Journal of Data Analysis and Critical Management, Volume 01, Issue 02, 2025

Large Language Models and Multimodal AI

Gimah Mathew*

Ladoke Akintola University of Technology

Abstract

Large Language Models (LLMs) have transformed natural language processing by demonstrating remarkable capabilities in text understanding, generation, and reasoning. Recent advances extend these models into multimodal AI, enabling the integration of multiple data modalities—such as text, images, audio, and video—into unified learning frameworks. Multimodal AI systems leverage LLMs to process and correlate information across modalities, enhancing context understanding, task flexibility, and human–computer interaction. These models find applications in image captioning, visual question answering, video summarization, conversational AI, and cross-modal retrieval. Despite their promise, challenges such as high computational requirements, alignment of heterogeneous modalities, interpretability, and ethical concerns remain. This paper explores the architecture, capabilities, applications, and limitations of LLMs and multimodal AI, highlighting their potential to enable more robust, context-aware, and interactive artificial intelligence systems.

Keywords: Large Language Models, Multimodal AI, Cross-Modal Learning, Transformers, Context-Aware AI, Image-Text Models, Conversational AI, Visual Question Answering, AI Alignment, Ethical AI.

Journal of Data Analysis and Critical Management (2025); DOI: 10.64235/dgd80j16

Introduction

Definition of Artificial Intelligence (AI) and Machine Learning (ML)

Artificial Intelligence (AI) refers to the development of computer systems capable of performing tasks that typically require human intelligence, such as reasoning, learning, perception, language understanding, and decision-making. Machine Learning (ML), a subset of AI, focuses on algorithms that enable systems to learn patterns from data and improve performance over time without being explicitly programmed. Modern AI systems heavily rely on ML techniques, particularly deep learning, to achieve high levels of accuracy and adaptability.

Overview of Large Language Models (LLMs)

Large Language Models (LLMs) are advanced deep learning models trained on massive text datasets to understand, generate, and reason about natural language. Built primarily on transformer architectures, LLMs contain billions (or even trillions) of parameters that capture linguistic patterns, semantic relationships, and contextual dependencies.

LLMs power applications such as:

- Text generation and summarization
- Machine translation
- Conversational agents and chatbots
- Code generation and reasoning tasks

Their ability to perform diverse tasks with minimal task-specific training has significantly advanced natural language processing (NLP).

Corresponding Author: Gimah Mathew, Ladoke Akintola University of Technology, e-mail: mathew@student.lautech.edu.ng

How to cite this article: Mathew, G. (2025). Large Language Models and Multimodal AI. Journal of Data Analysis and Critical Management, 01(2):108-115.

Source of support: Nil

Conflict of interest: None

Introduction to Multimodal AI

Multimodal AI refers to systems capable of processing and integrating information from multiple data modalities, including text, images, audio, video, and sensor data. Unlike unimodal systems that operate on a single type of input, multimodal models learn cross-modal representations that enable richer understanding and interaction.

Examples include:

- Image captioning (linking visual and textual information)
- Visual question answering

- Speech-to-text and audio-visual recognition
- Video analysis with textual explanations

Multimodal AI enhances contextual awareness by combining complementary information from different sources.

Importance of LLMs and Multimodal Systems in Modern AI Applications

LLMs and multimodal systems have become central to modern AI due to their:

- **Scalability:** Ability to generalize across tasks.
- **Context-awareness:** Deep understanding of complex inputs.
- **Human-like interaction:** Natural communication via text, speech, and visual understanding.
- **Cross-domain applicability:** Use in healthcare, education, entertainment, robotics, and business analytics.

These systems enable more intuitive and intelligent human–computer interaction, moving AI closer to general-purpose capabilities.

Objective

The objective of this study is to explore the capabilities, architectural foundations, challenges, and societal impact of Large Language Models and multimodal AI systems. By understanding their strengths and limitations, we can better evaluate their role in shaping the future of intelligent, interactive, and context-aware AI technologies.

Large Language Models (LLMs)

Definition and Characteristics

Large Language Models (LLMs) are advanced AI systems trained on massive text corpora to learn the structure, meaning, and patterns of human language. By leveraging deep neural networks with billions of parameters, LLMs can understand context, generate coherent text, and perform complex reasoning tasks.

Key characteristics include:

- **Scale:** Large parameter sizes and training on vast datasets.
- **Pre-training:** Learning general language representations from unlabeled data.
- **Fine-tuning:** Adapting pre-trained models to specific tasks using labeled data or instruction-based learning.

These features allow LLMs to perform multiple tasks without being explicitly programmed for each one.

Architecture and Techniques

- **Transformer Architecture:** LLMs are primarily based on the transformer architecture, which enables efficient parallel processing of sequential data.
- **Self-Attention Mechanism:** Self-attention allows the model to weigh the importance of different words in a sentence relative to each other, improving contextual understanding and long-range dependency modeling.

Pre-Training and Fine-Tuning

- **Pre-training (Unsupervised/Self-Supervised Learning):** The model learns general language patterns using large-scale text datasets.
- **Fine-Tuning (Supervised/Transfer Learning):** The model is adapted for specific downstream tasks such as sentiment analysis, translation, or question answering.

These techniques make LLMs highly flexible and adaptable across domains.

Examples of LLMs

- **GPT series (OpenAI):** A family of generative models designed for text completion, reasoning, and conversational AI.
- **BERT:** A bidirectional encoder model designed for deep contextual understanding of language.
- **RoBERTa:** An optimized variant of BERT with improved training strategies.
- **LLaMA:** A family of efficient open-weight LLMs designed for research and scalable deployment.
- **T5:** A model that frames all NLP tasks in a unified text-to-text format.
- **PaLM:** A large-scale language model designed for advanced reasoning and multilingual tasks.

Capabilities

LLMs demonstrate a wide range of natural language capabilities, including:

- **Natural Language Understanding (NLU):** Interpreting meaning, sentiment, intent, and context in text.
- **Natural Language Generation (NLG):** Producing coherent and contextually relevant text responses.
- **Text Summarization:** Condensing large documents into concise summaries.
- **Machine Translation:** Converting text between languages.
- **Question Answering:** Providing accurate responses based on provided context or general knowledge.

Multimodal AI

Definition and Scope

Multimodal AI refers to artificial intelligence systems that can process, understand, and integrate multiple types of data simultaneously. Unlike unimodal models that operate on a single data type (e.g., text only), multimodal systems combine and correlate information across different modalities such as:

- Text
- Images
- Audio
- Video
- Sensor data (e.g., IoT, robotics inputs)

By learning joint representations across modalities, these systems can interpret complex real-world information more effectively, mimicking how humans combine sight, sound, and language for understanding.

Motivation and Benefits

- **Richer Understanding of Context:** Integrating multiple data sources improves contextual awareness. For example, combining an image with descriptive text enables deeper scene understanding than either modality alone.
- **Enhanced Human–AI Interaction:** Multimodal AI supports natural communication methods such as voice commands paired with visual input, improving accessibility and usability.
- **Beyond Single-Modality AI:** Many real-world tasks inherently require multiple modalities—such as video captioning, medical diagnosis using scans and reports, or autonomous driving using cameras and sensors. Multimodal AI enables these advanced applications by bridging information across domains.

Examples of Multimodal Models

- **CLIP:** Learns joint representations of images and text by aligning visual and textual embeddings, enabling zero-shot image classification and cross-modal retrieval.
- **DALL·E:** Generates images from textual descriptions using transformer-based generative modeling.
- **Imagen:** A high-fidelity text-to-image generation model that combines large language understanding with diffusion-based image synthesis.
- **Flamingo:** Designed for few-shot learning across vision and language tasks.
- **Gato:** A unified multimodal agent capable of handling text, vision, robotics control, and other tasks within a single model framework.

Applications of LLMs and Multimodal AI

Large Language Models (LLMs) and multimodal AI systems are transforming industries by enabling context-aware, interactive, and cross-modal intelligence. Their applications span language processing, vision, healthcare, education, and accessibility.

Natural Language Processing (NLP)

- **Conversational AI (Chatbots and Virtual Assistants):** LLMs power advanced conversational systems capable of understanding context, maintaining dialogue, and generating human-like responses. These systems are widely used in customer support, enterprise solutions, and personal productivity tools.
- **Content Generation and Summarization:** LLMs assist in drafting articles, generating reports, summarizing long documents, and creating creative content such as stories or marketing copy.
- **Question-Answering and Information Retrieval:** By understanding natural language queries, LLMs can retrieve relevant information, provide explanations, and synthesize knowledge from large datasets.

Computer Vision and Cross-Modal Applications

- **Image Captioning and Understanding:** Multimodal models generate descriptive captions for images and analyze visual content in context.
- **Text-to-Image and Image-to-Text Generation:** Systems like DALL·E and CLIP enable cross-modal generation and alignment between text and visual representations.
- **Video Understanding with Textual Context:** Multimodal AI can analyze video content and provide summaries, detect events, or answer questions based on both visual and textual cues.

Healthcare

1. **Analyzing Medical Records and Imaging Data:** LLMs can process clinical notes, while multimodal models combine text with medical images (e.g., X-rays, MRIs) for comprehensive analysis.
2. **Multimodal Diagnostics:** Integrating textual patient history, imaging scans, and sensor data (e.g., wearable devices) enhances diagnostic accuracy and decision support systems.

Education and Accessibility

- **Personalized Tutoring Systems:** LLMs enable adaptive learning platforms that tailor explanations and exercises to individual student needs.

- **AI-Assisted Learning with Multimodal Content:** Combining text, visuals, and audio enhances comprehension and engagement in digital learning environments.

Accessibility Tools for Differently-Abled Users

- Text-to-speech systems for visually impaired users
- Image description tools powered by multimodal AI
- Real-time captioning and speech recognition for hearing-impaired users

Challenges and Limitations

Despite their impressive capabilities, Large Language Models (LLMs) and multimodal AI systems face significant technical, ethical, and practical challenges. Addressing these limitations is essential for responsible and reliable deployment.

Data and Computational Requirements

- **Need for Massive Datasets:** LLMs and multimodal models require enormous amounts of high-quality text, image, audio, and video data for effective training. Collecting, cleaning, and maintaining such datasets is resource-intensive and may raise privacy concerns.
- **High Training and Inference Costs:** Training large models demands substantial computational power, specialized hardware (e.g., GPUs/TPUs), and significant energy consumption. Even after training, large-scale deployment for inference can incur high operational costs.

These requirements limit accessibility and raise concerns about environmental sustainability.

Bias and Ethical Concerns

- **Propagation of Societal Biases:** Models trained on large-scale internet data may inherit and amplify biases related to gender, race, culture, or socioeconomic status.
- **Misinformation and Hallucinations:** LLMs can generate plausible but factually incorrect or fabricated information (often called "hallucinations"). In multimodal systems, incorrect interpretations of visual or audio inputs can further compound misinformation risks.

These issues highlight the need for robust bias mitigation, fact-checking mechanisms, and ethical AI governance.

Multimodal Integration Challenges

- **Aligning Information Across Modalities:** Combining text, images, audio, and other data types requires effective cross-modal alignment. Ensuring that representations correspond accurately across modalities is technically complex.
- **Handling Noisy or Incomplete Data:** Real-world multimodal data may be inconsistent, missing, or noisy (e.g., blurred images, distorted audio, incomplete text), which can degrade model performance.

Robust multimodal fusion techniques are necessary to manage these challenges.

Interpretability and Explainability

- **Black-Box Nature of Large Models:** LLMs and multimodal systems often function as complex, opaque neural networks, making it difficult to understand how decisions are made.
- **Explaining Cross-Modal Decisions:** Providing transparent explanations becomes more challenging when outputs depend on interactions between multiple data modalities.

Improving explainability is critical for building trust, ensuring accountability, and enabling safe deployment in high-stakes domains such as healthcare and law.

Recent Advances and Techniques

Rapid progress in Large Language Models (LLMs) and multimodal AI has been driven by innovations in training strategies, generalization methods, and cross-modal architectures. These advances have significantly improved adaptability, performance, and real-world applicability.

Fine-Tuning and Instruction Tuning

- **Fine-Tuning:** After large-scale pre-training, models are adapted to specific tasks using labeled datasets. This improves performance in domains such as legal analysis, medical text processing, or customer service automation.
- **Instruction Tuning:** Models are trained on datasets formatted as instructions and responses, improving their ability to follow natural language commands and perform diverse tasks.
  - Enhances usability in conversational AI systems.
  - Improves alignment with user intent and expectations.

These techniques allow LLMs to exhibit more controlled and task-specific behavior.

Few-Shot and Zero-Shot Learning

- **Zero-Shot Learning:** The model performs tasks without explicit task-specific training by leveraging general knowledge learned during pre-training.

- **Few-Shot Learning:** The model adapts to new tasks using only a small number of examples provided in the prompt.

These capabilities reduce the need for extensive labeled datasets and demonstrate the strong generalization ability of modern LLMs.

Cross-Modal Contrastive Learning

- Contrastive learning aligns representations from different modalities by maximizing similarity between related pairs (e.g., image and caption) and minimizing similarity between unrelated pairs.
- A prominent example is **CLIP**, which learns joint embeddings for text and images.
- Enables zero-shot image classification, cross-modal retrieval, and improved multimodal reasoning.

Multimodal Transformers and Attention-Based Fusion

- **Multimodal Transformers:** Extend the transformer architecture to handle multiple data types simultaneously by learning shared or coordinated representations.
- **Attention-Based Fusion:** Attention mechanisms dynamically weigh contributions from different modalities (e.g., text vs. image) based on relevance to the task.
- These architectures improve cross-modal reasoning and contextual understanding, enabling tasks such as visual question answering, video summarization, and multimodal dialogue systems.

Future Directions

As Large Language Models (LLMs) and multimodal AI systems continue to evolve, research is moving toward more unified, efficient, and responsible intelligent systems. The following directions highlight key areas shaping the future of this field.

Unified Multimodal AI Agents Capable of Reasoning and Planning

- Development of generalist AI agents that can seamlessly process text, images, audio, video, and sensor data within a single architecture.
- Integration of reasoning, long-term memory, and planning capabilities to support complex decision-making tasks.
- Movement toward models similar to Gato, which demonstrate the feasibility of unified architectures handling diverse tasks.
- These agents may support advanced applications such as interactive robotics, intelligent assistants, and real-time problem solving.

Improved Efficiency and Sustainability of Large Models

- Reducing computational and energy demands through model compression, sparsity, and efficient training strategies.
- Development of lightweight multimodal models suitable for deployment on edge devices.
- Incorporating sustainability principles to balance performance with environmental impact, ensuring scalable AI growth.

Integration with Robotics and Autonomous Systems

- Combining multimodal AI with robotics to enable perception-driven action and adaptive control.
- Enhancing autonomous vehicles and drones by integrating vision, language understanding, and sensor data for better situational awareness.
- Enabling embodied AI systems capable of interacting intelligently with physical environments.

Ethical Frameworks and Regulatory Considerations

- Establishing guidelines to ensure safe, fair, and transparent deployment of LLMs and multimodal systems.
- Addressing risks such as bias, misinformation, privacy concerns, and misuse.
- Encouraging collaboration between researchers, industry, and policymakers to develop **robust governance frameworks** for advanced AI technologies.

Conclusion

Large Language Models (LLMs) and multimodal AI systems represent a significant leap forward in artificial intelligence. By enabling advanced language understanding, text generation, reasoning, and cross-modal integration, these technologies have expanded the boundaries of what AI systems can achieve. Multimodal models further enhance this capability by combining text, images, audio, video, and sensor data, allowing for richer contextual awareness and more natural human–AI interaction.

The impact of these systems is already visible across industries. In healthcare, they support multimodal diagnostics and clinical decision assistance. In education, they power personalized tutoring and interactive learning tools. In the creative domain, they enable content generation, design assistance, and multimedia production. Their versatility positions them as transformative technologies across business, science, and society.

However, realizing their full potential requires addressing critical challenges, including ethical concerns, bias mitigation, misinformation risks, interpretability, and sustainability. Responsible development, transparent reporting, and regulatory oversight are essential to ensure safe and equitable deployment.

In conclusion, LLMs and multimodal AI stand as the next frontier in general-purpose intelligent systems, offering unprecedented adaptability and capability. With continued innovation and responsible governance, they have the potential to shape a future where AI systems are more powerful, inclusive, and aligned with human values.

References

Omonua, Lucky Alfred, Ewomazino Daniel Akpor, and Wilfred Oristesan Olley. "Mass media coverage of women and gender inequality." African Journal of Social Sciences and Humanities Research 6, no. 4 (2023): 150-165.

Orhewere, John Agbavbiose, and Wilfred Oritsesan Olley. "The construction of climate change discourse in online news comment sections: A critical discourse analysis." African Journal of Social Sciences and Humanities Research 6, no. 4 (2023): 75-84.

Jabed, M. M. I., Gupta, A. B., Ferdous, J., Islam, M., & Akter, S. (2022). Self-Supervised Learning for Efficient and Scalable AI: Towards Reducing Data Dependency in Deep Learning Models. International Journal of Intelligent Systems and Applications in Engineering, 10(3s), 317–.

Santos, C. (2022). Self-supervised representation learning: Investigating self-supervised learning methods for learning representations from unlabeled data efficiently. Journal of AI-Assisted Scientific Discovery, 2(1).

Routhu, K. K. (2018). Reusable Integration Frameworks in Oracle HCM: Accelerating Enterprise Automation through Standardized Architecture. International Journal of Scientific Research & Engineering Trends, 4(4).

Cao, Y.-H., Sun, P., Huang, Y., Wu, J., & Zhou, S. (2022). Synergistic self-supervised and quantization learning. ArXiv Preprint.

Miller, J. D., Arasu, V. A., Pu, A. X., Margolies, L. R., Sieh, W., & Shen, L. (2022). Self-supervised deep learning to enhance breast cancer detection on screening mammography. ArXiv Preprint.

Routhu, K. K. (2019). Hybrid machine learning architecture for absence forecasting within Oracle Cloud HCM. KOS Journal of AIML, Data Science, and Robotics, 1(1), 1-5.

Routhu, K. K. (2019). Conversational AI in Human Capital Management: Transforming Self-Service Experiences with Oracle Digital Assistant. International Journal of Scientific Research & Engineering Trends, 5(6).

Turrisi da Costa, V. G., Fini, E., Nabi, M., Sebe, N., & Ricci, E. (2022). solo-learn: A Library of Self-supervised Methods for Visual Representation Learning. Journal of Machine Learning Research, 23, 1–6.

Ozsoy, S., Hamdan, S., Arik, S. Ö., & Erdogan, A. T. (2022). Self-supervised learning with an information maximization criterion. In Advances in Neural Information Processing Systems.

Haresamudram, H., Essa, I., & Plötz, T. (2022). Assessing the state of self-supervised human activity recognition using wearables. ArXiv Preprint.

Barbalau, A., Ionescu, R. T., Georgescu, M.-I., et al. (2022). SSMTL++: Revisiting self-supervised multi-task learning for video anomaly detection. ArXiv Preprint.

Lemkhenter, A., & Favaro, P. (2022). Towards sleep scoring generalization through self-supervised meta-learning. ArXiv Preprint.

Zhang, C. (2022). A survey on masked autoencoder for self-supervised learning. ArXiv Preprint.

Kranthi Kumar Routhu. (2020). Intelligent Remote Workforce Management: AI, Integration, and Security Strategies Using Oracle HCM Cloud. KOS Journal of AIML, Data Science, and Robotics, 1(1), 1–5. https://doi.org/10.5281/zenodo.17531257

Routhu, K. K. (2020). Strategic Compensation Equity and Rewards Optimization: A Multi-cloud Analytics Blueprint with Oracle Analytics Cloud. Available at SSRN 5737266.

Olley, Wilfred Oritsesan, and Francisca Chinazor Alajemba. "Audience's perception of social media as tools for the creation of fashion awareness." The International Journal of African Language and Media Studies 2, no. 1 (2022): 141.

Wilfred, Olley Oritsesan, EWOMAZINO DANIEL AKPOR, and OBINNA JOHNKENNEDY CHUKWU. "APPLICATION OF AGENDA SETTING, MEDIA DEPENDENCY, AND USES AND GRATIFICATIONS THEORIES IN THE MANAGEMENT OF DISEASE OUTBREAK IN NIGERIA." Euromentor 12, no. 3 (2021).

Ate, Andrew Asan, Ewomazino Daniel Akpor, Wilfred Oritsesan, Sadiq Oshoke Akhor, Edike Kparoboh Frederick, Joseph Omoh Ikerodah, Abdulazeez Hassan Kadiri et al. "Communication and governance for cultural development: Issues and platforms." Corporate & Business Strategy Review 3, no. 2 (2022): 151-158.

Routhu, K. K. (2019). AI-Enhanced Payroll Optimization: Improving Accuracy and Compliance in Oracle HCM. KOS Journal of AIML, Data Science, and Robotics, 1(1), 1-5.

Olley, Wilfred Oritsesan, Ewomazino Daniel Akpor, Dike Harcourt-Whyte, Samson Ighiegba Omosotomhe, Afam Patrick Anikwe, Edike Kparoboh Frederick, Evwiekpamare Fidelis Olori, and Paul Edeghoghon Umolu. “Electoral violence and voter apathy: Peace journalism and good governance in perspective.” Corporate Governance and Organizational Behavior Review 6, no. 3 (2022): 112-119.

Olley, Wilfred Oritsesan, and Francisca Chinazor Alajemba. “Audience’s perception of social media as tools for the creation of fashion awareness.” The International Journal of African Language and Media Studies 2, no. 1 (2022): 141.

Abdulazeez, Isah, Wilfred O. Olley, and PhD2&Abdulazeez H. Kadiri. “CHAPTER THIRTY ONE SELF-AFFIRMATIVE DISCOURSE ON SOCIAL JUDGEMENT THEORY AND POLITICAL ADVERTISING.” Discourses on Communication and Media Studies in Contemporary Society (2022): 258.

Polu, A. R., Buddula, D. V. K. R., Narra, B., Gupta, A., Vattikonda, N., & Patchipulusu, H. (2021). Evolution of AI in Software Development and Cybersecurity: Unifying Automation, Innovation, and Protection in the Digital Age. Available at SSRN 5266517.

Bitkuri, V., Kendyala, R., Kurma, J., Mamidala, V., Enokkaren, S. J., & Attipalli, A. (2021). Systematic Review of Artificial Intelligence Techniques for Enhancing Financial Reporting and Regulatory Compliance. International Journal of Emerging Trends in Computer Science and Information Technology, 2(4), 73-80.

Attipalli, A., Enokkaren, S., BITKURI, V., Kendyala, R., KURMA, J., & Mamidala, J. V. (2021). Enhancing Cloud Infrastructure Security Through AI-Powered Big Data Anomaly Detection. Available at SSRN 5741305.

Singh, A. A. S., Tamilmani, V., Maniar, V., Kothamaram, R. R., Rajendran, D., & Namburi, V. D. (2021). Predictive Modeling for Classification of SMS Spam Using NLP and ML Techniques. International Journal of Artificial Intelligence, Data Science, and Machine Learning, 2(4), 60-69.

Kothamaram, R. R., Rajendran, D., Namburi, V. D., Singh, A. A. S., Tamilmani, V., & Maniar, V. (2021). A Survey of Adoption Challenges and Barriers in Implementing Digital Payroll Management Systems in Across Organizations. International Journal of Emerging Research in Engineering and Technology, 2(2), 64-72.

Rajendran, D., Namburi, V. D., Singh, A. A. S., Tamilmani, V., Maniar, V., & Kothamaram, R. R. (2021). Anomaly Identification in IoT-Networks Using Artificial Intelligence-Based Data-Driven Techniques in Cloud Environmen. International Journal of Emerging Trends in Computer Science and Information Technology, 2(2), 83-91.

Attipalli, A., BITKURI, V., KURMA, J., Enokkaren, S., Kendyala, R., & Mamidala, J. V. (2021). A Survey of Artificial Intelligence Methods in Liquidity Risk Management: Challenges and Future Directions. Available at SSRN 5741342.

Routhu, K. K. (2021). AI-augmented benefits administration: A standards-driven automation framework with Oracle HCM Cloud. International Journal of Scientific Research and Engineering Trends, 7(3).

Routhu, K. K. (2021). Harnessing AI Dashboards in Oracle Cloud HCM: Advancing Predictive Workforce Intelligence and Managerial Agility. International Journal of Scientific Research & Engineering Trends, 7(6).

Gupta, A. K., Polu, A. R., Narra, B., Buddula, D. V. K. R., Patchipulusu, H. H. S., & Vattikonda, N. (2024). Leveraging deep learning models for intrusion detection systems for secure networks. *Journal of Computer Science and Technology Studies*, *6*(2), 199-208.

Narra, B., Buddula, D. V. K. R., Patchipulusu, H., Vattikonda, N., Gupta, A., & Polu, A. R. (2024). The integration of artificial intelligence in software development: Trends, tools, and future prospects. *Available at SSRN 5596472*.

Achuthananda, R. P., Bhumeka, N., Dheeraj Varun Kumar, R. B., Hari Hara, S. P., & Navya, V. (2024). Evaluating machine learning approaches for personalized movie recommendations: A comprehensive analysis. *J Contemp Edu Theo Artific Intel: JCETAI-115*.

Waditwar, P. (2024) The Intersection of Strategic Sourcing and Artificial Intelligence: A Paradigm Shift for Modern Organizations. Open Journal of Business and Management, 12, 4073-4085. doi: 10.4236/ojbm.2024.126204.

Bitkuri, V., Kendyala, R., Kurma, J., Mamidala, J. V., Attipalli, A., & Enokkaren, S. J. (2024). A Survey on Blockchain-Enabled ERP Systems for Secure Supply Chain Processes and Cloud Integration. *International Journal of Technology, Management and Humanities*, *10*(04), 126-135.

Mamidala, J. V., Bitkuri, V., Attipalli, A., Kendyala, R., Kurma, J., & Enokkaren, S. J. (2024). Machine Learning Approaches to Salary Prediction in Human Resource Payroll Systems. *Journal of Computer Science and Technology Studies*, *6*(5), 341-349.

Waditwar, P. (2024) AI for Bathsheba Syndrome: Ethical Implications and Preventative Strategies. Open Journal of Leadership, 13, 321-341. doi: 10.4236/ojl.2024.133020

Attipalli, A., Kendyala, R., Kurma, J., Mamidala, J. V., Bitkuri, V., & Enokkaren, S. J. (2024). Privacy Preservation in the Cloud: A Comprehensive Review of Encryption and Anonymization Methods. *International Journal of Multidisciplinary on Science and Management IJMSM*, *1*(1).

Tamilmani, V., Maniar, V., Singh, A. A., Kothamaram, R. R., Rajendran, D., & Namburi, V. D. (2024). A Review of Cyber Threat Detection in Software-Defined and Virtualized Networking Infrastructures. *International Journal of Technology, Management and Humanities*, *10*(04), 136-146.

Singh, A. A. S., Kothamaram, R. R., Rajendran, D., Deepak, V., Namburi, V. T., & Maniar, V. (2024). A Review on Model-Driven Development with a Focus on Microsoft PowerApps. *International Journal of Humanities, Science Innovations and Management Studies*, *1*(1), 43-56.

Gangineni, V. N., Tyagadurgam, M. S. V., Pabbineedi, S., Penmetsa, M., Bhumireddy, J. R., & Chalasani, R. (2024). AI-Powered Cybersecurity Risk Scoring for Financial

Institutions Using Machine Learning Techniques (Approved by ICITET 2024). *Journal of Artificial Intelligence & Cloud Computing*.

S. R. Sagili, C. Goswami, V. C. Bharathi, S. Ananthi, K. Rani and R. Sathya, "Identification of Diabetic Retinopathy by Transfer Learning Based Retinal Images," 2024 9th International Conference on Communication and Electronics Systems (ICCES), Coimbatore, India, 2024, pp. 1149-1154, doi: 10.1109/ICCES63552.2024.10859381.

S. R. Sagili and T. B. Kinsman, "Drive Dash: Vehicle Crash Insights Reporting System," 2024 International Conference on Intelligent Systems and Advanced Applications (ICISAA), Pune, India, 2024, pp. 1-6, doi: 10.1109/ICISAA62385.2024.10828724.

S. R. Sagili, S. Chidambaranathan, N. Nallametti, H. M. Bodele, L. Raja and P. G. Gayathri, "NeuroPCA: Enhancing Alzheimer's disorder Disease Detection through Optimized Feature Reduction and Machine Learning," 2024 Third International Conference on Electrical, Electronics, Information and Communication Technologies (ICEEICT), Trichirappalli, India, 2024, pp. 1-9, doi: 10.1109/ICEEICT61591.2024.10718628.

S. R. Sagili, V. K, B. Puli, P. Sundaramoorthy, M. R and K. N V, "Advancing Cervical Cancer Identification using Generative-based Adversarial Networks: An Integrative Learning Methodology," 2025 6th International Conference for Emerging Technology (INCET), BELGAUM, India, 2025, pp. 1-5, doi: 10.1109/INCET64471.2025.11140170.

Routhu, K. K. (2024). Beyond Automation: AI-Powered Employee Engagement Journeys in Oracle HCM Cloud. *KOS Journal of AIML, Data Science, and Robotics*, *1*(1), 1-6.

Routhu, K. K. (2024). The future of HCM: Evaluating Oracle's and SAP's AI-powered solutions for workforce strategy. *Journal of Artificial Intelligence, Machine Learning & Data Science*, *2*(2), 2942-2947.