

ANALYSIS OF MODERN ARTIFICIAL INTELLIGENCE SYSTEMS WITH EMPHASIS ON GENERATIVE MODELS AND ETHICAL CHALLENGES

Vijetha K

Department of AI&ML
C.Byregowda Institute of Technology
Kolar 563101
vijetha.cbit@gmail.com

Jyothi N

Department of AI&ML
C.Byregowda Institute of Technology
Kolar 563101
jyothinagaraj@gmail.com

Abstract: Artificial Intelligence (AI), combined with machine learning and computer vision, is rapidly reshaping the way humans interact with technology by enabling intelligent automation and evidence-based decision systems. These technologies are enhancing efficiency and performance across major sectors such as healthcare, transportation, and education by delivering faster, more adaptive, and user-centered services. In particular, deep learning has become a core driver of recent progress, supported by both extensive theoretical advancements and large-scale real-world applications. At the same time, generative AI has introduced powerful methods for producing text, images, and source code through modern architectures such as transformers, generative adversarial networks (GANs), and diffusion-based models. These innovations are significantly influencing domains like natural language processing, software development, and immersive digital environments. Despite these advancements, critical challenges remain regarding ethical use, transparency, accountability, and data governance. This study provides a detailed review of current AI trends, assesses generative models using established evaluation metrics such as BLEU and ROUGE, and explores their broader technological and societal impacts. Furthermore, it highlights the importance of developing strong ethical frameworks to ensure that AI systems are deployed responsibly, securely, and transparently in practical applications.

Keywords: Intelligent systems, Automated learning systems, Advanced neural learning architectures, Generative AI, Transformers, GANs, Diffusion Models, Natural Language Processing, Ethics in AI, BLEU, ROUGE

I. INTRODUCTION

Learning-based systems have become a cornerstone of modern technological advancement, influencing numerous sectors and reshaping everyday life. AI systems are widely applied in autonomous systems, medical diagnostics, financial analytics, and interactive platforms. Their growing adoption highlights their importance in improving efficiency and enabling intelligent decision-making. Beyond technical development, AI introduces important societal and ethical considerations. Ensuring fairness, accountability, and safety is important given the growing integration of these systems into critical applications.

Computer vision, a major subfield of AI, enables machines to interpret visual data and perform functions such as visual object identification and pattern recognition. Its applications span industries including manufacturing, automotive systems, and energy management.

The emergence of generative AI marks a significant transition from traditional predictive models. Technologies such as transformers, adversarial generative models, and probabilistic autoencoder-based frameworks, and diffusion models enable the creation of original content, such as written and visual outputs, and simulations. These systems have expanded AI applications in creative and technical domains.

Developments are redefining human-computer interaction and expanding the scope of immersive technologies.

II. LITERATURE SURVEY

Research in artificial intelligence has progressed at a remarkable pace, where computational intelligence techniques, machine learning algorithms, and deep neural network architectures now serve as the core foundation of contemporary intelligent systems. Existing studies indicate that these approaches significantly enhance operational performance and enable more accurate, data-driven decision-making across diverse application areas.

In particular, hierarchical learning models have shown notable success in domains such as speech recognition, image analysis, and predictive modeling. Their ability to process large-scale, complex, and high-dimensional datasets has contributed to major improvements in real-world applications. More recently, generative artificial intelligence has emerged as an important research direction, focusing on the creation of realistic and contextually meaningful outputs. Modern architectures such as transformer-based networks, generative adversarial networks (GANs), and diffusion-based models have expanded capabilities in natural language generation, image synthesis, and automated software code production.

Despite these advancements, current research also identifies several limitations, including biased outputs, reduced model interpretability, and growing concerns related to data privacy and security. Although evaluation measures such as BLEU and ROUGE are widely used, they are often insufficient to fully capture

Additionally, the utilization of generative AI with augmented and virtual reality is enhancing interactive environments, particularly in education, gaming, and training. semantic relevance and contextual understanding. Overall, prior work underscores the need for improved evaluation frameworks, stronger ethical governance, and well-defined regulatory mechanisms to ensure the responsible, transparent, and secure deployment of artificial intelligence technologies.

III. METHODOLOGY

This research follows a structured methodology to investigate recent progress in intelligent systems and generative modeling approaches. The study is organized into several stages to enable a detailed examination of both performance-related and ethical dimensions of modern AI systems.

The process begins with a thorough review of existing literature to capture current trends, technological advancements, and notable contributions in the field of AI. Information is gathered from scholarly journals, technical documentation, and credible industry sources to build a strong theoretical foundation for the study.

Following this, key generative AI architectures are identified for analysis, including transformer-based models, generative adversarial networks (GANs), and diffusion-based frameworks. These models are compared in terms of their structural design, working principles, and practical application areas.

The performance of models generating textual outputs is measured using established evaluation metrics such as BLEU and ROUGE. For image-based and other non-text outputs, qualitative assessment methods are applied to evaluate factors like visual quality, consistency, and relevance.

In addition, the study incorporates an ethical assessment component that examines issues related to bias, fairness, transparency, interpretability, and data privacy. This ensures a balanced evaluation that considers both technical effectiveness and responsible deployment.

Finally, findings from the technical analysis and ethical evaluation are combined to present a holistic understanding of current AI systems, along with their strengths, limitations, and broader implications.

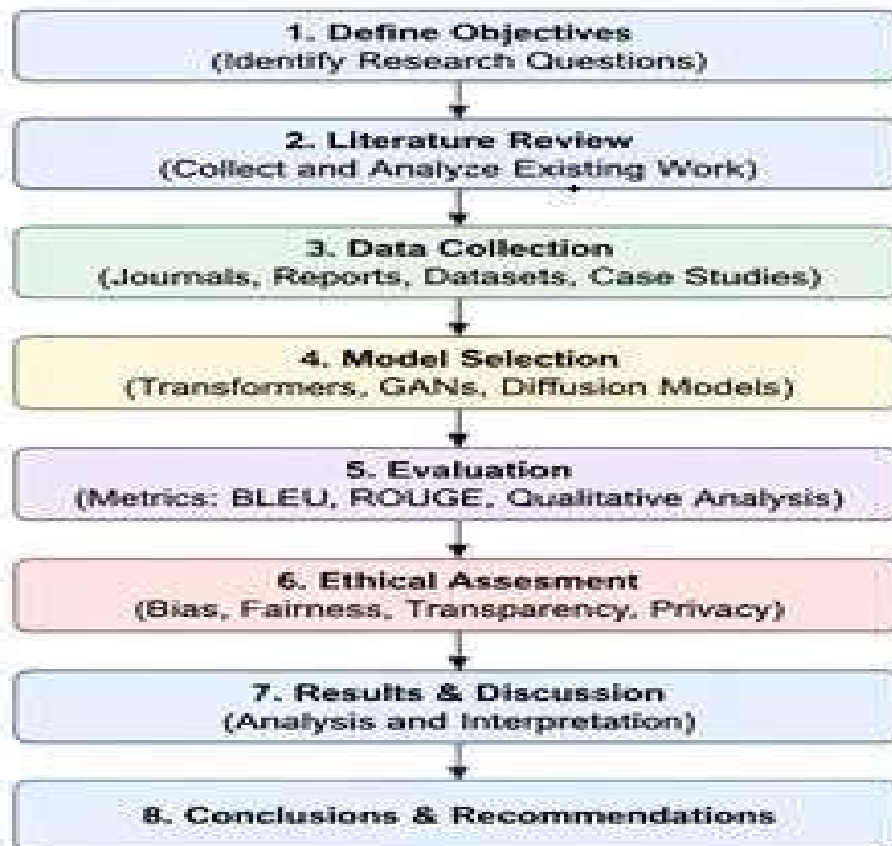


Fig. 1. Research methodology flowchart.

IV. DATA COLLECTION

A multi-source data collection approach is used in this study to ensure that the analysis of artificial intelligence and generative models is based on complete and reliable information. This approach helps in covering both technical developments and real-world applications from different perspectives.

The primary source of information includes academic research papers published in peer-reviewed journals, conference proceedings, and other scholarly platforms. These sources provide verified and high-quality knowledge about Intelligent learning algorithms, Deep neural networks, computer vision, and generative AI techniques. Trusted digital libraries and research databases are used to access these studies.

Along with academic material, industry reports and technical white papers are also considered. These documents help in understanding how AI systems are being used in real applications, along with the challenges faced in practical environments and current technological trends.

In addition, datasets and benchmarking resources are obtained from trusted online repositories to support experimental analysis. These datasets enable systematic evaluation of model performance using widely accepted metrics such as BLEU and ROUGE.

Furthermore, case studies are incorporated to examine the practical deployment of artificial intelligence systems across various sectors, including healthcare, transportation, and education. This approach provides deeper insight into both the strengths and limitations of AI when applied to real-world environments.

To ensure a comprehensive ethical perspective, relevant policy documents, legal standards, and regulatory frameworks are also analyzed. These materials offer essential guidance regarding data privacy, algorithmic fairness, transparency, and the responsible utilization of artificial intelligence technologies.

Finally, all collected information is thoroughly examined, refined, and systematically organized to ensure consistency, reliability, and accuracy. The resulting curated dataset forms the foundation for subsequent analysis and interpretation within the study.

V. PROPOSED FRAMEWORK

The proposed framework integrates literature analysis, model evaluation, and ethical assessment.

The study is initiated with research problem formulation objectives and identifying key evaluation criteria. Generative models are analyzed across multiple tasks, including text, image, and code generation.

Performance is measured based on established metrics including BLEU and ROUGE, along with qualitative assessments for realism and consistency. Ethical considerations, including bias detection and transparency analysis, are incorporated into the framework. The final output provides a balanced evaluation of technical performance and societal implications.

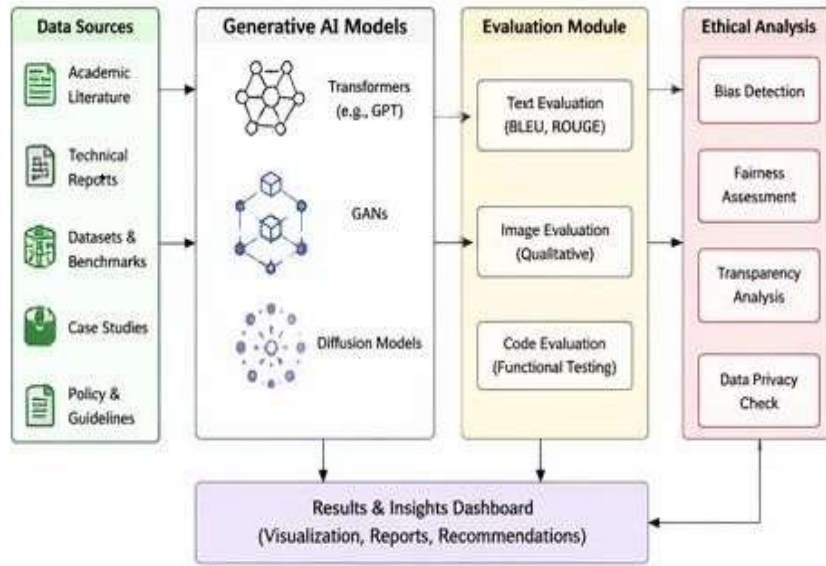


Fig. 2. Proposed architecture for generative AI analysis.

VI. RESULTS AND DISCUSSION

The evaluation results indicate that generative artificial intelligence models demonstrate strong performance across a broad spectrum of tasks. Architectures based on transformer models are capable of generating highly coherent and contextually relevant textual outputs with considerable accuracy. In a similar manner, generative adversarial networks (GANs) and diffusion-based approaches produce images that are both visually realistic and structurally consistent, while automated code generation systems can create syntactically valid and functionally executable programs.

However, despite these advantages, certain limitations remain evident. At times, generated outputs may contain inaccuracies or logical inconsistencies.

Additionally, model behavior can be influenced by biases embedded within the training datasets, which may affect fairness, reliability, and overall trustworthiness of results. Moreover, commonly used evaluation metrics are often insufficient to capture deeper semantic relationships and full contextual understanding within generated outputs. From an ethical perspective, concerns related to transparency, interpretability, and data privacy remain significant.

Overall, the findings indicate that although current AI systems demonstrate strong capabilities, further improvements are necessary to enhance their robustness, reliability, and trustworthiness in real-world applications.

TABLE I
TEST CASE RESULTS FOR GENERATIVE AI MODELS

Test Case ID	Task Type	Evaluation Method	Performance Metrics	Result Summary
TC-01	Text Generation	BLEU, ROUGE	BLEU: 0.82 ROUGE: 0.85	High coherence and contextual accuracy
TC-02	Text Summarization	ROUGE	ROUGE: 0.83	Key information retained with minimal redundancy
TC-03	Image Generation	Qualitative Analysis	Visual Consistency: High	Realistic output with minor inconsistencies
TC-04	Code Generation	Functional Testing	Accuracy: 92%	Correct syntax and logic with successful execution
TC-05	Bias Detection	Fairness Evaluation	Bias Score: Moderate	Minor bias detected in certain responses
TC-06	Robustness Testing	Relevance & Coherence	Relevance: Moderate	General response with limited depth

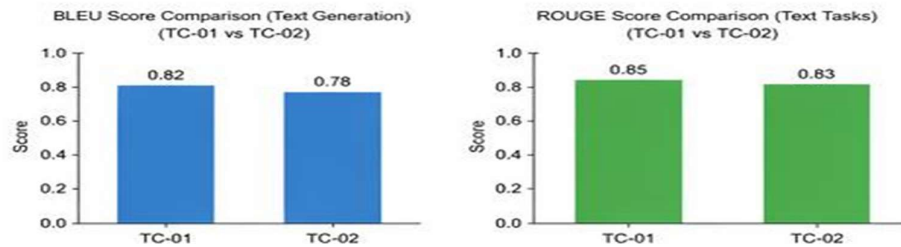


Fig. 3. BLEU and ROUGE score comparison for text generation tasks.

Test Case Analysis

A series of test cases was undertaken to verify model performance across multiple tasks. Results show strong performance in structured applications such as text generation and code synthesis. Summarization tasks demonstrate effective information compression, while image generation produces realistic outputs with minor inconsistencies. Bias detection tests reveal moderate bias levels, indicating the influence of training data. Robustness testing shows that models can handle ambiguous inputs but may produce generalized responses.

Overall, the test cases confirm the effectiveness of generative AI models while highlighting areas for improvement.

Challenges

Although the field has advanced significantly, several challenges persist. Data bias continues to affect model fairness and reliability. The lack of interpretability in complex models reduces transparency and trust. Evaluation methods are limited in capturing true semantic quality. High computational requirements restrict accessibility and scalability. Additionally, models may struggle with ambiguous inputs, affecting robustness. Moral concerns, including information privacy and lack of regulatory frameworks, further complicate AI deployment.

VII. CONCLUSION

This study provides a comprehensive investigation of contemporary artificial intelligence systems, with particular emphasis on generative modeling techniques and their practical applications in real-world environments. The findings indicate that these models achieve robust performance across multiple domains, as demonstrated through both quantitative evaluation metrics and qualitative analysis approaches. Nevertheless, a number of issues continue to exist, particularly concerning bias in model behavior, limited transparency in decision-making processes, and shortcomings in current evaluation approaches. Resolving these concerns is important for improving the dependability and trust level of AI-based systems.

Future work should aim to develop models that are easier to interpret, more resilient, and aligned with ethical principles. Advancing these aspects will play a important role in ensuring the responsible and sustainable deployment of artificial intelligence in practical environments.

VIII. FUTURE ENHANCEMENTS

Future research efforts should focus on designing more sophisticated evaluation frameworks capable of better capturing semantic comprehension and contextual relevance in outputs generated by artificial intelligence systems. Alongside this, enhancing fairness through effective bias mitigation strategies and the incorporation of diverse, well-balanced datasets is crucial for developing more inclusive and equitable models.

Improving the interpretability of models is equally important, as it can strengthen transparency and increase user trust in AI-driven applications. In addition, reinforcing model robustness is essential to ensure reliable performance when handling uncertain, ambiguous, or noisy input data. The integration of artificial intelligence with emerging technologies such as augmented reality (AR) and virtual reality (VR) also presents significant opportunities to expand the range of real-world applications.

Ultimately, establishing clear ethical guidelines and comprehensive regulatory frameworks will be vital in directing the responsible design, deployment, and utilization of artificial intelligence systems

REFERENCES

- [1]. N. Maslej et al., "Artificial Intelligence Index Report 2024," Stanford Institute for Human-Centered AI, 2024.
- [2]. C. Leiter et al., "NLLG Quarterly arXiv Report 09/24: Influential AI Papers," arXiv, 2024.
- [3]. R. Harries, C. Lawson, and P. Shapira, "Generative AI in Science: Applications, Challenges, and Emerging Questions," arXiv, 2025.
- [4]. O. Aydin et al., "Generative AI in Academic Writing: A Comparative Study of LLMs," arXiv, 2025.
- [5]. Z. Zhang, "A Comprehensive Overview of Generative AI: Technologies, Applications, and Challenges," *Neurocomputing*, vol. 632, 2025.
- [6]. A. Kumar, "Generative AI Revolution: Adoption and Performance Implications," *Journal of Business Research*, 2025.
- [7]. S. Sengar et al., "Generative Artificial Intelligence: A Systematic Review and Applications," *Multimedia Tools and Applications*, 2025.
- [8]. J. P. Andersen, "Generative AI in Research Practices and Integrity," *Technological Forecasting and Social Change*, 2025.
- [9]. Y. Liu, "Global Adoption of Generative AI: Trends and Impact," *World Development*, 2026.
- [10]. G. Ramkumar, M. Alluwaici, T. Vijetha, G. H, A. Venkatraman and R. M, "A Reliable System Performance Through NeuToMa with Spiking Neural Network Transformation in Network on Chip (NoC) Communication," 2026 International Conference on Communication, Computing and Emerging Technologies (IC3ET), Vasai, India, 2026, pp. 215-219, doi: 10.1109/IC3ET64989.2026.11467379.
- [11]. V. R, P. R. Kumar, D. Das, G. H, W. Hanbin and E. A, "Machine Learning-Driven Predictive Maintenance for Wind Turbines," 2026 International Conference on Electric Power and Renewable Energy (EPREC), Durg, India, 2026, pp. 1-6, doi: 10.1109/EPREC66546.2026.11412064.
- [12]. A. Tiwari, J. D. Kumar, W. Syed and G. H, "Efficient FPGA-Based CNN Accelerator for Real-Time Traffic Sign Classification on PYNQ-Z2," 2026 IEEE International Conference on AI Engineering and Innovations (AIEI), NIT Jamshedpur, India, 2026, pp. 1-4, doi: 10.1109/AIEI69164.2026.11496958.
- [13]. G. Ramkumar, M. Alluwaici, T. Vijetha, G. H, A. Venkatraman and R. M, "A Reliable System Performance Through NeuToMa with Spiking Neural Network Transformation in Network on Chip (NoC) Communication," 2026 International Conference on Communication, Computing and Emerging Technologies (IC3ET), Vasai, India, 2026, pp. 215-219, doi: 10.1109/IC3ET64989.2026.11467379.
- [14]. Girish H., Srinivas S., Kavya M., ...Kabbur S. "Swift Autonomous Fire Suppression Robot", 16th International Conference on Advances in Computing Control and Telecommunication Technologies Act 2025
- [15]. Girish H., Soujanya S., "A Comprehensive Cadence-based Design Flow for RISC-V UART SOC IN SCL 180NM", 16th International Conference on Advances in Computing Control and Telecommunication Technologies Act 2025

- [16]. Girish H., Rakesh S., Prasanna Raj C.P., Singh A.S., Yathisha L, "Deep Learning for Diabetic Retinopathy: Automated Detection and Classification from Retinal Images", 16th International Conference on Advances in Computing Control and Telecommunication Technologies Act 2025
- [17]. Girish H., Shyam Sundar V., Priyanka C., Renuka Hebballi R.S, "RISC V Steel Processor Implementation for Communication between Two Chips using AXI Lite", 16th International Conference on Advances in Computing Control and Telecommunication Technologies Act 2025
- [18]. M. H N, G. H and N. K N, "Simulation and Verification of SPI Master and Slave Communication," 2025 6th International Conference on Smart Electronics and Communication (ICOSEC), Trichy, India, 2025, pp. 1531-1534, doi: 10.1109/ICOSEC67334.2025.11459684. <https://ieeexplore.ieee.org/document/11459684>
- [19]. S. S. V, P. C and G. H, "Design and Analysis of a D-Band LC Bandpass Filter and Low-Noise Amplifier for 6G Front-End Applications," 2026 IEEE International Conference on Interdisciplinary Approaches in Technology and Management for Social Innovation (IATMSI), Gwalior, India, 2026, pp. 1-6, doi: 10.1109/IATMSI68868.2026.11466038. <https://ieeexplore.ieee.org/document/11466038>
- [20]. "A Comprehensive Cadence-based Design Flow for RISC-V UART SOC IN SCL 180NM", Girish H, Soujanya, Grenze International Journal of Engineering and Technology, 2025
- [21]. "Swift Autonomous Fire Suppression Robot" Girish H et. Al., Grenze International Journal of Engineering and Technology, 2025
- [22]. "Deep Learning for Diabetic Retinopathy: Automated Detection and Classification from Retinal Images", Girish H, Cyril Prasanna Raj P, Grenze International Journal of Engineering and Technology, 2025
- [23]. "RISC V Steel Processor Implementation for Communication between Two Chips using AXI Lite" Shyam Sundar V, Dr Girish H, Priyanka C, Renuka S Hebballi Grenze International Journal of Engineering and Technology, 2025
- [24]. Sujatha.S, Mahendra BM, & Dr. Girish H. (2024). Smart Irrigation System for Optimal Water Conservation. Revista Electronica De Veterinaria, 25(1), 2613 - 2619. <https://doi.org/10.69980/redvet.v25i1.1311>
- [25]. Mamatha K R, Thejaswini S, Rashmi N, & Dr. Girish H. (2024). Laser Fencing Surveillance System With Email Alert. Educational Administration: Theory and Practice, 30(2), 1539-1544. <https://doi.org/10.53555/kuey.v30i2.7919>
- [26]. Thejaswini S, Mamatha K R, Rashmi N, & Dr. Girish H. (2024). Performance Evaluation of AI Models in Early Heart Disease Diagnosis. Educational Administration: Theory and Practice, 30(2), 1531-1538. <https://doi.org/10.53555/kuey.v30i2.7918>
- [27]. Sushma K Sattigeri, Seema Srinivas, Shridhar Kabbur, Girish H, "Revolutionizing Urban Waste Management: An AI-Powered Autonomous Solution for Smart Cities", Tuijin Jishu/Journal of Propulsion Technology ISSN: 1001-4055, Vol. 45 No. 02 (2024)
- [28]. Girish H, Vajjala Sai Sree Hari,Ponde Lakshmi Narasimha,P N Shreya,Dadigala Bhargavi, "THE DESIGN AND VERIFICATION OF A 32-BIT RISC V PROCESSOR USING VEDIC MATHEMATICS", International Journal of Creative Research Thoughts (IJCRT), ISSN:2320-2882,

Volume.12, Issue 5, pp.1130-1135, May 2024, Available at :<http://www.ijcrt.org/papers/IJCRTAB02172.pdf>

- [29]. [IoT based home automation using Blynk cloud](#), J. Chetan Naik, B. S. Deeksha, H. Girish, Gurusharan Kashetty, T. G. Shivapanchakshari, Manohar Sai Polakam, Journal: [AIP Conference Proceedings](#), AIP Conf. Proc. 2965, 020006 (2024), <https://doi.org/10.1063/5.0215427>, Published: July 2024
- [30]. IoT based fire alarm system using Blynk app, T. Bhargava, H. Kishor Kumar, Pramod Mahalingappa Godi, H. Girish, J. Chetan Naik, Journal: [AIP Conference Proceedings](#), AIP Conf. Proc. 2965, 020007 (2024), <https://doi.org/10.1063/5.0212325>, Published: July 2024