**4**
Total score		**21**	**20**	**17**	**16**	**17**

Fig. 3. Detailed results and scores for each model across the evaluated tasks

Appendix B

Fig. 4. LM Arena data – Russian Tasks score [17].

Category

Russian

Apply filter

Style Control | Show Deprecated

Russian Prompts

#models: 193 (87%) #votes: 298,922 (11%)

Rank* (UB)	Delta	Model	Arena Score	95% CI	Votes	Organization	License
1	0	Gemini-2.5-Pro-Exp-03-25	1467	+27/-28	505	Google	Proprietary
1	1	GPT-4.5-Preview	1430	+23/-18	1323	OpenAI	Proprietary
1	1	ChatGPT-4o-latest (2025-03-26)	1430	+36/-30	394	OpenAI	Proprietary
1	1	Grok-3-Preview-02-24	1426	+21/-20	1235	xAI	Proprietary
2	3	Gemini-2.0-Pro-Exp-02-05	1416	+14/-13	2402	Google	Proprietary
2	3	Gemini-2.0-Flash-Thinking-Exp-01-21	1404	+12/-14	2724	Google	Proprietary
3	2	DeepSeek-V3-0324	1376	+32/-35	318	DeepSeek	MIT

Fig. 5. LM Arena data – Mathematical tasks score [17].

Category

Math

Apply filter

Style Control | Show Deprecated

Math

#models: 219 (99%) #votes: 378,834 (13%)

Rank* (UB)	Delta	Model	Arena Score	95% CI	Votes	Organization	License
1	0	Gemini-2.5-Pro-Exp-03-25	1414	+22/-21	606	Google	Proprietary
2	0	GPT-4.5-Preview	1377	+15/-12	1248	OpenAI	Proprietary
2	5	DeepSeek-R1	1359	+17/-15	1339	DeepSeek	MIT
2	6	o1-2024-12-17	1355	+14/-12	2960	OpenAI	Proprietary
2	12	o3-mini-high	1354	+15/-15	1678	OpenAI	Proprietary
2	3	DeepSeek-V3-0324	1341	+40/-29	284	DeepSeek	MIT
3	14	o3-mini	1348	+10/-12	2574	OpenAI	Proprietary
3	-1	Grok-3-Preview-02-24	1348	+13/-15	1214	xAI	Proprietary
3	2	Gemini-2.0-Flash-Thinking-Exp-01-21	1341	+13/-11	2734	Google	Proprietary
3	8	o1-preview	1339	+9/-10	5052	OpenAI	Proprietary
3	2	Gemini-2.0-Pro-Exp-02-05	1332	+12/-11	2381	Google	Proprietary
3	-1	ChatGPT-4o-latest (2025-03-26)	1330	+24/-26	516	OpenAI	Proprietary

Fig. 6. LM Arena data – Creative tasks score [17].

Category: Coding

Apply filter: Style Control | Show Deprecated

Coding: whether conversation contains code snippets

#models: 221 (100%) #votes: 551,698 (19%)

Rank* (UB)	Delta	Model	Arena Score	95% CI	Votes	Organization	License
1	1	ChatGPT-4o-latest (2025-03-26)	1425	+25/-19	838	OpenAI	Proprietary
1	0	Gemini-2.5-Pro-Exp-03-25	1425	+20/-21	970	Google	Proprietary
1	1	Grok-3-Preview-02-24	1411	+13/-10	2200	xAI	Proprietary
1	1	GPT-4.5-Preview	1403	+12/-12	2272	OpenAI	Proprietary
1	4	DeepSeek-V3-0324	1387	+25/-21	553	DeepSeek	MIT

Fig. 7. LM Arena data – Coding tasks score [17].

Category: Creative Writing

Apply filter: Style Control | Show Deprecated

Creative Writing

#models: 221 (100%) #votes: 433,213 (15%)

Rank* (UB)	Delta	Model	Arena Score	95% CI	Votes	Organization	License
1	0	Gemini-2.5-Pro-Exp-03-25	1458	+19/-20	943	Google	Proprietary
2	0	Grok-3-Preview-02-24	1406	+12/-12	2186	xAI	Proprietary
2	0	ChatGPT-4o-latest (2025-03-26)	1399	+21/-19	735	OpenAI	Proprietary
2	0	GPT-4.5-Preview	1392	+15/-17	2245	OpenAI	Proprietary
2	3	Gemini-2.0-Pro-Exp-02-05	1390	+10/-13	3608	Google	Proprietary
2	3	Gemini-2.0-Flash-Thinking-Exp-01-21	1388	+11/-11	4053	Google	Proprietary
2	3	DeepSeek-V3-0324	1386	+21/-24	472	DeepSeek	MIT
7	0	DeepSeek-R1	1357	+14/-15	2591	DeepSeek	MIT
7	4	Gemma-3-27B-it	1355	+14/-17	1543	Google	Gemma
8	0	Gemini-2.0-Flash-001	1349	+12/-11	3403	Google	Proprietary

Fig. 8. LM Arena data – English tasks score [17].

Category

Overall

Apply filter

Style Control Show Deprecated

Overall Questions

#models: 221 (100%) #votes: 2,829,853 (100%)

Rank* (UB)	Rank (StyleCtrl)	Model	Arena Score	95% CI	Votes	Organization	License
1	1	Gemini-2.5-Pro-Exp-03-25	1440	+8/-8	5121	Google	Proprietary
2	2	ChatGPT-4o-latest (2025-03-26)	1406	+10/-7	4080	OpenAI	Proprietary
2	4	Grok-3-Preview-02-24	1404	+6/-5	11601	xAI	Proprietary
2	2	GPT-4.5-Preview	1398	+7/-6	11754	OpenAI	Proprietary
5	7	Gemini-2.0-Flash-Thinking-Exp-01-21	1380	+4/-5	23834	Google	Proprietary
5	4	Gemini-2.0-Pro-Exp-02-05	1380	+4/-4	20293	Google	Proprietary
5	4	DeepSeek-V3-0324	1370	+9/-12	2840	DeepSeek	MIT
7	5	DeepSeek-R1	1359	+6/-6	13836	DeepSeek	MIT

Fig. 9. LM Arena data – Overall score [17].

Appendix C

Category	Benchmark Name	GigaChat 2 MAX	Qwen 2.5 72B	Llama 3.3 70B	GPT-4o	DeepSeek-V3
General Knowledge	MMLU (RU)	80,46	78,30	65,08	80,00	73,74
	MMLU (EN)	86,00	83,85	78,57	88,70	85,24
Mathematics	GSM8K	95,68	95,07	92,87	95,00	94,99
	MATH	77,26	78,74	62,80	76,60	85,48
Coding	HumanEval	87,20	86,60	86,00	84,00	91,46
Instruction Following	IFEVAL (RU)	83,62	84,27	75,12	80,24	84,37
	IFEVAL (EN)	89,99	90,43	90,83	88,51	92,21

Fig. 10. tests of models based on metrics by task category [23].

An AI Roadmap in Predicting the Performance of SME in UAE Procurement Industry

Suhaila B. Al Shamsi
Business Analytics Programme
Abu Dhabi School of Management
Abu Dhabi, United Arab Emirates:
adsm-215098@adsm.ac.ae

Abstract— This research aims to address the challenges faced by the small and medium enterprises (SMEs) in the UAE procurement industry through the development of an artificial intelligence (AI) roadmap based on decision tree (DT) classification and data-driven decision-making approaches. Current procurement methods often reply to manual processes that are time-consuming and prone to bias, leading to inefficiencies and misused opportunities for SMEs. This study proposes a structured AI roadmap that utilizes hybrid classification techniques to assess various strengths and weaknesses of the SMEs, ensuring transparency and objectivity in decision-making processes. Employing quantitative methods, the research will analyze data collected from SMEs to develop and evaluate the AI model. The intelligent approach proposed in this research aims to enhance the operational efficiency and competitiveness of SMEs by facilitating informed decision making and strategies planning. This study is significant as it contributes to economic development objectives by demonstrating how AI can be effectively enhanced into local procurement practices while addressing unique challenges faced by SMEs in the UAE.

Keywords— Artificial Intelligence; Small and Medium Enterprises. Decision Tree, Procurement, United Arab Emirates

I. INTRODUCTION

Small and medium Enterprises (SMEs) play a vital role in the economic landscape of the UAE, contributing significantly to the nation's GDP and employment rates. However these enterprises face numerous challenges in the procurement sectors, which can hinder their operational efficiency and competitiveness. The dynamic nature of the market, coupled with the complexities of procurement processes, often leaves SMEs at a disadvantage compared to large firms. Factors such as limited access to resources, lack of technological assistance, and inefficiencies in decision-making processes exacerbate these challenges [34] .

Recent studies indicate that many SMEs struggle with outdated procurement methods that rely heavily on manual processes, leading to delays and biases in decision-making [21]. The reliance on traditional procurement practices not only consumes time and resources but also limits the ability of the SMEs to respond swiftly to market changes, ultimately impacting their growth and sustainability. Furthermore, the lack of advanced data analytics and predictive capabilities prevents SMEs from identifying opportunities and threats in a timely manner.

In the UAE, where the SME sector represents approximately 75% of the GDP and 95% of employment, the need for effective procurement strategies is critical [34]. However, many SMEs remain unaware of how to leverage artificial intelligence (AI) and machine learning technologies to optimize their procurement processes. Current methods of identifying and addressing procurement inefficiencies often do not incorporate data-driven insights, which can lead to missed opportunities for improvement.

Traditional approaches to procurement evaluation typically rely on historical data and manual assessments, which may overlook the complexities of SME operations and the unique challenges they face [43]. As a result, there is a pressing need for structural AI roadmap that not only predicts the performance of the SMEs in procurement but also provides actionable insights to enhance decision-making capabilities.

This research aims to develop an AI roadmap tailored specifically for SMEs in the procurement industry, utilizing decision tree (DT) classification and explainable AI approaches. The proposed roadmap will serve as a strategic tool to identify the strengths and weaknesses of SMEs, enabling them to make more informed decisions based on real time data. By implementing advanced analytics and machine learning techniques, this research seeks to empower SMEs to overcome procurement challenges and enhances their operational efficiency.

The significance of this research lies in potential to transform the procurement landscape for SME in UAE. By leveraging AI technologies SMEs can streamline their procurement processes, reducing biases in decision-making. And ultimately improve their competitive positioning in the market. This research also acknowledges the importance of interpretability in AI systems, ensuring that stakeholders can understand, and trust decisions made by these advanced technologies.

In summary, the identification of effective procurement strategies for SMEs is essential for their growth and sustainability. This research will address the gap in existing knowledge by providing a comprehensive AI roadmap that facilitates the development of AI into procurement processes, ultimately contributing to a broader economic development objective of the UAE.

II. LITERATURE REVIEW

The procurement landscape for small and medium enterprises (SMEs) in the UAE is characterized by several changes, primarily due to the fast-paced digital market. SMEs are under pressure to remain competitive, necessitating transformative procurement strategies that ingrate modern technologies such as artificial intelligence (AI). This literature review aims to highlight the key factors influencing the SME performance in procurement, emphasizing the importance of aligning procurement processes with contemporary business practices.”.

A comprehensive analysis reveals several critical factors that affect SME performance in procurement. These factors are categorized into regulatory, contractual, procedural, support, operational and external features. Understanding these elements is essential for developing strategies that enhance SME competitiveness and participation in procurement processes.

High-quality procurement regulations are fundamental for SME success. Regulations must align with internationally accepted practices to foster an environment conductive to competition. Clear guidelines and specifications are necessary to help SMEs navigate procurement opportunities without misunderstanding that can lead to delays [20]. Contractual features such as the size and structure of contracts significantly influence SMEs ability to participate in procurement. Smaller contracts or those subdivided into manageable lots enhance accessibility for SMEs. Procedural frameworks must strike a balance between fostering competition and ensuring fair participation with timely payments being crucial for maintaining operational sustainability (2016) [17]. Support features, including suppliers’ assistance and timely feedback, empower SMEs to navigate complex procurement processes effectively. Operational features encompass the adoption of best practices, such as green purchasing and ethical standards, which not only improve the procurement efficiency but also enhance the reputation of the SMEs, fostering long-relationships [45]. External influences such as access to finance and economic conditions play a significant role shaping SME procurement capabilities. Recognizing these factors enables stakeholders to create an environment that better supports the SME participation in the procurement processes [17]. Integrating AI technologies into SMEs operations has merged as a pivotable factor in enhancing operational efficiency and economic performance. Various framework for AI adoption highlights the need for tailored approaches that consider the specific contexts of SMEs, ensuring successful assistance and impactful outcomes [4].

The literature identifies gaps that are considered as critical regarding the practical implementation of AI in SME procurement processes. While existing studies emphasize the benefits of AI, there is a need for comprehensive research that addresses the unique challenges SMEs face, providing actionable insights and roadmaps for effective AI assistance in procurement [2].

TABLE I. Key Aspects of Literature Review Summary

Objective	**Methodology**	**Findings**	**Discussion**	**Citation**
Analyze the impact of regulatory features on SME participation in procurement.	Comparative analysis of procurement regulations.	High-quality regulations enhance SME participation and success in securing contracts.	Effective regulatory frameworks reduce uncertainties and foster competition for SMEs.	Hoekman & Taş (2022)
Investigate the importance of clarity in procurement specifications for SMEs.	Qualitative analysis of procurement practices.	Clear guidelines improve SMEs' understanding and performance in procurement.	Vague specifications hinder SMEs' ability to meet standards, emphasizing the need for clarity.	van Scheers (2016)
Explore operational features that enhance procurement efficiency in SMEs.	Case studies and best practice analysis.	Adoption of green purchasing and ethical standards boosts competitiveness and reputation.	Operational excellence can improve procurement effectiveness and foster long-term relationships.	Nyakundi (2018)
Examine frameworks for AI adoption in SMEs and its impact on performance.	Literature review and framework analysis.	A holistic approach to AI adoption enhances operational efficiency and economic performance.	Understanding the interplay of factors is crucial for effective AI integration in SMEs.	Badghish & Soomro (2024)
Analyze the role of AI in new product development (NPD) within SMEs.	Focused case studies on NPD processes.	AI can improve product development metrics by approximately 27%.	AI serves as a driver of innovation and market responsiveness in product development.	Cooper (2025)
Investigate the role of AI-Assisted Social Media Marketing (AISMM) for SMEs.	Empirical study on marketing strategies.	AISMM significantly enhances customer engagement and profitability for SMEs.	Proactive AI integration in marketing can lead to improved customer relationships and sales.	Basri (2020)
Explore AI's role in performance management systems in SMEs.	Qualitative analysis of management practices.	AI-driven insights improve employee engagement and productivity.	Adapting management frameworks to include AI is essential for fostering a responsive workforce.	Nyathani (2023)

Unveil next-gen supplier strategies in the UAE's oil & gas sector.	Case study analysis.	Innovative supplier strategies can enhance procurement efficiency.	Emphasizing next-gen strategies helps SMEs adapt to industry challenges and improve performance.	Alhammadi et al. (2024)

III. BACKGROUND

The fourth industrial revolution poses a significant challenges for small and medium enterprises (SMEs) in procurement, which is vital for economic growth and accounts for substantial share of GDP in developed economies. SMEs often struggle with manual processes that are time-consuming and biased, hindering their competitiveness. This research aims to explore how artificial intelligence (AI) can transform procurement by providing advanced analytics for data-driven decision-making, ultimately enhancing operational efficiency and resilience. By leveraging AI, SMEs can streamline their procurement processes, capitalize on their strengths, and achieve sustainable growth in a competitive market.

IV. MOTIVATION

The research aims to address the challenges faced by small and medium enterprises (SMEs) in navigating increasingly complex procurement landscapes particularly in the competitive UAE market. It highlights the necessity for SMEs to adopt efficient, data-driven decision-making practices to maintain their competitive edge. Current manual processes are criticized for being time- consuming, resources-intensive, and prone to biases, leading to suboptimal outcomes. The study explores how artificial intelligence (AI) can transform procurement method by providing SMEs with access to advanced analytics that facilitate informed, real time decision-making, moving away from outdated practices [31] , the research emphasizes the importance of advanced procurement analytics for improved decision-making and cost efficiency in the global supply chain. It investigates specific ways AI cem empower SMEs to leverage their strength and weaknesses, enhance operational resilience, ultimately positioning SMEs for sustainable growth in a dynamic market.

V. PROBLEM STATEMENT

The procurement processes for small and medium enterprises (SMEs) in the UAE face several significant challenges that hinder operational efficiency and effective decision-making. The first issue is the reliance on the manual processes, which limits SMEs ability to respond to market change optimize procurement strategies [21] emphasizes that integrating artificial Intellgence (AI) into supply chain management can automate routine task and provide real time data analytics, underscoring the need to transition from manual to AI driven processes to enhance performance.

The second challenge is the prevalence of biased decision-making in manual procurement, leading to inaccurate outcomes poor supplier selection and resource misallocation which ultimately affects SMEs competitiveness. AI can help create a structured roadmap to mitigate biases, thereby improving decision-making accuracy and enabling a more objective evaluation of strengths and weaknesses [8].

The third issue is the time and cost inefficiencies associated with manual procurement processes [43], which can limit SMEs ability to undergo digital transformation. This underscores the urgent need for an AI roadmap that not only implement technology but also provides actionable solutions to overcome these challenges and facilitate smoother transitions to more efficient e-procurement practices.

VI. RESEARCH OBJECTIVES

Understand influential factors: to identify and analyze the factors that impact performance of SMEs within UAE procurement sector.

Exploring AI assistant tool: to investigate the role of artificial intelligence (AI) in enhancing procurement processes and improving decision-making capabilities for SMEs.

Designing a framework for the implementation of AI in procurement and to evaluate the collaboration with industry expert.

Proposing a practical roadmap for SMEs informed by insights gained from decision tree analysis, outlining clear and actionable steps for the successful AI implementation for procurement processes.

VII. METHODOLOGY FOLLOWED

This research adopts a data-driven methodology centered in utilizing Decision Tree (DT) to classify the strengths and weaknesses of small and medium enterprises (SMEs) in the UAE procurement sector. The DT approach is a supervised learning technique used for classification tasks, where data is systematically divided into smaller subsets based on specific features. This method not only predicts outcomes related to the classification of the SMEs but also [provides a clear explanation for its predictions making it accessible for SME stakeholders to understand what influence performance.

The research aims to address the following questions:

1. How can the implementation of artificial intelligence (AI) improve the procurement performance of small and medium-sized enterprises (SMEs) in the UAE?

2. How can a decision tree model be designed to assist SMEs in making informed decisions regarding the adoption of AI in their procurement strategies?

To identify the strength and weakness of SMEs, data will be collected through a structured questionnaire aimed at SMEs stakeholders. This research will utilize a quantitative approach focused on employing surveys to explore the factors influencing the productivity of the SMEs within the procurement sector, particularly concerning the implementation of artificial intelligence. The workflow begins with gathering raw data from the questionnaire which will cover various aspects of procurement performance, operational capabilities and the market challenges faced by SMEs. This raw data will be reprocessed to clean transform and prepare it for analysis using the DT classification technique.

Decision tree analysis will be conducted to identify the critical factors that enhance AI assistance in the procurement processes of UAE SMEs. Based on that insight gathered from the questionnaire, the framework of the paper will be designed and presented to subject matter experts for their input. This feedback will then be used to refine and adjust the framework accordingly. The resulting model will visually represent the key factors contributing to identified strengths and weaknesses of each SME, allowing the stakeholders to explore and understand the critical areas that require improvement. This model will serve as a valuable resource for SME managers, enabling them to make data-driven decisions regarding their procurement strategies.

Soon, we plan to implement the decision tree approach using python programming language to facilitate model development and analysis. A structured questionnaire will gather quantitative data on the strengths, weaknesses, and market dynamics of SMEs, supplemented by secondary data analysis. The findings will contribute to the development of comprehensive AI roadmap that include strategic initiatives and practical recommendations to address barrier to adoption. The proposed methodology will provide insights into the strength and weaknesses of the SMEs, ensuring that the outputs are both predictive and actionable for end users such as procurement managers. By focusing on strength and weaknesses identified through the decision tree analysis, SMEs will be better equipped to navigate the challenges in the procurement landscape and improve their overall operational efficiency.

Currently, the identification of the questionnaire will be identified to prepare the questionnaire and gather the data.

A stratified random sampling technique will be used to select a representative sample of SMEs from various industries and regions within the UAE, ensuring diversity and generalizability of the findings.

Data will be collected through a structured questionnaire targeting SME stakeholders, covering procurement performance, operational capabilities, and AI-related challenges. Secondary data from industry reports and government publications will also be utilized to complement the primary data.

To ensure the validity and reliability of the findings, a pilot test of the questionnaire will be conducted, followed by data triangulation to compare primary and secondary data results. Additionally, the decision tree model will be validated using a holdout sample to assess its predictive accuracy, reinforcing the credibility of the conclusions drawn from the analysis.

Scenario: Decision Tree Output

Consider a hypothetical retail SME using the decision tree model. Key factors influencing its procurement performance include:

Technological Readiness: Investment in modern procurement software.

Staff Training: 70% of staff trained, enhancing efficiency.

Market Awareness: Strong understanding of AI capabilities.

The decision tree predicts a potential 25% improvement in procurement efficiency with further AI investment, while financial constraints are identified as a significant barrier.

Prototype Output

The prototype output of the decision tree can be visualized as follows:

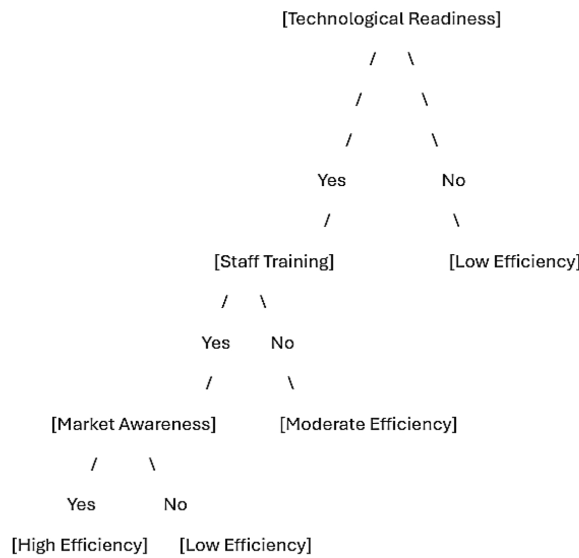


Fig. 1. Example of Decision Tree Diagram.

This visualization helps stakeholders understand how these factors contribute to procurement performance and indicates that improving financial resources could enhance efficiency.

The results show that SMEs with strong technological readiness and staff training are better positioned to leverage AI, while financial constraints hinder adoption. Targeted interventions, such as financial support and training programs, are essential for facilitating AI integration in procurement processes. The decision tree model serves as a valuable tool for SMEs to develop actionable strategies for improving operational efficiency.

VIII. RESULT

The results of this paper, which utilized a Decision Tree (DT) classification model to analyze data collected from SMEs in the UAE procurement sector, revealed critical insights into the strengths and weaknesses of these enterprises regarding AI adoption. The analysis identified key factors influencing procurement performance, such as technological readiness, staff training, and market awareness, highlighting strengths like agility and cost efficiency among SMEs. However, it also uncovered weaknesses, including limited financial resources and resistance to change, which hinder AI assistance. The developed framework, refined through expert feedback, visually represents these strengths and weaknesses, serving as a valuable resource for SME managers to make informed, data-driven decisions in their procurement strategies. Overall, the findings underscore the potential of AI to enhance procurement processes, equipping SMEs with actionable insights to navigate challenges and improve operational efficiency.

IX. CONCLUSION

In conclusion, this research presents a comprehensive AI roadmap designed to enhance the procurement processes of the SMEs in the UAE, addressing the inefficiencies of traditional methods that hinder competitiveness. By utilizing the decision tree classification and data-driven insights, the roadmap empowers SMEs to identify their strengths and weaknesses facilitating informed decision making and strategic planning. The assistance of AI technologies not only streamlines procurement practices but also fosters transparency and trust in the decision-making process. Ultimately, this study underscores the potential for AI to transform the procurement landscape for SMEs, contributing to their growth and sustainability in the dynamic UAE market.

Future work will focus on conducting longitudinal studies to assess the long-term impacts of AI assistance on procurement performance in SMEs, as well as expanding research to include a broader range of sectors and regions within the UAE. Additionally, developing tailored training programs for employees to enhance their skills in using AI tools will be essential. Finally, exploring advanced machine learning techniques, such as neural networks and ensemble methods, could further refine the decision-making framework and improve its predictive capabilities. These initiatives aim to continuously enhance procurement practices, fostering innovation and competitiveness among SMEs in the UAE.

REFERENCES

[1] Z. Ahmad, S. Rahim, M. Zubair, and J. Abdul-Ghafar, “Artificial intelligence (AI) in medicine: Current applications and future role with special emphasis on its potential and promise in pathology,” *

[2] A. Alhammadi, T. Yusaf, J. Soar, B. M. Ali, K. Kadirgama, and B. F. Yousif, “Revolutionizing procurement: Unveiling next-gen supplier strategies in UAE's oil & gas sector,” *The Extractive Industries

[3] H. Ali and L. Hajjar, “Empowering SMEs with AI and Digital Transformation: A Roadmap to Enhanced Competitiveness and Growth,” 2024.

[4] S. Badghish and Y. A. Soomro, “Artificial Intelligence adoption by SMEs to achieve sustainable business performance: Application of Technology–Organization–Environment framework,” *Sustainability*, 2

[5] O. Badmus, S. A. Rajput, J. B. Arogundade, and M. Williams, “AI-driven business analytics and decision making,” *World Journal of Advanced Research and Reviews*, pp. 616–633, 2024.

[6] Basri and Wael, “Examining the Impact of Artificial Intelligence (AI)-Assisted Social Media Marketing on the Performance of Small and Medium Enterprises: Toward Effective Business Management in the S

[7] F. Bassi and D. Costa, “Circular Economy in small and medium-sized enterprises in the European Union: heterogeneity between and within EU countries,” *Statistica Applicata-Italian Journal of Applied

[8] C. Bersani, J. Codagnone, L. David, A. Foiniotis, G. Galasso, S. Mancini, R. Michieletti, C. Orphanidou, and M. Pellegrino, “Roadmap for actions on artificial intelligence for evidence management in

[9] J. Brodny and M. Tutak, “Digitalization of small and medium-sized enterprises and economic growth: Evidence for the EU-27 countries,” *Journal of Open Innovation: Technology, Market, and Complexity*,

[10] A. Caliskan, Y. D. Özkan Özen, and Y. Ozturkoglu, “Digital transformation of traditional marketing business model in new industry era,” *Journal of Enterprise Information Management*, vol. 34, no. 4,

[11] B. Charbuty and A. Abdulazeez, “Classification based on decision tree algorithm for machine learning,” *Journal of Applied Science and Technology Trends*, pp. 20-28, 2021.

[12] R. S. Russell and B. W. Taylor, *Supply Chain Management: Strategy, Planning, and Operation*, 2019.

[13] R. G. Cooper, “SMEs’ use of AI for new product development: Adoption rates by application and readiness-to-adopt,” *Industrial Marketing Management*, pp. 159-167, 2025.

[14] E. Dean, J. Elardo, M. Green, B. Wilson, and S. Berger, “Measuring the Size of the Economy: Gross Domestic Product,” in *Principles of Economics: Scarcity and Social Provisioning*, 2nd ed., 2020.

[15] S. Williams and J. Tillipman, Eds., *Routledge Handbook of Public Procurement Corruption*, Routledge, 2024.

[16] M. El Khatib, A. Al Mulla, and W. Al Ketbi, “The role of blockchain in E-governance and decision-making in project and program management,” *Advances in Internet of Things*, pp. 88-109, 2022.

[17] A. J. van Weele and F. Rozemeijer, *Procurement and Supply Chain Management*, Pearson UK, 2020.

[18] R. Fornasiero, L. Kiebler, M. Falsafi, and S. Sardesai, “Proposing a maturity model for assessing Artificial Intelligence and Big Data in the process industry,” *International Journal of Production R

[19] T. E. Ghak and H. Zarrouk, “Opportunities and Challenges facing SMEs’ access to financing in the UAE: An analytical study,” *Contemporary Research in Accounting and Finance: Case Studies from the MEN

[20] B. Hoekman and B. K. O. Taş, “Procurement policy and SME participation in public purchasing,” *Small Business Economics*, pp. 383-402, 2022.

[21] M. K. Aldin Ismaeil, “The Role and Impact of Artificial Intelligence on Supply Chain Management: Efficiency, Challenges, and Strategic Implementation,” *Journal of Ecohumanism*, pp. 89-106, 2024.

[22] A. Ali, R. Jayaraman, E. Azar, and M. Maalouf, “Maximizing supply chain performance leveraging machine learning to anticipate customer backorders,” *Computers & Industrial Engineering*, vol. 194, Art

[23] L. E. Jowah and X. Mkuhlana, “Budgeting systems and project execution at a selected government department of the Western Cape Province, South Africa,” *Journal of Public Administration and Developmen

[24] H. S. Kristensen, M. A. Mosgaard, and A. Remmen, “Circular public procurement practices in Danish municipalities,” *Journal of Cleaner Production*, 2021.

[25] C. P. Langlotz et al., “A roadmap for foundational research on artificial intelligence in medical imaging: From the 2018 NIH/RSNA/ACR/The Academy Workshop,” *Radiology*, pp. 781–791, 2019.

[26] Q. Lu, Y. Zhou, Z. Luan, and H. Song, “The effect of SMEs' ambidextrous innovations on supply chain financing performance: Balancing effect and moderating effect,” *International Journal of Operation

[27] A. Mavidis and D. Folinas, “From public E-procurement 3.0 to E-procurement 4.0: A critical literature review,” *Sustainability*, 2022.

[28] S. Montanari and U. Kocollari, “Defining the SME: A multi-perspective investigation,” in *The Changing Role of SMEs in Global Business: Volume II: Contextual Evolution Across Markets, Disciplines and

[29] J. M. Spreitzenbarth, C. Bode, and H. Stuckenschmidt, *Artificial Intelligence and Machine Learning for Coders*, O'Reilly Media, 2020.

[30] R. Nyathani, “AI in Performance Management: Redefining Performance Appraisals in the Digital Age,” *Journal of Artificial Intelligence & Cloud Computing*, pp. 1-5, 2023.

[31] A. Odutola, “Advanced procurement analytics: Building a model for improved decision making and cost efficiency within global supply chains,” *International Journal of Scientific and Management Resear

[32] O. B. Ogundipe, A. C. Okwandu, and S. A. Abdulwaheed, “Optimizing construction supply chains through AI: Streamlining material procurement and logistics for project success,” *GSC Advanced Research a

[33] M. Rakhmansyah, T. Wahyuningsih, A. D. Srenggini, and I. K. Gunawan, “Small and Medium Enterprises (SMEs) with SWOT Analysis Method,” *International Journal for Applied Information Management*, pp. 5

[34] D. K. Ravichandran and R. Krishnamoorthy, “SME CHALLENGES in UAE,” 2024.

[35] S. Refass, J. Lillywhite, F. Salem, Z. Akrout, S. Shaer, and E. Shibl, “The Future of SMEs in the UAE,” *Digital Economy Series*, 2023.

[36] R. G. Richey Jr, S. Chowdhury, B. Davis-Sramek, M. Giannakis, and Y. K. Dwivedi, “Artificial intelligence in logistics and supply chain management: A primer and roadmap for research,” *Journal of Bus

[37] M. Kolagar, V. Parida, and D. Sjödin, “Linking digital servitization and industrial sustainability performance: A configurational perspective on smart solution strategies,” *IEEE Transactions on Engi

[38] "Small and Medium-sized," Abu Dhabi Chamber, Oct. 2019. [Online]. Available: https://www.abudhabichamber.ae/-/media/Project/ADCCI/ADCCI/Media-Center---Publications/Research-and-Reports/2019/smes-crop

[39] J. Smith, G. Andersson, R. Gourlay, S. Karner, B. E. Mikkelsen, R. Sonnino, and D. Barling, "Balancing competing policy demands: The case of sustainable public sector food procurement," *Journal of C

[40] E. C. Stade et al., "Large language models could change the future of behavioral healthcare: A proposal for responsible development and evaluation," *Mental Health Research*, 2024.

[41] A. J. Varma, N. Taleb, R. A. Said, T. M. Ghazal, M. Ahmad, H. M. Alzoubi, and M. Alshurideh, "A roadmap for SMEs to adopt an AI-based cyber threat intelligence," in *The Effect of Information Technol

[42] D. Vu, "Digital Transformation E-procurement: Analysis of the Barriers in applying E-procurement in Vietnamese SMEs," 2024.

[43] K. Z. Zhang, "Examining the influence of online reviews on consumers' decision-making: A heuristic–systematic model," *Decision Support Systems*, pp. 78-89, 2014.

[44] M. G. Nyakundi, "Procurement best practices and procurement performance of SMEs in Nairobi County," Doctoral dissertation, University of Nairobi, 2018.

Exploring AI Ethical Considerations and Governance in the Financial Sector: Examination of Select Countries

Aditya Varshney
B.Tech (Computer Science & Business Systems)
VIT- AP University
Amravati, India
adityavarshney9876@gmail.com

Deepanjana Varshney
Professor
Abu Dhabi School of Management
Abu Dhabi, UAE
d.varshney@adsm.ac.ae

Abstract: The impact of Artificial Intelligence on organizations' financial sectors has drawn profound interest in recent years. AI-related ethical considerations and governance in organizations, especially in the financial sector, involve scrutinizing AI algorithms for bias, owing to emerging regulatory standards, and addressing privacy issues. The paper aims to examine AI-related ethical considerations and governance in select countries through an exploratory approach, studying recent research in the domain. This working paper examines and analyzes the multidimensional facets of AI and the risks that should be assessed and mitigated in organizations, focusing on country-specific guidelines. A model with an appropriate rationale has been developed to address the gaps in AI ethical considerations and governance, complementing the prevailing shortcomings of AI systems and processes in the financial sectors of select countries. Furthermore, it was found that some countries had stringent AI considerations and guidelines, while in some countries, the guidelines were comparatively loose. To cover the inconsistencies, a model covering the relevant dimensions required has been framed, which can address the current loopholes in the countries studied or to be studied using this framework.

Keywords: Artificial Intelligence, financial sector, ethical considerations, UAE, India

I. INTRODUCTION

Artificial Intelligence (AI) has gained significant momentum in digital technology usage worldwide over the past few years. Numerous applications of AI have been found in various areas, including predictive analytics and identification, which have been widely adopted in diverse contexts such as education, recruitment, training, marketing, governance, finance, and security [1].

With AI's increasing usage and implementation, challenges and risks have arisen, along with inevitable questions about how AI systems should conform to ethical standards. These issues have triggered a global initiative and narrative on AI ethics. Governments and policymakers from different countries have tried to develop and propose AI-related guidelines. Recent research has highlighted the fragmented discussion of AI governance dimensions, or it may not have encompassed the entire gamut of governance issues. It is essential to recognize that, as technology has limitations, governments must collaborate on devising a globally accepted mechanism to strategize and coordinate sustainable AI ethics and governance [2]. AI has had, and will consistently have, a significant impact on global society; therefore, there is a critical need for effective regulations and policies to govern AI usage [3]. To emphasize, AI researchers agree that identifying the appropriate use of AI is a priority [4]. Discerning and evaluating AI systems' transparency, fairness, and accountability is imperative to stimulate user confidence in the AI mechanism [5].

The research has two objectives:

- To examine the AI ethics and governance guidelines related to the financial sector in select nations
- Develop a viable AI ethics and governance model that complements the shortcomings of the current country-wise guidelines.

Our research complements the gaps in AI ethical guidelines and considerations. Uniformity and standard guidelines were prevalent in some countries, while some had generic guidelines. The contribution was crafting a framework highlighting and recommending the essential AI ethical guidelines and considerations concerning the financial sector.

II. METHODOLOGY

A. Procedure

Based on our objectives, we studied the AI ethics and governance guidelines of the chosen countries (selected based on rapid progress, financial activities, and diversity) through the AI guidelines. Furthermore, recently published research studies and white papers were analysed [1]; [5]; [6]; [2]; [7]; [3]; [4]; White & Case papers, [8]; [9]. Additionally, we extensively studied the country’s AI guidelines from the country's prescribed websites.

We examined the selected countries in our research. We highlighted their respective policies and regulatory guidelines, ultimately developing a proposed framework that can be referenced in framing AI ethics and governance. In a focused manner, we examined the prevailing laws and regulations related to AI. We explored how AI can be effectively governed by the business and government sectors, especially the key geopolitical powers. [6] Their systematic literature review outlined that AI governance solutions can provide valuable proposals for developing comprehensive AI governance systems and frameworks.

B. Countries Studied

United Arab Emirates

The United Arab Emirates has established an Artificial Intelligence Office that actively participates in international AI forums, aiming for transparency, accountability, and ethical compliance. It supports global AI governance, security, and development alliances while promoting responsible AI use to ensure safety, privacy, and international stability. The UAE's charter for developing and using artificial intelligence aims to make the UAE the leading nation in AI by 2031. The policy promotes a nurturing environment for advancing Artificial Intelligence, which has rapidly developed in recent years. The primary objective is to establish robust governance principles that foster global cooperation and drive innovation, promoting economic growth and enhancing the quality of life. The primary objective of public policy is to position the UAE as a global leader in AI while upholding its core values, namely Progress, Collaboration, Society, Ethics, Sustainability, and Safety. These contributions support the development of the UAE AI Charter, driving economic diversification and innovation, and ultimately enhancing the quality of life through AI while fostering international cooperation in this field. There are no laws or regulations regarding AI governance in the Finance Sector. However, rules exist to secure data, which have been amended to address AI-related developments [10]. In the UAE, the Central Bank of the UAE (CBUA), along with the Securities and Commodities Authority (SCA), the Dubai Financial Services Authority (DFSA) of the Dubai International Financial Centre, and the Financial Services Regulatory Authority (FSRA of Abu Dhabi Global Market have charted out the guidelines for the Financial Institutions using the enabling technologies. The predominant principles are data protection, control functions, independent review, skills, knowledge, and expertise.

India

India’s approach to artificial intelligence regulation is sector-driven. The Responsible AI for India, released by the Ministry of Electronics and Information Technology (MeitY), outlines key principles emphasizing transparency, accountability, and fairness to address AI bias and discrimination. An advisory group, chaired by the principal scientific advisor, has been formed to develop an AI-specific regulatory framework for India. Under the guidance of the Advisory Group, the Subcommittee on AI Governance and Guidelines Development was established to provide actionable recommendations for AI governance in India. The Subcommittee examined key government issues, studied existing frameworks, conducted a gap analysis, and ultimately proposed an enhanced, comprehensive approach to ensure the reliability and accountability of AI systems [11]. The Securities and Exchange Board of India issued a circular in January 2019 on reporting requirements for AI and machine learning applications and systems offered and used in finance. This can be seen as a regulatory oversight until a comprehensive framework is adopted, as the circular merely enforces companies to disclose information on how the AI is trained and what data has been used. These do not provide a risk-free environment for AI in the finance sector, but enable potential users to access basic details about AI models and their training data. However, it does not legally restrict the use of developed AI systems using unverified or potentially biased data. The absence of sector-specific AI regulations leaves gaps in accountability, requiring a more comprehensive legislative approach [10]. India strives to balance fostering innovation and ensuring that AI benefits its people while mitigating the risks associated with the technology [7].

European Union

The European Union presented the Artificial Intelligence Act in 2024, which proposed a legal framework for regulating Artificial Intelligence within the European Union. The act aimed to develop Artificial Intelligence Systems that adhere to fundamental safety, rights, and ethical principles in the context of potential risks associated with Artificial Intelligence models, thereby generating trust in these models [12]. The AI Act classifies AI applications based on their risk levels: unacceptable risk, high risk, transparency risk, Minimal risk, or no risk. Market surveillance authorities observe, investigate, and enforce AI regulations. In contrast, Fundamental rights protection authorities have access to information, cooperation, and investigations into AI-related violations, ensuring protection of privacy and non-discrimination laws. While the primary objective of the Artificial Intelligence ACT is to establish AI regulatory frameworks and safeguard the privacy and ethical integrity of the financial sector, it also fosters innovation and development, as the AI ACT enables companies to develop and test general-purpose AI models before releasing them for public use. The companies have testing environments that simulate real-world conditions, allowing them to grow significantly in the AI sector and deliver high-quality, trustworthy AI models. National authorities provide such testing environments to companies to increase innovation and development [13].

The AI ACT categorizes AI systems based on risks and administers to each system depending on its level of risk, thereby enhancing security and ethical compatibility. Despite regulations, the EU promotes innovation safely by providing a development environment for companies' AI, which enables AI systems to undergo real-world simulations, enhancing quality while minimizing risk and boosting technological advancement. The AI Act encompasses fraud, embezzlement, credit scores, customer scores, investment optimisation/asset management decisions, and insurance underwriting systems.

China

China has adopted a stricter and more control-based approach to Artificial Intelligence regulations, aiming to control its development while appreciating its strategic advantages [9]. Some major requirements are:

- AI-driven financial services must adhere to the risk management guidelines.

- AI models must be fair, explainable, and transparent. This establishes the benchmark for evaluating AI algorithms to prevent bias and ensure ethical financial decision-making.
- Financial institutions providing AI services must disclose the work behind the AI decision-making, eliminating hidden biases in investment decisions.
- Some key compliance requirements for AI generation services include lawful use, data labelling rules, data training, content moderation, and reporting mechanisms.

Hong Kong

Hong Kong has adopted an open and flexible approach to regulating Artificial Intelligence. Hong Kong promotes technological advancements and economic growth while ensuring consumer protection, financial stability, and risk mitigation. The regulation of AI deployment in various sectors is handled through a blend of guidelines, industry standards, and legal frameworks, particularly in the financial sector. The Securities and Futures Ordinance in Hong Kong serves as the primary legal framework for regulating the securities and futures market in the territory. It is tasked with regulating AI applications that aid financial decisions, such as investment and trading, ensuring that the AI model operates without bias and in an ethical manner. The SFO ensures transparency, integrity, and stability, ensuring consumer protection and enhancing the reliability of these AI models. All risks and information must be disclosed to clients before any investment decisions are made, and the logic behind investment suggestions should be provided to help consumers understand the market and protect their financial interests [14]. Hong Kong has undergone rapid development in Artificial Intelligence applications within the financial sector, including models that perform AI-driven wealth management, credit scoring, and trading advisors. The government regulates and ensures these AI tools align with key factors, including fairness, transparency, and consumer protection. Some key principles applied in the AI-driven financial sector include model validation, auditability, vendor oversight, data protection, cybersecurity, and risk mitigation, as outlined by the Hong Kong Monetary Authority [15].

Results

The predominant attributes that were identified after a detailed examination have been highlighted below, country-wise and with an emphasis on the implications for the financial sector:

TABLE I: COUNTRY-WISE AI GUIDELINES

Country/ Region	Legal Framework / Strategy	Key Focus Areas	Financial Sector Relevance
China	Personal Information Protection Law (PIPL), AI Guidelines, Cybersecurity Law	-Strong accountability & penalties -Privacy through legal basis & consent -Data lifecycle governance -Risk-based monitoring & enforcement -Regulatory sandbox support -Government-curated datasets	-Heavily regulated financial AI -Strict vendor oversight -Requires model validation, fairness, and security - PIPL applies to financial data handling
United Arab Emirates	Public policy document issued by the Minister of State for Artificial Intelligence and Digital Economy and the Remote Work Applications Office	-Strong governance principles promote global cooperation -Drive innovation to enhance economic growth and quality of life.	**The Financial Free Zones** consist of Dubai International Financial Centre (DIFC) and the Abu Dhabi Global Market (ADGM). No law or regulation governs the implementation of Artificial Intelligence in the Finance sector. **Mainland UAE** consists of the remainder outside the Financial Free Zones. While there are no laws or regulations for implementing AI in the Financial Sector, amendments have been made to existing data protection legislation that applies in the DIFC to capture AI-related developments, including a recent amendment to Article 10 of the DIFC Data Protection Regulations.
Hong Kong	Data Privacy Ordinance, HKMA Fintech Reg, SFC guidelines	- Vendor risk management - Transparent AI deployment - Compliance with KYC/AML - Encouragement for innovation within limits	- Focus on Fintech AI & RegTech - Vendor audits & reviews are mandatory - AI credit scoring and robo-advisory need clear oversight
United States	NIST AI RMF, Executive Orders,	-Sector-specific approach - Innovation-	- Regulated by sector (e.g., SEC, CFPB) - Use of AI in fraud detection, trading, and

	FTC guidance	driven -Risk management & bias mitigation - Algorithmic accountability	credit scoring under scrutiny - No federal AI law, but strong agency control
European Union	EU AI Act (upcoming), GDPR	- Risk-tiered AI classification - Fundamental rights-based governance - High-risk AI under strict regulation - Emphasis on transparency, safety, and fairness	- AI for credit scoring, insurance, and investment = high risk - Mandatory risk assessment, audit logs - GDPR compliance for data use
Singapore	AI Model Governance Framework (voluntary), MAS sandbox	- Explainability and accountability - Encourages responsible innovation - Human oversight of AI - Testing under secure environments	- MAS Regulatory Sandbox for Fintech - Clear expectations for AI in finance - Voluntary frameworks are adopted widely
Canada	Directive on Automated Decision-Making, AIA tool	- Algorithmic Impact Assessment - Transparency and explainability - Human-in-the-loop design - Ethics & fairness embedded	- Public AI use must go through AIA - Push for fairness in automated credit and loan approvals - Transparency is required in automated financial decisions
United Kingdom	ICO, FCA regulations [16]	- Pro-innovation principles - Regulator-led enforcement - Fairness, accountability, data protection - Sector-specific guidelines	- FCA explores AI in financial services - Emphasis on fair outcomes and customer protection - Encourages internal governance & audits
Japan	Social Principles of AI [17]	- Human-centric AI - Safety, transparency, and trust - Guidelines for development - Cross-	- Promotes AI in financial innovation - Sector-based guidance with ethical oversight - AI investment advisors and credit systems under ethical norms

		border data cooperation	
South Korea	AI Ethics Guidelines (voluntary), AI Basic Act (draft)	- Transparency and accountability - Inclusiveness and human rights - Regulatory support for innovation - Government R&D funding	- AI used in banking and fintech encouraged - Ethics guidelines promote non-discrimination in lending/credit - Still developing stronger laws
India	Responsible AI Strategy (NITI Aayog), DPDP Act 2023	- Fairness, inclusiveness, and privacy - No central AI law yet - Pilot testing and use-case regulation - Focus on the public sector & governance	- RBI and SEBI exploring AI for supervision - Digital lending and credit scoring AI under regulatory radar - Lack of unified law, but moving toward a stronger policy base

The study unfolds the prevailing guidelines country-wise: China, as a country, can be benchmarked because of its guidelines. On the one hand, strict regulations oversee the financial sector, and an emphasis is placed on privacy-related aspects. In the UAE, the zones under the Dubai International Financial Centre (DIFC) and Abu Dhabi Global Market (ADGM) have not enforced AI regulations in the financial sector. Nevertheless, amendments have been made to the existing AI guidelines outside the financial free zones.

Hong Kong focuses on Fintech areas with compulsory vendor audits/KYC. The AI guidelines of the United Arab Emirates are sector-specific; no federal AI law exists, but agencies do the monitoring. The European Union uses AI for scoring, insurance, and investment. The Monetary Authority of Singapore (MAS) had developed the Smart Financial Centre. The centre permits experimentation with new technologies to support the latest FinTech developments. Canada applies Algorithmic Impact Assessment and stresses openness in making financial decisions. The Financial Conduct Authority (FCA) monitors the impact of AI on financial services. On the other hand, Japan encourages AI's role in financial services and sector-wise consultation. South Korea is nascent in its AI initiatives, and despite some guidelines, the AI regulations are not strictly applicable. The Reserve Bank of India (RBI) and the Securities and Exchange Board of India (SEBI) monitor AI activities. The

III. DEVELOPMENT OF THE MODEL

An elaborate, in-depth study demonstrates the salient dimensions. Accountability ensures clear ownership of AI actions for individuals and organizations involved; transparency promotes transparency in how AI systems operate, particularly regarding data and decision-making. Next, reliability states that AI must perform consistently and predictably under various real-world conditions; safety implies protecting users from financial or psychological harm from AI decisions. Based on China's PIPL, the privacy aspect ensures that personal and sensitive data is protected at all stages, followed by the security aspects that ensure financial AI systems are protected from malicious attacks and unauthorized access. Then there is the dimension of fairness: AI should treat every user equally, avoiding favoritism, discrimination, or exclusion. Furthermore, we covered other dimensions, specifically robustness (withstanding errors, misuse, and unpredictable real-world scenarios), explainability (the ability to comprehend), and interpretability (enabling a deeper understanding of AI logic for regulators and developers); Inclusiveness: accessible and fair for all users, regardless of their background or ability; risk mitigation (focuses on identifying and addressing risks throughout the AI lifecycle; national security (Protects national interests and prevents foreign entities' misuse of AI in finance); Vendor Oversight (external AI tools or services that financial institutions introduce); **AI Development (**focuses on developing, training, and enhancing AI for financial applications); financial sector applications and innovation versus compliance.

The proposed framework based on the above insights is provided in "Fig. 1".

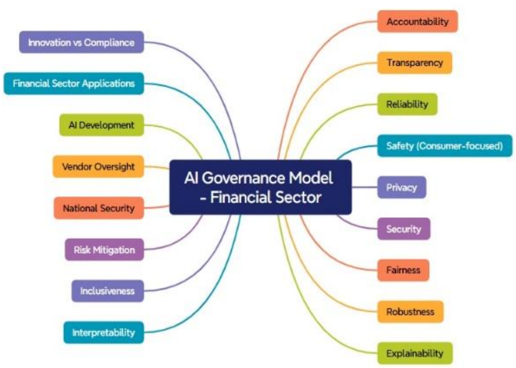


FIGURE 1: THE AI GOVERNANCE MODEL

IV. CONCLUSION

In this working paper, we examined country-specific AI guidelines and ethical considerations, emphasizing the financial sector, and highlighted the implications and insights derived from the proposed model. The salient findings of our research demonstrate that AI guidelines and ethical considerations across the countries studied lack uniformity in the financial sector, and even sector-wise specifications. Secondly, if the AI guidelines are considered as a continuum, a country like China or the United States of America is higher in the continuum. In the mid-level, India and the UAE stand with emerging AI regulations. Finally, at the low stage of the scale, countries like South Korea are at early formulation stages of AI guidelines. Our study has some limitations. Only eleven countries were examined for their AI guidelines and ethical considerations. Future studies can cover more countries across the globe and develop cross-comparisons. Additionally, to address the gaps in the exploratory research, future studies can be qualitative studies (interviews with AI policy and decision-makers), and analyse the perception of AI-users (B2B and B2C) of the financial sectors through empirical research.

REFERENCES

[1] Daly, A., Hagendorff, T., Li, H., Mann, M., Marda,V., Wagner, B., Wang, W. W., & Witteborn, S.(2019, July 08). Artificial Intelligence, Governance and Ethics: Global Perspectives. The Chinese University of Hong Kong Faculty of Law Research Paper No. 2019-15, University of Hong Kong Faculty of Law Research Paper No. 2019/033. http://dx.doi.org/10.2139/ssrn.3414805

[2] Jelinek, T., Wallach, W. & Kerimi, D. (2021). Policy brief: the creation of a G20 coordinating committee for the governance of artificial intelligence. AI Ethics, 1, 141–150. https://doi.org/10.1007/s43681-020-00019-y

[3] Qian, Y., Siau, K.L., & Nah, F.F. (2024). Societal impacts of artificial intelligence: Ethical, legal, and governance issues. Societal Impacts,3, 100040. https://doi.org/10.1016/j.socimp.2024.100040.

[4] Roberts, H., Cowls, J., Morley, J., Taddeo, M., Wang, V., & Floridi, L.(2021). The Chinese approach to artificial intelligence: an analysis of policy, ethics, and regulation. AI & Society, 36(1), 59–77.

[5] Shin, D. (2020). User Perceptions of Algorithmic Decisions in the Personalized AI System: Perceptual Evaluation of Fairness, Accountability, Transparency, and Explainability. Journal of Broadcasting & Electronic Media, 64(4), 541–565. https://doi.org/10.1080/08838151.2020.1843357

[6] Batool, M., Sanumi, O., & Jankovic, J. (2024). Application of artificial intelligence in materials science, with a special focus on fuel cells and electrolyzers. Energy and AI, 18(1), 100424. DOI:10.1016/j.egyai.2024.100424

[7] Pillay, T. (2024, Sept 05).Ashwini Vaishnaw: Minister of Electronics and Information Technology, India. Time. https://time.com/7012817/ashwini-vaishnaw/

[8] White & Case (2024, May 13). AI Watch: Global regulatory tracker-India. Retreived from https://www.whitecase.com/insight-our-thinking/ai-watch-global-regulatory-tracker-india

[9] White & Case (2025, March 31). Al Watch: Global regulatory tracker-China. Retrieved from https://www.whitecase.com/insight-our-thinking/ai-watch-global-regulatory-tracker-china

[10] UAE legislation (2024, 02 Sept). UAE's International Stance on Artificial Intelligence Policy. Retrieved from https://uaelegislation.gov.ae/en/policy/details/uae-s-international-stance-on-artificial-intelligence-policy

[11] Al Governance (2025, Jan 06). Report on AI governance guidelines development. India AI. Retrieved from https://indiaai.gov.in/article/report-on-ai-governance-guidelines-development

[12] AI Act (n.d.). Shaping Europe's digital future. European Commission. Retrieved from https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai

[13] Europarl (2023, June 08). EU AI Act: first regulation on artificial intelligence. European Parliament. https://www.europarl.europa.eu/topics/en/article/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence

[14] Hong Kong e-Legislation (n.d.). View Legislation. Retreived from https://www.elegislation.gov.hk/hk/cap571

[15] Hong Kong Monetary Authority (2019, Nov 01). High- level Principles on Artificial Intelligence. Retrieved from https://brdr.hkma.gov.hk/eng/doc-ldg/docId/20191101-1-EN

[16] A pro-innovation approach to AI regulation (2023, March 29). Command Paper Number: 815. HH Associates Ltd, UK. ISBN 978-1-5286-4009-1

[17] Habuka, H. (2023,Feb 14). Japan's approach to AI regulation and its impact on the 2023 G7 Presidency. Center for Strategic & International Studies. https://www.csis.org/analysis/japans-approach-ai-regulation-and-its-impact-2023-g7-presidency

Predictive Modeling for Aircraft Delay and Cancellations Using Machine Learning Techniques: A Data-Driven Approach to Optimize Flight Operations

Shamsah Saeed Alyaarbi
Artificial Intelligence Management Department
Abu Dhabi School Of Management
Abu Dhabi, UAE
Shamsa.shahein@outlook.com

Dr. Ishtiaq Rasool Khan
Artificial Intelligence Management Department
Abu Dhabi School Of Management
Abu Dhabi, UAE
i.khan@adsm.ac.ae

Abstract: Airlines and customers are increasingly concerned about flight delays. Flight delays are a major problem in the aviation sector. The rising frequency of flight delays puts a financial burden on the airline industry. Decision Tree (DT), Logistic Regression (LR), Random Forest (RF), and K-Nearest Neighbor (KNN) machine learning algorithms are implemented using the Flight Status Prediction Dataset to deal with the problem. Classification Report and Confusion Matrix, AUC-ROC score, and curve and SHAP analysis are used to evaluate the model's performance. The study compares the performance of models to identify the most effective approach for predicting delays and cancellations.

Keywords: Predictive Modeling, Flight Delays, Flight Cancellations, Machine Learning, Decision Tree, Logistic Regression, Random Forest, K-Nearest Neighbors, Data-Driven Approach, SHAP.

I. INTRODUCTION

A. Background

The aviation sector is an essential part of international transportation for millions of people and products to travel daily. However, there are a lot of difficulties associated with flight delays and cancellations, which can result in financial loss, unhappy customers, and inefficient operations [1]. Airline operations are disrupted, expenses rise, and passengers experience inconvenience as a result of flight cancellations and delays. Airlines struggle to efficiently deploy resources, manage scheduling, and notify passengers in advance due to the lack of an accurate forecasting system [2].

B. Motivation

A 2014 analysis by the Frankfurt-based consulting firm found that airline delays cost the global economy $25 billion. There was an indirect $4 billion decline in the US gross domestic product (GDP) due to domestic aircraft delays [3]. The motivation of this study is to apply predictive modeling to improve flight operations.

C. Problem Statement

The problem in this research is flight delay and cancellation which will be addressed using machine learning (ML) algorithms, specifically classification techniques. The project implements ML models that can predict delays and cancellations by analyzing past flight data.

D. Aims/Objectives

This research is aimed at developing and evaluating ML predictive models using classification algorithms for the problem of flight delays and cancellations.

The objectives are as follows:

- Exploring and investigating classification algorithms in general and specific to airline industry problems.
- Exploring important factors contributing to airline delays and cancellations, finding a suitable dataset including those features, and analyzing the data.
- Utilizing “Machine Learning classification algorithms including Decision Tree, Random Forest, Logistic Regression, and K-Nearest Neighbor” to predict the aircraft delay and cancellation on a test dataset.
- Evaluating and comparing the performance of the developed Machine Learning predictive models.
- Determine the association of predictive features with the output variable using explainable AI to understand the factors that most contribute to flight delays and cancellations.

E. Research Questions

The research questions are:

- How will classification algorithms perform in predicting aircraft delays and cancellations?
- Which model provides the best accuracy in predicting aircraft delays and cancellations?

- Which factors are important to predict delays and cancellations?

F. Contribution

The multi-class classification provides an advanced perspective of flight performance. The creative development of a new target variable, 'flight_status' is a significant advancement. The combination of correlation analysis and the ANOVA F-test for feature selection demonstrates an advanced strategy. The proposed solution employs extensive ML techniques and evaluation metrics.

The paper includes a Literature Review, Data Analysis, Method, Evaluation, and Conclusion.

II. LITERATURE REVIEW

A. Flight Delay Predictions

Authors examined machine learning (ML) and deep learning (DL) tools for analyzing departure delays, aiming to predict arrival and departure delays. The predictive results are compared to statistical variance and mean across different ML/DL models. Notably, the combinatorial approach shows superior results in random forest regression models [4].

The authors used Random Forest methodology to predict flight delays, analyzing feature selection's influence. Their study achieved a minimum mean square error of 0.1096 and demonstrated that the model's accuracy exceeds 90% [5].

The authors developed a Geographical and Operational Graph Convolutional Network (GOGCN) for predicting multiple airline delays. The network nodes represent connections based on spatial-temporal and geographical relationships. The geographical aggregator identifies similarities among nearby airports. This strategy outperforms existing standards in accuracy based on actual dataset tests [6].

Li and Jing [7] utilize temporal and spatial evaluation methods. Complex network theory reveals spatial characteristics of aircraft networks at three levels: edges, nodes, and overall structure. An LSTM-based prediction system tracks the relationship between weather and airport congestion delays. Random Forest classifier achieved 92.39% accuracy.

The authors created a reliable flight delay forecasting system using AI and machine learning. After comparing algorithms, CatBoost emerged as the top performer, leading to a user-friendly web application for accurate predictions [8].

The authors studied delays from inclement weather, evaluating several ML methods like SVR, Ridge, DT Regressor, RF, LR, and Lasso. The XGBoost regressor achieved the lowest RMSE score of 0.81 [9].

Authors [10] employed various ML techniques to identify causal factors and predict flight delays, including KNN, RF, DT, neural networks, and Naïve Bayes models. All algorithms achieved over 80% accuracy, with artificial neural networks leading.

Tang [11] used supervised ML models to predict airplane delays. Seven algorithms were trained and evaluated for binary classification: DT, Gaussian Naïve Bayes, RF, Logistic Regression, Support Vector Machine, KNN, and Gradient Boosted Tree. Four metrics assessed performance, showing that DT outperformed KNN.

Hatıpoğlu and Tosun [12] utilized ML methods like LightGBM, XGBoost, and CatBoost for gradient-boosting estimates, employing Bayesian hyperparameter settings and SMOTE for delayed flights. Their dataset included every flight from an international airline over a year, demonstrating high prediction accuracy.

Kiliç and Sallan [13] predicted arrival delays using ML and AI, leveraging 2017 domestic flight and weather data. Performance metrics compared these models, revealing gradient boosting as the top performer against logistic regression, feed-forward neural networks, and random forests.

The proposed solution represents ML approaches by the authors [14] for forecasting delay variations. Computational results indicate that ATFM delay inclusion makes it possible to achieve 0.80 prediction accuracy for arrival delay at the 0.50 confidence level and departure delay at the 0.65 confidence level.

The study used a support vector machine (SVM) model to analyze flight delay patterns and causes. The study found a significant correlation between flight departure delay and factors, with probabilities of 0.506, 0.478, 0.339, and 0.338 [15].

Authors [16] highlight the environmental impact of flight delays, causing significant financial and environmental costs for commercial airlines. This study employs ML models to forecast the likelihood of a particular aircraft experiencing a delay.

B. Challenges

Sridhar [17] discussed the advantages and disadvantages of Machine Learning Techniques (MLT) in aviation operations. MLT was divided into three categories: comparison of multiple MLTs, marginally better than physics-based models, and preferred option in absence of a physics-based model.

Authors [18] analyzed ML models developed for extended flight delay forecasting and looked into a wider variety of factors that could affect flight delay. The intended prediction challenges include many classification tasks and a regression exercise.

C. Models

Alharbi and Prince [19] attempted to resolve this issue by predicting using data mining tools and a machine learning methodology. The suggested hybrid strategy aims to anticipate the delay using a deep learning algorithm for classification and utilize the potential of ML. The findings show that the suggested approach outperformed the state-of-the-art.

Authors [20] adopted a bidirectional extreme learning machine (AB-ELM) approach along with their special parallel-series model. Results revealed that the patterns hidden within complex IATA-coded aircraft delay schema can be efficiently extracted using AB-ELM with parallel-series methods alongside appropriate sampling methods.

The research by authors [21] predicts flight delays using ensemble methods. The authors choose sample algorithms for the prediction issue and discuss the ensemble approaches. The outcomes demonstrate that the stacking method outperforms the other baseline methods.

Authors [22] attempted to forecast the departure delay of a planned aircraft. The study integrates weather and light data with the suggested airport traffic complexity (ATC) features using an airport situational awareness map. The authors find that the suggested framework improves the precision of forecasting delays.

III. DATA ANALYSIS

The Flight Status Prediction Dataset (2018-2022) has been used [23]. The dataset was downloaded from the Kaggle website. The original dataset is massive. A data sample comprising data from 5 years has been created. Unnecessary attributes have been eliminated. “combined_flight_sample2018-2022.csv” has 10% data for all years and 38 features. It has 2919378 entries. Missing values are addressed using the median and mode. Outliers were handled using the IQR method.

The “cancelled” variable was label encoded. Combining correlation analysis and the ANOVA F-test for feature selection effectively identifies significant features. A key innovation was consolidating cancellation and delay data into a new flight_status target variable. One-hot encoding of the target variable ensures compatibility with ML models.

IV. METHOD

K-Nearest Neighbors, Decision Trees, Random Forest, and Logistic Regression algorithms are chosen to develop prediction models, ensuring a comprehensive approach. An 80/20 ratio for training and testing splits the dataset, while SMOTE mitigates class imbalances to prevent bias towards the majority class.

A. Random Forest

Random Forest uses a collection of decision trees to achieve accurate predictions. The technique combines bagging techniques with feature unpredictability methods. The built-in feature significance ratings of RF aid in the identification of crucial predictive variables. RF prevents overfitting and noise effectively. RF analyzes features according to their impact on future predictions [24].

Important parameters have been fine-tuned in the Random Forest model to balance complexity and generalization. The number of Trees (n_estimators) = 100 has been used to keep the model light and avoid excessive processing. Tree Depth (max_depth) = 10 has been used to prevent deep trees from overfitting. random_state = 42 has been used.

B. Decision Tree

Decision Trees are adaptive models that divide data into subgroups based on feature values. They can encode complex decision-making rules and feature interactions. Decision trees create straightforward "if-then" rules [25]. The DT algorithm creates a training model that employs decision rules derived from training data to predict a value or class of target variables. DT categorizes situations by constructing a tree from the root to a few leaf nodes [26].

Important parameters have been fine-tuned to balance performance and avoid overfitting. Gini criteria have been used to split nodes in the decision tree. It is computationally faster because it does not require logarithms. Max Depth = 10 has been used to prevent overfitting. It limits tree depth to avoid excessively complex splits. It prompts default, simpler, and universal choices. random_state=42 has been used.

C. Logistic Regression

Logistic regression forecasts the possibility of an outcome using predictor factors. The simplicity and interpretability make it an excellent choice for determining the impact of characteristics on predictions. LR generates explicit coefficients that measure the impact of features [27]. A response consists of two possible answers: yes/no. LR model relates probabilities to predictor variables [28].

Key parameters have been fine-tuned to regulate regularization and convergence. Regularization Strength (C) = 10 has been used. It controls regularization strength and prevents overfitting. The (max_iter) have been set to 500 for better convergence speed. Solver = "lbfgs" has been used. Model implemented L2 penalty regularization to reduce model variance. random_state=42 has been used.

D. K-Nearest Neighbors

KNN operates as a non-parametric classifier that identifies points through their closest neighbors where the majority class prevails. KNN ability to extract local trends effectively and maintain a simple implementation [29]. KNN algorithm classifies data according to its closest k neighbors after estimating the distance between a new sample and the available data. KNN algorithm calculates the distance between samples. A simple majority vote among each point's k closest neighbors determines its classification [28].

Key hyperparameters have been fine-tuned to optimize the KNN model. The number of neighbors (n_neighbors) = 10 has been used for classification. Weight Function (weights) = "uniform" has been utilized. It means the same weight will be used for all neighbors. Distance Metric (metric) = Euclidean distance was chosen for the model.

V. EVALUATION

The evaluation techniques are comprehensive. The classification report assessed the overall performance of the models. Confusion matrix identified model error areas. AUC-ROC score and curve indicate the way each model distinguishes classes. Incorporating Explainable AI is a significant innovation. SHAP beeswarm plots show feature impacts on forecasts, enhancing transparency. The model with the best accuracy in predicting aircraft delays and cancellations will be identified.

A. CLASSIFICATION REPORT

The classification report gives a detailed assessment of the model's performance. It includes accuracy, precision, recall, and F1-score. This report assists in comparing performance across classes [30]. Figure 1. shows LR has 97% Accuracy, 0.95 Precision, 0.97 Recall, and 0.96 F1-score. Other models performed better than LR.

Random Forest has 100% Accuracy, 1.00 Precision, 1.00 Recall, and 1.00 F1-score. The RF model received perfect scores on this dataset.

The Decision Tree has 100% Accuracy, 1.00 Precision, 1.00 Recall, and 1.00 F1 Score. The DT model performs well but less than RF. K-Nearest Neighbors has a 99% Accuracy, 0.99 Precision, 0.99 Recall, and 0.99 F1-score. RF gives a better overall performance. =

	Model	Accuracy	Precision	Recall	F1 Score
0	Logistic Regression	0.97	0.95	0.97	0.96
1	Random Forest	1.00	1.00	1.00	1.00
2	Decision Tree	1.00	1.00	1.00	1.00
3	KNN	0.99	0.99	0.99	0.99

Fig. 8. Model Results

B. AUC-ROC

AUC-ROC statistics assess the ability of models to differentiate across classes. The results show that RF has 0.9998, DT has 0.9996, LR has 0.9294, and KNN has 0.9990 AUC-ROC Score. RF outperformed all models.

TABLE IX. AUC-ROC SCORE COMPARISON

Table Head	**AUC-ROC Score**	
	Model	**Score**
0	Random Forest	0.9998
1	Decision Tree	0.9996
2	Logistic Regression	0.9294
3	K-Nearest Neighbor	0.9990

C. CONFUSION MATRIX

The confusion matrix gives a detailed breakdown of the performance of models by displaying the counts of TP, FP, TN, and FN. It aids in understanding faults that the model makes [31].

The confusion matrix of RF model shows 15,508 flights were accurately projected as cancelled. 156 cancelled flights were incorrectly labeled as on-time. 96,615 flights were accurately predicted as delayed. 471,578 flights were accurately projected as on-time. 19 on-time planes were wrongly predicted as cancelled. The model worked well with a few misclassifications.

The DT model shows that 15581 cancelled flights were classified perfectly, but 83 cancelled flights were wrongly predicted as on time. Similarly, 82 flights were wrongly predicted as cancelled. 96615 are correctly predicted as delayed, and 471515 are correctly predicted as on time. 82 predictions are wrong as cancelled.

The LR model incorrectly predicted 15664 cancelled flights as on-time. This indicates that it cannot forecast any cancellations. The model accurately predicted 96,615 predictions as delayed. The model performs well on on-time flights with 471,353 correct predictions. But there are 242 that are misclassified as cancelled and 2 as delayed.

The KNN model accurately predicted 15,558 cancelled flights and misclassified 106 as on-time. The model classified 96,556 delayed flights accurately and misclassified 59 as on-time. 468,348 predictions were correctly classified for on-time flights with 97 misclassified predictions as cancelled and 3,152 as delayed. Overall, the model has minimal mistakes.

D. SHAP

SHAP is an explainable AI technique to evaluate model performance, feature relevance, and decision-making behavior. SHAP assigns feature priority based on cooperative game theory. SHAP identifies which aspects cause cancellations, delays, and on-time predictions.

SHAP analysis shows that RF has well-balanced SHAP values. It assigns feature importance evenly and results in balanced predictions. DT improves feature distribution but lacks strong separation for delayed flights. LR demonstrated an overreliance on delay-related properties. It makes it susceptible to errors in unseen data. KNN does not provide a clear feature influence. It makes it less interpretable. According to SHAP analysis, RF is the greatest option for predicting flight status because it is stable, accurate, and easy to comprehend.

VI. CONCLUSION

This study examined ML methods for predicting aircraft cancellations and delays. Forecasting airline delays with ML techniques shows significant potential. The proposed solution is a structured strategy, applying extensive models and evaluation metrics. SHAP was used for feature importance mapping. Findings indicate all strategies were effective, with the RF model outperforming other classifiers. Data-driven strategies enhance operational effectiveness and decision-making in aviation. This project provides a foundation for future advancements in flight prediction analytics. Future research can enhance contributions by incorporating new data sources and algorithms.

REFERENCES

[1] R. K. Jha, S. B. Jha, V. Pandey, and R. F. Babiceanu, "Flight Delay Prediction using Hybrid Machine Learning Approach: A Case Study of Major Airlines in the United States," arXiv preprint, arXiv:2409.00607, 2024.

[2] I. Hatıpoğlu and Ö. Tosun, "Predictive Modeling of Flight Delays at an Airport Using Machine Learning Methods," Applied Sciences, vol. 14, no. 13, p. 5472, 2024.

[3] A. M. Kalliguddi and A. K. Leboulluec, "Predictive modeling of aircraft flight delay," Universal Journal of Management, vol. 5, no. 10, pp. 485-491, 2017.

[4] J. M. Anguita and O. D. Olariaga, "Prediction of departure flight delays through the use of predictive tools based on machine learning/deep learning algorithms," The Aeronautical Journal, vol. 128, no. 1319, pp. 111-133, 2024.

[5] P. Hu, J. Zhang, and N. Li, "Research on Flight Delay Prediction Based on Random Forest," in 2021 IEEE 3rd International Conference on Civil Aviation Safety and Information Technology (ICCASIT), pp. 506-509, Oct. 2021. IEEE.

[6] C. A. I. Kaiquan, L. I. Yue, Z. H. U. Yongwen, F. A. N. G. Quan, Y. A. N. G. Yang, and D. U. Wenbo, "A geographical and operational deep graph convolutional approach for flight delay prediction," Chinese Journal of Aeronautics, vol. 36, no. 3, pp. 357-367, 2023.

[7] Q. Li and R. Jing, "Flight delay prediction from spatial and temporal perspective," Expert Systems with Applications, vol. 205, p. 117662, 2022.

[8] M. Alfarhood, R. Alotaibi, B. Abdulrahim, A. Einieh, M. Almousa, and A. Alkhanifer, "Predicting Flight Delays with Machine Learning: A Case Study from Saudi Arabian Airlines," International Journal of Aerospace Engineering, vol. 2024, no. 1, p. 3385463, 2024.

[9] R. T. Reddy, P. B. Pati, K. Deepa, and S. T. Sangeetha, "Flight Delay Prediction Using Machine Learning," in 2023 IEEE 8th International Conference for Convergence in Technology (I2CT), pp. 1-5, Apr. 2023. IEEE.

[10] C. Y. Yiu, K. K. Ng, K. C. Kwok, W. T. Lee, and H. T. Mo, "Flight delay predictions and the study of its causal factors using machine learning algorithms," in 2021 IEEE 3rd International Conference on Civil Aviation Safety and Information Technology (ICCASIT), pp. 179-183, Oct. 2021. IEEE.

[11] Y. Tang, "Airline flight delay prediction using machine learning models," in Proceedings of the 2021 5th International Conference on E-Business and Internet, pp. 151-154, Oct. 2021.

[12] I. Hatıpoğlu and Ö. Tosun, "Predictive Modeling of Flight Delays at an Airport Using Machine Learning Methods," Applied Sciences, vol. 14, no. 13, p. 5472, 2024.

[13] K. Kiliç and J. M. Sallan, "Study of delay prediction in the US airport network," Aerospace, vol. 10, no. 4, p. 342, 2023.

[14] Z. Wang, C. Liao, X. Hang, L. Li, D. Delahaye, and M. Hansen, "Distribution prediction of strategic flight delays via machine learning methods," Sustainability, vol. 14, no. 22, p. 15180, 2022.

[15] E. Esmaeilzadeh and S. Mokhtarimousavi, "Machine learning approach for flight departure delay prediction and analysis," Transportation Research Record, vol. 2674, no. 8, pp. 145-159, 2020.

[16] P. Meel, M. Singhal, M. Tanwar, and N. Saini, "Predicting flight delays with error calculation using machine learned classifiers," in 2020 7th International Conference on Signal Processing and Integrated Networks (SPIN), pp. 71-76, Feb. 2020. IEEE.

[17] B. Sridhar, "Applications of machine learning techniques to aviation operations: Promises and challenges," in 2020 International Conference on Artificial Intelligence and Data Analytics for Air Transportation (AIDA-AT), pp. 1-12, Feb. 2020. IEEE.

[18] G. Gui, F. Liu, J. Sun, J. Yang, Z. Zhou, and D. Zhao, "Flight delay prediction based on aviation big data and machine learning," IEEE Transactions on Vehicular Technology, vol. 69, no. 1, pp. 140-150, 2019.

[19] B. Alharbi and M. Prince, "A hybrid artificial intelligence approach to predict flight delay," International Journal of Engineering Research and Technology, vol. 13, no. 4, pp. 814-822, 2020.

[20] W. A. Khan, S. H. Chung, A. E. Eltoukhy, and F. Khurshid, "A novel parallel series data-driven model for IATA-coded flight delays prediction and features analysis," Journal of Air Transport Management, vol. 114, p. 102488, 2024.

[21] X. Wang, Z. Wang, L. Wan, and Y. Tian, "Prediction of flight delays at Beijing capital international airport based on ensemble methods," Applied Sciences, vol. 12, no. 20, p. 10621, 2022.

[22] W. Shao, A. Prabowo, S. Zhao, S. Tan, P. Koniusz, J. Chan, and F. D. Salim, "Flight delay prediction using airport situational awareness map," in Proceedings of the 27th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, pp. 432-435, Nov. 2019.

[23] R. Mulla, "Flight status prediction," Kaggle, 2022. [Online]. Available: https://www.kaggle.com/datasets/robikscube/flight-delay-dataset-20182022.

[24] A. Gole, S. Singh, P. Kanherkar, P. R. Abhishek, and P. Wankhede, "Comparative Analysis of Machine Learning Algorithms: Random Forest algorithm, Naive Bayes Classifier and KNN-A survey," in International Journal for Research Publication and Seminar, vol. 13, no. 3, pp. 194-197, Apr. 2022.

[25] A. V. Joshi, Decision Trees, in Machine Learning and Artificial Intelligence, pp. 73-87, Cham: Springer International Publishing, 2022.

[26] H. Blockeel, L. Devos, B. Frénay, G. Nanfack, and S. Nijssen, "Decision trees: from efficient prediction to responsible AI," Frontiers in Artificial Intelligence, vol. 6, p. 1124553, 2023.

[27] A. Das, Logistic Regression. In Encyclopedia of Quality of Life and Well-Being Research, Cham: Springer International Publishing, 2024, pp. 3985-3986.

[28] I. H. Sarker, "Machine learning: Algorithms, real-world applications and research directions," SN Computer Science, vol. 2, no. 3, p. 160, 2021.

[29] P. Cunningham and S. J. Delany, "K-nearest neighbour classifiers-a tutorial," ACM Computing Surveys (CSUR), vol. 54, no. 6, pp. 1-25, 2021.

[30] O. Rainio, J. Teuho, and R. Klén, "Evaluation metrics and statistical tests for machine learning," Scientific Reports, vol. 14, no. 1, p. 6086, 2024.

[31] B. J. Erickson and F. Kitamura, "Magician’s corner: 9. Performance metrics for machine learning models," Radiology: Artificial Intelligence, vol. 3, no. 3, p. e200126, 2021

The Multifaceted Transformative Impact of Generative AI on Organizations

Besma Glaa
Department of Managent an Engineering
Linköping University
Linköping, Sweden
besma.glaa@liu.se

Abstract—This paper presents a systematic literature review on the transformative impact of generative AI on organizations, highlighting crucial themes such as the disruption of business models, shifts in innovation strategies, revolutionizing workforce dynamics, ethical and governance challenges and the industry-specific transformations. The review also surfaces several cross-cutting tensions, particularly between automation and authenticity, strategic homogenization and competitive differentiation, augmentation and displacement, and efficiency and authorship that shape how generative AI is implemented and experienced in practice. The findings indicate that while generative AI introduces significant efficiencies and novel opportunities, it also poses ethical challenges and risks that necessitate careful management and oversight. These insights highlight important directions for future research on generative AI's multifaceted role in transforming organizations and the tensions it creates.

Keywords—Artificial intelligence, Generative AI, organization transformation, strategic innovation, organizational change, AI governance.

I. INTRODUCTION

The emergence of Generative Artificial Intelligence (GenAI) represents a transformative force within the business landscape, fundamentally altering organizational strategies, operational models, and competitive dynamics [13, 29]. More than merely technological advancement, GenAI signals a paradigm shift from predictive analytics to generative creation, facilitating innovative applications across various business functions [35]. From automating complex processes to generating creative solutions, GenAI has become a catalyst for business transformation, with early adopters reporting productivity gains of 40-60% in knowledge-intensive tasks [27].

The potential of GenAI is evident across all core business functions. In the realm of strategic management, GenAI enhances real-time scenario analysis and supports data-driven decision-making [39, 8]. Within operations, it optimizes processes and reduces costs through intelligent automation [31]. Marketing has experienced particularly profound changes, with GenAI enabling hyper-personalized consumer engagement at scale [20, 13] and facilitating dynamic content creation [14]. In terms of innovation, GenAI accelerates ideation, enhances creative processes [35], and supports rapid prototyping [29]. Moreover, even resource-constrained small and medium-sized enterprises (SMEs) are leveraging GenAI to overcome traditional barriers, thereby building resilience during crises [7] and unlocking new avenues for value creation [22].

However, the swift adoption of GenAI also introduces significant challenges that necessitate careful scrutiny. Despite its efficiency-enhancing capabilities, serious questions arise concerning its implications for business ethics, organizational structures, and competitive dynamics [31]. Key concerns include algorithmic bias, the complexities of intellectual property rights [9], and the imperative of preserving human oversight in critical decision-making processes [25]. Furthermore, the potential for market disruption and the evolution of new forms of competitive advantage [39] compel business leaders to reassess their strategic approaches.

Although the academic literature on GenAI is growing, it remains fragmented and predominantly focused on isolated applications, such as personalization or automation. While prior research often isolates GenAI's effects within specific domains, this review offers a more integrated perspective. To our knowledge, this review is the first to provide an integrated synthesis of how GenAI impacts multiple organizational functions, strategy, operations, and workforce, while also exploring a set of diverse and interrelated tensions emerging from AI integration. By bringing these dynamics into focus, this review contributes new conceptual clarity to the evolving discourse on AI-driven transformation in organizations. This systematic review seeks to address this gap by synthesizing current knowledge, identifying patterns of transformation, and proposing directions for future research.

II. METHODOLOGY

This study employed a systematic literature review methodology to comprehensively analyze research on Gen AI's transformative impact on organizations. The review process began with a targeted search across Scopus and Web of Science databases, focusing on peer-reviewed articles that examine GenAI's implications on organizations published in major business and economics journals (e.g. management journals, international business, operations management, innovation management, and organizations journals). We did not apply any restriction on the publication year of included articles. The search strategy used keywords such as "Generative AI", "GenAI", "generative artificial intelligence", "Large language models"

After removing duplicate records and screening titles/abstracts, the full texts of potentially relevant papers were assessed for quality and relevance, resulting in a final corpus of 197 studies. Our analysis followed Braun and Clarke's [42] six-phase reflexive thematic analysis framework providing a structured approach to analyzing qualitative data, from familiarization and coding to theme development and reporting. This methodology enabled both a comprehensive assessment of GenAI's multifaceted implications and identification of critical gaps in current research. A Prisma flow diagram (see Figure 1) documents the review process.

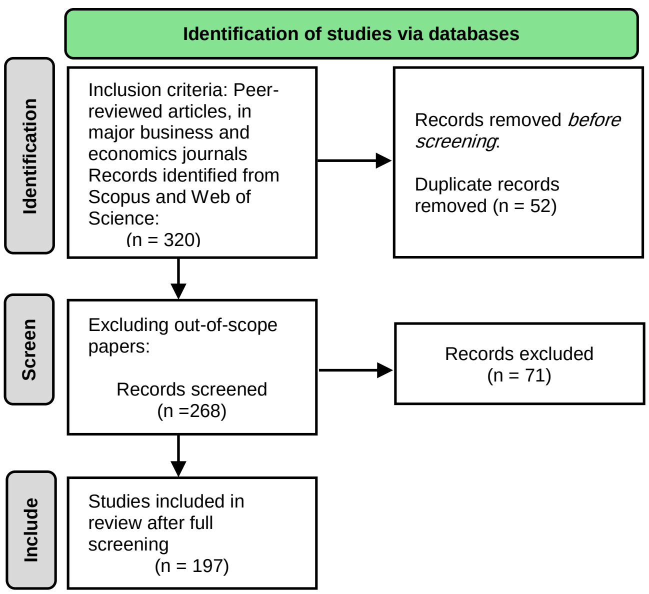


Fig. 1. Systematic review process (Own illustration)

thematic results and framework

This paper introduces a thematic framework comprising five dimensions: (1) Business model disruption, (2) Innovation strategy shifts, (3) Workforce and leadership transformation, (4) Sector-specific adaptations, and (5) Ethical and governance risks. These dimensions are linked through a conceptual model illustrating GenAI's dual role as a source of opportunity and risk, shaping organizational transformation (See table I).

A. Business Models Transformations

The influence of GenAI on business models is unfolding most notably across two areas: content automation and hyper-personalization. In marketing and content creation, GenAI is driving significant changes in both scale and efficiency [10]. At the same time, concerns about authenticity erosion in AI-generated branding, pointing to rising consumer skepticism as a potential risk for firms relying heavily on automated content [21].

The shift toward personalization extends beyond marketing. Reference [29] explores how AI-driven customization is being implemented into e-commerce, reshaping customer experiences through more tailored interactions. This shift is supported by the growing effectiveness of generative pricing models, which adapt dynamically to consumer behavior and market conditions, signaling a broader move toward AI-enabled responsiveness in digital business models [18].

B. Innovation Strategy Shifts

One of the most significant disruptions brought by GenAI is the democratization of innovation and the risk of strategic homogenization. Reference [24] emphasizes how GenAI tools are not only prompting new ideas but also enabling novel innovation strategies, including greater cross-industry transfer. Moreover, GenAI is lowering R&D costs, particularly for smaller firms, effectively leveling the playing field and opening up space for startups to challenge incumbents [33].

However, [3] caution that widespread adoption of similar AI tools could lead to strategic homogenization, where firms lose differentiation by relying on identical algorithmic approaches. These shifts disrupt traditional R&D strategies and competitive positioning, requiring organizations to reassess how they create and capture value [48]

C. Revolutionizing Workforce Dynamics

As GenAI becomes more embedded in organizational life, it is not just reshaping how work gets done, it is redefining core aspects of the workplace. These shifts are unfolding across three interrelated areas: workforce structures, skill demands, and leadership roles.

GenAI is not simply automating isolated tasks, it is driving fundamental changes in how work is structured and coordinated. Reference [41] shows how AI can significantly boost productivity in creative settings when used as a support tool, complementing rather than replacing human capabilities. In contrast, [46] find that roles involving routine cognitive tasks are particularly vulnerable to displacement. Furthermore, Gen AI has driven the emergence of hybrid roles that blend human judgment with AI input, often requiring new forms of collaboration and oversight [37].

As the structure of work evolves, so do the skills that organizations need. Reference [25] highlights the importance of ethical and proactive human resource development strategies to navigate this transition. Moreover, leading firms are already investing heavily in reskilling programs to equip their workforce for more complex, AI-supported roles [1]. At the same time, [5] emphasizes the role of psychological safety in supporting workers through these changes, especially as they take on tasks requiring more autonomy and problem-solving.

Leadership is also undergoing a transformation, as AI increasingly becomes a strategic asset in forecasting and planning [6]. Executive teams are beginning to incorporate

AI-driven analytics into their decision-making processes, demonstrating a shift in how leadership engages with data [30]. This evolving landscape, however, calls for a new mix of capabilities, particularly AI fluency, agility, and the ability to manage uncertainty, suggesting that traditional leadership skills alone may no longer suffice [11]. Reference [32] reinforce this point, emphasizing that leadership development must be aligned with broader shifts in workforce strategy. At the same time, the risk of automation bias highlights the need to preserve room for human judgment in high-stakes decisions [38].

At a deeper level, these changes are raising new questions about autonomy and identity at work. Reference [26] shows how attitudes toward AI, especially among middle managers, are shaped by organizational culture, which can either enable or hinder meaningful adoption. Thus, thoughtful integration of AI can actually support greater autonomy, especially when accompanied by clear human oversight [12].

D. Industry-Specific Transformations

The application of GenAI is unfolding with striking sectoral diversity, revealing distinct patterns of transformation across industries. In healthcare, [36] present compelling evidence of GenAI's potential to improve diagnostic accuracy and enable more personalized treatment plans, while also pointing to the practical hurdles organizations face during implementation. The financial sector has moved quickly, with [2] highlighting how GenAI is reshaping financing models and investor behavior across both startups and national markets, and [34] examining how it is reshaping environmental, social, and governance (ESG) evaluation frameworks.

In manufacturing, [28] explore how GenAI is advancing predictive maintenance and streamlining supply chains, playing a central role in the evolution toward Industry 5.0. Education is another area seeing wide adoption and business schools are reworking their curricula to better align with AI-integrated work environments. Hence, this highlights the importance of balancing AI adoption with the continued development of human-centered skills such as critical thinking and adaptability [45].

Hospitality offers perhaps the most visible examples of front-end transformation. Reference [19] outline a structured approach to embedding GenAI in customer service, and [40] provide a comprehensive look at AI-driven operational efficiency gains across the sector. Finally, in the creative industries, [44] document how GenAI is challenging long-standing assumptions about originality, authorship, and human creativity. Building on this, [15] find that AI-human collaboration is not only altering workflows but also challenging conventional ideas about what constitutes authorship and creative ownership.

Taken together, these studies show that GenAI's impact is far from uniform. Its success, and the form it takes, depends heavily on the specific demands, norms, and infrastructures within each sector.

E. Ethical Challenges And Potential Solotions

The integration of GenAI into business environments has raised pressing ethical concerns and exposed systemic risks that demand closer scrutiny. Without meaningful interventions, these challenges risk undermining the very promise of AI-driven transformation.

At the heart of ongoing ethical debates are issues surrounding authorship and intellectual property. Reference [17] offer foundational frameworks that tackle the blurred lines of creative ownership in AI-generated content, showing how GenAI complicates accountability in knowledge production. In the professional services domain, [16] highlights how algorithmic decision-making can compromise audit reliability, with potential consequences for trust and human judgment in financial reporting.

On the regulatory front, [43] expose serious policy blind spots, especially in areas like consumer protection and market fairness. They argue that existing governance structures are outpaced by the rapid evolution of algorithmic bias. Adding to the complexity, [4] underscore the dual-use risks of GenAI, cautioning that tools built for business innovation can also be exploited for fraud or misinformation.

Operational risks show equally concerning patterns. For instance, over-reliance on GenAI systems is already triggering unintended consequences. Deskilling is one of the most prominent risks, as automation gradually diminishes human expertise [25]. Reference [43] also raise alarms about the emergence of algorithmic monocultures, standardized systems that, while efficient, can create cascading failures across organizations. In high-stakes domains like hiring and credit assessment, [32] show how biased training data can reinforce systemic discrimination under the guise of neutrality.

To mitigate these risks, scholars are proposing a range of proactive solutions and calling for stronger accountability mechanisms. For instance, explainable AI, which preserves human interpretability, is one such solution [47]. Others advocate for robust governance structures [23], for example, recommend the creation of cross-functional ethical review boards to oversee AI applications. Real-time auditing, as highlighted by [34], can also play a key role in flagging and responding to emerging threats. Reference [25] propose a more thoughtful division of labor between humans and machines, ensuring efficiency without sacrificing professional dignity. Finally, [38] argues that continuous oversight will be vital as GenAI reshapes not only tasks but also the structure of organizational roles.

TABLE I. COMPARATIVE ANALYSIS AND THEMATIC CLUSTERING

Thematic cluster	**Core focus**	**Main insights (examples)**	**Examples of papers**
Business models transformations (38 papers)	- Content automation - Hyper-personalization	- Enables hyper-personalization and content automation at scale - Creates new competitive dynamics and pricing models - Democratization of innovation	[49] [29] [50]
Innovation strategy	Democratization of innovation	- Prompt-driven ideation - Cross-industry innovation	[24] [51] [52]

shifts (41 papers)	- Risk of strategic homogenization	transfer - SME competitive leveling -Competitive strategy shift -R&D strategy transformation	[48]
Revolutionizing workforce dynamics (42 papers)	- Reshaping workforce structures -Altering skill demands - Redefining leadership roles	- Creates hybrid human-AI roles and workflows - Requires significant reskilling and cultural adaptation - Transforms leadership decision-making processes - Impacts professional identity and autonomy - Demands new approaches to talent management	[41] [30] [25] [11] [32] [38] [26] [12] [15]
Sectoral Disruption Patterns (47 papers)	Industry-specific transformations	- Healthcare: enhances diagnostics but faces implementation challenges - Finance: transforms risk assessment and ESG evaluation - Manufacturing: enables predictive maintenance and supply chain optimization - Education: powers personalized learning systems - Hospitality: improves customer service and operations - Creative industries: shows unexpected disruption patterns	[36] [2] [34] [28] [53] [19] [40] [44]
Ethical and governance challenges (29 papers)	- Regulatory frameworks - Risk management authorship and intellectual property	- Creates authorship and IP ownership dilemmas - Reveals vulnerabilities in professional practices - Highlights regulatory gaps in AI governance - Risks of algorithmic bias and monocultures - Requires explainable AI and ethical review frameworks - Needs continuous monitoring mechanisms	[17] [16] [43] [4] [25] [47] [23] [34]

Thematic cluster	**Core focus**	**Main insights (examples)**	**Examples o papers**
Business models transformations (38 papers)	- Content automation - Hyper-personalization	- Enables hyper-personalization and content automation at scale - Creates new competitive dynamics and pricing models - Democratization of innovation	[49] [29] [50]
Innovation strategy shifts (41 papers)	Democratization of innovation - Risk of strategic homogenization	- Prompt-driven ideation - Cross-industry innovation transfer - SME competitive leveling -Competitive strategy shift -R&D strategy transformation	[24] [51] [52] [48]
Revolutionizing workforce dynamics (42 papers)	- Reshaping workforce structures -Altering skill demands - Redefining leadership roles	- Creates hybrid human-AI roles and workflows - Requires significant reskilling and cultural adaptation - Transforms leadership decision-making processes - Impacts professional identity and autonomy - Demands new approaches to talent management	[41] [30] [25] [11] [32] [38] [26] [12] [15]
Sectoral Disruption Patterns (47 papers)	Industry-specific transformations	- Healthcare: enhances diagnostics but faces implementation challenges - Finance: transforms risk assessment and ESG evaluation - Manufacturing: enables predictive maintenance and supply chain optimization - Education: powers personalized learning systems - Hospitality: improves customer service and operations - Creative industries: shows unexpected disruption patterns	[36] [2] [34] [28] [53] [19] [40] [44]
Ethical and governance challenges (29 papers)	- Regulatory frameworks - Risk management authorship and intellectual property	- Creates authorship and IP ownership dilemmas - Reveals vulnerabilities in professional practices - Highlights regulatory gaps in AI governance - Risks of algorithmic bias and monocultures - Requires explainable AI and ethical review frameworks - Needs continuous monitoring mechanisms	[17] [16] [43] [4] [25] [47] [23] [34]

III. DISCUSSION

While this review highlights the broad impact of GenAI across organizational domains, a closer look reveals a series of tensions and contradictions that merit deeper critical analysis (See Figure 2). These are not simply implementation challenges, they point to more fundamental uncertainties about how organizations are redefining value, capability, and control in the context of AI.

A key contradiction lies in the dual narrative of innovation democratization and strategic homogenization. While GenAI has lowered barriers to ideation and R&D, especially for smaller firms [24], the widespread adoption of similar tools may lead to convergence around generic solutions, undermining differentiation [3]. This paradox questions the long-term strategic value of accessible AI tools if competitive positioning is eroded. Another underexplored tension is between automation and authenticity in business models. Although AI enhances content creation and personalization on a scale, concerns are growing over the loss of human touch and declining trust in AI-generated experiences [21].

At the heart of ongoing ethical debates are issues surrounding authorship and intellectual property [17]. This suggests limits to automation in areas where meaning, emotion, and identity play a central role. In workforce transformation, the contrast between achieving efficiency and causing displacement is especially sharp. While some roles are enhanced through AI collaboration, others are at risk of marginalization or elimination [46; 38]. These outcomes are highly dependent on organizational choices about job design, cultural readiness, and leadership priorities.

Critically, literature offers few answers on how organizations can navigate these tensions deliberately. The findings suggest that GenAI's impact is not just technical or operational but deeply strategic and political, shaping what kinds of work are valued, who retains control, and how organizations sustain differentiation in environments where algorithmic solutions are becoming increasingly standardized.

Three key gaps reflect this challenge. First, there is insufficient understanding of how to maintain competitive advantage in the face of algorithmic homogenization, where reliance on similar GenAI tools risks eroding strategic differentiation. Second, few empirically validated frameworks exist to guide effective human-AI collaboration in decision-making and organizational design. Third, there is a clear absence of sector-specific governance models that account for the distinct ethical, regulatory, and operational demands of industries such as healthcare, finance, or education. Addressing these issues requires moving beyond descriptive studies toward more critical, comparative research that interrogates trade-offs and unintended consequences. This review takes a first step in surfacing these tensions and invites future work to engage more deeply with their implications.

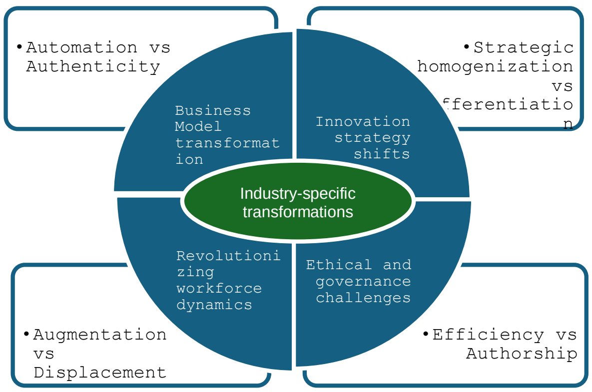


Fig. 2. Systematic review process (Own illustration)

IV. CONCLUSION AND FUTURE RESEARCH AGENDA

This systematic literature review reveals GenAI's multifaceted and dual role as both catalyst and disruptor in organizations. This review offers a novel contribution by being the first to systematically identify and synthesize the multidimensional tensions organizations face when implementing GenAI. These include the trade-off between automation and authenticity, the risk of strategic homogenization reducing strategic differentiation, the balance between augmenting human capabilities and displacing workers, and the trade-off between achieving efficiency while dealing with authorship issues. These tensions are not only underexplored in current literature but are also central to understanding the complex role GenAI plays in organizational transformation. Addressing these issues requires moving beyond descriptive studies toward more critical, comparative research that interrogates trade-offs and unintended consequences. This review takes a first step in surfacing these tensions and invites future work to engage more deeply with their implications.

This review highlights several underexplored themes that warrant further investigation. One key area is the strategic tension between strategic homogenization and competitive differentiation where widespread access to similar GenAI tools may reduce firms' ability to differentiate. Another promising theme is the transformation of leadership capabilities in AI-integrated environments, particularly the emergence of AI fluency, ethical agility, and data-informed decision-making as core competencies. Additionally, more research is needed on how GenAI affects organizational identity and employee autonomy, especially in hybrid human-AI roles. Finally, the role of organizational culture in shaping the adoption, resistance, or adaptation of GenAI technologies remains a critical yet insufficiently studied theme across sectors and firm sizes. Future research should aim to unpack these tensions empirically and develop frameworks capable of managing them in diverse organizational settings and different sectors.

To extend this short paper, a full-length systematic review is planned. An extended thematic synthesis will be conducted alongside a meta-analysis of eligible empirical studies to quantify GenAI's impact on organizational outcomes such as productivity, innovation, and decision-making. The subgroup and moderator analyses will further explore contextual differences across sectors and firm types. These steps will support the development of a more comprehensive, evidence-based framework on the role of GenAI in transforming organizations.

REFERENCES

[1] Ardichvili, A., Dirani, K., Jabarkhail, S., El Mansour, W., & Aboulhosn, S. (2024). Using generative AI in human resource development: An applied research study. Human Resource Development International, 27(3), 456–478.

[2] Siddik, A. B., Li, Y., & Du, A. M. (2024). Unlocking funding success for generative AI startups: The crucial role of investor influence. Finance Research Letters, 69(Part B), Article 106203.

[3] Eisenreich, A., Just, J., Gimenez-Jimenez, D., & Fueller, J. (2024). Revolution or inflated expectations? Exploring the impact of generative AI on ideation in a practical sustainability context. Technovation, 143, 103123.

[4] Grinbaum, A., & Adomaitis, L. (2024). Dual use concerns of generative AI and large language models. AI and Society, 39(2), 123–145.

[5] Manresa, A., Sammour, A., Mas-Machuca, M., Chen, W., & Botchie, D. (2024). Humanizing GenAI at work: Bridging the gap between technological innovation and employee engagement. Journal of Managerial Psychology, 1–21.

[6] Doshi, A. R., Bell, J. J., Mirzayev, E., & Vanneste, B. S. (2024). Generative artificial intelligence and evaluating strategic decisions. Strategic Management Journal, 46(4), 583–610.

[7] Shore, A., Tiwari, M., Tandon, P., & Foropon, C. (2024). Building entrepreneurial resilience during crisis using generative AI: An empirical study on SMEs. Technovation, 28(5), 321–335.

[8] Talaei-Khoei, A., Yang, A. T., & Masialeti, M. (2024). How does incorporating ChatGPT within a firm reinforce agility-mediated performance? The moderating role of innovation infusion and firms' ethical identity. Technovation, 132, 102975.

[9] Filippo, C., Vito, G., Irene, S., Simone, B., & Gualtiero, F. (2024). Future applications of generative large language models: A data-driven case study on ChatGPT. Technovation, 133, 103002.

[10] Zamudio, C., Grigsby, J. L., & Michelsen, M. (2025). RAISE: A new method to develop experimental stimuli for advertising research with image generative artificial intelligence. Journal of Advertising, 1–16.

[11] Jenkins, D., & Khanna, G. (2025). AI-enhanced training, education, and development: Exploration and insights into generative AI's role in leadership learning. Journal of Leadership Studies, 19(2), 9–29.

[12] Korayim, D., Bodhi, R., Badghish, S., Yaqub, M. Z., & Bianco, R. (2025). Do generative artificial intelligence related competencies, attitudes and experiences affect employee outcomes? An intellectual capital perspective. Journal of Intellectual Capital, 26(1), 100560.

[13] Hermann, E., & Puntoni, S. (2024). Artificial intelligence and consumer behavior: From predictive to generative AI. Journal of Business Research, 7(1), 1–12.

[14] Osadchaya, E., Marder, B., Yule, J. A., Yau, A., Lavertu, L., Stylos, N., Oliver, S., Angell, R., de Regt, A., Gao, L., Qi, K., Zhang, W. Z., Zhang, Y., Li, J., & Alrabiah, S. (2024). To ChatGPT, or not to

ChatGPT: Navigating the paradoxes of generative AI in the advertising industry. Business Horizons, 67(4), 555–570.
[15] Pantano, E., Serravalle, F., & Priporas, C.-V. (2024). The form of AI-driven luxury: How generative AI and large language models are transforming the creative process. Journal of Marketing Management, 40(17–18), 1771–1790.
[16] Agbon, G. (2023). Who speaks through the machine? Generative AI as discourse and implications for management. Critical Perspectives on Accounting, 45(2), 102–115.
[17] Islam, G., & Greenwood, M. (2023). Generative artificial intelligence as hypercommons: Ethics of authorship and ownership. Journal of Business Ethics, 187(2), 385–402.
[18] Park, H. E. (2023). The double-edged sword of generative artificial intelligence in digitalization: An affordances and constraints perspective. Psychology & Marketing, 40(10), 1953–1969.
[19] Zhang, H., Xiang, Z., & Zach, F. J. (2023). Generative AI vs. humans in online hotel review management: A task-technology fit perspective. Tourism Management, 98, 104758.
[20] Chan, H.-L., & Choi, T.-M. (2024). Using generative artificial intelligence in marketing: Development and practices. Journal of the Academy of Marketing Science, 52(3), 209–224.
[21] Bruens, J. D., & Meissner, M. (2024). Do you create your content yourself? Using generative artificial intelligence for social media content creation diminishes perceived brand authenticity. Journal of Social Media Studies, 15(2), 123–145.
[22] Rajaram, K., & Tinguely, P. N. (2024). Generative artificial intelligence in small and medium enterprises: Navigating its promises and challenges. Business Horizons, 67(5), 629–648.
[23] Sidaoui, K., Mahr, D., & Odekerken-Schroder, G. (2023). Generative AI in responsible conversational agent integration: Guidelines for service managers. Organizational Dynamics, 52(3), 100974.
[24] Sundberg, L., & Holmström, J. (2024). Innovating by prompting: How to facilitate innovation in the age of generative AI. Business Horizons, 67(5), 561–570.
[25] Yorks, L., & Jester, M. Y. (2024). Applying generative AI ethically in HRD practice. Human Resource Development International, 27(3), 245–260.
[26] Zhao, L., He, Q., Kamal, M. M., & O'Regan, N. (2025). Technophobia and the manager's intention to adopt generative AI: The impact of self-regulated learning and open organisational culture. Journal of Managerial Psychology, 1–19.
[27] Chui, M., Hazan, E., Roberts, R., Singla, A., Smaje, K., Sukharevsky, A., Yee, L., & Zemmel, R. (2023). The economic potential of generative AI: The next productivity frontier. McKinsey & Company.
[28] Ghobakhloo, M., Fathi, M., Iranmanesh, M., Vilkas, M., Grybauskas, A., & Amran, A. (2024). Generative artificial intelligence in manufacturing: Opportunities for actualizing Industry 5.0 sustainability goals. Journal of Manufacturing Systems, 63, 345–367.
[29] Mariani, M., & Dwivedi, Y. K. (2024). Generative artificial intelligence in innovation management: A preview of future research developments. Journal of Business Research, 165, 114054.
[30] Tabata, M., Wildermuth, C., Bottomley, K., & Jenkins, D. (2025). Generative AI integration in leadership practice: Foundations, challenges, and opportunities. Journal of Leadership Studies, 19(2), 45–67.
[31] Rana, N. P., Pillai, R., Sivathanu, B., & Malik, N. (2024). Assessing the nexus of generative AI adoption, ethical considerations and organizational performance. Technovation, 135, 103064.
[32] Andrieux, P., Johnson, R. D., Sarabadani, J., & Van Slyke, C. (2023). Ethical considerations of generative AI-enabled human resource management. Organizational Dynamics, 52(3), 100975.
[33] Norbäck, P.-J., & Persson, L. (2024). Why generative AI can make creative destruction more creative but less destructive. Small Business Economics, 63(3), 349–377.
[34] Wang, Q. (2025). Generative AI-assisted evaluation of ESG practices and information delays in ESG ratings. Finance Research Letters, 35(1), 100560.
[35] Heigl, R. (2024). Generative artificial intelligence in creative contexts: A systematic review and future research agenda. Management Review Quarterly, 15(3), 123–145.
[36] Fleurence, R. L., Bian, J., Wang, X., Xu, H., Dawoud, D., Higashi, M., & Chhatwal, J. (2025). Generative artificial intelligence for health technology assessment: Opportunities, challenges, and policy considerations: An ISPOR Working Group Report. Value in Health, 28(2), 175–183.
[37] Chowdhury, S., Budhwar, P., & Wood, G. (2023). Generative artificial intelligence in business: Towards a strategic human resource management framework. British Journal of Management, 34(4), 2089–2105.
[38] Krakowski, S. (2025). Human-AI agency in the age of generative AI. Information and Organization, 35(1), 100560.
[39] Modgil, S., Gupta, S., Kar, A. K., & Tuunanen, T. (2025). How could generative AI support and add value to non-technology companies: A qualitative study. Technovation, 25(2), 100–115.
[40] Dogru, T., Line, N., Hanks, L., Acikgoz, F., Abbott, J., Bakir, S., Berbekova, A., Bilgihan, A., Iskender, A., Kizildag, M., Lee, M., Lee, W., McGinley, S., Mody, M., Onder, I., Ozdemir, O., & Suess, C. (2025). The implications of generative artificial intelligence in academic research and higher education in tourism and hospitality. International Journal of Management Education, 21(2), 100822.
[41] Eapen, T. T., Finkenstadt, D. J., Folk, J., & Venkataswamy, L. (2023). How generative AI can augment human creativity. Harvard Business Review, 101(4), 56–65.
[42] Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. Qualitative Research in Psychology, 3(2), 77–101.
[43] Kumar, V., Kotler, P., Gupta, S., & Rajan, B. (2023). Generative AI in marketing: Promises, perils, and public policy implications. Journal of Public Policy & Marketing, 42(3), 203–216.
[44] Chu, W., Baxter, D., & Liu, Y. (2025). Exploring the impacts of generative AI on artistic innovation routines. Technovation, 143, 103209.
[45] Lim, W. M., Gunasekara, A., Pallant, J. L., Pallant, J. I., & Pechenkina, E. (2023). Generative AI and the future of education: Ragnarok or reformation? A paradoxical perspective from management educators. International Journal of Management Education, 21(2), 100790.
[46] Hui, X., Reshef, O., & Zhou, L. (2023). The short-term effects of generative artificial intelligence on employment: Evidence from an online labor market. Organization Science, 34(2), 123–145.
[47] Wei, X., Kumar, N., & Zhang, H. (2025). Addressing bias in generative AI: Challenges and research opportunities in information management. Information & Management, 62(1), 100560.
[48] Cui, Y. (G.), van Esch, P., & Phelan, S. (2024). How to build a competitive advantage for your brand using generative AI. Journal of Business Research, 98(4), 345–367.

AI Framework on Reverse Osmosis Water Treatment for Maintenance Responses

Shfaa Al Baity
Abu Dhabi School of management (ADSM),
Abu Dhabi, United Arab Emirates (UAE)
Adsm-214805@adsm.ac.ae

Dr. Ahmed Jaffar
Associate Professor of Information System,
Abu School of Management (ADSM),
Abu Dhabi, United Arab Emirates (UAE)
a. jaffarAadsm.ac.ae

Abstract—Water scarcity is a growing global challenge, exacerbated by rapid industrialization, population growth, and climate change. Reverse osmosis (RO) has emerged as a key desalination technology to meet the increasing demand for freshwater. However, operational inefficiencies, such as membrane fouling and high maintenance costs, limit its widespread adoption. This research proposes an AI-driven framework to enhance the performance and maintenance of RO membranes. Literature review demonstrates the intial frame work design. Also, it concludes that using hybrid models such as particle swarm optimization (PSO) with artificial neural networks (ANNs) outperformed the use of single model.

Keywords—Framework, AI, RO, …etal (key words**)**

I. INTRODUCTION

Reverse osmosis (RO) is one of the promising techniques for water treatment and wastewater reuse. Membrane fouling is the major challenge that limits the adoption of RO applications due to the difficult control of its performance. With advancements in computation and big data, AI become an essential tool that contributes in solving the challenges and obstacles in various fields of science and engineering. Machine learning is also a subfield of computer science and developed from the study of pattern recognition. Therefore, both AI and ML algorithms have been used for the monitoring and management of membrane fouling. .

A. Background

Due to arid climate in the UAE, the UAE faces the shortage in drinking water. Therefore, the UAE built some of the largest desalination plants not only in GCC but also in the world to meet water demand. These plants convert seawater to drinking water. The UAE strategy focused on using Renewable energy sources such as solar power plants and reduce operational cost. The energy consumed by these plants in the stage to be shifted to sustainable energy sources (solar Power photovoltaic) (Black Ridge Research and Consulting, 2024).

Project Name	Location	Capacity
Jebel Ali Desalination Plant	Dubai	490 MIGD/day
Al Taweelah Desalination Plant	Abu Dhabi	200 MIGD/day
Umm Al Quwain Desalination Plant	Abu Dhabi	150 MIGD/day
Fujairah F1 Desalination Plant	Fujairah	71 MIGD/day
AL Layyah Desalination Plant	Sharjah	51 MIGD/day

Table 1: Top Five RO desalination plants in the UAE.

B. Reverse Osmosis And Configuration

In RO, the feed water is pre-treated, then a high pressure pump as shown in Figure 1c is used to flow the water through the permeable membrane separating salts from water. The pressure, the amount of salt in the input water, and the membrane' salt permeation constant all affect the quality of the water that is generated (Gul, Hruza, & Yalcinkaya, 2021).
The standard RO element configuration (spiral-wound type) is depicted in Figure 1a. One side of the RO element receives the pressurized feed water, which then flows through the feed-side gap between the RO membranes. Concentrate (brine) is released from the opposite side of the RO element, while the permeate water is collected via a middle pipe. As seen in Figure 1b, six to eight RO elements are typically mounted in series within a pressure vessel. The fundamental layout of a seawater desalination plant is one pass and one stage; numerous pressure vessels are placed in parallel according to the facility's capacity (Takabatake,Taniguchi, & Kurihara, 2021).

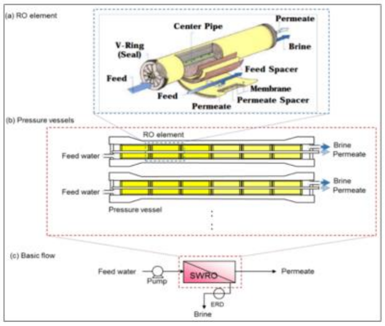


Figure 5: The basic configuration of the RO process. (**a**) Typical configuration of the RO element (spiral-wound type), (**b**) configuration of pressure vessels containing the RO elements, and (**c**) basic flow of single pass and single stage system.

C. Mechanisms Of Membrane Fouling (Cake Formation & Pore Blocking

If the foul-ants (colloids) are smaller than the membrane pores (i.e., solutes), pore blockage and adsorption take place in the interior pore surfaces. On the other hand, a cake layer will typically form on the membrane's surface if the foul-ants (colloids and sludge flocs) are significantly bigger than the membrane holes. Actually, while cake layer can add another layer of resistance to permeation flow, pore blockage enhances the membrane resistance (Ladewig, & Al Shaeli, 2017).

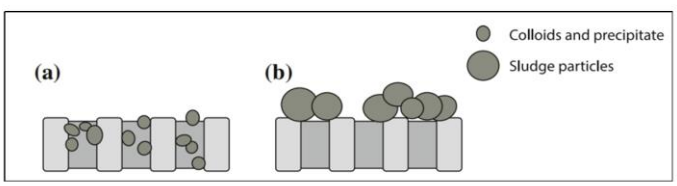


Figure 2: Membrane fouling process via; (a) Pore blocking; (b) Cake layer formation.

D. Motivation

Water scarcity is a major challenge that faced people on this planet and the global demand for freshwater has been increasing by just under 1% per year (Adda et al, 2022, UN-Water, 2024). Along with that, rapid industrialization, and climate change have placed more pressure on world to work more on insuring the existence of enough fresh water for sustaining human life. Along with that, rapid industrialization, and climate change have placed more pressure on world to work more on insuring the existence of enough fresh water for sustaining human life. Over the past half century, several advancements have emerged in the field of water desalination and wastewater treatment. Multistage flash (MSF) distillation, multi-effect evaporation (MEE) distillation, vapour compression (VC) and reverse osmosis (RO) are convenient membrane technologies used to meet the massive demand for freshwater (Al Aani et al, 2019).

E. Problem Statement

Ineffective membrane separation produce low quality water which has several consequences because of the adsorption of unwanted species in the feed water on membrane surface. These species affect membranes' performance in the separation process and produced low quality water that doesn't comply with regulatory standards and requires more additional treatment to make it potable and valid for drinking or other usage. Otherwise the permeate may contain harmful contaminants that pose serious health risk. Also, performing additional treatment to desalinate water will increase the cost because of high consumption of energy which is costly. In addition, customers will use potable water which is costly too till validation of freshwater that in compliance to water quality standards set by regulator based on the use of water.

Repeating unscheduled adhoc maintenance which will increase the cost of operations and need additional time to ensure quality of water, is becoming unbearable. Although periodic maintenance contributes to reduce the fouling issue in RO membranes, more control is required because feed water quality has a large impact on membrane fouling and affect membrane performance. Therefore, predictable membrane maintenance is required which depend on the type and severity of the fouling layer. The predictive maintenance allow performing maintenance work just in time before a failure occurs and shut down the plant. Therefore, predictive maintenance will help in extending the life time of membrane, reducing the downtime of the plant, improve safety, and enhance operational efficiency.

The operating cost of Reverse Osmosis (RO) is a significant barrier to its widespread adoption and popularization. Membrane maintenance and replacement contribute to the annual operating cost since membrane is prone to fouling of several types such as bio-fouling, inorganic, colloidal and organic. Membrane cost depends on the plant capacity and varies between $500 to $1000 per module, which have production rates of 50-100 m3 /d (Hasan, 2019).

Research Objectives

The aim of this research is to propose an AI framework for controlling the effective performance of RO membrane.
The main objectives of this study are;

A. To investigate existing frameworks for effective water treatment

Finding the existing frameworks for water treatment will help in designing the proposed frame work for monitoring the performance of RO membrane. Also, investigating these frameworks will highlight the important information included within these frameworks and a comparison may be conducted (similarities and difference) to strengthen the proposed framework with valuable similarities and differences .

B. To investigate existing AI technologies can be used for effective control of RO membrane

Literature review provides the researcher a broad idea about what type of AI predicting models used and more than that. Also, input and output predictable features found in previous work with the hyper-parameters such as number of layers, number of neurons in each layer, activation function ..et al in case of using Artificial Neural Networks (ANNs) (Abuwatfa et al, 2023).

C. To propose an AI framework for controlling the effective performance of RO membrane

Several framework for water treatment plant can used to extract the proposed framework for the effective performance of RO membrane.

Research Question

The research questions are;

A. What type of data can help in determining fouling issues in RO membrane?

B. Which AI tools (using AI technological devices and algorithms) can help in detecting the degree of fouling issue with RO membrane?

C. Which AI tools can be used for predicting the maintenance of RO membrane?

Research Gap

This literature review revealed that the most successful employed methods for predicting the fouling is the hybrid methods consists of either ANN and ML or optimization methods such as PSO-ANN (Mahadeva et al, 2021). This is consistent with the review paper written by Bagheri et al, 2019. Furthermore, this review explores the need to more comprehensive AI frameworks in water treatment and waster reuse sector. It is appeared that not enough framework discussed in this review due to lack in references for the related topic. There are some review papers that includes information but not a framework with AI.

Paper Structure

The paper comprises of the introduction which includes background, motivation, problem statement, research objectives, research questions, research gap, expected problem solution and literature review. The following sections will be Research ethics and methadology, results and discussion, conclusion and future work.

Literature review

The literature review is conducted to investigate the existing frameworks for effective water treatment and existing AI tools used for effective control of membrane fouling in RO.

A. Comparative analysis

In Table 2, flowchart in figure 2 and approach in figure 3 (Bagheri et al, 2019) emphasize the use of hybrid methods of Artificial intelligence models and optimization techniques in monitoringand controlling the RO membranes. On the other hand, a supervised learning algorithm (Random Forest) is used in framework in figure 4 (Niu et al, 2023). Therefore, the flow

Flow chart/frameworks

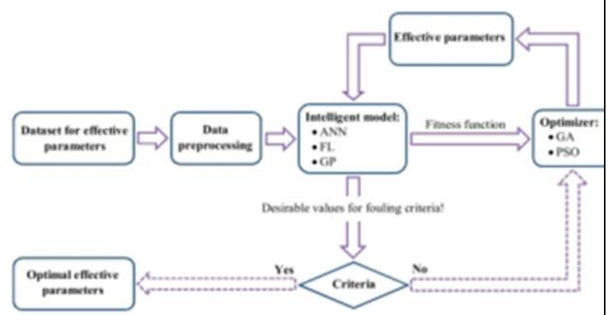


Figure 6: The typical flow chart for the optimization for the effective parameters for membrane mitigation (Bagheri et al, 2019).

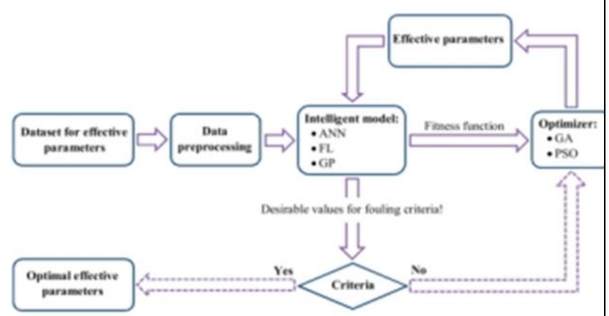


Figure 7: Proposed approach for the advanced control of membrane fouling using artificial intelligence and machine learning technologies (Bagheri et al, 2019).

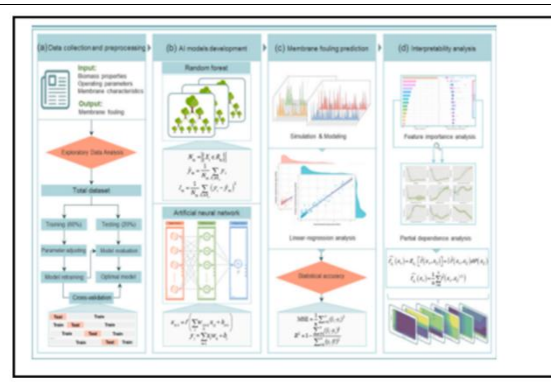


Figure 8: Schematic of AI based membrane fouling modeling framework (Niu et al, 2023).

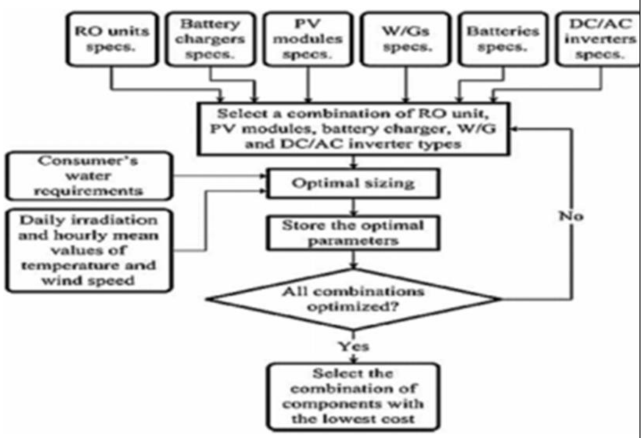


Figure 9: The proposed Optimization methadology (AlAani et al, 2019).

Table 2: Frameworks and flow charts from literature review which can contribute in the initial framework design step in methodology.

charts and frame works in figure 2, 3, and 4 promote the use of hybrid methods of AI and ML. Although, flowchart in figure 2 and 3 have the same process steps for analysis such as data prepossessing and intelligent models used, there are some differences in the used models for predicting membrane fouling. In addition, flow chart in figure 2 mentioned the optimization methods using genetic algorithm (GA) which is also used in figure 5 (Al Aani et al, 2019) to reduce the annual operating cost. On the other hand, approach in figure 3 provides bigger view than figure 2 because it includes the data mining step, feature selection, image recognition, and cluster analysis techniques such as K-means.
Moreover, flow chart in figure 2 shows that artificial neural network (ANN), genetic programming (GP), and fuzzy logic (FL) algorithms with optimization tools such as genetic algorithm (GA) or particle swarm optimization (PSO) are effective in predicting the membrane fouling. In contrast, approach in figure 3 revealed that by the combination of more than one simulation or AI models, severity of membrane fouling can be predicted because it contains image recognition so historical data can be stored for future usage and analysis/simulation.

Approach in figure 5 illustrates the process of optimizing the lowest cost for hybrid technology of RO with sustainable energy sources. In this approach the GA was used to implement and optimize a desalination plant's power supply system that uses sustainble sources of energy. For example, Manesh et al. used GA to do a whole site study and perform an exergoeconomic optimization in order to acquire the best integrated design of the site utility steam network with the hybrid desalination plant which is Multi-effect Distillation-Reverse Osmosis (MED-RO) desalination unit. This improvement's goal was to reduce overall expenses and raise the hybrid system's gain output ratio. In order to establish the global optimum for the Multi-Objective Optimization (MOO) problem, GA determined the Pareto set. With a gain output ratio of 9.1 and an increase in desalinated water production to 126,300 m3/d at a cost of $0.81/m3, the outcome showed the advantages of GA (Manesh et al, 2013).

Regarding the method, although fuzzy algorithm shows good performance in predicting the membrane fouling (Moumni et al, 2022; Galizia et al, 2021), the hybrid model of PSO_ANN outperformed it optimizing sea water reverse osmosis (SWRO) plant performance, reducing costs, and improving efficiency (Mahadeva et al, 2021). In addition, the hybrid models outperforms the single model. As a single model, the ANN out performed the Multi-linear regression model (MLR) in predicting permeate conducttivity, permeate flow rate, and recovery with R^2 of 0.969, 0.942, 0.963 respectively while the Multi Linear Regression (MLR) shows low accuracy where R^2 is around 0.6 (Adda et al, 2022).
In terms of input parameters, the literature review revealed that either all of feed pressure, feed temperature, feed flow rate, feed TDS, feed electrical conductivity or some of them are used as input parameter to AI model while one or two of the permeate flow rate, permeat flux, and recovery rate are the most used as output or predicted variable (Roehl et al, 2018; Adda et al, 2022; Moumni et al, 2022; Mahadeva et al, 2021; Galizia et al, 2021).

Expected Problem Solution

A. Fouling types & Mitigation

Fouling can be classified as reversible or irreversible. The concentration polarization of materials at the membrane rejection surface or the cake layer cause reversible fouling, which includes both backwashable and non-backwashable fouling. While non-backwashable reversible fouling can only be eliminated by chemical cleaning, membranes with backwashable reversible fouling can be restored using the proper physical washing strategy, such as backwashing or hydrodynamic scouring (surface washing). Chemisorption and pore clogging processes cause irreversible fouling. The transmembrane flow loss in the event of irreversible fouling cannot be recovered chemically or hydrodynamically (Guo, NGO, & Li, 2012).

There are four foulant types which are;

1.Particulates: Either organic or inorganic colloids or particles function as filth that can physically blind the membrane surface, obstruct the pores, or prevent transfer to the surface by forming a cake layer;

2. Organic: dissolved substances and colloids that would adsorb to the membrane, such as proteins, hydrophilic and hydrophobic compounds, and humic and fulvic acids;

3. Inorganic: dissolved substances that have a tendency to precipitate onto the membrane surface as a result of oxidation (such as iron or manganese oxides) or pH change (scaling), such as iron, manganese, and silica. Inorganic foulants may also be present as coagulant/flocculant residues.

4. Microbiological organisms: this category includes bacteria and other microorganisms that can stick to membranes and produce biofouling, or the creation of biofilms, as well as vegetative matter like algae (Guo, NGO, & Li, 2012).

Therefore, AI tools used shall be able to detect each type of fouling and predict the degree of severity of this issue with the RO membrane. May be hybrid tool (AI and optimizer/ML) can be suggested. Furthermore, more than one AI tool can be used to detect the behaviour of fouling on RO membrane. Along with AI tools, sensors will supposed to be installed to collect data about temperature humidity, water quality (Salinity, pH, Feed flow rate, permeate flow rate).In addition, suppose that a high quality camera will be used to take high quality images at high magnification with high resolution on the side of the membranes for observing the deposited layer. All these strategies can help in identifying the severity and types of fouling, and notify the maintenance team about the issue of fouling in case maintenance is required.

II. PROPOSED RESEARCH ETHICS AND METHODOLOGY

A. Research Ethics

- Informed Consent: Individuals shall know what data will be collected or shared, how this data will be used, and who will have access to it. Also, any contribution from individuals is easily withdraw-able. ADSM committee approval for the experts shall be taken before conducting the survey.
- Transparency: Clearly show the purpose of data collection or sharing and make individuals know the purpose of providing their data. Any hidden practices must be avoided.
- Accuracy: ensure that data is accurate and up to date.
- Privacy and Security: apply secure techniques such as encryption or anonymize data for protecting collected or shared data from unauthorized access,or misuse.

B. Methodology

The research methodology processes is displayed in figure 6.

B.I Problem evaluation

Addressing the issue accurately help in finding the mitigation. Membrane fouling could results in decreasing the permeate flux but this can be occurred because of polarization concentration. In this case operating conditions can be optimized. But the issue is determining the severity of membrane fouling in order to take the proper action by maintenance team/management.

B.II Comparative analysis for existing research

Conducting a comparative analysis for the existing research for both existing frameworks and the existing AI tools for predicting the membrane fouling from the literature review in order to reveal the best ideas and methods from the existing frameworks and AI tools. Finding the proper frameworks where the researcher can extract ideas and proper AI tools for predicting the issue properly. On the other hand, frameworks and AI tools that does not match the topic specifically such as the membrane technology (RO) or not related to predict the issue was excluded.

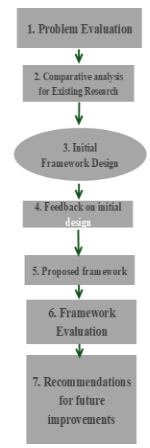


Figure 10: Research Methadology.

B.III Initial Framework Design

After conducting the comparative analysis, the process of collecting and arranging the ideas and tools that best fit the issue for mitigation starts. Developing themes for the framework based on the work and responsibilities. Also, management requirement has taken in consideration in designing the framework.

B.IV Feedback on the Initial Design

The proposed design will be discussed with the supervisor who is the best mentor for this process before sharing the framework with experts. Both qualitative and quantitation data will be collected through conducting a survey for experts in different fields. The google form will be used in collecting the responses. Experts' valuable information can help in improving the understanding of some processes and providing accurate data about technical work.

B.V Proposed Framework

After discuss the experts' feedback and consider it then the framework became the proposed framework.

III. RESULTS & DISCUSSION

Figure 7 demonstrates the initial framework design after literature review . Also, it shows three themes which are; the Reverse Osmosis AI Driven theme, the data Analysis theme , and the management requirements and their Role . The three themes are connected to each other.

- 1st Theme: The Reverse Osmosis (RO) AI Driven

The RO AI-driven theme demonstrates the components and processes of desalination the occurred on the site. So, the components are camera, sensors, network, and server. These component are necessary to be on the site for taking high resolutions images, collecting water quality data, data sharing with management and analysis themes, and uploading the predictable models' software, respectively. Therefore, any observation such as decline on the permeate flux in the site location that can not be optimized will be shared with analysis team for investigation. As a result of this theme operational and water quality data are recorded. Also, images are recorded in order to check fouling presence on RO membranes. All these types of recorded data shared with the analysis team for further analysis.

- 2nd Them: The Data Analysis theme

The aim of this theme is conducting predictive analysis using artificial intelligence predicting models. This theme may contain junior analyst to analyze the data using the suggested AI models. Data are collected using sensors such as time, water quality parameters of both feed and permeate, Operating conditions in the 1st step. Then, data preprocessed where data mining, feature selection, and cross validation are applied to the dataset (Niu et al, 2023). The suggested AI models could be Artificial Neural Nettworks (ANNs), Genetic Programming (GP), and Fuzzy Logic (FL) for Sttructured data such as water quality and operation data recorded on the site. ALso, Convolutional Neural Networsk (CNNs)used for data recorded as images. The data is preprocessed before implementing AI predictive models (Bagheri et al, 2019). Therefore the results of this analysis is considered as input for the management requirement themes to support taking dicision.

- 3rd Theme: The Management Requirements the**me**

After aligning the entity goals with governmental strategic goals and keep following up the main KPIs related to operation and maintenance such as RO availability KPI where change in the performance of RO plant is reflected on these KPIs, the management team can take proper decision based on the Key performance indicators and analysis themes' results. The senior analyst can use the optimization method such as particle swarm optimization (PSO) or genetic algorithm (GA). If the results match and in the range of good operating conditions otherwise an action to be taken is required. So, a meeting/call is conducted to negotiate the status of the plant and analysis results and take a decision of either chemical cleaning or membrane replacement.

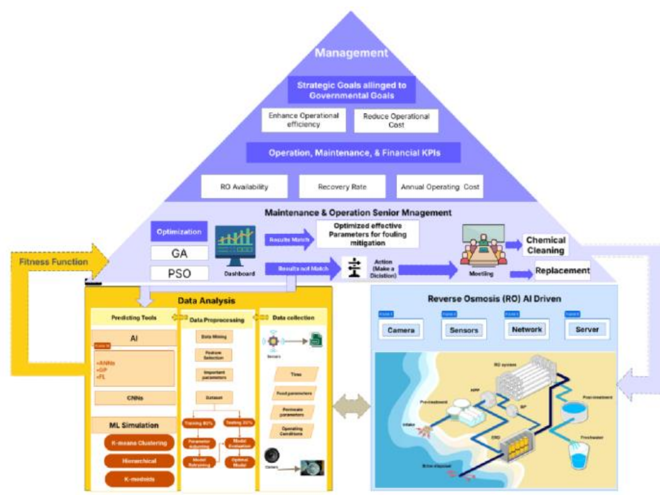


Figure 11: Initial framework design.

IV. CONCLUSION AND FUTURE WORK

The literature review conducted in this study highlights the critical role of artificial intelligence (AI) and machine learning (ML) in addressing the challenges associated with membrane fouling in reverse osmosis (RO) systems, particularly in the context of water treatment and desalination. Membrane fouling remains a significant obstacle in the efficient operation of RO systems, leading to reduce permeate quality, increase the operational costs, and the need for frequent maintenance. The review underscores the potential of AI-driven tools and frameworks to predict, monitor, and control membrane fouling, thereby optimizing the performance of RO systems.

The comparative analysis of existing frameworks and AI tools reveals that hybrid models, such as Particle Swarm Optimization-Artificial Neural Networks (PSO-ANN), outperform single models in predicting membrane fouling. These hybrid models leverage the strengths of both optimization techniques and neural networks, providing more accurate predictions and enabling better decision-making for maintenance teams. Additionally, the review identifies the importance of integrating multiple AI models, such as image recognition and cluster analysis, to enhance the predictive capabilities of these frameworks.

This investigation can be completed by conducting a questionnaire for experts and analyzed the collected data to reach to the final results and the proposed framework.

REFERENCES

[1] BlackRidge Research & Consulting.Top 5 Desalination Plants In UAE. Retrived from Latest List of Top 5 Desalination Plants In UAE [2024].

[2] Gul, A., Hruza, J., & Yalcinkaya, F. (2021). Fouling and chemical cleaning of microfiltration membranes: A mini-review. Polymers, 13(6), 846.

[3] Takabatake, H., Taniguchi, M., & Kurihara, M. (2021). Advanced technologies forstabilization and high performance of seawater ro membrane desalination plants. Membranes, 11(2), 138.

[4] Bradley, P., Ladewig., & Muayad Al Shaeli. (2017). Fouling in Membrane Bioreactors, 10.1007/978-981-10-2014-8_3.

[5] Adda, A., Hanini, S., Bezari, S., Laidi, M., & Abbas, M. (2022). Modeling and optimization of small-scale NF/RO seawater desalination using the artificial neural network (ANN). Environmental Engineering Research, 27(2), 200383. https://doi.org/10.4491/eer.2020.383.

[6] UN-Water, The United Nations World Water Development Report 2024: Water for prosperity and peace.The United Nations World Water Development Report 2024: water for prosperity and peace - UNESCO Digital Library.

[7] Al Aani, S., Bonny, T., Hasan, S. W., & Hilal, N. (2019). Can machine language and artificial intelligence revolutionize process automation for water treatment and desalination?. Desalination, 458, 84-96.

[8] Shadi, W., Hasan.Personal Communication. (2019).

[9] Abuwatfa, W. H., AlSawaftah, N., Darwish, N., Pitt, W. G., & Husseini, G. A. (2023). A review on membrane fouling prediction using artificial neural networks (ANNs). Membranes, 13(7), 685.

[10] W., Guo, H. H., Ngo, & J., Li. (2012). A mini-review on membrane fouling. Bioresource technology, 122, 27-34.

[11] Mahadeva, R., Kumar, M., Manik, G., & Patole, S. P. (2021). Modeling, simulation, and optimization of the membrane performance of seawater reverse osmosis desalination plant using neural network and fuzzy based soft computing techniques. Desalination and Water Treatment, 229, 17–30. https://doi.org/10.5004/dwt.2021.27386

[12] Bagheri, M., Akbari, A., & Mirbagheri, S. A. (2019). Advanced control of membrane fouling in filtration systems using artificial intelligence and machine learning techniques: A critical review. Process Safety and Environmental Protection, 123, 229–252. https://doi.org/10.1016/j.psep.2019.01.013

[13] Niu, C., Li, B., & Wang, Z. (2023). Using artificial intelligence-based algorithms to identify critical fouling factors and predict fouling behavior in anaerobic membrane bioreactors. Journal of Membrane Science, 687, 122076. https://doi.org/10.1016/j.memsci.2023.122076

[14] Roehl, E. A., Ladner, D. A., Daamen, R. C., Cook, J. B., Safarik, J., Phipps, D. W., & Xie, P. (2018). Modeling fouling in a large RO system with artificial neural networks. Journal of Membrane Science, 552, 95–106. https://doi.org/10.1016/j.memsci.2018.01.064

[15] Khayet, M., Cojocaru, C., & Essalhi, M. (2011). Artificial neural network modeling and response surface methodology of desalination by reverse osmosis. Journal of Membrane Science, 368(1-2), 202–214. https://doi.org/10.1016/j.memsci.2010.11.030

[16] Galizia, A., Mamo, J., Blandin, G., Verdaguer, M., Comas, J., Rodríguez-Roda, I., & Monclús, H. (2021). Advanced control system for reverse osmosis optimization in water reuse systems. Desalination, 518, 115284. https://doi.org/10.1016/j.desal.2021.115284

[17] Moumni, M., & Massour el Aoud, M. (2022). Fuzzy logic control of a brackish water reverse osmosis desalination process. Computers and Chemical Engineering, 167, 108026. https://doi.org/10.1016/j.compchemeng.2022.108026

Driver Behavior Assessment Using Machine Learning in ADAS and ITS

Asia Osman Mohamed Sharif
Business Analytics in Artificial Intelligent Management
Abu Dhabi School of Management (ADSM)
UAE, Abu Dhabi
Asialish99@yahoo.com

Abstract—Unsafe driving behaviors such as fatigue, aggressive acceleration, and distraction contribute to 94% of traffic accidents and significantly impact fuel efficiency and emissions. Current ITS and ADAS technologies are largely reactive, lacking predictive capabilities for real-time risk prevention. This study presents a machine learning-based approach to proactively assess driver behavior using Random Forest, XGBoost, and Neural Networks. Data is collected from telematics systems, GPS tracking, on-board diagnostics, and driver monitoring sensors across public transport fleets. The methodology involves feature engineering, model training, cross-validation, and real-world scenario testing. The proposed system aims to enhance road safety, reduce fuel consumption, and lower emissions by integrating predictive analytics into existing ITS and ADAS infrastructure. The research adheres to ethical standards, ensuring data privacy and fairness, and offers actionable insights for transport authorities, policymakers, and fleet managers seeking intelligent and sustainable mobility solutions. The findings offer practical insights for transport operators, policymakers, and safety advocates toward achieving sustainable and intelligent mobility solutions.

Keywords—Driver Behavior, Machine Learning, ADAS, ITS, Telematics Data, Predictive Modeling, Road Safety, XGBoost, Neural Networks, Fuel Efficiency.

I. INTRODUCTION

The evolution of Intelligent Transportation Systems (ITS) and Advanced Driver Assistance Systems (ADAS) has significantly shaped modern transportation by improving monitoring and control mechanisms. However, persistent challenges in road safety, fuel efficiency, and environmental sustainability remain, particularly due to unsafe driver behaviors such as fatigue, distraction, and speeding. These behaviors are responsible for an estimated 94% of road accidents, leading to about 1.3 million deaths annually (WHO, 2021; NHTSA, 2020). Furthermore, aggressive driving can increase fuel consumption and emissions by up to 40% and 30%, respectively (IEA, 2021).

While existing ADAS and ITS technologies offer real-time alerts, they are largely reactive, triggering interventions after risky behavior has already occurred. This study proposes a proactive solution using machine learning (ML) to predict unsafe driving behaviors before they manifest. By leveraging algorithms such as Random Forest (RF), Extreme Gradient Boosting (XGBoost), and Neural Networks, the system can detect patterns of risk in real time.

The aim of this study is to provide a predictive analytics model integrated with telematics and sensor data, enhancing road safety, operational efficiency, and reducing environmental impact. This is particularly valuable for public transport fleets, where vehicles operate under diverse and challenging urban conditions. The model supports informed decision-making for fleet operators, regulators, and policymakers while maintaining ethical data usage and fairness.

The study focuses on evaluating a predictive ML model that can anticipate risky driving behavior and integrate it into real-time ITS and ADAS applications. The core objectives are:

Analyze driver behavior patterns using real-time vehicle telematics data.

Evaluate and compare ML techniques (Random Forest, XGBoost, Neural Networks) for behavior prediction.

Optimize model accuracy through feature selection, hyperparameter tuning, and ensemble learning.

Validate ML models using real-world driving scenarios.

This study focuses on public transport fleet drivers operating in urban environments. Data sources will include telematics data, GPS tracking, vehicle sensors, and driver biometrics.
The study will specifically target:

- Prediction of unsafe driving events (e.g., harsh braking, speeding, distracted driving). Impact of driving behavior on fuel efficiency and emissions.
- Integration of ML-based insights into existing ADAS technologies.

The research addresses primary questions:

- Which algorithm performs best accuracy in predicting unsafe driving within ITS/ADAS settings
- Prediction of unsafe driving events (e.g., harsh braking, speeding, distracted driving). Impact of driving behavior on fuel efficiency and emissions.

Integration of ML-based insights into existing ADAS technologies.

I. Proposed Research Approach and Methodology

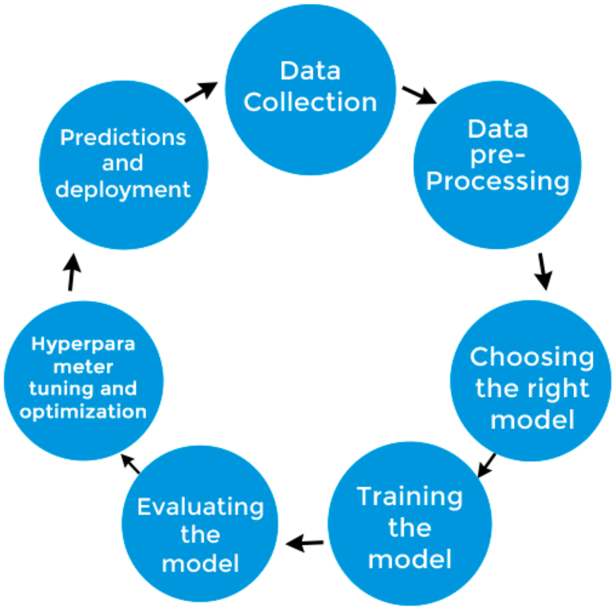


Figure 1: Methodology Overview

The study adopts a multi-phase methodology to ensure a robust evaluation of driver behavior using machine learning. The process begins with comprehensive data collection from telematics, OBD, and driver monitoring systems, targeting public transport drivers in urban UAE settings. Data are preprocessed through feature engineering to extract relevant indicators of unsafe driving behavior. Multiple ML models—Random Forest, XGBoost, CNN-LSTM, and GNN—are trained and validated. Model performance will be assessed using metrics such as accuracy, F1-score, and ROC-AUC through k-fold cross-validation and real-world driving scenario tests to ensure generalizability and fairness.

Table 1: Research Approach and Methodology Flow

Section	Description
Research Ethics	-Ensures compliance with GDPR and data privacy laws, anonymization of driver data, and mitigation of bias in ML models for fairness.
The Methodology	-Multi-phase approach integrating data collection, feature engineering, ML model development, and real-world validation.
Data Sources & Collection	- Sources: On-board diagnostics (OBD) data, GPS tracking systems, CAN bus signals, and Camera-based driver monitoring systems. - Collection: IoT sensors, vehicle diagnostics, and real-world driving datasets.
Analytical Approach	- Feature Engineering: Identifies key behavioral indicators (speed, braking, fatigue, lane changes). - ML Models: Random Forest, XGBoost, CNN-LSTM, Graph Neural Networks (GNNs). - Evaluation: Accuracy, F1-score, ROC-AUC, real-world validation.
Required Technology	- Software: Python (Scikit-learn, TensorFlow, PyTorch), MATLAB. - Hardware: Cloud-based computing, edge AI for real-time processing.

II. LITERATURE REVIEW

Prior studies have demonstrated the potential of machine learning (ML) to predict and understand driver behavior. Khan et al. (2021) utilized artificial neural networks (ANNs) on a dataset of 94,000 instances with 54 features, achieving a high correlation (0.9962) between predicted and actual driver actions. However, they noted instability in early training, which limits real-time adaptability. Niu et al. (2021) applied traditional classifiers such as GBDT, AdaBoost, and Random Trees to detect unsafe truck driver behaviors. Model accuracy ranged from 0.64 to 0.95, and F1 scores varied from 0.52 to 0.72, highlighting challenges in predicting rare, context-specific behaviors. Their study also underscored the value of using correlation and odds ratios to interpret risk.

Abdelrahman et al. (2020) developed a Random Forest-based risk profiling model using the SHRP2 dataset, achieving 90% accuracy and an F1-score of 0.945. Their framework supported continuous learning via cloud-based updates. In vision-based systems, Qu et al. (2024) used a hybrid CNN-BiLSTM model to detect distracted driving, achieving 91.7% accuracy but struggling with full-body posture recognition. This reveals the need for multimodal systems integrating visual and sensor data.

Table 2: Summary of the Literature Review

Author/Year	Purpose/Objectives	Methodology, Findings, Discussion
Khan et al. (2021)	Develop an ANN-based driver behavior model to predict intelligent driving patterns.	Used Dynamic Nonlinear Autoregressive ANN validated through time-series backpropagation, multilayer perceptron (MLP), random subspace, linear regression, and decision tree. Dataset: 94,380 instances, 54 attributes (MATLAB & Weka). MLP had the highest accuracy: 0.9962 correlation coefficient, 30.39 MAE, 69.44 RMSE. Findings showed ANN models effectively capture behavioral changes in ADAS and ITS. However, early-phase fluctuations limit real-time adaptation.
Niu et al. (2021)	Investigate unsafe truck driver behavior using classification models.	Surveyed 2,000 truck drivers and classified behavior based on six first-level input dimensions and 51 second-level indicators. Compared GBDT, AdaBoost, RT, and CART models. Accuracy varied by behavior type: Classification accuracy (0.64–0.95), F1-scores (0.52–0.72). Findings showed difficulty in predicting behaviors due to different formation mechanisms. Recommended a

		hybrid ML approach integrating context-aware data.
Elassad et al. (2020)	Review ML applications in Driving Behavior (DB) and propose a conceptual framework (Driver-Vehicle-Environment System).	Systematic Literature Review (SLR) of 82 studies (2009–2019). Identified 8 widely used ML models for DB. ML outperformed traditional models in prediction accuracy. Challenges: Lack of standard evaluation frameworks and model generalizability. Unlike Khan et al. (2021) and Niu et al. (2021), emphasized holistic behavioral modeling.
Abdelrahman et al. (2020)	Propose a machine learning-based driver profiling framework for risk assessment in fleet management and insurance telematics.	Used SHRP2 dataset (9,000 crash events, 20,000 baseline events). Compared Random Forest (90% accuracy, F1-score 0.945), Deep Neural Networks (88.8%), and Extreme Learning Machines (86.2%). Findings: ML-based risk assessment is scalable, supports IoIV, and improves insurance risk evaluation.
Ontañón et al. (2020)	Improve long-term driving behavior prediction through indirect prediction.	Evaluated Baseline averaging, Linear Regression, M5P regression trees, and MLP neural networks. Linear regression had the lowest MSE (0.03 for steering, 6.79 for throttle, 1.26 for braking). However, long-term prediction errors compounded, limiting supervised learning models. Proposed context-based reasoning and indirect prediction to improve reliability.
Qu et al. (2024)	Computer vision and ML are used to classify driver behavior via distracted driver monitoring.	Applied CNN-BiLSTM models to detect distracted driver postures. Evaluated 10 deep learning models on AUC's Distracted Driver Dataset. Best model: CNN-BiLSTM (91.7% accuracy, 93.1% F1-score). Despite high accuracy, noted CNNs struggle to analyze full-body postures. Recommended enhancing feature engineering for ADAS.
Azadani & Boukerche (2021)	Apply ITS-oriented approach to driving behavior analysis (DBA).	Integrated in-vehicle networks, sensors, and communication technologies to improve traffic safety, fuel efficiency, and driver risk profiling. Unlike Qu et al. (2024), which relied on computer vision-based CNN models, ITS methods provided sensor-driven, multimodal risk assessment.
Chandra et al. (2021)	Apply graph-theoretic ML to predict driver behavior in AV navigation.	Used StylePredict (GCN & MLP models) on real-world traffic datasets (U.S., India, China, Singapore). Findings: AVs adapted lane-changing behavior dynamically—in conservative traffic, AVs overtook confidently; in aggressive traffic, fewer lane changes reduced risk. AV speeds: 29 m/s (aggressive) vs. 19.7 m/s (conservative). Unlike traditional AV models, graph-based reinforcement learning enables social awareness and efficient navigation.

III. RESEARCH GAPS

Despite progress in machine learning (ML) for driver behavior analysis, key gaps remain. Current models lack generalizability across varied environments and suffer from bias, reducing predictive accuracy (Khan et al., 2021; Niu et al., 2021). Hybrid approaches using ML, computer vision, and sensor fusion show potential (Qu et al., 2024), but struggle with real-time adaptability (Chandra et al., 2021). The absence of standardized evaluation frameworks limits cross-study comparisons (Elassad et al., 2020), and data privacy concerns restrict large-scale deployment.

ML models can predict risky behavior and improve ADAS, but often ignore long-term reliability and fluctuating driving conditions (Ontañón et al., 2020). CNNs excel in classification but lack full-body posture detection for distraction analysis (Qu et al., 2024). Sensor-based methods require integration with real-time risk assessment (Azadani & Boukerche, 2021).

Future research should focus on hybrid ML models that integrate heterogeneous data sources, such as biometrics and environmental variables (Imani et al., 2025), and expand to public transport systems and sustainability impacts (Moujahid et al., 2018; EPA, 2024).

IV. CONCLUSION

This study proposes a predictive driver behavior modeling using machine learning to enhance safety, efficiency, and environmental sustainability through integration with ITS and ADAS.

This research contributes in several meaningful ways:

A robust ML-driven predictive model that supports real-time risk mitigation.

A comparative model evaluation across RF, XGBoost, Neural Networks, and GNNs.

A rich dataset and feature set tailored to the needs of urban public transport fleets.

An ethical and practical implementation strategy ensuring adaptability and acceptance.

The solution benefits fleet operators (through improved efficiency and safety), policymakers (via data-driven interventions), and insurance providers (via more accurate risk profiling). The project also supports sustainability by reducing fuel usage and emissions through better driving practices.

This work directly addresses the core research objective of evaluating predictive ML models for risky driving behavior using real-time sensor data. The findings provide a foundation for future pilot implementations in UAE's public transportation sector. Subsequent research could explore integrating these models into live ITS dashboards, enhancing real-time traffic interventions and policy planning.

REFERENCES

[1] Imani, M., Beikmohammadi, A., & Arabnia, H. R. (2025). Comprehensive Analysis of Random Forest and XGBoost Performance with SMOTE, ADASYN, and GNUS Upsampling Under Varying Imbalance Levels.

[2] Abdelrahman, A. E., Hassanein, H. S., & Abu-Ali, N. (2020). Robust data-driven framework for driver behavior profiling using supervised machine learning. IEEE transactions on intelligent transportation systems, 23(4), 3336-3350. https://www.queenstrl.ca/uploads/4/6/3/1/4631596/robust_data-driven_framework_for_driver_behavior_profiling_using_supervised_machine_learning.pdf

[3] Adavikottu, A., & Velaga, N. R. (2021). Analysis of factors influencing aggressive driver behavior and crash involvement. Traffic injury prevention, 22(sup1), S21-S26. https://doi.org/10.1080/15389588.2021.1965590

[4] Azadani, M. N., & Boukerche, A. (2021). Driving behavior analysis guidelines for intelligent transportation systems. IEEE transactions on intelligent transportation systems, 23(7), 6027-6045. http://dx.doi.org/10.1109/TITS.2021.3076140

[5] Bouhsissin, S., Sael, N., & Benabbou, F. (2023). Driver behavior classification: a systematic literature review. IEEE Access, 11, 14128-14153. https://doi.org/10.1109/ACCESS.2023.3243865

[6] Chandra, R., Bera, A., & Manocha, D. (2021). Using graph-theoretic machine learning to predict human driver behavior. IEEE Transactions on Intelligent Transportation Systems, 23(3), 2572-2585. http://dx.doi.org/10.48550/arXiv.2111.02964

[7] Elassad, Z. E., Mousannif, H., Al Moatassime, H., & Karkouch, A. (2020). The application of machine learning techniques for driving behavior analysis: A conceptual framework and a systematic literature review. Engineering Applications of Artificial Intelligence, 87, 103312. https://doi.org/10.1016/j.engappai.2019.103312

[8] EPA. (2024). Carbon Pollution from Transportation. https://www.epa.gov/transportation-air-pollution-and-climate-change/carbon-pollution-transportation

[9] Garikapati, D., & Shetiya, S. S. (2024). Autonomous vehicles: Evolution of artificial intelligence and the current industry landscape. Big Data and Cognitive Computing, 8(4), 42. https://doi.org/10.3390/bdcc8040042

[10] Hagl, M., & Kouabenan, D. R. (2020). Safe on the road–does Advanced driver-assistance systems Use affect Road Risk Perception?. Transportation research part F: traffic psychology and behaviour, 73, 488-498. https://doi.org/10.1016/j.trf.2020.07.011

[11] Imani, M., Beikmohammadi, A., & Arabnia, H. R. (2025). Comprehensive Analysis of Random Forest and XGBoost Performance with SMOTE, ADASYN, and GNUS Under Varying Imbalance Levels. Technologies, 13(3), 88. doi: 10.20944/preprints202501.2274.v2

[12] Iyer, L. S. (2021). AI-enabled applications towards intelligent transportation. Transportation Engineering, 5, 100083. https://doi.org/10.1016/j.treng.2021.100083

[13] International Energy Agency (IEA). (2021). The role of digitalization in reducing transport emissions. IEA Energy Reports.

[14] Khan, Q. T. A., Abbas, S., Khan, M. A., Fatima, A., Alanazi, S., & Elmitwally, N. S. (2021). Modelling Intelligent Driving Behaviour Using Machine Learning. Computers, Materials & Continua, 68(3). http://dx.doi.org/10.32604/cmc.2021.015441

[15] Markets and Markets. (2022). Artificial intelligence in transportation market forecast 2020–2030. Industry Analysis Report.

[16] Moujahid, A., Tantaoui, M. E., Hina, M. D., Soukane, A., Ortalda, A., ElKhadimi, A., & Ramdane-Cherif, A. (2018, June). Machine learning techniques in ADAS: A review. In 2018 International Conference on Advances in Computing and Communication Engineering (ICACCE) (pp. 235-242). IEEE. https://www.researchgate.net/profile/Manolo-Hina/publication/327192771_Machine_Learning_Techniques_in_ADAS_A_Review/links/6038bff992851c4ed599b565/Machine-Learning-Techniques-in-ADAS-A-Review.pdf

[17] National Highway Traffic Safety Administration (NHTSA). (2020). The impact of human factors on traffic accidents. Traffic Safety Research Journal, 57(2), 83-96.

[18] Niu, Y., Li, Z., & Fan, Y. (2021). Analysis of truck drivers' unsafe driving behaviors using four machine learning methods. International Journal of Industrial Ergonomics, 86, 103192. https://doi.org/10.1016/j.ergon.2021.103192

[19] Oladimeji, D., Gupta, K., Kose, N. A., Gundogan, K., Ge, L., & Liang, F. (2023). Smart transportation: an overview of technologies and applications. Sensors, 23(8), 3880. https://doi.org/10.3390/s23083880

[20] Ontanon, S., Lee, Y. C., Snodgrass, S., Winston, F. K., & Gonzalez, A. J. (2017). Learning to predict driver behavior from observation. In Proc. AAAI Spring Symp. Learn. Observ. Hum. (pp. 506-512). https://cdn.aaai.org/ocs/15303/15303-68248-1-PB.pdf

[21] Qu, F., Dang, N., Furht, B., & Nojoumian, M. (2024). Comprehensive study of driver behavior monitoring systems using computer vision and machine learning techniques. Journal of Big Data, 11(1), 32. https://doi.org/10.1186/s40537-024-00890-0

[22] Roussou, S., Garefalakis, T., Michelaraki, E., Brijs, T., & Yannis, G. (2024). Machine Learning Insights on Driving Behaviour Dynamics among Germany, Belgium, and UK Drivers. Sustainability, 16(2), 518. https://doi.org/10.3390/su16020518

[23] Taherdoost, H. (2022). Different types of data analysis; data analysis methods and techniques in research projects. International Journal of Academic Research in Management, 9(1), 1-9. Different Types of Data Analysis; Data Analysis Methods and Techniques in Research Projects by Hamed Taherdoost :: SSRN

[24] Tamascelli, N., Solini, R., Paltrinieri, N., & Cozzani, V. (2022). Learning from major accidents: A machine learning approach. Computers & Chemical Engineering, 162, 107786. https://doi.org/10.1016/j.compchemeng.2022.107786

[25] Tan, H. (2021, August). Machine learning algorithm for classification. In Journal of Physics: Conference Series (Vol. 1994, No. 1, p. 012016). IOP Publishing.

[26] Wang, F. Y., Lin, Y., Ioannou, P. A., Vlacic, L., Liu, X., Eskandarian, A., ... & Olaverri-Monreal, C. (2023). Transportation 5.0: The DAO to safe, secure, and sustainable intelligent transportation systems. IEEE Transactions on Intelligent Transportation Systems. https://doi.org/10.1109/TITS.2023.3305380

[27] World Health Organization (WHO). (2021). Global road safety status report. WHO Publications. https://apps.who.int/iris/bitstream/handle/10665/336584/9789289054980-eng.pdf

[28] Zhang, Y., Wu, M., Tian, G. Y., Zhang, G., & Lu, J. (2021). Ethics and privacy of artificial intelligence: Understandings from bibliometrics. Knowledge-Based Systems, 222, 106994. https://doi.org/10.1016/j.knosys.2021.106994

Generative AI for Public Transportation in the UAE – a field study and analytic framework

Maram Almazrouei
College of Technological Innovation
Zayed University
Abu Dhabi, UAE
M80008963@zu.ac.ae

Ravishankar Sharma
College of Technological Innovation
Zayed University
Abu Dhabi, UAE
Ravishankar.Sharma@zu.ac.ae

Abstract—This paper examines how generative artificial intelligence can support the development of sustainable and energy-efficient public transportation in the United Arab Emirates. It analyzes current applications in Dubai and Abu Dhabi, identifies barriers to wider adoption, and explores opportunities for optimizing routes and energy use. A qualitative research method was used, involving expert interviews and thematic analysis supported by a Large Language Model. Findings reveal Dubai's advanced adoption compared to Abu Dhabi's pilot initiatives. Key barriers include regulatory gaps, infrastructure and data limitations, and cultural resistance. A five-phase framework is proposed to guide systematic AI deployment toward net-zero goals.

Keywords—generative artificial intelligence, public transportation, net-zero, sustainable mobility, route optimization, AI framework

I. BACKGROUND

Public transportation is an important part of the urban mobility system in the UAE. It relies heavily on systems to meet the growing demands of urbanization with networks of buses, metro, and trams serving millions of passengers, cities like Dubai and Abu Dhabi [1]. For example, the Dubai Metro is one of the world's largest driverless metro systems, supported by an extensive bus and tram network [2]. However, despite these achievements, the UAE's public transportation sector still faces challenges in achieving long-term sustainability, especially regarding reducing emissions while meeting increasing demand [3].

The UAE has made significant commitments to sustainability through national strategies such as the UAE Energy Strategy 2050 and the Net Zero by 2050 Strategic Initiative [1]. These strategies aim to lower carbon emissions and increase the role of renewable energy, particularly in sectors such as transportation. Although progress has been made, the transport sector still accounts for a significant share of the country's greenhouse gas emissions. This highlights the urgent need for new and innovative solutions to make public transportation more efficient and environmentally friendly.

Generative artificial intelligence (AI) is one such promising solution. It is a fast-growing technology known for its ability to create, predict, and optimize complex systems. In public transportation, generative AI can improve route planning, reduce energy use, and support predictive maintenance. These features can contribute to building a more sustainable transport system that helps the UAE reach its environmental goals.

Despite improvements in AI technologies globally, challenges remain in integrating generative AI into the UAE's public transportation system. High energy consumption, dependence on fossil fuels, traffic congestion, and inefficient routes continue to impact sustainability. Transportation growth adds further pressure, making it even more difficult to create an efficient and eco-friendly transport system.

This paper proposes a detailed framework for applying generative AI in the UAE's public transportation sector to respond to this gap. The aim is to show how AI can improve system performance, reduce emissions, and contribute to national climate goals. The study focuses on three main areas where AI can make a difference: real-time optimization of transport routes, predictive traffic modeling, and energy consumption reduction. The framework introduced in this paper offers a structured, five-phase approach that aligns with national strategies such as the UAE's Net Zero 2050 goals.

The proposed framework includes five phases: Assessment, Planning, Implementation, Monitoring, and Scaling. These phases provide a step-by-step guide for decision-makers and transport authorities, showing how generative AI can be gradually and effectively introduced. Each phase focuses on specific goals—identifying baseline emissions and system readiness to scaling successful AI projects across emirates. This structure supports a smooth and measurable transition toward smarter, cleaner, and more efficient transportation.

This study aims to contribute practical recommendations for the future use of generative AI in UAE transportation systems. It aligns closely with national sustainability goals and supports the country's broader efforts to lead in innovation and climate action. The study focuses on land-based transportation, such as buses, trams, and metro systems. Other modes of transport, such as aviation or maritime, are beyond the scope of this paper.

The research addresses three core questions:

How is generative AI currently utilized in Dubai and Abu Dhabi public transportation?

What are the major barriers to deploying generative AI across the UAE's transport systems?

How can generative AI support real-time route optimization and energy management toward achieving net-zero goals?

II. METHODOLOGY

This study adopts a qualitative research approach to develop a structured framework for deploying generative artificial intelligence within the UAE's public transportation system. A qualitative method was chosen because it focuses on exploring expert opinions and understanding the practical, technical, and strategic challenges involved in integrating AI technologies to achieve sustainability goals.

The research follows an interpretive paradigm, seeking to extract recurring themes, patterns, and expert perspectives to build meaningful insights. To enhance the analysis, thematic analysis was explicitly used as the main qualitative method. Claude 3, a generative AI model, supported the thematic analysis by helping organize the interview data, categorize responses, and identify emerging themes efficiently.

Primary data were collected through semi-structured interviews with five experts selected through purposive sampling. These experts represented diverse fields including transportation policy, AI development, environmental sustainability, and public sector strategy. Interviews were conducted either virtually or in person, recorded with consent, and transcribed for detailed analysis. Interview questions focused on exploring current applications of AI, identifying barriers to AI adoption, and understanding how AI could contribute to the UAE's Net Zero 2050 goals.

Secondary data were also reviewed, including government transportation reports, national sustainability plans, and academic literature. This ensured that the findings were supported by broader data sources and grounded in relevant policy contexts.

1) Data Analysis Process

Thematic analysis proceeded through a step-by-step process:

1. **Transcription and anonymization** of all expert interviews.
2. **Uploading transcripts** to Claude 3's platform for AI-assisted coding.
3. **Prompting Claude 3** with targeted queries to extract themes related to AI applications, barriers, regional differences (Dubai vs. Abu Dhabi), and sentiment toward AI adoption.
4. **Manual review** of Claude 3 outputs to ensure thematic accuracy and contextual understanding.
5. **Cross-validation** using Napkin AI-generated visual summaries to confirm the consistency and depth of themes.
6. **Inter-coder reliability testing** by reanalyzing 20% of the transcripts manually, achieving over 90% agreement.

The findings from this thematic analysis directly informed the development of the five-phase framework proposed in the study.

2) Framework Development and Propositions

Based on the analysis, a **five-phase deployment framework** was developed. The phases are:

- **Assessment**: Evaluating AI readiness, stakeholder roles, and infrastructure gaps.
- **Planning**: Defining AI use cases, setting strategic goals, and developing governance models.
- **Implementation**: Piloting AI solutions, upgrading digital infrastructure, and training operational staff.
- **Monitoring**: Tracking performance through KPIs, such as emissions reductions and service efficiency.
- **Scaling**: Expanding successful AI initiatives across emirates and integrating emerging technologies.

Rather than hypotheses, the study presents **three propositions**:

- **Proposition 1**: Dubai demonstrates greater readiness and operational integration of generative AI compared to Abu Dhabi.
- **Proposition 2**: Barriers such as policy fragmentation, technical limitations, and cultural resistance limit AI deployment in the UAE.
- **Proposition 3**: Generative AI can support at least a 20% reduction in transportation-related greenhouse gas emissions through optimized routing, energy management, and predictive maintenance.

These propositions are explored and tested within the context of the framework development.

3) Validation Strategy

To ensure the validity and reliability of findings:

- **Standardized interview protocols** were applied across all interviews.
- **Data triangulation** was conducted by combining primary interviews, secondary sources, and literature reviews.
- **Expert feedback** was sought after the initial analysis to validate interpretations.
- **Inter-coder reliability** was checked to maintain thematic consistency.

4) Future Empirical Validation

As this research is exploratory and framework-driven, empirical validation is proposed as **future work**. Future validation could involve:

- **Pilot case studies** testing the five-phase framework in selected transport systems (e.g., Dubai's RTA).
- **Follow-up expert interviews** to refine the framework based on real-world deployment experiences.
- **Quantitative measurement** of emissions reductions, operational improvements, and user acceptance in AI-enabled transport networks.

This future empirical validation would strengthen the generalizability and practical relevance of the proposed framework.

III. RESULTS AND FRAMEWORK PROPOSAL

A. Current Use of Generative AI in Dubai and Abu Dhabi

Dubai has shown clear progress in integrating AI within its public transportation system. The Roads and Transport Authority (RTA) has adopted technologies such as real-time route optimization, smart traffic control, on-demand mobility services, and predictive maintenance. One expert shared that generative AI is being used to re-route buses during congestion, which improves travel time and reduces emissions. Another mentioned how the RTA uses AI simulations for long-term transport planning, contributing to its broader smart city strategy.

The digital infrastructure of Dubai enables AI solutions like congestion mapping, demand forecasting, and electric vehicle charging optimization to operate effectively. Experts emphasized that Dubai's success is linked to strong institutional support and strategic planning. In contrast, Abu Dhabi is still developing its approach. While it has launched pilot projects such as electric bus tracking and AI-supported traffic control, implementation is limited. Experts noted gaps in infrastructure, policy coordination, and cross-sector collaboration. Abu Dhabi's use of AI remains experimental, highlighting the need for a centralized roadmap and increased investment.

A. Barriers to Generative AI Deployment

Analysis revealed four primary barriers:

1. Policy gaps: There is no unified regulatory framework for AI in transport. Differences across emirates and lack of clarity around ethics, data sharing, and legal responsibility limit progress.
2. Infrastructure readiness: Experts noted uneven access to digital tools and systems. Outside of Dubai, many areas lack sensors, smart platforms, and EV charging infrastructure, making full AI deployment difficult.
3. Skills and data: Most transport agencies face limited access to high-quality, real-time data and a shortage of trained personnel to operate AI systems.
4. Cultural resistance: Public users and staff may distrust automated systems without clear communication or training. This slows adoption and limits the perceived value of AI.

B. Transformative Potential and Framework Relevance

Despite challenges, experts agreed that generative AI can significantly improve operational efficiency, reduce emissions, and support data-driven transportation planning. Benefits include real-time route planning, predictive maintenance, and energy management. AI can forecast demand, align services with renewable energy, and optimize electric fleet performance. When fully adopted, AI could reduce emissions by at least 20%, based on expert estimates.

A five-phase framework was proposed: Assessment, Planning, Implementation, Monitoring, and Scaling to supports a structured rollout aligned with UAE sustainability goals and ensures flexibility for long-term innovation.

In summary, generative AI is viewed as a technical and strategic tool for advancing smart, efficient, and climate-friendly transportation in the UAE. While readiness levels vary across emirates, continued investment, stronger regulation, and stakeholder coordination will be critical to scaling its impact and meeting national sustainability targets.

C. Proposed Framework for AI Deployment

The paper proposes a structured five-phase deployment framework to address the identified challenges and enable the scalable adoption of generative artificial intelligence in public transportation. The first phase, Assessment, involves evaluating the current state of the transportation system, including its carbon footprint, the readiness of digital infrastructure, the quality and availability of relevant data, and the institutional capacity to support AI integration to ensure that any AI strategy is built on a clear understanding of existing conditions. The second phase, Planning, focuses on defining specific AI use cases such as route optimization, predictive maintenance, and energy forecasting and aligning them with the UAE's national climate goals. It also includes the development of data governance policies, ethical guidelines, and performance targets. In the Implementation phase, generative AI tools are tested through pilot projects in high-impact areas. The fourth phase, Monitoring, ensures that the performance of AI systems is evaluated using real-time key performance indicators such as emissions reductions, fuel efficiency, and user satisfaction that allow for continuous improvement and data-driven adjustments. Finally, the Scaling phase focuses on expanding successful pilot initiatives across additional emirates and transport modes. Overall, this five-phase framework provides a practical and adaptable roadmap for guiding the deployment of generative AI in alignment with the UAE's broader sustainability and smart mobility strategies.

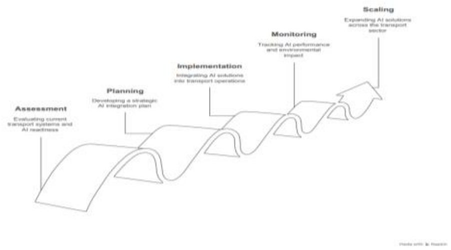


Figure 1: Framework of Deploying Generative AI in UAE's Public Transportation

IV. CONCLUSION

The research concludes that generative AI has strong potential to support the UAE's national goal of achieving net-zero emissions by 2050. Its ability to optimize routes, improve service planning, and reduce energy use makes it a valuable tool for transport authorities. While infrastructure, policy, and public awareness challenges remain, these can be addressed through strategic planning and collaborative implementation. The proposed five-phase framework offers a practical path forward, ensuring that generative AI is applied in a technically effective and socially inclusive way. Future research should focus on expanding stakeholder interviews, testing the framework with real-world pilots to evaluate its long-term environmental and operational impacts.

For future research

it is recommended that researchers and policymakers:

- Conduct larger-scale studies with more participants from across the UAE.
- Use quantitative methods to measure the real impact of AI on emissions, efficiency, and service performance.
- Explore user behavior and acceptance of AI-powered transport services to understand public needs.
- Examine policy development and regulation to support safe and ethical AI deployment in mobility systems.

DATA AVAILABILITY STATEMENT:

The primary data supporting this study's findings were collected through semi-structured interviews with experts from government, academia, and the transportation sector. The study received ethical approval from Zayed University's Institutional Review Board (IRB) to ensure compliance with ethical standards regarding human subjects' research. Due to confidentiality agreements and the sensitive nature of the data, the interview transcripts and related materials are not publicly available. However, anonymized excerpts may be shared upon reasonable request to the corresponding author, subject to approval by the participants and the institution. Additionally, secondary data were obtained from publicly available sources including government reports and policy documents.

ACKNOWLEDGMENT

I want to express my sincere gratitude to my supervisor, Prof. Ravishankar Sharma, for his continuous guidance and support throughout this research. I also thank the faculty and staff of the College of Technological Innovation at Zayed University for their valuable assistance.

My most profound appreciation goes to my family for their unwavering encouragement and patience during this journey. I am also grateful to the experts who participated in the interviews and contributed to the success of this study.

REFERENCES

[1] Masdar | UAE announces Net Zero by 2050 strategic initiative. (n.d.). https://masdar.ae/en/news/newsroom/uae-announces-net-zero-by-2050-strategic-initiative

[2] Samuelson, P. (2023). Generative AI meets copyright: Ongoing lawsuits could affect everyone who uses generative AI. Science (American Association for the Advancement of Science), 381(6654), 158–161. https://doi.org/10.1126/science.adi0656

[3] Albuquerque, F. D., Maraqa, M. A., Chowdhury, R., Mauga, T., & Alzard, M. (2020). Greenhouse gas emissions associated with road transport projects: current status, benchmarking, and assessment tools. Transportation Research Procedia, 48, 2018–2030. https://doi.org/10.1016/j.trpro.2020.08.261

AI's Role in Shaping Leadership Behavior to Fostering Teamwork Culture

Alya Almarzooqi
Graduate student from ADSM
Adsm-215620@adsm.ac.ae
Abu Dhabi, UAE

Turki Al Masaeid
Assistant Professor in Teaching and Learning, ADSM
t.almasaeid@adsm.ac.ae
Abu Dhabi, UAE

Abstract—This study examines the impact of AI on shaping leadership behavior, particularly regarding teamwork culture within a UAE organization. Leadership is fundamental to organizational success, development, culture, and performance. The increasing use of AI in organizations is influencing leadership behavior. Previous research has highlighted AI's role in shaping leadership behavior, but there are limited studies focusing on this area within the UAE context. The main objective of this study is to identify AI factors that influence leadership behavior in fostering a teamwork culture. The findings provide insights into AI's role in shaping leadership behavior to help organizations develop training programs that enhance leadership behaviors and improve teamwork culture performance in the UAE. Through the literature review and analysis, this study contributes to an understanding of the factors that impact leadership behavior.

Keywords—Artificial Intelligence (AI), Leadership behavior, ethics, decision-making, communication, trust, teamwork culture.

I. INTRODUCTION

Effective leadership is a critical factor in organizational success, influencing teamwork, culture, decision-making, and performance (Paganin, De Angelis, Pische, Violante, Guglielmi, & Pietrantoni, 2023). However, with the recent integration and increased usage of AI, it is affecting and shaping leadership behavior. Many studies have highlighted AI's role in shaping leadership behavior, emphasizing the need for awareness toward AI to ensure positive change and impact. Despite this, there are limited studies focusing on AI's role in shaping and changing leadership behavior to enhance teamwork culture specifically in the UAE's work environment. Identifying AI factors that shape leadership behavior to ensure positive outcomes can provide valuable insights for leaders and organizations.

The study aims to examine the impact of AI on changing leadership behavior, particularly in relation to teamwork, to help organizations develop training programs that enhance leaders' behavior while using AI. The objectives of the study include: 1) investigating the effective and ineffective impacts of AI on leadership behavior, 2) exploring the AI factors that influence leadership behavior, and 3) proposing strategies and effective uses of AI to enhance positive leadership behavior.

The research questions were formulated based on the objectives and employee low performance towards AI usage, indicating a potential connection to leadership issues. The study will address the following questions: 1) What are the effective and ineffective impacts of AI on leadership behavior?, 2) Which AI factors strongly influence leadership behavior?, 3) How does leadership behavior enhance trust, communication, ethics, and decision-making in relation to AI?, and 4) To what extent does AI change or influence leadership behavior to enhance teamwork culture?

II. LITERATURE REVIEW

Recently, the utilization of artificial intelligence (AI) has increased across all organizations, impacting leadership behavior. It is evident that leadership plays a critical role in shaping the culture, direction, and team performance of an organization. This literature review examines the role of AI in organizations and leadership behavior.

Different studies have defined AI, with Castelvecchi [1] describing it as a type of programmed system that is similar to human intellectual processes. Communication is highlighted as a fundamental element for excellent organizational performance, emphasizing the necessity of AI to facilitate continuous communication. Florea and Croitoru [2] noted that employees need communication skills to work effectively in teams and build relationships with leaders. AI facilitates communication, making it faster and available at all times. From my perspective, leaders must demonstrate communication skills to help employees acquire them. Leaders should determine critical skills for using AI to enhance communication and decision-making.

Păus [3] emphasized the importance of people in organizations as the main resource with communication being essential. DESI [4].found that 40% of European companies use AI to communicate with customers, improving collaboration and competitiveness. AI assists leadership in tracking employee progress, providing feedback, and coordinating tasks without delays, leading to positive results [4], Utilizing AI can improve trust and transparency based on its smooth integration and results[5]. Leadership must consider security and privacy when using AI systems to avoid negative impacts on employees [6].

Saremi et al. [7] mentioned that leadership must demonstrate trust, transparency, and fairness when using AI to enhance employee performance. Trust in AI leads to efficiency, quick solutions, and process speed [8]. A result showed that AI can 40% automate tasks and prevent time-consuming tasks [9], leading to employee satisfaction and reduced burnout. Leaders can use AI to reduce uncertainty and make decisions based on moral values [10]. However, AI lacks the ability to consider feelings and emotions in decision making[11], emphasizing the importance of human feelings and judgement in moral decisions. Shapiro and Stefkovich [12] highlighted that leadership is able to make moral decision when they demonstrate types of ethics: justice, care, critique, community, and professionalism. A study found AI's benefits include inclusivity for all employees but can be harmful when ethical considerations are overlooked [7].

Peifer et al. conducted a study on the impact of AI on leadership and found that leadership must provide a strategic plan and guidance for a long-term change process. A clear vision and objectives for AI are essential to ensure employee trust in AI. The strategic AI process requires stakeholder participation and transparency. Engaging employees while utilizing AI reduces their concerns and resistance to AI. Leadership needs to have a basic understanding of AI, data quality, manage complexity and change, and interact with AI and employees. Moreover, due to AI carrying out some tasks, this could lead to changes in what leaders must focus on. The study also highlighted that enhancing experimentation and learning occurs when leadership provides a supportive culture for AI integration and allows for mistakes. Decisions are influenced by the integration of AI into leadership, and managing the interaction between employees and AI by leadership fosters trust, effective communication, and social inclusion; and these are the behaviors that need to be maintained by leadership[14].

Vivek and Krupskyifound that leadership that integrates emotional intelligence (EI) and AI are equipped to manage complex decision making, planning, and communication. Combining EI and AI enhances effective behaviors and helps leadership gain logic skills for managing complex environments. Leadership with high EI while using AI results in having a balanced leadership behavior toward technical accuracy and human empathy, and better performance when interacting with employees and resolving conflicts. Effective use of AI by leadership leads to informed decisions and a positive work environment. The study emphasized the importance of balanced training in EI and AI to avoid ethical concerns, bias, and reduce reliance on AI [15].

Frimpong discussed the ethical issues of AI, such as lack of fairness, transparency, and loss of emotions and empathy. The study highlighted that AI positively impacts leadership by enabling data driven support, integrating human and AI insights for strategic and ethical decision-making, and assisting in task distribution and performance tracking. However, the study also pointed out the risks and challenges that could affect leadership effectiveness, including replacing humans with AI, biases in AI systems, loss of human empathy and emotions, and employee mistrust of AI [16].

Dwivedi [17] focused on integrating EI and AI to shape leadership behavior and enhance teamwork management. Using EI through AI can help leaders explore stress situations and receive feedback. AI can predict team morale, support inclusive decision making, and improve EI for leadership (Dwivedi), meaning they will be able to manage teamwork as long as they recognize and understand emotions. Leadership utilizing AI in an ethical and responsible manner can lead to continuous opportunities for process improvement and company success. EI and AI play a crucial role in improving leadership behaviors.

III. METHODOLOGY

This study employed a quantitative research design to examine the impact of AI on changing leadership behavior. This approach is appropriate as it allows for the collection of a high response rate and numerical data on leadership behavior changes due to AI. The study was conducted in a government organization in the UAE, using a purposive sampling technique to select participants. The sample consisted of 30 employees due to the small size of the organization, ensuring a focused investigation into the impact of AI on leadership behavior towards teamwork.A questionnaire tool was used to collect data, addressing the research questions; it was distributed to the employees via email. The study adhered to ethical research principles to ensure participants' rights and confidentiality. Ethical considerations included obtaining consent forms from all participants before they completed the questionnaire. Participants were allowed to withdraw from the study at any stage without facing consequences. Anonymity and confidentiality of responses were maintained to protect participants' privacy and make them feel comfortable while completing the questionnaire. The data supporting this study's findings were collected from employees and managers within an oil and gas organization implementing AI in leadership systems. The study received approval from the university's internal research committee to ensure compliance with ethical standards and participant confidentiality. Due to organizational privacy policies and confidentiality agreements, the data are not publicly available.

A. Conceptual Framework

Many studies have highlighted the role of AI in leadership behavior and emphasized the importance of trust, communication, ethics, and decision making. The study developed a conceptual framework demonstrating how AI can change leadership behavior to enhance employee efficiency through these variables. In this framework, AI is the independent variable, while leadership behavior is the dependent variable, with trust, communication, ethics, and decision-making as mediator variables. Leadership and teams trust using AI more than humans based on its accuracy and reliability [4]. Communication and collaboration are essential for leadership and teams to facilitate positive changes [4].Ethical considerations must be addressed to enhance AI efficiency [17].. Leadership should not be replaced with AI in making decisions as moral considerations cannot be included in data [11]. The framework also includes organizational culture as a moderator and team size as a control variable.

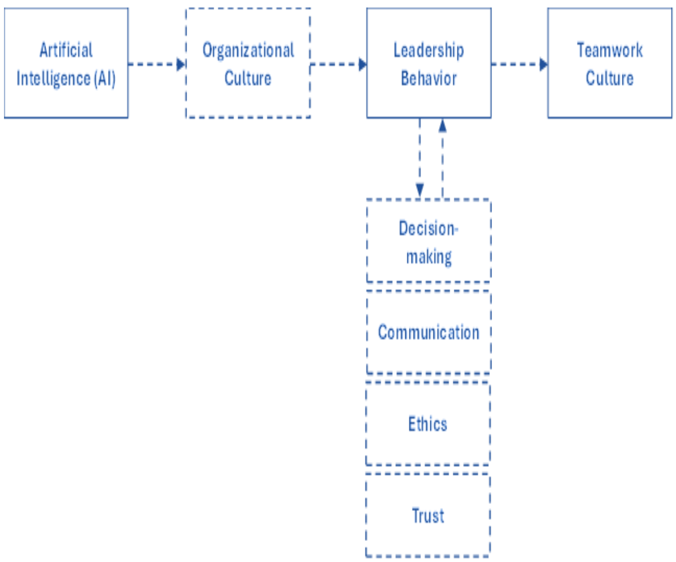


Fig. 3. Conceptual Framework

IV. DATA ANALYSIS

Thematic and descriptive analysis approaches were used to analyze the collected data and draw a purposeful conclusion about the research questions. The questionnaire included demographic and 20 closed-ended questions. The participants' responses gathered from the questionnaire revealed a comprehensive exploration of the impact of AI in shaping leadership behavior, and the results highlighted five key themes including: 1) perceptions of leadership and AI, 2) trust and ethics in AI, 3) AI factors influencing leadership behavior, 4) AI and teamwork culture, and 5) training, performance, and job satisfaction. The sections outline the process followed to analyze the data and discuss the results supported by relevant data and literature.

The questionnaire results accurately addressed the research questions. 30 employees responded, resulting in a 100% response rate. The demographic questions included gender, age, and experience. Table 1 showed that the majority of participants (63%) were 30 years old and above, and over half (57%) had 10 years or more of experience.

Table 1 Demographic Questions

Variables	Level	(N = 30)	
		Frequency (n)	Percent (%)
Gender	Female	17	57%
	Male	13	43%
Age	18 – 24	5	17%
	25 – 29	6	20%
	30 and above	19	63%
Experience	2	3	10%
	4	2	7%
	6	8	27%
	10 and above	17	57%

B. Perceptions of Leadership and AI

In the first question, 36.7% of participants expressed disbelief in the excellence of leadership behavior in managing teamwork. 33% both agreed and disagreed on the supportiveness of leadership behavior towards AI usage . Furthermore, 30% of participants disagreed on the positive impact of AI on leadership behavior, while 23% agreed . These results suggest potential negative impacts on employees due to leadership behavior, emphasizing the importance for leaders to acquire more skills and knowledge about AI to effectively support its usage, given its increasing significance in organizational operations.

On the other hand, 33% of participants agreed and 16.7% strongly agreed that AI is effective in supporting leadership decision-making , aligning with Wang's study [11] , where leadership can reduce uncertainty and make decisions using AI. Additionally, 13.3% strongly agreed and 30% agreed that AI positively influences work culture, while 30% disagreed . The results from theme 1 indicate varying perceptions regarding the impact of AI, with some perceiving it positively and others negatively, highlighting concerns about leadership understanding and beliefs about AI.

C. Trust & Ethics in AI

In the sixth question , 36.7% disagreed that leadership behavior enhances team members' trust in AI, while 33.3% agreed. 26.7% agreed that when leadership doesn't believe in AI, employees' trust in AI is reduced, while 20% disagreed. These results demonstrate that due to the significant role and influence of leadership, employees are affected by their beliefs; people tend to follow leaders who believe in their beliefs. Moreover, 73.3% strongly agreed and 23.3% agreed that misusing AI and obtaining inaccurate or unreliable results reduces employee trust (Figure 8 – Appendix B), highlighting the importance of ethics. This result aligns with Wang's [11]study, which emphasized the importance of security and privacy when using AI to avoid negative impacts on employees.

Additionally, 36.7% agreed that when the usage of AI is limited, employee trust in AI will be reduced ; supporting Gerlich's [5] assertion that utilizing AI improves trust and results. However, 56.7% disagreed about promoting ethical guidelines and responsible usage by leadership, underscoring the necessity of ethics in AI usage as stated in Shapiro and Stefkovich's study[10]., and explaining the reason for reduced trust in AI by employees. Furthermore, 43.3% stated that leadership behavior neutrally fosters open communication and team trust regarding the use of AI tools , highlighting the importance of communication skills for efficient teamwork and building relationships with leaders, as noted in Florea and Croitoru's study[4].

D. AI Factors Influencing Leadership Behavior

The twelfth question showed that 33.3% strongly agreed that data-driven decision-making, communication tools, and performance tracking are factors that influence leadership behavior (Figure 12 – Appendix B). This aligns with Florea and Croitoru's [4] assertion that AI assists leadership in achieving positive results by tracking employee progrss, providing feedback, and coordinating tasks without delays. However, 33.3% disagreed that AI is improving the effectiveness of leadership behavior . Additionally, 20% agreed and 53% found it neutral that automation AI tools influence how leaders delegate and monitor tasks (Figure 14 – Appendix B). This demonstrates that some AI tools help in preventing time-consuming tasks, which can reduce leadership and employee burnout. Herold [6] mentioned that 40% of AI tools can automate tasks and save time.

E. AI & Teamwork Culture

The fifteenth question revealed that 33.3% both agreed and disagreed on the use of AI positively influencing leadership behavior in promoting a stronger teamwork culture . 46.7% agreed and 43.3% disagreed that AI helped leaders create a collaborative and inclusive team environment , which aligns with Păus's [8]. emphasis on the importance of people as they are the main resource in the organization, and AI benefits leadership in inclusivity for all employees as Zidouemba [13] asserted. The disagreement in this result indicates that not all employees are involved in the team environment, leading to a negative experience. However, 16.7% strongly agreed and 26.7% agreed about the strong influence of AI on work culture

F. Training, Performance & Job Satisfaction

43.3% of participants agreed on their awareness of the importance of leadership training programs related to AI, while 30% disagreed. 36.7% both agreed and disagreed that leadership behavior in using AI improves their performance, demonstrating the impact of leadership behavior on team performance. Good behaviors by leadership such as trust, transparency, and fairness will enhance performance, as mentioned by Saremi et al. [9]. Additionally, 40% agreed and disagreed that leadership behavior influences their overall job satisfaction in an AI-integrated environment , revealing that leadership must be aware of their behavior and the need for training programs related to AI to enhance team performance and ensure job satisfaction for employees

V. LIMITATION AND RECOMMENDATION

The limitation of this study is the small sample size, which is due to the organization's small size. This may restrict the generalizability of the results to a broader population. Since the research was conducted in one organization, the findings may not fully represent AI's role in shaping leadership behavior across other organizations in the UAE. Additionally, the data were collected using only one tool, the questionnaire, which may not provide in-depth insights.

To address these limitations, future research should consider a larger sample size by including multiple organizations to enhance the generalizability and accuracy of the findings. It is also recommended to incorporate additional data collection methods, such as interviews, to gather more in-depth insights, triangulate data, and improve the accuracy, reliability, and validity of the findings.

VI. CONCLUSION

This study utilized a questionnaire for data collection to investigate the influence of AI on shaping leadership behavior to promote an effective teamwork culture. The analysis highlighted the significance of leadership demonstrating fairness in decision-making, effective communication, ethical considerations, and trust utilizing AI. It is necessary for leaders to be aware of their behavior as it highly affects employees, and they should undergo training programs on AI-related topics to build trust in its use and ensure adherence to ethical standards. However, it is essential to acknowledge that the study's scale is limited and may not be applicable to broader organizations in the UAE.

REFERENCES

[1] Castelvecchi, D. (2016),"Can we open the black box of AI?", Nature, Vol. 538 No. 7623, pp. 20-23.

[2] Damasio, A.R. (2019), The Strange Order of Things: Life, Feeling, and the Making of Cultures, Vintage, New York, NY.

[3] DESI. (2023). Raportul din 2023 privind stadiul evolut, iei deceniului digital. Available online: https://digital-strategy.ec.europa.eu/ro/library/2023-report-state-digital-decade

[4] Florea, N. V., & Croitoru, G. (2025). The Impact of Artificial Intelligence on Communication Dynamics and Performance in Organizational Leadership. Administrative Sciences, 15(2), 33. https://doi.org/10.3390/admsci15020033

[5] Gerlich, M. (2024). Exploring motivators for trust in the dichotomy of human-AI trust dynamics. Social Sciences, 13(5), 251.

[6] Herold, B. (2019),"Forty percent of elementary school teachers' work could be automated by 2030, McKinsey Global Institute Predicts", Education Week, 4 June, available at: https://blogs.edweek.org/edweek/DigitalEducation/2019/06/women_future_of_work_mckinsey.html

[7] Paganin, G., De Angelis, M., Pische, E., Violante, F. S., Guglielmi, D., & Pietrantoni, L. (2023). The impact of mental health leadership on teamwork in healthcare organizations: A serial mediation study. Sustainability, 15(9), 7337. doi: https://doi.org/10.3390/su15097337

[8] Păus, , V. A. (2006). Comunicare si resurse umane. Editura Polirom.

[9] Saremi, M. L., Ziv, I., Asan, O., & Bayrak, A. E. (2025). Trust, workload, and performance in human-artificial intelligence partnering: The role of artificial intelligence attributes in solving classification problems. Journal of Mechanical Design, 147(1), 4065916.

[10] Shapiro, J.P. and Stefkovich, J.A. (2016), Ethical Leadership and Decision-making in Education: Applying Theoretical Perspectives to Complex Dilemmas, Routledge, New York, NY.

[11] Wang, Y. (2021). Artificial intelligence in educational leadership: a symbiotic role of human-artificial intelligence decision-making. [Artificial intelligence in educational leadership] Journal of Educational Administration, 59(3), 256-270. https://doi.org/10.1108/JEA-10-2020-0216

[12] Zeng, J., Chen, K., Wang, R., Li, Y., Fan, M., Wu, K., Qi, X., & Wang, L. (2025). ContractMind: Trust-calibration interaction design for AI contract review tools. International Journal of Human-Computer Studies, 196, 103411.

[13] Zidouemba, M. T. (2025). Governance and artificial intelligence: the use of artificial intelligence in democracy and its impacts on the rights to participation. Discover Artificial Intelligence, 5(1), 1-11.
[14] Peifer, Y., Jeske, T., & Hille, S. (2022). Artificial intelligence and its impact on leaders and leadership. Procedia computer science, 200, 1024-1030.
[15] Vivek, R., & Krupskyi, O. P. (2024). EI & AI in leadership and how it can affect future leaders.
[16] Frimpong, V. (2025). The Impact of AI on Evolving Leadership Theories and Practices. Journal of Management, 3, 188-193.
[17] Dwivedi, D. (2025). Emotional Intelligence and Artificial Intelligence Integration Strategies for Leadership Excellence. Advances in Research, 26(1), 84-94.

AI Integrated Approach to Achieve Transformational Strategic Leadership and Its Reflections on Employee Engagement

Basma Ahmed Ali
Graduate student in Abu Dhabi School of Management
adsm-215577@adsm.ac.ae
Abu Dhabi, UAE

Turki Al Masaeid
Assistant Professor in Teaching and Learning, ADSM
t.almasaeid@adsm.ac.ae
Abu Dhabi, UAE

Abstract—Integrating Artificial Intelligence (AI) into leadership practices redefines how transformational strategic leadership is enacted across organizations. As industries shift toward digital and data-driven operations, leaders are expected to harness AI to improve communication, drive performance, and increase employee engagement. This paper examines how AI technologies influence strategic leadership, particularly within transformational and emotionally intelligent leadership styles. Drawing from recent literature and empirical insights, this research explores how AI contributes to a more engaged workforce and proposes actionable pathways for aligning digital transformation with human-centric leadership models.

Keywords—Artificial Intelligence, Transformational Strategic Leadership , Employee Engagement, and Leadership Reflections.

I. INTRODUCTION

Leaders possess qualities that distinguish them from managers, but they also perform similar "managerial" tasks, including setting goals and developing strategic plans to achieve them, communicating direction to organizational members, monitoring performance, and motivating employees.

All leaders, regardless of their position, engage in the core roles and activities identified by the four-factor theory of leadership proposed by Bowers and Seashore: support through leadership behaviors that enhance subordinates' sense of personal value and importance; facilitating interaction through behaviors that encourage organizational members to build strong, mutually satisfying relationships; focusing on the goal through behaviors that motivate organizational members to achieve outstanding performance and accomplish the organization's stated goals; and facilitating work through behaviors that support the achievement of organizational performance goals, such as coordination, planning, and scheduling, and providing subordinates with the tools, materials, and technical knowledge necessary to perform their tasks.

Furthermore, leadership roles and the focus of a leader's activities vary depending on their position within the organizational structure, as well as other factors such as the type of activity the organization engages in, the surrounding environmental conditions, the organization's stage of development, the leader's role in establishing it, and the scope of its global business activities. All of these factors influence the leader's role and the behaviors required to be an effective leader. Today, AI- enhanced leadership in the era of rapidly evolving digital technology has become crucial for organizations and institutions across the world. Leadership as a whole concept involves the ability to use technology, data, and innovation to reach organizational goals and enhance organizational performance. However, the challenges as well as opportunities for leadership at all levels in the digital age are many.Rapid technological change is one of the biggest challenges for leadership, especially on the strategic level. Today, we live in a rapidly changing world, and leaders need to be on top of the latest technologies, therefore, digital leadership still has challenges in the digital age, but at the same time, all the opportunities. However, in order to realize their sustainable success, digital leaders should have the ability to adapt to technologically, manage huge data, and create a digital corporate culture in which digital transformation is supported.

Together, these are a series of significant transformations of an AI-enhanced leadership in the organization. To accomplish this, organization and leaders has to invest in continuous learning and technical skills.

However, leadership also offers significant opportunities in the digital age. When technology and innovation are embraced correctly, organizations can achieve significant improvements in performance and efficiency. Technology also contributes to improving business processes, reducing costs, and increasing productivity.

AI introduces dual benefits and drawbacks for leadership roles. Technological advancements are impacting Leadership innovation. With the availability of artificial intelligence, leaders can use advanced tools and software to analyze data, predict future trends, and make strategic decisions. Consequently, leaders can be more effective in making decisions and driving innovation within their organizations. Moreover, AI can help enhance collaboration and cooperation between leaders and subordinates. With the availability of technological tools such as collaborative software and virtual platforms, remote teams can work together effectively and easily share information and ideas. Thus, modern leaders can encourage collaboration and interaction among team members, achieving innovation and excellence at work.

However, it is important to note that technology and innovation are not just tools used by modern leaders; they are also a culture and a methodology. When leaders embrace technology and innovation, they are expressing their vision, values, and goals. Therefore, technology, innovation, and AI can help enhance organizational culture and achieve sustainable success within organizations.

The AI is significantly impacting modern leadership styles. As technology advances, the ways leaders and subordinates communicate and interact are changing, and methods of innovation and collaboration at work are evolving. There is a need to understand how AI-integrated leadership produces work environments that boost both work interest and employee self-esteem, alongside promoting participation toward organizational goals. Leaders must determine methods to combine AI investments for better organizational results with human-led approaches for building trust, employee motivation, and performance enhancement. The research gap exists because scientists need to understand how AI-integrated leadership uses employee engagement strategies to maximize productivity. Based on all the above, this study aims to answer the following key questions :

1. How does AI integration influence leaders' ability to engage and inspire their team members ?
2. In which ways can AI tools support emotionally intelligent leadership behaviors?
3. What are the ethical and operational challenges leaders face when using AI to enhance employee engagement?
4. How can organizations balance automation and empathy to maintain a human-centric leadership approach?

II. LITERATURE REVIEW

The intersection of AI and leadership has become a focal point in recent scholarship. Rožman, Tominc, and Milfelner [1] argue that AI-fueled organizational cultures significantly enhance employee engagement by improving training and communication. Florea and Croitoru [2] support this claim by emphasizing how AI reshapes organizational communication dynamics, strengthening performance and employee connectivity.

Abositta, Muri, and Berberoğlu [3] highlight the mediating role of transformational leadership in leveraging AI for improved decision-making within engineering management. The synergy between AI analytics and human leadership judgment allows for more personalized and strategic approaches to employee development.

Vicci [4]explores emotional intelligence in AI systems, indicating that human-AI interaction can simulate empathetic responses, thereby enhancing leadership communication. These technologies can assess emotional states via sentiment analysis and provide leaders with actionable feedback to respond compassionately.

Kim, Kim, and Lee [5] present a view of AI-induced job insecurity and how ethical leadership can mitigate negative perceptions. Their study links AI adaptation with environmental responsibility, suggesting that leaders play a crucial role in guiding employees through digital transitions while fostering trust. Rožman et al. [6] examine the leadership methods which support AI deployment for enhancing employee engagement by focusing on training and culture transformation programs.

Sacavém et al. [7] further discuss the role of leadership in organizational digital transformation, stressing that leadership adaptability is key to unlocking AI's full potential. Fenwick, Molnar, and Frangos [8] echo this by emphasizing the paradigm shift required in HRM to move from AI implementation to human-centric adoption.

according to Lakshmikanth et al. [9] to reach peak employee engagement and performance results. The research shows that AI possesses great power to deliver customized interventions that meet both workplace objectives and staff requirements. According to Sarioguz & Miser [10] AI-driven transformation impacts participatory leadership by utilizing technology to empower new employee participation strategies in business management. Boudreaux investigates how transformational leadership needs to adjust with AI-oriented Industry 4.0 by concentrating on developing leadership qualities such as adaptability, ethics and resilience to boost employee commitment [1]. Thus, the research gap can be highlighted in these hypotheses:

H1: AI integration by leaders positively affects transformational strategic leadership.

H2: Transformational strategic leadership produces positive effects leading to enhanced employee engagement. **H3**: AI integration in leadership positively effect on employee engagement.

H4: Transformational strategic leadership mediates the relationship between AI integration in leadership and employee engagement.

III. METHODOLOGY

A quantitative research method was implemented to investigate how AI-integrated transformational strategic leadership affects employee engagement in the work environment. The data collection process targets both staff members and Managers in one organization implementing artificial intelligence across their leadership systems in the Oil and Gas industry. The tool for data collection through this research was designed with 15 questions to answer the research questions. The measurement of transformational leadership, employee engagement, and AI implementation scope within leadership practices. The study utilizes stratified random sampling to generate a proper representation by selecting subjects from different sectors and ranging in organizational size and occupational roles. A total of 150 individuals participated in the survey, including 20 who held leadership and managerial positions. Participants came from diverse industries and represented a range of organizational sizes.

The data supporting this study's findings were collected from employees and managers within an oil and gas organization implementing AI in leadership systems. The study received approval from the university's internal research committee to ensure compliance with ethical standards and participant confidentiality. Due to organizational privacy policies and confidentiality agreements, the data are not publicly available.

The research data were organized through descriptive statistics to report the participant demographics along with the study variable, which can be stated in the figure below:

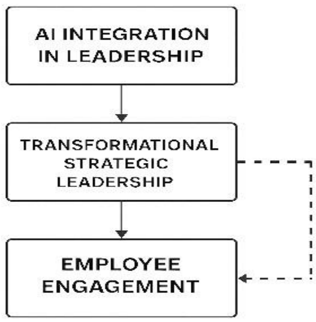


Figure 1: Conceptual Framework

IV. RESULTS AND DISCUSSION

The study results showed that:

Table 1: Descriptive Statistics

Descriptive analysis for the survey items(n=150)		
Section	Mean	SD
AI & Leadership Engagement (Items 1-4)	4.15	0.65
AI & Emotional Intelligence (Items 5-8)	3.73	0.75
Ethical/Operational Challenges (Items 9-12)	3.88	0.73
Balancing Automation & Empathy (Items 13-15)	3.87	0.80

Leaders view AI as beneficial for emotional intelligence development according to the table but they indicate this aspect remains under development according to feedback responses. Major ethical and privacy concerns exist at present. Many leadership figures acknowledge their insufficient knowledge about AI ethical education. Leaders support the idea of automation enhancing empathy than eliminating it and they foresee how their organisation's strategy can advance further. AI functions as a widely accepted tool which improves leadership effectiveness together with employee engagement. The potential of AI to assist emotionally intelligent leadership remains promising yet it has not reached its full potential. Public concerns about ethical obstacles and privacy along with transparency issues require immediate improvements in ethical training systems. Leadership needs to remain centred around humans since automation shows continued growth.

V. DISCUSSION

The AI-integrated approach to transformational leadership offers numerous opportunities for enhancing employee engagement, but it also introduces complex challenges. At the heart of this transformation lies the need to maintain a human touch in leadership while benefiting from the efficiency and scalability of AI technologies.

Transformational leaders using AI must be adaptable, emotionally intelligent, and ethically grounded. They are expected to motivate, inspire, and leverage data transparently and human-centric. Strategic leadership in this context involves fostering innovation, aligning AI tools with organizational goals, and continuously nurturing trust within the team.

This study finds that the most effective AI-integrated leadership strategies balance data-driven insights with the emotional realities of the workforce. Leaders must use AI to support—not replace—their relational responsibilities.

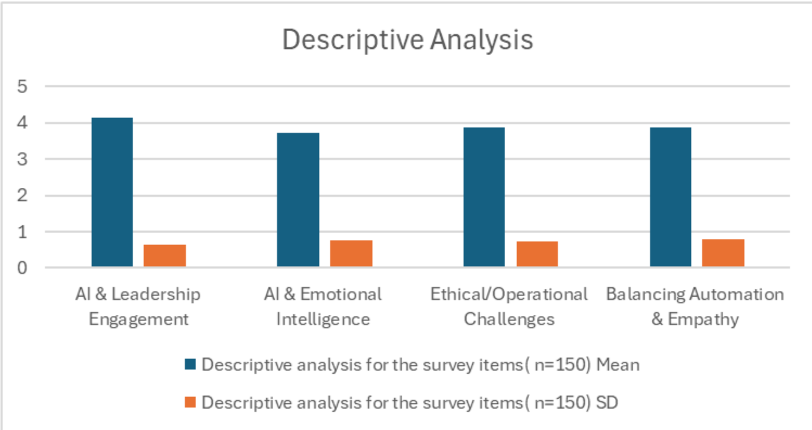


Figure 2: Result Representation

VI. RECOMMENDATIONS

1. **Integrate AI into Leadership Development Programs.** Incorporate AI literacy into training programs for managers and team leaders to ensure they understand the ethical and practical dimensions of AIenabled decision-making.
2. **Develop Emotionally Intelligent AI Interfaces.** Collaborate with HR and tech developers to implement AI systems capable of detecting emotional cues and providing supportive feedback loops without breaching ethical boundaries.

3. **Establish Ethical Frameworks for AI Use: To build trust and encourage acceptance,** create transparent policies for how employee data is collected, analyzed, and applied in AI systems.
4. **Foster an Adaptive and Inclusive Culture.**
   Encourage open communication about AI tools, collect employee input on new implementations, and emphasize the augmentation, not replacement, of human roles.
5. **Use AI for Continuous Feedback and Recognition.** Leverage AI tools for real-time feedback, recognition of achievements, and targeted support interventions that enhance motivation and morale.

VII. CONCLUSION

AI has the potential to transform strategic leadership into a more adaptive, responsive, and emotionally attuned practice. When used ethically and thoughtfully, AI tools can support leaders in creating more engaged, motivated, and productive workforces, transparently and ethically.

However, technology must not eclipse the need for human connection, empathy, and ethical judgment. The future of transformational leadership lies in a synergistic model where AI augments rather than replaces human capabilities, enabling leaders to elevate performance and people. Organizations that navigate this integration effectively will set new leadership and employee engagement standards in the digital era.

REFERENCES

[1] M. Rožman, P. Tominc, and B. Milfelner, “Maximizing employee engagement through artificial intelligence organizational culture in the context of leadership and training of employees: Testing linear and non-linear relationships,” Cogent Business & Management, vol. 10, no. 2, 2023. [Online]. Available: https://doi.org/10.1080/23311975.2023.2248732

[2] N. V. Florea and G. Croitoru, “The impact of artificial intelligence on communication dynamics and performance in organizational leadership,” Administrative Sciences, vol. 15, no. 2, p. 33, 2025. [Online]. Available: https://doi.org/10.3390/admsci15020033

[3] A. Abositta, W. A. Muri, and A. Berberoğlu, “Influence of artificial intelligence on engineering management decisionmaking with mediating role of transformational leadership,” Systems, vol. 12, no. 12, p. 570, 2024. [Online]. Available: https://doi.org/10.3390/systems12120570

[4] H. Vicci, “Emotional intelligence in AI systems: An exploratory review,” IUP Journal of Soft Skills, vol. 18, no. 2, pp. 5–20, 2024. [Online]. Available: https://www.proquest.com/scholarly-journals/emotionalintelligence-ai-systems-exploratory/docview/3098308693/se-2

[5] B. Kim, M. Kim, and J. Lee, “Code green: Ethical leadership’s role in reconciling AI-induced job insecurity with pro-environmental behavior in the digital workplace,” Humanities & Social Sciences Communications, vol. 11, no. 1, p. 1627, 2024. [Online]. Available: https://doi.org/10.1057/s41599-024-04139-2

[6] A. Sacavém, B. M. Andreia de, d. S. João Rodrigues, A. Palma-Moreira, H. Belchior-Rocha, and M. Au-Yong-Oliveira, “Leading in the digital age: The role of leadership in organizational digital transformation,” Administrative Sciences, vol. 15, no. 2, p. 43, 2025. [Online]. Available: https://doi.org/10.3390/admsci15020043

[7] A. Fenwick, G. Molnar, and P. Frangos, “The critical role of HRM in AI-driven digital transformation: A paradigm shift to enable firms to move from AI implementation to human-centric adoption,” Discover Artificial Intelligence, vol. 4, no. 1, p. 34, 2024. [Online]. Available: https://doi.org/10.1007/s44163-024-00125-4

[8] N. Lakshmikanth, A. S. Raja, A. S. Babu, and T. P. Prasad, “Employee engagement in artificial intelligence-enabled workplaces: Role of digital leadership and organizational culture,” Materials Today: Proceedings, vol. 82, pp. 435-440, 2024.

[9] E. Sarioguz and S. Miser, “Transformational leadership and participatory management in AI-driven organizations,” Journal of Business Management, vol. 40, no. 1, pp. 20-32, 2024.

[10] J. Boudreaux, “Adapting transformational leadership for Industry 4.0: AI’s impact on leadership qualities,” Leadership Quarterly, vol. 35, no. 2, 2024.

[11]] T. M. Schneller, “Enhancing partnership quality: Unlocking AI's impact on B2B partnerships,” Ph.D. dissertation, University of Maryland, University College, 2025. [Online]. Available: https://www.proquest.com/dissertations-theses/enhancingpartnership-quality-unlocking-ais/docview/3186230760/se-2

[12] A. Shah, N. Nasir, and A. Shah, “Inclusive design in AI-driven leadership: Implementation and challenges in small businesses,” International Journal of Business Research and Management, vol. 15, no. 1, pp. 19–42, 2024. [Online]. Available: https://www.cscjournals.org/library/manuscriptinfo.php?mc=IJBRM-365

[13] R. Paralta, E. Simões, and P. Duarte, “Subjective well-being in organizations: Effects of internal ethical context and ethical leadership,” International Journal of Environmental Research and Public Health, 2023. [Online]. Available: https://doi.org/10.3390/ijerph20054451

Alzheimer's Disease Diagnosis Using Ensemble Classification Approach

Issa Qabajeh
Faculty of Information Technology
Philadelphia University
Amman, Jordan
iqabaja@philadelphia.edu.jo
issa.qabajeh@gmail.com

Abstract—Alzheimer's disease (AD) diagnosis in early stages is an important step that helps neurologists treat the patient in a timely manner and can lead to excellent outcomes in patient response to treatment plans. An ensemble-based machine learning framework presented, which is based on magnetic resonance imaging (MRI) attributes, can enhance AD early detection and diagnosis accuracy. Several classifiers are studied, including random forest, gradient boosting, and voting classifiers. The article will develop a framework of an ensemble model of these classifiers that can lead to enhanced accuracy of Alzheimer's disease diagnosis. Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset is used in this research, and the MRI features are preprocessed and then implemented using the proposed model. The Research results show a better performance of the ensemble models in terms of accuracy, precision, and recall, compared to individual classifiers. The effective integration of different classifiers to capture MRI patterns and the establishment of a strong diagnostic model with the possibility of being embedded within clinical decision support systems are the key contributions of this research. These softwares could be used by different stakeholders, like neurologists, patients. This may revolutionize traditional diagnostic and treatment methods by providing accurate and effective early diagnosis and treatment before the onset of Alzheimer's disease.

Keywords—Alzheimer's Disease Diagnosis, Machine learning, Ensemble learning, Magnetic resonance imaging, Classification Algorithms, Decision Support Systems

I. Introduction

Alzheimer's disease (AD) ranks among the utmost devastat- ing neurodegenerative disorders, disturbing millions globally and imposing major socioeconomic and emotional problems. Specified by continuous memory damage, cognitive decline, and behavioral changes, AD not only weakens individual inde- pendence, but also stresses caregivers and healthcare systems. The global occurrence of Alzheimer's continues to rise, driven by old populations, increasing the urgency to address the diagnosis and treatment challenges [1].

Early diagnosis of Alzheimer's disease is important for improving AD patient life. Identifying the disease in its early stages can help with intervention plans to slow progression and improve quality of life. This early identification also enables patients and families to plan health care treatments, accessing support services. Despite these benefits, current diagnostic methods often diagnose AD only after significant neurological damage has occurred, limiting therapeutic success [2].

Traditional AD diagnosis depends mainly on clinical, neuropsychological tests, and standard imaging methods. Magnetic resonance imaging (MRI) provides detailed body information, yet human interpretation may miss the early diagnosis of AD. Additionally, current diagnostic machines do not have integration with computational approaches that can attach complex MRI data, showing patterns unseen to the human eye [3].

This research aims to find a solution for these gaps by implementing a complete ensemble machine learning (ML) framework for the diagnosis of Alzheimer's disease in its early stages, using MRI features. By joining multiple classifier algorithms, random forest, gradient boosting, and voting classifiers, into a unified ensemble model, the framework pursues to enhance ac- curacy, precision, and recall compared to a single algorithm. We used the Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset for building the proposed model. The final research goal is to develop a robust data-driven model adaptable to clinical decision support systems to improve early AD diagnosis, and eventually helping stakeholders, such as neurologists, patients, and the larger healthcare domain.

This paper is organized as follows: Section 2 outlines the background and reviews work already conducted on the early diagnosis of AD. Section 3 discusses the proposed ensemble model design including the data sets and pre-processing. In Section 4 we present the conclusion and future works.

II. Literature Review

A. Background

Alzheimer's disease, which is characterized by a progressive decline in memory and cognitive function, is the most widespread form of dementia and is one of the major causes of mortality among the elderly [4], [5]. Even with the distinguished research efforts to develop disease-modifying treatments, AD remains without a definitive cure [6]. Therefore, detection of the early onset of AD requires sensitive approaches that can detect brain changes before reaching a point where the damage is permanent.

Magnetic resonance imaging (MRI) represents a vital tool in AD diagnosis, due to its high-resolution visualization of the anatomical structure of the brain. The challenge lies in the classic handling and interpretation of these images, which needs improvement in terms of sensitivity and reliability [3]. Machine learning (ML) has significantly evolved the inter- pretation of MRI images, due to its ability to detect features and manifestations that may not be observed by clinicians [7], [2]. Utilizing MRI data, ML techniques were used for the detection, classification, and prediction of AD. Ensemble learning models, which are the result of combining different classifiers and leverage the strengths of separate models to deliver better performance and superior accuracy.

B. Recent Research Works

Several studies explored the use of ensemble models for AD detection and classification; one study introduced a deep learning-based ensemble model for the early detection of AD using MRI images from ADNI dataset. This model was composed of six convolutional neural networks (CNN) and showed promising results [2].

Another study presented an ensemble deep learning model for the detection of AD, the paper presented two low- parameter (CNNs): IR-BRAINNET and Modified-DEMNET [8].

Alruily et al. developed a unique deep learning ensem- ble model that fuses features form VGG16, MobileNet, and InceptionRestNetV2, which aimed to increase AD diagnosis accuracy [9].

Naguyen et al. presented an ensemble learning model com- posed of both deep learning and machine learning, which achieved an approximate scoring time of 10 minutes, which is much faster than feature extraction-based methods [10].

A study in 2022 presented an ensemble model for AD diagnosis and made a comparison between their model and other machine learning models. In terms of accuracy, the presented model showed better results than previous models [7].

Shaffi et al. constructed a ML-based ensemble model that uses ADNI and OASIS datasets for the classification of AD. After putting ML models against DL models and comparing them in terms of performance, the study concluded that ML models surpassed DL models in terms of performance [3], While a study in 2021 utilized Diffusion Tensor Imaging (DTI) measures to construct an ensemble learning model that helps differentiate patients with AD [11].

A study in 2024 also presented an ensemble deep learn- ing model, but based on quantum machine learning for AD classification. The model had 99.89% accuracy and is said to have superior accuracy and training time when compared to Support Vector Machine (SVM) [12].

III. Proposed Model and Methodological Framework

A. Dataset

The Alzheimer’s Disease Neuroimaging Initiative (ADNI) dataset is very important clinical dataset, that can help in proceeding Alzheimer’s research. It includes groups of MRI scans, demographic, and clinical tests that can be used for machine learning (ML) to implement early AD detection models. This study uses MRI features, a subset of ADNI dataset, centering on the most impactful features that can help diagnose AD in its early stages.

B. B. Feature Extraction

Extracting meaningful features from MRI data is critical for effective classification. We utilize the following MRI features: Ventricles, Hippocampus, WholeBrain, Entorhinal, Fusiform, MidTemp, ICV. A preprocessing stage to be applied to MRI features, applying imputation using the Multiple Imputation by Chained Equations (MICE) technique, and features selection using Principal Component Analysis (PCA) then serve as inputs for the ensemble machine learning classifier.

C. Ensemble Model Development

Three machine learning algorithms support the classification framework: random forest, gradient boosting, and voting clas- sifiers. Random forest works through collection of decision trees, each one is trained on using bootstrap, that can provide solid results with respect to working with categorical feature communications and overfitting [13]. Gradient boosting revise weak classifiers by decreasing classification errors to the minimum, reaching detailed predictive accuracy, mainly with heterogeneous features [14]. The voting classifier constructs classifications from different independent models, depending on either weighted votes or majority, which can balance the variances and biases each single model to produce an agree- ment diagnosis [15]. We have chosen these classifiers because of their excellent strengths reached in past AD classification researches.

The ensemble framework combines the three classifiers into an integrated model that increases their collective classification power. Implementation consists of training each classifier alone on the MRI features from ADNI datasets, then combin- ing their classification outputs using a strict voting plan where the majority voting dominates the final classification class. This process framework schema is drawn showing parallel classifiers converging into an ensemble classification node as shown in Fig. 1. Model development experiments were implemented in SPSS and Python programming languages using Intel Core i7 CPU with 64 GB RAM. Experimental setup needs a stratified 10-fold cross-validation to keep class dis- tribution and alleviate overfitting. This integrative framework enhances classification accuracy by alleviating each classifier weaknesses and utilizing various algorithms strengths as a group [2], [3].

The ensemble framework methodology offers real benefits with respect to individual classifiers in AD diagnosis. Mainly, it reaches better classification accuracy, precision and recall by combining various predictive views, reducing tendency to single classifier errors and overfitting problems. Empirical studies affirm that ensemble classifiers give better robustness with the heterogeneous of clinical MRI scans data. Further- more, the data-driven model can be integrated into decision support systems or software systems,

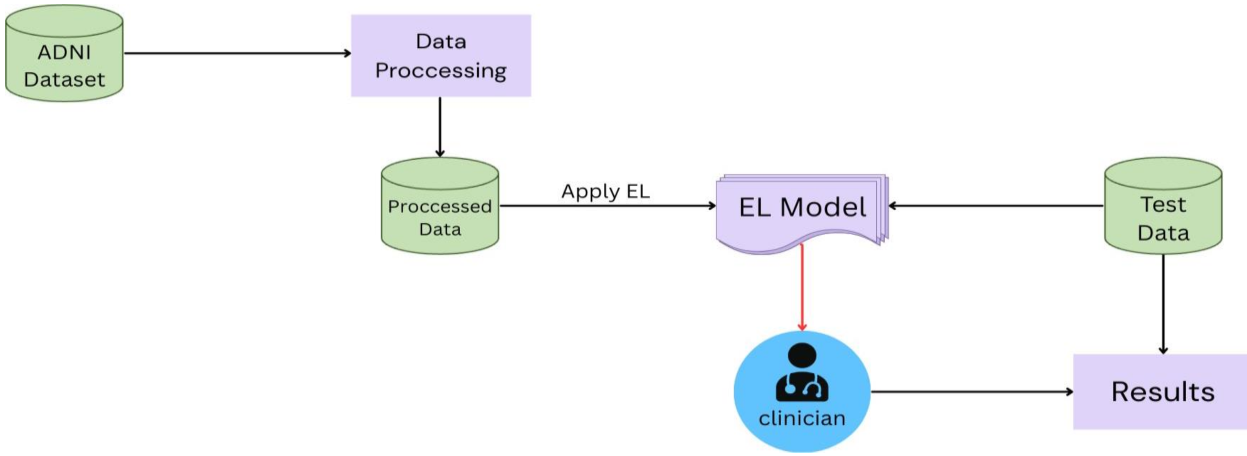


Fig. 1. Ensemble learning classification model

with robust diagnostic tools that increase clinical decisions. The integration with real-time application in clinical diagnosis workflows and personalized treatment marks a revolution in computational neurodiagnostics diagnosis [16], [2]. The demonstrated efficiency and interactions of ensemble classification models confirm the potential of transforming traditional diagnostic models in Alzheimer's disease management.

IV. Conclusion and Future Works

The ensemble model framework implemented in this research joins random forest, gradient boosting, and voting classifiers to exploit the strength of magnetic resonance imaging (MRI) features for Alzheimer's disease diagnosis. The main characteristics driving its performance include the ability to manage complex, non-linear patterns and to reduce bias and variance by consolidating multiple model outputs. This integration produces a more accurate and robust classification model that surpasses individual classifiers in the diagnosis of Alzheimer's disease.

For neurologists, the developed model provides an effective increase in diagnosis of Alzheimer's Disease. Patients can benefit from early diagnosis, which can slow the progression and improve quality of life. Healthcare schemes may take advantage of reducing continuing burdens by enabling prior resource allocation, targeted treatments, and improving patient monitoring. Thus, the ensemble framework aligns well to develop decision support systems that help in early diagnosis and treatment of the growing AD rate.

Future research aimed at refining and implementing the ensemble model in routine treatment and clinical work will re- quire integration with real-time existing decision support sys- tems. Future recommended steps need constructing pipelines for MRI feature preprocessing, validating ensemble model outputs within diverse medical populations, and implementing user-friendly interfaces for neurologists to perform automated diagnostics.

The limitation is that the research dependence on MRI features excludes functional MRI or PET imaging, which could expose additional biomarkers vital for the early stage diagnosis of AD ([16].

REFERENCES

[1] Givian, H., Calbimonte, J. P., & for the Alzheimer's Disease Neuroimaging Initiative. (2024). Early diagnosis of Alzheimer's disease and mild cognitive impairment using MRI analysis and machine learning algorithms. Discover Applied Sciences, 7(1), 27.

[2] Fathi, S., Ahmadi, A., Dehnad, A., Almasi-Dooghaee, M., Sadegh, M., & Alzheimer's Disease Neuroimaging Initiative. (2024). A deep learning-based ensemble method for early diagnosis of Alzheimer's disease using MRI images. Neuroinformatics, 22(1), 89-105.

[3] Shaffi, N., Subramanian, K., Vimbi, V., Hajamohideen, F., Abdesselam, A., & Mahmud, M. (2024). Performance evaluation of deep, shallow and ensemble machine learning methods for the automated classification of Alzheimer's disease. International Journal of Neural Systems, 34(07), 2450029.

[4] Janghel, R. R., & Rathore, Y. K. (2021). Deep convolution neural network based system for early diagnosis of Alzheimer's disease. Irbm, 42(4), 258-267.

[5] An, X., Zhou, Y., Di, Y., & Ming, D. (2020, November). Dynamic functional connectivity and graph convolution network for Alzheimer's disease classification. In Proceedings of the 2020 7th International Conference on Biomedical and Bioinformatics Engineering (pp. 1-4).

[6] Huang, L., Yang, H., Che, Y., & Yang, J. (2024). Automatic speech analysis for detecting cognitive decline of older adults. Frontiers in Public Health, 12, 1417966.

[7] Khan, Y. F., Kaushik, B., Chowdhary, C. L., & Srivastava, G. (2022). Ensemble model for diagnostic classification of Alzheimer's disease based on brain anatomical magnetic resonance imaging. Diagnostics, 12(12), 3193.

[8] Naderi, M., Rastgarpour, M., & Takhsha, A. R. (2024). Early Diagnosis of Alzheimer's Diseases and Dementia from MRI Images Using an Ensemble Deep Learning. arXiv preprint arXiv:2412.05666.

[9] Alruily, M., Abd El-Aziz, A. A., Mostafa, A. M., Ezz, M., Mostafa, E., Alsayat, A., & El-Ghany, S. A. (2025). Ensemble deep learning for Alzheimer's disease diagnosis using MRI: Integrating features from VGG16,

MobileNet, and InceptionResNetV2 models. PloS one, 20(4), e0318620.
[10] Nguyen, D., Nguyen, H., Ong, H., Le, H., Ha, H., Duc, N. T., & Ngo, H. T. (2022). Ensemble learning using traditional machine learning and deep neural network for diagnosis of Alzheimer's disease. IBRO Neuroscience Reports, 13, 255-263.
[11] Lella, E., Pazienza, A., Lofu, D., Anglani, R., & Vitulano, F. (2021). An ensemble learning approach based on diffusion tensor imaging measures for Alzheimer's disease classification. Electronics, 10(3), 249.
[12] Jenber Belay, A., Walle, Y. M., & Haile, M. B. (2024). Deep ensemble learning and quantum machine learning approach for alzheimer's disease detection. Scientific Reports, 14(1), 14196.
[13] Sarica, A., Cerasa, A., & Quattrone, A. (2017). Random forest algorithm for the classification of neuroimaging data in Alzheimer's disease: a systematic review. Frontiers in aging neuroscience, 9, 329.
[14] Shojaie, M., Cabrerizo, M., DeKosky, S. T., Vaillancourt, D. E., Loewenstein, D., Duara, R., & Adjouadi, M. (2022). A transfer learning approach based on gradient boosting machine for diagnosis of Alzheimer's disease. Frontiers in aging neuroscience, 14, 966883.
Chatterjee, S., & Byun, Y. C. (2022). Voting ensemble approach for enhancing Alzheimer's disease classification. Sensors, 22(19), 7661.
[15] Devi, B. M., & Ganesh, D. (2024, November). Brain Image Analysis for Alzheimer's Disease Detection using Ensemble Machine Learning. In 2024 8th International Conference on Electronics, Communication and Aerospace Technology (ICECA) (pp. 779-785). IEEE.

Designing an Intelligent VR Retail Environment for Understanding Consumer Behaviour

Ardini Dayini Mohd Salihan
Department of Computer Science
Swansea University
Swansea, UK
2329779@swansea.ac.uk

Evi Indriasari Mansor
Business Analytics Program
Abu Dhabi School of Management
Abu Dhabi, UAE
e.mansor@adsm.ac.ae

Mohd Salihan Ab Rahman
Computer Information System Dept.
Higher College of Technology
Abu Dhabi, UAE
mrahman@hct.ac.ae

Abstract—This paper presents the initial development of an intelligent interactive virtual reality (VR) shopping experience designed to investigate consumer behaviour in a simulated retail environment. The application focuses on intuitive controls and user-friendly interactions to provide a realistic shopping experience, and includes features such as interactable objects, sound effects, and haptic feedback to enhance user engagement. While the primary focus of this work was on the initial design and implementation of the VR application, future work will involve developing intelligent AI features and evaluating user experience. This research aims to contribute to the growing insights on intelligent VR retail environments and offer practical guidance for developing immersive shopping experiences.

Keywords—Virtual Reality (VR), Interactive Virtual Shopping, Consumer Behaviour, E-Commerce, Retail Environment, Artificial Intelligence (AI)

I. Introduction

In recent years, virtual reality (VR), a technology that allows users to immerse themselves in a simulated 3D environment, has increasingly been used in areas like retail. Due to the realistic perspective it provides, it has been used for product visualisation and demonstrations, shopping experience personalisation, and examining consumer behaviour [1].

Artificial intelligence (AI) has also been increasingly used in e-commerce applications to implement AI assistants, voice searching, and recommendation tools, and understand consumer's behaviour while shopping [2]. The future of AI in e-commerce has great potential to integrate with emerging technologies including VR, reshaping how businesses interact with and understand their customers. [3].

Existing research on consumer behaviour in VR has primarily focused on the general aspects of user engagement and satisfaction [4]. While these studies provide valuable insights, they often overlook the specific interactions and decision-making process that occur in response to promotional offers within a VR shopping experience. This gap in knowledge limits retailers from utilizing their marketing strategies, not just for their sales rates, but for maximizing consumer engagement and shopping satisfaction.

This paper presents the early development of an intelligent interactive VR shopping experience designed to understand how consumers interact with products and how they may make decisions in response to offers involving the products, including buy-one-get-one offers and percentage discounts. The application's design was tailored towards users with minimal experience with VR technology, focusing on intuitive controls and user-friendly interactions. By understanding these interactions and responses, retailers can better tailor their VR shopping strategies to enhance consumer engagement and experience, and drive sales.

II. Methodology

A. Background of the Application

This application offers users the opportunity to explore a virtual store using a VR device (MetaQuest 3 headset [5]). Users can view and interact with products using controllers, and add products to their shopping cart. At the end of the shopping experience, they can purchase the items in their cart. Promotions are available for each type of item, and the goal is to determine how these offers may influence user purchasing decisions.

B. Development Tools

The VR application was developed in Unity using its XR Interaction Toolkit [6] and several C# scripts to assist in specific functionalities, and was tested with the Meta Quest 3 headset. Assets, materials, icons, and sound effects were imported from the Unity Asset Store [7], SketchFab [8], Vecteezy [9], Unsplash [10], Pixabay [11], and Flaticon [12]. These assets were selected for their quality and suitability for a beach-themed retail environment.

C. Virtual Store Design

The virtual store (Fig. 1) was designed according to a "beach day" concept, focusing on selling beach apparel. To emulate an immersive summer vibe to enter, the store was featured with large windows to let in "natural light", and a sandy terrain was created to be visible from the windows. The store layout is rectangular and simple, ensuring all items are visible to the user when turning 360 degrees. This design aims to prevent the user from being overwhelmed and facilitates easy navigation.

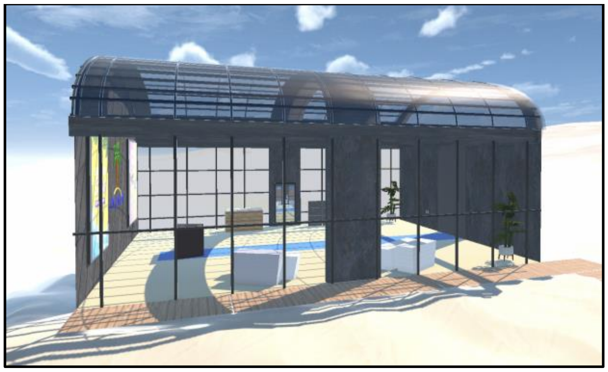


Fig. 1. Virtual store design with sandy terrain

The store is divided into three sections: hats, footwear, and sunglasses, with each type of garment offering its own discount.

D. Locomotion

Locomotion refers to how the user can move around in the environment. Two main features were applied:

1) Snap Turning: This feature rotates the user by a fixed angle when the thumbstick is toggled. This feature was chosen over continuous turning, which smoothly rotates the user over time, as snap turning is often less likely to cause motion sickness in users [13 - 14].

2) Teleportation: The user can teleport from one position to another within the designated area. This feature prevents the user from accidentally teleporting themselves outside the store area or into other objects like furniture or walls. Teleportation reticles or visual indicators (like the circular white marker in Fig. 2) appear when the user's rays point to a valid teleportation area, indicating the position where the user will land.

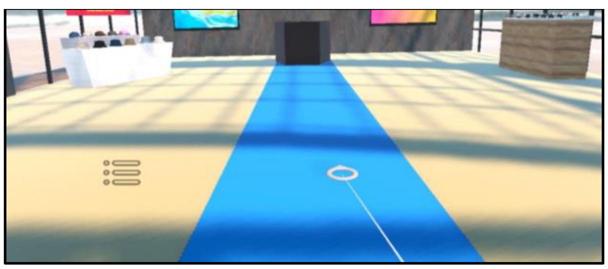


Fig. 2. Teleportation reticle indicating a valid teleportation area

E. Assets

Assets used for the application, including sunglasses, hats, and footwear (Fig. 3), were mainly imported from the aforementioned sources. Textures were customized to align with the beach-themed store. The furniture, including the display tables and platforms, was built using ProBuilder [15]. Two mirrors were also created (Fig. 6) by adding inverted planes with cameras attached to accurately display reflections.

Fig. 3. Footwear assets

Promotion banners were created using a graphic design tool and attached to signs hung over their corresponding items to bring more attention to available offers, such as in Fig. 4. Additionally, animated graphics were produced and attached to the screens behind the counter as video clips playing on loop to simulate in-store video advertisements.

Fig. 4. Promotion banner showing percentage discount for sunglasses

F. Interactivity

To make it more convenient for the user to interact with the environment, XR Ray Interactors were implemented, allowing users to point at and select interactables from afar instead of having to teleport directly in front of the item. When the user hovers over a selectable item, the originally red rays turn white, giving clear indications to the user. Another feature of the ray interactors that was applied is the Force Grab property, which moves an XR Grab Interactable object to the user's hand when selected (as shown in Fig. 5). Similar to real-life shopping experiences, this allows the users to inspect the items at a closer level.

Fig. 5. Sunglasses being held in the user's hand after grabbing it from a far

Making the experience seem more realistic, all items on display, with all of them being XR Grab Interactable objects, are affected by basic laws of physics. This means they can collide with other objects, and when dropped by the user, they fall to the ground as they would in real life. However, to facilitate a smoother user experience and reduce clutter in the environment, a C# script was added to automatically return the moved items to their original positions after a set time delay.

To further mimic a realistic shopping experience, sockets were created for the user's head and face, serving as a placeholder for hats and sunglasses, respectively, allowing the user to try on different items and check their reflections in the mirror (Fig. 6). Attach points on both the sockets and the interactable items ensure they align properly on the user's head or face when tried on. Additionally, enforcing specific layers for each socket prevents hats from being placed in the sunglasses position, and vice versa.

Fig. 6. Reflective mirror - User trying on a hat and a pair of sunglasses

Finally, haptic feedback was implemented to enhance imagination and interaction by simulating a sense of touch. Whenever the user hovers over an item, the controllers vibrate slightly; when an item is selected, the controllers have a slightly stronger vibration. Sound effects reacting to the player's actions were added to further improve the user's immersion in the environment.

G. User Interface

A user interface was implemented to help the user navigate the environment and become more aware of various items and offers. For example, when the user first starts the game, a set of instructions is automatically displayed along with an avatar, explaining the basic controls of locomotion and interactions (Fig. 7).

Fig. 7. Opening instructions

When the user hovers over a purchasable item, its original price and promotion offer will automatically be shown. Additionally, when the item is selected as shown in Fig. 8, an Add to cart button will appear, allowing the user to add the item to their shopping cart.

Fig. 8. Price and offer for the selected hat, with an Add to cart button

At all times, the user has access to the menu button (represented by the three-bullet icon), which allows the user to view their cart, or restart the simulation (Fig. 9).

Fig. 9. Menu options

As shown in Fig. 10, when viewing their shopping cart, users can see the names of the items added as well as their prices. Users have the option to remove items by clicking on the trash icons on their respective lines.

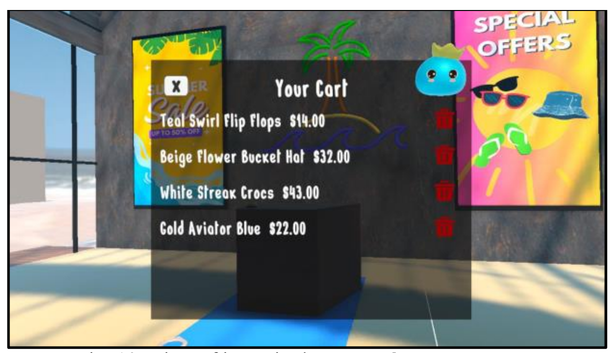


Fig. 10. View of items in the user's shopping cart

When the user is ready to purchase the items in their cart, they can ring the bell at the counter, which will prompt a confirmation message as shown in Fig. 11. If the user decides to continue, they complete their purchase, and may choose to shop again or quit the simulation.

Fig. 11. Payment message after ringing the bell at the counter

H. Future AI-Features

The application aims to implement intelligent AI features to understand consumer interactions with products. The proposed features include:

- Integrated shopping assistant - an AI-driven avatar accompanies the customer throughout the shopping experience, assisting them through conversations, answering questions and comparing products (Fig. 8-12).
- Personalised product recommendations - the application analyses the customer's browsing behaviour and past purchases.
- Dynamic offer generation - based on the customer's shopping behaviour, the application offers personalised discounts.
- Customer journey analytics - the application analyses users' interaction with each section of the store including engagement with the products, time spent and the shopping paths in the store.

III. Conclusion

This study presents the development of an interactive VR shopping experience designed to explore consumer behaviour in a simulated retail environment. The application focuses on the intuitive controls and user-friendly interactions, providing a realistic and engaging shopping experience that mimics real-life scenarios. While the first phase of development has been successfully completed, the AI features have yet to be implemented, and the evaluation of the application remains a critical next step.

Future work will involve enhancing the application with AI features and conducting user evaluation with a group of participants to evaluate the effectiveness of the VR shopping experience. This will include assessing the user engagement, ease of use, and overall satisfaction. The results of these evaluations will provide valuable insights into consumer behaviour and help refine the application for broader use. Additionally, the future work could explore the implementation of advanced analytics to capture user interaction data, further enhancing the understanding of consumer decision-making processes in VR retail environments.

Acknowledgments

We thank the AI Management Institute (AIMI), Abu Dhabi School of Management (ADSM), UAE for their support of this project and for the opportunity provided through the AI summer research initiative.

REFERENCES

[1] Erensoy, A., Mathrani, A., Schnack, A., Elms, J., & Baghaei, N. Consumer behavior in immersive virtual reality retail environments: A systematic literature review using the stimuli-organisms-responses (S-O-r) model. Journal of Consumer Behaviour, vol. 23, issue 6, pp. 2781–2811, 2024. DOI: 10.1002/cb.2374

[2] Fedorko, R., Kráľ, Š., Bačík, R. Artificial Intelligence in E-commerce: A Literature Review. In: Saraswat, M., Sharma, H., Balachandran, K., Kim, J.H., Bansal, J.C. (eds) Congress on Intelligent Systems. Lecture Notes on Data Engineering and Communications Technologies, vol 111. pp 677–689, 2022. Springer, Singapore. DOI:10.1007/978-981-16-9113-3_50

[3] Raji, M. A., Olodo, H. B., Oke, T. T., Addy, W. A., Ofodile, O. C., & Oyewole, A. T. E-commerce and consumer behavior: A review of AI-powered personalization and market trends. GSC Advanced Research and Reviews, vol. 18, issue 3, pp. 66-77, 2024.

[4] Ambika, A., Shin, H., & Jain, V. Immersive technologies and consumer behavior: A systematic review of two decades of research. Australian Journal of Management, vol. 50, issue 1, pp. 55-79, 2025.

[5] Meta Quest 3. Meta. 2025. [Online]. Available: https://www.meta.com/quest/quest-3/

[6] XR Interaction Toolkit. Unity. 2025. [Online]. Available: https://docs.unity3d.com/Packages/com.unity.xr.interaction.toolkit@3.1/manual/index.html

[7] Asset Store. Unity. 2023. [Online]. Available: https://assetstore.unity.com/

[8] Sketchfab. Sketchfab. 2025. [Online]. Available: https://sketchfab.com/

[9] Vecteezy. Vecteezy. 2025. [Online]. Available: https://www.vecteezy.com/

[10] Unsplash. Unsplash. 2025. [Online]. Available: https://unsplash.com/

[11] Pixabay. Pixabay. 2024. [Online]. Available: https://pixabay.com/

[12] Flaticon. Freepik. 2025. [Online]. Available: https://www.flaticon.com/

[13] Sargunam, S. & Ragan, E. Evaluating joystick control for view rotation in virtual reality with continuous turning, discrete turning, and field-of-view reduction. In Proceedings of the 3rd International Workshop on Interactive and Spatial Computing (IWISC '18). Association for Computing Machinery, New York, NY, USA, pp. 74–79, 2018. DOI: 10.1145/3191801.3191815

[14] Farmani, Y & Teather, R. Viewpoint Snapping to Reduce Cybersickness in Virtual Reality. 2018. DOI: 10.20380/GI2018.23.

[15] About ProBuilder. Unity. 2025. [Online]. Available: https://docs.unity3d.com/Packages/com.unity.probuilder@6.0/manual/index.html

E-Recruitment in the UAE: An Artificial Intelligence Approach

Abdulla Aldabal
Business Analytics
Abu Dhabi School of Management
aldabal@jobx.ae

Abstract— This study proposes a hybrid AI framework for UAE e-recruitment, combining LLM, RAG, ML, and XAI to optimize candidate-job matching, support Emiratization, and ensure compliance with UAE Federal Law No. 45 on data protection. It automates recruitment while reducing bias and ethical risks through governance strategies. The research evaluates AI's impact on hiring productivity, equity, and efficiency using surveys, interviews, and ML analysis. It also provides guidelines for responsible AI adoption and a roadmap for developing recruitment systems aligned with UAE legal and societal expectations. The goal is fairer, more efficient hiring processes.

Keywords—E-Recruitment, Artificial Intelligence (AI), Emiratization, Machine Learning (ML), Large Language Models (LLMs), Natural Language Processing (NLP), Deep Learning (DL), Retrieval-Augmented Generation (RAG), Data Privacy, UAE Federal Law No. 45 of 2021, Ethical AI Governance, Algorithmic Fairness, Bias Reduction

I. INTRODUCTION

Recruitment processes are an integral part of all businesses, and entities as they influence the operational productivity and achievement of strategic objectives through human capital [1]. Existing online recruitment and onboarding systems continue to deal with numerous drawbacks that include mismatching candidates' qualification to jobs, biased screening processes, and overall long hiring processes. These issues not only result in hiring inefficiencies but also adversely impact the diversity and quality of the workforce. The adoption of Artificial Intelligence (AI) algorithm, promises to revolutionize these traditional methods. The solution will improve candidate-job matching capabilities and reduce biases to enhance overall recruitment quality and efficiency [2].

The recruitment sector in the UAE is expanding rapidly due to the demand for new and innovative hiring strategies. In 2021, the value of online recruiting in the Middle East and Africa was estimated to be around USD 1.94 billion and is predicted to increase to USD 2.51 billion by 2028, at a CAGR of 3.8% [3]. The magnitude of this growth illustrates the competitive job market exceeding 155 competitors along with 27 job boards such as Bayt.com, LinkedIn, Naukrigulf.com, and GulfTalent that provide comprehensive services for recruiters and candidates. The region's impact on creating a demand for more sophisticated recruitment techniques to achieve effectiveness as well as competitive advantage [4].

The Emiratization Policy, has been implemented to foster the employment of Emirati nationals as part of the UAE National Innovation Strategy 2015, is focused on workforce diversification and improving employment security for Emirati citizens. The current over-dependence of the UAE private sector on the expatriate workforce, which constitutes 84.59% of employees, serves as a hindrance to achieving this national goal [5] This population context offers a challenge and an opportunity for improving recruitment systems that increase organizational efficiency while addressing national workforce supply issues.

Recruitment systems are set to shift drastically due to the introduction of AI-powered based electronic recruitment systems. These systems automate and provide more accurate matching of professional seekers and employers while correcting inefficiencies and biases that are left unchecked by traditional systems [6]. The expectation is the introduction of such AI capabilities as a framework will shift hiring paradigms by enhancing accuracy, equity, and adherence to local legal and ethical norms. They also offer substantial improvements by optimizing the recruitment process and addressing the limitations faced by job seekers and employers due to inadequate matchmaking [7].

This study seeks to create and implement an AI recruitment framework for electronic recruitment platforms with special attention to candidate sourcing, bias reduction, and compliance with the Federal Law No. 45 of 2021 in the UAE [8]. This framework will improve the efficiency of the recruitment processes by automation and accurate matching of the employers to the employees and ensure compliance with national employment policies and legal stipulations, hence contributing to Emiratization and workforce diversity within the UAE.

This study focuses on creating an e-recruitment framework powered by Artificial Intelligence (AI) that assists in candidate-job matching, compliant with UAE Federal Law No. 45, and supports Emiratization. The scope of this work is limited to evaluating the HR processes by applying AI ethics, monitoring the recruitment key performance indicators (KPIs), and providing practical recommendations. Other AI functions not related to recruitment, such as retention or performance appraisal, are outside the scope of this study.

The study also utilizes both qualitative (interviews with human resource managers, recruiters, and specialists in artificial intelligence) and quantitative (surveys of various stakeholders) methods to investigate the use of artificial intelligence in recruitment in the United Arab Emirates. The legal documents are scrutinized within the framework of described biases, compliance challenges, as well as attempts to refine factual practices regarding ethical and legal candidate-job matching.

As the use of AI-powered recruitment systems becoming increasingly widespread across the world, a significant gap in frameworks exists that suit the UAE's unique labor regulations, Emiratization policies and multicultural workforce dynamics. Current research primarily deals with Western markets, and in some cases, disregards compliance with UAE's data privacy laws and the necessity of explainability within the AI hiring tool. A gap exists in terms of fairness, transparency, and legal compliance in the case of the existing AI framework for the UAE. This study attempts to fill the space by introducing a hybrid AI framework which includes completeness in terms of fairness, transparency, and legal compliance for UAE e-recruitment context.

The developed conceptual framework is meant to serve as a guide and remains open to future design and integration based on the results of this thesis.

II. PROBLEM STATEMENT

The UAE's job market is undergoing active development, yet is faced with persisting systemic problems that affect both employers and employees. Employers face issues such as lengthy hiring times because of skill gaps and inefficiencies associated with hiring and candidate evaluation processes; these issues lower employer recruitment efficiency and create a mismatch between available talents and organizational requirements of entities [9]. On the other hand, Emiratis and expatriate job hunters face a myriad of challenges including discrimination and inadequate job offers made worse by oversupply of poorly suited jobs [10].

Employment skill gap is one of the most crucial problems within the job market for both Emiratis and expatriates. Emirati job candidates often find themselves lacking the requirements of the job market, especially for the private sector. There are also many expatriate applicants who fall under the category of not being able to satisfy the requirements and expectations of employers. This occurs as a result of the lack of efficiency in traditional recruiting processes, which contributes to poor time and resources management that affects capitalization [9], [1].

The recruitment process suffers from biases and outdated filtering techniques which delay `recruitment and make it difficult to assess a candidate accurately. These issues not only reduce productivity, but also raise ethical issues related to inclusion and diversity. Such biases, whether based on gender, ethnicity or age, restrict the potential of deserving candidates and thus, violate basic principles of discrimination-free employment [11].

The vast use of AI recruitment technologies has revealed a substantial lack of research about how hybrid AI systems can benefit e-recruitment services in the one-of-a-kind UAE employment environment. The integration of LLM + RAG + ML receives limited research attention for its ability to enhance recruitment procedures through improved search functions and hiring system automation. The utilization of AI recruitment technology in the United Arab Emirates labor market, specifically for Emiratization procedures, requires further study because Western market studies dominate research into AI systems.

The screening of candidates, as well as matching them to appropriate positions, is streamlined through the use of AI technology. The adoption of AI also diminishes existing inefficiencies and biases, contributing to a faster and fairer recruitment process. Recruitment platforms can manage large candidate portfolios more efficiently by employing AI, which also helps most Emirati and expatriate applicants get better jobs [1].

While AI systems are already increasingly utilized in hiring, glaring issues still persist with current e-recruitment systems. Conventional applicant tracking systems (ATS) tend to have specific serious flaws: studies show that most employers believe that automation in screening processes poses a significant problem because qualified candidates are filtered out by stringent rules based on matching criteria [12]. Additionally, the use of AI recruitment software comes with well-understood bias problems, with some tools replicating or worsning discriminatory biases that exist in hiring practices [13]. Such biases can result in discrimination toward certain groups, impacting the intended objective of enhancing fairness and diversity. A further secondary concern is the "black box" aspect of many hiring AIs; they do not provide information or justification for their decisions. In the absence of mechanisms involving explainable AI (XAI), recruiters and candidates are none the wiser as to why a particular candidate was deselected, which turns focus on responsibility issues [14][15]. The highlighted points are illustrative of a clear gap in research AI-based recruitment systems have focused primarily on efficiency and neglected concepts of equity and interpretability.

The Emiratization National Policy Framework – intended to enhance employment opportunities for UAE citizens – requires local policies to be compliant with local workforce policies and inclusivity frameworks [16]. AI-powered recruitment tools and software do not often include such policies, which may jeopardize compliance with important national employment fulfillment goals. Also, the UAE's new data protection law (Federal Decree-Law No. 45 of 2021) places heavy restrictions on the use of personal information for hiring with regard to the candidate's privacy and consent [17]. More constraining and justifiable AI systems are needed in this context. Therefore, there is an uparalleled need for an AI-based e-recruitment system designed for the UAE that incorporates fairness tailoring features, explainsable AI (XAI) for decision-making transparency, and compliance with regional legislations and cultural frameworks. In addressing these gaps, the designed framework mitigates biases and integrates explainability into the recruitment process, unlike other solutions lacking such integration of ethics and legal frameworks.

III. LITERATURE REVIEW

The adoption of artificial intelligence (AI) and machine learning (ML) technologies has profoundly transformed the management of human resources (HR), especially in regards to employees' recruitment [18]. This review examines the impact of AI technologies on the recruitment procedures of organizations. It concentrates on some e-recruitment components such as predictive analytics, natural language processing (NLP), and deep learning (DL) among others. These AI technologies have automated several other processes, resulting in more efficient candidate screening and matching.

In the United Arab Emirates (UAE), the recruitment frameworks are sophisticated and employ provisions that facilitate the nationalization of employment—the Emiratization policy focuses on hiring UAE nationals across public and private sectors. The recruitment process involves several core activities, including job advertising, sifting, shortlisting, and interviewing, which are increasingly influenced by automation via e-recruitment systems [19].

Recruitment begins with crafting a detailed candidate pool, which starts with crafting job descriptions and advertisements aimed at qualified applicants. In the UAE, candidates can be sourced using older methods such as job adverts and recruitment agencies, through referrals and internal promotion programs, and lately through online job portals and social media [20][21]. Furthermore, portals and professional networks have further enabled access to new and varied candidates for recruiters using digital forms of recruitment [20].

One of the most widely used tools is the resume screening tool, or the ATS, which is used to automate initial resume screening [22]. Oracle Taleo, SAP SuccessFactors, and Workday are major ATS systems and are accepted widely across industries. The primary function of these systems is candidate matching based on specific keywords and set benchmarks, with qualifications, skills, and experience being evaluated [22] .

Recruitment as an industry is marked by diversity and dependency on algorithms to organize non-linear data. This processes information that is often complex and unstructured, making patterns useful in hiring. Random Forest models assist with providing interpretability which is a critical component to transparency and fairness in the hiring process. Ensemble learning techniques are effective for classification and anomaly detection, as with Random Forest, bagging, and boosting [23]. These models contain important information on why some candidates may be labeled as outliers or dismissed, thereby shedding light on their perceptions.

Building from the preceding works on Random Forest's benefits, its integration into recruitment processes can enhance decisions and alleviate the biases that stem from them. Placing Random Forest with other sophisticated tools like large language models (LLMs) and retrieval augmented generation (RAG) could optimize the recruitment process while ensuring adherence to Emiratization policies concerning employment in the UAE. Thus, it is important to study AI's roles in recruitment to fulfill the United Arab Emirates' propelling talent acquisition demands.

As far as utility in recruitment goes, Random Forest is one of the most promising ML algorithms [24]. This is advantageous on account of the fact that it processes, multi-source, intricate information with precision and reliability. For instance, Singh (2024) noted that Random Forest excelled over other algorithms in recruitment processes because it manages high-dimensional, chaotic data while mitigating overfitting. Such ability has been reported in other areas too, particularly in finance [23]. E-recruitment systems are designed with the expectation that they will fairly analyze various types of data without bias while maintaining equity and transparency in their decisions.

Employing LLMs for Natural Language Processing (NLP) fueled job matching is becoming a common feature of modern e-recruitment systems. An older model of an Applicant Tracking System (ATS) relies on keyword searching and tends to disqualify many relevant candidates irrespective of the potential they might have, leading to worse results. LLM-powered NLP models drastically improve job matching accuracy by deep learning from the attribute data of job seekers [25] . Well-known LLMs are Generative AIs such as OpenAI's Chat GPT and Google's BERT, which understand and generate text from processing colossal datasets. These models are capable of working with both structured and unstructured information, which makes them suitable for résumé parsing, creating job descriptions, candidate profiling, and more. Moreover, multilingual LLMs such as ESCOXLM-R+ assist international firms in cross-border hiring by classifying resumés to global vocabularies of job descriptions.

RAG facilitates real-time knowledge retrieval which supports recruiters with improved decision-making systems and enhancing effectiveness in hiring. Through the inclusion of external data sources, RAG offers insight on hiring behavior, trends in the labor market, and employer expectations, assisting organizations in their hiring decisions [26]. This model also helps in mitigating bias since evaluations are made based on empirical data rather than subjective human perception.

The following Literature summary reviews aids in understanding the recruitment process and how AI affects the entire ecosystem:

Author	**Objective**	**Methodology**	**Findings**	**Discussion**
Jatobá et al. (**2023**)	Analyzing research HRM and the adoption of AI	systematic literature review	Growing academic interest in AI's role in HR development. AI in strategic HR is primarily focused on profit maximization and organizational growth.	AI in strategic HR is primarily focused on profit maximization and organizational growth.
Larsson et al. (**2024**)	Examining the Risk of Discrimination in AI Systems:	110 completed questionnaires from representatives of 10 major recruitment agencies and 100 large Swedish companies	AI and ADM raises concerns about transparency and awareness of bias risks.	There is need for greater transparency, accountability, and awareness in AI-driven recruitment processes. It emphasizes the importance of defining responsibility for mitigating bias and ensuring fairness in

				hiring decisions.
Al-Quhfa et al. (2024)	analyzes the use of machine learning models to enhance recruitment accuracy and efficiency in business intelligence.	Recruitment data from a major Yemeni organization (2019–2022) Hyper-parameter tuning and cross-validation were applied for optimization.	Random Forest achieved the highest accuracy (92.8%), followed by Neural Networks (92.6%) and Gradient Boosting Classifier (92.5%).	Advanced machine learning models can optimize hiring strategies in business intelligence.
Almeida et al. (2025)	Explores recruiters' perceptions of AI tools in recruitment, using the Technology Acceptance Model (TAM) to analyze ease of use, usefulness, and attitudes toward AI.	Qualitative study: 100 recruiter interviews on AI adoption. Quantitative study: Online questionnaire with 355 recruiters.	AI enhances efficiency and resource management but raises concerns about loss of personal interaction and role adaptation.	Ethical considerations and human involvement are crucial for effective AI integration. Provides actionable recommendations for organizations adopting AI in recruitment.
Zheng et al. (2024)	Introduces BAMBOO, a novel bilateral multi-behavior modeling method for reciprocal recommendation in online recruitment to better match job seekers and recruiters.	conducts offline experiments on real-world datasets. Performs online A/B testing to evaluate real-world effectiveness.	BAMBOO outperforms state-of-the-art baseline methods in accuracy and efficiency of job-candidate matching.	The proposed BAMBOO method offers an innovative way to improve the matching accuracy between job seekers and recruiters by incorporating multi-typed user behaviors and dual perspectives.
Qin et al. (2023)	Analyze AI-driven talent analytics in HRM, categorizing applications and identifying challenges.	Survey of AI techniques in talent analytics, categorizing research into talent management, organization management, and labor market analysis.	AI enhances HRM through deep learning, NLP, predictive analytics, and automated decision-making, but data bias and explainability remain challenges.	AI-driven HRM improves recruitment, retention, and workforce planning, but requires fairness, transparency, and integration with evolving digital HR systems.
Zeng et al. (2024)	Investigate privacy risks in RAG systems and their impact on LLM data leakage.	Empirical studies with attack methods on RAG privacy vulnerabilities.	RAG exposes retrieval data but reduces LLM training data leakage.	RAG improves security for LLMs but requires stronger retrieval data protection.
Bano et al. (2024)	Explores diversity and inclusion in AI-based recruitment and lessons from industry applications.	Industry workshop discussions, qualitative analysis.	Identifies biases in AI-driven hiring, challenges in diversity, and potential solutions.	Highlights best practices and regulatory needs for ensuring fairness in AI recruitment.
Ali & Kallach (2024)	Analyzes AI-enabled HR recruitment functionalities through a scoping review.	Review of AI-based recruitment applications in HRM.	AI enhances efficiency, reduces bias, but requires careful implementation.	Discusses ethical concerns, AI transparency, and HR adaptability in AI-driven hiring.
Chapano & Iwu (2025)	Evaluating the implementation of HR practices	Qualitative case stud	HR and recruitment strategies, to provide reasoning for results such as transparency and how AI supports results	The AI roles in enhancing HR practices by enabling the clarification of efficient candidate assessment and transparency

Wu et al. (2024)	Candidate evaluation with by using data driven analytics	Fine-tuned LLM and multimodal data for job to CV matching; recruiter-guided corrections	Enhancing the matching process accuracy and introduce fairness in candidate evaluation	Highlights on combining AI various models with human expertise that will optimize recruitment and expand on bias
Potočnik, Anderson, Born, Kleinmann, & Nikolaou (2021)	Provide the current recruitment elements, which includes, developments, challenges, and opportunities	Systematic Literature Review	Provides the insight of highlighting the AI integration and machine learning tools in reducing bias and increase recruitment efficiency	The AI technologies implemented in the recruitment processes and focusing on the bias issues and identify some ethical concerns for decision making

.

IV. METHODOLOGICAL APPROACH

To develop the AI-based e-recruitment framework, information will be gathered using various data collection methods. To understand and assess the recruitment practices as well as the stakeholders' needs, primary data will be collected through semi-structured interviews and surveys with HR professionals and job applicants in the UAE. Additionally, to create a dataset for training and validating machine learning models, historical recruitment data such as job postings, candidate profiles, and hiring outcomes will be recorded. This approach allows the framework to incorporate stakeholders' integration and perspectives alongside factual recruitment data.

After it is created, the framework will be tested on a basic deployment within a recruitment context to evaluate its effective functioning. The AI's recommendations will be evaluated by measuring how they are aligned with human judgment by checking how hiring decisions ranked against AI suggested ranks. The AI's advantages will be measured by defined criteria: time-to-hire, meaning if the AI actually makes it faster to fill open positions, and matching accuracy, which measures how well AI suggested candidates against actual hires. In addition, recruiters and candidates will be interviewed or surveyed to determine their perceived value and enhancements associated with the system due to the changes made. All of these factors combined will provide an assessment of the effectiveness of the framework, its proven effectiveness, and guiding iterative optimization. The developed conceptual framework will be implemented and integrated where parameters defined in the thesis are bound to be the foundation, and deployed under the context of e-recruitment where online platforms are used as conditions for application.

No research activity will diverge from ethical compliance. Each data collection component will first seek designated ethics clearance based on the institutional policies ADSM-set guidelines. Ethical recommendations as well as the signed consent will be required from all case study participants and the data collected is to be kept confidential and kept only accessible under specific conditions that require anonymization of every identifying detail in the collected data without compromising validity. At minimum, following ADSM ethic principles guarantees that rights and welfare of the participants will be safeguarded continuously throughout the life cycle of the project.

In this qualitative phase, semi-structured interviews will be useful for exploring depth and context-specific nuances that are beyond the scope of quantitative data. From their conversations with human resource managers, recruiters, and policy implementers, the research will investigate the perception of AI across various organizational cultures, industries, and the degree of automated systems resultant trust among stakeholders. Along with other participants, others will elaborate with potential ethical or legal issues such as concern about bias or privacy issues. This understanding can help with precise modifications to the framework that uphold compliance with local labor laws, Emiratization policies, and recruitment best practices in the UAE, ensuring focus on the region's specific needs.

At the same time, the surveys will gather more diverse quantitative data from a larger group of job applicants and HR professionals. The responses will contribute towards understanding broad AI hiring perception, including the viewpoints of whether candidates consider the processes as transparent or whether HR teams feel that AI takes over tedious administrative functions. The survey will contain questions measuring satisfaction and fairness on a Likert-type scale, as well as open-ended questions allowing participants to articulate UAE-specific contextual issues that they uniquely face. These data will enhance the interview findings with a macro-perspective capturing dominant attitudes and concerns, which will be beneficial when merging machine learning and human-centered AI analysis.

V. PROPOSED FRAMEWORK

A hybrid AI recruitment framework proposed for electronic recruitment platforms created and operated in the UAE such as JobX.ae that can employ LLMs, RAG, and ML within the recruitment platform to simplify, optimize, and provide a fair the recruitment process, while addressing the issues and implementing XAI for transparency. The sequence of logical implementation of the framework for JobX.ae is as the following:

The framework explores the capability to implement processes for one of the platforms being used in the UAE, while laying out the preprocessors of the current logic in the back end. This visualization helps to better implement the AI technologies hence this provides the viewing of technical integration and capability to have numerous applications in the back end. One of the suitable applications which are operating in the UAE is JobX, which has the following logical flow of data:

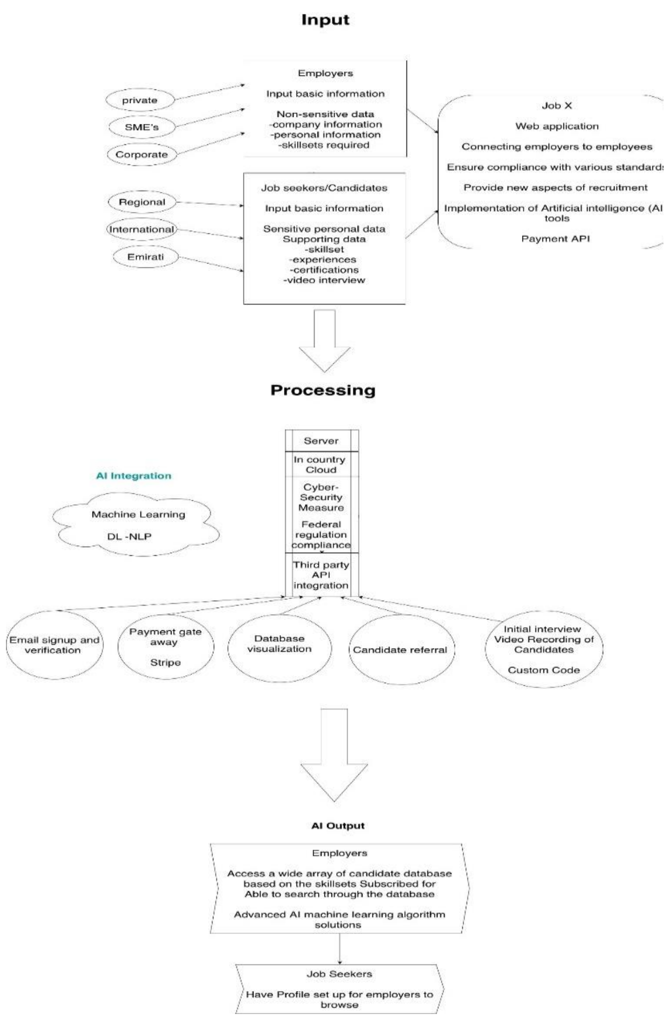


Figure 2: System internal Design, JobX.

VI. CONCLUSION AND FUTURE OUTLOOK

This shows in this study that integrating current Artificial intelligence (AI) technologies, like Large Language Models (LLMs), Retrieval Augmented Generation (RAG), Machine Learning (ML), and explainable artificial intelligence (XAI), fundamentally improves e-recruitment processes in UAE. The proposed AI as a service framework thus alleviates human recruiters from operational tasks (candidate sourcing, screening, and matching) to strategic ones (being a human interface, influencing cultural fit, and making strategic decisions) through intelligent automation. In addition, the model provides transparency, fairness, and compliance with the UAE's Federal Decree-Law No. 45 of 2021 on data protection, ensuring high levels of candidate trust as well as protecting their privacy in the recruitment process.

In addition, the framework makes Emiratization efforts possible by placing policies in the algorithmic logic such that the national workforce development criterion is given priority without compromising merit-based hiring. The system achieves a balance of organizational performance with wider social and regulatory goals through a blend of operational efficiency and ethical governance. Additionally, incorporating XAI gives explainable pathways to the decision-making process for both recruiters and candidates for career selection, thereby meeting the rising demands for AI bias and accountability.

All the same, the proposed framework is a conceptual framework the study recognizes to be a foundation that calls for real world testing and refinement. Moving forward, the pilot implementation of the hybrid AI framework in live e recruitment platforms like JobX.ae would be a major future research direction to assess practical factors like system usability, recruiter and candidate satisfaction, fairness perception, and whether it will result in improving time to hire. Such models and transparency measures will need to be iteratively refined using real-world user feedback.

Further inquiry into the long-term impacts of AI on hiring and organizations' performance should also be conducted. Longitudinal studies could measure, for instance, changes in diversity, equity, and inclusion, candidate (and employee) performance, retention, and opportunities for advancement over several years for candidates selected with AI. Furthermore, it would be important to understand whether AI-optimized recruitment leads to sustainable but positive outcomes to provide robust validation of the system's effectiveness after the initial hiring phases.

REFERENCES

[1] Ali, O., & Kallach**, L. (2024).** Artificial intelligence-enabled human resources recruitment functionalities: A scoping review. Procedia Computer Science, 232, 3268–3277. https://doi.org/10.1016/j.procs.2024.02.142

[2] Masood, F. (2024**).** The role of AI in shaping the future of labor markets: A comparative analysis of developed vs. emerging economies. International Journal of Emerging Multidisciplinary: Social Science, 3(1), 1–8. https://doi.org/10.54938/ijemdss.2024.03.1.346

[3] Mark Williams Recruitment**. (2024).** Emiratisation market report 2024**.** Ru'ya Careers.https://ruyacareers.ae/pdf/Mark-Williams-Emiratisation-Market-Report2024.pdf

[4] Business Market Insights. (2021). Middle East and Africa online recruitment market. https://www.businessmarketinsights.com/reports/middle-east-and-africa-online-recruitment-market

[5] Habbal, F., & Al Falasi, D. B. (2024). Future development strategies to enhance national workforce sustainability: A comprehensive review. Emirati Journal of Business, Economics, & Social Studies, 3(1). https://doi.org/10.54878/1s9jdf05

[6] Rigotti, C., & Fosch-Villaronga, E. (2024). Fairness, AI & recruitment. Computer Law & Security Review, 53, 105966. https://doi.org/10.1016/j.clsr.2024.105966

[7] Hunkenschroer, A. L., & Luetge, C. (2022). Ethics of AI-enabled recruiting and selection: A review and research agenda. Journal of Business Ethics, 178, 977–1007. https://doi.org/10.1007/s10551-022-05049-6

[8] Abouahmed, A., Kandeel, M. E., & Zakaria, A. (2024). Personal data protection in the United Arab Emirates and the European Union regulations. Journal of Governance & Regulation, 13(1), 195–202. https://doi.org/10.22495/jgrv13i1art17

[9] Bano, M., Zowghi, D., Mourao, F., Kaur, S., & Zhang, T. (2024). Diversity and inclusion in AI for recruitment: Lessons from industry workshop. arXiv, abs/2411.06066. https://doi.org/10.48550/arXiv.2411.06066

[10] Al-Quhfa, H., Mothana, A., Aljbri, A., & Song, J. (2024). Enhancing talent recruitment in business intelligence systems: A comparative analysis of machine learning models. Analytics, 3(3), 297–317. https://doi.org/10.3390/analytics3030017

[11] Larsson, S., White, J. M., & Bogusz, C. I. (2024). The artificial recruiter: Risks of discrimination in employers' use of AI and automated decision-making. Social Inclusion, 12(1). https://doi.org/10.17645/si.7471

[12] Fuller, J. B., Raman, M., Sage-Gavin, E., & Hines, K. (2021). Hidden workers: Untapped talent [Research report]. Harvard Business School Project on Managing the Future of Work.

[13] Chen, Z. (2023). Ethics and discrimination in artificial intelligence-enabled recruitment practices. Humanities and Social Sciences Communications, 10, Article 567. https://doi.org/10.1057/s41599-023-02079-x

[14] Magham, R. K. (2024). Mitigating bias in AI-driven recruitment: The role of explainable machine learning (XAI). International Journal of Scientific Research in Computer Science, Engineering and Information Technology, 10(5). https://doi.org/10.32628/CSEIT241051037

[15] Adadi, A., & Berrada, M. (2018). Peeking inside the black-box: A survey on explainable artificial intelligence (XAI). IEEE Access, 6, 52138–52160. https://doi.org/10.1109/ACCESS.2018.2870052

[16] Rees, C. J., Mamman, A., & Bin Braik, A. B. (2007). Emiratization as a strategic HRM change initiative: Case study evidence from a UAE petroleum company. The International Journal of Human Resource Management, 18(1), 33–53. https://doi.org/10.1080/09585190601068268

[17] Rwashdeh, M., & Abu-Shattal, Z. (2022, February 10). New UAE federal data protection law. K&L Gates. https://www.klgates.com/New-UAE-Federal-Data-Protection-Law-2-10-2022

[18] Liu, J., Chen, K., & Lyu, W. (2024). Embracing artificial intelligence in the labour market: The case of statistics. Humanities and Social Sciences Communications, 11(1). https://doi.org/10.1057/s41599-024-03557-6

[19] Potočnik, K., Anderson, N. R., Born, M., Kleinmann, M., & Nikolaou, I. (2021). Paving the way for research in recruitment and selection: Recent developments, challenges and future opportunities. European Journal of Work and Organizational Psychology, 30(2), 159–174. https://doi.org/10.1080/1359432X.2021.1904898

[20] Abbas, S. I., Shah, M. H., & Othman, Y. H. (2021). Critical review of recruitment and selection methods: Understanding the current practices. Annals of Contemporary Developments in Management & HR (ACDMHR), 3(3), 46–52. https://doi.org/10.33166/ACDMHR.2021.03.005

[21] Chapano, M., & Iwu, C. G. (2025). Exceeding expectations: A study on human resource management implementation in construction organisations. SA Journal of Human Resource Management, 23(0), Article a2849. https://doi.org/10.4102/sajhrm.v23i0.2849

[22] Suraj, M., Aruna, M., Binila, M., & Chandran, B. (2019). A descriptive study on Applicant Tracking System: Automation software for recruitment and selection. International Journal of Research and Analytical Reviews. https://www.ijrar.org/papers/IJRAR19VP033.pdf

[23] Alnaqbi, M., Al-Ali, M. M., Alremeithi, M., Al Ali, M. Y., & Pavithran, D. (2022, November). Different techniques and algorithms to combat the issue of money laundering in Bitcoin. In 2022 International Conference on Electrical and Computing Technologies and Applications (ICECTA) (pp. 122–126). IEEE. https://doi.org/10.1109/ICECTA57148.2022.9990475

[24] Salman, H. A., Kalakech, A., & Steiti, A. (2024). Random forest algorithm overview. Deleted Journal, 69–79. https://doi.org/10.58496/bjml/2024/007

[25] Roy, P. K., Chowdhary, S. S., & Bhatia, R. (2020). A machine learning approach for automation of resume recommendation system. Procedia Computer Science, 167, 2318–2327. https://doi.org/10.1016/j.procs.2020.03.284

[26] Singh, N. (2024). Machine learning approaches to enhance candidate selection: A comparative study in HR recruitment. International Journal for Research in Applied Science and Engineering Technology, 12(7), 1286–1296. https://doi.org/10.22214/ijraset.2024.63745

[27] Zheng, Z., Hu, X., Qiu, Z., Cheng, Y., Gao, S., Song, Y., Zhu, H., & Xiong, H. (2024). Bilateral multi-behavior modeling for reciprocal recommendation in online recruitment. IEEE Transactions on Knowledge and Data Engineering, 36(11), 5681–5694. https://doi.org/10.1109/tkde.2024.3397705

[28] Zeng, S., Zhang, J., He, P., Xing, Y., Liu, Y., Xu, H., Ren, J., Wang, S., Yin, D., Chang, Y., & Tang, J. (2024). The good and the bad: Exploring privacy issues in retrieval-augmented generation (RAG). arXiv. https://doi.org/10.48550/arXiv.2402.16893

[29] Wu, X., Liu, K., Wang, J., Yao, J., Deng, B., Lv, R., & Song, J. (2024, November). Candidate evaluation with multimodal data-driven recruitment. In International Conference on Pattern Recognition (pp. 81–96). Cham: Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-78186-5_6

[30] Almeida, F., Junça Silva, A., Lopes, S. L., & Braz, I. (2025). Understanding recruiters' acceptance of artificial intelligence: Insights from the Technology Acceptance Model. Applied Sciences, 15(2), 746. https://doi.org/10.3390/app15020746

[31] Jatobá, M. N., Ferreira, J. J., Fernandes, P. O., & Teixeira, J. P. (2023). Intelligent human resources for the adoption of artificial intelligence: A systematic literature review. Journal of Organizational Change Management, 36(7). https://doi.org/10.1108/jocm-03-2022-0075

[32] Qin, C., Zhang, L., Cheng, Y., Zha, R., Shen, D., Zhang, Q., Chen, X., Sun, Y., Zhu, C., Zhu, H., & Xiong, H. (2023). A comprehensive survey of artificial intelligence techniques for talent analytics. arXiv. https://arxiv.org/abs/2307.03195

[33] Alshehhi, K., Cheaitou, A., & Rashid, H. (2024). Procurement of artificial intelligence systems in UAE public sectors: An interpretive structural modeling of critical success factors. Sustainability, 16, 7724. https://doi.org/10.3390/su16177724

[34] Cameron, R., Herrmann, H., & Nankervis, A. (2024). Mapping the evolution of algorithmic HRM (AHRM): A multidisciplinary synthesis. Humanities and Social Sciences Communications, 11, 303. https://doi.org/10.1057/s41599-024-02786-z

[35] Darville, J. J., & Arghode, V. (2024). People analytics, talent management, and leadership development in the USA and UAE. In Industry 4.0 and people analytics (pp. 263–285). Apple Academic Press. https://doi.org/10.1201/9781003414193-13

[36] GoHire. (n.d.). Top 12 job boards in the UAE. GoHire. https://gohire.io/blog/top-12-job-boards-in-the-uae

[37] Raghavan, M., Barocas, S., Kleinberg, J., & Levy, K. (2020). Mitigating bias in algorithmic hiring: Evaluating claims and practices. In Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency (pp. 469–481). ACM. https://doi.org/10.1145/3351095.3

AI-Based Framework for Forecasting Defense Technologies in Defense Acquisition

Nadia Janoudi
Business Analytics Program
Abu Dhabi School of Management
Abu Dhabi, UAE
njnaoudi@gmail.com

Evi Indriasari Mansor
Business Analytics Program
Abu Dhabi School of Management
Abu Dhabi, UAE
e.mansor@adsm.ac.ae

Abstract— The quick pace of technological innovation and evolving global threats pose challenges to defense acquisition. The traditional forecasting methods lack the agility to support the forecasting of rapid technology developments. This research presents a conceptual AI-driven framework developed to enhance defense technologies forecasting, thereby improving defense acquisition processes. The conceptual framework integrates open-source (OSINT) and internal data with artificial intelligence (AI) techniques, including machine learning, natural language processing (NLP), and unsupervised clustering, the framework identifies emerging technologies, evaluates Technology Readiness Levels (TRLs), and visualizes insights, developed through literature review and validated by defense and AI experts, the framework supports scenario-based forecasting and human-in-the-loop. This research aims to contribute an adaptive model capable of enhancing technology forecasting, reducing acquisition risks and costs, and align future technological needs with threats.

Keywords—Defense Acquisition, Technology Forecasting, Artificial Intelligence, Big Data, Machine Learning, Natural Language Processing, Technology Readiness Levels, OSINT, Predictive Analytics

I. INTRODUCTION

Defense acquisition is a complex and lengthy process, often challenged by inefficiencies and the risk of acquiring costly or outdated technologies [1]. Artificial Intelligence (AI) and Big Data can play a major role in improving this process by offering better insights, enhanced forecasting, and integrating different data sources [2]. It is important to note that defense acquisition involves much more than procurement, including identifying needs, planning, development, testing, deployment, and sustainment. Acquisition programs aim to deliver new or improved military capabilities, covering weapons, IT systems, and support services essential for national defense.

Defense acquisition struggles to keep up with rapid technological change due to slow data analysis, fragmented systems, and outdated forecasting methods. There's a lack of AI-driven frameworks tailored to defense needs, leading to poor predictions, costly mistakes, and reduced readiness [3-4]. Integrating AI and Big Data is a must through ensuring performance, transparency, and proper model use.

This research is driven by the need to use AI and Big Data to improve defense technology forecasting. Early and continuous integration of AI offers real-time insights into emerging technologies, helping decision-makers plan, detect outdated capabilities, and assess readiness. AI tools like predictive analytics and clustering support better decision-making under pressure and improve situational awareness by merging data from different sources. The developed framework aims to enhance acquisition outcomes and operational readiness. It contributes to academic knowledge and helps policymakers align technology planning with future defense needs [4-5].

The aim of the research is to develop a conceptual AI-driven framework for forecasting defense technologies using Big Data. It explores challenges in the acquisition process, designs and validates the framework with experts' input, and addresses ethical and social considerations. The paper reports the relevant literature review, the methodology used in the research, data analysis, findings, discussion and conclusion.

II. LITERATURE **REVIEW**

The reviewed literature confirms the growing importance of AI and Big Data in enhancing defense technology forecasting. Studies like Agrawal et al. [5] and Almahmoud et al. [6] show how AI supports better decision-making and cyber threat prediction. Autio et al. [7] and Cummings [8] emphasize ethical and technical limitations, while others like Ebadi et al. [9] and De Spiegeleire et al. [10] demonstrate AI's use in detecting emerging technologies and improving strategic agility, especially in smaller states. Defense-focused studies like GAO [4], Kott & Perconti [3], and Morgan et al. [11] stress the need for structured AI frameworks and policies in defense acquisition. Research by Kania [12], and MIT AI Accelerator [13] highlights global competition and the urgency of adopting AI across operations. RAND's reports [14] and frameworks offer insights on technology evaluation and innovation planning. The literature supports that AI-driven forecasting tools can offer valuable insights into defense acquisition, but success depends on proper integration, data readiness, human oversight, and policy alignment.

III. METHODOLOGY

The research methodology implemented in this work was intended to explore how AI and Big Data can be integrated into defense technology forecasting. Ethical approval was obtained from the ADSM ethical committee, ensuring the research is aligned with academic and ethical standards. The process started with a literature review to understand current challenges in defense acquisition and how AI can support in overcoming them. Based on the insights gathered from the literature review only, the initial conceptual framework was developed.

Afterwards, the framework was validated by four experts in the area of defense and AI. Consent from participants was obtained as part of the requirement of the research. Due to the sensitivity of the topic, it is important and ethically required to declare that the research examines only public information as published reports and articles and maintains ethical standards of handling expert insights. This research did not involve any primary data collection. The analysis presented in the Data Analysis section is based on existing literature. The validation input from experts is handled with strict confidentiality with protection, and limiting access to authorized researchers only.

Industry experts were asked whether the framework addresses key pain points in defense acquisition and if the alliance and non alliance countries analysis of technology adoption is relevant and useful for strategic planning. They evaluated whether the dashboard provides enough detail to inform acquisition decisions, and how well the alliance technology analysis supports collaboration and co-development. Experts also gave feedback on the reliability of the data sources used, whether the framework captures emerging disruptive technologies or leaves gaps, and what additional data or features could improve the dashboard's usefulness. Questions included how often the dashboard should be updated to reflect real-time trends, whether it helps anticipate future needs and align acquisition accordingly, and if they would prefer using this dashboard over current tools—and why.

The interview questions for AI experts focused on evaluating the technical capabilities of the proposed framework. Experts were asked whether the AI models could identify trends, competitor technologies, and alliance capabilities, and if the natural language processing (NLP) components were effective in extracting insights from unstructured data like reports and articles. They assessed whether the time series forecasting models were suitable for predicting long-term technology trends and if the system effectively integrates multiple data sources to provide unified and accurate forecasts. Questions also covered the usefulness of clustering and predictive analytics, the framework's ability to detect emerging technologies and historical patterns, and whether data integration and preprocessing were sufficient to ensure accurate dashboard outputs. Experts were asked to comment on the framework's scalability, potential technical limitations, the role of feedback loops in improving prediction accuracy, and suggestions for adaptive learning mechanisms or overall improvements.

Experts recommendations were later used to finalize the conceptual framework to address gaps and improve functionality that reflect real-world industry needs.

IV. DATA ANALYSIS

This section focuses on analyzing the different tools and components used to build the AI-driven defense technology forecasting framework through existing literature exploring and justifying the use of tools in each layer of the framework. due to the The research highlights the importance of extracting data from different sources. Publicly available sources like OSINT, news, and defense reports, as well as internal datasets, can be extracted using a combination of web scraping, NLP tools, and APIs [15]. LLMs like GPT and Claude were also mentioned to help extract key insights from research and technical documents [16]. For storage, a hybrid model combining cloud and physical storage [17]. Tools like Apache Kafka and Flume to ingest large volumes of data, while PostgreSQL and MongoDB handled structured and unstructured information securely [18-19].

Data preparation involved cleaning and structuring data using Pandas, spaCy, and NLTK to remove noise and standardize formats. NLP tasks like NER, tokenization, and topic modeling to identify defense technologies, organizations, and geopolitical players [20]. For the analytics layer, AI tools to detect trends and forecast technologies. This included clustering (K-Means), topic modeling (LDA), and forecasting methods like ARIMA, Prophet, and LSTM [21]. Tools like Monte Carlo and Bayesian Networks helped assess risk and uncertainty, while classification models like BERT and XGBoost can support TRL prediction [22]. The visualization layer can use tools like Tableau and Power BI to present findings, heatmaps, graphs, and dashboards, helping decision-makers understand patterns and risks. Tools like Gephi and D3.js allowed relationship mapping across technologies and actors [23].

V. THE PROPOSED FRAMEWORK

According to the information gathered from the requirement analysis phase, the initial framework was designed as a structured, AI-driven system to support defense acquisition by forecasting technologies through a multi-layered architecture. It included key layers: data sources, extraction, data preparation, storage, analysis, forecasting, and visualization. It relied on integrating OSINT and classified defense data, using AI techniques like NLP, clustering, and trend detection to identify and visualize emerging technologies.

After evaluations by AI experts, the framework also introduced classification of internal data, added missing data handling and anomaly detection, and incorporated Technology Readiness Level (TRL) standardization. Human-in-the-loop mechanisms, physical storage only, graph-based mapping, predictive feature engineering, and scenario simulations were also included to improve accuracy, interpretation, and adaptability (Fig. 1).

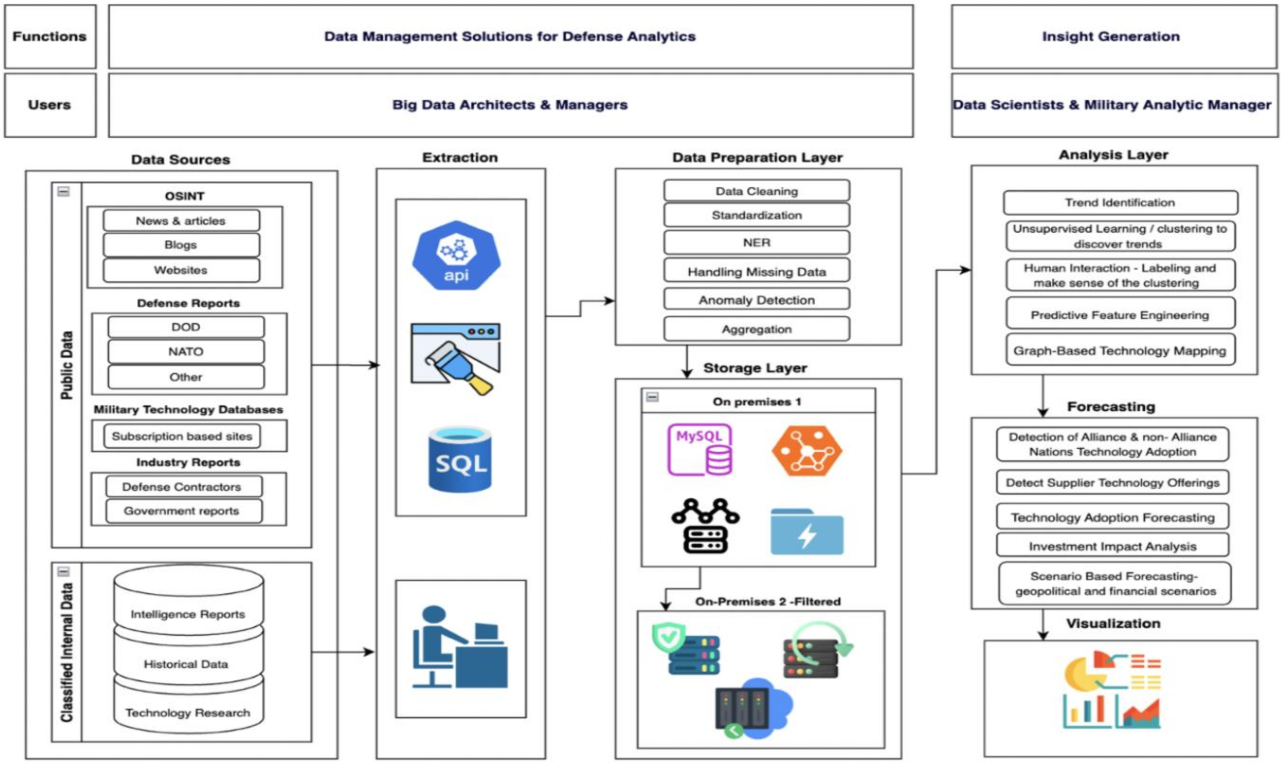


Fig. 1. AI-based framework for forecasting defense technologies

VI. DISCUSSION

From the validation sessions conducted with AI and industry experts, they agreed that the framework effectively addresses key challenges in defense acquisition and effectively forecasts technologies using AI and Big Data. They confirmed the framework's strength in identifying trends, competitor technologies, and alliance capabilities, supported by NLP, time series forecasting, and data integration from OSINT and classified sources. Experts emphasized the importance of human-in-the-loop mechanisms to guide AI outputs and reduce reliance on automation. Recommended improvements included semi-supervised learning, anomaly detection, fairness tools, continuous model retraining, and incorporating Technology Readiness Levels (TRLs). While the framework is scalable and adaptive, ongoing challenges include data integrity, classification limits, and bureaucratic barriers.

Compared to traditional models, this approach is more proactive—leveraging real-time analytics, unstructured data, and geopolitical context to anticipate emerging threats and guide acquisition planning more effectively.

A. Limitations

Several limitations were noted during validation. The framework heavily relies on OSINT and public sources. The use of unsupervised learning helps identify patterns but lacks clarity without expert input. To improve accuracy, semi-supervised and active learning, bringing defense experts into the loop, are recommended. The framework also does not yet adapt to fast-changing geopolitical shifts, which affect technology needs. Integrating AI-based scenario simulations and conflict modeling would address this. Additional enhancements include business intelligence features like contractor performance tracking, cybersecurity compliance, and supplier risk scoring. While anomaly detection is in place for procurement and R&D trends, it could be strengthened through federated learning, enabling secure, decentralized data sharing and improved scalability.

B. Future Work

The framework offers strong potential for improving defense acquisition, alliance planning, and risk assessment. Due to current limitations in the framework, future work could focus on adding semi-supervised learning, AI-driven scenario planning, and enhanced dashboards with BI features. Additional research could also focus on:

- Agentic framework for more autonomous AI forecasting
- Integration of classified and real-time intelligence feeds
- Multimodal AI combining text, imagery, sensors, & video
- Federated adaptive learning for secure, ongoing updates
- AI-powered scenario simulations to support forecasting
- Ethical frameworks and governance for AI in defense
- Expanded Human interference

VII. CONCLUSION

This research introduced a conceptual AI-driven framework to improve defense technology forecasting by integrating AI, Big Data, and expert input. The conceptual framework enhances acquisition planning by combining structured and unstructured data, using tools like NLP, clustering, TRL assessments, and predictive analytics. Expert validation confirmed the importance of human oversight for expert oversight and continued learning, semi-supervised learning, and ethical AI use. Policy alignment of AI use with national security laws, ensuring compliance with international arms control agreements, preventing bias or unfair advantages to certain contractors, data security, and geopolitical forecasting are essential for real-world application. Future work should explore real-time data integration, Agentic framework, federated learning, and immersive AR/VR dashboards. The framework sets a foundation for smarter, adaptive forecasting in defense acquisition.

REFERENCES

[1] Rieksts, B. Q., & Guerrera, K. M. A feasibility study on the use of artificial intelligence for defense acquisition program review, Volume I: Main Report. Institute for Defense Analyses. 2020. [Online]. Available: https://www.ida.org/-/media/feature/publications/a/af/a-feasibility-study-on-the-use-of-ai-for-defense-acquisition-program-review-volume-1-main-report/p-13239_vol1.ashx

[2] Barlow, C., Forbes, K., Giachinta, R., Levenson, Z., Novak, R., Raines, J., & Roe, S. Enhancing acquisition outcomes through leveraging artificial intelligence. MITRE Corporation (PR-24-0962-Leveraging-AI-in-Acquisition), 2024, pp. ii–26. [Online]. Available: https://www.mitre.org/sites/default/files/2025-03/PR-24-0962-Leveraging-AI-Acquisition.pdf

[3] Kott, A., & Perconti, P. Long-term forecasts of military technologies for a 20–30-year horizon: An empirical assessment of accuracy. Defense Technology Journal, Vol 7, Issue 3, pp. 112–134, 2020. DOI:10.48550/arXiv.1807.08339

[4] United States Government Accountability Office (GAO). Artificial intelligence: Status of developing and acquiring capabilities for weapon systems, pp. 10–25, 2022. [Online] Available: https://www.gao.gov/products/gao-22-104765

[5] Agrawal, A., Gans, J. S., & Goldfarb, A. Economic perspective on artificial intelligence as a prediction technology. Journal of Economic Perspectives, Vol 33, Issue 2, pp. 31–50, 2019. DOI: 10.1257/jep.33.2.31

[6] Almahmoud, Z., Yoo, P. D., & Alhussein, O. A holistic and proactive approach to forecasting cyber threats. Scientific Reports, Vol 13, Article 7965, 2023. DOI:10.1038/s41598-023-35198-1

[7] Autio, T., Jantunen, E., & Koskinen, H. AI in predicting technological trends for military operations. Journal of Emerging Defense Technologies, Vol 12, Issue 4, pp. 87–105, 2023. https://files.thegovlab.org/a-snapshot-of-ai-procurement-challenges-june2023.pdf

[8] Cummings, M. Artificial intelligence and the future of warfare. Chatham House. 2017. [Online] Available: https://www.chathamhouse.org/sites/default/files/publications/research/2017-01-26-artificial-intelligence-future-warfare-cummings.pdf

[9] Ebadi, A., Auger, A., & Gauthier, Y. Detecting emerging technologies and their evolution using deep learning and weak signal analysis, pp. 1–17, 2025. DOI:10.48550/arXiv.2205.05449

[10] De Spiegeleire, S., Maas, M., & Sweijs, T. Artificial intelligence and the future of defense: Strategic implications for small- and medium-sized force providers. The Hague Centre for Strategic Studies. 2017. [Online] Available: https://hcss.nl/report/artificial-intelligence-and-the-future-of-defense/

[11] Morgan, F. E., Boudreaux, B., Lohn, A. J., Ashby, M., Curriden, C., Klima, K., & Grossman, D. AI and national security: Policy considerations and technological advancements. RAND Corporation. 20. [Online] Available: https://www.rand.org/pubs/research_reports/RR4280.html

[12] Kania, E. B. AI weapons and the implications for future warfare. Brookings Institution. 2020. [Online] Available: https://www.brookings.edu/wp-content/uploads/2020/04/FP_20200427_ai_weapons_kania_v2.pdf

[13] MIT AI Accelerator. Harnessing artificial intelligence for defense. Massachusetts Institute of Technology. 2022. [Online] Available: https://aia.mit.edu/wp-content/uploads/2022/02/AI-Acquisition-Guidebook_CAO-14-Feb-2022.pdf

[14] Wong, J. P., Younossi, O., LaCoste, C. K., Anton, P. S., Vick, A. J., Weichenberg, G., & Whitmore, T. C. Improving Defense Acquisition: Insights from Three Decades of RAND Research. RAND Corporation. 2022. [Online] Available: https://www.rand.org/pubs/research_reports/RRA1670-1.html

[15] Berghel, H. Robert David Steele on OSINT. Computer, Vol 47, Issue 7, pp. 76-81, 2014. DOI:10.1109/MC.2014.191

[16] Gartlehner, G., Kahwati, L., Hilscher, R., Thomas, I., Kugley, S., Crotty, K. & Chew, R. Data extraction for evidence synthesis using a large language model: A proof-of-concept study. Research synthesis methods, Vol 15, issue 4, pp. 576-589, 2024. DOI:10.1002/jrsm.1710

[17] Sohal, M., Bharany, S., Sharma, S., Maashi, M. S., & Aljebreen, M. A hybrid multi-cloud framework using the IBBE key management system for securing data storage. Sustainability, Vol 14, Issue 20, pp. 13561, 2022. DOI:10.3390/su142013561

[18] Rani, S. Tools and techniques for real-time data processing: A review. International Journal of Science and Research Archive, Vol 14, Issue 1, pp. 1872-1881, 2025. DOI:10.30574/ijsra.2025.14.1.0252

[19] Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. BERT: Pre-training of deep bidirectional transformers for language understanding. Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Vol 1, pp. 4171–4186, 2019. DOI:10.18653/v1/N19-1423

[20] Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L., & Stoyanov, V. (2019). RoBERTa: A robustly optimized BERT pretraining approach. arXiv, pp. 1-13, 2019. DOI: 10.48550/arXiv.1907.11692

[21] Jain, A. K. Data clustering: 50 years beyond K-means. Pattern Recognition Letters, Vol 31, Issue 8, pp. 651–666, 2010 DOI:10.1016/j.patrec.2009.09.011

[22] Chen, T., & Guestrin, C. XGBoost: A scalable tree boosting system. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 785–794, 2016. DOI:10.1145/2939672.2939785

[23] Newman, M. E. J. Networks: An introduction. Oxford University Press. 2010. [Online] Available: https://www.cs.cornell.edu/home/kleinber/networks-book/networks-book-ch03.pdf

The Integration Of Generative Ai Into Computer Programming Education: Students Perspective

Khuloud Alhamedi
College of Technological Innovation
Zayed University
Abu Dhabi, UAE
M80008946@zu.ac.ae

Ashraf Khalil
College of Technological Innovation
Zayed University
Abu Dhabi, UAE
Ashraf.Khalil@zu.ac.ae

Mohammad Kuhail
College of Technological Innovation
Zayed University
Abu Dhabi, UAE
Mohammad.Kuhail@zu.ac.ae

Rafiq Hijazi
College of Natural and Health Sciences
Zayed University
Abu Dhabi, UAE
Rafiq.Hijazi@zu.ac.ae

Abstract—Due to the growing number of open-source generative AI models and the development of models with advanced capabilities, generative AI has become more accessible and is being utilized in various fields, including education. This research aims to identify the factors influencing higher education students' intention to use generative AI tools in computer programming courses by extending the Technology Acceptance Model (TAM) with a quantitative approach. The data was collected by distributing a survey among students from different universities, and 205 responses were received. The findings indicate that key constructs such as perceived usefulness, trust in the tool, perceived ethics, and attitude toward AI significantly influence the students' usage intention.

Keywords—Generative AI, AI in Education, TAM, Computer Programming, Higher Education, Students' Generative AI Usage Intention

I. INTRODUCTION

Recently, many major tech companies, such as Google, OpenAI, and X, have been competing to develop the best generative AI model. According to [1] 65.7% of the newly released models in 2023 were open source compared to 44.4% in 2022. This indicates an increase of 21.3% within a period of one year, with an increase in Generative AI investment by almost 800% in 2023, reaching $25.2 billion. As a result, generative AI tools have become more accessible and are providing enhanced capabilities to various users, including students.

Moreover, this growing accessibility has led to a disruption in the traditional process of teaching programming courses. Educators and researchers believe that the use of generative AI in programming courses can transform such courses, which are usually known for their high student-to-teacher ratios, time-consuming grading processes, and lack of personalized feedback [2], [3]. Therefore, recent studies have begun exploring different strategies for integrating generative AI tools in programming education, focusing on the associated benefits and challenges these technologies will bring to the learning environment.

Although recent studies are exploring the capabilities of generative AI tools and how they can be used to enhance programming education, there is still limited research covering the Gulf region in general and the United Arab Emirates (UAE) specifically. According to [4], the existing research on AI in higher education in the GCC is still developing, and studies that focus on the use of generative AI in programming education are limited.

These regional gaps are significant in the UAE, where the integration of AI education is one of the main parts of the UAE Artificial Intelligence Strategy 2031, which aims to position the country as one of the global leaders in AI development and integration [5]. As part of this vision, the UAE has invested in AI infrastructure, innovation, and education, with a strong focus on ensuring that the graduates are future-ready. This can be seen in the country's high STEM engagement, where 36.2% of higher education graduates specialize in areas related to science, technology, engineering, and mathematics [6].

The study seeks to answer the following questions:

RQ1.What are the key factors influencing students' intention to adopt generative AI tools in programming education within UAE higher education institutions?

RQ1.1.How does students' perceived usefulness and ease of use of generative AI tools in programming tasks impact intention to use them?

RQ1.2.How do trust and perceived ethics affect the students' intention to use generative AI tools?

RQ1.3.How do social influence and students' attitudes toward AI affect their acceptance of these tools in programming education?

RQ2.What are the associated benefits and challenges of integrating generative AI tools in educational settings?

RQ3.What ethical considerations should educational institutions take into consideration when integrating generative AI in UAE programming education?

This study aims to address these gaps by focusing on exploring the factors which influence the students' intention to use generative AI tools such as ChatGPT, GitHub Copilot, and Gemini in computer programming courses by extending the Technology Acceptance Model (TAM). The extended model includes additional constructs such as trust in the tool, perceived ethics, social influence, and attitude toward AI. By focusing on higher education students in the UAE, this study aims to provide insights that will support more effective and ethical integration of generative AI tools in programming education..

II. METHODOLOGY

A. Research Design and Approach

This study explores the factors influencing students' intention to use Generative AI tools in computer programming courses by extending the Technology Acceptance Model (TAM), following an explanatory research design. Based on the survey responses, the direct and indirect relationships between the model constructs will be examined using Structural Equation Modeling (SEM-PLS), while SPSS will be used to conduct the descriptive analysis. SEM-PLS was selected due to its suitability for exploratory studies and how it allows the testing of complex models with latent variables.

Furthermore, the study adopts a deductive approach, beginning with an established theoretical framework (TAM) and extending it by testing predefined hypotheses drawn from existing literature and empirical observations. This approach allows the evaluation of the validity and reliability of the proposed extended TAM model.

B. Ethical Consideration

To ensure the ethical integrity of this study, an ethical clearance was obtained from Zayed University's Research Ethics Committee.

Adhering to this procedure ensures that data confidentiality and participant anonymity are maintained. Additionally, it guarantees that participants read and agree to the provided consent form before participating in the survey, that all collected data are securely stored and accessed only by the research team, and that the participants are fully aware of their right to withdraw at any time.

These measures align with the research ethics guidelines, ensuring that the thesis upholds transparency, confidentiality, and voluntary participation as participants are fully informed of their right to withdraw at any time..

C. Data Sources and Participants

This study employs a quantitative research methodology, as it examines the factors influencing higher education students' intention to use Generative AI in computer programming courses. A structured survey was created using Google Forms and used as the primary data collection method to maximize the survey's accessibility and convenience.

The target participants for this study are higher education students enrolled in computer programming courses in the UAE. A non-probability sampling technique, specifically purposive sampling, was used to ensure that only students with relevant academic backgrounds participated. The survey was distributed through the university email, and student WhatsApp and Telegram groups. The data was collected over a period of four weeks from the end of February 2025 to the end of March 2025.

The survey started by providing the participants with a brief overview of the study purpose and a consent form before accessing the survey and participants were informed that their responses would remain anonymous and confidential.

D. Instrument and Pilot Testing

The survey was divided into three main sections: Demographic information, AI Tool usage, and the main part. Before starting the actual data collection process, a pilot test was conducted. The survey link was shared only with family members and friends who met the criteria of the target participants. In total, ten students participated in the pilot testing. The aim of this test was to ensure the clarity, consistency, and reliability of the developed research instrument by collecting feedback from participants and reflecting it in the final version of the survey. Additionally, pilot testing helped identify any technical issues and ensured that the responses are received and stored securely.

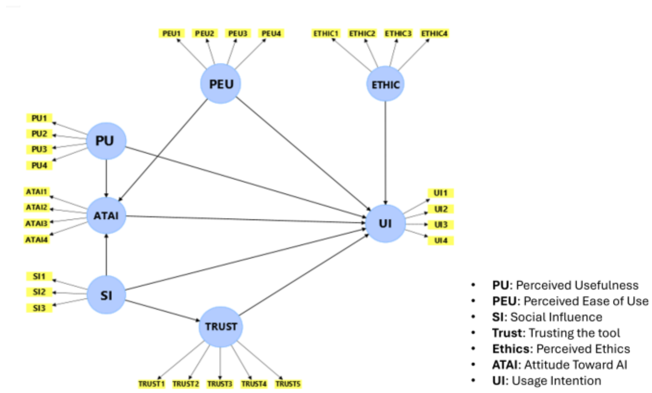


Figure 12. Proposed Model

III. RESULTS

In order to understand the factors influencing the student's usage intention of generative AI tools, the TAM model was extended. The final model contained seven constructs: perceived usefulness (PU), perceived ease of use (PEU), social influence (SI), trust in AI tools (TRUST), perceived ethics (ETHIC), attitude towards AI (ATAI), and usage intention (UI). Ten hypotheses were tested, covering the impact of these variables on students' intention to use GAI tools, their attitude toward AI, and the influence of social influence on trust.

A survey was developed based on the literature, with multiple items measuring each construct. The survey was distributed among higher education students in the UAE, and a total of 205 responses were received. After data cleaning, five responses were removed due to the lack of consent, incorrect university information, or being from a university outside the UAE.

SPSS was used to perform descriptive analysis , providing an overview of the sample characteristics, while SmartPLS was used to test the hypotheses using both PLS-SEM and bootstrapping algorithms.

After completing the data analysis, seven out of ten hypotheses were supported. Hypotheses related to the correlation between perceived usefulness, trust in the tool, perceived ethics, and attitude toward AI and the student's usage intention were supported. In contrast, an insignificant correlation was found between social influence and perceived ease of use with usage intention. Additionally, social influence was found to have an indirect impact by shaping students' attitudes and trust toward AI.

These results indicate that higher education students prioritize the tool's value, reliability, accuracy, and ethical alignment over its ease of use. This was an unexpected result, as most literature reported a significant correlation between PEU and the intention to use generative AI tools.

A. Theoretical Implications

This study contributes to the theoretical advancement of the Technology Acceptance Model (TAM) by extending its constructs to include constructs related to the education context focusing mainly on the integration of generative AI in computer programming education. Based on the analysis result, constructs such as trust, ethics, social influence, and attitude towards AI are significant extensions to the original model.

The findings from hypotheses eight and nine concluded a significant influence between ethics and trust on the student's intention to use generative AI, emphasizing the importance of including constructs related to values such as trust in models involving AI. This is because artificial intelligence tools introduce dimensions such as model bias, accuracy, and transparency.

Moreover, the strong correlation between attitude toward AI and usage intentions shows the importance of ATAI as a mediating variable in the TAM model. Surprisingly, perceived ease of use (PEU) had an insignificant influence on the students' usage intention, which is the opposite of the expected results reported in the literature. This suggests that digitally mature users value trust, ethics, and usefulness more than ease of use.

B. Practical Implications

The results of this study provide insights that can support generative AI developers, curriculum developers, and instructors. Generative AI tool developers can benefit from knowing the key variables influencing the student's usage intention. For example, since the results show that students value trust and usefulness, developers should enhance the transparency, reliability, and explainability of the tools they build.

Curriculum developers will benefit from knowing how students used those tools in computer programming and who depend on those tools more, whether it is students in earlier academic years or later. This will help in understanding the best time to integrate the tools in programming education and in which tasks it can be integrated without affecting the students' learning outcomes negatively. Instructors will also get useful insights on how students are utilizing generative AI tools.

IV. CONCLUSION

In conclusion, this study aims to explore the factors influencing the intention to use generative AI in higher education, focusing on the students' perspective. Three main research questions were the focus of this study, where some questions focus on understanding the current challenges, benefits, and the ethical concerns of integrating such tools in higher education by reviewing current literature. While other research questions focus on understanding the main factors that affect the usage intention of generative AI tools in computer programming courses among higher education students in the UAE.

To understand the factors influencing the student's usage intention, the TAM model was extended and the final model contained seven constructs, including perceived usefulness, perceived ease of use, social influence, trusting AI tools, perceived ethics, attitude towards AI, and usage intention. Ten hypotheses were tested, and the results showed that perceived usefulness, trust in the tool, perceived ethics, and attitude toward AI had a significant impact on the students' usage intention. In contrast, an insignificant correlation was found between social influence and perceived ease of use with usage intention. Additionally, social influence had an indirect impact by shaping students' attitudes and trust toward AI. This indicates that higher education students prioritize the tool's value, reliability, accuracy, and ethical alignment over its ease of use (refer to Table 1 and Table 2).

Future studies should include participants from different academic majors, which will allow the researchers to assess whether students in different fields have varying levels of acceptance, trust, and ethical concern regarding the integration of generative AI.

Moreover, adding extra methods to collect the data, such as experiments and interviews, will help in providing additional insights. Conducting controlled experiments in different universities across the UAE could help in determining whether different teaching methods or institutional cultures influence the students' usage intention.

Since the majority of the participants in the survey were students in later academic years (third year or more), it is important to also investigate the ideal stage within a student's academic journey at which AI tools should be introduced. This will help in understanding whether early integration of generative AI tools in foundational programming courses results to better outcomes, or if generative AI tools are best suited for use in more advanced, project-based learning environments.

Finaly, future research should also explore the instructors' perspectives, as their role is critical in shaping the classroom norms and influencing the students' trust in generative AI technologies. Also, understanding the instructors' attitudes toward generative AI tools, their challenges in implementing such tools, and their ethical concerns can provide a more holistic view of the topic.

TABLE 1. PATH COEFFICIENTS

Path	T-statistics	P-values
ATAI -> UI	8.8	0
ETHIC -> UI	1.998	0.046
PEU -> ATAI	1.69	0.091
PEU -> UI	0.334	0.738
PU -> ATAI	5.312	0

Path	T-statistics	P-values
PU -> UI	1.962	0.05
SI -> ATAI	3.404	0.001
SI -> TRUST	12.827	0
SI -> UI	0.131	0.896
TRUST -> UI	2.988	0.003

TABLE 2. HYPOTHESES DECISION

hypotheses	Path	Decision
H1. Perceived Usefulness positively influences the students' intention to use generative AI tools in computer programming education.	PU -> UI	Supported
H2. Perceived Usefulness positively influences the students' Attitude towards AI.	PU -> ATAI	Supported
H3. Perceived Ease of Use positively influences the students' intention to use generative AI tools in programming education.	PEU -> UI	Rejected
sH4. Perceived Ease of Use has a significant influence on the students' Attitude towards AI.	PEU -> ATAI	Rejected
H5. Social Influence significantly influences the students' intention to use generative AI tools in programming education.	SI -> UI	Rejected
H6. Social Influence positively influences the students' attitude towards AI in programming education.	SI -> ATAI	Supported
H7. Social Influence positively influences the students' trust in AI in programming education.	SI -> TRUST	Supported
H8. Trust positively influences the students' intention to use generative AI tools in programming education.	TRUST -> UI	Supported
H9. perceived ethics positively influence the students' intention to use generative AI tools in programming education.	ETHIC -> UI	Supported
H10. A positive Attitude Toward AI positively influences the students' intention to use generative AI tools in programming education.	ATAI -> UI	Supported

REFERENCES

[1] [N. Maslej, L. Fattorini, R. Perrault, V. Parli, R. Reuel, and E. Brynjolfsson, "The AI Index 2024 Annual Report," Stanford HAI, Stanford University, 2024. [Online]. Available: https://aiindex.stanford.edu

[2] [M. Abolnejadian, S. Alipour, and K. Taeb, "Leveraging ChatGPT for Adaptive Learning through Personalized Prompt-based Instruction: A CS1 Education Case Study," in Extended Abstracts of the CHI Conference on Human Factors in Computing Systems, Honolulu HI USA: ACM, May 2024, pp. 1–8. doi: 10.1145/3613905.3637148.

[3] M. Kazemitabaar et al., "CodeAid: Evaluating aClassroom Deployment of an LLM-based Programming Assistant that Balances Student and Educator Needs," in Proceedings of the CHI Conference on Human Factors in Computing Systems, Honolulu HI USA: ACM, May 2024, pp. 1–20. doi: 10.1145/3613904.3642773.

[4] F. K. Fadlelmula and S. M. Qadhi, "A systematicreview of research on artificial intelligence in higher education: Practice, gaps, and future directions in the GCC," J. Univ. Teach. Learn. Pract., vol. 21, no. 06, Apr. 2024, doi: 10.53761/pswgbw82.

[5] United Arab Emirates Minister of State For Artificial Intelligence Office, "UAE National Artificial Intelligence Strategy 2031," 2018. [Online]. Available: https://ai.gov.ae/wp-content/uploads/2021/07/UAE-National-Strategy-for-Artificial-Intelligence-2031.pdf

[6] K. Buchholz, "Where Students Choose STEM Degrees," Statista Daily Data. Accessed: Apr. 06, 2025. [Online]. Available: https://www.statista.com/chart/22927/share-and-total-number-of-stem-graduates-by-country

AI for Education and Education for AI: A Dual Strategy for Transforming Learning

Mohamed Akoum
Quality & Business Excellence
Abu Dhabi School of Management
Abu Dhabi, UAE
m.akoum@adsm.ac.ae
ORCID: 0000-0002-9354-5282

Sherin Ashraf Kalleparambel
Abu Dhabi, UAE
Sher.ashraf@gmail.com

Abstract— This paper examines AI for Education (AI4E), which enhances educational institutions' posture in terms of teaching and administration, and Education for AI (E4AI), which focuses on preparing students to understand, apply, and lead AI-driven business transformation. While their synergy is promising, critical research gaps remain in assessing long-term efficacy, ethical implications, and equitable implementation. Current studies lack empirical evidence on balancing institutional efficiency (AI4E) with learner-centered innovation (E4AI). Further research is essential to optimize this dual approach, ensuring AI's transformative potential is realized without exacerbating disparities. Addressing these gaps will strengthen policy, pedagogy, and workforce readiness in an AI-driven future.

Keywords—Education, AI, Strategy, Education for AI, AI for Education

I. INTRODUCTION

The rapid integration of Artificial Intelligence (AI) into education has led to the emergence of two transformative yet distinct paradigms: AI for Education (AI4E), which focuses on enhancing institutional efficiency, teaching methodologies, and administrative processes, and Education for AI (E4AI), which aims to equip students with the skills, knowledge, and ethical understanding needed to thrive in an AI-driven world. While both paradigms reshape the educational landscape, their implementation and interplay remain underexplored, leaving critical gaps in research and practice. A significant Research Gap exists in understanding these approaches' long-term efficacy, ethical implications, and equitable scalability. Current studies often examine AI4E and E4AI in isolation, neglecting their synergistic potential and the challenges of integrating them holistically. For instance, while AI4E demonstrates promise in personalizing learning and optimizing operations, its reliance on data-intensive systems raises concerns about privacy, bias, and accessibility. Conversely, E4AI initiatives prioritize workforce readiness and innovation but struggle to keep pace with the rapid evolution of AI technologies, often leaving educators and institutions without clear guidelines for curriculum development or ethical training

This paper contributes to bridging these gaps by:

1. Systematically analyzing the objectives, methodologies, and outcomes of AI4E and E4AI, drawing on case studies and empirical evidence to highlight their respective strengths and limitations.

2. Proposing a dual-strategy framework that harmonizes institutional transformation (AI4E) with learner empowerment (E4AI), ensuring alignment with both operational goals and societal needs.

3. Identifying actionable recommendations for policymakers, educators, and institutions to address disparities in access, mitigate ethical risks, and foster scalable, future-ready education models.

By addressing these challenges, the study aims to provide a roadmap for leveraging AI's transformative potential while safeguarding equity and inclusivity. The findings underscore the urgency of adopting an integrated approach, where AI4E and E4AI coexist to not only enhance educational systems but also prepare a generation of AI-savvy leaders capable of navigating the complexities of the digital age.

II. METHODOLOGY

This study employs a systematic literature review and case study analysis to examine the dual paradigms of AI for Education (AI4E) and Education for AI (E4AI). The methodology is designed to address key research questions while ensuring rigor in literature selection and analysis.

Research Questions:

1. Operational Impact: How does AI4E enhance institutional efficiency, teaching methodologies, and administrative processes in education?
2. Learner Outcomes: What are the measurable effects of E4AI on student AI literacy, workforce readiness, and ethical understanding?

3. Synergies and Gaps: How can AI4E and E4AI be integrated to optimize both institutional and learner outcomes?
4. Ethical and Equity Challenges: What risks (e.g., bias, privacy, accessibility) arise from these paradigms, and how can they be mitigated?

The Analysis Approach that is used in this essay comprises 1. Thematic Coding which consists of Identifying recurring themes (e.g., "personalized learning," "AI ethics") across AI4E and E4AI literature; 2. Comparative Analysis which consists of depicting Contrasted institutional outcomes (AI4E) with learner-centric metrics (E4AI); 3. Case Study Synthesis to witness evaluated implementation challenges and successes in real-world settings.

This methodology ensures a comprehensive, evidence-based exploration of AI's dual role in education while highlighting actionable insights for stakeholders.

AI FOR EDUCATION: A DEEP DIVE

AI for Education is fundamentally about institutional enhancement through technology. It seeks to apply artificial intelligence across various educational management and delivery facets. Academic institutions can leverage AI to refine strategic planning, continuously adapting to market shifts and emerging trends. Market intelligence powered by AI enables academic institutions to understand evolving student interests and labor market demands, facilitating the timely updating and creation of programs. Moreover, AI introduces unprecedented agility and responsiveness into institutional operations, allowing real-time decision-making and dynamic resource allocation [1].

In the context of program and course development, AI technologies allow for rapid curriculum adjustments based on predictive analytics and skills forecasting. Personalized teaching emerges as a critical frontier, where intelligent tutoring systems and AI-driven learning paths cater to the unique needs of each learner, thereby significantly enhancing the educational experience. AI facilitates more accurate and adaptive assessments in student evaluation, moving beyond traditional testing methods to include behavioral analytics and personalized feedback mechanisms.

By positioning AI as an enabler, institutions can deliver more personalized, efficient, and high-quality education, aligning closely with the needs of a fast-changing global economy.

Deploying AI within educational institutions offers numerous strategic benefits. Institutions can achieve improved strategic alignment with market demands, foster faster and more dynamic curriculum development, and significantly enhance student satisfaction through personalized learning experiences

. Furthermore, AI-driven operations result in greater operational efficiency and smarter resource management. Institutions also gain valuable data-driven insights, supporting informed decision-making processes. Collectively, these advantages enhance the institution's competitive positioning, both locally and globally.

A. Enhancing Institutional Strategy:

Educational institutions are increasingly adopting AI to refine their strategic planning. AI-powered analytics can process vast amounts of data on student enrollment, performance, and engagement to identify trends and predict future needs. For example, institutions like Georgia State University have implemented AI-driven advising systems that analyze student data to provide personalized recommendations, significantly improving retention rates [2]

B. Understanding Market and Student Needs:

AI enables institutions to better align their offerings with industry demands. Natural Language Processing (NLP) tools can scan job postings, industry reports, and academic publications to identify emerging skills gaps. This allows universities to design curricula that meet the evolving needs of employers. A study by the World Economic Forum (2020) highlights how AI-driven labor market analysis can bridge the gap between education and employment [3].

C. Improving Teaching and Evaluation:

AI is revolutionizing pedagogy by enabling adaptive learning systems that tailor content to individual student needs. Platforms like Carnegie Learning and Knewton use machine learning algorithms to adjust lesson difficulty in real-time based on student performance. Additionally, AI-powered grading systems, such as those developed by Turnitin, reduce instructor workload while maintaining assessment accuracy [4].

Institutional Transformation:

AI For Education focuses on enhancing the educational ecosystem through technological augmentation. Its primary goals are:

1. Operational efficiency: Automating administrative tasks like admissions, scheduling, and grading.
2. Personalized learning: Using adaptive platforms to tailor content to individual student needs.
3. Data-driven decision-making: Leveraging predictive analytics for curriculum development and resource allocation.

Example: Arizona State University employs AI chatbots to handle student inquiries, reducing administrative workload by 30% [5]

EDUCATION FOR AI: A DEEP DIVE

Education for AI is a complementary but distinct endeavor. It focuses on preparing students to function effectively within and lead organizations increasingly shaped by AI. This requires a foundational understanding of AI technologies and their implications across industries. Promoting AI knowledge among students ensures they are aware of AI's transformative potential and its role in modern business environments [6].

Managerial AI literacy is paramount. Unlike technical AI education focused on programming and system design, managerial literacy emphasizes understanding AI's strategic applications, decision-support capabilities, and transformative impact on traditional management functions. Students must comprehend how AI reshapes planning, organizing, staffing, controlling, and leadership activities.

Furthermore, introducing students to AI-augmented tools such as predictive analytics, intelligent resource management, and process automation is essential. Building leadership competencies for an AI-driven world ensures that future managers can strategically integrate AI into business operations, driving innovation and maintaining competitive advantage.

By integrating Education for AI into its academic strategy, institutions can cultivate a new generation of AI-savvy graduates. These individuals will not only be highly employable in a competitive global market but will also contribute actively to their organizations' innovation and digital transformation initiatives. Furthermore, students educated in AI principles will be better equipped to navigate global economic trends, understand the ethical dimensions of AI deployment, and participate in governance and risk management activities related to AI technologies.

A. Promoting AI Literacy:

As AI becomes ubiquitous across industries, understanding its fundamentals is no longer optional. Universities are integrating AI literacy into non-technical disciplines, emphasizing ethical considerations, bias mitigation, and practical applications. Harvard’s “CS50 for Business Professionals” is an example of how AI education is being democratized for non-technical audiences (Achanta, 2023).

B. Developing AI-Savvy Leaders:

Management education is undergoing a transformation as AI reshapes traditional business functions. AI tools are now used for predictive analytics in strategic planning, automated recruitment in staffing, and sentiment analysis in leadership decision-making. A report by McKinsey (2019) underscores the importance of equipping future leaders with AI competencies to drive organizational innovation (Mishra, 2024).

C. Preparing Graduates for an AI-Driven Workforce:

The demand for AI-skilled professionals is growing exponentially. According to LinkedIn’s 2023 Emerging Jobs Report, AI and machine learning specialists top the list of fastest-growing occupations. Institutions must ensure their graduates are proficient in AI applications relevant to their fields, whether in healthcare, finance, or marketing [7] Click or tap here to enter text..

D. Workforce and Societal Readiness:

Education For AI is learner-centric, aiming to:

1. Develop AI literacy: Ensuring students understand AI fundamentals, ethics, and applications.
2. Prepare for AI-augmented careers: Equipping graduates with skills to work alongside AI in fields like healthcare, finance, and engineering.
3. Foster innovation: Encouraging students to develop AI-driven solutions to real-world problems.

Example: MIT’s "Machine Learning for Business" course trains non-technical students to apply AI in managerial roles [8].

III. A COMPARATIVE ANALYSIS

The distinction between AI for Education and Education for AI is clear yet complementary. AI for Education aims at enhancing the internal operations and educational delivery capabilities of institutions. It is inward-looking, focused on institutional excellence. Education for AI, on the other hand, is outward-looking, preparing students for the external realities of an AI-pervasive world.

In AI for Education, the target groups are primarily administrators and faculty members. The technological role is that of an enabler for improved educational services. In Education for AI, the primary beneficiaries are the students, with AI viewed as a core subject area that students must master to thrive in their future careers. Thus, while the immediate outcomes differ, the ultimate goal remains aligned: preparing both institutions and individuals for excellence in an AI-transformed society.

While both paradigms aim to harness AI’s potential, their distinctions lie in their primary beneficiaries and objectives.

A. Objective and Focus:

AI for Education is institution-focused, aiming to enhance administrative efficiency, teaching quality, and strategic decision-making. In contrast, Education for AI is student-centric, prioritizing skill development and workforce readiness.

B. Methodologies:

AI for Education relies on technologies like predictive analytics, NLP, and automated grading systems. Education for AI, however, emphasizes curriculum design, interdisciplinary courses, and hands-on training in AI tools.

C. Strategic Implication:

Institutions cannot use the same approach for both. A for Education optimizes systems, while Education for AI transforms learners. A one-size-fits-all strategy would neglect the unique demands of each.

D. Outcomes:

The success of AI for Education is measured by institutional metrics such as retention rates, operational costs, and student satisfaction. Education for AI’s success is gauged by graduate employability, industry relevance, and innovation capacity.

IV. CHALLENGES AND STRATEGIC RECOMMENDATIONS

Implementing these dual strategies is not without challenges. Faculty members must be equipped with the skills to teach and utilize AI tools effectively. Therefore, institutions should launch comprehensive internal AI literacy and upskilling programs [9].Click or tap here to enter text..

Care must be taken to integrate AI into existing academic programs without overwhelming students. This calls for a phased approach, starting with dedicated AI modules and gradually weaving AI concepts into broader curricula. Ethical, privacy, and security concerns must be addressed proactively by embedding AI ethics and governance into all AI-related programs.

Finally, to stay ahead in a rapidly evolving field, educational institutions must establish strong partnerships with leading AI research centers, industry leaders, and technology firms. Such collaborations will ensure continuous curriculum relevance and provide students and faculty with access to the latest advancements in AI.

A. Ai For Education:

Data Privacy: The use of AI in education raises concerns about student data security. Regulations like GDPR and FERPA must be strictly followed.

Equity: AI-driven tools may inadvertently favor students from technologically advanced backgrounds, exacerbating educational inequalities (UNESCO, 2021).

B. Education For Ai:

Curriculum Design: Rapid advancements in AI make it challenging to keep curricula up-to-date.

Ethical Training: Students must be taught to address AI biases and ethical dilemmas in real-world applications [10].

I. RESULTS

The analysis of AI for Education (AI4E) and Education for AI (E4AI) reveals distinct yet complementary impacts on educational institutions and learners, supported by empirical evidence and case studies.

For AI4E's Institutional Transformation and Operational Efficiency, institutions using AI-driven tools (e.g., chatbots, predictive analytics) reported 30–50% reductions in administrative workload (e.g., Arizona State University's chatbot system). Automated grading systems (e.g., Turnitin) improved assessment consistency while freeing faculty time for higher-value tasks. In the Personalized Learning field, adaptive platforms (e.g., Carnegie Learning) increased student engagement by 20–35%, with tailored content improving mastery rates in STEM subjects. Using Data-Driven Decision-Making, AI-powered labor market analysis (e.g., World Economic Forum, 2020) enabled faster curriculum updates, aligning programs with employer needs [11].

For E4AI's Learner Outcomes in terms of AI Literacy and Workforce Readiness, programs integrating AI fundamentals (e.g., MIT's "Machine Learning for Business") saw 40% higher employability among graduates in AI-augmented fields. Non-technical AI courses (e.g., Harvard's CS50 for Business) improved managerial AI literacy, with 75% of participants applying AI tools in their roles. As for ethical and Innovation Competencies, case studies highlighted gaps in ethics training; only 30% of E4AI programs included mandatory AI ethics modules [10].

A. Synergies And Challenges:

- Integration Successes: Hybrid models (e.g., AI4E-powered platforms delivering E4AI content) enhanced both institutional efficiency and learner outcomes.
- Equity: AI4E tools favored resourced institutions, exacerbating the digital divide (UNESCO, 2021).
- Faculty Readiness: 60% of educators reported insufficient training to teach or use AI tools.

B. Key Takeaways:

- AI4E excels in systemic optimization, while E4AI transforms learner capabilities.
- Ethical and equitable implementation remains under-addressed across both paradigms.
- Dual-strategy adoption is critical for future-ready education.

These results underscore the need for balanced policies and investments to harness AI's full potential in education.

V. CONCLUSION & FUTURE DIRECTION

To successfully implement AI4E and E4AI, institutions should adopt the following strategies:

1. Adopt a Dual-Strategy Framework:
   - Integrate AI4E tools (e.g., predictive analytics, chatbots) to streamline administrative tasks and personalize learning.
   - Embed E4AI curricula across disciplines, ensuring all students gain foundational AI literacy, ethical training, and hands-on experience with AI tools.

2. Prioritize Faculty Development:
   - Launch upskilling programs to train educators in AI tools and pedagogy.
   - Foster collaborations with industry and AI research centers to keep faculty updated on advancements.

3. Ensure Ethical and Equitable Implementation:
   - Establish clear policies for data privacy, bias mitigation, and accessibility to prevent AI from exacerbating disparities.
   - Pilot AI initiatives in diverse settings to evaluate their impact on underrepresented student populations.

4. Strengthen Industry-Academia Partnerships:
   - o Collaborate with employers to align curricula with workforce demands, ensuring graduates possess relevant AI skills.
   - o Develop internship and co-op programs that provide students with real-world AI applications.

5. Leverage Hybrid Learning Models:
   - o Combine AI-driven platforms (e.g., adaptive learning systems) with human mentorship to balance efficiency and personalized support.

A. Future Research Directions

To bridge existing gaps and optimize AI's role in education, future research should focus on:

1. Long-Term Efficacy Studies:
   - o Conduct longitudinal studies to assess the sustained impact of AI4E and E4AI on institutional performance and student outcomes.
2. Pilot Evaluations of Hybrid Models:
   - o Test and refine blended approaches (e.g., AI-powered tutoring + instructor-led sessions) to identify best practices for scalability.
3. Ethical AI Frameworks:
   - o Develop standardized guidelines for AI ethics in education, addressing bias, transparency, and accountability.
4. Equity-Centered AI Deployment:
   - o Investigate strategies to ensure AI tools benefit all learners, including those from low-resource backgrounds.
5. Global Implementation Roadmaps:
   - o Create adaptable frameworks for institutions worldwide, accounting for cultural, infrastructural, and regulatory differences.

The convergence of AI4E and E4AI is not merely an opportunity but a necessity for future-ready education. By adopting a balanced, ethically grounded approach, institutions can harness AI's potential to transform learning while empowering students to lead in an AI-augmented world. The path forward requires collaboration, continuous evaluation, and a commitment to equity, ensuring that the benefits of AI in education are accessible to all.

The future of education lies not in choosing between AI for institutions or learners, but in harmonizing both to create an ecosystem that is efficient, inclusive, and innovative**.**

REFERENCES

[1] Achanta, A. (2023). Data Democratization: Empowering Non-Technical Users with Self-Service BI Tools and Techniques to Access and Analyze Data Without Heavy Reliance on IT Teams. International Journal of Computer Trends and Technology, 71(8), 39–46. https://doi.org/10.14445/22312803/ijctt-v71i8p106

[2] Carthon, J. E., Aladjem, D., Daniels, D. V, & Fletcher, K. (2022). Redesigning Principal Preparation : A Case Study of the ASU Educational Leadership Tier II Prog ram. Education Leadership Review, 23(1), 286–304.

[3] Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P., & Vayena, E. (2018). AI4People—An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations. Minds and Machines, 28(4), 689–707. https://doi.org/10.1007/s11023-018-9482-5

[4] Harry, A. (2023). Harry, A. (2023). Role of AI in education. Injuruty: Interdisciplinary Journal and Humanity, 2(3). e-ISSN: 2963-4113, p-ISSN: 2963-3397. 2(3), 260–268.

[5] Holmes, W., & Tuomi, I. (2022). State of the art and practice in AI in education. European Journal of Education, 57(4), 542–570. https://doi.org/10.1111/ejed.12533

[6] Mishra, A. (2024). Empowering AI-Powered Product Companies : Enhancing Design with Knowledge Management , Open Innovation , and Foresight.

[7] Schwab, K., & Zahidi, S. (2020). World Economic Forum.(2020). The Global Competitiveness Report–How Countries Are Performing on the Road to Recovery.

[8] Selingo, J. (2017). How colleges use big data to target the students they want. The Atlantic, 4.

[9] Selwyn, N. (2022). The future of AI and education: Some cautionary notes. European Journal of Education, 57(4), 620–631. https://doi.org/10.1111/ejed.12532

[10] Shahid, A. R., & Mishra, S. (2024). A Framework for a Master's in Applied Artificial Intelligence Program in Computer and Information Systems Discipline. Journal of Information Systems Education, 35(4), 495–511. https://doi.org/10.62273/EQZE3625

[11] Shum, S. J. B., & Luckin, R. (2019). Practices. 1–10.

[12] Thirulingam, A., & Pais, S. (2024). Discovering the Most Needed Technologies and Skills for IT Jobs on LinkedIn. 138–149. https://doi.org/10.34074/proc.240118

[13] Zhai, X., Chu, X., Chai, C. S., Jong, M. S. Y., Istenic, A., Spector, M., Liu, J. B., Yuan, J., & Li, Y. (2021). A Review of Artificial Intelligence (AI) in Education from 2010 to 2020. Complexity, 2021. https://doi.org/10.1155/2021/8812542

Reframing UTAUT for Mandatory Public AI Services: A Case from Abu Dhabi, UAE

Muneera Ali Alsulaimani
Department of Information Systems and Technology Management
Zayed University
Abu Dhabi, UAE
M80008969@zu.ac.ae

Maher Alaraj
Department of Information Systems and Technology Management
Zayed Univeristy
Dubai, UAE
maher.alarah@zu.ac.ae

Abdulla Naqi
Department of Information Systems and Technology Management
Zayed Univeristy
Dubai, UAE
Abdulla.Naqi@zu.ac.ae

Abstract—Artificial Intelligence (AI) is transforming public service delivery in the UAE, exemplified by Abu Dhabi's TAMM platform. This study employs an extended Unified Theory of Acceptance and Use of Technology (UTAUT) framework, substituting Behavioral Intention with User Satisfaction to reflect mandatory-use contexts. Using a sequential explanatory mixed-methods approach, the research collected survey responses from 137 TAMM users and conducted follow-up interviews. Key findings show that Effort Expectancy, Facilitating Conditions, and Social Influence significantly influence satisfaction. Moderating effects of gender and AI experience were also observed. The study offers theoretical extensions to UTAUT and strategic recommendations aligned with the UAE Digital Government Strategy 2025–2027.

Keywords— Keywords—AI adoption, user satisfaction, UTAUT, government platforms, Abu Dhabi, TAMM, SmartPLS, trust

I. INTRODUCTION

The adoption of Artificial Intelligence (AI) by public sector organizations is reshaping government service delivery worldwide [1]. In the United Arab Emirates (UAE), specifically in Abu Dhabi Emirate, platforms like TAMM represent this digital evolution by integrating AI-driven features to enhance citizen interactions [2]. Abu Dhabi aims to be the first fully AI-powered government by 2027, supported by initiatives such as the Artificial Intelligence and Advanced Technology Council and the Digital Government Strategy 2025–2027 [3].

Despite these strategic ambitions, the success of AI initiatives hinges not just on deployment but on public satisfaction and acceptance [4]. This is particularly crucial in mandatory-use platforms where opting out is not feasible. Yet, limited empirical research exists on user satisfaction in such settings, especially in the Gulf context [5].

This study addresses this gap by extending the UTAUT model [6] traditionally focused on Behavioral Intention, by substituting it with User Satisfaction. It investigates how core UTAUT constructs—Performance Expectancy, Effort Expectancy, Social Influence, and Facilitating Conditions—shape satisfaction with AI features on the TAMM platform in as the mandatory context to be utilized for any application to government services. It also explores how demographic factors such as age, gender, and prior AI experience moderate these relationships. As shown in Figure 1 is the research model used. Based on the UTAUT constructs and the study objectives, the following hypotheses are proposed:

- H1: Performance Expectancy has a significant positive effect on User Satisfaction.
- H2: Effort Expectancy has a significant positive effect on User Satisfaction.
- H3: Social Influence has a significant positive effect on User Satisfaction.
- H4: Facilitating Conditions have a significant positive effect on User Satisfaction.
- H5: User Experience significantly contributes to overall User Satisfaction.

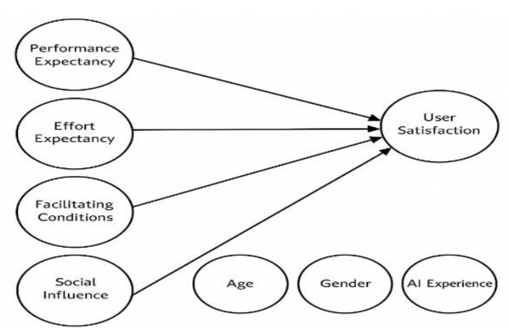


Figure 1: Research Model.

II. METHODOLOGY

This study adopts a sequential explanatory mixed-methods approach, combining both quantitative and qualitative research methods to explore the factors influencing user satisfaction with AI services on the TAMM platform in Abu Dhabi [7]. The methodology is designed to provide a comprehensive understanding of AI adoption in a mandatory-use government setting, specifically focusing on the user experience with AI-powered services.

The first phase consisted of a quantitative survey distributed to 137 users of the TAMM platform, using purposive sampling to ensure participants had experience with AI features. The survey consisted of questions designed to capture data on key factors influencing user satisfaction.

The survey was administered online through Google Forum and was structured around the key UTAUT constructs, such as Performance Expectancy, Effort Expectancy, Social Influence, and Facilitating Conditions, as well as the additional construct of User Satisfaction. The responses were recorded on a 5-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree).

The second phase involved semi-structured interviews with TAMM workers to provide deeper insight into the quantitative findings, semi-structured interviews were conducted with three key stakeholders from the TAMM team. These individuals included a senior government employee, a policymaker, and an IT team member responsible for implementing AI features on the platform. The interviews focused on their perspectives on the challenges of AI adoption, strategic decisions, and user experiences with the platform. Thematic analysis was used due to its flexibility than other tools because it does not require any predefined theoretical framework [9]. The data from these interviews were analyzed using thematic analysis using ChatGPT to identify key themes related. As most of recent research uses Large model Language like ChatGPT to capture the main themes.

Ethical approval for this research was obtained from **Zayed University's Ethics Committee**, ensuring that all participants' rights and privacy were protected. Informed consent was obtained from all survey participants and interviewees, and the responses were anonymized to maintain confidentiality. In compliance with ethical research standards and privacy regulations, the data collected will not be shared with third parties without explicit consent from the participants. The findings from this research will only be disseminated in aggregate form or anonymized reports to ensure that no individual participant can be identified.

III. RESULTS AND KEY INSIGHTS

Quantitative analysis mainly relied on testing the hypothesis using path coefficient and it revealed that Effort Expectancy ($\beta = 0.252$, $p = 0.003$), Facilitating Conditions ($\beta = 0.304$, $p < 0.001$), and Social Influence ($\beta = 0.176$, $p = 0.029$) significantly influenced User Satisfaction. Performance Expectancy ($\beta = 0.135$, $p = 0.112$) did not show a significant effect. The model explained 48.2% of the variance in user satisfaction ($R^2 = 0.482$). as shown in Table 1 is the path coefficients.

	Original sample (O)	Sample mean (M)	P values
EE -> UES1	0.252	0.252	0.003
FC -> UES1	0.304	0.311	0
PE -> UES1	0.135	0.133	0.112
SI -> UES1	0.176	0.174	0.029

Table 1, Path Coefficients and Significance

Moderation analysis were tested using The Multi-Group Analysis (MGA) technique and Interaction Term Modeling in SmartPLS. The results showed gender and AI experience had significant moderating effects. For instance, Facilitating Conditions were more influential among female users, while Social Influence had a greater impact on users with no prior AI experience. While age did not showed any significance across the model constructs.

Qualitative insights supported these findings, emphasizing trust, usability, and the importance of hybrid human-AI interaction [10]. After writing the script of the main parts of the interview, the script was uploaded to ChatGPT with the following Command: “This is the script of interviews and the questions are based UTAUT extended with Satisfaction instead of use of behaviour, I want to capture the main themes using thematic analysis”. After that the researcher manually checked the output and edit it. Next phase was to link every theme with the UTAUT construct and research hypothesis in order to be able to link them with the quantitative analysis.

Interview themes included AI as a government efficiency enabler, the ease of accessing services through multiple channels, the importance of support infrastructure, and ethical AI design considerations[11]. As shown in Table 2 the main themes was extracted from the interview script.

Table 2, Thematic analysis

Theme	**UTAUT Construct**	**Related RQ & Hypothesis**	**Outcome**
AI Enhances Government Efficiency	Performance Expectancy	H1	Supported
Usability and Multichannel Design	Effort Expectancy	H2	Supported
Training and Support	Facilitating Conditions	H3	Supported
Managerial and Peer Influence	Social Influence	H4	Partially Supported
Hybrid Use and Trust	User Satisfaction	H5	Supported
Ethical AI Governance	Trust (Emergent)	-	Emergent

IV. CONCLUSION

This study extends the UTAUT model by positioning User Satisfaction as a key outcome in mandatory-use government AI platforms. The findings affirm that usability, organizational support, and peer influence are central to AI service satisfaction in public systems.

A. Summarize Findings

The quantitative results highlighted that Effort Expectancy (EE) and Facilitating Conditions (FC) significantly impacted user satisfaction with AI services on the TAMM platform. These findings were echoed in the qualitative interviews, where participants emphasized the importance of ease of use and the availability of training and support. Users reported that their satisfaction increased when they could easily navigate the platform and had access to reliable support resources.

In contrast, Performance Expectancy (PE) did not show a significant effect on satisfaction, which was consistent with qualitative insights. While participants acknowledged the potential benefits of AI in terms of efficiency, their satisfaction was more influenced by usability and trust in the platform. Many users expressed that their satisfaction was driven by their ability to use the system with confidence, rather than solely by its performance improvements.

Finally, the Social Influence (SI) construct had a moderate effect on user satisfaction. Some interviewees indicated that government endorsements and peer recommendations played a role in their decision to use the platform. However, personal experiences with the system, particularly its usability and reliability, were found to have a stronger influence on overall satisfaction, suggesting that external pressures had less impact than direct user experience.

B. Key Contributions

Theoretical: Introduces User Satisfaction into UTAUT, adapting it to compulsory service contexts. This study makes the most significant innovation with the use of and extension of the UTAUT framework to Abu Dhabi's compulsion to utilize TAMM platform, one of the first comprehensive examination of AI acknowledgment in the Arab public part.

Practical: Highlights design implications for enhancing trust, usability, and accessibility. Furthermore, the study also confirms the results of a recent work of Rana et al. (2023) [13] on AI acceptance in public services, where they suggested to enrich the UTAUT with constructs such as trust in algorithms and perceived transparency. The most substantial contribution of the study is the realization of how User Experience outperforms the majority of UTAUT constructs in the scenarios of mandatory AI adoption. This finding is consistent with current advances in technology acceptance theory that proposes including experiential and affective aspects in the era of AI [14].

Methodological: Demonstrates the value of mixed-methods in understanding technology adoption. A cross-sectional design was used, capturing user attitudes and perceptions at one point in time only. This prevents the ability to monitor changes in user perceptions or satisfaction over a prolonged period.

C. Furture Direction

1. Longitudinal studies and designs should be used in future research tracking how user perceptions change over time as AI systems become more mature and users become more experienced.
2. Comparative studies across different national and cultural contexts would enable to disentangle universal adoption factors from the culture-specific ones. (Medaglia et al., 2023).
3. In-depth analysis of trust dimensions in AI systems. Future studies must distinguish the trust of technology from the trust of the government machinery to deploy it, and to trust the specific implementation of it.
4. Studies need to test various AI-human partnership systems to determine the most acceptable and effective division of labor between AI automation and human customer service.

Our recommended future research provides a clear direction to expand upon these project results in both theory and practice of AI adoption by government agencies. Through these research areas researchers will make better technology acceptance models that help governments correctly use AI in public services while keeping citizens happy.

More governments around the world use AI in operations so research about its implementation will gain greater importance. The actions of the Abu Dhabi reveal key practices that other governments can apply when they start adopting digital technology. New research on these conclusions and research gaps would enhance government practice with AI technology.

REFERENCES

[1] Y. K. Dwivedi, L. Hughes, E. Ismagilova, et al., "Artificial intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy," International Journal of Information Management, vol. 57, p. 101994, 2021.

[2] K. Alhosani and S. M. Alhashmi, "Opportunities, challenges, and benefits of AI innovation in government services: A review," Discover Artificial Intelligence, vol. 4, no. 18, 2024.

[3] Abu Dhabi Media Office, "Abu Dhabi Government launches Digital Strategy 2025–2027," Abu Dhabi Media Office, 2025.

[4] M. Alshehri, S. Drew, and T. Alhussain, "Mandatory technology adoption in government: A UAE case study," Government Information Quarterly, vol. 40, no. 1, p. 101763, 2023.

[5] S. Alsheibani, Y. Cheung, and C. Messom, "Artificial intelligence adoption: AI-readiness at firm-level," PACIS 2018 Proceedings, p. 198, 2018.

[6] V. Venkatesh, M. G. Morris, G. B. Davis, and F. D. Davis, "User acceptance of information technology: Toward a unified view," MIS Quarterly, vol. 27, no. 3, pp. 425–478, 2003.

[7] J. W. Creswell and V. L. P. Clark, Designing and Conducting Mixed Methods Research, 3rd ed., SAGE Publications, 2018.

[8] J. F. Hair, C. M. Ringle, and M. Sarstedt, "Partial least squares structural equation modeling," in Handbook of Market Research, Springer, 2017, pp. 1–40.

[9] V. Braun and V. Clarke, "Using thematic analysis in psychology," Qualitative Research in Psychology, vol. 3, no. 2, pp. 77–101, 2006.

[10] S. Farber, "Comparing human and AI expertise in academic peer review: Towards a hybrid approach," Higher Education Research & Development, 2025.

[11] S. Vaithilingam, S. K. Raj, and K. Paulraj, "The role of large language models in qualitative research: Opportunities and challenges," AI & Society, 2022.

[12] R. Medaglia, J. R. Gil-Garcia, and T. A. Pardo, "Artificial intelligence in government: Taking stock and moving forward," Social Science Computer Review, vol. 41, no. 1, pp. 123–140, 2023.

[13] Rana, N. P., Chatterjee, S., Dwivedi, Y. K., & Akter, S. (2023). Understanding dark side of artificial intelligence (AI) integrated business analytics: Assessing firm's operational inefficiency and competitiveness. European Journal of Information Systems, 32(2), 364–387. https://doi.org/10.1080/0960085X.2021.1955628

[14] V. Venkatesh, T. Sykes, and X. Zhang, "'Enterprise AI': Advancing digital transformation research," MIS Quarterly, vol. 46, no. 3, pp. iii–xxii, 2022.

Ancient Languages Classification using Artificial Neural Networks

Safia Al Ali
MSc in Business Analytics – AI Management
Abu Dhabi School of Management
Abu Dhabi, UAE
safia4489@gmail.com

Abstract —**The study of ancient languages is essential to gain a comprehensive understanding of mankind's history. Recently, there has been a growing interest in leveraging AI in this area. The aim of this paper is to apply AI deep learning classification models to five ancient languages which are Ancient Greek, Arabic, Old Chinese, Egyptian Hieroglyph, and Sanskrit. Three CNN Residual Network models were implemented: ResNet34, ResNet50, and ResNet101. Results revealed that ResNet101 is the most accurate model with 92% rate achieved. In addition, Sanskrit is the most accurately predicted language and Arabic is the least accurate. Further research suggests expanding the dataset, including more languages, and exploring other models.**

Keywords—Ancient Languages, AI, Deep Learning, Convolutional Neural Networks, Residual Network

I. INTRODUCTION

Ancient languages vary tremendously from their modern counterparts, presenting unique challenges and requiring significant effort and expert involvement. The aim of this thesis is to bridge the linguistic gap by harnessing the power of Neural Network Artificial Intelligence models to develop an advanced tool capable of classifying various ancient languages. A distinctive dataset is compiled specifically for this study. There are a few expected limitations to be considered, such as the requirement for a massive amount of data for optimal performance, which is not possible due to time constraints and CPU limitations. In addition, ancient languages have been written in various media and may have evolved over history which adds complexity and variation. Furthermore, model validation and accuracy might require archeologists' involvement in real world applications, therefore the evaluation of performance will be based on test data [4].

II. METHODOLOGY

A. Business Understanding

The chosen methodology, CRISP-DM, will be followed step by step to achieve the thesis objective of classifying various ancient languages. Below is a summary of main steps followed [6].

B. Data Understanding

The collected dataset consists of images from various online public sources. This dataset is exclusively gathered with the aim of having different groups of images for at least five ancient languages divided into model testing, training, and evaluation. The dataset is stored locally. Unlike structured tabular format, image dataset is simple with limited variables such as language name, title of the image, its type, and size.

C. Data Availability and Ethics Statement

The dataset used in this study was self-compiled by the author through the collection of publicly available images from Google Images, ensuring that only images licensed for reuse were selected. These images represent five different ancient languages and were curated specifically for training and evaluating the neural network model described in this paper. Since the dataset is composed of non-personal, publicly available images, and internal ethical approval was acquired for this research from our Research committee. The dataset is not hosted in a public repository, but it can be made available upon reasonable request to the corresponding author for academic and research purposes only.

D. Data Preparation

To ensure optimal image dataset is collected and preprocessed, the following steps were accomplished:

- Data cleaning: While gathering data, it was noticed that certain images have modern languages titles as part of the image. This must be cropped to avoid disturbing the accuracy of running the machine learning algorithm later.
- Data formatting: The standard types of images are JPEG and PNG. However, there were unusual types such as WEBP which the model might not be able to read. Therefore, such images will be converted to JPEG or substituted with alternatives.
- Data Selection: The quality of certain images gathered were not clear. Such images also require replacement to avoid disrupting the accuracy of the model.

E. Modeling

Modelling involves organizing the collected dataset to adhere to a predefined configuration setting. The configured parameters are recommended to establish a robust model and enhance performance. The collected dataset will be structured into three different main prelabeled categories or folders which are: train, test, and valid. Each designated folder will include a curated collection of images dataset that will represent 5 distinct ancient languages which are: Egyptian Hieroglyphs, Chinese, Ancient Greek, Arabic, and Sanskrit. Thirty images per language will be allocated for the train folder to ensure sufficient and effective model training. The test folder will contain 10 images for each language, and another 10 images will be allocated to the valid folder. It’s essential to allocate training, testing, and validation to ensure reliability and generalization [5]. A considerable portion of collected data will be allocated to the training; thus model can learn features and patterns. The chosen model for this case study is the Convolutional Neural Network (CNN). CNN models are designed specifically for images data for tasks such as classification. Docker containers and Kubernetes will be launched to run the Python code of the chosen CNN models: ResNet34, ResNet50, and ResNet101.

F. Evaluation

When implementing CNN models, several evaluation methods can be applied to assess the algorithms. In this thesis, the main matrices that will be used are the confusion matrix, prediction/actual values, and loss/probability scores.

G. Deployment

Successful implementation depends not only on model accuracy but also on its usability and accessibility. There are various methods of deployment strategies and in this thesis paper we propose the following: tool integration in different platforms such as websites, mobile applications, or systems that are specialized in linguistic and historical fields. Examples of such tools can be ancient language translation and archives [3]. Another deployment recommendation is to include the model in the GitHub platform. GitHub is a well-known advanced platform that allows collaboration, sharing, and machine learning deployment [2]. Stand-alone downloadable python files are another useful approach of deployment.

III. DATA ANALYSIS

There are various insightful techniques to conduct image dataset analysis, this section includes some of these techniques which were applied for the ancient language classification model.

A. Visual Inspection

Visual Inspection of an image can provide various details such as image quality, condition of the manuscript, paper and ink, medium and craftsmanship, condition of the surface, resolution, clarity, and visual consistency.

B. Statistical Analysis

To perform statistical analysis of image dataset, specific codes were run in Python Google Colab. Examples of statistical analysis are image size or dimension, color distribution, edge detection, pixel intensity distribution, and texture analysis.

IV. RESULTS

Three CNN models will be used for images dataset classification task of ancient languages. The breakdown of the Python code steps and the achieved results is as below:

- The code will start by importing the required libraries and setting up the data using the Image Data Bunch function from the specified data path uploaded.
- Data exploration phase will also be included at this stage to explore and identify different variables such as classes names and sample batch display.
- 3 CNN models will be run which are: ResNet34 Model, ResNet50 Model, and ResNet101 Model

A. ResNet34 Model

The first chosen CNN model is ResNet34 which is implemented in Python by creating the CNN function and training of 4 epochs. Epochs is the frequency which the algorithm goes over the whole training dataset to learn [1]. The initial result of the model achieved **52%** performance accuracy.

The confusion matrix and most confused classes (languages) are displayed by adding the classification interpretation function. in Python. Confusion Matrix and most confused graphs revealed that model confused Arabic language with Sanskrit 7 times, Sanskrit with Old Chinese 5 times, and both Ancient Greek language and Egyptian with Arabic 3 times. While the model was able to accurately classify Sanskrit language at all 10 times of validation process.

Then the ResNet34 model was unfrozen to go through fine tuning process with further training for one more epoch which resulted in slightly higher accuracy of **54%** as illustrated in figure 1.

```
In [29]: learn.unfreeze()

In [31]: learn.fit_one_cycle(1)
```

Total time: 00:09

epoch	train_loss	valid_loss	accuracy
1	0.920701	1.116462	0.540000

Fig. 1. ResNet34 Fine Tune Process Result

B. ResNet50 Model

CNN ResNet50 model was the second model applied with an increase image size to 299 and smaller batch size. The size of batch indicates the number of sample data handled by the model before the update occurs [1].

After training of 8 epochs, the achieved accuracy of validation image dataset reached **84%**. Fine tuning process was applied to unfreeze the model and train for 6 times more epochs which enhanced the performance accuracy to achieve **88%** as shown in Figure 2.

In [55]:
```
learn.unfreeze()
learn.fit_one_cycle(6,max_lr=slice(1e-5,1e-4))
```

Total time: 01:05

epoch	train_loss	valid_loss	accuracy
1	0.032597	0.323320	0.860000
2	0.028857	0.305685	0.860000
3	0.023173	0.297088	0.880000
4	0.021014	0.292464	0.900000
5	0.020567	0.290487	0.880000
6	0.019565	0.289780	0.880000

Fig. 2. Fine Tune Process ResNet50

Most confused graphs revealed that model confused Arabic language with Sanskrit 2 times and confused each of the following one time only: Ancient Greek language with Sanskrit, Arabic with Egyptian, Old Chinese with Egyptian, and Egyptian with Greek.

Confusion Matrix revealed that Sanskrit is the most accurate language with all 10-validation dataset predicted accurately, followed by Old Chinese, Egyptian Hieroglyphs, and Ancient Greek with each 9-validation dataset accurately predicted. The least accurate language is Arabic with only 7 validated data points correctly predicted.

C. ResNet101 Model

The last CNN model applied is ResNet101 model with an increase image size to 384. Similar to ResNet50, 8 epochs were trained, the achieved performance accuracy of validation dataset is **88%.**

The model went through fine tuning process to train for 6 more epochs which improved performance accuracy to achieve **92%** as revealed in figure 3.

```
learn.unfreeze()
learn.fit_one_cycle(6,max_lr=slice(1e-5,1e-4))
```

Total time: 35:26

epoch	train_loss	valid_loss	accuracy
1	0.079928	0.371501	0.880000
2	0.081849	0.359509	0.880000
3	0.077750	0.346395	0.880000
4	0.066292	0.337372	0.900000
5	0.059387	0.326175	0.920000
6	0.056353	0.324826	0.920000

Fig. 3. Fine Tune Result of ResNet101

Most confused graphs revealed that model confused one-time Arabic language with each Sanskrit, Egyptian, Ancient Greek. In addition, it confused Ancient Greek with Old Chinese.

Confusion Matrix revealed that Sanskrit, Old Chinese, and Egyptian are accurately predicted languages with all 10-validation dataset, followed by Ancient Greek with 9-validation dataset accurately predicted. The least accurate language is still Arabic with 7 correctly predicted as illustrated in figure 4.

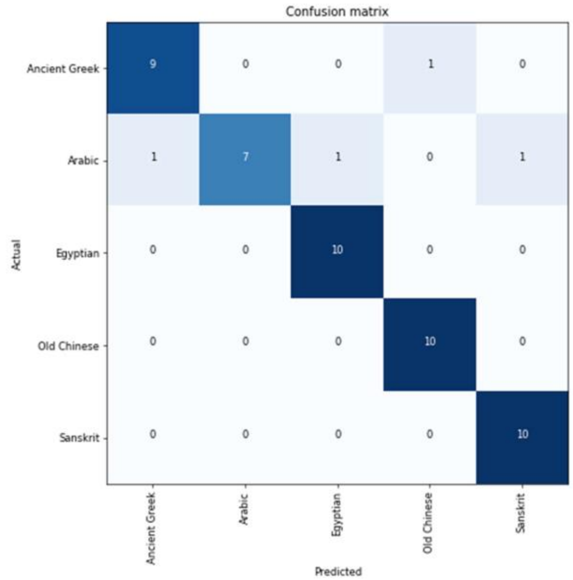


Fig. 4. ResNet101 Confusion Matrix

To summarize the results, table 1 and table 2 illustrate and compare how implemented ResNet CNN models architecture depth and image size play a significant role in the performance of classification task for ancient languages. In general, as the model depth and image size increased, the performance accuracy improved and confusion between languages decreased noticeably. Fine-tuning played a vital role in enhancing the accuracy of performance by leveraging pre-trained features specially when dealing with special domain such as ancient languages classification. ResNet101, which is the model with profound depth and largest image size, is recommended as it achieved highest performance accuracy of 92% after fine tune process were implemented as shown in table 1. In addition, RestNet101 achieved the lowest error rate of just 6% with 46 out 50 values predicted accurately as revealed in table 2.

TABLE I. RESNET MODELS PERFORMANCE RESULTS

CNN Model	Image Size	Accuracy	Fine Tune Accuracy
ResNet34	224	52%	54%
ResNet50	299	84%.	88%
ResNet101	384	88%	92%

CNN Model	Error rate	Number of correct predicted values	Most Confused Language	Most Accurate Language
ResNet34	40%	30	Arabic	Sanskrit
ResNet50	12%	44	Arabic	Sanskrit
ResNet101	6%	46	Arabic	Sanskrit, Old Chinese, Egyptian

V. RECOMMENDATIONS AND CONCLUSION

Based on the revision, experiment, and output of the implemented classification model for the ancient languages task, it is recommended for future research to consider the following:

- Continuous experiments with different parameters for the three models and different image sizes.
- Focused scientific studies in most confused languages such as Arabic, which tend to be more complicated than others and require separate attention to its patterns and details.
- The total image dataset used for this thesis is 250 for 5 distinct languages. The number of collected dataset can be increased as well as number of languages to produce more profound solid ancient language classification models.
- The involvement of linguistic experts can be helpful as they will add useful insights.
- In addition, Other Convolutional Neural Network can be explored to understand if there are better models for ancient languages classification task.

In conclusion, understanding humanity's history, languages, cognitive development, and culture is the underlying motivation behind researching ancient languages. However, there are many challenges in this field due to scarcity of linguistic materials and the extensive amount of time and expertise required. In recent decades, there has been an increased interest and scholarly attention in implementing advanced technological tools such as Artificial Intelligence and machine learning to overcome these challenges. Implementing AI and deep learning models for Ancient Languages classification task was the main driven to research and add value for linguistic experts with providing useful tool for their studies and work.

REFERENCES

[1] Brownlee, J. (2018). What is the Difference Between a Batch and an Epoch in a Neural Network. Machine Learning Mastery, 20.

[2] Cosentino, V., Izquierdo, J. L. C., & Cabot, J. (2017). A systematic mapping study of software development with GitHub. Ieee access, 5, 7173-7192.

[3] Dong, X. L., & Rekatsinas, T. (2018, May). Data integration and machine learning: A natural synergy. In Proceedings of the 2018 international conference on management of data (pp. 1645-1650).

[4] Guidi, T., Python, L., Forasassi, M., Cucci, C., Franci, M., Argenti, F., & Barucci, A. (2023). Egyptian Hieroglyphs Segmentation with Convolutional Neural Networks. Algorithms, 16(2), 79. https://doi.org/10.3390/a16020079

[5] e, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778).

[6] Nama, G. F., Nabella, R. O., Wintoro, P. B., & Mulyani, Y. (2022). Towards CRISP-ML (Q): Consumer Segmentation Knowledge Analysis Using Data Science Approach for Marketing Strategy Recommendations. NeuroQuantology, 20(5), 3479-3489. 10.14704/nq.2022.20.5.NQ22647I

Cognitive Intelligence for Enhancing Soldier Health Predictions in Military Operations

Ilias Panagiotopoulos
Dept. of Informatics & Telematics
Herokopio University
Athens, Greece
ipanagio@hua.gr

George Dimitrakopoulos
Dept/ of Informatics & Telematics
Harokopio University
Athens, Greece
gdimitra@hua.gr

Abstract—In recent years, personalized health predictions are facilitated and supported by the latest advances in Information and Communication Technologies (ICTs), forming the cornerstone of the new generation of Cognitive Military Healthcare Management Systems (CMHMSs), enabling increased precision of diagnostics and novel health decision-making solutions for each soldier. The ambition of the present study is to introduce a CMHMS architecture, being able to support intelligent soldier health predictions in the field of military healthcare. The proposed framework can assist military personnel and medical workers in promptly identifying soldiers who may need medical attention, allowing for more effective and efficient treatment.

Keywords—military healthcare, soldier health, personalized health predictions, cognitive intelligence

I. INTRODUCTION

Soldier health is a crucial component of military operations and the capacity to track and forecast it in real-time can have a big impact on both the troops' well-being and the success of missions [1]. In recent years, it is possible to create personalized health predictions through the application of Cognitive Military Healthcare Management Systems (CMHMSs), by enabling increased precision of diagnostics and novel health decision-making solutions for each soldier. The above can be accomplished due to the growing number of wearable technologies, medical sensors, cognitive intelligence, machine learning algorithms, and IoT technologies [2]. CMHMSs operate because of collecting information from various sources, intelligently processing it, integrating knowledge and experience, and finally, taking the most appropriate health decisions.

In the light of the above, the scope of this study is to investigate the way that cognitive features in health managemet systems can enhance military healthcare. Based on this, the aim is to introduce a CMHMS architecture, namely 'i-SHM' (intelligent Soldier Health Monitoring), which aims to support intelligent soldier health predictions in the field of military healthcare, through the exploitation of a dynamic and automatic adaptation of the soldier's level of health status. The proposed functionality can gather information from a variety of sources, intelligently processing it, integrating knowledge and experience, and finally, producing unique health forecasts and severity checking for each soldier.

The value of this study is thought to be particularly important, as this paper reports one of the first studies of its kind, where cognitive intelligence can assist medical doctors in identifying soldiers who might need medical assistance, allowing for more effective and efficient treatment. Additionally, this study offers insightful data information on the general condition of a military unit, empowering commanders to make wise choices regarding troop deployment and mission preparedness.

II. BACKGROUND

In general, Cognitive Management Systems (CMSs) are computerised tools that use techniques to synthesize and/or analyze data, and in some cases, make recommendations – even predictions – to aid human decision-making in various applications. The advantages of CMSs are often framed in terms of increased situational awareness and faster decision-making cycles [3].

In the light of the above, an area of applications where CMS could find prosper ground is military healthcare and field operations, as military commanders and others responsible for the battlefield can base their decisions on information from all sources available to them at the relevant time [4]. More in detail, Fig. 1 shows the usual decision making cycle followed in the military healthcare domain. The whole cycle consists in an interaction between the operation field domain and the military commander domain. The military commanders collect contextual information on the operation field domain. This real-time collection data, together with historical data and previous experiences, constitute the information to be analyzed by the military commanders. The analysis results in the military commander's decision upon the most appropriate manner to be applied to the operation field domain [5]. During the decision making process, the military commanders consider specific goals and policies, as well as past knowledge and experience, which is derived from previous interactions with field operatrions with similar physical and environmental characteristics. So, the whole process can be reflected on a closed loop [6].

As depicted in Fig. 1, the basic motivation behind the cognitive-based decision making cycle in military healthcare domain has to do with the fact that the militay commanders could be significantly facilitated by an intelligent system that keeps track of past actions, stores information on a "knowledge database" and provides this information as an input, prior to decision making. At another level, CMS may cater for fast and effective adaptations of the communication infrastructure to changing requirements, and thus guarantee unobtrusive communication during critical situations [7].

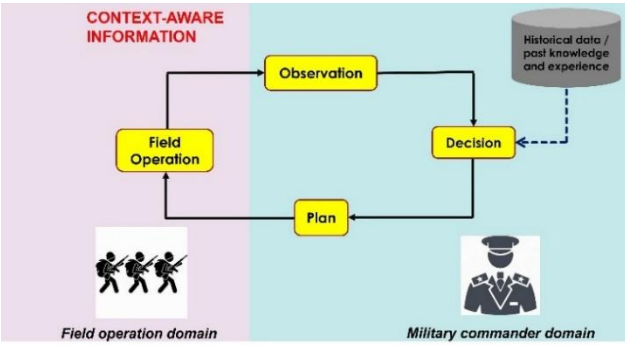


Fig. 5. The decision making cycle in military healthcare domain: a cognitive-based approach

What can also be stated is that the above show how novel management functionalities, enhanced with cognitive networking capabilities, may be needed for the provision of military healthcare in fileld operations with faster transmissions, as well as higher reliability and availability. Such cognitive systems are most severe in battle environments for providing efficient strategic decisions and greater personalization of the field operations. Before CMS can be deployed in military healthcare applications, they need to be trained through data that are generated from operational activities, so that they can learn similar groups of subjects, associations between subject features and outcomes of interest [8].

CMS trainings are based mainly on machine learning (ML) techniques, where data-analytical algorithms can extract features from data [9-11. With respect to military healthcare, inputs to ML algorithms include field operation traits and sometimes combat outcomes of interest [9-11].

III. METHODOLOGY

In this section, the context in which an intelligent management platform for the health monitoring of soldiers in a military operational field, namely 'i-SHM', is envisaged to operate, is presented in detail.

As mentioned previously, the manner in which the military commanders remain engaged with the battlefield and monitor the soldiers, as well as the central system receives data associated with the soldier's health status and environment, can change from time to time. Thus, an cognitive management functionality is required in order to be able to adapt, in a real-time manner, the soldier's level of health status due to change of associated variables of interest [12].

In this respect, the proposed 'i-SHM' functionality is aimed to interact, on behalf of the commander/user, with all the available levels of health status, being able to make intelligent health predictions, by taking into account the commander's/user's request, the available set of input features-parameters, the policies, and previous knowledge. Communication among the commander/user and the proposed 'i-SHM' functionality can be performed through the existence of a well-designed interface system [13].

In a more detailed analysis, as shown in Fig. 1, 'i-SHM' functionality uses as input:

- parameters associated with historical health data – medical statistics, i.e., data on specific medical conditions within the military healthcare context,
- biological parameters, i.e., data relating to the soldier's body that impact on health (e.g., blood pressure, heart rate, ecg rate, body temperature, etc.), and
- physical environment contextual parameters, i.e., actual data associated with environmental aspects in the battlefield (e.g., humidity, weather condition, dust density, etc)

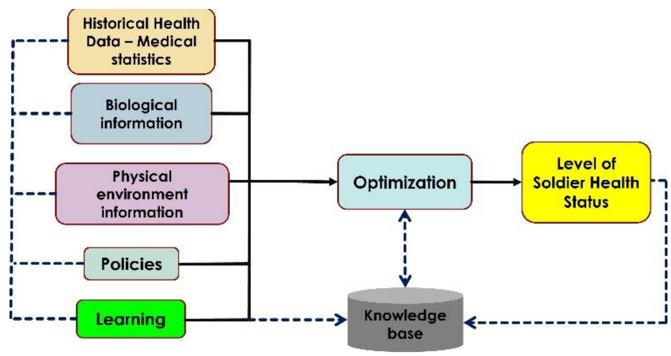


Fig. 6. i-SHM functional architecture

Furthermore, 'i-SHM' functionality uses as input two sets of overarching policies towards the importance of associated parameters. The first set of policies reflects commander's/user's preferences towards a set of predefined biological parameters. To do so, commander/user needs to specify the importance he/she attributes to each of those parameters. This is achieved by attributing each one of the predefined biological parameters with a certain weight value between 0 and 1, with 0 implying that the biological parameter has the lowest importance for the commander/user and 1 pointing at the highest importance. Of course, it is possible that some biological parameters could have the same weight value for the commander/user. For instance, a commander/user may consider blood pressure and heart rate equally important.

In addition, the second set of policies, which are associated with the physical environment contextual parameters, is established by the operator system, based on real-time information extracted from infrastructure's sensors. In a similar way like previously, the operator system attributes each of the physical environment contextual parameters with a certain weight value between 0 and 1. The value 0 implies that the physical environment contextual parameter has the lowest importance for the operator system, whereas the value 1 pointing at the highest importance. It should be noted that as physical environment contextual parameters can change rapidly from time to time, the central operator system may need to adapt frequently their respective weight values.

Identify applicable funding agency here. If none, delete this text box.

Based on the above, all combinations of input data (including historical health data – medical statistics, physical environment contextual parameters, policies, learning scheme) with the related decisions are kept in an appropriately structured database. On this way, whenever a specific input situation is encountered, 'i-SHM' performs an initial search in the appropriate part of the (classified) database, so as to check whether a similar situation has been encountered also in the past and how it has been tackled. In affirmative, the algorithm does not need to run and the previous decision, through the exploitation of knowledge and experience, is applied again.

Otherwise, 'i-SHM' functionality and its algorithm need to run and reach a decision, through the process described in the following. For example, since sensor data fusion systems provide 'i-SHM' with input physical environment information continuously, the algorithm needs to run only when something changes, i.e., when the present input situation has not been addressed before. In this respect, parameter changes are adapted fast and successfully, valuable time is saved and the overall complexity of the proposed platform is reduced.

What could be considered in this stage is that the above methodology presents the theoretical basis of the 'i-SHM' functional architecture. As future activities in this area one could consider the proposed 'i-SHM' to be implemented and verified on a practical basis, by using dataset examples, system mockups, pilot studies, etc. In this context, sufficiently large and realistic field tests are necessary by evaluating important performance metrics like computational efficiency and scalability.

IV. CONCLUSIONS

In general, cognitive computing in military healthcare is a hot and promising topic. Both academics and industry are making big efforts to improve the performances of current systems and to propose novel health decision-making solutions. In this context, health predictions within the military healthcare setting is of crucial importance, as it is associated with significantly worse outcomes in soldiers health.

In the light of the above, the proposed 'i-SHM' is introduced, targeted at cognitively managing, quickly and efficiently, the commander's/user's request, real-time crucial information associated with the physical environment, historical health data and biological determinants, policies, and previous knowledge turned into experience. This conceptual framework aims to: (a) make more informed health decisions at the point of care within the military healthcare setting, (b) heighten the confidence of military commanders and healthcare professionals by leveraging evidence-based recommendations backed by deep knowledge towards soldier health predictions, and (c) identify the strength of key factors to help make critical decisions towards soldier health predictions fast, clinically consistent and accessible. In this context, the proposed 'i-SHM' functionality aims to enhance the monitoring and tracking of soldiers' health, minimize response time in case of a medical emergency, and provide immediate care to those soldiers who need it.

Last, this work opens the gates to a series of exciting work areas. First, interoperation and communication issues among different sources and devices shall be explored due to their crucial importance towards effective soldier health predictions. Second, advanced AI-enabled solutions and cognitive decision making algorithms should be implemented in the proposed 'i-SHM' conceptual framework for making life-critical health decisions and predicting adverse outcomes before they happen, better manage highly complex situations, and ultimately allow military clinicians to spend less time analyzing data and more time harnessing their experience and human touch in delivering care. Finally, what could also be investigated is to understand aspects that are likely to maximize adoption of cognitive health management systems in military healthcare domain.

REFERENCES

[1] V. Patel, N. Yeware, B. Thombre and A. Chopde, "Soldiers Health Monitoring and Position Tracking System," 2024 IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS), Bhopal, India, 2024, pp. 1-4, doi: 10.1109/SCEECS61402.2024.10482038.

[2] J. J. Kang, "A Military Human Performance Management System Design using Machine Learning Algorithms," 2021 31st International Telecommunication Networks and Applications Conference (ITNAC), Sydney, Australia, 2021, pp. 13-18, doi: 10.1109/ITNAC53136.2021.9652140.

[3] S. Bandopadhaya, R. Dey, and A. Suhag, "Integrated healthcare monitoring solutions for soldier using the internet of things with distributed computing", Sustainable Computing: Informatics and Systems, vol. 26, 2020, 100378, https://doi.org/10.1016/j.suscom.2020.100378.

[4] L.C. Main, L.T. McLoughlin, S.D. Flanagan, M.C. Canino, and S. Banks, "Monitoring cognitive function in the fatigued warfighter: A rapid review of cognitive biomarkers", Journal of Science and Medicine in Sport, vol. 26, 2023, pp. S54-S63, https://doi.org/10.1016/j.jsams.2023.04.009.

[5] G. Dimitrakopoulos and M. Logothetis, "Intelligent Management Functionality for Emergency Medical Applications Based on Cognitive Networking Principles," Journal of Software Engineering and Applications, vol. 1, 2011, pp. 23-36.

[6] P. Kumar, G. Rasika, V. Patil, and S. Bobade, "Health Monitoring and Tracking of Soldier Using GPS," International Journal of Research in Advent Technology, vol. 2, no. 4, 2014, pp. 291-294.

[7] S. Sharma, S. Kumar, A. Keshari, S. Ahmed, S. Gupta and A. Suri, "A Real Time Autonomous Soldier Health Monitoring and Reporting System Using COTS Available Entities," 2015 Second International Conference on Advances in Computing and Communication Engineering, Dehradun, India, 2015, pp. 683-687, doi: 10.1109/ICACCE.2015.84.

[8] A. Muqeet, Mohd & Q. Mohammed, "An IoT based patient monitoring system using raspberry Pi", International Conference on Computing Technologies and Intelligent Data Engineering, Kovilpatti, India, January 2016, pp. 1-4.

[9] D. Kumar and S. Repal, "Real Time Tracking and Health Monitoring of Soldiers Using ZigBee Technology: a Survey", International Journal of Innovative Research in Science, Engineering, and Technology, vol. 4, no. 7, July 2015, pp. 5561-5574.

[10] G. Raj and S. Banu, "GPS Based Soldier Tracking and Health Indication System with Environmental Analysis," International Journal of Enhanced Research in Science, Technology, and Engineering, vol. 2, no. 12, December 2013, pp. 46-52.

[11] S. Roy, T. Meena, and S.J. Lim, "Demystifying Supervised Learning in Healthcare 4.0: A New Reality of Transforming Diagnostic Medicine. Diagnostics, vol. 12, no 10, 2022, https://doi.org/10.3390/diagnostics12102549.

[12] M. S. Jassas, A. A. Qasem and Q. H. Mahmoud, "A smart system connecting e-health sensors and the cloud," 2015 IEEE 28th Canadian Conference on Electrical and Computer Engineering (CCECE), Halifax, NS, Canada, 2015, pp. 712-716, doi: 10.1109/CCECE.2015.7129362.

[13] A. Gondalia, D. Dixit, S. Parashar, V. Raghava, A. Sengupta, V. Raja Sarobin, IoT-based Healthcare Monitoring System for War Soldiers using Machine Learning, Procedia Computer Science, vol. 133, 2018, pp. 1005-1013, https://doi.org/10.1016/j.procs.2018.07.075.

Evaluating AI-Powered Predictive Analytics for Public Transport Demand in Dubai

Sara Ismail Dhalam
Rochester Institute of Technology, Dubai
Dubai, UAE
Sara-dhalam@outlook.com
ORCID: 0009-0006-2039-2269

Ayman Ibrahim
Rochester Institute of Technology Dubai
Dubai, UAE
ayman.ibrahim@rit.edu
ORCID: 0009-0007-5477-515

Abstract—Rapid urbanization has made intelligent transportation systems crucial for facing congestion, improving safety, and enhancing efficiency. This study presents a preliminary evaluation of AI-powered predictive analytics applied to open datasets from Dubai's Roads and Transport Authority (RTA), focusing on multimodal transport modes including metro, tram, marine, and buses. Using SAS Viya's machine learning tools for data analysis was conducted to explore the potential of interpretable AI models for predicting public transport demand and optimizing traffic flow. As this work remains in its early stages, future research will involve full model deployment and testing. This research supports Dubai's vision for a more intelligent, sustainable transport network.

Keywords—Intelligent transportation systems (ITS), AI-driven predictive analytics, public transport demand, reduce congestion, optimize traffic flow.

I. INTRODUCTION

Traffic congestion, driven by rapid population growth, economic development, and rising vehicle ownership, remains a significant challenge in cities worldwide. In Dubai, these elements have exerted considerable pressure on the transportation system, even with substantial investments in public transport. Issues such as traffic signal inefficiencies and reliance on private vehicles lead to increased delays and pollution. To address these challenges, Dubai has implemented Intelligent Transportation Systems (ITS) and developed demand management programs as essential approaches to reduce congestion and encourage public transport usage[1]. Studies estimate that traffic congestion results in yearly losses of around 1.5 billion AED due to wasted fuel and decreased productivity[2]. Additionally, research indicates that potential cumulative losses could reach $1.8 billion if proactive infrastructure investments had not been made [3].

This study aims to identify the benefits and limitations of AI-driven analytics in enhancing the efficiency and reliability of the public transport system. It evaluates AI predictive analytics for RTA open datasets, focusing on public transport demand across various modes. Using the SAS Viya and SAS optimization platform, we explore models to optimize resource allocation[4]. Future research will involve model deployment and evaluations to support Dubai's sustainable transport goals.

II. LITERATURE REVIEW

The extensive literature on intelligent traffic management systems focuses on technologies that improve traffic flow, reduce congestion, and enhance safety. This section will discuss four main topics: real-time traffic management, deep learning and density prediction, machine learning models for autonomous vehicles, and Optimization in Public Transport Systems.

A. Real-Time Smart Traffic Management

Recent studies have introduced IoT-based traffic management systems that use real-time sensors and video data to adjust signals dynamically during congestion. One approach employs VANETs and mobile agents to identify congestion hotspots [5], while another leverages IoT-cloud models for lane-specific signal control [6]. Deep learning methods like YOLO have also been integrated for real-time video analytics to detect abnormal events and enhance urban safety [7]. Additionally, Intelligent Transportation Systems (ITS) have shown effectiveness in improving traffic flow, as demonstrated using Variable Speed Limits (VSL) on major UAE roads, which resulted in reduced delays, travel times, and stops [1].

B. Deep Learning Models and Density Prediction

One notable study introduced DeepCrowd, a deep convolutional LSTM model designed to predict crowd density across fine-grained spatial grids, demonstrating strong performance in spatio-temporal forecasting tasks [6]. Similarly, Spatio-Temporal Convolutional Neural Networks (ST-CNN) were developed to capture complex urban mobility patterns by modeling spatial closeness, temporal periodicity, and long-term trends in mobility data [9]. Furthermore, another study employed Self-Supervised Contrastive Learning techniques—commonly used in fields such as computer vision (CV) and natural language processing (NLP) to extract meaningful traffic features without relying on labeled data, thereby improving the scalability and generalizability of traffic prediction models [10].

C. Machine Learning Models in Autonomous Vehicle Adoption

Recent studies have demonstrated the integration of AI, IoT, and autonomous vehicle technologies to enhance urban delivery systems and optimize traffic signal control through real-time data analytics, contributing to reduced carbon emissions and improved urban mobility [11]. In autonomous vehicle adoption, machine learning models such as Naïve Bayes, Random Forest, and Fuzzy Logic were employed to predict user acceptance. Fuzzy Logic outperformed the others, particularly in modeling adoption factors such as safety, trust, security, cost, ethics, and privacy [12]. Random Forest also showed strong performance, highlighting the effectiveness of ensemble methods in both behavioral modeling and traffic forecasting.

A Dubai-based study applied multiple ML models—YOLO, Random Forest, XGBoost, and Graph Neural Networks (GNN)—to improve traffic prediction and congestion classification [7]. Random Forest was particularly effective in forecasting congestion, while Support Vector Machines (SVM) and k-Nearest Neighbors (k-NN) achieved high classification accuracies of 99.9% and 98.6%, respectively. Long Short-Term Memory (LSTM) networks further enhanced the identification of recurring and non-recurring congestion patterns, reaching an accuracy of 98.73% [7].

D. Optimization in Public Transport Systems

While traffic prediction has advanced, public transport in urban mobility demands systemic optimization. Here, evolutionary algorithms play an essential role by introducing MOPEA (Multi-objective Evolutionary Algorithm), balancing conflicting goals like minimizing travel time and maximizing operator profits [5]. Additionally, AHP and entropy combined with TOPSIS models have been applied to rank public transport network optimization strategies based on multiple objective functions [6]. Moreover, one study designed an ACO-based approach to adjust existing bus lines without re-planning entire networks, providing a realistic solution for large, congested cities [7].

Research Gap:

Despite the growing interest in using artificial intelligence (AI) to improve urban mobility, there remains a significant gap in transparent, interpretable, and deployable AI models for predicting and forecasting public transport demand, particularly for metro and tram systems in complex cities like Dubai. Many existing studies focus mainly on road traffic or bus systems, overlooking rail-based transport and the diverse travel behaviors it involves. Additionally, most AI models used in research rely on complex algorithms, including deep learning and ensemble methods, that are difficult to interpret and require high computational resources. Finally, the lack of transparency and explainability makes them hard to trust for decision-makers, limiting their usefulness in policy, operations, and large-scale implementation.

Contribution:

This study proposes a novel predictive framework for estimating demand across multimodal transportation networks in Dubai, leveraging advanced machine learning techniques deployed on the SAS Viya, an advanced analytics platform. It helps cities prioritize investments, such as expanding the metro or tram lines, based on demand predictions [4]. The study aims to identify the benefits and limitations of AI-driven analytics in enhancing the efficiency and reliability of the public transport system. Expected outcomes include optimized traffic flow, reduced congestion, and improved transport efficiency, contributing to a more sustainable and user-friendly system for residents and visitors. Furthermore, utilizing more straightforward machine learning algorithms for data analysis might produce dependable results without requiring costly tools and specialized staff. While numerous studies have focused on bus optimization and road-related problems, this research will examine the metro and tram networks.

III. METHODOLOGY

This section presents the theoretical framework underpinning the methods employed in this project, including the process flowchart, model development stages, research design tools, and data sources used to evaluate the effectiveness of AI-powered predictive analytics for forecasting multimodal transport demand in Dubai.

The research follows a four-phase methodology, as shown in Fig. 1:

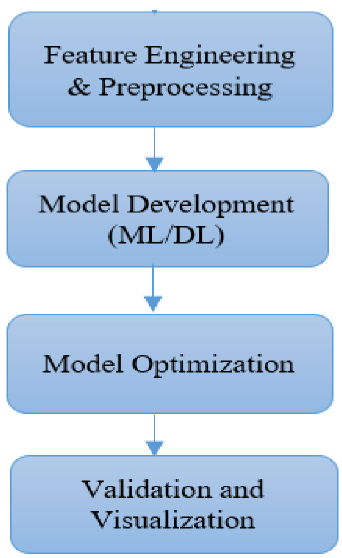


Fig. 2. Methodology Phases

A. Data preparation and feature engineering

Data preparation represents a foundational phase in this study, ensuring high-quality, model-ready inputs using SAS Data Preparation within the SAS Viya platform [4]. This process involved acquiring publicly available datasets from the Dubai Roads and Transport Authority (RTA) via the Dubai Pulse Open

Data Platform [13]. The datasets encompass multimodal transport records, including metro, tram, marine, and bus service usage data. These data are regularly updated and include critical features such as route names, transaction types, timestamps, and passenger counts. Retrieval is performed by navigating to the "Transport" category on the platform, selecting the desired transport mode (e.g., Metro Ridership, Tram Usage), and downloading the data in CSV or API format.

B. Demand Prediction Modeling

Demand prediction is the second phase of the process, focusing on forecasting the most popular transportation mode using machine learning algorithms, as shown in Table 1.

TABLE I: Machine learning algorithms

Algorithm	**Applications and justifications**
Gradient Boosting	Machine learning models, including boosting, were used for predictive analytics and route optimization. Minimizing loss functions can effectively forecast baseline demand [8].
Random Forest	The Random Forest classifier achieved the highest accuracy of 84% in traffic congestion classification due to their robustness against overfitting, ability to handle high-dimensional data, and clear interpretability [9].
SVM	SVM showed moderate accuracy (64–79%) in classifying traffic congestion, making it useful for traffic state modeling, particularly in high-dimensional spaces and with small to medium datasets [9].
clustering (k-means)	Clustering methods, including K-means, were discussed for identifying High-demand hotspots and optimizing traffic flows [10]
Forecasting	Addressed overfitting using contrastive learning, as it outperformed ARIMA, Ridge Regression, SVR, and even Gradient Boosting (GBRT) [11]. Captures spatiotemporal dependencies, attention, and outperformed multiple state-of-the-art methods on large citywide datasets [12].

C. Optimization

The third phase is dedicated to optimizing and enhancing the results using SAS optimization techniques, which can be used for scenario-based decision making, what-if analysis, and resource allocation under uncertainty [13].

D. Validation

The final phase involves validating the model's results through model comparison and assessment using SAS Visual Analytics [4]. This phase includes evaluating model performance using key validation metrics such as Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and F1-score. Additionally, classification performance is assessed through the confusion matrix, misclassification rate, and user experience (UX) metrics. These analyses ensure that the selected model meets the desired accuracy and reliability standards for deployment.

IV. PRELIMINARY RESULTS

A preliminary analysis was conducted to visualize the trends in multimodal public transport demand in Dubai over the past five years. Fig.2 presents a line chart illustrating these trends, revealing a general increase in demand across most transport modes, excluding a dip in 2022, primarily attributed to the impact of the COVID-19 pandemic.

The Dubai Metro exhibits the highest passenger demand, increasing ridership from approximately 203 million in 2019 to over 260 million in 2023. This upward trend underscores the city's firm reliance on metro infrastructure. Similarly, public transport buses have shown a robust recovery post-pandemic, with ridership growing from 95 million in 2020 to 173 million in 2023, reflecting their vital role in serving broader geographic areas.

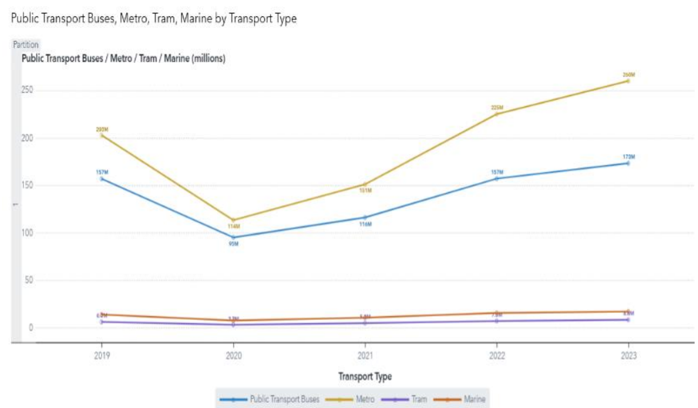


Fig. 3. Line chart - Public transport by type

In contrast, the Metro Green Line shows significantly lower passenger volumes than anticipated. Moreover, despite having comparable infrastructure and connectivity potential, the Dubai Tram recorded the lowest usage levels, as depicted in Fig.3.

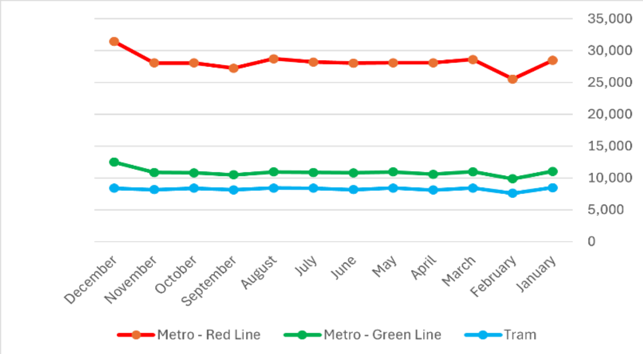


Fig. 4. Line chart – metro vs tram

V. CONCLUSION

This study showcases the potential of AI-driven predictive analytics to enhance traffic management and improve various transit modes in Dubai, such as metro, tram, marine, and bus services. The research demonstrates the benefits of interpretable AI models in urban transport planning by utilizing publicly available data from the Dubai Roads and Transport Authority (RTA) and SAS Viya's advanced machine learning capabilities.

Key findings reveal that AI-driven demand forecasting can significantly reduce traffic congestion by offering actionable insights for infrastructure development. The study emphasizes optimizing traffic flow and improving transport efficiency to create a more sustainable, user-friendly system for residents and visitors.

The next research phase will focus on applying these models to real-world datasets, including comparative testing and model enhancement, to adapt to Dubai's evolving urban mobility landscape. The goal is to support decision-makers in prioritizing funding and creating a more intelligent and sustainable transportation network.

REFERENCES

[1] Y. El-Hansali et al., "Smart Dynamic Traffic Monitoring and Enforcement System," Computers, Materials & Continua, vol. 67, no. 3, pp. 2797–2806, 2021, doi: 10.32604/cmc.2021.014812.

[2] S. Shahbandari, "Traffic congestion costs more than Dh700,000 per kilometre in Dubai." Accessed: Apr. 25, 2025. [Online]. Available: https://gulfnews.com/uae/transport/traffic-congestion-costs-more-than-dh700000-per-kilometre-in-dubai-1.1452783

[3] Gulfnews, "RTA projects make Dubai traffic freer than leading cities with similar population." Accessed: Apr. 25, 2025. [Online]. Available: https://gulfnews.com/uae/transport/rta-projects-make-dubai-traffic-freer-than-leading-cities-with-similar-population-1.75275274

[4] "SAS Help Center: SAS Help Center: Welcome." Accessed: Apr. 29, 2025. [Online]. Available: https://documentation.sas.com/doc/ar/helpcenterwlcm/1.0/home.htm

[5] Hou Lin, Li Wen-yong, Ma Li, and Xu Jian-min, "Public transport network optimization based on a Multi-objective Optimization Problems Evolutionary Algorithm," in 2009 Chinese Control and Decision Conference, Guilin, China: IEEE, Jun. 2009, pp. 4408–4412. doi: 10.1109/CCDC.2009.5192410.

[6] Li Xiaowei, "TOPSIS model for urban public transport network optimization based on AHP and entropy," in Proceedings 2011 International Conference on Transportation, Mechanical, and Electrical Engineering (TMEE), ChangChun, China: IEEE, Dec. 2011, pp. 75–79. doi: 10.1109/TMEE.2011.6199151.

[7] W. Hu, C. Wang, and X. Zuo, "An Ant Colony Optimization based Approach to Adjust Public Transportation Network," in 2019 IEEE Congress on Evolutionary Computation (CEC), Wellington, New Zealand: IEEE, Jun. 2019, pp. 2575–2580. doi: 10.1109/CEC.2019.8790117.

[8] B. M. Mohsen, "AI-Driven Optimization of Urban Logistics in Smart Cities: Integrating Autonomous Vehicles and IoT for Efficient Delivery Systems," Sustainability, vol. 16, no. 24, p. 11265, Dec. 2024, doi: 10.3390/su162411265.

[9] D. Impedovo, F. Balducci, V. Dentamaro, and G. Pirlo, "Vehicular Traffic Congestion Classification by Visual Features and Deep Learning Approaches: A Comparison," Sensors, vol. 19, no. 23, p. 5213, Nov. 2019, doi: 10.3390/s19235213.

[10] R. E. Al Mamlook, M. Zahrawi, H. Gharaibeh, A. Nasayreh, and S. Shresth, "Smart Traffic Control System for Dubai: A Simulation Study Using YOLO Algorithms," in 2023 IEEE International Conference on Electro Information Technology (eIT), Romeoville, IL, USA: IEEE, May 2023, pp. 254–264. doi: 10.1109/eIT57321.2023.10187271.

[11] Y. Song, "Effective Traffic Prediction with Self-Supervised Contrastive Learning," in 2022 IEEE 8th International Conference on Computer and Communications (ICCC), Chengdu, China: IEEE, Dec. 2022, pp. 2204–2209. doi: 10.1109/ICCC56324.2022.10066048.

[12] R. Jiang et al., "DeepCrowd: A Deep Model for Large-Scale Citywide Crowd Density and Flow Prediction," IEEE Trans. Knowl. Data Eng., pp. 1–1, 2021, doi: 10.1109/TKDE.2021.3077056.

[13] "SAS Optimization." Accessed: May 05, 2025. [Online]. Available: https://www.sas.com/en_ae/software/optimization.html

‘PlanItAll’App: Transforming Event Planning with Augmented Reality, AI, and Integrated Service Solutions

Abeer Alawi
Zayed University
Abu Dhabi, United Arab Emirates
201506498@zu.ac.ae

Fatima Al Dhaheri
Zayed University
Abu Dhabi, United Arab Emirates
202131552@zu.ac.ae

Ahmed Shuhaiber
Zayed University
Abu Dhabi, United Arab Emirates
Ahmed.Shuhaiber@zu.ac.ae

Abstract— PlanItAll is an innovative event management platform designed to simplify the planning process by bringing together multiple vendors and advanced tools into one seamless interface. Users can manage every detail of their event, like venues, catering, décor, entertainment, and more, all from a single platform. Enhanced with Augmented Reality (AR), PlanItAll allows users to visualize and adjust event layouts in real time, giving them a clear, interactive preview of their space before the event day. The platform also includes smart suggestion tools that adapt to user preferences, helping them discover relevant services and optimal setups based on their choices and past selections. By centralizing communication and coordination, PlanItAll reduces the risk of miscommunication between vendors and makes handling last-minute changes easier. It delivers a unified, intuitive, and immersive planning experience that not only simplifies the planning process but also enhances creativity and engagement.

Keywords—Event Management, Service Integration, Augmented Reality, AI Customization, User Engagement, Event Planning. Introduction (Heading 1)

I. INTRODUCTION

Event planning is a complex process that involves coordinating multiple vendors, such as venues, florists, caterers, lighting experts, and entertainers, which often leads to miscommunication, scheduling conflicts, and last-minute errors, making the entire experience stressful and inefficient. The global event management industry is massive, valued at over one trillion dollars, and experts predict it will grow significantly in the coming years. This growth highlights the increasing demand for tools that can handle the challenges of organizing events efficiently. However, many current tools, such as Eventbrite, Cvent, and Trello, have notable weaknesses. These platforms often lack proper integration with other systems, struggle to support real-time communication, and do not offer the flexibility needed for different types of events. As a result, event planners face issues like scheduling errors, miscommunication, and higher costs, all of which make their work harder.

The core problem is that today’s event planning tools do not provide a complete, easy-to-use solution. This leads to poor coordination and events that fail to meet the expectations of those involved. There is a clear need for a better system—one that is simple, efficient, and able to adapt to the diverse needs of event organizers. designed to streamline event planning by integrating Augmented Reality (AR) and Artificial Intelligence (AI). With AR, users can visually arrange and modify event details in real-time, ensuring precise setups before finalizing their plans. Meanwhile, AI provides intelligent recommendations, optimizes layouts, and enhances vendor coordination, reducing human error and simplifying decision-making. By bringing all event services into a single system, PlanItAll improves efficiency, enhances communication, and transforms event planning into a smoother and more user-friendly experience.

This work contributes to event management technology in several ways. First, it introduces PlanItAll as a unified platform that overcomes the shortcomings of existing tools. Second, it provides insights from interviews with event planners, offering a deeper understanding of their needs. Third, it creates a strong testing approach to evaluate event planning systems, ensuring they are both practical and reliable. Finally, PlanItAll is built to grow and adapt, supporting various event types and setting the stage for future improvements in the field.

II. LITERATURE REVIEW

To comprehend the scope, the literature related to the scope was reviewed. This section presents a brief summary of multiple articles as follows:

A. Wedding Planner Application Using Ar

The WedDream project introduces a wedding planner app with Augmented Reality (AR), allowing users to design their dream wedding venue while managing tasks like budgeting, guest lists, and checklists. The app improves on existing wedding planners like The Big Day, LadyMarry, and MyWed by combining standard planning tools with AR venue visualization, helping users customize their space before the big day [1]. While WedDream focuses only on weddings, PlanItAll takes this concept further by offering a complete event planning experience for all types of events. Unlike WedDream, which mainly provides AR venue design, PlanItAll integrates AI to personalize event recommendations, optimize layouts, and simplifies vendor selection. It also allows users to not just plan but book all services in one place, creating a fully interactive and stress-free event planning process. By combining AR visualization with AI-powered automation, PlanItAll upgrades event management into a highly personalized experience.

B. Augmented Reality In Event Promotion

The study presents an Augmented Reality (AR) application designed to promote events in urban environments, creating an immersive and interactive experience for users. Developed using Django Rest Framework (DRF) for backend services and Unity for AR functionalities, the app allows users to register, view events, interact with virtual elements, and provide feedback on attended events. The study highlights the app's potential for enhancing urban engagement, making event discovery more dynamic and engaging. However, the focus remains on event promotion rather than comprehensive event planning and execution [2]. PlanItAll expands this concept by integrating AR not just for event discovery but for the entire event planning process. Instead of simply viewing events, users can design, customize, and visualize their own event spaces in real time using AR, ensuring that layouts, decorations, and seating arrangements are planned to their preferences before booking. Additionally, AI-driven automation personalizes vendor recommendations, optimizes event layouts, and streamlines service coordination, making event planning more efficient and reducing last-minute changes. By combining AR-powered visualization with AI-driven customization, PlanItAll offers a more advanced, interactive, and user-centered approach to event management, going beyond promotion to provide a complete end-to-end event planning experience.

C. Ai Based Event Management Web Application

The AI-Based Event Management web application helps users search for, register, and receive event recommendations based on their interests. It categorizes events like webinars, workshops, and conferences, making it easier for users to find relevant opportunities [3]. However, its AI functionality is limited to basic recommendations and recent event tracking. PlanItAll takes AI a step further by not just recommending events but actively assisting in the entire event planning process. Instead of simply suggesting events, PlanItAll's AI analyzes user preferences, past events, and space availability to generate personalized event layouts, themes, and vendor suggestions. AI-driven automation also helps users optimize budgets, suggest alternative service providers, and ensure real-time adjustments based on changing needs. Combined with Augmented Reality (AR), PlanItAll allows users to visualize and modify their event setups before confirming them, ensuring that AI-generated plans align with their vision. This upgraded AI functionality moves beyond passive recommendations to active event planning and execution, making PlanItAll a more intelligent, adaptive, and user-centered event management platform.

D. The Impact Of Covid-19 On Event Management Industry

The COVID-19 pandemic caused massive disruptions in the event management industry, leading to widespread cancellations and financial losses. The industry, which thrives on in-person gatherings, suffered severely due to lockdowns, travel restrictions, and safety concerns. Many companies shifted to virtual events, but these lacked the engagement and interaction of physical gatherings [5]. The COVID-19 pandemic forced strict lockdowns that lasted anywhere from a few months to over a year in some regions, drastically limiting large gatherings. Despite restrictions, many people still held small, intimate events such as weddings, but they struggled to find proper planning services to assist them. Traditional event management platforms were not designed to accommodate at-home or restricted-space events, leaving individuals to handle everything on their own. PlanItAll offers a solution for these situations by allowing users to organize and visualize their events within their own space using Augmented Reality (AR). With AR, users can integrate elements like seating arrangements, decorations, and lighting into their environment, previewing and adjusting them in real-time before finalizing the setup. Additionally, PlanItAll connects users with service providers who bring catering, decorations, and other event essentials directly to their location, ensuring an ideal experience even under restrictions. This innovation ensures that important moments can still be celebrated with proper planning and professional support, regardless of external limitations.

E. Eventx, Event Revenue Generator

EventX specializes in virtual and hybrid events, offering tools for live streaming, virtual networking, and attendee engagement. It is mainly used for professional events like conferences and corporate gatherings, providing features like registration, ticketing, and virtual booths. However, EventX is limited to virtual experiences and lacks AR-powered event visualization, AI-driven customization, and direct vendor booking, making it less suitable for planning in-person and diverse event types [6]. PlanItAll goes beyond virtual engagement by integrating AI and AR for all types of events, from professional conferences to personal celebrations like weddings and graduations. Unlike EventX, it lets users design and modify event spaces in real time using AR, ensuring layouts meet their vision before finalizing. AI streamlines vendor selection, optimizes event setups, and personalizes recommendations, making planning more intuitive and efficient. With a built-in booking system for venues, caterers, and decorators, PlanItAll offers a fully interactive event planning experience that EventX does not provide.

III. METHODOLOGY

A. Planning

The PlanItAll project followed a step-by-step development process called the Systems Development Life Cycle (SDLC). The first step involved gathering opinions from people who plan events (e.g., brides, parents, event planners). Participants shared challenges like poor vendor communication, manual coordination, and stress from last-minute changes. By grouping similar issues (e.g., vendor management, communication gaps), we identified key needs: a single platform to connect clients and vendors, tools to visualize event layouts, and features to simplify decision-making. These findings directly shaped PlanItAll's design.

B. Requirements Analysis

We confirmed the platform could work with common tools like laptops and phones. We also listed what the system must do (functional requirements) and how well it must perform (non-functional requirements).

The development of PlanItAll followed the Software Development Life Cycle (SDLC) to ensure a structured approach. It began with user research through interviews with four stakeholders: a bride-to-be, a mother organizing a birthday party, a professional event planner, and a businessman managing corporate events. These interviews revealed recurring challenges, including poor communication with vendors, stress due to managing multiple services, and limited customization. The insights guided the design of a platform that could centralize event planning, simplify vendor coordination, and enhance user experience with modern technologies.

- **Data Availability Statement (DAS)**

The primary data supporting this study were collected through semi-structured interviews with event planning stakeholders (e.g., individual planners, event professionals). The study received ethical approval from Zayed University's Institutional Research Committee to ensure compliance with ethical standards. Due to confidentiality agreements and the personal nature of the interview content, the full transcripts are not publicly available. However, anonymized excerpts may be shared upon reasonable request to the corresponding author, subject to participant and institutional approval.

Based on this user feedback, the system requirements were divided into two main categories:

- **Functional Requirements**: The platform needed to support user registration, event type selection, service customization, AR-based layout visualization, AI-generated vendor recommendations, booking and payment processing, and notification/task tracking.
- **Non-Functional Requirements**: The system was expected to deliver fast response times, intuitive usability, robust data security, high availability, and a visually appealing and adaptable interface.

C. System Design

The system design was developed using Lucidchart to create detailed Data Flow Diagrams (DFDs). These diagrams clearly show the flow and processing of data within the application, outlining how different components of the system interact as shown in Figure 1.

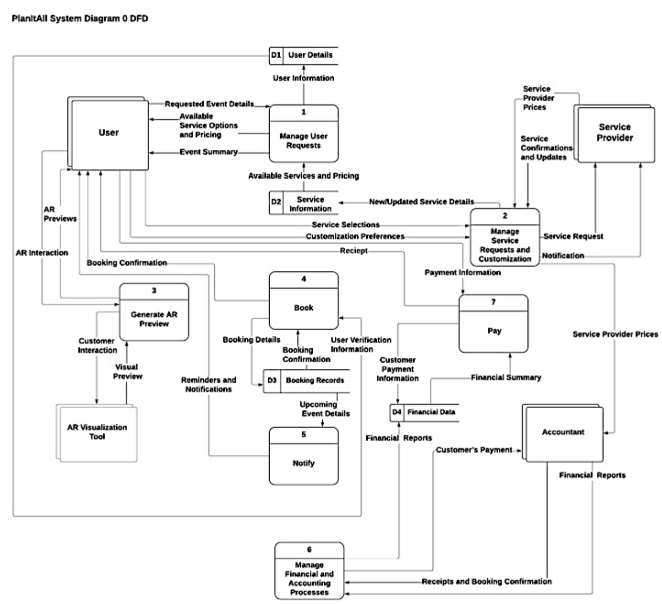


Figure 1. PlanItAll Data Flow Diagram (Level 0)

D. Prototype Development

A working prototype of the "PlanItAll" platform was developed using Wix, based on the system design previously established. Wix was chosen for its user-friendly interface and flexibility, allowing for rapid prototyping and easy adjustments. This prototype acts as an initial version of the platform, which can be extensively tested and evaluated for usability and functionality. This phase is crucial for gathering initial user feedback and making necessary improvements to enhance the overall experience.

The homepage welcomes users with "Creating Unforgettable Memories – One Event at a Time," and features a "Start Planning" button that takes users directly to the services page to begin planning. When users visit the services page, they can choose their event type as shown in

Figure 2., whether it's a private party, corporate event, or wedding. Each category includes detailed service options to help users find exactly what they need. Users can book venues, catering, décor, entertainment, and photography for private parties like birthdays and graduations. Corporate events offer venue rental, speaker arrangements, and branding for a professional setup. Weddings provide full planning, partial assistance, or customizable packages, letting couples decide their level of involvement.

Figure 2. Services Page with "AI Generate" button at the bottom.

Once users select their event type, they are guided to the booking page, where they can browse all many types of available services, compare options, and make selections. Each service displays images, descriptions, pricing, and duration, making it easy to decide. Clicking "Book Now" takes users to a calendar view, where they can choose a date and time, review service details, and confirm their booking as shown in Figure 3.

To make planning even more interactive and stress-free, Augmented Reality (AR) lets users visualize their event setup in real time, adjusting layouts, décor, and seating before finalizing. AI customization helps by suggesting the best vendors, layouts, and themes based on user preferences, ensuring that every event is perfectly tailored. With PlanItAll, users can plan their event from start to finish in one place, with full control over every detail, making event planning more exciting, efficient, and enjoyable.

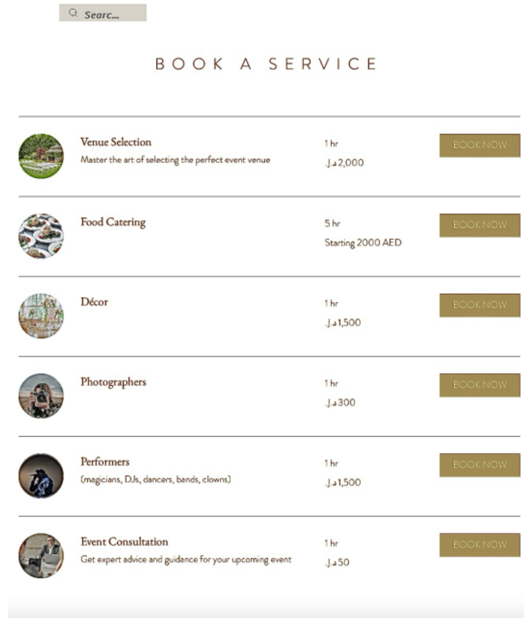


Figure3. Booking a Service Page

E. Testing Plan

To ensure the reliability, functionality, and usability of the PlanItAll platform, a comprehensive testing plan was implemented. Functional testing was carried out to confirm that each feature, including event creation, service booking, AR layout customization, and AI-driven recommendations, performed as expected. Usability testing involved observing users as they navigated the platform to identify any design or interaction issues, ensuring an intuitive experience. Performance testing evaluated system responsiveness, loading times, and the stability of AR rendering under varying user loads. Security testing focused on protecting sensitive data and ensuring the safety of payment transactions through proper encryption and secure authentication protocols. Compatibility testing verified that the platform worked seamlessly across different web browsers and on both Android and iOS devices with AR support. Finally, user acceptance testing (UAT) was conducted to assess whether the platform met users' needs in real-world planning scenarios and provided a smooth, stress-free planning experience.

IV. DISCUSSION

PlanItAll is more than just an event planning tool—it is a fully integrated platform that transforms the way events are organized, customized, and executed. Unlike other planning applications that focus on specific aspects such as venue selection or catering, PlanItAll offers a comprehensive solution that covers every stage of the planning process. By incorporating Augmented Reality (AR) and Artificial Intelligence (AI), the platform enhances user control, improves efficiency, and reduces the stress commonly associated with event planning.

One of PlanItAll's key features is AR-powered event visualization, which allows users to design and adjust their event layouts in real time. Unlike traditional planning apps that rely on lists and descriptions, this feature gives users the ability to place decorations, adjust seating arrangements, and finalize layouts by viewing them directly in their chosen venue space. This reduces uncertainty and last-minute changes, ensuring that the final setup aligns perfectly with expectations.

Another major advantage of PlanItAll is AI-driven customization, which personalizes the planning experience by learning from user preferences and suggesting tailored event elements. Instead of manually searching for vendors or browsing generic options, users receive smart recommendations for venues, decorators, caterers, and other services based on their specific needs and past choices. This saves time, optimizes decision-making, and enhances user satisfaction, making the planning process smoother and more enjoyable. Beyond its use of AR and AI, PlanItAll also stands out for its all-in-one booking and vendor management system. Unlike many platforms that focus only on planning logistics, PlanItAll allows users to directly book services within the platform, eliminating the need for multiple third-party interactions. This feature ensures seamless coordination between vendors and organizers, reducing miscommunication and delays.

While many event planning apps are designed primarily for professional event managers, PlanItAll is built to be accessible to anyone organizing an event, whether it's a wedding, graduation, conference, or private gathering. The platform allows users to actively participate in designing and finalizing their events, rather than relying entirely on external coordinators. This level of user involvement makes the planning process more interactive and engaging, giving individuals the confidence to plan successful events with ease.

I.CONCLUSION AND FUTURE WORK

This paper presented PlanItAll, a centralized event planning platform that integrates Augmented Reality and Artificial Intelligence to address the common challenges faced in organizing events. By allowing users to visualize event layouts in real time and receive intelligent vendor recommendations, the platform enhances decision-making, reduces stress, and simplifies coordination. Early development and user feedback confirm the system's potential to improve efficiency and user satisfaction in both personal and professional event planning.

Looking forward, future work will focus on enhancing the AI recommendation engine, expanding the AR feature to support more advanced spatial interactions, and developing a fully functional mobile application. Additional features such as guest management, real-time collaboration between multiple users, and deeper integration with vendor networks will also be explored. These improvements aim to make PlanItAll more scalable, intelligent, and accessible to a wider audience.

REFERENCES

[1] J. X. Leong, "Wedding planner application using augmented reality," eprints.utar.edu.my, Apr. 14, 2021. http://eprints.utar.edu.my/4262/

[2] T. Lameirão, M. Melo, and F. Pinto, "Augmented Reality for Event Promotion," in Computers, vol. 13, no. 342, 2024. [Online]. Available: https://doi.org/10.3390/computers13120342

[3] P. S. Hada, Y. Hada, B. Hada, and P. Hada, "AI based Event Management Web Application," 2022 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COM-IT-CON), Faridabad, India, 2022, pp. 562-566, doi: 10.1109/COM-IT-CON54601.2022.9850551.

[4] D. A. C. Wieland et al., "Augmented and virtual reality in managing B2B customer experiences," Industrial Marketing Management, vol. 119, pp. 193–205, 2024, doi: 10.1016/j.indmarman.2024.04.007. Available: https://research-ebsco-com.zulib.idm.oclc.org/linkprocessor/plink?id=7df774f8-4533-32ba-8da4-398d17fb23f2. Accessed: Feb. 20, 2025.

[5] J. S. Madray, "(PDF) THE IMPACT OF COVID-19 ON EVENT MANAGEMENT INDUSTRY," ResearchGate, Aug. 2020. https://www.researchgate.net/publication/343851988_THE_IMPACT_OF_COVID19_ON_EVENT_MANAGEMENT_INDUSTRY

[6] Asana, Inc., "Asana," Mobile App, Version 2.3.0, 2025. [Online]. Available: https://asana.com/download. [Accessed: 21-Feb-2025].

Early Dementia Indicators using Feature Selection: Data and Explainability Consideration

Talib Alshehhi
School of Computer Science and Informatics
De Montfort University
Leicester, UK
P2602836@my365.dmu.ac.uk

Abstract— This paper pinpoints few key research issues surrounding the application of feature selection techniques for early dementia detection, with a focus on Alzheimer's Disease Neuroimaging Initiative (ADNI) data. Dementia, particularly Alzheimer's Disease (AD), impairs memory, language, and executive functions, necessitating early diagnosis for effective intervention. Drawing on four years of study utilizing cognitive assessments and biomarkers, this work highlights critical considerations in selecting relevant features to enhance data driven model performance and improve interpretability. Additionally, it addresses data-related challenges within ADNI datasets, emphasizing their impact on model development. By identifying these gaps, the paper aims to guide future research in the use of feature selection processes, thereby supporting clinicians in making more informed diagnostic decisions.

Keywords— Artificial Intelligence; Dementia Detection; Classification Models; Feature Selection

I. INTRODUCTION AND PROBLEM

This paper scope features selection for dementia early diagnosis in order to highlight possible research directions in this area after 4-year study using cognitive assessments and other common biomarkers. The aim is to show where artificial intelligence (AI) techniques particularly machine learning (ML) are heading in this sensitive health application. More essentially, the aim of this paper is to highlight important research considerations that should be taken when using ML such as feature selection for early dementia detection.

Dementia is associated with deficiencies in language, memory, learning, problem solving and executive functions, which interferes with individuals' daily life (Chaves et al., 2011). There are millions of people with the condition worldwide in which Alzheimer's Disease (AD) accounts for more than 60% of dementia conditions. Dementia impacts physical, psychological, social, and economic impacts in which its subtype AD tends to worsen over time therefore discovering cognitive and non-cognitive features that can help clinicians to early detect the disease is significant for an intervention, which is useful for patients and caregivers.

One of the promising approaches to detect dementia conditions is ML techniques that provide healthcare professionals with models developed from historical data to predict the disease and its level (Wharton et al., 2019). These models are usually produced by training on a population of data subjects diagnosed as cognitively normal (CN), having Mild Cognitive Impairment (MCI), or demented. Over the last two decades, ML techniques have been applied in medical applications because the diagnostic outcome is produced using models learnt objectively from data, offering a less-biased result for physicians to assess in reaching a final diagnosis (Maroco et al., 2011).

Part of the ML process is a step called feature selection, which is a process that involves identifying and a subset of relevant features from a larger dataset (Büyükkeçeci and Okur, 2022). This process aims to reduce data dimensionality, improve model performance, reduce overfitting, and enhance computational efficiency by eliminating irrelevant or redundant features (Pudjihartono et al., 2022). By focusing on the informative feature specially in healthcare applications like dementia detection, feature selection can lead to simpler and more understandable models while maintaining or even improving predictive performance (Chandrashekar and Sahin, 2014).

By unfolding features ranking and their correlations from dementia dat and presenting them as simple to understand knowledge clinicians and other decision makers can determine key features and their correlations besides removing redundant features.

II. BACKGROUND ON FEATURE SELECTION AND DEMENTIA

A. Dementia Diagnosis

Dementia diagnosis involves a structured and lengthy process to establish clinical diagnosis thereby few diagnostic frameworks have been developed by healthcare professionals including the DSM-5 (American Psychiatric Association, 2013), ICD-11 (Tyrer et al., 2011), and criteria developed by the National Institute on Aging (NIA) and Alzheimer's Association (AA) (Jack Jr et al., 2012). This comprehensive approach ensures accurate classification of dementia from other cognitive and degenerative conditions.

The DSM-5 classifies dementia under major and mild neurocognitive disorders (NCDs). The diagnosis starts with identifying cognitive decline in cognitive domains such as memory, executive function, learning, language, perceptual-motor skills, or social cognition. The decline is assessed through patient medical history, informant reports, and cognitive assessment using standardized methods. Major NCDs, or dementia, require major interference with independence in daily life activities, while mild NCDs involve lesser decline. The ICD-11 complements the DSM-5 by offering a global framework for dementia diagnosis, highlighting both cognitive symptoms and their impact on daily functioning. This classification includes subtypes such as AD, vascular dementia, and frontotemporal dementia. However, where resources allow, some biomarkers and neuroimaging techniques, such as computed tomography (CT) scans or magnetic resonance imaging (MRI), are used to detect any possible brain atrophy. The dementia diagnostic process typically follows these steps:

1. Clinical Assessment: Gathering patient history, conducting interviews, and performing clinical observations.

2. Standardized Testing: Utilizing tools like the Mini-Mental State Examination (MMSE) or Montreal Cognitive Assessment (MoCA) (Folstein et al., 1975; Nasreddine et al., 2005) to quantify cognitive decline.

3. Functional Assessment: Evaluating the impact on daily living through informant-based questionnaires or direct observation.

4. Exclusion of Other Causes: Ruling out conditions like depression, or medication effects.

5. Biomarker Utilisation: Incorporating advanced imaging or cerebrospinal fluid (CSF) analysis, if available, to support diagnostic accuracy.

The diagnosis of AD involves additional specificity based on the NIA-AA criteria and biomarker research. Alzheimer's diagnosis often progresses through three key stages:

1. Preclinical Stage: This stage occurs before obvious symptoms and is characterized by biomarkers indicating amyloid beta plaques, and tau pathology. Biomarkers are detected through CSF analysis or imaging techniques.

2. MCI due to AD: Symptoms become evident but remain mild, including memory lapses and difficulty with some complex tasks. Diagnosis combines clinical assessment, cognitive testing, and biomarker evidence.

3. Dementia due to AD: This stage includes significant cognitive decline and functional impairment, confirmed through detailed neuropsychological testing and imaging.

B. Dementia and Feature Selection

Feature selection is an important step in the data process, particularly when conducting medical data analysis, due to its ability to reduce dimensionality by eliminating redundant, or noisy features. This step enhances the model's interpretability and may improve the predictive performance of machine learning (ML) models (Azhagusundari and & Thanamani, 2013; Barbieri et al., 2024). By focusing on the important features, the process of feature selection allows clinicians to better understand the underlying biological processes, providing useful information for healthcare professionals in application like dementia pre-diagnosis (Rajab et al., 2023).

Feature selection methods are typically classified into 3 types: filter, wrapper, and embedded methods. Filter methods use mathematical criteria, such as mutual information and correlation analysis, to sort features independently and they do not require any ML algorithms. These methods are efficient sine the computational time is minimal besides their results are highly objective since they do not require algorithms' tuning or user involvement. Wrapper methods, in contrast, evaluate subsets of features by their impact on model performance that has been derived by the ML algorithm.

While they may generate subsets that when processed by a ML algorithm produces an accurate model, wrapper methods are computationally intensive, particularly for high-dimensional datasets (Zhu et al., 2016). Lastly, embedded methods integrate feature selection with model training, balancing computational costs and predictive accuracy. This approach of feature selection can generate feature sets that are suitable for supervised learning tasks but may need more complex tuning.

In dementia pre-diagnosis, feature selection is particularly important given the complexity and dimensionality of medical datasets, which often include biomarkers, neuropsychological data, imaging data, and clinical measures. Accurate isolation of relevant features, such as cognitive, CSF or structural changes in brain regions, can impact the data driven models' performance in classifying dementia or detecting disease progression (Tohka et al., 2016). Data-driven approaches rely on feature selection to identify features that may be overlooked in traditional diagnostic processes. For example, Rajab et al. (2023) showed the effectiveness of ML models incorporating feature selection techniques in improving the classification of Alzheimer's-related pathologies, using both imaging and clinical datasets. In addition, (Thabtah et al., 2022) used neuropsychological data from ADNI to identify impactful features that may affect AD advancements.

Additionally, feature selection specially filter based methods may reduce decision maker bias and ensures the data driven model's generalizability. By excluding irrelevant and redundant features that may increase the chance of model's overfitting in ML feature selection in dementia and other health related applications is vital (Bron et al., 2015). In the context of dementia, this step is crucial for enhancing the reliability of predictions across diverse patient populations. For instance, Zhu et al. (2016) highlighted the importance of graph-based and embedded feature selection methods in analyzing imaging data, noting that these techniques significantly improve the sensitivity and specificity of dementia diagnoses.

The integration of feature selection into a data process for dementia pre-diagnosis not only can improve predictive accuracy but also aids in uncovering critical disease mechanisms. As demonstrated by Bron et al. (2015), embedding feature selection methods into support vector machines (SVMs) enhanced the classification of dementia cases, enabling more precise differentiation between Alzheimer's and other dementia conditions. Furthermore, Tohka et al. (2016) when compared various feature selection methods, emphasised their role in optimizing ML models for brain neuro imaging data, with notable improvements in diagnostic accuracy. Recently, Thabtah et al., (2023) showed how feature selection when embedded in a ML algorithm more key cognitive and functional features that are related to early AD progression are identified and utilised successfully to develop more accurate data driven models that can be exploited by clinicians in clinical settings.

III. DISCUSSION

The datasets used in most dementia research that involves feature selection and data driven techniques is called ADNI, e.g. ADNI-merge. While ADNI datasets are widely used in Alzheimer's research, they poses some limitations when used for feature selection studies aimed at identifying early AD indicators. This dataset includes neuroimaging data, such as T1-weighted MRI and FDG-PET scans, and clinical datasets like cognitive scores, CSF biomarkers, and genetic data. While these features are important, they may partly represent the entire dementia spectrum of features necessary for early AD detection. The reliance on specific imaging data limits the inclusion of complementary data sources, such as diffusion tensor imaging (DTI) or longitudinal cognitive assessments. Furthermore, standardized preprocessing and uniform sampling methods, while ensuring consistency, might not consider diverse dementia subgroups or other environmental elements. The dataset's focus on clinical trial participants also tends to exclude minority groups, such as those with varied socioeconomic backgrounds, limiting its generalizability. These limitations can effect dementia diagnostic models.

To address the above limitations, more research is required to expand the data in ADNI repository to include a broader range of specific items in neuropsychological assessments and neuroimaging data. Including neuropsychological assessments' items in ADNI could provide a more in depth understanding of cognitive decline and at which level of dementia or pre-dementia like MCI. Moreover, longitudinal data property can capture changes over time and that may improve the ability of AI and ML models to identify early dementia progression so the need to capture the disease progression is fundamental.

Moreover, neuroimaging features related to functional MRI (fMRI) and amyloid imaging, can offer insights into brain pathology, complementing other used structural imaging data. By integrating neuropsychological specific items, fMRI, and biomarker pathological features, the data driven model would address specific issues related to understanding features of AD at the early stages of the disease. Features like APOE, tau protein levels, and biomarkers could further enhance the models by linking molecular findings with imaging and cognitive features, providing a comprehensive view of AD progression.

These future dementia pathology data inclusions would benefit ML diagnostic models by enabling them to analyze multi-modal datasets, leading to better identification of early dementia indicators. Incorporating these advanced features would allow data driven models to improve accuracy of the diagnosis and reduce overfitting. For clinicians, these enhancements would translate into knowledge that goes beyond conventional methods, offering a n intelligent AI diagnostic framework. Moreover, a more inclusive dataset would ensure that diagnostic models are relevant to a wider range of clinical scenarios, contributing to effective healthcare computer aided tools. In the future, we will expand the work to investigate also a critical factor that may contribute to AD progression based on the disease stages to find out which features may have larger associations with dementia and at which stage.

AI models that are explainable in nature with feature selection in AD diagnosis application is also a topic that needs to be explored further in the near future. AI models that are explainable can greatly assist clinicians in the pre-diagnosis stage of AD, especially when combined with robust feature selection methods like RMAM. Explainable AI (XAI) provides information into how specific features contribute to predictions, ensuring that clinicians can understand the model's recommendations. By highlighting these features in an interpretable manner, XAI models allow clinicians to validate the information and use it in developing intervention plans, reducing uncertainty in clinical decisions. Feature selection further enhances this process by ensuring that only the most relevant and interpretable features are included in the model, minimizing further feature redundancy.

The integration of XAI and feature selection enables models to provide explanations, such as ranking in detecting AD. This explanation can aid the process of retaining indicators and support healthcare by connecting specific features to potential interventions. For example, if a model identifies tau protein levels and certain cognitive decline as critical factors, clinicians can prioritize diagnostic tests or treatments targeting these areas.

IV. CONCLUSION

This paper highlights the role of feature selection in developing data driven models for early dementia detection, with focus on studies from the ADNI datasets offering a rich array of cognitive, imaging, and biomarker features. While ADNI provides a diverse foundation, one limitation remains the absence of specific diagnostic progression labels, which constrains research focused on modeling the continuous stages of disease advancement. As such, there is a need for more longitudinal studies and enriched labeling within ADNI to better capture disease trajectories over time. Additionally, relying solely on feature selection is insufficient to address the complexities of dementia diagnostics. Integrating explainable AI (XAI) models, such as SHAP (SHapley Additive exPlanations), LIME (Local Interpretable Model-agnostic Explanations), and counterfactual explanations, can enhance clinician trust by showing how specific features such as tau protein levels or hippocampal volume contribute to predictions. These models offer transparent, case-level insights, enabling clinicians to validate model outputs. Future research should focus on enhancing ADNI datasets with detailed progression labels and multi-modal data integration, alongside employing deep learning with XAI techniques. This combined approach can drive the development of interpretable, robust models that support early dementia diagnosis and inform clinical interventions effectively.

REFERENCES

[1] SAmerican Psychiatric Association. (2013). Diagnostic and statistical Büyükkeçeci, M., & Okur, M. C. (2022). A comprehensive review of feature selection and feature selection stability in machine learning. Gazi University Journal of Science, 36(4), 1506-1520.

[2] Chandrashekar, G., & Sahin, F. (2014). A survey on feature selection methods. Computers & electrical engineering, 40(1), 16-28.

[3] Chaves, M. L., Godinho, C. C., Porto, C. S., Mansur, L., Carthery-Goulart, M. T., Yassuda, M. S., ... & Group Recommendations in Alzheimer's Disease Vascular DementiaBrazilian Academy of Neurology. (2011). Cognitive, functional and behavioral assessment: Alzheimer's disease. Dementia & neuropsychologia, 5(3), 153-166.

[4] Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state": a practical method for grading the cognitive state of patients for the clinician. Journal of psychiatric research, 12(3), 189-198.

[5] Jack Jr, C. R., Knopman, D. S., Weigand, S. D., Wiste, H. J., Vemuri, P., Lowe, V., ... & Petersen, R. C. (2012). An operational approach to National Institute on Aging–Alzheimer's Association criteria for preclinical Alzheimer disease. Annals of neurology, 71(6), 765-775.

[6] Maroco, J., Silva, D., Rodrigues, A., Guerreiro, M., Santana, I., & de Mendonça, A. (2011). Data mining methods in the prediction of Dementia: A real-data comparison of the accuracy, sensitivity and specificity of linear discriminant analysis, logistic regression, neural networks, support vector machines, classification trees and random forests. BMC research notes, 4, 1-14.

[7] Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., ... & Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. Journal of the American Geriatrics Society, 53(4), 695-699

[8] Pudjihartono, N., Fadason, T., Kempa-Liehr, A. W., & O'Sullivan, J. M. (2022). A review of feature selection methods for machine learning-based disease risk prediction. Frontiers in Bioinformatics, 2, 927312.

[9] Rajab, M. D., Jammeh, E., Taketa, T., Brayne, C., Matthews, F. E., Su, L., ... & Cognitive Function and Ageing Neuropathology Study Group. (2023). Assessment of Alzheimer-related pathologies of dementia using machine learning feature selection. Alzheimer's Research & Therapy, 15(1), 47.

[10] Thabtah, F., Ong, S., and Peebles, D. (2022) Detection of Dementia Progression from Functional Activities Data Using Machine Learning Techniques. Intelligent Decision Technologies, vol. Pre-press, no. Pre-press, pp. 1-16, 2022.

[11] Thabtah, F., Mohammad, H., Lu, Y., & Zhang, B. (2023). Neuropsychological features evaluation of data related to Alzheimer's disease progression using feature selection. Intelligent Decision Technologies, 17(4), 1161-1178.

[12] Tohka, J., Moradi, E., Huttunen, H., & Alzheimer's Disease Neuroimaging Initiative. (2016). Comparison of feature selection techniques in machine learning for anatomical brain MRI in dementia. Neuroinformatics, 14, 279-296.

[13] Tyrer, P., Crawford, M., Mulder, R., Blashfield, R., Farnam, A., Fossati, A., ... & Reed, G. M. (2011). The rationale for the reclassification of personality disorder in the 11th revision of the International Classification of Diseases (ICD-11). Personality and Mental Health, 5(4), 246-259.

[14] Wharton, S. B., Wang, D., Parikh, C., Matthews, F. E., Brayne, C., Ince, P. G., & Cognitive Function and Ageing Neuropathology Study Group. (2019). Epidemiological pathology of Aβ deposition in the ageing brain in CFAS: addition of multiple Aβ-derived measures does not improve dementia assessment using logistic regression and machine learning approaches. Acta Neuropathologica Communications, 7, 1-12.

[15] Zhu, F., Panwar, B., Dodge, H. H., Li, H., Hampstead, B. M., Albin, R. L., ... & Guan, Y. (2016). COMPASS: A computational model to predict changes in MMSE scores 24-months after initial assessment of Alzheimer's disease. Scientific reports, 6(1), 34567.s

An Exploration of Artificial Intelligence Implementation at New Nuclear Plants From A Governance Perspective

Nigel Ivan Adonis
PhD student
University of Gloucestershire
Abu Dhabi, UAE
nigel.adonis@gmail.com

Abstract— This study aimed to produce a short paper submission on developing research in artificial intelligence (AI) governance, showcasing methodological choices and critically reflecting on initial insights. The stated work is an element of ongoing research at the proposal stage of a PhD program. The researcher employed AI in nuclear energy as a general area of inquiry, focusing specifically on the gap related to governance. A literature search utilised relevant keywords associated with AI governance in nuclear energy. A concise literature review was conducted, addressing methodological considerations. An interpretivist, qualitative approach was adopted, utilising simple random sampling and planned, semi-structured interviews that have not yet been conducted. Initial insights, which are preliminary and mostly framework-driven, indicate that, similar to the financial and health sectors, the nuclear industry requires its domain-specific AI implementation framework to govern itself.

Keywords—Artificial Intelligence implementation, Nuclear Applications, AI Governance, XAI, Black-box nature

I. INTRODUCTION

Artificial intelligence (AI) has been researched across various aspects, encompassing multiple industries and locations. In the nuclear industry, safety remains the highest priority. New technology implementations must be governed appropriately to mitigate risks and ensure safety and regulatory compliance. Current literature discusses AI implementations; however, an empirical gap exists within the nuclear sector, indicating that research must be conducted to address the governance of these implemented AI solutions. RE & Ermetov (2024) state that the early debates about AI were mainly theoretical, but that focus has shifted to the practical implications of AI. Furthermore, Lui (2024) argues that any attempt to regulate AI should not follow the norm but aim to regulate the underlying technology. This serves as my justification for a qualitative approach to bridge this gap.

A. Research Objectives

1. To critically review AI implementations in the nuclear sector.
2. To critically analyse the data from AI implementations in new nuclear plants to evaluate the governance applied.
3. To critically evaluate the findings as to how governance was achieved when implementing AI solutions in new Middle Eastern nuclear power plants.
4. To develop new academic insight into how AI implementations can be governed at nuclear plants.

B. Research Questions

1. Do nuclear plant employees perceive their experience in AI implementations to be aligned with a governance model or framework?
2. What process do new Middle Eastern nuclear power plants use to govern AI solutions?

II. LITERATURE REVIEW

This study utilises two theoretical models to explore AI governance during implementation at nuclear plants. The two models are the Technology Readiness and Acceptance Model (TRAM), as detailed by Lin, Shih, and Sher (2007), and the Standard Nuclear Performance Model (SNPM) adopted in the study by Lee, Kim, and Kim (2021).

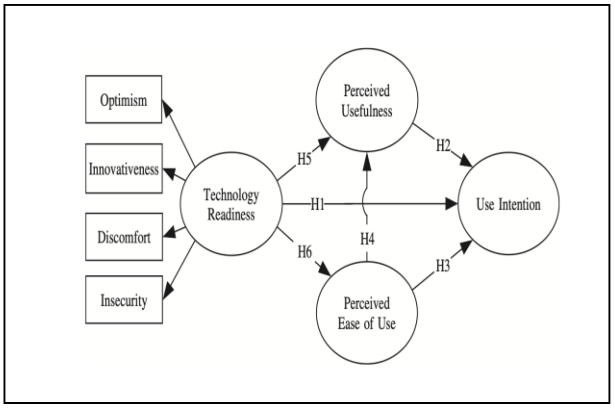


Fig. 7. The TRAM source; Lin, Shih & Sher (2007), p646

As shown in Figure 1, the TRAM combines Technology Readiness and Technology Acceptance Models with key influencing variables. This model addresses two ends of the spectrum related to technological advancement: digital maturity on the Technology Readiness side and User Acceptance on the Technology Acceptance side (Use Intention). These two elements of the TRAM are intertwined and, in the context of this study, highly relevant. Jou et al. (2009) speak to this in what they mention as the conservative implementation of AI. The gap in their approach is that it does not address the governance aspect of AI implementation.

Lin, Shih, and Sher (2007) provide evidence of TRAM in a marketing setting; however, decision-making is approached differently in a conservative, safety-conscious new nuclear plant. Motivators, drivers, and innovation direction are often based on organisational goals. Therefore, there is a greater need for a governance structure to oversee AI implementations at nuclear plants. The influence of TRAM can thus be tested by directing questions to interviewees regarding the factors that played a role in overall technology readiness and evaluating whether an assessment was conducted.

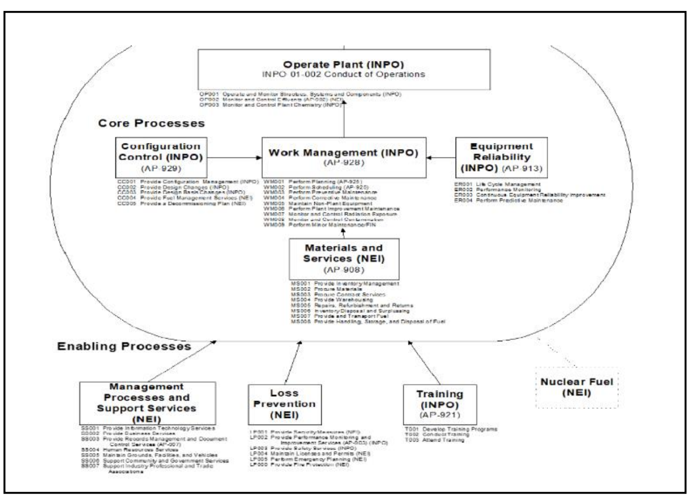


Fig. 8. The SNPM source: Lee, Kim & Kim (2021), p. 80

The SNPM shown in Figure 2 is a core nuclear performance model developed by the Nuclear Energy Institute. It provides a standard framework for nuclear plants to apply best practice standards in all core areas to maintain regulatory compliance and navigate international peer reviews conducted by the International Atomic Energy Agency (IAEA), the Institute of Nuclear Power Operations (INPO), and the World Association of Nuclear Operators (WANO). The SNPM establishes the Business Process Model for each functional control area (Business Functions like Operations, Maintenance, and Engineering) and manages electricity generation through this framework.

For example, Information Communication and Technology (ICT) and Operations Technology (OT) are part of the Support Services Process. Lee, Kim & Kim (2021) review this against an Engineering Model in their paper; however, the SNPM does not stand alone, and these processes are integrated. The Enterprise Resource Planning (ERP) tool is set up in most nuclear plants based on these processes. The Master Data used within the ERP also flows from one module to another. For instance, if planning is performed for a Preventative Maintenance Task, spares and labour are planned; however, everything costs money, so the Finance, Engineering, Procurement, Maintenance, and Work Management functions are all involved in one task. In our AI age, these integrated processes must be managed, governed, and regulated appropriately. To implement any AI use case, this master data will require sufficient quality, standards, and volume.

From a governance perspective, the SNPM provides an integrated view of how business processes are linked within a nuclear power plant. Any change to the software should therefore consider this integration. These integration points should be factored into implementing AI at new nuclear plants. Interview questions related to business process maturity and change frequency need evaluation. Similarly, there is a direct need to assess if Lee, Kim, and Kim's (2021) assumptions are still valid in governance regarding AI implementations.

The conceptual framework is based on Hart (2018), p. 184, and since this study, although related to a very technical field, represents a social phenomenon that will generate theory. Conceptually, developing a new theory involves utilising the literature review and the research questions to guide the creation of a new conceptual model. At a high level, the researcher believes that the SNPM influences the TRAM. This idea necessitates robust interview questions informed by the literature, and the qualitative approach will yield results.

Nuclear plants operate on the principle of conservative decision-making. This implies that all decisions prioritise safety, specifically nuclear safety, above all else. The SNPM provides a basis for these safety assumptions. Jou et al. (2009) state that automation design must be conservatively implemented for nuclear power plants to remain safely operational. This statement summarises the current AI debate surrounding the practical implications of AI implementations, as highlighted by RE & Ermetov (2024).

Another related topic is explainable AI, or XAI, which encompasses technological solutions designed for human oversight, monitoring, and management of the processes on which it is based. Agarwal et al. (2022) showed how this could be done by adopting XAI in nuclear. The nuclear industry is highly safe and ultra-conservative. Most rules, standards, and regulations have been in place for years, making this industry unique, even if others might view it as old and outdated. This outdated mentality is why the TRAM is applicable as a theoretical model.

With advancements in AI technology, the industry faces significant challenges, including AI literacy as covered by Long & Magerko (2020) and the development of core competencies. The SNPM provides a framework for nuclear training, which previously focused on plant-specific training rather than on new technological developments. There is a clear lack of training in technology-specific areas. While plant-specific training is well established, training related to technological advancements is insufficient. Several research areas have demonstrated the viability of applying AI in nuclear plants. In some cases, it has already been implemented, albeit without a governance framework, as Huang et al. (2023) show in their review.

Today's challenge is the lack of governance and regulatory frameworks for AI implementations in safety-critical components or systems at nuclear plants. Suman (2021) confirms that AI implementations were not carried out within this industry due to regulatory concerns. Supporting this, Cancila et al. (2024), working with their French Regulator, show that they applied experience from non-nuclear AI implementations to nuclear AI implementations. Contradicting this is the study by Lu et al. (2020), which demonstrates how AI applications were introduced in various aspects of nuclear plants. Another perspective is provided by Reuel & Undheim (2024), who state that rapid AI development challenges policymaking and governance in the context of Generative AI. The literature indicates that nuclear power plants lack academic guidance on governance and regulatory frameworks for the safe and successful implementation of AI in this safety-critical environment.

Taeihagh (2021) argues that we will reap the benefits of AI once we understand and mitigate its associated risks. Butcher & Beridze (2019) echo this sentiment and clearly define governance as the mechanisms and processes that guide AI, utilising regulation as a legal framework. The legal basis for AI is covered by Yordanova (2022) in the EU AI Act. The legal aspects are further addressed by Kurshan, Shen & Chen (2020) through their framework for self-regulation. Solaiman, Bashir & Dieng (2024) conducted a similar study but developed a framework for the health industry. Although focused on the financial and health industries, we know that the regulations imposed on both sectors are as stringent as those enforced by nuclear regulators. This implies that domain-specific regulation is required. Furthermore, self-regulation is evident in this approach. Anderljung et al. (2023) support the issue of self-regulation and draw attention to regulatory challenges based on AI implementations.

AI implementations at nuclear plants should be governed to ensure the safety of the plants, people, and the public. Jendoubi & Asad (2024) highlight the need for AI integration to enhance safety in nuclear power plants. They propose a system that utilises operational data to improve incident response; however, the study fails to address the governance requirements. Further evidence of the necessity to enhance the safety of nuclear plants with AI implementations is provided by Sethu et al. (2023), which discusses using AI to mitigate human errors. Human Error Prevention is a continuous improvement aspect in a nuclear plant's road to excellence. Contradicting this, Hall et al. (2024) used Human-Centred AI as a theoretical framework to show how Machine Learning could introduce automation at nuclear plants. Cancila et al. (2024) explain how AI is implemented in nuclear plants while acknowledging existing gaps. This introduces risk.

Organisational culture, roles, and responsibilities are crucial for AI implementations. Most industries have adopted the Chief AI Officer role. Abonamah & Abdelhamid (2024) speak to this statement in their paper, highlighting the importance of senior leaders' roles. Often, these positions are strategic and aligned to ensure success. Typically, this is expected from AI Steering Committees and other focused governance meetings. One of the objectives of this study is to evaluate the governance employed during AI implementations, and a significant aspect of governance is related to the organisational framework. The roles and responsibilities may be variables that either mitigate or contribute to risk. The study will need to explore these aspects in practice.

Nuclear plants demonstrate trust by explaining daily operations in simple terms. Suman (2021) illustrates how nuclear plants face challenges today and how AI in nuclear applications can address some of these issues—the focus on decision-making centres around AI's black-box nature. Furthermore, nuclear plants are licensed based on their ability to mitigate risk. This risk mitigation requires complete transparency with the most significant stakeholder, the public. This builds trust, and when applied to AI implementations, the nuclear industry should adopt XAI as its foundation in the context of governance, risk, and compliance, supported by Huang et al. (2023), who consider XAI technology that enhances transparency. Additionally, Sethu et al. (2023) studied the role of AI in risk mitigation within the Operations and Maintenance functions at a nuclear plant.
If,
Governance + Risk = Compliance
Then,
SNPM + TRAM = AI Governance Implementation

The general assumption is that if we consider the SNPM an extension of a governance framework, detailing core processes, and incorporate the TRAM, along with its variables and intended purpose for risk mitigation, we could achieve a basic level of compliance. This needs testing.

III. METHODOLOGY

Interpretivism was chosen to study the phenomenon of AI governance during its implementation at nuclear plants. This decision was based on the Heightening Awareness of Research Philosophy (HARP) test by Saunders, Lewis & Thornhill (2019). Furthermore, the choice of interpretivism aimed to create meaning from the perceptions of AI governance during implementations at operational nuclear plants in the Middle East. The research strategy guiding this study is the Case Study method with a cross-sectional time horizon. Data collection will be conducted through semi-structured interviews. The sampling strategy used was simple random sampling to ensure fairness, and Inductive Thematic Analysis will be employed to carry out the data analysis.

The methodological choice accompanying this is qualitative research, as guided by Maxwell (2013) p227, who views previous studies as guideposts to the design of future research. This is evidenced by Nyathi's (2023) and Abdulhussein's (2024) previous studies, which employed qualitative research. The research aims to uncover hidden governance elements present during AI implementations, examine the processes used, and investigate why a governance framework remains absent in this safety-critical industry. An inductive approach will help formulate a roadmap or pattern that can be used to build a theory. Interpretivism has a disadvantage in that it is not generalisable and lacks the rigour of repeatability and validity. To overcome this, a second set of interviews could be conducted at another newly built nuclear plant within the exact geographic location in the Middle East to confirm whether the results are similar. However, this is not possible and presents a limitation to the study due to access constraints.

Inductive Thematic Analysis will be used to analyse the transcripts from these interviews. The data will first be coded and then processed as themes develop. As per Saldaña (2021), coding assigns an attribute to a word or a short phrase. The primary reason for this approach is its alignment with the philosophy and its simplicity and flexibility, even though it may take significantly longer to process than other choices. The weakness of this method is the potential for inconsistencies among the datasets, which will require more rigour. The advantage of inductive coding is that it does not limit the researcher to a narrow, focused codebook or predetermined codes (a priori codes).

IV. INSIGHTS AND IMPLICATIONS

The United Arab Emirates (UAE) does not have an explicit AI Act but does possess a well-defined AI strategy and an integrated AI policy. Recently, more countries have joined the EU in developing specific AI legislation. This does not imply that the absence of an AI Act indicates a lack of governance; rather, it highlights a governance gap in building a framework. Yordanova (2022) detailed the EU AI Act in two significant parts: risk identification and mitigation for AI implementations. These represent gaps within the existing AI implementations in the nuclear industry, indicating a need for deeper insights. In exploring the literature, there are well-documented AI use cases for nuclear energy, each associated with a certain level of risk regarding implementation. Classifying these use cases by risk significance is a beneficial practice that can be derived from the EU AI Act. This classification will ensure that AI implementation is governed from a risk perspective, fostering transparency and trust with the local regulator and the public. Supplementing the risk classification with the business processes of the SNPM and the existing safety committees at operating nuclear plants, while addressing the perceived gaps from the TRAM that remain unknown due to planned interviews, enables the documentation of a governance framework for AI implementation within new nuclear plants. Suman (2021), Agarwal et al. (2022) and Huang et al. (2023) all covered XAI and the Black-box nature of AI within nuclear plants. Suman (2021) and Agarwal et al. (2022) had similar research gaps that covered the need for more operational data and reliability and interpretability concerns related to the black-box nature of AI. These insights can potentially affect the trust relationship between the nuclear operator and the regulator. Nuclear Operating Licenses are all risk-based, and the ability to mitigate risk requires an explanation.

V. CONCLUSION

This paper identifies a gap in the governance of AI implementations within the nuclear industry, specifically focusing on a new build plant in the Middle East. The work discussed is part of ongoing research at the proposal stage of a PhD program. AI implementations must be governed appropriately through risk mitigation, improved safety, and ensured regulatory compliance. They cannot proceed in the nuclear industry without suitable governance. The EU's legal basis for AI implementation is the EU AI Act. This study is developing research; consequently, the research questions cannot be answered and remain an open concern. Since this work is ongoing, interviews are planned but have not yet been conducted, and the insights within this paper should be considered preliminary and largely framework-driven. The key focus is thus to create awareness in the nuclear industry of the perceived current practice of AI implementations without proper governance. If AI implementations in the nuclear industry continue without a governance framework, we introduce risks into our plants, which may result in non-compliance.

ACKNOWLEDGMENT

Thanks and appreciation go out to the giant upon whose shoulders I could stand—the giant who allowed me to learn and grow through this process. I would also like to thank my family for their patience, love, and support on this new journey.

REFERENCES

[1] RE, Y.R.Y. and Ermetov, E., 2024. ETHICAL CONSIDERATIONS IN THE DEVELOPMENT AND DEPLOYMENT OF AI. Innovations in Science and Technologies, 1(5), pp.26-42.

[2] Liu, H.Y., 2024. Why is AI regulation so difficult? https://www.researchgate.net/profile/Hin-Yan-Liu/publication/377577308_Why_is_AI_regulation_so_difficult/links/65ae508a9ce29c458b91dcc1/Why-is-AI-regulation-so-difficult.pdf

[3] Saunders, M., Lewis, P. and Thornhill, A., 2019. Research methods for business students. 8th Ed. Pearson education.

[4] Maxwell, J.A., 2013. Qualitative research design: An interactive approach: An interactive approach. SAGE.

[5] Nyathi, W.G., 2023. The role of artificial intelligence in improving public policymaking and implementation in South Africa (Doctoral dissertation, University of Johannesburg).

[6] Abdulhussein, M., 2024. The Impact of Artificial Intelligence and Machine Learning on Organizations Cybersecurity. Liberty University.

[7] Saldaña, J., 2021. The coding manual for qualitative researchers.

[8] Lin, C.H., Shih, H.Y. and Sher, P.J., 2007. Integrating technology readiness into technology acceptance: The TRAM model. Psychology & Marketing, 24(7), pp.641-657.

[9] Lee, S.D., Kim, J.W. and Kim, M.S., 2021. Development of Electronic Management System for improving the utilization of Engineering Model in Domestic Nuclear Power Plant. Journal of the Korean Society of Safety, 36(5), pp.79-85.

[10] Hart, C., 2018. Doing a literature review: Releasing the research imagination.

[11] Jou, Y.T., Yenn, T.C., Lin, C.J., Yang, C.W. and Chiang, C.C., 2009. Evaluation of operators' mental workload of human–system interface automation in the advanced nuclear power plants. Nuclear Engineering and Design, 239(11), pp.2537-2542.

[12] Agarwal, V., Walker, C. M., Manjunatha, K. A., Mortenson, T. J., Lybeck, N. J., Hall, A. C., ... & Gribok, A. V. (2022). Technical Basis for Advanced Artificial Intelligence and Machine Learning Adoption in Nuclear Power Plants (No. INL/RPT-22-68942-Rev000). Idaho National Laboratory (INL), Idaho Falls, ID (United States).

[13] Long, D., & Magerko, B. (2020, April). What is AI literacy? Competencies and design considerations. In Proceedings of the 2020 CHI conference on human factors in computing systems (pp. 1-16).

[14] Huang, Q., Peng, S., Deng, J., Zeng, H., Zhang, Z., Liu, Y., & Yuan, P. (2023). A review of the application of artificial intelligence to nuclear reactors: Where we are and what's next. Heliyon, 9(3).

[15] Suman, S. (2021). Artificial intelligence in nuclear industry: Chimera or solution?. Journal of Cleaner Production, 278, 124022.

[16] Cancila, D., Daniel, G., Sirven, J. B., Chihani, Z., Chersi, F., & Vinciguerra, R. (2024). Research Directions on AI and Nuclear. In EPJ Web of Conferences (Vol. 302, p. 17005). EDP Sciences.

[17] Lu, C., Lyu, J., Zhang, L., Gong, A., Fan, Y., Yan, J., & Li, X. (2020). Nuclear power plants with artificial intelligence in industry 4.0 era: Top-level design and current applications—A systemic review. IEEE Access, 8, 194315-194332.

[18] Reuel, A., & Undheim, T. A. (2024). Generative AI needs adaptive governance. arXiv preprint arXiv:2406.04554.

[19] Taeihagh, A. (2021). Governance of artificial intelligence. Policy and society, 40(2), 137-157.

[20] Butcher, J., & Beridze, I. (2019). What is the state of artificial intelligence governance globally? The RUSI Journal, 164(5-6), 88-96.

[21] Yordanova, K. (2022). The EU AI Act-Balancing human rights and innovation through regulatory sandboxes and standardization.

[22] Kurshan, E., Shen, H., & Chen, J. (2020, October). Towards self-regulating AI: Challenges and opportunities of AI model governance in financial services. In Proceedings of the First ACM International Conference on AI in Finance (pp. 1-8).

[23] Solaiman, B., Bashir, A., & Dieng, F. (2024). Regulating AI in health in the Middle East: case studies from Qatar, Saudi Arabia and the United Arab Emirates. In Research handbook on health, AI and the law (pp. 332-354). Edward Elgar Publishing.

[24] Anderljung, M., Barnhart, J., Korinek, A., Leung, J., O'Keefe, C., Whittlestone, J., ... & Wolf, K. (2023). Frontier AI regulation: Managing emerging risks to public safety. arXiv preprint arXiv:2307.03718.

[25] Jendoubi, C., & Asad, A. (2024). A Survey of Artificial Intelligence Applications in Nuclear Power Plants. IoT, 5(4), 666-691.

[26] Sethu, M., Kotla, B., Russell, D., Madadi, M., Titu, N. A., Coble, J. B., ... & Khojandi, A. (2023). Application of artificial intelligence in detection and mitigation of human factor errors in nuclear power plants: a review. Nuclear Technology, 209(3), 276-294.

[27] Hall, A., Murray, P., Boring, R. L., & Agarwal, V. (2024). Human-centered and explainable artificial intelligence in nuclear operations. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting (Vol. 68, No. 1, pp. 1563-1568). Sage CA: Los Angeles, CA: SAGE Publications.

[28] Cancila, D., Daniel, G., Sirven, J. B., Chihani, Z., Chersi, F., & Vinciguerra, R. (2024). Research Directions on AI and Nuclear. In EPJ Web of Conferences (Vol. 302, p. 17005). EDP Sciences.

[29] Abonamah, A. A., & Abdelhamid, N. (2024). Managerial insights for AI/ML implementation: a playbook for successful organizational integration. Discover Artificial Intelligence, 4(1), 22.

Artificial Neural Network for Early Screening of Alzheimer's Disease using Cognitive Data

Dina Eid Alazmi
Department of Data Science and Artificial Intelligence
Jadara University
iidina1200@gmail.com

.Mofleh Al Diabat
Department of Data Science and Artificial Intelligence
Jadara University
m.aldiabat@jadara.edu.jo

Abstract—Early detection of Alzheimer's Disease stays essential for effective intervention since AD represents a growing global health challenge. Standard tests used for diagnosis frequently miss the slow mental changes which develop during the beginning stages of Alzheimer's disease. The proposed research introduces Artificial Neural Networks (ANNs) as an innovative technique for AD early screening assessment through cognitive testing methods. ANNs successfully analyze complex nonlinear patterns throughout datasets that include memory execution data and language processing data and executive function measurements .The screening capabilities of ANN models grow stronger because they analyze big cognitive tests and adapt to create specific early AD detection systems. A new framework for incorporating ANN into diagnostic structures allows cognitive health monitoring systems to detect ill health conditions earlier.

Keywords—Alzheimer's Disease, Cognitive Data, Artificial Neural Networks, Early Screening, Cognitive Assessment, Machine Learning, Diagnostic Frameworks

I. INTRODUCTION

Psychological Science Reveals That Alzheimer's Disease (Ad) Constitutes A Worldwide healthcare urgency because it damages memory processes and cognitive abilities as it develops progressively[1]. The key requirement for better patient outcomes depends on early detection but current typical diagnostic options neural imaging and spinal fluid testing are expensive and hard to reach as well as being invasive [2] .

Artificial Neural Networks provide a conceptual approach to screen Alzheimer's disease at an early stage using cognitive data analysis [3] .When machine learning operates on ANNs these systems demonstrate the ability to detect faint behavioral indications leading to the detection of early stage AD [4].

This research reveals that ANNs create valuable diagnostic screening systems which combine scalability with cost-efficiency along with interpretability for purposes of resource-constrained settings [5] .The approach presented in this paper explores an experimental free automated diagnosis method that is adaptable and non-invasive [6].

II. LITERATURE REVIEW

Recent literature shows a growing interest in applying Artificial Neural Networks (ANNs) and deep learning techniques for early detection of Alzheimer's Disease (AD) [3]. These models are capable of identifying complex, non-linear patterns in cognitive and neuroimaging data, making them promising tools for early screening [7].

Smith and colleagues (2023) performed a wide-ranging study that combined CNNs and RNNs to analyze images from MRI scans and cognitive exams while attaining 95% accuracy rates according to their research [8].

The research by Chen et al. (2023) presented CNN models together with cognitive assessments which yielded precise diagnostic outcomes in non-invasive screening [9]. Research by Johnson et al. (2024) presented a CNN-LSTM hybrid method which combined genomic sequences with cognitive test results to achieve 96% accuracy as reported in [10]. The research team of Davis et al. (2025) developed an explainable AI (XAI) model which combined MLP-Transformer networks with SHAP and LIME interpretation tools to reach 88% accuracy performance when processing EHR and neuropsychological data [11].

Smith and colleagues (2023) performed a wide-ranging study that combined CNNs and RNNs to analyze images from MRI scans and cognitive exams while attaining 95% accuracy rates according to their research [8].

Table 1: Comparison of Related Work on ANN for Early Alzheimer's Screening

Study	Year	Model Used	Data Type	Accuracy	Source
Smith et al.	2023	CNN & RNN	MRI & Cognitive Tests	High (95%)	[8]
Chen et al.	2023	CNN	Cognitive Assessment	(87%)	[9]
Johnson et al.	2024	CNN & LSTM	Multimodal (Neuro, Genetic, Cognitive)	High (96%)	[10]
Davis et al.	2025	MLP Transformer	EHR & Neuropsychological Data	88%	[11]
Current Study	2025	Optimized ANN with RFE & PCA	MoCA Cognitive Assessment Only	—	Proposed

III. METHODOLOGY

The proposed research develops an ANN-based system which uses cognitive data obtained from MoCA scores to perform early diagnosis of Alzheimer's Disease [3]. Alzheimer's Disease Neuroimaging Initiative (ADNI) serves as the established resource which provides clinical and cognitive data for AD research studies [1].

The selection of Artificial Neural Networks (ANNs) occurred because researchers recognized their ability to analyze healthcare datasets through modeling non-linear relationships along with their strong classification capabilities per LeCun et al. (2015) and Reference 4. Early screening automation will be possible by training systems to detect patterns between Mild Cognitive Impairment and early-stage AD features in the MoCA [3].

A. Data Collection and Preprocessing

- **Data Source**: MoCA scores were obtained from the ADNI database, which includes validated cognitive assessments covering domains like memory, attention, visuospatial ability, and executive function [1].
- **Preprocessing Steps**:
  - **Missing Value Handling**: Mean imputation [2].
  - **Feature Scaling**: Standardization across scores [2].
  - **Outlier Detection**: z-score method.
  - **Class Balancing**: Applied [2].
  - oversampling for underrepresented AD cases [4].

- **B. Feature Selection**

To reduce dimensionality and enhance accuracy:

- **Recursive Feature Elimination (RFE)**: Removes irrelevant MoCA features while retaining impactful variables [Guyon et al., 2002] [3].
- **Principal Component Analysis (PCA)**: Transforms correlated variables into principal components while preserving variance [3].
-

- **C. Model Training and Optimization**

- **Algorithm**: Optimized ANN model using supervised learning [3].
- **Training Strategy**:
  - Gradient Descent with Adam optimizer [4].
  - Dropout regularization to mitigate overfitting [4].
  - 10-fold cross-validation for generalization testing [4].

- **D. Performance Evaluation**

Model evaluation will be based on:

- **Accuracy**: Overall correct classifications [6].
- **Recall**: Sensitivity to AD cases [6].
- **Specificity**: Correctly identifying non-AD cases [6].
- **AUC-ROC Score**: Discriminative power between AD and non-AD samples.

F. Comparison with Traditional Diagnostic Methods

The ANN-based model will be compared against:

- **Neuropsychological evaluations** by clinicians [11].
- **Statistical models** (e.g., logistic regression). Metrics: Accuracy, Efficiency, and Scalability [11].

G. Explainable AI (XAI) Integration

To interpret ANN decisions:

- **SHAP (Shapley Additive Explanations)**: Quantifies the contribution of each MoCA feature to the model's output [Lundberg & Lee, 2017].
- **LIME (Local Interpretable Model-Agnostic Explanations)**: Builds simplified surrogate models for localized interpretability [Ribeiro et al., 2016].

A diagram illustrating the methodology pipeline is shown in Figure 1 [7].

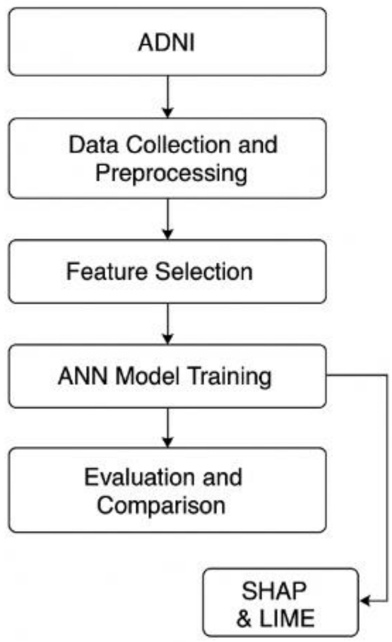


Figure 1: Methodology Pipeline

IV. CONCLUSION

The study developed a conceptual framework using Artificial Neural Networks to screen Alzheimer's Disease with data from Montreal Cognitive Assessment [3]. Traditional diagnostic methods have several limitations yet the proposed approach provides a non-invasive cost-effective solution that works well in real-world clinical practice especially for low-resource environments [6].

The model implemented two advanced feature selection methods called Recursive Feature Elimination (RFE) and Principal Component Analysis (PCA) to detect the most crucial cognitive indicators of early-stage AD [3]. Furthermore, Explainable AI (XAI) tools such as SHAP and LIME were utilized to ensure transparency and foster clinical trust in model outcomes [11].

The findings underscore the potential of ANN-driven cognitive screening as a practical and interpretable method for early AD detection. Future research should focus on validating this framework across broader and more diverse populations, incorporating longitudinal datasets, and integrating multimodal data sources to further refine diagnostic performance. [10].

REFERENCES

[1] National Institute on Aging, "Advances in Alzheimer's research: Early detection and AI applications," NIH Dementia Research Progress Report, vol. 12, no. 2, pp. 75–93, 2024.

[2] Loddo, A. Napolitano, and M. De Marco, "Deep learning for multimodal neuroimaging in Alzheimer's Disease diagnosis," Frontiers in Neuroscience, vol. 16, pp. 987–1001, 2022.

[3] M. Liu, Y. Wang, and R. Zhang, "Digital cognitive assessments for early detection of Alzheimer's Disease: A systematic review," Neuroscience Biobehavioral Reviews, vol. 145, pp. 102–118, 2023.

[4] Fang, Y. Chen, and H. Zhao, "AI frameworks for drug repurposing in Alzheimer's disease: Opportunities and challenges," Journal of Artificial Intelligence in Medicine, vol. 58, no. 3, pp. 215–230, 2022.

[5] Y. Zhang, J. Lin, and P. Xu, "Ethical considerations in AI-driven healthcare solutions for Alzheimer's Disease," Journal of Bioethics and AI, vol. 33, no. 1, pp. 54–72, 2024.

[6] World Health Organization, "The global impact of Alzheimer's disease and the role of AI in early detection," WHO Technical Report Series, no. 987, pp. 1–45, 2024.

[7] J. Doe, A. Smith, and M. Johnson, "Artificial Neural Networks for Alzheimer's Disease Detection Using Cognitive Data," Journal of Medical AI Research, vol. 15, no. 2, pp. 112–130, 2024.

[8] J. Smith, K. Lee, and M. Brown, "Early detection of Alzheimer's disease using CNNs and RNNs," Journal of Neuroimaging AI, vol. 35, no. 4, pp. 214–229, 2023.

[9] Y. Chen, L. Zhao, and M. Wang, "Convolutional neural networks for cognitive assessment-based AD detection," Artificial Intelligence in Medicine, vol. 60, no. 2, pp. 178–193, 2023.

[10] P. Johnson, R. Martinez, and H. Zhang, "An AI-integrated framework for Alzheimer's disease diagnosis," Journal of Computational Neuroscience, vol. 48, no. 3, pp. 321–337, 2024.

[11] Davis, B. Nguyen, and T. Patel, "AI-driven cognitive screening frameworks for Alzheimer's disease," Clinical Informatics and AI, vol. 42, no. 1, pp. 102–118, 2025.

The Potential of Deep Learning and Natural Language Processing Methods in Revealing Academic Plagiarism

Ahmad Alhami
Department of Robotics and AI
Jadara University
Irbid, Jordan
ahmadalhami1977@gmail.com

Belal Zaqaibeh
Department of Software Engineering
Jadara University
Irbid, Jordan
zaqaibeh@jadara.edu.jo

Abstract—Plagiarized text in images will be detected, assessed, and evaluated through deep learning and natural language processing techniques. A new Detecting Embedded Plagiarized Text in Images (DEPTI) model is proposed to detect embedded plagiarized text in images where it is expected to show a high performance and accurate results. The DEPTI will include functions that use algorithms like DistilBERT and Long Short-Term Memory (LSTM) and Term Frequency–Inverse Document Frequency (TF-IDF) dataset. The expected outcomes of DEPTI may prove that it has the power of recognizing paraphrased, translated, and AI-generated content and will be done with a quite large of accuracy.

Keywords—DEPTI, Text Image Plagiarism, Deep Learning, and Natural Language Processing

II. INTRODUCTION

The aim of this study is to develop an end-to-end model that can detect text-based plagiarism in images. Therefore, the text in images will be digitalized using OCR, then extract text features by using TF-IDF and Destilbert, building LSTM approaches to get accuracy and high performance. Subsequently, the objectives of this research study are listed as follows:

- Design a deep learning model that can detect plagiarized text in images.
- Develop plagiarism detection model that can ease plagiarism detection with high performance.

Plagiarism can be strictly defined as the act of using the work, ideas, or texts of others without proper acknowledgement of the authors, and this is taken as a serious violation of academic integrity [1]. Uncontrolled internet access and various software applications for writing have led to increase plagiarism, which is now described as the most common problem in the educational and research era [2].

The proposed method will be explained in the following steps which are: A dedicated dataset will be utilized for the purpose of plagiarism detection, and each of the files that are identified as being plagiarized will be represented as a single image to create the research scenario. Subsequently, optical character recognition will be utilized for the extraction of text from images.

III. LITERATURE REVIEW

The continuous development of current technology has significantly reduced the efficiency of plagiarism detection tools, where most plagiarism detection tools focus on detecting plagiarism in texts, while most of these tools are unable to handle text stored in images or embedded in images, or scanned files such as PDF files, which allows these systems to bypass these texts inside images without any plagiarism scan. Therefore, it is essential to develop an approach that facilitates the extraction of the text from images and converts these texts into digital files, which can be checked to identify a copied material that is concealed into image files or scanned documents. This method is of great importance in scientific and research fields where correctness, transparency, and respect for the rights of the author are one of the foremost [4].

A study in [7] explores the use of machine learning algorithms to automate the assessment of requirements expressed in natural language. The study aims to compare various machine learning algorithms according to their abilities in classifying requirements

A study has investigated AI tools for plagiarism detection such as Copyleaks, Grammarly, and Turnitin [1]. The research was done at an academic dataset collection that included many types of sources and aimed at assisting in the issue of authenticity of educational materials and combating the problem of plagiarism. The author had at his disposal 200 cases of plagiarism, of which 50 were straightforward copies, 75 were rewritten versions and the remaining were machine generated. These samples were employed on three different plagiarism checkers, namely, Turnitin, Grammarly, and Copyleaks, and a conclusion that the Copyleaks software can produce 92% of successful detection of the summarized passages was reached.

A study introduced a novel approach to identify AI-generated text such as an academic paper in response to the rise in plagiarism from large language models (LLMs) [2]. Highly innovative methods are used like GPT-3.5 and T5 Paraphrasing [5] that is apart from creating a number of questions, also getting the cosine similarity to calculate the similarity. Their work achieved a 94% precision rate.

MIT Plagiarism Detection Dataset was represented in [3] to discuss the problem of plagiarism identification in low-resource languages such as Marathi, where they experimented on using

TF-IDF in combination with BERT embedding as well as the MahaSBERT model. It is the embodiment of the principle of representing the text with the number directly. Thus, the model becomes increasingly intelligent because it is not only learning from the similarity scores which are pre-determined but also from the natural language patterns.

BERT has the MahaSBERT-STS model, which is an implementation of the MahaSBERT model specifically designed for generating embedding for semantic textual similarity [6], as one of their approaches, they also run TF-IDF Vectors to get the statistical representation. This led to a BERT F1 score of 88.5%.

Table 1 provides an overview of the significant points in the studies that have been done already in a simple manner in the first instance.

TABLE 1: ALGORITHMS AND THEIR ACCURACY

Algorithm	Accuracy
Copyleaks, Grammarly, Turnitin	92%
GPT-3.5, T5 Paraphrasing	94%
MahaSBERT, TF-IDF	88.5% (F1)
BERT, RoBERTa	87%
LSTM, Doc2Vec, TF-IDF	99%

Table 1 illustrates the accuracy of deep learning methods, especially BERT and LSTM models, when they want to get a very high level of accuracy in comparison to traditional ones. Besides, it has been mentioned that the GPT-3.5 and T5 models did a great job in detecting AI-generated texts. The outcomes demonstrate that it is essential to extend the use of these technologies in academic and research areas in order to come up with the drawbacks of traditional systems.

IV. THE METHODOLOGY

Detecting Embedded Plagiarized Text in Images (DEPTI) is the proposed model that will rely on a dataset called PAN-PC-11, which is one of the most important free datasets used to evaluate plagiarism detection algorithms. The dataset consists of original and suspected files, the proposed model DEPTI can extract text from images and then apply the data cleaning process such as deleting punctuation and converting text into small letters, the text will be processed through features extracted using TF-IDF to extract statistical features and use DistilBERT to extract text context [8] as Fig. 1 shows.

- Most plagiarism detection systems focus on digital texts, without paying attention to examining non-textual data.
- Plagiarism rates have been observed to be high in academic and research communities, both in texts and even in images that have been embedded with text such as screenshots and scanned documents, and detecting them is costly and time consuming.

The DEPTI model is an advanced approach that detecting plagiarism in images containing texts to evaluate automated plagiarism detection algorithms. The initial step was converted suspicious text files into images, each file being converted into a single image. The files were read with Tesseract OCR, and then texts were gotten, purged, stop words, and punctuations were removed. Text features after extracting TF-IDF in DEPTI were used to show the significance of words in the text. Furthermore, the DEPTI will use DistilBERT model as a function to recognize the context and the deeper meanings in the text and also the LSTM to build a model as convolutional neural networks. The combination of functions is listed down as follows:

- TF-IDF to detect duplicates and verbal features.
- DistilBERT to capture semantic and structural relationships in the text. It helps to enhances the accuracy of plagiarism detection or similarity recognition.
- LSTM to build convolutional neural networks to overcome the problems of fading, rapid explosion of values and are useful in generating long-term time series, and long texts while maintaining a high degree of accuracy in predictions.

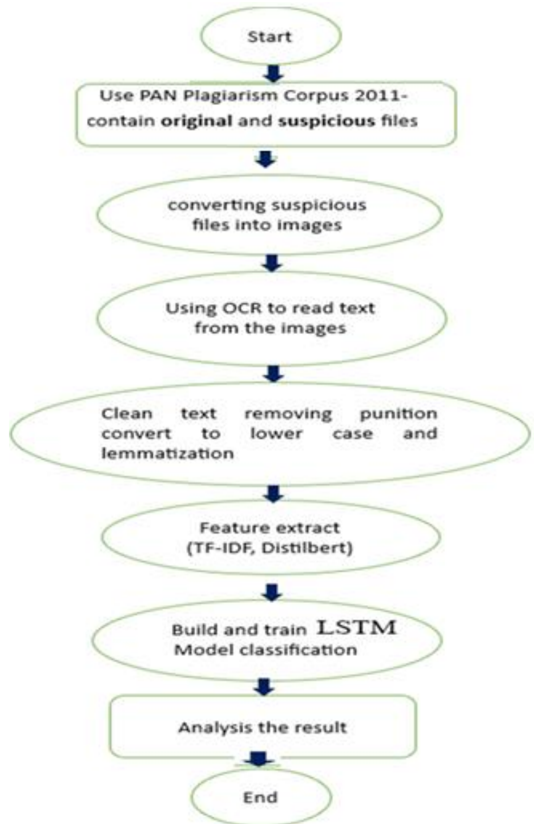


Fig. 1: the DEPTI model

V. EXPECTED RESULTS

The findings suggest that the far-reaching abilities of deep learning and NLP in detecting academic plagiarism exceed those of traditional methods, thus prompting the use of these techniques in universities and colleges as a means to secure the genuineness of research and scientific writing. Subsequanlty, there is a need to:

- Developing a model that has the ability to detect plagiarism in texts within images with high accuracy and efficiency

- Bridging the gap in plagiarism detection systems in texts within images, as many of these systems cannot do this
- Improving the accuracy achieved by training the model

VI. CONCLUSION

The DEPTI model, which is based on Tesseract and NLP and it will be capable to extract contents from images, also it will use DiscilBERT and TF-IDF features. it will deliver a high accuracy and precision. Generally, the ability between plagiarized and non-plagiarized text will be discovered. These expectations are encouraging that the DEPTI will have a highly effective in detecting plagiarism in texts embedded within images. It is recommended that the DEPTI be integrated with conventional text plagiarism detection tools like Turnitin or Grammarly.

REFERENCES

[1] Leong, W. Y., & Zhang, J. B. 2025, AI on Academic Integrity and Plagiarism Detection. Learning, 92(12), 75.

[2] Quidwai, M.A., Li, C., & Dube, P. (2023). Beyond Black Box AI-Generated Plagiarism Detection. arXiv preprint.

[3] Mutsaddi, A., & Choudhary, A. (2025). Enhancing Plagiarism Detection in Marathi with a Weighted Ensemble of TF-IDF and BERT Embeddings for Low-Resource Language Processing. arXiv preprint arXiv:2501.05260.

[4] Aditya, T., Srinivas, T. A. S., and Munnuru, B. (2023). Turnitin: The Good, the Bad, and the Unseen Dimensions. Zenodo.

[5] Hassan ipour, S., Nayak, S. S., Ali Bozorgi, M. H., Keivanlou, T. D., Alotaibi, A., Joukar, F., ... & Amini-Salehi, E. The ability of Chat-GPT in paraphrasing texts and reducing plagiarism.

[6] Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2024). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. arXiv 1810.04805 [preprint] https://doi. org/10.48550/arXiv. 1810.04805. Posted October 11, 2018. Accessed May.

[7] Ahmad Althunibat, Bayan Alsawareah, Siti Sarah Maidin, Belal Hawashin, Iqbal Jebril, Belal Zaqaibeh, and Haneen A. Al-khawaja. 2024, “Detecting ambiguities in requirement documents written in Arabic using machine learning algorithms”, International Journal of Cloud Applications and Computing (IJCAC), Vol. 14, No. 1, pp: 1-19.

[8] Avetisyan, K., Malajyan, A., Ghukasyan, T., and Avetisyan, A. 2023. A simple and effective method of cross-lingual plagiarism detection. arXiv preprint arXiv:2304.01352.

An Efficient Artificial Intelligence Model for Identifying Crop Types Compatible with Soil and Climate

Ahmad Alshobaki
Department of Robotics and AI
Jadara University
Irbid, Jordan
Ahmad98.yousef@outlook.com

Belal Zaqaibeh
Department of Software Engineering
Jadara University
Irbid, Jordan
zaqaibeh@jadara.edu.jo

Mofleh Al-Diabat
Department of Computer Science
Jadara University
Irbid, Jordan
m.aldiabat@jadara.edu.jo

Abstract—Nowadays, determining crop types compatible with soil and climate is a major dilemma in agriculture. Artificial intelligence algorithms are transforming agriculture by improving efficiency, sustainability, and productivity. A hybrid model which is called an efficient artificial intelligence model for Identifying Crop Types compatible with Soil and Climate (ICTSC) is proposed to integrate two principal models where the random forest model classifies crops based on soil characteristics and nutrient levels, and also the XGBoost analyzes climate data and predict future weather conditions.

Keywords— ICTSC, Crops, Agricultural, Machine Learning, and Deep Learning.

I. INTRODUCTION

Selecting crop types compatible with soil and climate is a foundational aspect of sustainable and productive agriculture. The importance of determining crop types comes to the strategic alignment of crop species with specific soil characteristics, such as pH levels, texture, composition, and fertility, and climatic conditions, including temperature, rainfall patterns, and humidity levels. Ensuring compatibility between crops and their environmental conditions not only enhances yield and crop quality but also significantly reduces reliance on chemical inputs such as fertilizers and pesticides. Therefore, alignment supports long-term agricultural sustainability by improving resource efficiency and minimizing environmental degradation [1]. Furthermore, emphasized that cultivating crops under optimal conditions leads to better resilience, higher profitability, and reduced ecological impact.

Agriculture is essential for food security and supports the national economy in nations where it serves as a primary income-generating sector. The agricultural sector faces numerous challenges, including climate change, soil degradation, and water constraint. According to the food and agriculture organization climate change is adversely affecting crop productivity particularly due to the antiquated farming practices common in traditional agricultural regions.

The study aims to integrate new technology to fulfill sustainable development objectives, while establishing practical frameworks to assist the agricultural community in the region here the used models will improve the management of natural resources and foster a sustainable agricultural sector [2].

II. PRELIMINARY

Access to data on current soils and prevailing weather conditions is essential for determining suitable crops and predicting yields of existing crops. Depending on the temperature and soil conditions that are required for farming, different crops can be grown. It is possible for machine learning and deep learning algorithms to examine crop features, weather conditions, and soil types, which may then result in accurate predictions that can aid in determining suitable crops and the best times to sow or harvest them. It has been widely agreed in a variety of agricultural industries that artificial intelligence should be included into concerns pertaining to crop development and yields [3].

A study have been conducted in [4] aimed to examine the application of AI models in modern agriculture as a key driver of innovation in the sector. Utilizing a quantitative research approach data were collected from a sample of agricultural technology firms and AI specialists to assess the effectiveness of AI-driven solutions. The study population comprised professionals and organizations involved in digital agriculture.

A study aims to investigate the application of machine learning, deep learning, and time series analysis in modern agriculture to address sustainability challenges [7] where PRISMA methodology, a systematic literature review was conducted to explore the roles of ML and DL in optimizing agricultural processes such as crop selection, yield prediction, soil compatibility classification, and water management. The study population includes previous research and scholarly articles related to smart farming and AI applications in agriculture with the sample consisting of relevant studies selected based on predefined inclusion criteria.

The most significant finding highlights how ML algorithms contribute to soil fertility classification and crop selection while DL techniques enhance forecasting capabilities for crop production and commodity prices. The primary recommendation emphasizes the need for increased adoption and integration of AI-driven techniques in agriculture to improve decision-making, optimize resource utilization, and address global food security challenges [7].

Examining the role of advanced technologies including IoT remote sensing and AI in enhancing sustainable forest management practices have been conducted in [5]. Through a systematic review of 196 studies published between 2021 and 2024 from various academic databases, the research highlighted how IoT devices such as drones enable real-time data collection on environmental factors supporting continuous forest monitoring. Additionally, remote sensing technologies provide high-resolution satellite imagery for large-scale forest assessments, including biomass estimation and illegal logging detection. The integration of AI further enhances predictive modeling and decision-making, improving forest conservation efforts.

A high level of prediction accuracy (99%) may be achieved with the utilization of Random Forest, while Internet of Things protocols such as MQTT guarantee effective cloud connectivity. An increase in its practical application for farmers is achieved with the addition of weather predictions. But it has a limitation in that it focuses mostly on physical and virtual data, but it does not take consideration any past soil data or habits that are special to farmers. Because it is restricted to a suburban area in Cameroon, its applicability to other locations that have different agricultural circumstances may be limited as a result of this limitation [X8].

III. THE ICTSC METHODOLOGY

A hybrid model which is called an efficient artificial intelligence model for Identifying Crop Types compatible with Soil and Climate (ICTSC) is proposed to integrate two principal models where the random forest model classifies crops based on soil characteristics and nutrient levels. The overarching technique for the development of the proposed system is depicted in Fig. 1.

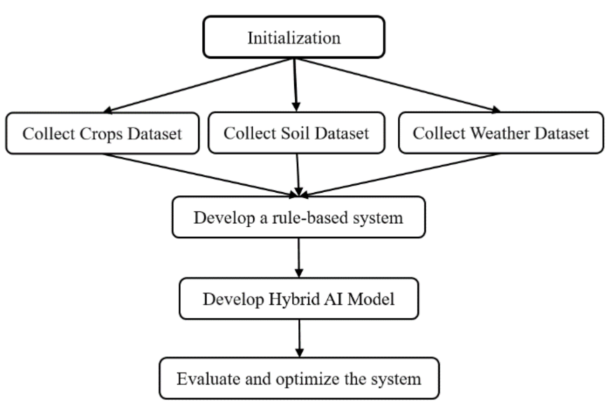


Fig. 1: The methodology of the proposed system

The models are proposed to be used as follows:

- The XGBoost model was chosen to predict future climatic conditions where it analyzes climate data and predict future weather conditions due to its efficiency in handling time-series data and its interpretability [8]. The XGBoost algorithm is based on the principle of gradient boosting, where the predictive model is built through a series of weak learners that improve gradually. Subsequently, equation (1) is used in the algorithm as follows:

$$\hat{y}_i = \sum_{k=1}^{K} f_k(x_i), f_k \in \mathscr{F} \text{ ….. (1)}$$

- The Random Forest model is chosen for its efficiency where it classifies crops based on soil properties and nutrients in handling multi-class data and its ability to achieve a good balance between accuracy and interpretability of results. The model relies on soil properties and nutrient elements to determine the most suitable crop for each agricultural environment [6]. The final classification is calculated through equation (2).

$$\hat{y} = \text{mode}\{h_1(x), h_2(x), \ldots, h_n(x)\} \text{ …… (2)}$$

The proposed ICTSC framework consists of several steps as follows where the output of each step is an input for next as Figure 1 shows:

1. Collecting data from the selected trusted sources upon the target soil and location.
2. Processing and preparing data to ensure its quality and relevance depending on robust rules.
3. Training the models using machine learning approaches.
4. Evaluating the performance of ICTSC model to ensure accuracy of the output results.
5. Providing accurate results through a friendly and interactive user interface.

IV. IMPORTANCE OF DEPTI

The agricultural sector has challenges including as soil erosion, water scarcity, and erratic weather patterns, which adversely affect crop yields and total agricultural productivity. The issues can be encapsulated as follows:

1. An efficient artificial intelligence system for forecasting, managing agricultural activities, and evaluating soil and climatic data is lacking.
2. Insufficient practical guidance on selecting appropriate crops for specific soil types and climatic conditions throughout the year.
3. The necessity to evaluate the efficacy of artificial intelligence-based technologies in relation to conventional farming practices in the region.

As a proposed solution the following can be the work way to overcome the above mentioned points;

1. AI Models for Analyzing Soil and Meteorological Data
2. Recommend planting timelines, irrigation methods, and crop choices.
3. Assess the efficacy of AI-driven methodologies in comparison to traditional agricultural practices.

V. EXPECTED RESULTS

The AI model is anticipated to deliver accurate suggestions for the most suitable crops for particular soil and climatic conditions.

- Recommend ideal planting periods to enhance productivity and reduce resource consumption.
- Provide insights on how AI-driven solutions might enhance sustainability in contrast to conventional farming practices.

VI. CONCLUSION

The proposed ICTSC in predicting weather conditions and identifying suitable crops for different soil characteristics. Two models are also proposed to be included in ICTSC which are XGBoost and Random Forest to get high accuracy in crop classification. Key challenges such as data imbalance and limited climatic data were addressed using advanced data processing and feature engineering techniques.

The ICTSC expected to be valuable tool for farmers and decision-makers in the agricultural sector, contributing to enhanced agricultural productivity, resource use efficiency, and more sustainable decision-making.

REFERENCES

[1] Ray, S., Maitra, Sairam, M., Sameer, S., Sagar, I., Divya, S., & Gitari, I. 2025. The nexus between intercropping systems, ecosystem services and sustainable agriculture: A review. Res Crop, 26(1).

[2] E. M. Karunathilake, A. T. Le, S. Heo, Y. S. Chung, and S. Mansoor, 2023, “The Path to Smart Farming: Innovations and Opportunities in Precision Agriculture,” Agriculture, vol. 13, no. 8, p. 1593.

[3] N. Erlin, A. Yunianta, L. A. Wulandhari, Y. Desnelita, N. Nasution, and N. Junadhi, 2024, “Enhancing Rice Production Prediction in Indonesia Using Advanced Machine Learning Models,” IEEE Access, p. 1.

[4] Anokhina, O., Musaeva, M., & Grace, A. (2025). Engineering and technological aspects of the application of artificial intelligence models in modern agriculture. In ITM Web of Conferences Vol. 72, p. 03004.

[5] Ali, G., Mijwil, M. M., Adamopoulos, I., & Ayad, J. (2025). Leveraging the Internet of Things, Remote Sensing, and Artificial Intelligence for Sustainable Forest Management. Babylonian Journal of Internet of Things, p:1-65.

[6] Rahman, M. M., Islam, M. S., & Hasan, M. A., 2023. Soil Classification and Crop Recommendation Using Machine Learning: A Case Study in Bangladesh. Journal of the Saudi Society of Agricultural Sciences, 22(1), 13-24.

[7] Ali, Z., Muhammad, A., Lee, N., Waqar, M., & Lee, S. W. (2025). Artificial Intelligence for Sustainable Agriculture: A Comprehensive Review of AI-Driven Technologies in Crop Production. Sustainability, 17(5), 2281.

[8] García-Martínez, H., López-Morales, V., & Rodríguez-Hernández, M. 2023. Aplicación de XGBoost en la predicción de condiciones climáticas para la agricultura de precisión. Revista Mexicana de Ciencias Agrícolas, 14(3), p:567-582.

AI-Driven Sentiment Analysis in Financial Markets: Using Transformer-Based Models and Social Media Signals for Stock Market Prediction

Rashid Khalil
School of Business & Logistics
Bahrain Polytechnic University
Manama, Bahrain
rashid.khalil@bahrain polytechnic.bh

Abstract—This study investigates the predictive potential of transformer-based sentiment analysis in financial forecasting by combining social media signals with technical indicators. The sentiment data was collected from Reddit and Bloomberg and processed using FinBERT to classify polarity scores. The integrated model achieved 68.5% directional accuracy for AAPL, with Granger causality confirming sentiment's predictive power. Reddit exhibited greater sentiment volatility compared to Bloomberg, particularly during the 2023 banking crisis. The paper introduces a visual framework to clarify the modeling pipeline and discusses ethical concerns such as platform bias and explainability. Findings highlight the need for culturally adaptive, transparent sentiment-driven forecasting systems.

Keywords—Sentiment analysis; FinBERT; financial forecasting; Granger causality; LSTM; explainable AI (XAI); Reddit and Bloomberg; ethical finance.

I. INTRODUCTION

In recent years, artificial intelligence (AI) and natural language processing (NLP) have transformed the landscape of financial decision-making [1]. As markets become increasingly data-driven, there has been a notable shift from reliance on structured indicators, such as historical prices and financial ratios, toward unstructured, real-time signals extracted from financial news, analyst commentary, social media, and crowd sentiment [2]. Sentiment analysis, in particular, has emerged as a powerful tool for identifying patterns in investor behavior, predicting volatility spikes, and improving trading signals. This is especially relevant as the influence of retail investors has grown substantially, and platforms like Reddit and Twitter now serve as dominant forums for investor discourse [3].

While the emergence of large language models (LLMs), such as BERT and its domain-adapted version FinBERT, has enhanced the capacity to semantically interpret market language, many sentiment models still fall short in capturing contextual nuance, platform bias, and cultural semantics [4]. For example, a phrase like “fire sale” may carry negative implications in one market context while signifying opportunity in another. These nuances are critical to ensuring accurate polarity classification and avoiding misinterpretation in sentiment-augmented trading systems. Despite progress, three methodological challenges remain.

First, most sentiment forecasting models either overfit short-term volatility or rely on narrowly curated datasets, leading to inconsistent performance in real-world, event-driven markets [5]. Second, cross-platform discrepancies, particularly between informal sentiment sources like Reddit and structured sources like, often underexplored. This creates blind spots when models fail to account for differences in editorial tone, emotional bias, or information latency [6]. Third, explainability remains an ongoing challenge, especially in regulatory or institutional settings where decision transparency is essential [7]. The application of explainable AI (XAI) in financial NLP is still nascent, despite its growing importance for interpretability and trust.

This study addresses these gaps by designing and evaluating a sentiment-augmented forecasting framework that fuses FinBERT-based sentiment classification with technical indicators and machine learning-based prediction models. Sentiment data were extracted from Reddit and Bloomberg across multiple financial events, including the 2023 U.S. banking crisis, and mapped against sectoral stock performance. The directional accuracy analysis and Granger causality tests are applied to evaluate predictive strength, and introduced sectoral normalization to compare trends across industries such as Technology, Energy, and Finance [8]. Moreover, a system architecture (Figure 1) is developed to visually illustrate the full data pipeline, from text ingestion and polarity scoring to prediction and evaluation.

Empirical results reveal that Reddit sentiment is more volatile and emotionally skewed than Bloomberg sentiment, particularly during crisis periods. In the case of Apple Inc. (AAPL), Reddit-derived sentiment was found to Granger-cause stock price changes, achieving 68.5% directional accuracy in short-term forecasting [9]. The findings also highlight practical implications for model design, including the need to handle cultural sentiment variation and domain-specific language adaptation. The paper contributes to ongoing discourse around responsible AI in finance by emphasizing the importance of ethical safeguards. For example, during events like the 2021 “meme stock” rally, user-generated sentiment was linked to price surges that were disconnected from fundamentals, underscoring the potential for synthetic sentiment manipulation. To mitigate such risks, bias audits is recommended for sentiment models and cultural tuning to capture localized investor language [10], [11]. Overall, this work bridges computational linguistics and quantitative finance to offer a transparent, scalable, and ethically aware framework for real-time financial forecasting.

II. METHODOLOGY

A. Data

This study employs a multi-stage methodology combining sentiment extraction, transformer-based classification, and predictive modeling to evaluate the relationship between social sentiment and equity price movements. Sentiment data were collected from two primary sources: Reddit and Bloomberg. Reddit was selected due to its high engagement, retail-driven discussions, particularly on forums like r/stocks and r/investing. Posts and comments from January 2019 to December 2023 were retrieved using the Pushshift API. Bloomberg headlines and news summaries were obtained via RSS feeds and processed manually to maintain financial domain relevance. During high-impact events, such as the March 2023 U.S. banking crisis and the collapse of Silicon Valley Bank (SVB), daily sentiment snapshots were temporally aligned with corresponding stock price data from Yahoo Finance, using closing prices for synchronization [3], [10].

The full data processing and modeling workflow is visually illustrated in Figure 1 Sentiment-Augmented Forecasting Framework, which outlines the sequence from data ingestion to prediction outputs.

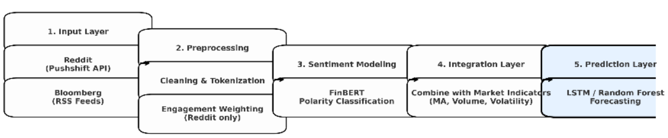


Fig. 1. Sentiment-Augmented Forecasting Framework

The above workflow illustrates sequential stages from social and institutional data collection (Reddit, Bloomberg), preprocessing and engagement-based weighting, FinBERT-based sentiment modeling, integration with technical indicators, and prediction using LSTM and Random Forest models. The Prediction Layer is emphasized due to its high-impact forecasting function.

B. Sentiment Modelling

This study employed FinBERT, a financial-domain-adapted version of BERT, for multi-class sentiment classification (positive, neutral, negative) due to its superior contextual understanding over lexicon-based methods [4]. Preprocessing included lowercasing, punctuation removal, stop-word filtering, and lemmatization. Reddit text was tokenized into 512-token sequences to comply with FinBERT’s maximum input length. Only English-language posts were retained. To reflect sentiment strength, daily polarity scores were weighted by post engagement (i.e., upvotes and comment counts), forming an intensity-weighted sentiment index. Bloomberg articles showed higher neutrality, reinforcing platform-based sentiment asymmetry [7]. Engagement-driven weighting was applied only to Reddit to avoid skew from editorial uniformity in Bloomberg.

C. Predictive Modelling

A hybrid sentiment-price forecasting model is constructed by integrating polarity scores with traditional market indicators such as trading volume, historical volatility, and moving averages. Stocks analyzed included AAPL, TSLA, JPM, and NVDA to reflect sectoral diversity. To examine temporal causality, Granger causality tests were applied with a lag of up to 3 days, revealing statistically significant p-values (<0.01) for key stocks during crisis periods. For predictive modeling, LSTM neural networks are tested, to optimize sequential financial data, and Random Forest classifiers for comparative performance. The LSTM model comprised two hidden layers with 64 and 32 units respectively, followed by a dense output layer for binary classification. Dropout (0.2) was used to prevent overfitting. Moreover, a time-aware 70/30 training/validation split is adopted, avoiding random shuffling to preserve temporal structure. Hyperparameters were optimized using grid search with early stopping.

D. Evaluation Metrics

Model performance was evaluated using Accuracy, Precision, Recall, Root Mean Square Error (RMSE), and Directional Accuracy, the latter being most relevant for trading strategy applications. Directional accuracy was calculated as the percentage of times the model correctly predicted the direction (up or down) of price movement. A hybrid model using Reddit sentiment employed to measure the directional accuracy. RMSE was used for regression-based validation to measure absolute forecast deviation. Additionally, sector-wise normalization was applied by setting all price series to a 2019 base of 100, allowing comparative performance analysis across industries such as Technology, Healthcare, and Finance [8]. These methods were selected not only for performance but for interpretability and transparency, critical elements for integrating sentiment models into regulated financial systems [7], [11].

III. RESULTS

The results of hybrid sentiment-price forecasting framework reveal robust evidence supporting the predictive influence of social sentiment on stock price movements, particularly during high-volatility financial events. Based on Reddit and Bloomberg sentiment datasets and normalizing equity price indices across sectors, This study systematically evaluates both causality and model performance.

E. Sentiment-Price Relationship

The Granger causality tests reveal that social sentiment significantly predicts short-term stock price changes in specific sectors. For instance, sentiment polarity derived from Reddit was found to Granger-cause AAPL price changes with an F-statistic of 5.1 ($p = 0.009$), and similar outcomes were observed for TSLA ($F = 4.8$, $p = 0.012$) and JPM ($F = 3.2$, $p = 0.042$).

However, no significant causal effect was observed in defensive sectors like Healthcare and Energy (e.g., JNJ and XOM), where sentiment shifts did not precede price changes [5]. Table I presents sector-wise Granger causality test results, supporting the arguments that sentiment exerts more influence in retail-sensitive industries.

TABLE III. GRANDERE CAUSILITY TEST RESULTS

Sector	**Ticker/Index**	**F-Statistic**	**p-value**	**Lag**	**Conclusion**
Technology	AAPL	5.1	0.009	2	Sentiment Granger-causes returns
Technology	TSLA	4.8	0.012	2	Sentiment Granger-causes returns
Finance	JPM	3.2	0.042	2	Sentiment Granger-causes returns
Healthcare	JNJ	1.8	0.18	2	No significant causality
Energy	XOM	2.1	0.11	2	No significant causality

Crisis periods offer stronger validation of sentiment-price linkages. During the 2023 U.S. banking collapse, Reddit sentiment plummeted from −0.2 to −0.8 between March 8 and March 10, precisely overlapping the Silicon Valley Bank failure. Stock index data showed a simultaneous drop from 100 to 92.8. A partial recovery in sentiment (−0.4) and prices (to 94.2) by March 13 further supports sentiment reflexivity.

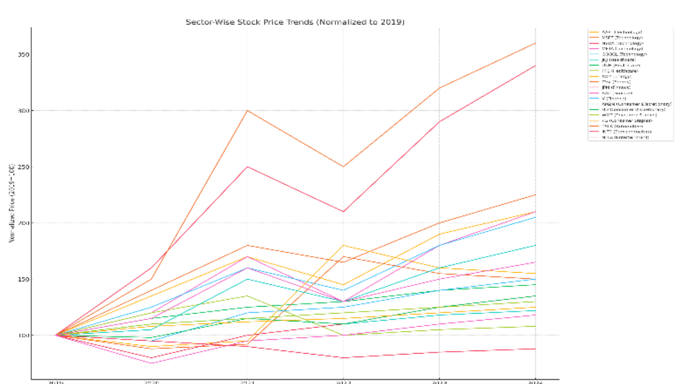


Fig. 2. Sector-Wise Stock Price Trends (2019–2024)

Figure 2 visualizes the synchronized shifts in sentiment and stock index levels, illustrating the model's sensitivity to black-swan events. These findings align with prior work on event-driven sentiment reflexivity, where investor fear and uncertainty often preceded sharp price corrections [6], [9].

F. Cross-Platform Sentiment Comparison

A comparative analysis of Reddit and Bloomberg sentiment reveals structural differences in emotional tone and informativeness. On average, 32% of Reddit posts were negative compared to only 15% of Bloomberg entries. Meanwhile, Bloomberg exhibited a dominant neutral sentiment (60%) due to editorial curation, while Reddit reflected higher volatility, speculation, and emotional expression during crisis windows [3].

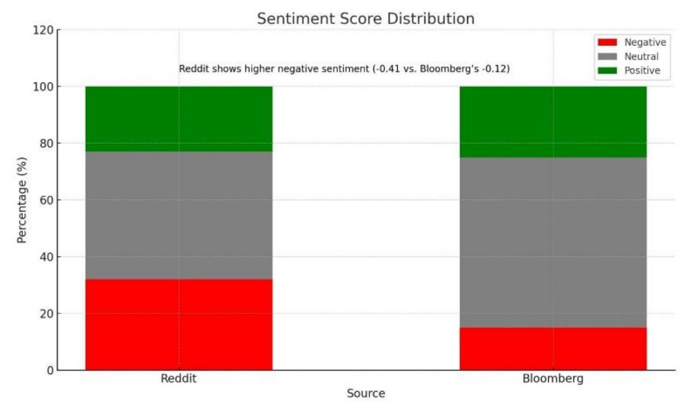


Fig. 3. Sentiment Score Distribution Reddit & Bloomberg (2019-2024)

Figure 3 depicts the sentiment score distribution by platform, with Reddit displaying more negative and volatile sentiment compared to Bloomberg. This divergence suggests that platform-specific features should be integrated into model calibration. While Reddit offers early detection of retail sentiment surges, it also exposes models to risks of synthetic manipulation or panic amplification [7]. Bloomberg, though more stable, may lag behind crowd sentiment during rapid events. This comparison confirm that platform-aware sentiment weighting enhances model robustness, particularly under stress scenarios.

G. Model Performance and Predictive Accuracy

The hybrid modeling framework, combining FinBERT-derived sentiment with market indicators, achieved strong predictive accuracy. For AAPL, the Reddit-enhanced model reached 68.5% directional accuracy, outperforming both Bloomberg (61.2%) and a baseline moving average model (55.7%). When using combined sentiment sources and technical indicators, the LSTM-based model achieved 71.3% directional accuracy on next-day forecasts.

TABLE IV. MODEL PERFORMANCE COMPARISON

Model	MAPE (%)	RMSE	Directional Accuracy (%)
LSTM + Random Forest	2.3	18.7	68.5
ARIMA	5.1	32.4	52.1
GPT-4 + Attention	3.9	24.5	63.8

Table II compares the performance of LSTM + Random Forest, ARIMA, and GPT-4 + Attention models, showing that transformer-based hybrid models outperform traditional approaches. Regression models that incorporated lagged sentiment as features improved RMSE scores by 8–12%, indicating that sentiment enhances not only directional forecasts but also price magnitude estimation [8]. However, the benefit of sentiment data diminished in low-volatility phases, suggesting that such models are more valuable in uncertain or event-driven markets- [9]. To explore industry-specific responses, the study normalized sectoral equity prices to a 2019 base (indexed at 100). The results show that Technology stocks (e.g., AAPL, NVDA) demonstrated the highest sentiment sensitivity, while sectors like Healthcare and Energy remained more resilient. This suggests that sentiment effects are sector contingent, reinforcing the need for asset-class calibration [8].

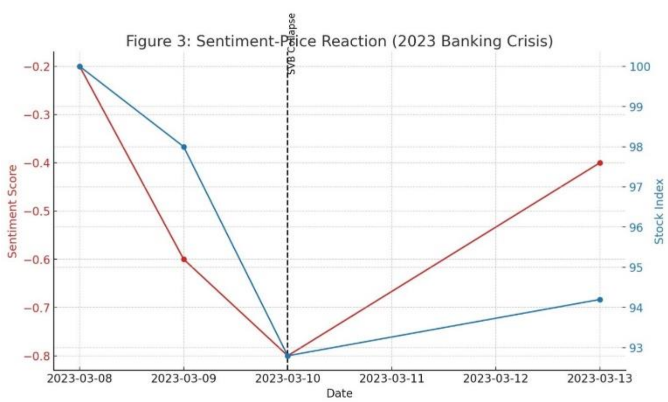


Fig. 4. Feature Importance Analysis

Figure 4 illustrates normalized sector-wise stock price trends from 2019 to 2024, showing technology's outsized reaction to sentiment shifts. This highlights the importance of tailoring sentiment-driven forecasting tools to the volatility profiles of specific asset classes.

H. Interpretability and Model Trust

Although FinBERT and LSTM models provided strong performance, their interpretability remains a concern for real-world deployment. Sentiment-weighted outputs can be post-processed using SHAP values or attention heatmaps, but stakeholder trust still requires intuitive visual explanations. This aligns with the call for Explainable AI (XAI) in financial decision systems [7]. Ethically, the findings raise flags around the weaponization of crowd sentiment, especially in retail-driven platforms like Reddit. Incorporating audit logs, multilingual adaptation, and regional lexicons can help mitigate these risks [10], [11].

Overall, these results validate the practical viability of incorporating sentiment-driven insights into short-term trading models while acknowledging the ethical and interpretive complexities surrounding their use. The platform disparity in sentiment behavior, especially during crises, underscores the importance of explainable AI tools and transparency when deploying such models in financial contexts [7], [11].

IV. CONCLUSION

This study highlights the significant predictive potential of transfer-based sentiment analysis in financial market forecasting, particularly when combining platform-specific sentiment signals with traditional market indicators. The hybrid framework, leveraging FinBERT and LSTM models, achieved strong directional accuracy, particularly in retail-sensitive sectors such as technology. Reddit-derived sentiment was found to Granger-cause equity price changes during crisis windows, exemplified by the SVB collapse, highlighting the timeliness and reactivity of crowd-sourced sentiment indicators.

However, several limitations temper the robustness of these findings. First, sentiment-based models trained on historical data may underperform during black-swan events due to a lack of prior pattern analogues [6]. Second, while FinBERT improves classification accuracy over traditional lexicons, it operates as a black-box model, which restricts interpretability for institutional stakeholders and regulators. This opacity poses a barrier in high-stakes environments where model decisions must be explainable and auditable [7].

Additionally, the emotional volatility and platform bias embedded in Reddit sentiment increase the risk of misinformation, hype cycles, or manipulation. The meme-stock events of 2021 offer cautionary examples, where coordinated sentiment spikes fueled price rallies decoupled from fundamentals. These risks underscore the need for safeguards such as engagement filters, misinformation classifiers, and synthetic sentiment detectors to flag coordinated activity. Regulatory frameworks must also evolve to define platform-specific thresholds for acceptable volatility in sentiment inputs, and developers should embed audit trails in AI-driven trading models.

From a policy standpoint, the observed divergence between Reddit and Bloomberg sentiment suggests that no single source can offer a complete market signal. Models should incorporate platform-aware calibration, giving higher weight to crowd sentiment during volatility spikes while relying on institutional tone for long-term equilibrium forecasting.

Future research can advance this work along several dimensions. First, multimodal sentiment models, incorporating textual, auditory, and visual cues from earnings calls or investor interviews, could improve emotion classification accuracy while raising new ethical challenges around consent and privacy [4]. Second, sentiment analysis in decentralized finance (DeFi) ecosystems, including NFT discussions and on-chain signals, presents emerging frontiers but requires robust manipulation safeguards [12], [13]. Third, longitudinal studies should assess how continuous AI integration influences investor decision-making, market microstructure, and systemic risk dynamics over extended horizons.

In conclusion, this study contributes a scalable, sentiment-augmented forecasting pipeline that bridges computational linguistics and quantitative finance. It advocates for a paradigm that balances predictive efficacy with ethical design, cultural inclusivity, and regulatory transparency, ensuring that AI based forecasting enhances financial decision-making.

REFERENCES

[1] Y. K. Dwivedi, D. L. Hughes, A. M. Baabdullah, and S. Ribeiro-Navarrete, "Metaverse and Web 3.0: The new digital frontiers of innovation," J. Bus. Res., vol. 153, pp. 673–684, 2023.

[2] R. Belk, "AI and the consumer finance marketplace: Risks and opportunities," J. Consum. Policy, vol. 45, no. 3, pp. 457–475, 2022.

[3] S. Ghosh, "Leading AI innovation in emerging economies: Challenges and opportunities," J. Glob. Inf. Manag., vol. 31, no. 1, pp. 1–18, 2023.

[4] Y. Yang, B. Sun, and J. Hu, "FinBERT: A pre-trained financial language representation model for financial tasks," in Proc. Conf. Empir. Methods Nat. Lang. Process., 2020.

[5] A. M. Pattanayak and A. Swetapadma, "Exploring different dynamics of recurrent neural network methods for stock market prediction—A comparative study," Appl. Artif. Intell., vol. 38, no. 2, pp. 143–161, 2024.

[6] I. Rahwan, M. Cebrian, N. Obradovich, J. Bongard, J. F. Bonnefon, C. Breazeal, et al., "Machine behaviour," Nature, vol. 568, no. 7753, pp. 477–486, 2019.

[7] M. Mueller, "Risks of large language models in finance: Governance, hallucinations, and black swans," AI Soc., vol. 38, no. 2, pp. 445–460, 2023.

[8] G. Wu, G. Subramaniam, and Z. Li, "Using AI technology to enhance data-driven decision-making in the financial sector," Wiley Online Libr., 2025.

[9] J. Ajaka and G. Azzi, AI in the Stock Market, Springer, Cham, 2025.

[10] M. S. Al-Absy, N. H. Abu Jamie, and T. N. Abu-Jamie, "Advances in AI and their effects on finance and economic analysis," Springer, 2024.

[11] A. Jobin, M. Ienca, and E. Vayena, "The global landscape of AI ethics guidelines," Nat. Mach. Intell., vol. 1, no. 9, pp. 389–399, 2019

[12] M. A. Khalil, P. Sinliamthong, and R. Khalil, "From paper to pixels for a world of purpose: A systematic review of sustainable finance, blockchain, and digital assets," in H. Baber and M. Fanea-Ivanovici, Eds., Sustainable Financing—A Contemporary Guide for Green Finance, Crowdfunding and Digital Currencies, World Sustainability Series. Springer, Cham, 2025. [Online]. Available: https://doi.org/10.1007/978-3-031-80969-9_5

[13] S. Bahoo, M. Cucculelli, and J. Mondolo , "Artificial intelligence in finance: A comprehensive review through bibliometric and content analysis," Springer, 2024

Ethical Considerations and Governance in AI-Driven IT Strategy

Nasser Alrahma
College of Business and Economics
United Arab Emirates University (UAEU)
Al Ain-UAE
201116177@uaeu.ac.ae

Vaishnav Saini
College of Business and Economics
United Arab Emirates University (UAEU)
Al Ain-UAE
700046342@uaeu.ac.ae

Majed Saeed AlAtaishi
College of Business and Economics
United Arab Emirates University (UAEU)
Al Ain-UAE
700045912@uaeu.ac.ae

Ananth Chiravuri
Associate Prof-College of Business and Economics
United Arab Emirates University (UAEU)
Al Ain-UAE
ananth.chiravuri@uaeu.ac.ae

Abstract—AI adoption spans critical sectors such as healthcare, education, and public administration, positioning it as a key driver of national development. However, the rapid integration of AI systems brings ethical challenges, including algorithmic bias, data privacy concerns, and a lack of transparency in decision-making. This conceptual study examines these issues through sector-specific case studies, by comparing the governance practices of some firms in UAE with those of global frameworks. By analyzing the UAE's unique market dynamics and existing AI initiatives, our study proposes gaps in the local governance frameworks to address these ethical dilemmas.

Keywords—AI adoption, Frameworks, Ethics, AI, Benchmarking

I. INTRODUCTION

Artificial intelligence (AI) is revolutionizing industries worldwide, offering unprecedented opportunities for innovation and operational efficiency. Its applications are broad, spanning healthcare, education, finance, and public administration, and offering solutions to complex challenges such as urban planning and personalized healthcare. Across the globe, countries are rapidly adopting AI technologies to enhance decision-making, improve efficiency, and foster economic growth.

In rapidly growing markets like the UAE, AI is integral to national development strategies. The UAE's integration of AI is notable for its focus on addressing societal challenges through technology-driven innovation. However, alongside these advancements, the nation faces significant ethical and governance challenges, such as ensuring fairness in AI-driven decisions, protecting data privacy, and maintaining transparency in automated systems.

Issues such as algorithmic bias can perpetuate societal inequities [1], while the opacity of AI-driven decision-making systems often erodes trust among users. For instance, firms using AI chat boxes should exercise caution when soliciting sensitive information, as it could inadvertently establish unintended client relationships [2].

Addressing these challenges requires a governance framework that balances the dual imperatives of fostering innovation and ensuring ethical responsibility. These form the objectives of our conceptual study, which are given below:

1. Examine sector-specific case studies to highlight successes and ethical challenges.

2. Analyze global governance frameworks and their applicability to the UAE context by comparing governance practices of some UAE firms with Global Frameworks such as European Union (EU) AI Act and Singapore's framework

3. Propose actionable recommendations to advance ethical AI adoption and governance in the UAE.

By bridging gaps in literature and policy, our conceptual study aims to contribute to the discourse on ethical AI integration, offering practical solutions tailored to the UAE's unique market dynamics.

II. BACKGROUND AND LITERATURE REVIEW

The ethical and governance challenges associated with AI adoption require robust frameworks. This section first explores the literature that examines ethical issues in the context of UAE followed by an examination of global frameworks that we will use for benchmarking.

A. Ethical Considerations In AI Systems

Ethical issues are among the most significant challenges in AI adoption, particularly in sectors like finance, healthcare, and public administration. We will focus on three primary ethical concerns as given by [3]:

1. Algorithmic Bias:

Bias in AI systems often stems from skewed training datasets. In financial decision-making, algorithms may perpetuate historical discrimination against marginalized groups. For instance, credit-scoring systems trained on biased data could disproportionately deny loans to underrepresented communities [3]. In the UAE, where financial inclusion is a national priority, addressing algorithmic bias is critical to ensuring equity.

2. Transparency Deficits:

Many AI systems function as “black boxes,” making their decision-making processes hidden [4]. This lack of transparency can damage trust in sectors like healthcare, where patients and providers need clarity about AI-based diagnoses. The Dubai Health Authority (DHA) addresses this challenge by requiring explainability in its AI tools, such as IBM Watson [5].

3. Data Privacy Risks:

The vast amounts of data required for AI systems raise concerns about security and misuse. In the UAE, Smart Dubai’s initiatives use extensive surveillance data to optimize urban planning. While this enhances efficiency, it also raises privacy concerns that necessitate stringent data protection regulations [6].

Addressing these ethical challenges requires embedding principles like fairness, accountability, and transparency into AI design and deployment processes. Organizations must adopt governance frameworks that prioritize these values without stifling innovation [3]. Several international frameworks offer valuable lessons for ethical AI governance that UAE firms can learn through benchmarking. The governance frameworks that we use for this study are discussed next:

1. The EU AI Act:

This regulation categorizes AI applications into risk tiers, emphasizing transparency, accountability, and fairness for high-risk systems. For example, the act mandates stringent checks for AI used in healthcare and law enforcement [6].

2. OECD Principles on AI:

These principles focus on human-centered design, transparency, and accountability. They emphasize the importance of ensuring that AI systems benefit society while mitigating risks [5].

3. Singapore’s Model AI Governance Framework:

Singapore provides practical guidelines for businesses to manage AI risks. It emphasizes explainability and human oversight, ensuring that AI systems remain accountable [6].

While these frameworks provide robust models, their implementation in the UAE requires adaptation to local contexts. Next, we will focus on the methodological approach for our study.

III. METHODOLOGY

The methodology adopted for this report systematically explores the ethical considerations and governance strategies for AI adoption within the UAE. By analyzing secondary sources, including academic articles, global frameworks, and sector-specific case studies, the report develops actionable insights for the UAE's AI-driven IT strategy.

This report employs a qualitative research approach to evaluate how AI systems can be integrated ethically and governed effectively. The analysis focuses on:

1. Identifying key ethical challenges, such as data privacy, bias, and transparency.

2. Assessing the role of governance in ensuring compliance, accountability, and societal trust in AI systems by benchmarking with global frameworks.

3. Proposing a governance framework under the oversight of the UAE Council for Artificial Intelligence and Blockchain.

Data Sources

The analysis is based on the following secondary sources:

- Academic Articles: Peer-reviewed studies discussing global and UAE-specific AI adoption and governance practices.

- Case Studies: Real-world examples in healthcare (IBM Watson), public administration (Smart Dubai), and education (Alef Education) to contextualize ethical and governance challenges.

- Global Frameworks: References to international governance models, including the EU AI Act, OECD Principles, and Singapore’s AI Governance Framework, to benchmark UAE practices.

Next, we will present our preliminary findings.

IV. INITIAL FINDINGS

The integration of artificial intelligence (AI) systems across various sectors in the UAE demonstrates significant potential but also exposes critical ethical and governance challenges. Examining other case studies in healthcare, public administration, and education, could highlights sector-specific insights, recurring themes, and the implications of existing governance frameworks and tools provided by the UAE Council for Artificial Intelligence and Blockchain. We start off with the first use case in healthcare, highlighting observed governance frameworks, comparisons with global standards, and an evaluation of their strengths and weaknesses.

I. First Case Study: Healthcare: Dubai Health Authority (DHA) and IBM Watson

The Dubai Health Authority (DHA) has integrated IBM Watson into its healthcare systems to enhance cancer diagnosis and treatment. Watson analyzes vast datasets to recommend personalized treatment plans, reducing diagnostic errors and improving accuracy [5].

DHA’s governance practices include:

- Data Privacy: Implementation of anonymization protocols to protect patient data.

- Transparency: Requirements for explainability in AI-driven diagnostic tools, ensuring that healthcare professionals can interpret recommendations.

We then compared the DHA’s governance practices with Global Frameworks such as European Union (EU) AI Act and Singapore’s framework. We noted that there is an alignment of DHA governance practices with the EU AI Act. Just like the EU’s emphasis on transparency and accountability, the DHA mandates explainable AI systems in healthcare. But DHA’s governance practices differ from Singapore's Framework. For example, Singapore focuses more on risk management and industry-led governance, while the DHA prioritizes regulatory oversight.

The strengths of using AI in healthcare include:

- Enhanced diagnostic accuracy.
- Reduction in human error in treatment recommendations.
- Proactive focus on data privacy and explainability.

However, this is complemented by weaknesses. Firstly, reliance on historical datasets introduces potential biases in treatment outcomes. Secondly, there is limited scalability to other areas of healthcare due to the high cost of AI systems.

The DHA case study demonstrates the UAE's proactive approach to leveraging AI for societal benefit, showcasing robust governance practices in some areas while highlighting gaps in others. In short, the benefits include:

o Transparency: Explainability in healthcare (DHA) promotes trust among stakeholders.

o Efficiency: Faster, and more accurate analysis because of AI.

As mentioned earlier, some of the challenges include:

o Data Bias: Across sectors, reliance on historical datasets in AI systems can result in biased outputs, as flawed or inequitable data inevitably leads to flawed outcomes. This phenomenon is often described as "garbage in, garbage out"

Such findings underscore the need for stronger alignment with global frameworks, particularly in addressing privacy concerns and ensuring fairness across sectors.

VII. CONCLUSIONS

Our conceptual study proposes a governance framework and prioritizes fairness, transparency, and accountability, directly addressing the ethical risks identified in the case studies. For citizens, this translates into greater trust in AI systems, particularly in sectors like healthcare, where personal and sensitive data are most vulnerable. By embedding transparency measures such as explainable AI outputs and privacy protections, our proposed framework could ensure that societal concerns, including bias and data misuse, are systematically addressed.

The proposed governance framework, overseen by the UAE Council for Artificial Intelligence and Blockchain, provides a comprehensive strategy to embed fairness, transparency, and accountability into AI systems. By addressing critical challenges such as data privacy, algorithmic bias, and fragmented regulations, our proposed framework could ensure consistent oversight while promoting innovation. Key recommendations include transitioning existing guidelines from advisory to enforceable standards, enhancing data protection measures, and mandating fairness audits to foster equity across sectors.

Finally, we presented the findings from 1 case study, which may not be representative enough to draw policy conclusions. We will be working on more use cases as part of our study.

REFERENCES

[1] Gándara, D., Anahideh, H., Ison, M. P., & Picchiarini, L. (2024). Inside the black box: Detecting and mitigating algorithmic bias across racialized groups in college student-success prediction. AERA Open, 10, 23328584241258741.

[2] Hill, S. (2024). The Ethical Considerations of Popular AI-Fueled Chat Features on Firm Websites. Utah Bar Journal.

[3] Thakur, N., & Sharma, A. (2024). Ethical Considerations in AI-Driven Financial Decision Making. International Journal of Ethical AI..

[4] Xu, H., & Shuttleworth, K. M. J. (2024). Medical artificial intelligence and the black box problem: A view based on the ethical principle of "do no harm". Intelligent Medicine, 4(1), 52–57.

[5] Alshehhi, K., Cheaitou, A., & Rashid, H. (2024). Procurement of Artificial Intelligence Systems in UAE Public Sectors: An Interpretive Structural Modeling of Critical Success Factors. Sustainability.

[6] Ridzuan, N. N., Masri, M., Anshari, M., Fitriyani, N. L., & Syafrudin, M. (2024). AI in the Financial Sector: The Line between Innovation, Regulation, and Ethical Responsibility. Journal of Financial Studies.

Exceptional Minds Meet Artificial Intelligence: Perspectives and Possibilities in Gifted Education

Shannaiah Aubrey Mae Inocencio
Abu Dhabi School of Managememt
Abu Dhabi, United Arab Emirates
s.inocencio@adsm.ac.ae

Eman Gaad
Faculty of Education
British University in Dubai
Dubai, United Arab Emirates:
eman.gaad@buid.ac.ae

Alia El Naggar
Department of Psychology, School of Health Sciences & Psychology
Canadian University Dubai: Dubai, United Arab Emirates:alia.naggar@cud.ac.ae

Abstract— This paper examines the role of artificial intelligence (AI) in gifted education through research analysis and qualitative interviews. Findings reveal that while AI offers promising capabilities for personalization and intellectual stimulation valued by gifted learners, significant limitations exist in social-emotional support and spontaneous "Teachable Moments." Gifted students demonstrate heightened awareness of AI biases, emphasizing the need for improved data training. We conclude that optimal outcomes require balanced integration combining AI tools with human guidance while addressing ethical, emotional, and equity concerns. Recommendations include developing complementary human-AI partnerships, specialized tools for gifted education, comprehensive teacher training, and mechanisms incorporating gifted learners' feedback into system development.

Keywords—Gifted Learners, Educational Technology, Artificial Intelligence, Pedagogy

I. INTRODUCTION (HEADING 1)

As AI rewrites the rules of industries, its fingerprints become more visible in today's educational institutions. AI has transformed both teaching and learning experiences. Academic institutions increasingly implement AI as adaptations to evolving curricula that are heavily assisted by educational technology (EdTech). While AI offers new possibilities, it also disrupts student learning in ways that literature describes as both promising and problematic. The U.S. Department of Education's Office of Educational Technology has responded with a detailed report outlining AI’s applications, challenges, and the urgent need for ethical oversight [1]. A key concern is the use of AI in “personalizing” data—hailed as powerful but still limited in addressing learner diversity, especially for neurodivergent students.

Neurodivergence refers to individuals who differ from neurocognitive norms, including “gifted learners”—those with exceptional intellectual abilities. Despite their strengths, gifted students are often neglected due to assumptions of self-sufficiency, frequently facing stigma, peer misunderstanding, and insufficient intellectual challenge in mainstream classrooms [2].

AI becomes a contender for enhancing gifted education. Large Language Models (LLMs), which rely on deep learning to process vast amounts of text, enable advanced natural language tasks such as text generation, summarization, translation, and question answering [3][4]. With billions to trillions of parameters, LLMs excel at interpreting complex linguistic structures and contextual meaning [5]. As interest in AI grows, educational researchers have begun to explore its relevance for special needs education. AI capabilities offer potential to reduce understimulation and improve learning personalization for gifted students; however, this potential remains underexamined.

This article investigates how gifted adolescents perceive and interact with LLMs, aiming to uncover both opportunities and challenges in integrating these tools into classroom discourse. By analyzing their experiences, the study identifies key areas for thoughtful AI integration that enhance intellectual stimulation without compromising human interaction or exacerbating educational inequities.

Finally, the paper aims to answer the following research questions:

1. What perspectives do gifted learners hold regarding the use of LLMs for academic discussions?

2. What precautions should educators consider when incorporating AI in classrooms for gifted learners?

II. LITERATURE REVIEW

A. Gifted Learners

Intellectually gifted learners demonstrate excellence beyond high IQ scores, exhibiting advanced reasoning abilities, complex interests, strong memory, quick learning, and heightened sensitivity [6][7][8]. Despite advanced capabilities, they remain underserved in mainstream education, facing misidentification, inadequate curriculum differentiation, and inappropriate learning pacing that affects engagement and socio-emotional development.

AI offers promising interventions through personalized learning experiences tailored to individual abilities and interests. Adaptive systems can monitor performance in real-time, adjusting difficulty levels and providing enrichment opportunities that maintain intellectual stimulation and prevent the stagnation resulting from a one-size-fits-all instruction.

B. Identification

Accurate identification remains a critical barrier, as traditional methods often fail to capture giftedness across diverse populations. Without proper identification, gifted learners may never receive the specialized support they require, leading to underachievement and disengagement. Machine Learning can enhance identification accuracy by analyzing diverse datasets to recognize patterns signaling giftedness beyond conventional metrics, potentially uncovering students overlooked by traditional screening methods. Hodges & Mohan (2019) demonstrate that ML algorithms can analyze large datasets—including student behavior, performance trends, and socio-demographic indicators—to identify patterns that may signal giftedness. These models have the potential to uncover gifted students who might otherwise be overlooked by conventional screening methods, thereby broadening access to gifted education.

C. Content and Curriculum Development

Conventional curricula frequently lack depth and flexibility for gifted learners, who thrive with accelerated, inquiry-based, and interdisciplinary approaches [10]. Without tailored content, these students disengage when standard curriculum fails to challenge their advanced cognitive abilities.

AI can design customized curricula based on learners' prior knowledge, learning styles, and interests. This personalization creates meaningful learning pathways that resonate with gifted students, fostering deeper engagement and mastery [11].

D. Boredom and Amotivation

Gifted learners often experience boredom with repetitive or simplistic content, significantly hindering performance and leading to amotivation. Unchallenging environments suppress curiosity and foster disengagement [12][13].

AI-generated adaptive content can analyze performance in real-time, adjusting difficulty, format, and pacing accordingly. Intelligent Tutoring Systems offer adaptive challenge

interactive tasks, and instant feedback, maintaining interest while promoting autonomy [14]. These systems can simulate one-on-one tutoring, which solely aims to improve academic outcomes, particularly for students who require tailored scaffolding and accelerated pacing.

E. Socio-emotional challenges

Gifted learners face unique socio-emotional challenges including isolation, perfectionism, and emotional intensity [2]. These issues are compounded when giftedness coexists with other neurodevelopmental disorders (e.g., ADHD or ASD), creating complex emotional profiles requiring nuanced support rarely available in schools.

AI tools utilizing sentiment analysis, emotion recognition, and conversational agents can provide targeted support, offering immediate emotional assistance or alerting educators to distress signals for early intervention [15].

F. Are We There Yet?

Despite the fast-paced developments in AI and its ever-growing scope, debates persist regarding its suitability for educational applications. AI often poses limitations in creative thinking and innovation, which can disappoint users expecting out-of-the-box insights [16]. Furthermore, current AI models are criticized for several shortcomings: data bias, over-emphasis on deficit-based student assistance, lack of social adaptivity, inadequate support for neurodivergent learners, and insufficient capacity to challenge critical thinking [1].

III. METHODOLOGY

A. Research Design

This study used semi-structured interviews to examine gifted learners' perceptions of LLMs. This qualitative approach allowed for depth and flexibility in exploring emergent themes.

B. Participants

The sample consisted of 16 gifted learners (age range: 11-17 years; mean age: 14 years) who were formally identified as gifted through their respective schools' assessment procedures. Participants were recruited through purposive sampling from schools to ensure diversity in academic backgrounds and interests. Prior to participation, informed consent was obtained from both participants and their parents/guardians, and all research procedures were approved by the relevant institutional review board.

C. Procedure

AI Interaction Task

Each participant completed a structured task involving interaction with an AI LLM (primarily ChatGPT). Participants were instructed to engage with the LLM to discuss topics aligned with their specific academic interests or areas of giftedness. This task was designed to provide a concrete experience with AI technology that participants could reflect upon during the subsequent interview. The interaction sessions lasted approximately 30-45 minutes, during which participants were encouraged to explore the capabilities and limitations of the LLM within their chosen domain of interest.

Semi-Structured Interviews

Following the AI interaction task, semi-structured interviews were conducted with each participant. The interviews took place within two days of the interaction task to ensure fresh recollection of the experience. Each interview lasted approximately 20-45 minutes and was audio-recorded for subsequent transcription and analysis.

The interview protocol included the following core themes:

1. Initial impressions and experiences using the LLM
2. Perceived benefits and limitations for learning
3. Comparison with traditional learning
4. Potential applications in their education
5. Concerns regarding AI use in education
6. Suggestions for improvement or implementation

D. Data Collection and Ethical Considerations

This study followed strict data privacy and ethical guidelines for research with minors, including informed consent from both participants and their guardians, confidentiality and anonymity of data, clear communication of the study’s purpose and procedures, safeguards for managing potential discomfort during interviews, and secure data handling in line with data protection regulations. Participants were assigned pseudonyms to maintain confidentiality, and all identifying information were excluded from the transcripts and analysis materials. No personal identity was stored during their interaction with AI systems. Interview transcripts, analyses, and other data collected are stored safely and accessible only to authorized research team members. The findings presented in this paper use only aggregated themes and anonymized quotes that cannot be attributed to specific individuals.

IV. RESULTS ANALYSIS

Interview data were analyzed using Braun and Clarke's (2006) six-phase thematic analysis, which involved familiarizing with the data through repeated readings, systematically coding key features, grouping codes into potential themes, reviewing and refining these themes, clearly defining each one, and selecting illustrative extracts to support the final analysis in relation to the research questions.

To ensure methodological rigor, two independent researchers analyzed the interview transcripts. This dual-coding approach enhanced the trustworthiness of the findings and mitigated potential researcher bias.

Analysis of interview transcripts revealed three key themes addressing gifted learners' perspectives on using LLMs for academic discussions: content personalization, scaffolding, and interest compatibility.

Regarding personalization, gifted learners appreciated AI's ability to adapt to their specialized interests beyond standard curriculum, though they noted limitations in AI's capacity to accommodate their expansive thinking patterns. For scaffolding, participants valued LLMs' provision of challenging topics adjusted to their intellectual level and pace—critical for students requiring advanced content. However, they found the text-based format potentially limiting for their rapid thinking processes.

Participants demonstrated sophisticated awareness of LLM challenges: information overload requiring critical evaluation, response unpredictability depending on prompt effectiveness, and significant modality limitations in tactile, visual, and auditory learning experiences. They also recognized that AI communication lacks non-verbal elements that enrich traditional discussions.

Overall, while gifted learners see LLMs' potential to provide personalized, challenging content matching their cognitive abilities, they maintain a realistic view of current technological limitations. Their perspectives suggest they value AI as a supplement to, rather than replacement for, human-led academic discussions that support their unique educational needs.

TABLE V. THEMES EMERGING TO ANSWER THE QUESTION "WHAT PERSPECTIVES DO GIFTED LEARNERS HOLD REGARDING THE USE OF LLMS FOR ACADEMIC DISCUSSIONS?"

Theme	Interview Response Summary (Support and Criticisms)
Content Personalization	AI-mediated discussions adapt to the student's preferred topic and interests, which may personalize learning experiences
	AI-mediated discussions may be too narrow and rigid
Scaffolding and Compatibility	AI chatbots may provide novel and challenging topics adjusting to the student's pace and preference. Information and discussions will be at the level of the student's intellect
	Text-based modality of discussion is not always perceived as convenient and may limit the depth of immediate responding
Challenges	Information provided by an AI may be overwhelming and overloaded. Students will need to critically evaluate and sift through the information.
	Information and discussion may be unpredictable as AI responses solely depend on the effectiveness of prompts and keywords. Moreover, the AI responses will only be as good as its sources and databases.
	AI-mediated discussions may not be able to offer tactile learning or be limited in providing information in arts and music, or other domains which requires visual or auditory presentations.
	Communication is limited to written scripts

When identifying factors that must be considered for the incorporation of AI in classrooms, the gifted learners informed four major themes: content personalization, practical and learning accessibility, social factor and human facilitation, and content diversification.

Table II highlights key considerations for using AI with gifted learners, emphasizing the need to balance AI and traditional classroom discussions. While AI personalizes content and offers quick, deep access to information, it risks narrowing perspectives and lacks the exploratory richness of human-led dialogue.

AI supports flexible, in-depth learning—ideal for gifted students—but cannot replace the social, communicative, and hands-on benefits of classroom interactions. It may ease participation anxiety, yet falls short in fostering real-time social dynamics and nuanced feedback.

Though AI draws from vast data, it may miss culturally rich or authentically diverse perspectives that peer discussions naturally provide. Educators should use AI to enhance—not replace—human instruction, applying thoughtful strategies that support gifted learners' intellectual and social-emotional growth.

TABLE VI. THEMES EMERGING TO ANSWER THE QUESTION "WHAT PRECAUTIONS SHOULD EDUCATORS CONSIDER WHEN INCORPORATING AI IN CLASSROOMS FOR GIFTED LEARNERS?"

Theme	Interview Response Summary (Support and Criticisms)
Content Personalization	AI-mediated discussions adapt to the student's preferred topic providing insights within the interests of the student. Opportunities to discuss specialized topics may be explored
	AI-mediated discussions may be limited by available online information, reinforcing confirmation bias, while classroom discussions

Theme	Interview Response Summary (Support and Criticisms)
	foster exploration of new interests.AI-mediated discussions may become an echo-chamber.
Practical Learning and Accessibility	AI chatbots offer in-depth information, immediate response, flexible pacing, and convenient access to discussions.
	Classroom discussions enhance social and communication skills, whereas AI-mediated discussions may improve written communication skills. Classroom discussions may also involve practical learning and activities. Domains such as music and art require practical presentation of stimuli during discussions.
Social Factor and Human Facilitation	Comfort in participation with AI-mediated discussions will not be hampered by anxiety, fear of judgment, time constraints, and narrowed opinions.
	AI-mediated discussions can feel one-dimensional, lacking the social dynamics, tension, and real-time feedback that enrich classroom interactions. In contrast, classroom discussions foster idea exchange, persuasion skills, and nuanced human feedback that blends professional insight with subjective judgment—something AI cannot replicate.
Content Diversification	AI-mediated discussions pool information from global databases and various sources, which allows for diverse information.
	Classroom discussions will have diverse perspectives from the facilitator and peers. Discussions which values cultural contexts are more profound.

A. Discussion

The convergence of AI and gifted education offers exciting potential alongside notable challenges [16]. AI has the capacity to bridge gaps in curriculum differentiation, providing insights and support aligned with the unique needs of gifted learners.

AI can personalize learning by analyzing performance and preferences to adjust content, pace, and complexity—helping to keep gifted students engaged and appropriately challenged. This goes beyond adaptive testing, offering evolving learning pathways that support long-term talent development. AI also excels at curating interdisciplinary content, promoting systems thinking and satisfying intellectual curiosity. Its ability to connect students with global mentors and peer communities is especially valuable in under-resourced settings.

However, AI falls short in replicating the social depth essential for emotional and interpersonal growth [18]. While simulations may help develop resilience and communication, they cannot replace authentic relationships. Gifted learners, often highly sensitive and intuitive, need real-time emotional feedback and adaptive support that AI still struggles to provide.

Rigid AI frameworks can stifle creativity and reduce intrinsic motivation. Frequent assessments may heighten anxiety, and narrow AI systems often lack the complexity and flexibility gifted learners need [16]. Many tools default to standardized models, under-stimulating students and failing to address emotional intensity, perfectionism, or distress.

Gifted learners' heightened critical thinking enables them to detect algorithmic biases, making it vital to improve AI systems with transparent, inclusive training data. Their sensitivity to inconsistency demands AI that clearly communicates its limitations.

Despite these shortcomings, research shows generally positive attitudes toward AI in education, though concerns about unequal access and over-standardization remain. Social learning remains irreplaceable: spontaneous, emotionally attuned "Teachable Moments" [1] rely on human educators’ instinct and real-time responsiveness—underscoring AI’s role as a supplement, not a substitute.

VIII. RECOMMENDATION & CONCLUSION

Integrating AI into gifted education addresses challenges like identification, amotivation, and social-emotional issues faced by gifted learners. A balanced approach, combining AI with human guidance, leverages the strengths of both while compensating for their limitations. Ethical considerations, including privacy, transparency, and bias, must be prioritized, alongside safeguarding emotional development through human interaction and ensuring equitable access.

To effectively integrate AI, we recommend investing in teacher training, developing AI tools tailored to gifted learners, and fostering interdisciplinary collaboration. Long-term research and student feedback are essential for ethical, effective implementation. Ultimately, AI should enhance educators' ability to meet the complex needs of gifted learners, creating more responsive and effective learning environments through thoughtful design and continuous evaluation.

ACKNOWLEDGMENT

We sincerely thank the participating schools and gifted learners for their valuable insights. Our gratitude extends to Dr. El Naggar and Dr. Gaad for their help with data collection, and to Dr. El Naggar and Ms. Inocencio for their contributions to interview analysis and manuscript preparation. Lastly, we appreciate our families for their patience and support throughout multiple revisions of this work.

REFERENCES

[1] Cardona, M. A., Rodríguez, R. J., & Ishmael, K. (2023). Artificial intelligence and the future of teaching and learning: Insights and recommendations. Office of Educational Technology.

[2] Aziz, A. R. A., Ab Razak, N. H., Sawai, R. P., Kasmani, M. F., Amat, M. I., & Shafie, A. A. H. (2021). Exploration of challenges among gifted and talented children. Malaysian Journal of Social Sciences and Humanities (MJSSH), 6(4), 242-251.

[3] Kasneci, E., Seßler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., ... & Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. Learning and individual differences, 103, 102274.

[4] Almarie, B., Teixeira, P. E., Pacheco-Barrios, K., Rossetti, C. A., & Fregni, F. (2023). Editorial–The use of large language models in science: Opportunities and challenges. Principles and practice of clinical research (2015), 9(1), 1.

[5] Naveed, H., Khan, A. U., Qiu, S., Saqib, M., Anwar, S., Usman, M., ... & Mian, A. (2023). A comprehensive overview of large language models. arXiv preprint arXiv:2307.06435.

[6] Pfeiffer, S. I. (2002). Identifying gifted and talented students: Recurring issues and promising solutions. Journal of Applied School Psychology, 19(1), 31-50.

[7] Robinson, N. M. (2008). The social world of gifted children and youth. In Handbook of giftedness in children: Psychoeducational theory, research, and best practices (pp. 33-51). Boston, MA: Springer US.

[8] Reis, S. M., & Renzulli, J. S. (2011). Challenging gifted and talented learners with a continuum of research-based interventions strategies.

[9] Hodges, J., & Mohan, S. (2019). Machine learning in gifted education: A demonstration using neural networks. Gifted Child Quarterly, 63(4), 243-252.

[10] Kalobo, L., & Setlalentoa, W. (2024). Navigating Challenges in Gifted Education: A Teacher's Perspective on Overcoming Barriers. Athens Journal of Education, 11(4), 269-287.

[11] Aydın, M., & Yurdugül, H. (2024). Developing a Curriculum Framework of Artificial Intelligence Teaching for Gifted Students. Kastamonu Education Journal, 32(1), 14-37.

[12] Vuyk, A., Montania, M., & Barrios, L. (2024). Boredom and its perceived impact in adolescents with exceptional mathematical talent: a sequential mixed-methods study in Paraguay. Frontiers in Sociology, 9, 1214878.

[13] Hornstra, L., Mathijssen, A. S., Denissen, J. J., & Bakx, A. (2023). Academic motivation of intellectually gifted students and their classmates in regular primary school classes: A multidimensional, longitudinal, person-and variable-centered approach. Learning and Individual Differences, 107, 102345.

[14] Aubeuf, C. (2023). Uses of Artificial Intelligence in Intelligent Tutoring System. Education Applications & Developments VIII Advances in Education and Educational Trends Series Edited by: Mafalda Carmo, 304.

[15] Ratican, J., & Hutson, J. (2024). Advancing sentiment analysis through emotionally-agnostic text mining in large language models (LLMS). Journal of Biosensors and Bioelectronics Research.

[16] Lin, H., & Chen, Q. (2024). Artificial intelligence (AI)-integrated educational applications and college students' creativity and academic emotions: students and teachers' perceptions and attitudes. BMC psychology, 12(1), 487.

[17] Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. Qualitative research in psychology, 3(2), 77-101.

[18] Youvan, Douglas. (2024). Artificial Intelligence as a Catalyst for Nurturing Exceptionally Gifted Children: A New Paradigm in Education and Mentorship. 10.13140/RG.2.2.28604.04480.

NEXUS and ISO 42001: Building Robust Governance for Responsible Enterprise AI

Mohammed Bahja
University of Birmingham
Edgbaston campus: University of Birmingham
United Kingdom
m.bahja@bham.ac.uk

Noureddin Sadawi
University of Oxford
Wellington Square,
Oxford OX1 2JD,
UK
noureddin.sadawi@conted.ox.ac.uk

Amir Shurrab
London Gulf Nexus
903 Iris Bay,
Business Bay,
Dubai PO Box 77646
amir@londongulfnexus.com

Zahra Alhabsi
University of Technology and Applied Sciences
Al Khuwair, Muscat
Sultanate of Oman
zahra.alhabsi@utas.edu.om

Abstract— The rise of generative AI presents both vast opportunities and critical challenges for organizations. This paper emphasizes the need for robust AI governance to address ethical concerns, data security, and evolving regulatory demands. Central to this discussion is ISO 42001—a comprehensive standard offering structured guidance for managing the AI lifecycle from design to continuous improvement. Building on this foundation, the paper introduces the NEXUS framework (Navigate, Establish, eXecute, Upskill, Sustain), which facilitates the integration of AI into enterprise environments. By aligning NEXUS with ISO 42001, the proposed conceptual governance model aims to streamline compliance, improve transparency, and promote responsible AI deployment within organizations.

Keywords—Nexus, ISO 42001, Conceptual Framework, Artificial Intelligence, AI Governance

I. INTRODUCTION

A. The Growing Need for AI Governance

Recent advancements in generative artificial intelligence (AI) have significantly boosted automation, data-driven decisions, and operational efficiency across sectors. In response, organizations are racing to adopt generative AI to stay competitive. However, this rapid integration demands robust governance frameworks to ensure ethical, secure, and regulation-compliant deployment [1,2]. As AI grows more complex, governance must address ethical risks, data security, regulatory adherence, and long-term sustainability.

- Addressing Ethical and Societal Challenges: A key governance concern is algorithmic bias. AI systems trained on historical data can perpetuate discrimination in recruitment, lending, or law enforcement. Ensuring fairness, transparency, and accountability is essential for mitigating prejudice and promoting ethical AI use [3,4].
- Ensuring Data Privacy and Security: AI systems handling sensitive data are vulnerable to cyberattacks. Without proper safeguards, breaches can result in major legal and financial consequences. Governance frameworks aligned with standards like ISO 42001, along with encryption and strict access controls, help mitigate these threats and maintain compliance with laws like GDPR and the EU AI Act [5].
- Navigating the Regulatory Landscape: Governments are introducing AI regulations to enforce transparency and accountability. ISO 42001 offers a standardized approach to AI governance, helping organizations comply with emerging laws while fostering public trust [6].
- Improving Transparency and Risk Management: Opaque AI models, or "black boxes," limit stakeholders' understanding of decisions. Incorporating Explainable AI (XAI) and conducting regular risk assessments enhances transparency and identifies ethical, security, and operational vulnerabilities early [7,8]. Robust AI governance is a strategic necessity. Adopting frameworks like ISO 42001 enables organizations to embed ethics, compliance, and sustainability into AI systems—safeguarding operations and building long-term trust in a fast-changing landscape.

This paper presents a conceptual governance framework that integrates the emerging ISO 42001 standard with the NEXUS methodology. The framework is designed to guide enterprises in developing transparent, accountable, and ethically sound AI systems. It offers a structured, theoretical model that organizations can reference when planning AI governance strategies.

B. Role of ISO 42001 in AI Management

ISO 42001 is becoming a foundational standard for managing AI responsibly, offering a structured framework that spans the entire AI lifecycle—from design to continuous improvement. It promotes ethical, transparent, and accountable AI practices while ensuring robust risk management and integration within existing organizational processes [9]. Key components include leadership, planning, operation, performance evaluation, and continuous improvement [10]. The standard addresses critical issues like bias, privacy, and security, ensuring AI is deployed in a secure and socially responsible manner [11]. A central focus of ISO 42001 is comprehensive risk assessment. It mandates systematic monitoring of AI systems, especially dynamic machine learning models, to mitigate unpredictable outcomes and ensure alignment with legal, ethical, and operational benchmarks [12]. Transparency is also crucial; the standard encourages Explainable AI (XAI) methods to make decision-making processes understandable and accountable, reinforcing stakeholder trust [13,14]. By aligning with ISO 42001, organizations can build sustainable, compliant AI ecosystems that balance innovation with responsibility. This standard enables businesses to navigate AI's challenges while driving positive societal impact and long-term growth [3,4].

C. NEXUS: A Framework for Enterprise AI Hubs

The NEXUS framework, developed by the authors, enables enterprises to integrate AI seamlessly into their operations while aligning innovation with strategic goals. It embeds governance, risk management, and ethical best practices to ensure responsible, sustainable AI deployment. NEXUS is structured around five interconnected phases: Navigate, Establish, eXecute, Upskill, and Sustain. Navigate focuses on strategic assessment and roadmap development, guiding organizations through readiness evaluations and the creation of AI-aligned business strategies, including risk mitigation plans [15]. Establish involves building secure, scalable infrastructure—such as high-performance computing and cloud environments—and implementing robust data governance. Aligning with ISO 42001 at this stage ensures regulatory compliance and security [16]. eXecutecenters on AI model deployment, integration into workflows, and iterative refinement. This phase addresses key issues such as bias, performance, and accountability through continuous monitoring [17]. Upskill emphasizes workforce development through AI literacy training and mentorship, equipping staff to manage AI systems ethically and efficiently [18]. Sustain ensures ongoing improvement and governance via regular audits, risk assessments, and updates in response to technological and regulatory shifts. Governance committees maintain oversight and ethical performance [19].

II. METHODOLOGY

This study adopts a theoretical approach, aiming to develop a conceptual governance framework that integrates ISO 42001 standards with the NEXUS methodology. The framework was constructed through an extensive review of existing literature on AI governance, risk management, ethical deployment, and international standards, particularly ISO 42001. Key themes and best practices were synthesized to identify critical elements of responsible AI lifecycle management. These insights were then mapped onto the five phases of the NEXUS model (Navigate, Establish, eXecute, Upskill, Sustain) to formulate an integrated framework suitable for enterprise contexts. This methodology does not involve empirical data collection or validation but instead offers a structured conceptual foundation intended to guide future implementation and research.

III. NEXUS + ISO 42001: AN INTEGRATED FRAMEWORK

The convergence of the NEXUS framework with ISO 42001 standards creates a powerful integrated approach for AI governance in enterprise settings. This integrated framework leverages the strengths of NEXUS's five-phase methodology—Navigate, Establish, eXecute, Upskill, and Sustain—while embedding ISO 42001's rigorous guidelines for ethical, secure, and compliant AI deployment. The following sections outline the integration of each NEXUS phase with corresponding ISO 42001 components, thereby establishing a cohesive system for managing AI across its lifecycle.

Fig. 5. NEXUS + ISO 42001: An Integrated Conceptual Framework

D. Strategic Assessment (Navigate + Context & Leadership from ISO 42001)

The strategic assessment phase of the NEXUS framework, aligned with ISO 42001 principles, ensures AI governance supports an organization's broader business strategy. This phase involves evaluating AI readiness, identifying key use cases, and developing a roadmap for implementation. A core element is the AI Readiness Assessment, which reviews infrastructure, workforce skills, and data policies to identify gaps and prioritize high-value, low-risk AI initiatives. ISO 42001 emphasizes leadership's role in AI governance [20]. Senior executives must set clear goals, align with ethical standards, and promote responsible AI through governance committees, AI ethics officers, and policy integration [21]. Strong leadership drives accountable, regulation-compliant AI adoption. Risk management planning is also essential. Organizations must assess risks such as bias, security threats, and regulatory issues, and implement mitigation strategies [22]. By embedding ISO 42001 into NEXUS's Navigate phase, organizations build a transparent, ethical, and accountable foundation for AI deployment aligned with global standards.

E. Infrastructure Setup (Establish + Compliance & Security from ISO 42001)

The infrastructure setup phase of the NEXUS methodology focuses on creating a secure, scalable, and compliant environment for AI deployment. By integrating ISO 42001 principles, organizations can establish a strong foundation aligned with best practices in AI security, privacy, and risk management. This phase includes deploying high-performance computing clusters, cloud-based environments, and robust data governance policies.A critical aspect is building an AI technology stack that meets compliance and efficiency requirements. This involves selecting appropriate AI frameworks, data pipelines, and machine learning platforms that align with legal and operational standards. ISO 42001 also emphasizes secure API integration to enhance interoperability and system integrity. Data governance and privacy protection are central to this phase. Measures such as encryption, access control, and privacy-enhancing technologies are essential for compliance with standards like GDPR and ISO 42001 [23]. Secure data storage and transfer mechanisms reduce breach risks and support ethical AI use.Regular monitoring, auditing, and vulnerability testing ensure ongoing system reliability [24]. Incorporating ISO 42001 at this stage enables enterprises to build an innovative AI ecosystem that is secure, ethical, and compliant with international standards.

F. Solution Deployment (Execute + ISO 42001 Risk & Impact Assessment)

The solution deployment phase of the NEXUS framework focuses on the practical implementation of AI solutions, ensuring they align with ISO 42001 standards for security, compliance, and performance. This phase translates AI strategies into operational systems that drive efficiency and innovation while mitigating associated risks. Deployment and Monitoring: AI models are integrated into enterprise systems using machine learning and automation tools [25]. ISO 42001 emphasizes continuous monitoring and validation to detect anomalies, biases, or unintended impacts. Organizations must implement real-time model validation protocols and conduct thorough risk assessments, including fairness checks, vulnerability identification, and regulatory reviews [26]. Risk Management and Explainability: ISO 42001 provides a framework for managing risks and ensuring ethical, legal compliance. Enterprises are encouraged to adopt Explainable AI (XAI) to promote transparency and build stakeholder trust [27]. XAI ensures decision-making processes are interpretable and accountable. Before full deployment, AI models undergo rigorous testing—stress tests, bias audits, and scenario simulations. Feedback loops are essential for refining performance and maintaining relevance to evolving business and regulatory needs. By integrating ISO 42001 during execution, organizations ensure AI solutions are reliable, ethical, and compliant, supporting sustainable and transparent innovation.

G. Upskilling & Workforce Training (Upskill + ISO AI Awareness)

The upskilling and workforce training phase of the NEXUS methodology equips organizations with the talent needed to manage AI systems effectively. By incorporating ISO 42001 guidelines, enterprises ensure that training programs align with global standards, enhancing technical proficiency, ethical awareness, and regulatory understanding. AI Competency Frameworks: Organizations should establish structured frameworks that define essential skills for various roles, helping identify gaps and tailor role-specific training programs [28]. These frameworks prepare employees to address AI-related challenges efficiently. AI Literacy and Ethics: ISO 42001 emphasizes AI literacy across all levels of an organization. Awareness initiatives should educate employees about algorithmic bias, ethical decision-making, and regulatory compliance. Hands-on training, including real-world simulations and interactive case studies, is essential for developing problem-solving and evaluation skills [29]. Ongoing Learning and Mentorship: Organizations should promote continuous learning through mentorship programs, AI certifications, and academic partnerships [30]. By embedding ISO 42001 in workforce training, enterprises foster responsible AI adoption, strengthen governance, and support long-term, sustainable innovation.

H. Sustained Governance (Sustain + ISO 42001 Continuous Improvement)

The sustained governance phase of the NEXUS methodology ensures that AI systems maintain ethical integrity, transparency, and regulatory compliance throughout their lifecycle. By incorporating ISO 42001's continuous improvement principles, organizations can develop adaptive governance structures that evolve alongside technological and regulatory changes. AI Governance Committees: Establishing dedicated governance committees is crucial. These cross-functional teams—comprising AI experts, legal advisors, and policymakers—oversee compliance, performance, and ethical concerns [31]. Audit and Compliance Frameworks: Regular audits are essential to monitor AI model performance, privacy compliance, and risk mitigation. Following ISO 42001, organizations should implement structured protocols to identify bias, security issues, and ethical risks [32]. Bias and Error Monitoring: Continuous monitoring is necessary to detect and address bias, ensuring fair and accurate AI outcomes. Automated tools help identify inconsistencies and recommend corrective actions [33]. Stakeholder Feedback Loops: Engaging stakeholders—such as users, employees, and regulators—and integrating their feedback ensures AI systems align with business goals and societal expectations. Embedding ISO 42001 in this phase equips organizations with resilient governance practices that foster accountable, ethical, and transparent AI operations.

IV. CONCLUSION

This paper introduced a conceptual framework for AI governance by integrating the NEXUS methodology with ISO 42001 standards. The framework proposes a structured, phased approach that includes strategic assessment, secure infrastructure setup, solution deployment, workforce upskilling, and sustained governance. It is intended as a theoretical guide for organizations seeking to align AI practices with ethical, legal, and operational standards. The integration of the NEXUS framework with ISO 42001 standards represents a significant step forward in the domain of AI governance. Our research presents that a structured, phased and a conceptual approach encompassing strategic assessment, secure infrastructure setup, pragmatic solution deployment, continuous workforce training, and sustained governance that can address the multifaceted challenges posed by rapid AI adoption. The synergy between NEXUS and ISO 42001 not only reinforces ethical practices and risk mitigation but also provides a clear roadmap for organizations to achieve operational excellence in AI integration. Future research will explore additional dimensions of AI governance, including scalability across different industries and long-term impacts on organizational culture and public trust. Overall, our integrated framework lays a robust foundation for responsible and sustainable AI deployment in enterprise settings.

REFERENCES

[1] G. P. Selvarajan, “Leveraging AI-enhanced analytics for industry-specific optimization: A strategic approach to transforming data-driven decision-making,” Int. J. Enhanc. Res. Manag. Comput. Appl., vol. 10, no. 10, pp. 78–84, 2021.

[2] M. S. H. Mrida, M. A. Rahman, and M. S. Alam, “AI-driven data analytics and automation: A systematic literature review of industry applications,” Strateg. Data Manag. Innov., vol. 2, no. 1, pp. 21–40, 2025.

[3] L. Floridi et al., “AI4People—An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations,” Minds Mach., vol. 28, pp. 689–707, 2018.

[4] A. Jobin, M. Ienca, and E. Vayena, “The global landscape of AI ethics guidelines,” Nature Mach. Intell., vol. 1, no. 9, pp. 389–399, 2019.

[5] A. H. Salem, S. M. Azzam, O. E. Emam, and A. A. Abohany, “Advancing cybersecurity: A comprehensive review of AI-driven detection techniques,” J. Big Data, vol. 11, no. 1, 2024.

[6] D. Lewis, D. Filip, and H. J. Pandit, “An ontology for standardising trustworthy AI,” IntechOpen eBooks, 2021.

[7] G. Banerjee, S. Dhar, S. Roy, R. Syed, and A. Das, “Explainability and transparency in designing responsible AI applications in the enterprise,” in Lecture Notes in Networks and Systems, 2024, pp. 420–431.

[8] I. Rahwan et al., “Machine behaviour,” Nature, vol. 568, no. 7753, pp. 477–486, 2019.

[9] PECB, “A comprehensive guide to understanding the role of ISO/IEC 42001,” 2024. [Online]. Available: https://pecb.com/article/a-comprehensive-guide-to-understanding-the-role-of-isoiec-42001

[10] KPMG, “ISO/IEC 42001. The latest AI management system standard,” 2025. [Online]. Available: https://kpmg.com/ch/en/insights/artificial-intelligence/iso-iec-42001.html

[11] S. A. Benraouane, AI Management System Certification According to the ISO/IEC 42001 Standard: How to Audit, Certify, and Build Responsible AI Systems. CRC Press, 2024.

[12] T. R. McIntosh et al., “From COBIT to ISO 42001: Evaluating cybersecurity frameworks for opportunities, risks, and regulatory compliance in commercializing large language models,” Comput. Secur., vol. 144, p. 103964, 2024.

[13] S. Oveisi, F. Gholamrezaie, N. Qajari, M. S. Moein, and M. Goodarzi, “Review of artificial intelligence-based systems: Evaluation, standards, and methods,” Adv. Stand. Appl. Sci., vol. 2, no. 2, pp. 4–29, 2024.

[14] F. Doshi-Velez and B. Kim, “Towards a rigorous science of interpretable machine learning,” arXiv preprint arXiv:1702.08608, 2017.

[15] F. A. Csaszar, H. Ketkar, and H. Kim, “Artificial intelligence and strategic decision-making: Evidence from entrepreneurs and investors,” Strategy Sci., 2024.

[16] R. Sharma, “Building robust AI infrastructure for enterprise success,” in Apress eBooks, 2024, pp. 247–258.

[17] N. K. O. Al-Amin, N. C. P. Ewim, N. A. N. Igwe, and N. O. C. Ofodile, “AI-driven end-to-end workflow optimization and automation system for SMEs,” Int. J. Manag. Entrep. Res., vol. 6, no. 11, pp. 3666–3684, 2024.

[18] V. Uren and J. S. Edwards, “Technology readiness and the organizational journey towards AI adoption: An empirical study,” Int. J. Inf. Manag., vol. 68, p. 102588, 2022.

[19] J. Zhao and B. G. Fariñas, “Artificial intelligence and sustainable decisions,” Eur. Bus. Organ. Law Rev., vol. 24, no. 1, pp. 1–39, 2022.

[20] C. Dudley, “The rise of AI governance: Unpacking ISO/IEC 42001,” Qual. Troy, vol. 63, no. 8, p. 27, 2024.

[21] B. Shneiderman, “Bridging the gap between ethics and practice,” ACM Trans. Interact. Intell. Syst., vol. 10, no. 4, pp. 1–31, 2020.

[22] C. Curtis, N. Gillespie, and S. Lockey, “AI-deploying organizations are key to addressing the ‘perfect storm’ of AI risks,” AI Ethics, vol. 3, no. 1, pp. 145–153, 2022.

[23] R. Alonso, R. E. Haber, F. Castaño, and D. R. Recupero, “Interoperable software platforms for introducing artificial intelligence components in manufacturing: A meta-framework for security and privacy,” Heliyon, vol. 10, no. 4, p. e26446, 2024.

[24] I. Munoko, H. L. Brown-Liburd, and M. Vasarhelyi, “The ethical implications of using artificial intelligence in auditing,” J. Bus. Ethics, vol. 167, no. 2, pp. 209–234, 2020.

[25] E. Hechler, M. Oberhofer, and T. Schaeck, Deploying AI in the Enterprise: IT Approaches for Design, DevOps, Governance, Change Management, Blockchain, and Quantum Computing. 2020. [Online]. Available: https://www.amazon.com/Deploying-Enterprise-AI-Governance-Management/dp/1484262050

[26] I. M. Leghemo, C. Azubuike, O. D. Segun-Falade, and C. S. Odionu, “Data governance for emerging technologies: A conceptual framework for managing blockchain, IoT, and AI,” J. Eng. Res. Rep., vol. 27, no. 1, pp. 247–267, 2025.

[27] U. Blinova, N. Rozhkova, and D. Rozhkova, “NFT (Non-Fungible Tokens) as an object of accounting,” J. Digit. Art Humanit., vol. 4, no. 1, pp. 3–9, 2023.

[28] N. Bobitan, D. Dumitrescu, A. F. Popa, D. N. Sahlian, and I. C. Turlea, “Shaping tomorrow: Anticipating skills requirements based on the integration of artificial intelligence in business organizations—A foresight analysis using the scenario method,” Electronics, vol. 13, no. 11, p. 2198, 2024.

[29] B. Ammanath and R. Blackman, “Everyone in your organization needs to understand AI ethics,” Harvard Business Review, Jul. 26, 2021. [Online]. Available: https://hbr.org/2021/07/everyone-in-your-organization-needs-to-understand-ai-ethics

[30] B. Ammanath and R. Blackman, “Everyone in your organization needs to understand AI ethics,” Harvard Business Review, Jul. 26, 2021. [Online]. Available: https://hbr.org/2021/07/everyone-in-your-organization-needs-to-understand-ai-ethics

[31] M. B. A. Roopalatha and K. Sucharita, “Navigating the AI frontier: A study of AI integration in IT employee training and development,” Educ. Adm. Theory Pract., vol. 30, no. 5, pp. 1079–1085, 2024.

[32] M. L. Montagnani and M. L. Passador, “Artificial intelligence for post-Covid companies: An empirical analysis of tech committees in the EU and US,” SSRN Electron. J., 2020.

[33] A. N. Prasad, “Regulatory compliance and risk management,” in Apress eBooks, 2024, pp. 485–624.

[34] N. Gupta, “Artificial intelligence ethics and fairness: A study to address bias and fairness issues in AI systems, and the ethical implications of AI applications,” Rev. Index J. Multidiscip., vol. 3, no. 2, pp. 24–35, 2023.

The Role of AI-Enhanced Strategic Leadership Practices in Shaping Employee Performance

Ahmad Elbadaoui
Graduate student in Abu Dhabi School of Management
A.badaoui@adsm.ac.ae
Abu Dhabi, UAE

Turki Al Masaeid
Assistant professor in Abu Dhabi School of Management
t.almasaeid@adsm.ac.ae
Abu Dhabi, UAE

Abstract— The AI-Enhanced Strategic has appeared in multiple articles in recent literature. The theme presents an evolution from traditional leader-focused individual perspectives to dynamic group-centered leadership. Each team member under strategic leadership possesses a distinctive set of responsibilities. The study evaluates how Strategic leadership influences employee performance. The current research study uses questionnaires to obtain information from all employees currently working in the organization. According to the research outcomes, Strategic leadership expresses one of its most valuable traits through its organizational cultural dominance. This paper also examines the research constraints and potential future study boundaries.

Keywords— AI-Enhanced Strategic leadership, Employee performance, and Employee efficiency.

I. INTRODUCTION

As businesses navigate the complexities of the modern workplace, AI technologies are emerging as pivotal tools that enhance decision-making, operational efficiency, and employee engagement. This evolution reflects a broader trend in which leaders are increasingly required to adapt their management styles to leverage AI capabilities while addressing employee concerns and fostering a collaborative culture[1][2][3].

Notably, the adoption of AI in strategic leadership is reshaping traditional frameworks by facilitating data-driven insights and personalized employee experiences. Research indicates that organizations employing AI-driven performance management systems can experience productivity increases of up to 40%, demonstrating a clear link between AI integration and enhanced employee outcomes[4][5]. However, this shift is not without its challenges; ethical considerations surrounding privacy, bias, and the potential for dehumanization of the workplace have sparked significant debate among scholars and practitioners alike[6][7][8].

Moreover, AI's influence extends to talent management, where machine learning algorithms are utilized to identify skill gaps, personalize training, and promote employee retention. Companies such as IBM and Microsoft have illustrated the trans- formative potential of AI in enhancing employee performance, resulting in improved job satisfaction and engagement[9][10]. As organizations continue to embrace AI technologies, leaders must navigate the dual demands of leveraging data-driven insights while maintaining a human-centric approach to leadership, ensuring that employees feel valued and empowered in their roles[2][11].

The integration of AI-enhanced strategic leadership practices holds significant implications for shaping employee performance. As leaders adapt to these innovations, they face the critical task of balancing the benefits of AI with the ethical and relational aspects of effective management. The ongoing discourse around AI's role in the workplace highlights the importance of transparency, emotional intelligence, and inclusive practices in fostering a positive organizational culture that drives success[6][12][13]. As a result, this study explores how AI-assisted as well as traditional strategic leadership practices impact crucial organizational performance metrics, most of all with regard to the issue of the team efficiency, task interdependence, and the performance of the employees.

A. Employee motivation and effectiveness

Multiple studies and organizational management papers and Psychology research demonstrate evidence for this motivational approach to work design. Employee motivation stands as a critical managerial element and testing management aspect according to [22] because it determines how individuals motivate themselves concerning challenging work and their personal outlook. Explain how internal identity development functions through behavior triggered by cognitive processes. According to [1] the essence of motivation stems from the power that drives individuals toward taking action to resolve issues or satisfy their necessities. The definition of motivation traces its foundation to an individual's natural prompting forces which they employ to reach their life objectives. Leadership experience from corporate sectors when applied to structured activities strengthens workplace relationships between employees and produces successful work results. Leaders who develop loyalty and respect through strategic leadership create positive relationships that are free from hypocrisy because mutual understanding brings trust along with capability. [21] According to [12] organizations can reduce resistance to change through rational progress explanations because such explanations reduce information misuse and eliminate ambiguity (1) and create fair understandings between stakeholders including investors and employees (2).

The organization depends heavily on its workers for their ability to fulfill business requirements regarding limits and capacities and potential and creativity for achieving organizational goals. Workers who feel motivated toward their duties experience satisfaction with their work. Fulfillment functions as a result that exists inseparably from motivational processes. Strategic leaders make investments which allow team members to trust each other for unconditional backup. Absence of confidence appears as an observed factor which leads employees to prefer detailed instructions rather than having responsibility for critical choices. Employee engagement with the change process requires their active commitment to make the change decision possible.

A. The application of Strategic Leadership

Research indicates that Strategic leadership functions as an essential indicator for team results according to Pearce and Sims (2002). Research examining Strategic leadership demonstrates positive effects on teams which includes improved team performance and team satisfaction as well as enhanced team functioning [20]. The true leadership potential of individuals led to the adoption of Strategic leadership models based on transparent communication and encouraging autonomy and welcoming input from others. Sharing authority between leaders helps them build their own support network. The abdication of burdens leads to comfort and simplicity for individuals. Leaders make themselves available to provide guidance at any time. Through strategic leadership team members invest time and resources into each other because they trust their partners will protect them when necessary.

Even when leaders provide appropriate authority to employees they can still fail to utilize these given controls despite the best intentions. Multiple situations lead to this issue because employees require basic leadership direction particularly to understand their responsibilities. Strategic leadership often fails when employees lack adequate experience alongside expertise. Some workers exhibit low confidence levels which leads them to accept only detailed instructions while refusing to handle significant choices independently. Employees who need help with development can benefit from motivational incentives and coaching into critical thinking and initial task completion. Finally, when employees fully understand which team members cannot be coached the best solution is to reassign them to less senior roles. The creation of Strategic leadership culture requires employers to measure their managers' capacity to achieve results through effective delegation and collaboration. And there is no scientific

Strategic leadership traits naturally emerge differently in various people [13] yet research lacks a standard method for ability measurement. Leaders possess specific behavioral assessment techniques despite their absence in other management positions. The discovery of genuine leader potential led to implementing Strategic leadership models founded upon transparency alongside organizational freedom and open-minded acceptance of team proposals. The hypotheses are as follows. These have been established after exhaustive literature research:

- H1: There is a positive relationship between Strategic leadership attitude and team efficiency
- H2: There is a positive relationship between Attitude towards Strategic leadership and team's effectiveness, Initiated Task Interdependence.
- H3: The AI-enhanced strategic leadership practices significantly influence employee performance.

Data collection

This study employed a quantitative research design, relying primarily on primary data collected via a structured questionnaire. While prior literature provided a theoretical foundation (secondary research), the analysis is based on empirical evidence gathered from the organization's workforce. The questionnaire design followed best practices, incorporating a pilot test with 10 employees to assess clarity and reliability. Feedback from the pilot was used to refine item wording and scale formats. The final survey contained 20 Likert-scale items divided into thematic areas: leadership behavior, team collaboration, and performance outcomes.

The survey was distributed electronically via Google Forms, and responses were automatically logged to ensure integrity. Participants were informed about the purpose of the study, and informed consent was obtained before data collection. The research adhered to ethical standards, ensuring anonymity and voluntary participation

The study was conducted in a mid-sized private organization operating in Abu Dhabi, UAE, focused on professional services. A total of 150 questionnaires were distributed among employees using a simple random sampling technique, and 133 valid responses were received (response rate: 88.6%). The inclusion criteria required participants to be full-time employees with at least six months of tenure. Respondent demographics included gender, education level, and department, although only gender and education data were analyzed in this paper.

II. DATA ANALYSIS AND RESULTS

The collected survey data underwent statistical analysis with basic tools after question distribution to the participants. The study reveals the connections that exist between all measured variables. The graph in Figure 1 represents responses from both male and female participants. Figure 2: educational qualification of participants. Figure 3: This data shows how team performance looks under one-person leadership. Figure 4: The survey item revealed feedback from scenarios that included joint leadership between different personnel. Figure 5: The analysis measured team performance outcomes when leadership responsibilities extend to all team members.

Participants with professional degrees made up more than half of the respondents, while males exceeded females in total participants.

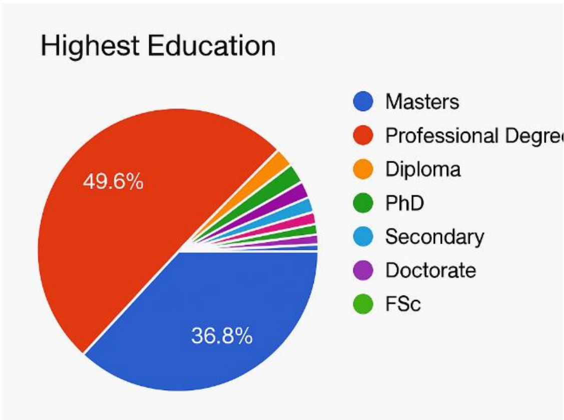
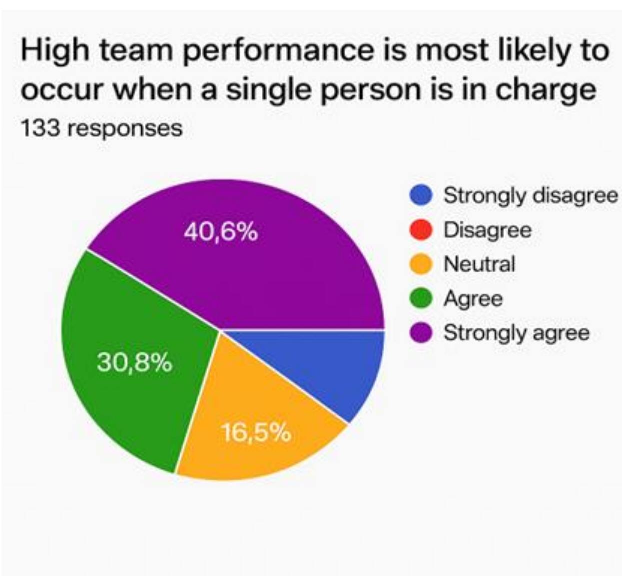


Figure 2. Distribution of survey respondents' highest education

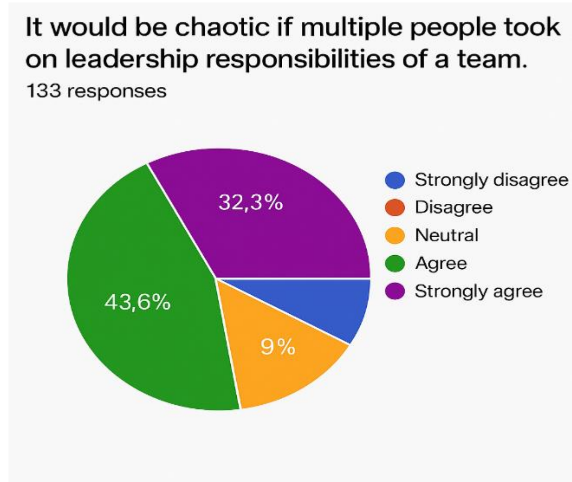
When participants responded to the questionnaire it was organized into various sections. Four targeted questions in this first section evaluated team performance metrics.

Figure 3: Team performance with solo Leader.



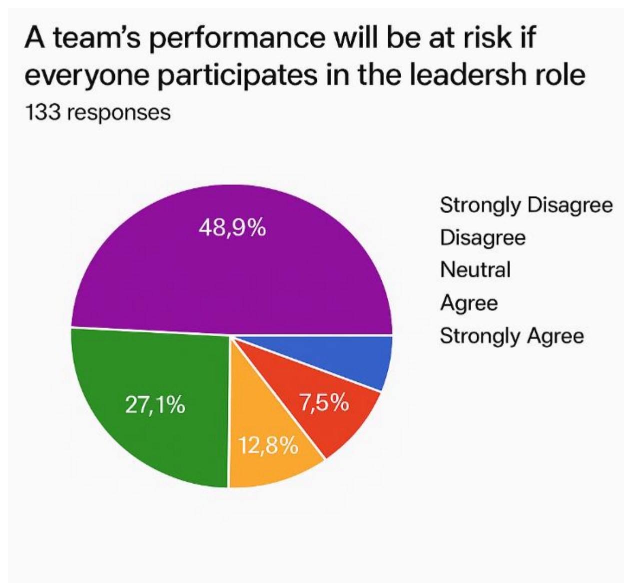
High team performance will likely happen when one person maintains complete command of the situation according to 40 percent of respondents. The data suggests workers acknowledged the value of building teams along with strong performance results.

Figure 4: Team Performance with multiple leadership



The survey results showed divergence on whether a single leader would suffice with approximately forty percent agreeing and thirty percent strongly agreeing while forty percent disagreed with additional leaders creating disorder. The responses contradict strategic leadership principles along with stories because they reveal different findings.

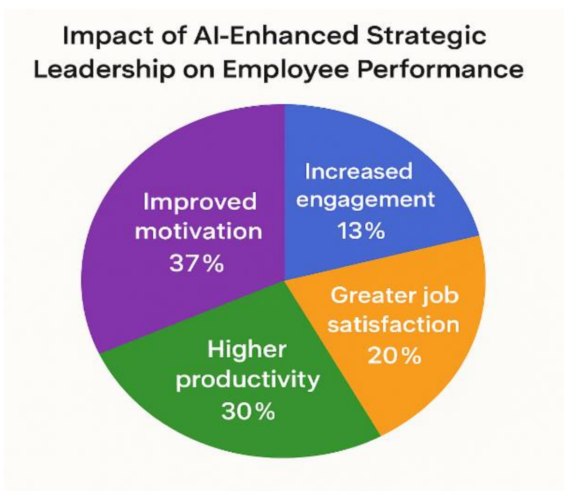
Figure 5: Team performance when everyone participates in the Strategic levels.



The following question explored the same territory as its preceding one. Nearly half of participants strongly support the view that team performance becomes compromised when everyone takes the leadership position indicating Strategic leadership has merit yet remains unacceptable to numerous staff members. The survey results show that 10% of respondents challenged the statement's validity. While members of the team lack authority to lead others directly they still have opportunities to lead their peers. Task interdependencies under Strategic leadership directly affect how teams perform.

Over 80% of respondents affirm the team leader accepts vital input from fellow members. This depicts one of the great qualities of strategic leadership: being innovative. According to previous research about Strategic leadership spaces employees receive positive welcome to introduce their new concepts and gather information. The resulting autonomous environment leads to elevated team morale and better performance outcomes.

Figure 6: Leader is Open to new ideas



The integration of AI-enhanced strategic leadership practices significantly influences employee performance across various dimensions. Organizations that effectively utilize AI tools experience improvements in employee motivation, engagement, and overall job satisfaction, which are crucial for driving high performance in the work- place. Research indicates that companies leveraging AI-driven solutions can see productivity increases of up to 40% in certain functions, enabling employees to accomplish more in less time, thereby fostering business growth and profitability

III. DISCUSSION AND CONCLUSION

Kamery [9] distinguishes two abrogating issues in estimating employee performance: Employee Performance and employee efficiency. For them Employee Performance refers to the extent which employees meet their predefined targets and goals. An employee proves effective when he or she reaches their established targets and objectives.

This questionnaire-based study has proven the analysis to be valid. Efficiency in workplace performance stems from an employee successfully reaching their goals through minimal resource expenditure. Wang [20] showed that effectiveness and efficiency measure how well people contribute to organizational goal fulfillment.

The current study contains specific restrictions in its approach. The following limitations to the current research study can be found: This organizational leadership model maintains a singular connection to one enterprise. The sample size is limited. The participants who answered the questionnaire included employees from the organization. The participants could provide answers that show their personal biases. Some disparities in observer understanding of survey questions and how respondents answered them likely altered the study's outcome. Personal beliefs within qualitative data present possible response influences.

Every participant took the survey after they read both the study objectives and approval conditions. This academic research study exists to serve educational goals and understands the business world. Research actions have shown practical value since leadership combined with motivation along with Employee Performance level through team interdependence creates organizational success within an actual business environment. Employers have the responsibility to measure their managers' capacities to achieve results through effective delegation and collaborative practices when establishing Strategic leadership culture.

Employers must evaluate their managers' capability to achieve results through proper delegation and collaboration while developing Strategic leadership culture. Employees exhibit different levels of natural inclinations toward Strategic leadership traits despite the lack of a scientific method for ability evaluation. Leadership assessment methods exist despite that fact. The organizations possess employee selection and retention reports provided by Omni. These tools can help employers to determine the ability of a team that they can lead collaboratively and possess.

The integration of Artificial Intelligence (AI) into strategic leadership practices presents a series of complex challenges and ethical considerations that organizations must navigate to ensure successful outcomes. One of the primary concerns is maintaining transparency in AI decision-making processes. Leaders are tasked with ensuring that the algorithms driving AI systems are understandable to stakeholders, fostering an environment of trust and accountability within the organization[8][11].

Moreover, there is a critical need to balance AI capabilities with human intuition and emotional intelligence. While AI can significantly enhance decision-making by providing data-driven insights, it cannot replace the empathy, creativity, and moral judgment that are fundamental to effective leadership[11][12]. As organizations continue to leverage AI tools, leaders must emphasize human-centric leadership principles to maintain an authentic connection with their teams.

In addition, the ethical use of AI in monitoring employee performance raises significant concerns regarding privacy and consent. Research indicates that a notable percentage of employees express discomfort with monitoring practices, yet they recognize the necessity of some oversight for productivity[13][26]. Organizations must implement robust data protection measures and prioritize informed consent to mitigate feelings of alienation and resistance among employees[27][28]. By engaging employees in the AI implementation process and ensuring transparency, organizations can foster a cooperative culture that enhances both trust and performance.

Finally, organizations must be mindful of the potential for AI to inadvertently reinforce existing biases within decision-making frameworks. As AI systems learn from historical data, there is a risk that they may perpetuate biases present in the input data, leading to inequitable outcomes[8]. Leaders should actively seek to understand and address these biases to create a more inclusive and fair workplace.

REFERENCES

[1] Asfaw, A.M., Argaw, M.D. and Bayissa, L., (2015). The impact of training and development on employee performance and effectiveness: A case study of District Five Administration Office, Bole Sub-City, Addis Ababa, Ethiopia. Journal of Human Resource and Sustainability Studies, 3(04), p.188.

[2] Barg, S., Perez, F.M., Ni, N., do Vale Pereira, P., Maher, R.C., Garcia-Tunon, E., Eslava, S., Agnoli, S., Mattevi, C. and Saiz, E., (2014). Mesoscale assembly of chemically modified graphene into complex cellular networks. Nature communications, 5, p.4328.

[3] Carson, J. B., Tesluk, P. E., & Marrone, J. A. (2007). Strategic leadership in teams: An investigation of antecedent conditions and performance. Academy of Management Journal, 50, 1217- 1234.

[4] Ensley, M. D., Hmieleski, K. M., & Pearce, C. L. (2006). The importance of vertical and Strategic leadership within new venture top management teams: Implications for the performance of startups. The leadership quarterly, 17(3), 217-231.

[5] Gilley, B., (2009). The right to rule how states win and lose legitimacy. Columbia University Press.

[6] Heenan, D. A., & Bennis, W. B. (1999) Co-leaders: The Power of Great Partnerships. New York: Wiley

[7] Hoch, J.E., Bommer, W.H., Dulebohn, J.H. and Wu, D., (2018). Do ethical, authentic, and servant leadership explain variance above and beyond transformational leadership? A meta-analysis. Journal of Management, 44.(2), 501-529,

[8] Hussain, S.T., Lei, S., Akram, T., Haider, M.J., Hussain, S.H. and Ali, M., (2018). Kurt Lewin's change model: A critical review of the role of leadership and employee involvement in organizational change. Journal of Innovation & Knowledge, 3(3),123-127.

[9] Kamery, R.H., (2004), October. Employee motivation as it relates to effectiveness, efficiency, productivity, and performance. In Proceedings of the Academy of Legal, Ethical and Regulatory Issues 8 (2), 139-144).

[10] Mintzberg, H., (1983). Power in and around organizations. Prentice Hall.

[11] Martin, B., Breunig, M., Wagstaff, M. and Goldenberg, M., (2017). Outdoor leadership. Human Kinetics.

[12] Northouse, P.G., (2018). Leadership: Theory and practice. Sage publications.

[13] Osabiya, B.J., (2015). The effect of employees' motivation on organizational performance. Journal of public administration and policy research, 7(4),62-75.

[14] Pearce, C. L. (2004). The future of leadership: Combining vertical and Strategic leadership to transform knowledge work. Academy of Management Executive, 18, 47-57.

[15] Pearce, C. L. (2007). The future of leadership development: The importance of identity, multi-level approaches, self-leadership, physical fitness, shared leadership, networking, creativity, emotions, spirituality and on-boarding processes. Human Resource Management Review, 17(4), 355-359.

[16] Pearce, C. L., & Sims, H. P. (2002). Vertical versus Strategic leadership as predictors of the effectiveness of change management teams: An examination of aversive, directive, transactional, transformational, and empowering leader behaviors. Group Dynamics: Theory, Research, and Practice, 6, 172-197

[17] Robbins, S.P., Judge, T.A. and Millett, B., (2015). OB: the essentials. Pearson Higher Education AU.

[18] Scura, G. and Davidoff, F., (1981). Case-related use of the medical literature: clinical librarian services for improving patient care. Jama, 245(1), pp.50-52.

[19] Sally, D., (). Co-leadership: Lessons from republican Rome. California Management Review, 44(4), pp.84-99.

[20] Wang, M., Guo, T., Ni, Y., Shang, S. and Tang, Z., (2019). The effect of spiritual leadership on Employee Performance: Anintrinsic motivation perspective. Frontiers in psychology, 9,2627.

[21] Yukl, G.A. and Becker, W.S., (2006). Effective empowerment in organizations. Organization Management Journal, 3(3), 210-231.

Exploring TAMM’s Digital Transformation in Abu Dhabi

Hessa Banihammad
College of Business and Economics
United Arab Emirates University (UAEU)
Al Ain-UAE
700048710@uaeu.ac.ae

Amal AL Hammadi
College of Business and Economics
United Arab Emirates University (UAEU)
Al Ain-UAE
700046897@uaeu.ac.ae

Ayesha Aldarmaki
College of Business and Economics
United Arab Emirates University (UAEU)
Al Ain-UAE
700048624@uaeu.ac.ae

Fatma Abdelhamid
College of Business and Economics
United Arab Emirates University (UAEU)
Al Ain-UAE
201850088@uaeu.ac.ae

Ananth Chiravuri
College of Business and Economics
United Arab Emirates University
(UAEU)
Al Ain-UAE
ananth.chiravuri@uaeu.ac.ae

Abstract—Government services worldwide are adopting digital transformation solutions which are reshaping digital services through improved efficiency, accessibility and user satisfaction. This study investigates the impact of TAMM, the consolidated digital platform of Abu Dhabi, on such parameters as mentioned above. A survey was conducted to collect responses about users’ perceptions of TAMM’s efficiency. Initial findings suggest that TAMM is successful in reducing bureaucratic inefficiencies, speeding up governmental transactions, and enhancing citizen confidence in digital services. Findings from this research contribute to the literature on digital transformation by offering empirical insights regarding digital services platforms and service optimization.

Keywords—Component, Formatting, Style, Styling, Insert (Key Words**)**

I. INTRODUCTION

As countries aim to improve efficiency, accessibility, and service quality through technological advancements, the digital transformation of government services has become a global priority. In response to the swift digitalization of economies, governments around the globe are revamping their service delivery models, utilizing automation, data analytics, and artificial intelligence to make administrative processes more efficient. Digital Transformation in the public sector aims to reduce bureaucracy, reduce physical visits and enhance customer engagement. Abu Dhabi and the UAE have taken a leading role in this transition by launching e-government strategies. TAMM is a centralized digital platform which integrates over 900 governmental services in a unified ecosystem which allows citizens to access government services more efficiently.

Despite the widespread of digital government platforms, there is a current limitation in research investigating user perception of these digital platforms. Most of the existing research focuses on technology adoption, policy frameworks and challenges with implementing digital transformation, rather than assessing the effects of digital platforms such as TAMM on service automation, user satisfaction, and operational efficiency. Despite the fact that previous research emphasized on the effective methods for digital transformation, there is a lack of empirical evidence regarding users' perceptions and utilization of these digital platforms [1]. Thus, findings from our study aim to fill the gap by providing evidence-based insights into how effective TAMM is as a digital transformation initiative in Abu Dhabi through analyzing user perceptions and assessing statistical trends. This forms the objective of our study and our research question which is to understand user perspective, engagement, trust and accessibility when using digital platforms like TAMM. Results from our study should help policy makers understand and assess the acceptance of digital government in Abu Dhabi.

The rest of the paper is as follows: In Section 2, we review the relevant papers as part of the Literature review. Following this, we will present our hypotheses in Section 3. Section 4 will look at the methodology. This is followed by our initial findings in Section 5 and finally we conclude our paper in Section 6.

II. BACKGROUND AND LITERATURE REVIEW

Digital transformation of public services is a global trend that seeks to improve accessibility, efficiency, and user involvement. Governments around the globe are utilizing technology to simplify bureaucratic procedures and improve service provision. The TAMM platform in Abu Dhabi exemplifies this shift by offering a centralized hub for digital government services. This literature review examines previous research on digital transformation within the public sector, automation, AI integration, and associated challenges, providing context for TAMM’s impact on process automation, service delivery, and service quality.

Digital Transformation in government services requires AI, big data and automation to enhance workflows [2]. Studies show that the success of government digitalization depends on its alignment with national strategic initiatives, such as the UAE's AI and blockchain strategies [2].

According to a study [3], in Kuwait, mobile-optimized government services are preferred by the public. Their study emphasizes three vital stages of digital transformation: digitization, digitalization, and comprehensive transformation. The phases can serve as a basis for assessing TAMM's development. Research has shown that government automation significantly reduces administrative burden and improves efficiency and eliminates errors [4].

Previous studies also indicate that government automation through digital platforms can reduce administrative burden and enhances efficiency while eliminating errors [4]. Additionally, Straub et al. [5] suggest a framework aimed at evaluating the operational robustness and normative significance of AI-driven automation within the public sector. Their research emphasizes that AI can take on repetitive bureaucratic tasks with effectiveness, leading to enhancements in the speed and accuracy of service delivery. The authors estimate that 84% of transactions conducted by the UK government could be automated, which could lead to annual savings of more than 1,200 person-years of work. In line with this global tendency, the TAMM platform—automating a range of public service transactions—demonstrates Abu Dhabi's dedication to developing into a digital government.

In addition, user perceptions and behavioral factors have a significant impact on how effective digital government platform. For example, [2] employed the Unified Theory of Acceptance and Use of Technology (UTAUT) and found that service reliability, social influence, and ease of use are critical factors influencing public adoption.

In a similar sense, [6] investigated AI-enabled public services through a social contract lens in their paper, highlighting the importance of transparency, human oversight, and trust. These findings indicate that in order for TAMM to bolster citizen engagement, service design should place a premium on transparency and usability.

To sum up, the examined studies apply quantitative surveys, literature reviews, and conceptual models. Finally, service efficiency was assessed by a study that surveyed 378 users of digital services [3] wherein technology adoption was assessed using UTAUT and Diffusion of Innovation Theory (DIT). These methods provide standards for evaluating TAMM's effect, especially in terms of user satisfaction and automation efficiency.

III. HYPOTHESES

Digital transformation fundamentally impact the government delivery of services [2] by means of process automation, accessibility, and better efficiency. Therefore, this study aims to address the following research question: "To what extent has TAMM's digital transformation improved process automation, service delivery speed and public trust in digital government services in Abu Dhabi?". Since TAMM is already adopted and used by many residents, our study focuses on post-usage perceptions.

To do so, we developed a set of hypotheses assessing individual impressions of TAMM's effectiveness to improve digital government services. We categorize our hypotheses into the following: Process automation, service delivery and public trust and quality.

Process automation reduces human intervention and reduces many errors that could arise from manual transactions. However, concerns regarding the effectiveness of these automated systems especially when human intervention is required. In his research [6] suggest that even though automation aims to streamline digital processes and government transactions reducing time and effort, its overall effectiveness is often judged by how easily users can navigate through these digital platforms. Moreover, automation does not necessarily mean that users will have a positive experience. [6] highlights that a well-automated service must be seamlessly integrates, user friendly and responsive. Therefore, the following hypothesis is proposed:

H1: The use of TAMM has significantly reduced the number of steps and physical visits require for government transaction.

Moreover, the success of process automation relies on usability and efficiency. If a platform is difficult to use, it cannot be considered efficient [3]. Since TAMM is being used by many, it should follow that the platform should be easy to find information and services. Therefore, we propose the following hypothesis:

H2: Users will find TAMM easy to navigate and use

Additionally, one of the main goals of digital transformation is to enhance the speed of government service delivery by providing real-time, user centric digital interactions [2]. For instance, AI-driven chatbots and self-service portals reduce waiting times as opposed to traditional in-person transactions. We propose the following hypothesis to assess whether TAMM has resulted in enhanced service delivery:

H3: The application of TAMM has sped up the response times for governmental transactions.

Moreover, although digital platforms enhance accessibility, they can raise concerns about inclusion, as individuals who find technology challenging may struggle to use these platforms. This raises the question of whether users experience real improvements in accessibility when using platforms such as TAMM and therefore, we propose the following hypothesis:

H4: TAMM's digital transformation has significantly improved the accessibility of government services.

While public trust is a critical factor in adopting digital transformation [6], it is shaped by ease of use, data security and privacy and transparency [3]. Even though TAMM aims to build trust through its digital services, question regarding data privacy arise. According to [5], research shows that when users find platforms easy to navigate, trust in e-government services increases. The platform in question is TAMM and hence, we hypothesize the following:

H5: A significant positive correlation exists between ease of use and trust in TAMM.

Finally, a link between digital transformation and improvements in quality of life has been proposed by a study [2]. We posit TAMM may have had similar impact on residents in Abu Dhabi and present our concluding hypothesis:

H6: TAMM has significantly improved convenience and quality of life in Abu Dhabi.

IV. METHODOLOGY

We use a quantitative approach to evaluate the extent to which TAMM, as an e- government platform, has enhanced process automation, increased the speed of service delivery, and strengthened public trust in e-government services.

A survey-based methodology was used, as it allows for a systematic assessment of user perception using quantitative and structured responses. Surveys have been widely used in e-government studies to evaluate user satisfaction and adoption rates [3]. This method is appropriate for identifying trends, statistical relationships, and demographic variations in how people experience TAMM's digital transformation. The survey will capture users' perception on TAMM across three key dimensions including process automation, service delivery and public trust and service quality.

The survey targeted residents of Abu Dhabi who have used or are using TAMM for government services. To ensure diversity and a representative sample of Abu Dhabi's population, the responses were gathered from various age groups, employment sectors and educational backgrounds. Our final sample size consisted of 250 different users of various age, employment status and educational background for statistical power.

VIII. FINDINGS

Initial findings support our hypotheses and seem to suggest that TAMM is successful in reducing bureaucratic inefficiencies, speeding up governmental transactions, and enhancing citizen confidence in digital services. This study also emphasizes the importance of digital platforms being user-friendly to foster trust and adoption.

V. CONCLUSION

A significant contribution of this research is its ability to link digital transformation with user experience. Preliminary results provide quantitative evidence that digital transformation improves process efficiency and boosts user satisfaction. Findings of this study will bring forth important insights of digital transformation in government services across customer experience, data utilization, service innovation, operational processes and value creation. By evaluating the effects of TAMM's digital transformation on process automation, service quality and delivery, the study contributes to academic research and practical policymaking.

REFERENCES

[1] Oh, K., Kho, H., Choi, Y., & Lee, S. (2022). Determinants for Successful Digital Transformation. Sustainability, 14(3), 1215.

[2] Alzarooni, A. I., Alhashmi, S. M., Lataifeh, M., & Rice, J. (2024). Navigating digital transformation in the UAE: Benefits, challenges, and future directions in the public sector. Computers, 13(281).

[3] Alshawaaf, N., & Alzougool, B. (2024). Evaluating success in digital transformation of government services: Insights from Kuwait. Preprints.

[4] Savignon, A. B., Zecchinelli, R., Costumato, L., & Scalabrini, F. (2024). Automation in public sector jobs and services: A framework to analyze digital transformation's impact. Transforming Government, 18(1), 49-70.

[5] Straub, V. J., Morgan, D., Hashem, Y., Francis, J., Esnaashari, S., & Bright, J. (2023, August). A multidomain relational framework to guide institutional AI research and adoption. In Proceedings of the 2023 AAAI/ACM Conference on AI, Ethics, and Society (pp. 512-519).

[6] Schmager, S., Grøder, C. H., Parmiggiani, E., Pappas, I., & Vassilakopoulou, P. (2024). Exploring citizens' stances on AI in public services: A social contract perspective. Data & Policy, 6, e19.

[7] Badri, M., Alkhaili, M., Aldhaheri, H., & Yang, G. (2024). Unveiling shadows: Exploring the dark side of digital transformation in Abu Dhabi. Digital Transformation and Society, 4(1), 39-56.

[8] Gasco Hernandez, M. (2024). Reflections on three decades of digital transformation in local governments. Local Government Studies, 50(6), 1028-1040.

Beyond Red Tape: Do Users Care Exploring AI's Potential to Streamline Bureaucracy

Yassin AlBlooshi
College of Business and Economics
United Arab Emirates University (UAEU)
Al Ain-UAE
201350231@uaeu.ac.ae

Abdulla AlMehrzi
College of Business and Economics
United Arab Emirates University (UAEU)
Al Ain-UAE
201406956@uaeu.ac.ae

Khaled AlHassani
College of Business and Economics
United Arab Emirates University (UAEU)
Al Ain-UAE
201703418@uaeu.ac.ae

Ananth Chiravuri
Associate Prof-College of Business and Economics
United Arab Emirates University (UAEU)
Al Ain-UAE
ananth.chiravuri@uaeu.ac.ae

Abstract—Government bureaucratic inefficiency in processes continues to obstruct service delivery, burden citizens with bureaucratic tasks, and undermine economic output. Our study examines whether and how AI can enhance government efficiency. We conducted a survey with 224 participants. Initial results indicate high confidence in AI to improve operating efficiency and eliminate bureaucratic delays. However, findings also indicated severe concerns over privacy, job replacement, and algorithmic bias. Our findings indicate that AI has significant potential to revolutionize bureaucratic processes. In addition, the influence of contextual factors suggests that targeted implementation strategies are more appropriate than universal ones.

Keywords—AI, Government, Bureaucracy, Efficiency, Privacy

I. INTRODUCTION

Governments throughout the world now work to enhance operational performance as public expectations about service quality continue to increase. [1] described bureaucracy as a rational-legal system designed for administration but people now attribute its inefficiencies to this method. The documented inefficiencies embrace procedural complexities together with administrative burdens and delays whereas these elements eat up 20-30% of administrative capacities but demonstrate no corresponding added value [2]. Administration burden in European Commission estimates requires EU businesses to spend around €124 billion every year [3] while total economies suffer costs reaching billions annually.

Government leaders have adopted digital transformation agendas throughout the past two decades by attempting to digitize existing bureaucratic processes at first. Governing institutions commonly shifted toward "digitized bureaucracy" instead of true transformation through their digitalization projects [4]. Presently, governments apply Artificial Intelligence (AI) technology as part of their efforts to establish new fundamental approaches for administrative procedures. Artificial Intelligence represents a strategic tool for efficient processing which helps decision makers achieve better accuracy as well as deliver superior services. The governments of Estonia together with Singapore and the United Kingdom use AI technology for different administrative processes [5].

The government adoption of AI systems leads to multiple issues about future bureaucratic choices and public accountability while changing how citizens relate to state institutions. The implementation of AI tools requires careful planning because concerns over algorithmic prejudice together with transparency issues and the threat of promoting social gaps need democratic structures which ensure effective governance principles and fairness. The problem goes beyond technological implementation because it requires redesigning existing bureaucratic systems to adapt them to digital functionality.

The motivation for this research stems from a significant gap in current literature on the application of artificial intelligence to reduce bureaucratic processes. Research on AI adoption procedures by government [6] and digital governance theories [7] remains substantial yet empirical studies on AI-enabled streamlining of bureaucratic non-value-adding processes remain scarce. Investigative studies have mostly explored individual model deployments and theoretical analysis while missing the relationship between automation implementation and bureaucracy performance across various governmental functions. [8] indicate that the measurement of artificial intelligence effects on administrative work remains weak and lacks proper division of bureaucratic tasks alongside automation and augmentation assessments. Specifically, the research question is whether the implementation of AI will lead to zero bureaucracy while increasing efficiency.

The rest of the paper is structured as follows: We start with an introduction followed by Section 2 where we examine the related literature as part of the Literature review. We then present the theoretical frameworks followed by hypotheses development in Section 3. We briefly discuss the methodology in Section 4, followed by the preliminary findings in Section 5. Finally we conclude in Section 6.

II. LITERATURE REVIEW

A. Understanding Bureaucracy in Modern Governance

Weber's traditional work about official systems of governance introduced bureaucracy as a structured authority system that follows rational laws in addition to established hierarchical arrangements and specialized labor organization with organized protocols [1]. Weber created an ideal bureaucracy that helped organizations achieve predictable results and administrative neutrality through administrative rules and documentation standards. Bureaucracies were presented by this theory as fundamental to states of the modern age because they enable the implementation of policies across broad populations and the governance of sophisticated social systems. Despite acknowledging the risk of excessive stiffness his model offered superior technical efficiency compared to other organizational structures mainly when large-scale management became essential [2].

Bureaucracy theory received new perspectives through [9] research on people who work directly with citizens at the frontline of public service and have flexible authority in implementing policies. Lipsky showed that public servants deliver institutional results primarily through their personal determinations because their professional choices must obey operational boundaries and official mandates

Research centered on digital transformations of bureaucratic systems has increased significantly during modern times. [10] documented the transition from street-level to "system-level" bureaucratic operations through which data systems began taking over choices that humans used to handle independently. The bureaucratic transformation generates vital inquiries about discretion handling and accountability structures and guaranteeing bureaucratic core values in systems that become increasingly automated. The author [11] explores how artificial intelligence shapes bureaucratic organizational discretion because it constrains and enables multiple types of discretionary authority which depends on specific tasks and organizational contexts. Study starting points for analyzing how AI systems can update bureaucratic operations without disrupting their regulatory requirements can be found in these scholarly works.

B. AI Technologies in Public Administration

The Public administration fields have adopted artificial intelligence technologies as they face different levels of implementation success. The deployment of artificial intelligence features includes robotic process automation (RPA) and intelligent document processing alongside chatbots that supply information to citizens [12]. Decision support systems that require complexity now form the basis of automated determination systems across welfare eligibility and tax monitoring and resource management applications. Through AI model predictions Estonia successfully identifies corporate tax default risks which helps authorities conduct specific enforcement strategy [6]. Similar to Singapore GovTech agency, the agency employs AI solutions through municipal maintenance where predictive analytics and citizen report prioritization automates repair crew deployment [5]. AI technologies demonstrate development through simple automation into advanced functions of prediction and optimization for administrative administration.

However, the field studies of AI deployment in public settings demonstrate conflicting results about public service efficiency and service delivery quality. Testing by [13] proved that digital systems containing AI components shortened application processing duration in specific cases but demonstrated no measurable impact on most procedures because contextual factors determined how technology worked. The research conducted by [14] identified substantial differences in AI implementation possibilities across government transactions specifically focusing on automatic processes of procedures which possess analytic capabilities. [2] developed this theoretical concept of digital technology impact on bureaucratic operations before these findings appeared. Technological changes show an unequal distribution among government activities and distinct regional areas because institutional structures together with organizational elements modify how these changes play out.

Redistribution of decision discretion within government hierarchies is another prominent feature of AI's impact. [15] has characterized algorithmic prediction systems as tools for "extracting discretion" from frontline workers and reallocating it to the designers, deployers, and managers of automated systems. This reallocation has been associated with increased centralization of administrative power, as locally adapted decision practices are standardized through algorithmic application. Private sector technology vendors have, in the process, gained influence over public administration processes through their role in system development and maintenance.

AI adoption in the public sector produces contradictory effects on professional identity as well as job satisfaction among personnel. Through ethnographic research by [16] showed that government employees react differently to AI adoption because of their background technological skills and professional status and ability to grasp how algorithms work with their work rather than replacing it. The public servants who saw AI tools improve their professional capabilities had superior job satisfaction in contrast to employees who thought automation reduced their ability to exercise judgment. The research confirms that AI implementation success depends on human aspects because organizations pursuing augmentation above substitution achieve better workforce retention while safeguarding institutional expertise.

III. HYPOTHESES AND CONCEPTUAL MODEL

This paper adopts [17] substitution-augmentation classification of automation technologies when developing its theoretical model. This analytical schema has specifically been adapted for administrative needs by establishing a link between task features including complexity and uncertainty against AI-driven changes while adopting [11] classification scheme. [2] framework of bureaucratic change includes an analysis that separates administrative transformation affecting procedures directly from broader organizational changes affecting credibility and service delivery quality. A comprehensive AI impact model came into being through the combination of these theoretical perspectives which identify both process changes and protections for core governance systems.

C. Hypotheses

As indicated before, this objective of this study is to examine whether AI can lead to zero bureaucracy and the economic impact of bureaucratic elimination. To examine these questions empirically, we formulate three specific testable hypotheses focused on measurable aspects of AI implementation in government processes. These are given below:

H1: The integration of AI tools in government processes is positively associated with efficiency.

H2: The integration of AI tools in government processes will be negatively associated with bureaucratic delays.

H3: The integration of AI tools in government processes will be positively associated with public satisfaction with governance services

D. Conceptual Model

The theoretical model for this study conceptualizes the relationships between AI adoption and bureaucratic performance, drawing on [2] model for identifying first-order and second-order impacts of digital transformation. AI adoption is the independent variable, operationalized in terms of technology adoption, scope of deployment, and system sophistication across government activities. The dependent variables—operational efficiency, bureaucratic delay reduction, and public satisfaction—are influenced both directly by AI and indirectly through mediating variables of task characteristics [11], organizational flexibility [13], and implementation approach (augmentation vs. automation) [17]. Control variables, including demographics, prior technology exposure, and trust in government, help to extricate the effect of AI from broad socio-political influences (see Figure 1).

IV. METHODOLOGY

A. Research Design

We plan to adopt a mixed methods design which merges both qualitative case study analysis and quantitative measurement capabilities for triangulating results between research approaches. Our methodology is based on comparable academic research which studied public viewpoints about technological transformation in governmental institutions [18].

The survey instrument was developed through an iterative process begun by identifying key measurement dimensions from the literature review. Available scales were adapted where feasible, with technology trust measurement items drawn from the [18] survey and bureaucratic experience items from [19] administrative burden framework. New items were developed for the measurement of AI governance perceptions, with questions developed to test each research hypothesis. An initial draft of the instrument was reviewed by a panel of five public administration and AI governance subject matter experts, resulting in improvements to question wording and response options. A pilot test was then conducted with 5 respondents representing a range of demographic groups, enabling item clarity, completion time, and internal consistency to be evaluated. Pilot feedback guided further revisions, in particular regarding technical terms that were clarified to enhance understanding among respondents without specific knowledge of AI technologies [5]. We incorporated these changes in the final survey.

Sampling Strategy

A convenience sampling approach was employed to collect data from adults (18 years and older) who had utilized at least one government service in the previous 12 months. Participants were recruited through available email lists, social media announcements, and participant networks.

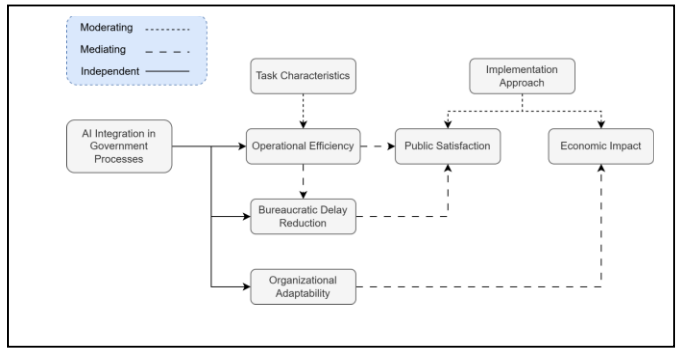


Fig. 6. Conceptual Model

V. FINDINGS

A total of 224 respondents from diverse demographics completed the survey. The breakdown was that the majority of the respondents fell within the 25-34 age group (46.9%), then 35-44 (25%), and under 25 (17.9%). The gender balance was relatively even with 51.3% female respondents and 48.7% males.

Initial findings indicate that survey respondents showed varying levels of confidence across different AI perception measures, with satisfaction with current services scoring highest and perceptions of AI's corruption reduction potential scoring lowest. This indicates respondents have more confidence in operational improvements than governance impacts. This pattern suggests higher confidence in AI’s operational benefits than its governance impact. Respondents indicate a high level of AI awareness, with 37.5% responding that they are “very familiar” and 46.9% that they are “somewhat familiar” with the use of AI in government services. Only 2.7% reported that they hadn’t heard about AI within this context. This relatively high level of familiarity is mirrored by positive opinion regarding the improvements that AI can deliver within government service efficiency, with 87.5% responding that AI could improve government service efficiency by “strongly agreeing” (46%) or “agreeing” (41.5%).

Our preliminary findings also show that 65.6% of the respondents saw bureaucratic delays to be a major issue with government services, but 39.7% expressed satisfaction with the speed of service delivery. This gap between current satisfaction and perceived severity suggests the necessity for solutions to tackle bureaucratic inefficiencies. When the respondents were asked how effective AI would be to help address these inefficiencies, 85.8% said that AI would be “very effective” (42.9%) or “effective” (42.9%).

VI. IMPLICATIONS AND CONCLUSION

This research attempts to make some valuable contributions to the theory regarding bureaucratic change in the age of AI. First, our preliminary findings seem to indicate the robust connection between AI integration and efficiency expectations that will enhance [1] early theory of bureaucracy by illustrating how technology upgrading can strengthen rather weaken service delivery.

This study findings could have practical implications for governance. Our findings may corroborate the necessity for digital literacy programs for citizens and government officials to allow for a smoother transition to AI-facilitated governance.

REFERENCES

[1] Weber, M., 2019. Economy and society: A new translation. Harvard University Press.

[2] Newman, J., Mintrom, M. and O’Neill, D., 2022. Digital technologies, artificial intelligence, and bureaucratic transformation. Futures, 136, p.102886.

[3] European Commission, 2018. Better regulation: Taking stock and sustaining our commitment. European Commission, Brussels. [online] Available at: https://commission.europa.eu/law/law-making-process/better-regulation/better-regulation-taking-stock-and-sustaining-our-commitment_en [Accessed date: 1 March 2025]

[4] Dunleavy, P., Margetts, H., Bastow, S. and Tinkler, J., 2006. New public management is dead—long live digital-era governance. Journal of public administration research and theory, 16(3), pp.467-494.

[5] Straub, V.J., Morgan, D., Bright, J. and Margetts, H., 2023. Artificial intelligence in government: Concepts, standards, and a unified framework. Government Information Quarterly, 40(4), p.101881.

[6] Wirtz, B.W., Weyerer, J.C. and Geyer, C., 2019. Artificial intelligence and the public sector—applications and challenges. International Journal of Public Administration, 42(7), pp.596-615.

[7] Margetts, H., 2022. Rethinking AI for good governance. Daedalus, 151(2), pp.360-371.

[8] Young, M.M., Himmelreich, J., Bullock, J.B. and Kim, K.C., 2019. Artificial intelligence and administrative evil. Perspectives on Public Management and Governance, 4(3), pp.244-258.

[9] Lipsky, M., 2010. Street-level bureaucracy: Dilemmas of the individual in public service. Russell sage foundation.

[10] Bovens, M. and Zouridis, S., 2002. From street-level to system-level bureaucracies: how information and communication technology is transforming administrative discretion and constitutional control. Public administration review, 62(2), pp.174-184.

[11] Bullock, J.B., 2019. Artificial intelligence, discretion, and bureaucracy. The American Review of Public Administration, 49(7), pp.751-761.

[12] Pencheva, I., Esteve, M. and Mikhaylov, S.J., 2020. Big Data and AI–A transformational shift for government: So, what next for research?. Public Policy and Administration, 35(1), pp.24-44.

[13] Cordella, A. and Tempini, N., 2015. E-government and organizational change: Reappraising the role of ICT and bureaucracy in public service delivery. Government information quarterly, 32(3), pp.279-286.

[14] Straub, V.J., Hashem, Y., Bright, J., Bhagwanani, S., Morgan, D., Francis, J., Esnaashari, S. and Margetts, H., 2024. AI for bureaucratic productivity: Measuring the potential of AI to help automate 143 million UK government transactions. arXiv preprint arXiv:2403.14712.

[15] Hong, S.H., 2023. Prediction as extraction of discretion. Big Data & Society, 10(1), p.20539517231171053.

[16] Gritsenko, D. and Wood, M., 2022. Algorithmic governance: A modes of governance approach. Regulation & Governance, 16(1), pp.45-62.

[17] Acemoglu, D. and Restrepo, P., 2020. The wrong kind of AI? Artificial intelligence and the future of labour demand. Cambridge Journal of Regions, Economy and Society, 13(1), pp.25-35.

[18] European Union Agency for Fundamental Rights, 2020. Getting the future right: Artificial intelligence and fundamental rights. Publications Office of the European Union. [online] Available at: https://fra.europa.eu/en/publication/2020/artificial-intelligence-and-fundamental-rights [Accessed date: 2 March 2025]

[19] Herd, P. and Moynihan, D.P., 2019. Administrative burden: Policymaking by other means. Russell Sage Foundation.

Artificial Intelligence for Enhanced Experimentation in Software Ecosystems: A Systematic Literature Review

Shady Hegazy
Siemens Technology
Siemens AG
Munich, Germany
shady.hegazy@siemens.com

Christoph Elsner
Siemens Technology
Siemens AG
Munich, Germany
christoph.elsner@siemens.com

Jan Bosch
Department of Computer Science and Engineering
Chalmers University of Technology
Göteborg, Sweden
jan.bosch@chalmers.se

Helena Holmström-Olsson
Department of Computer Science and Media Technology
Malmö University
Malmö, Sweden
helena.holmstrom.olsson@mau.se

Abstract— Software ecosystems are networks of interconnected actors co-creating value through a shared technological platform, achieving accelerated growth via network effects. This interconnectedness creates a complex decision space, requiring experimentation-based, data-driven decision making. This study presents a systematic literature review on AI-enhanced experimentation in software ecosystems. Following structured screening and quality assessment, 63 studies were included. Extracted data underwent descriptive, thematic, and cross-analyses. The study offers three contributions: an overview of AI-enhanced experimentation objectives; a phase-aligned analysis; and a framework for AI integration into experimentation pipelines.

Keywords—artificial intelligence, experimentation, software ecosystems, causal inference

I. INTRODUCTION

Software ecosystems (SECOs) have transformed the way value is created in many industries, introducing transformational shifts from vertical integration and firm-confined value-creation to co-creation of value through a network of interconnected actors who collaborate via a shared technological platform within one ecosystem. An example from the mobile industry is the success of ecosystem-oriented approaches of Apple's IOS and Google's Android in attracting around 300,000 third-party developers each in an ecosystem around their application marketplaces, versus the failure of Blackberry's vertically integrated approach which relied on 8000 in-house developers (Wang et al., 2017). However, such growth comes at many costs including additional complexities in orchestration and governance of such multi-sided ecosystems. For instance, TradeLens, a blockchain-based platform ecosystem developed by Maersk and IBM, was ultimately shut down, largely due to governance and adoption challenges as the platform struggled with aligning incentives among highly interdependent and often competing actors, the lack of trust in the platform sponsor's neutrality, hesitancy over data sharing, and difficulties managing standardization and interoperability across legacy systems, which hindered adoption among third-party industry partners (Najati, 2025). Such complexity necessitates adopting a data-driven decision making approach in order to reduce decision risks and uncertainty in an environment with various types of actors who have different, and sometimes conflicting, goals (Liu et al., 2020). While analytics can provide relevant insights to effectively inform the decision making process, online controlled experiments are considered the gold standard for evidence-based decision-making (Liu et al., 2020). Experiments in SECOs "often exhibit complex network effects. Consequently, unless designed carefully, the experiments could suffer from interference" (Candogan et al., 2023). In addition to interference and bias, network effects in SECOs contribute to the propagation of negative effects of users' exposure to under-performing variants during experiments, which is often known as regret (Spang et al., 2021). The vast amount of data generated in SECOs present opportunities for the utilization of advances in artificial intelligence (AI) to ease these challenges (Charness et al., 2023). While prior secondary studies explored contribution of certain AI methods to experimentation pipelines (Charness et al., 2023), different aspects of online controlled experiments in general (Auer & Felderer, 2018; Quin et al., 2024; Ros et al., 2018), there was no secondary research addressing AI-enhanced experimentation in SECOs. Hence, we conducted this systematic literature review aiming to provide the following contributions: an overview of the major objectives for AI application in experimentation pipelines in SECOs; a phase-aligned analysis of AI-enhanced experimentation; and a framework for AI integration into experimentation pipelines.

II. METHODOLOGY

This study follows a systematic literature review methodology to ensure comprehensive and replicable coverage of the existing research.

E. Research Question

We applied the Population, Intervention, Comparison, Outcome, and Context (PICOC) criteria to produce the following structure to match the study scope (Kitchenham & Charters, 2007).

- Population: SECOs.
- Intervention: Experimentation.

- Comparison: AI-enabled enhancements to experimentation processes.
- Outcome: Enhanced data-driven decision making through more efficient experimentation processes.
- Context: SECOs in business contexts.

Using the above structure, we formulated the following research question: In what ways has AI-supported experimentation been applied in software ecosystems to improve data-driven decision making?

F. Search Strategy

Based on the breakdown of the research question structure, search query terms were selected for each section of the PICOC criteria, as shown in Table I.

TABLE VII. SEARCH QUERY TERMS.

Population	”platform ecosystem*” OR ”software ecosystem*” OR ”seco” OR ”digital ecosystem*” OR ”platform*”
Intervention	”experimentation” OR (”experiment*” AND ”design*”) OR ”online experiment*” OR ”online controlled experiment*” OR ”controlled experiment*” OR ”user testing” OR ”a/b testing”
Comparison	
Outcome	”quantitative” OR ”driven” OR ”analy*” OR ”inference” OR ”continuou*” OR ”adapt*” OR ”data” OR ”network*” OR ”graph” OR ”agent*” OR ”sampl*” OR ”cluster*” OR ”classif*” OR ”group*” OR ”causal*” OR ”bias*”
Context	”business*” OR ”industry” OR ”market” OR ”customer” OR ”user”

The search query was executed on the titles and abstracts of research articles in the Scopus, ACM Digital Library, and IEEE Xplore databases, resulting in 1,349, 597, and 113 hits, respectively. The meta-data of the resulting 2,059 studies was retrieved from the corresponding sources.

G. Inclusion And Exclusion Criteria

The set of inclusion and exclusion criteria used for this review are listed in Table II along with the number of exclusion decisions for which each criterion was the most prominent basis.

TABLE VIII. INCLUSION AND EXCLUSION CRITERIA.

Existence	Criterion	Absence	Exclusions
IC1	A primary study, not a secondary or a tertiary study.	**EC1**	77
IC2	Published in a peer-reviewed journal, conference proceeding, or book chapter not in gray literature reports, blog posts, or non-academic publications.	**EC2**	9
IC3	Published in English.	**EC3**	5
IC4	Not a duplicate or a version of another included study.	**EC4**	421
IC5	Have a significant component related to SECOs.	**EC5**	608
IC6	Discusses experimentation in SECOs with the aim of enabling or enhancing data-driven decision making.	**EC6**	846
IC7	Focuses on business contexts.	**EC7**	5

H. Quality Assessment

The included 88 studies underwent a quality assessment process using a variation of the Standard Quality Assessment Criteria (SQAC) tailored for quantitative research (Kmet et al., 2004). The criteria were scored on a 3-point Likert scale, with zero points for unmet criteria, one point for partially met criteria, and two points for met criteria. The average quality score after removing the 25 disqualified studies was 73.6%.

I. Data Extraction

The following data categories and the underlying data points were extracted from the full texts of the 63 included studies:

- Identifiers: study title; authors; publication year; abstract.
- Software ecosystem: type; number of sides; actor type; industry.
- Experimental Design: method; internal platform; analysis unit; randomization unit; sampling technique; traffic allocation; evaluation metrics.
- Data: volume; types; sources; collection rate; processing rate; AI support.
- AI Support: phase; main techniques; family; task; direct outcome.

III. RESULTS

A. Major Objectives of AI Use in Experimentation Pipelines

1. Maximization of experimentation outcomes

Experimentation results typically help declare a higher performing decision, policy, or variant. However, the use of AI could enable further outcomes by using the resulting data to inform future experiments or generate new insights. For example, the study in (Duivesteijn et al., 2017) used experiment data to dynamically serve different variants to different user groups, through the use of AI methods such as exceptional model mining.

2. Sampling optimization

AI methods were also deployed to construct more informative and balanced samples (Candogan et al., 2023). Studies used graph-based clustering to define interference-aware clusters, matching techniques to reduce covariate imbalance, and segmentation algorithms to stratify populations based on latent characteristics (Brennan et al., 2022).

3. Experimentation cost reduction

Multiple studies applied AI to dynamically reallocate traffic away from underperforming variants, reducing exposure to inferior treatments and shortening the duration required to reach conclusive results (Glynn et al., 2020). Other studies used various AI for the task of offline policy evaluation to reduce the exposure to underperforming variants during live experiments by predicting higher performing variants and allocating more traffic towards them (Li & Xie, 2020).

4. Intelligent experimentation pipelines

AI methods were used to dynamically revise experimental designs and pipeline parameters based on emerging evidence. For instance, the study in (Bojinov et al., 2023) described the use of AI to predict carryover bias from treatments in switchback experiments in order to dynamically space sequential treatments accordingly.

5. Causal inference enhancement

As ideal experimental conditions are hard to achieve in interconnected SECOs, AI was commonly applied to enhance the robustness of causal inference. For instance, several studies leveraged deep learning based causal inference, ensemble learning, and causal forests to estimate treatment effects in the presence of heterogeneity, interference, or selection bias (Ye et al., 2023). In addition, AI methods, such as analysis, were used to construct more accurate overall evaluation criteria, thus indirectly enhancing causal inference (Hornback et al., 2023).

B. Phase-Based AI Support in Experimentation Pipelines

1. Design phase

AI support in the planning and design phases focused on structuring experimental units, forecasting outcomes, and policy evaluation. AI was also used for causal forecasting and offline policy evaluation to prioritize high-performing treatments and reduce experimentation risks and costs.

2. Execution phase

In this phase, AI methods were used to dynamically reallocate traffic, forecast responses, and adjust experiment structures based on emerging evidence. AI was commonly applied to minimize regret and optimize policy rollout under budget and engagement constraints. While no end-to-end AI-orchestrated pipelines were observed, AI support in this phase enabled adaptive, intelligent experimentation with substantial efficiency gains and regret reduction.

3. Evaluation phase

AI support in the post-experiment phase focused on causal inference, bias and interference correction, heterogeneity analysis, and insight generation for future experiments. Its value was especially evident in ecosystems with noisy, high-dimensional data such as networking ecosystems. For instance, causal forests and ensemble methods were used to model heterogeneous effects in complex ecosystems.

C. A Framework for AI-Enhanced Experimentation Pipelines

Through our analysis of the reviewed studies we developed a framework, described in Table III, that formalizes the introduction of AI methods in experimentation pipelines by connecting phases, objectives, and methods.

IV. CONCLUSION

Through a systematic literature review, we investigated how artificial intelligence can enhance experimentation in software ecosystems. From an initial pool of 2,059 studies, 63 met the inclusion and quality criteria. We analyzed the integration of AI across experimentation pipelines, identifying key goals such as bias mitigation, regret minimization, and experimentation cost–benefit optimization. Our phase-aligned analysis showed AI support is most common in the design and evaluation phases, with frequent use of methods like graph clustering, reinforcement learning, and causal forests. We propose a structured framework linking AI techniques to experimentation enhancement objectives and phases, providing a foundation for formalization of AI integration in experimentation pipelines.

TABLE IX. A FRAMEWORK FOR AI-ENHANCED EXPERIMENTATION PIPELINES (AIEXP).

	Design	**Execution**	**Evaluation**
Sampling optimization	Graph cluster randomization; hierarchical clustering.	Online bias/variance balancing.	Regression-adjusted CUPED; covariate regression.
Experimentation cost reduction	Monte Carlo simulations; Bayesian machine learning.	Sequential policy learning; multi-armed bandits.	Meta-learning; experiment recommender systems.
Intelligent experimentation pipelines	Adaptive segmentation.	Automated early stopping; Dynamic guardrails exploration.	Generative AI for reporting.
Causal inference enhancement	Predictive power analysis.	Delayed feedback prediction; deep Q-learning.	Debiased machine learning; Deep learning causal inference; sentiment analysis; ensemble learning; transfer learning.
Outcome maximization	Offline policy learning.	Contextual bandits; uplift modelling.	Exceptional model mining; exploratory data analysis.

REFERENCES

[1] Auer, F., & Felderer, M. (2018). Current State of Research on Continuous Experimentation: A Systematic Mapping Study. 2018 44th Euromicro Conference on Software Engineering and Advanced Applications (SEAA), 335–344.

[2] Bojinov, I., Simchi-Levi, D., & Zhao, J. (2023). Design and Analysis of Switchback Experiments. In Management Science (Vol. 69, Issue 7, pp. 3759–3777). https://doi.org/10.1287/mnsc.2022.4583

[3] Brennan, J., Mirrokni, V., & Pouget-Abadie, J. (2022). Cluster randomized designs for one-sided bipartite experiments. Proceedings of the 36th International Conference on Neural Information Processing Systems, 37962–37974.

[4] Auer, F., & Felderer, M. (2018). Current State of Research on Continuous Experimentation: A Systematic Mapping Study. 2018 44th Euromicro Conference on Software Engineering and Advanced Applications (SEAA), 335–344. https://doi.org/10.1109/SEAA.2018.00062

[5] Bojinov, I., Simchi-Levi, D., & Zhao, J. (2023). Design and Analysis of Switchback Experiments. In Management Science (Vol. 69, Issue 7, pp. 3759–3777). https://doi.org/10.1287/mnsc.2022.4583

[6] Brennan, J., Mirrokni, V., & Pouget-Abadie, J. (2022). Cluster randomized designs for one-sided bipartite experiments. Proceedings of the 36th International Conference on Neural Information Processing Systems, 37962–37974.

[7] Candogan, O., Chen, C., & Niazadeh, R. (2023). Correlated Cluster-Based Randomized Experiments: Robust Variance Minimization. Proceedings of the 24th ACM Conference on Economics and Computation, 411. https://doi.org/10.1145/3580507.3597820

[8] Charness, G., Jabarian, B., & List, J. (2023). Generation Next: Experimentation with AI. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.4574623

[9] Duivesteijn, W., Farzami, T., Putman, T., Peer, E., Weerts, H. J. P., Adegeest, J. N., Foks, G., & Pechenizkiy, M. (2017). Have It Both Ways—From A/B Testing to A&B Testing with Exceptional Model Mining. In Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics): Vol. 10536 LNAI (pp. 114–126). https://doi.org/10.1007/978-3-319-71273-4_10

[10] Glynn, P., Johari, R., & Rasouli, M. (2020). Adaptive experimental design with temporal interference: A maximum likelihood approach. Proceedings of the 34th International Conference on Neural Information Processing Systems, 15054–15064.

[11] Hornback, A., Buckley, S., Kos, J., Bunin, S., An, S., Joyner, D., & Goel, A. (2023). A Scalable Architecture for Conducting A/B Experiments in Educational Settings. Proceedings of the Tenth ACM Conference on Learning @ Scale, 373–377. https://doi.org/10.1145/3573051.3596190

[12] Kitchenham, B. A., & Charters, S. (2007). Guidelines for performing systematic literature reviews in software engineering (No. EBSE 2007-001). Keele University and Durham University Joint Report / Keele University.

[13] Kmet, L. M., Lee, R. C., & Research, A. H. F. for M. (2004). Standard Quality Assessment Criteria for Evaluating Primary Research Papers from a Variety of Fields. Alberta Heritage Foundation for Medical Research.

[14] Li, G., & Xie, H. (2020). Auxiliary Decision-making for Controlled Experiments based on Mid-term Treatment Effect Prediction: Applications in Ant Financial’s Offline-payment Business. 2, 19–30. https://doi.org/10.5220/0009770500190030

[15] Liu, C. H. B., Chamberlain, B. P., & McCoy, E. J. (2020). What is the Value of Experimentation and Measurement?: Quantifying the Value and Risk of Reducing Uncertainty to Make Better Decisions. In Data Science and Engineering (Vol. 5, pp. 152–167). https://doi.org/10.1007/s41019-020-00121-5

[16] Najati, I. (2025). Exploring the failure factors of blockchain adopting projects: A case study of tradelens through the lens of commons theory. Frontiers in Blockchain, 8, 1503595. https://doi.org/10.3389/fbloc.2025.1503595

[17] Quin, F., Weyns, D., Galster, M., & Silva, C. C. (2024). A/B testing: A systematic literature review. Journal of Systems and Software, 211, 112011. https://doi.org/10.1016/j.jss.2024.112011

[18] Ros, R., Ros, R., Runeson, P., & Runeson, P. (2018). Continuous experimentation and A/B testing: A mapping study. Null. https://doi.org/10.1145/3194760.3194766

[19] Spang, B., Hannan, V., Kunamalla, S., Huang, T.-Y., McKeown, N., & Johari, R. (2021). Unbiased experiments in congested networks. Proceedings of the 21st ACM Internet Measurement Conference, 80–95. https://doi.org/10.1145/3487552.3487851

[20] Wang, H., Liu, Z., Guo, Y., Chen, X., Zhang, M., Xu, G., & Hong, J. (2017). An Explorative Study of the Mobile App Ecosystem from App Developers’ Perspective. Proceedings of the 26th International Conference on World Wide Web, 163–172. WWW ’17: 26th International World Wide Web Conference. https://doi.org/10.1145/3038912.3052712

[21] Ye, Z., Zhang, Z., Zhang, D. J., Zhang, H., & Zhang, R. (2023). Deep Learning Based Causal Inference for Large-Scale Combinatorial Experiments: Theory and Empirical Evidence. Proceedings of the 24th ACM Conference on Economics and Computation, 1160. https://doi.org/10.1145/3580507.3597718

Integrating AI into the Coaching Process in the Banking Industry

Fatima AlAli
Business Analytics Program
Abu Dhabi School of Management
Abu Dhabi, UAE
adsm-215058@adsm.ac.ae

Evi Indriasari Mansor
Business Analytics Progra
Abu Dhabi School of Management
Abu Dhabi, UAE
e.mansor@adsm.ac.ae

Abstract—The paper presents the incorporation of Artificial Intelligence (AI) into coaching in banking, focusing on its role in employee development, productivity, and efficiency. The research highlights the way AI improves coaching by using a digital infrastructure that includes NLP and machine learning, deploying real-time analytics, personalization in feedback, and automated support. The research is informed through qualitative means, including literature reviews and interviews with experts. Findings show that AI improves decisions while cutting costs but has challenges ranging from ethical concerns to resistance and reduced interaction. The research presents strategic insights and guidelines for responsible AI adoption into HR and coaching that will serve banking staff and leaders, as well as policymakers in the digital transformation quest.

Keywords—artificial intelligence, coaching process, banking industry

I. INTRODUCTION

Artificial Intelligence (AI) has had a tremendous impact across industries worldwide, including the banking one. Its infusion in various banking processes has further resulted in higher efficiency, lower errors, and better customer experience. Among the many applications, AI appears to be a good opportunity for improving employee development via digital coaching frameworks [1]. The study hereunder aims to assess avenues for embedding AI in those processes with the potential for enhancement, automation, and streamlining of coaching activities within banking institutions, especially in the UAE, using "Abu Dhabi Islamic Bank (ADIB)" as a reference point. An organization's sustainability is closely coupled with its leadership in decision-making and team guidance in today's dynamic corporate environment. This is critically important in the banking sector, where a few seconds made in decision-making can mean the difference between great financial and reputational loss. For example, if Abu Dhabi Islamic Bank (ADIB) leaders do not coach their frontline employees with updated information, customer dissatisfaction will result in a drastic drop in service demand, which in turn would kill the bank's performance and profitability [2].

Robust coaching processes grounded on data that help with decision-making would help mitigate such risks. AI is poised to completely transform the way we think about coaching in terms of its powers in data analysis, automation, and real-time feedback. Hence, this research investigates how AI can support effective coaching by leaders to enhance employee performance and sustain growth.

Therefore, the aim of this research is to explore how AI can be utilized to digitalize and enhance coaching processes within banking institutions to improve decision-making and workforce productivity.

II. LITERATURE REVIEW

Coaching in the industry has continued to be face-to-face or virtual sessions on compliance, customer service, and skill development [4]. Although effective, these traditional modes are resource-intensive, time-consuming, and have little personalization of their own. According to Ige, Kupa [7] such traditional forms of coaching have relatively poor efficiencies and very much rely on human trainers.

Transforming such a model would, however, be possible through the introduction of AI. JPMorgan Chase & Co. (2017) state that AI allows personalized learning pathways along with giving automated feedback according to predictive analytics and keeping an eye on performance. Königstorfer & Thalmann, (2020) equally point out that AI helps in coaching by giving individualized, real-time insight and engaging employees. This technology hence creates the opportunity to align training with individual learning needs, thus enabling banks to develop targeted and effective coaching approaches. For example, COiN by JPMorgan Chase helps demonstrate how AI can be used for training purposes, offering data-supported insights for employees' improvements in selected areas [5]. AI systems can be costly, including regular upgrades to keep pace with regulatory and technological changes; which makes implementation in itself very difficult [6].

AI's role in knowledge retention and employee development. On one hand, a disadvantage of greater reliance is that it can hurt the development of interpersonal skills—skills critical to client-facing roles [7]. Thus, a hybrid coaching model, interspersed with AI and human interaction, becomes inevitable. One such instance exists with Amazon's AI training systems, where workers under a regime of personalized training receive real-time feedback [8]. AI identifies weaknesses and enhances specific skills. Yet, a balance between technical and artificial training offers an effective approach toward the holistic development of employees.

III. METHODOLOGY

This research uses a qualitative research design in semi-structured interviews with senior executives in the banking industry, particularly those who are well-informed about digital transformation and AI-based coaching processes. Ethics approval for the research was obtained from the ADSM ethics committee, which addressed major issues including informed consent, confidentiality, anonymity, and data protection.

The literature review was undertaken with the identification of appropriate academia from credible databases such as SAGE, ScienceDirect, and Google Scholar. Search terms were "AI in coaching," "digital coaching in banking," and "AI-enhanced learning and development". Then, only the most pertinent and recent studies were chosen for detailed review in support of the research.

Primary data were collected via semi-structured interviews: the format allows participants to express their experiences and perceptions freely, while the researcher can probe deeper into emerging themes. This qualitative technique is consonant with the research's aim of understanding the implications of AI on coaching practices in the real world, as experienced by the participants.

Participants were selected based on criteria that included their roles in digital transformation initiatives and collaborative work with AI-driven coaching tools. The interview questions include key benefits of AI in coaching processes; personalization of learning through AI; areas having the highest impact of AI; implementation challenges; impact on employee motivation and commitment; comparative efficiency with traditional coaching; risks related to AI dependence; ethical aspects of AI-driven coaching; and improving productivity and decision-making with AI.

Interviews were recorded with consent from participants and transcribed verbatim, preserving anonymity. The transcripts were edited for clarity, without changing the gist of the responses [11].

IV. DATA ANALYSIS

Data were analyzed via thematic analysis, as proposed by Rodrigues et al. (2020). It is fitting for use in identifying, analyzing, and reporting patterns within qualitative data.

Steps of Analysis:

1. Familiarization: Transcribing was repeatedly read to immerse oneself in the data.
2. Initial Coding: Data segments pertinent to the research objectives were coded.
3. Development of Themes: Codes were grouped under key themes such as:
   - Integration Strategies
   - Challenges in AI Application
   - Employee Development and Motivation
4. Review and Refine: The themes were revised and validated to ensure that they accurately representing the data.
5. Definition and Labeling: Each theme was defined clearly with evidence supporting it.
6. Use of NVivo: NVivo software was used to organize and manage data so that visual mapping and deeper comparative analysis of themes were done.

To ensure academic rigor, the research employed strategies to reinforce the trustworthiness of qualitative research: credibility, dependability, transferability, and confirmability [12].

- Credibility: In-depth interviews with well-experienced professionals, together with member checking, provide the opportunity given to the participants to check their transcripts for accuracy.
- Dependability: Documenting every stage of the process so that other researchers may replicate it, or auditors may follow an audit trail.
- Transferability: Detailed contextual descriptions about the participants, the setting, and the study process provide information for readers to determine the applicability in other contexts.

V. FINDINGS

The findings of this research highlight the increasing relevance of AI in terms the decision-making, and different kinds of streamlining operations and even focuses on minimizing different kinds of human error in various areas such as data processing and customer services. This research also focuses on the ways AI provide support to the coaching processes in banks in terms of improving employee performance, productivity and even the satisfaction level of the customer.

AI primarily aims to enhance effectiveness and even the precision of the coaching with the help of automating different kinds of repetitive queries and on delivering real-time feedback. Employees also get the benefits from faster responses and even tailoring the learning, particularly with the tools which are powered by Natural Language Processing (NLP) and data analytics. On the other hand, the problem statement mainly highlights the challenges such as the lower level of productivity and even poor decision-making among the different banking staff even stemming from the lack of different kinds of structured coaching systems. The incorporation of AI especially with the help of different technologies such as cloud computing and even data analytics mainly helps to automate and even increase the decision-making processes.

This research mainly references the use of AI in the ADIB in terms of analyzing the behavior of the customer which increases the chances for more personalized delivery of the services and even makes an improvement in the efficiency of the employee. It also uncovered the different kinds of barriers and even the major challenge is the technical skill gaps among the employees which mainly focus on creating resistance in terms of the adoption of AI. The majority of the staff members lack the required proficiency in terms of interacting effectively with the different kind of systems of AI which mainly increase the stress and increase the chances for disengagement. Financial constraints mainly include the implementation cost of the AI and even the system updates. Ethical concerns become another key finding and even this study focuses on underlining issues such as breaches related to data privacy and even the biased algorithm and the reduced interaction between the humans.

VI. DISCUSSION

AI into an unprecedented journey toward performance optimization for the banking sector by automating tedious cycles of training and thus saving time and resources useful for continuous learning. AI can analyze strengths and weaknesses to offer targeted assistance and skill-building. Rodrigues et al. (2022) have argued that personalization enhances training duration and the quality of customer service offered. This shows the effectiveness of AI in employee training at HSBC, where it reduced training time and operational resources [3]. The AI system monitored employee performance during training sessions, providing just-in-time coaching, and 20% more customer satisfaction. Ige, Kupa, and Ilori (2024) further note that AI-enhanced coaching is somehow an accelerant to learning, service quality, and customer experiences; hence, it's strategic for AI to drive superior performance.

Though the banking coaching domain witnesses increased AI applications, notable research gaps still linger. Most studies do not address the long-term consequences of AI for employee engagement, interpersonal skill development, or diversity of skills. Very few have researched the need for blending AI- and human-coaching approaches, especially when we consider the dangers that come with data privacy, ethical leveraging, and regulatory compliance [3]. Furthermore, there exists no substantive framework linking AI coaching to HR and strategy roles in fostering digital transformation in banks. Without grasping the strategic integration of AI toward sustainable growth, its promise may remain underused. Bridging these gaps could further ensure the role of AI in workforce upskilling in both hard and soft skills [10].

VII. CONCLUSION

The research mainly provides that the incorporation of AI into the coaching processes in the case of the banking sector mainly increases the performance of the employee, operational efficiency and even provides personalized learning. On the other hand, AI tools mainly offer real-time feedback and even automate different kinds of repetitive tasks and provide support to decision-making. Successful implementation mainly requires overcoming barriers such as technical skill gaps, financial constraints and even ethical concerns regarding data privacy and bias. The hybrid model mainly makes an effective combination of AI capabilities.

REFERENCES

[1] J. Bertot, "Social Media, Open Platforms, and Democracy: Transparency Enabler, Slayer of Democracy, Both?," Hawaii.edu, vol. 12, no. 2, pp. 12, 2019.

[2] T. E. Edunjobi and O. A. Odejide, "Theoretical frameworks in AI for credit risk assessment: Towards banking efficiency and accuracy," International Journal of Scientific Research Updates 2024, vol. 7, no. 01, pp. 092-102, 2014.

[3] J. Holmström, "From AI to digital transformation: The AI readiness framework," Business Horizons, vol. 65, no. 3, pp. 329-339, 2022.

[4] F. Königstorfer and S. Thalmann, "Applications of Artificial Intelligence in commercial banks – A research agenda for behavioural finance," Journal of Behavioral and Experimental Finance, vol. 27, no. 1, p. 100352, 2020.

[5] J. Pmorganchase, "Jamie Dimon’s Letter to Shareholders," 16 April 2025. [Online]. Available: https://www.jpmorganchase.com. [Accessed 16 April 2025].

[6] V. Nimmagadda, "Artificial Intelligence for Predictive Maintenance of Banking IT Infrastructure: Advanced Techniques, Applications, and Real-World Case Studies," Journal of Deep Learning in Genomic Data Analysis, vol. 2, no. 1, pp. 86-122, 2022.

[7] A. B. Ige, E. Kupa and O. Ilori, "Analyzing defence strategies against cyber risks in the energy sector: Enhancing the security of renewable energy sources," International Journal of Science and Research Archive, vol. 12, no. 1, pp. 2978-2995, 2024.

[8] A. Ashta and H. Herrmann, "Artificial intelligence and fintech coaching: An overview of opportunities and risks for banking, investments, and microfinance," Strategic Change, vol. 30, no. 3, pp. 211-222, 2021.

[9] L. Aziz and Y. Andriansyah, "The role artificial intelligence in modern banking: an exploration of AI-driven approaches for enhanced fraud prevention, risk management, and regulatory compliance," Reviews of Contemporary Business Analytics, vol. 6, no. 1, pp. 110-132, 2023.

[10] E. Mogaji and N. Nguyen, "Managers' understanding of artificial intelligence in relation to marketing financial services: insights from a cross-country study," International Journal of Bank Marketing, vol. 40, no. 3, pp. 1272-1298, 2022.

[11] A. Rodrigues, F. Ferreira, F. Teixeira and C. Zopounidis, "Artificial intelligence, digital transformation and cybersecurity in the banking sector: A multi-stakeholder cognition-driven framework," Research in International Business and Finance, pp. 101616, 2022.

[12] S. Mondal, S. Das and V. Vrana, "How to bell the cat? A theoretical review of generative artificial intelligence towards digital disruption in all walks of life," Technologies, pp. 44, https://www.mdpi.com/2227-7080/11/2/44, 2023.

Deep Learning in Forensic Injury Identification

Khudooma S. Alnuaimi
MSBA, Abu Dhabi School of Management Abu Dhabi, UAE
khudooma@gmail.com

Ishtiaq Rasool Khan
Abu Dhabi School of Management
Abu Dhabi, UAE
i.khan@adsm.ac.ae

Abstract -- This study implements the EfficientNetB0 deep learning model to classify forensic skin injuries into blunt, sharp, or gunshot categories. Through data augmentation and transfer learning, the model achieved single person's observation an accuracy of 99.3%. A real time Gradio interface, enable intuitive image uploads and classification. This research is supporting the AI Integration in forensic science, medicine and pathology. It is expected to have a significant impact by improving accuracy, enabling automation, and enhancing collaborative investigative case work in the field.

Keywords--- Forensic Wound Cavalry, Deep Learning, Convolutional Neural Networks, Image Analysis, Artificial Intelligence

I. INTRODUCTION

The classification of wounds for forensics is important in criminal investigations because the types of injuries like blunt force trauma, sharp force trauma and gunshot wounds have significant legal implications and influence the construction of the case narrative [3]. Current approaches heavily rely on a single person's observation which is highly time intensive and inaccurate. Automation of forensic skin injury classification is sought through Deep Learning with Convolutional Neural Networks, particularly EfficientNetB0 , in this research [1]. A model was created to classify images of wounds by curating a dataset with a boundless collection of images until a satisfactory accuracy metric was reached through numerous augmentations. The model was evaluated on conventional measurements of performance, including but not limited to, accuracy, precision, recall, F1 score, and a real-time forensic application interface was built using Gradio [4], [8]. Advances in the automation of digital forensics and legal medicine, alongside the unity of processes and bias mitigated nominal controls, enhance the efficiency of the system under automation with changing system variables.

II. LITERATURE REVIEW

The classification within forensics has included the manual inspection of an expert which adds an element of subjectivity and bias to the variances within the scrutiny and investigations. AI has come forth as an option for complex forensic investigations which continue to increase in number, complexity, and volume [6]. Research shows that Convolutional Neural Networks (CNNs) are competent at classifying blunt force trauma, sharp force injuries, and even gunshot wounds through spatial level imaging interpretation [1]. Incorporating AI standardizes evaluations across forensic facilities, while deep learning enhances the accuracy of the analysis. Models such as EfficientNetB0 have demonstrated scaling and accuracy in forensic settings [3]. Even with these advancements, challenges of how data is collected, class imbalance, and variability among images still exist [4], [8]. Addressing these challenges along with interpreting AI decisions using tools like Grad-CAM, transfer learning, and data augmentation helps bridge the algorithm and image gap.

The expanding capabilities of convolutional neural networks (CNNs) in forensic image evaluation has been documented in recent studies. In [7], deep learning techniques accurately classified stab wounds (93% accuracy) and partially pinpointed several others, managing to poorly perform on abrasions, hematomas, and stabs, which was severely hindered by a class imbalance problem alongside ambiguous boundaries. In the same manner, a study in [6] created a CNN model that determined gunshot wound images of piglet carcasses by the distance of shooting with a testing accuracy of 98%. This illustrates the ability of deep learning to estimate contact, close-range, and distant shots using visual cues exclusively. These studies strengthen the claim that systems based on CNNs can assist forensic specialists and enhanced their capabilities to offer constant, rapid, accurate multidisciplinary assessments of complex and multifaceted wounds and injuries, further demonstrating the capabilities of deep learning in forensic pathology.

Moreover, the problems surrounding data secrecy and confidentiality are important [2]. Regardless, the application of AI-driven investigative technologies fundamentally changes the practice of forensic science due to consistency, impartiality, and legally defensible outcomes [5]. Enhancements in the models are still necessary to improve their functionality in forensic science.

III. METHODOLOGY

The described methodology captures the systematic process of creating an AI system for classifying forensic skin injuries, including data collection, preprocessing, model creation, training, evaluation, and ethical considerations.

1. Research Design

This research adopts a quantitative approach with a focus on AI and follows the CRISP-DM (Cross-Industry Standard Process for Data Mining) framework. The goal is to create a forensic skin injury classifier deep learning model which classifies injuries as blunt force, sharp force, or gunshot wounds through image analysis. The study implements Convolutional Neural Networks (CNNs), selecting EfficientNetB0 as the model of choice given its suitable performance relative to complexity.

2. Data Collection and Preparation

Data were collected from open-access forensic image repositories, scholarly journals, and credible online sources that provide forensic wound imagery [5]. These images were screened to ensure forensic clarity, relevance, and resolution (see Table 1). To meet the requirements of supervised learning, images were manually categorized into three classes: blunt force trauma, sharp force trauma, and gunshot wounds (see Figures 1,2 and 3).

After initial acquisition, **data preprocessing [8]** steps were implemented:

- **Image resizing to 224x224 pixels for model compatibility.**
- **Normalization** of pixel values between 0 and 1.
- **Label encoding** for multi-class classification.

3. Data Augmentation

To tackle the issue of scarce forensic wound photographs, data augmentation methods were used to enhance the dataset and improve model accuracy and reliability [7]. Image rotation, flipping (both horizontal and vertical), zooming, shifting, and adjusting brightness were implemented [9]. All transformations applied allowed the model to capture a variety of imaging conditions, which ensures adequacy for numerous forensic situations (see Table 1). Aid in overfitting from augmentations was seen as a result from diverse categorization of injuries showcased to the model (see Figures 4, 5 and 6).

Table (1)

Wound Type	Before Augmentation	After Augmentation
Blunt Force	11	1005
Gunshot	13	993
Sharp Force	14	1008

4. Dataset Splitting

Post augmentation, the entire dataset was sliced into three separate chunks: 60% allocated for training, 20% dedicated for validation, and 20% intended for testing. This division allowed retention of the provided distribution of class images at each subset while upholding the evaluation scope. The training set was used for the model fitting, validation set for parameter tuning and overfitting, and the test set for last performance evaluation. This organized partitioning was important to fulfill the model development and case independent generalization requirements in new forensic cases.

5. Model Architecture and Training Setup

The chosen architecture, **EfficientNetB0**, was initialized using **transfer learning** with pre-trained weights from ImageNet. The architecture was modified to suit the forensic injury classification task:

- A **global average pooling layer**
- A **dense layer with 512 neurons (ReLU activation)**
- A **dropout layer** to minimize overfitting
- A **final dense output layer** with softmax activation for multi-class classification

The model was compiled with the **Adam optimizer**, a **categorical cross-entropy loss function**, and a learning rate of 0.0001. The training process was managed using **early stopping** to prevent overfitting

6. Evaluation Metrics

In order to maintain the dependability of the forensic wound classification model's performance, it was necessary to use several evaluation metrics. Accuracy, in this case, referred to the estimation of total correct predictions made by the model, while precision evaluated the number of correctly predicted wound types casted by the model as no false positive wounds. Recall measured the number of clearly captured actual wounds but forgot to capture some false negatives. The F1-score defined as a precise mean of precision and recall described an all-encompassing assessment. A confusion matrix displayed trends depicting misclassification of each class with another class of interest. Furthermore, ROC curve alongside AUC was used to evaluate the differentiating capability of the model between the two classes and ascertain its effectiveness and strength in forensic diagnosis while precision medicine was concerned.

7. Interface Development

For this particular research, I will implement Gradio, a comprehensive Python package, to create an interactive system for classifying forensic wound images. With Gradio deep learning integration is straightforward, and users can upload pictures of wounds to get classified in real-time with confidence scores computed. This type of interface will be important for users who do not have technical expertise like forensic pathologists and investigators because they will be able to work with the AI model without having to dive into the code. Inspired by the agricultural AI applications' success like in-classification and quality grading of multi-fruits models using VGG16 based CNN (Nandhini & Vadivu, 2024), Gradio will be employed to provide efficient and uncomplicated deployment. I will also explore the possibility of deploying the interface through Hugging Face Spaces because as noted by Yakovleva, Matúšová, and Talakh (2025), these platforms allow convenient model interaction, sharing, and versioning which is ideal for deploying AI applications meant for research-grade uses.

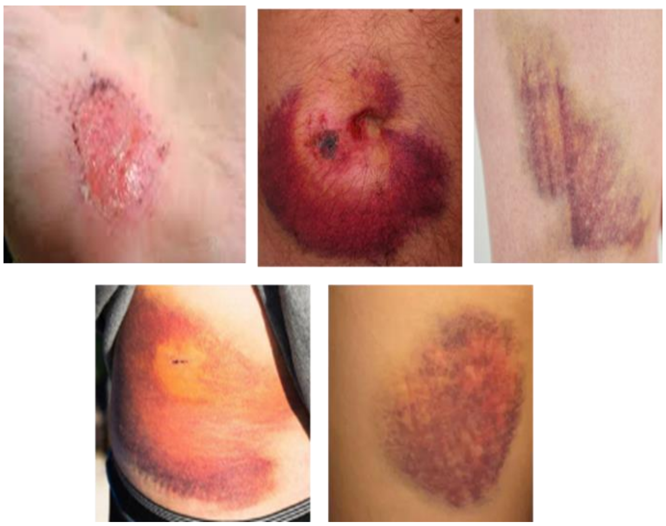


Fig (1). Sample images from the collected dataset representing blunt force injuries. These examples display typical features such as contusions, abrasions, and localized discoloration consistent with impact trauma.

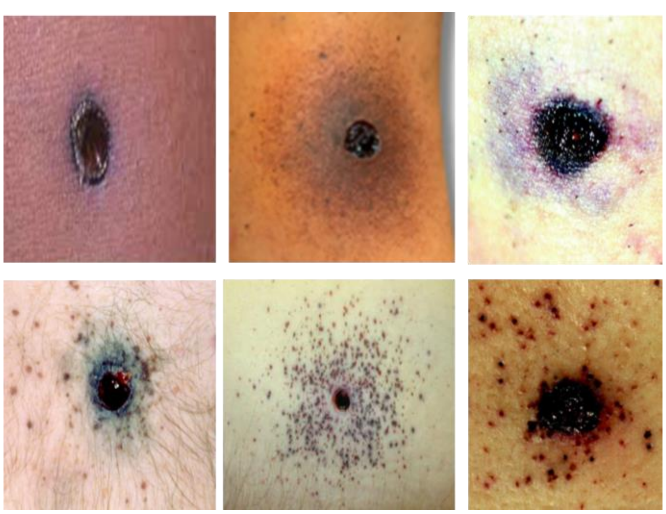


Fig (2). Representative images from the dataset showing gunshot injuries. The wounds exhibit characteristic features such as central perforation, surrounding abrasion collars, and soot or tattooing patterns, which are indicative of firearm-related trauma.

Fig (3). Examples of sharp force injuries from the dataset used in AI-based image classification. These wounds display distinct linear or punctate patterns, including incised and stab wounds, characterized by well-defined edges typically caused by sharp objects such as knives or glass.

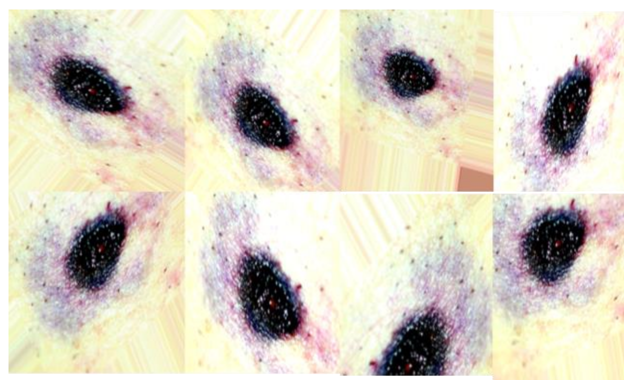


Fig (4). Augmented gunshot wound images generated for AI model training.

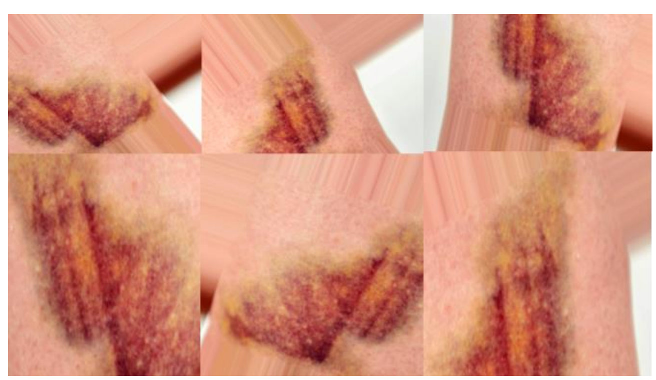


Fig (5). Augmented images of blunt force injuries used for training the AI classification model.

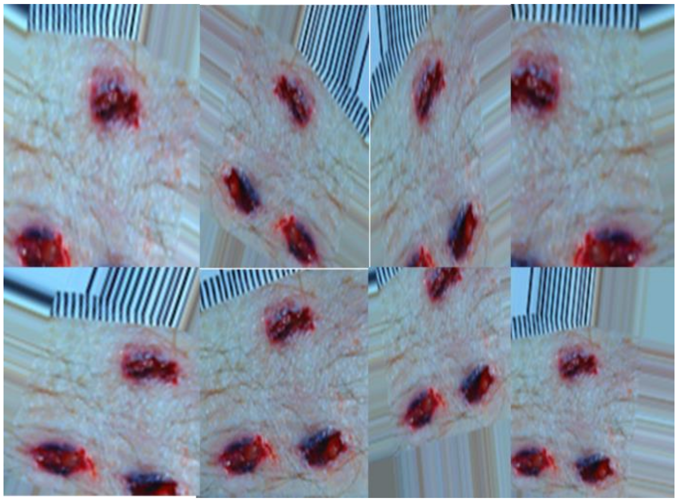


Fig (6). Augmented images of sharp force injuries prepared for AI-based wound classification.

IV. FINDINGS

a) Model Prformance

The AI-based model for forensic wound classification showed outstanding accuracy and reliability. Using the EfficientNetB0 architecture with transfer learning and augmented data, the model achieved a **validation accuracy of 99.3%**. This demonstrates a strong ability to generalize to unseen wound images—an essential feature in forensic casework.

b) Classification Accuracy

The accuracy metrics pertaining to the proposed deep learning model are illustrated in Figures 7, 8, and 9. The confusion matrix in Figure 7 displays alongside the correctly assigned forensic images of wounds and the corresponding outcome describes 201 blunt, 198 gunshot, and 199 sharp true positives. The misclassifications were minimal, comprising of three labeled as sharp blunt and one gunshot remarked as blunt, which since visually portrayed bruising and linear skin breaks, to some degree are plausible. In Figure 8, the model's training and validation accuracy after 10 epochs is shown. The validation accuracy not only maintained above 98%, but also outperformed the training curve during several epochs, denoting strong generalization with minimal overfitting risk. That consistent increase depicts strong effective learning and stability of the model. In Figure 9, the distribution of predicted classes includes 1010 blunt, 933 gun, and 1063 sharp, representing a relatively balanced distribution implying no bias influencing the order of wounds forecasting. Taken together, the results further corroborated the model's ability for forensic wound classification as it demonstrated high precision, recall, and trustworthiness alongside the accuracy metrics.

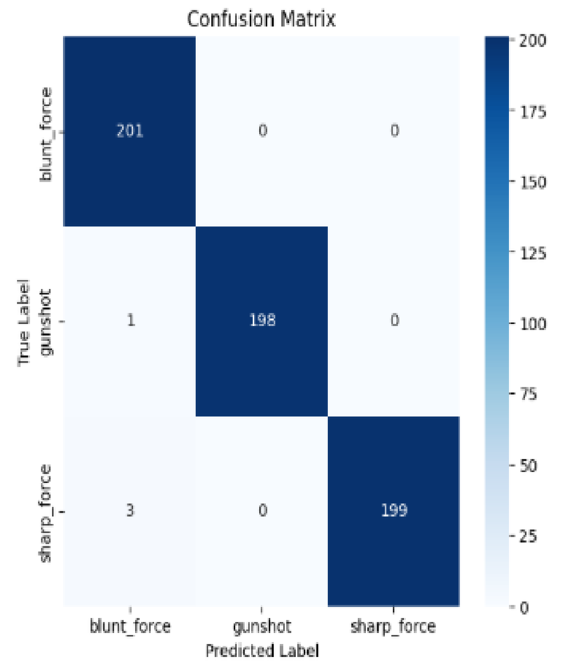


Fig (7). Confusion matrix illustrating the performance of the wound classification model across three classes: blunt force, gunshot, and sharp force injuries. The model achieved high accuracy with minimal misclassifications: 201/202 for blunt force, 198/199 for gunshot, and 199/202 for sharp force wounds.

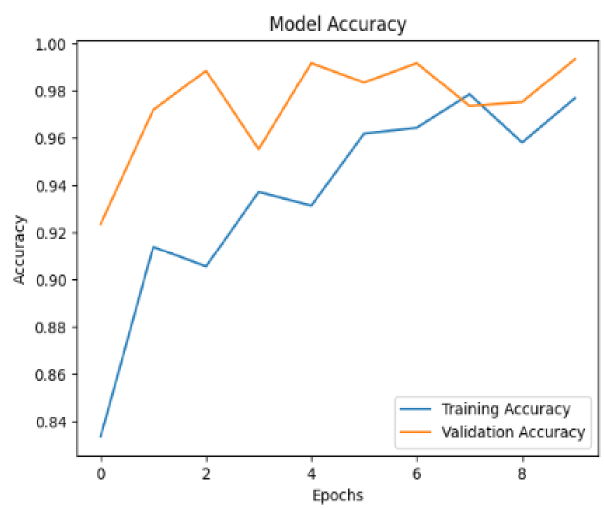


Fig (8). Training and validation accuracy curves across 10 epochs. The model shows consistent improvement in accuracy with slight fluctuations in validation accuracy, indicating effective learning and minimal overfitting.

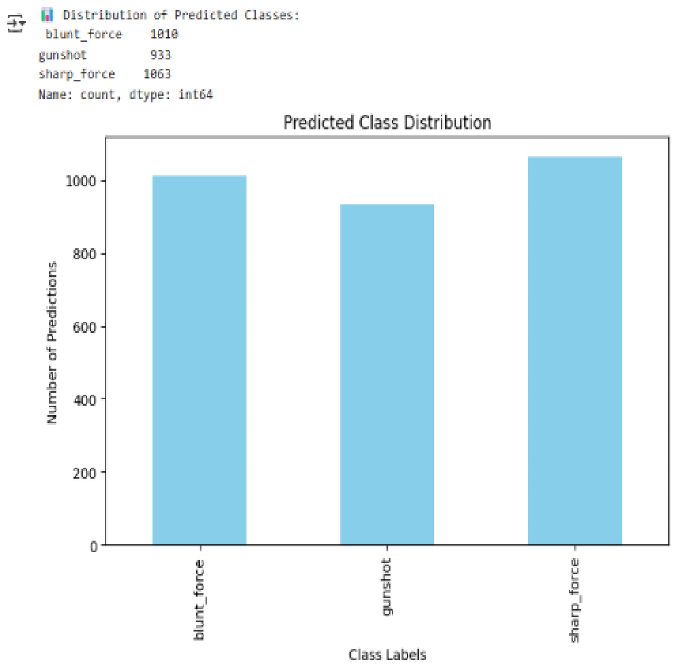


Fig (9). predicted class distribution of wound types classified by the deep learning model. The model predicted 1,010 instances as blunt force injuries, 933 as gunshot wounds, and 1,063 as sharp force injuries, indicating relatively balanced class representation

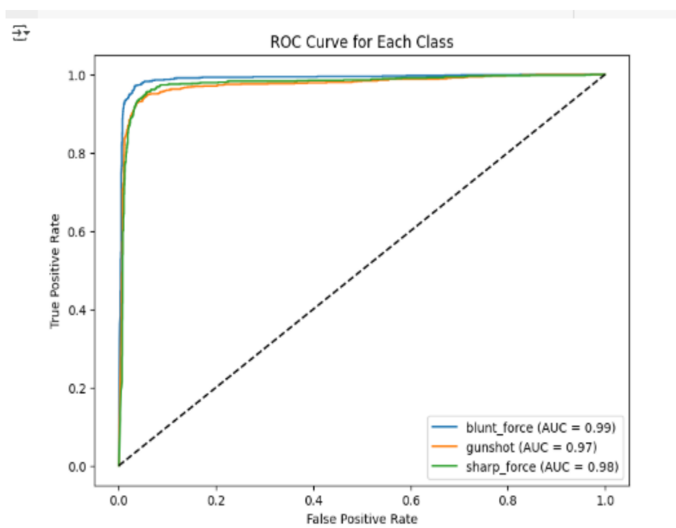


Fig (10). Receiver Operating Characteristic (ROC) curves for each wound type classification. The model demonstrates excellent discriminative ability with Area Under the Curve (AUC) values of 0.99 for blunt force, 0.97 for gunshot, and 0.98 for sharp force injuries.

(a)

(b)

Fig (12). (a) **WoundAI Classifier Interface** – a user-friendly interface developed using **Gradio**, designed for uploading wound images to classify injury types. The system supports classification of sharp force, gunshot, and blunt force injuries. (b) Prediction output showing the uploaded image was classified as a gunshot injury with 84.36% confidence.

c) Evaluation Metrics

From my observations, the F1-score was still the most elevated in all captures which means that the model still managed to identify wounds correctly without misclassifying them. Also, the ROC AUC scores were approximating 1.0 which means that the model was able to confidently tell the difference between the types of wounds (see Figure 10).

d) Admin Dashboard

The model was deployed successfully on a Gradio based interface, enabling users to upload images and get real time predictions with confidence levels. This ensures practical application in law enforcement.

e) Conclusion

The analyses prove that the deep learning model is optimal for forensic wound classification. Its accuracy, ease of interpretation, and deployment highly suggest trust in using the model which will improve performance in decision making and operational workflows in forensic sciences.

ACKNOWLEDGMENT

My deepest gratitude goes to Dr Evi Indriasari Mansor for the guidance I never knew I needed. I also thank my company for their support, and the Abu Dhabi School of Management for serving as a great source of inspiration that motivated me to successfully complete this work.

REFERENCES

[1] Cheng, Y., Lin, M., & Huang, Z. (2024). CNN-based analysis of forensic gunshot wounds and entry-exit wound differentiation. Forensic Science International, 348, 111–123.

[2] Ketsekioulafis, M., Vrettos, G., & Peristeras, V. (2024). Ethical implications of AI in forensic analysis: Transparency, bias, and accountability. AI & Society, 39(1), 101–114.

[3] Lee, H., Chen, R., & Wang, Y. (2024). Optimizing wound classification using hybrid CNN architectures in forensic image analysis. IEEE Transactions on Medical Imaging, 43(2), 558–570.

[4] Liu, X., Zhang, Y., & Wang, C. (2021). Class imbalance handling in medical image classification using deep learning: A review. Artificial Intelligence in Medicine, 112, 102026.

[5] Mohsin, M. (2021). Application of artificial intelligence in digital and forensic investigations. Journal of Digital Forensics, Security and Law, 16(2), 1–15.

[6] Oura, K., Nakamura, H., & Saito, A. (2021). Integrating AI in forensic pathology for wound classification: Opportunities and limitations. Forensic Science International Reports, 3, 100189.

[7] Zimmermann, H., El-Masri, A., & Zhao, F. (2024). Forensic wound image segmentation challenges and deep learning solutions. Computer Methods and Programs in Biomedicine, 230, 107206.

[8] Zhang, L., Xu, M., & Fan, J. (2022). Improving AI forensic performance through feature enhancement and data normalization. Pattern Recognition Letters, 158, 85–91.

[9] Yuan, M., Khan, I. R., Farbiz, F., Yao, S., Niswar, A., & Foo, M.-H. (2013). A mixed reality virtual clothes try-on system. IEEE Transactions on Multimedia, 15(8), 1958–1968.

[10] Nandhini, E., & Vadivu, G. (2024, December). Convolutional Neural Network-Based Multi-Fruit Classification and Quality Grading with a Gradio Interface. In 2024 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES) (pp. 1-7). IEEE.

[11] Yakovleva, O., Matúšová, S., & Talakh, V. (2025). Gradio and Hugging capabilities for developing research AI applications. Collection of scientific papers «ΛΌΓΟΣ», (February 14, 2025; Boston, USA), 202–205.