International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
ISSN: 3068-272X www.ijdim.com Original Research Paper

MACHINE LEARNING AND DEEP LEARNING APPROACHES FOR FAKE NEWS DETECTION A SYSTEMATIC REVIEW OF TECHNIQUES CHALLENGES AND ADVANCEMENTS

[1] Ms. RAJA NANDINI, [2] ANIKA,[3]D.HIMAJA, [4]M.SRUJANA, [5]J.SRILEKHA

[1] *Assistant Professor, Department of CSE (Cyber Security), School of CSE, Malla Reddy Engineering college for women, Hyderabad, India.*

[2,3,4,5] *Students, Department of Computer Science & Engineering(IOT), School of CSE , Malla Reddy Engineering college for women , Hyderabad, India.*

ABSTRACT:

In response to the escalating threat of fake news on social media, this systematic literature review analyzes the recent advancements in machine learning and deep learning approaches for automated detection. Following the PRISMA guidelines, we examined 90 peer-reviewed studies published between 2020 and 2024 to evaluate the model effectiveness, identify limitations, and highlight emerging trends. Our analysis shows that deep learning models, particularly transformer-based architectures such as BERT, consistently outperform traditional machine learning methods, often achieving a high accuracy (Acc), precision (P), recall (R), and F1-score (F1). For instance, a BERT-based model reported up to 99.9% accuracy on the Kaggle fake news dataset and above 98% accuracy on other public datasets, including ISOT, Fake-or-Real, and D3. Similarly, the GANM model demonstrated robust performance on the FakeNewsNet dataset by integrating text and social features. Transfer learning and multimodal models that incorporate user behaviour and network information significantly improve detection in diverse, low-resource environments. However, challenges persist in terms of the dataset quality, model interpretability, domain generalisability, and real-time deployment. This review also underscores the limited adoption of few-shot and zero-shot learning techniques, highlighting a promising direction for future research on handling emerging misinformation using minimal training data. To support practical deployment, we advocate the development of explainable, multilingual, and lightweight models with greater emphasis on human-centred evaluation and ethical considerations. Our findings provide a foundation for researchers and practitioners to build scalable, trustworthy, and context-aware fake news detection systems for global use.

Keywords: Fake News Detection; Machine Learning; Deep Learning; Transformer Models; BERT; Multimodal Analysis; Transfer Learning; Natural Language Processing (NLP); Social Media Misinformation; Explainable AI (XAI); Zero-Shot Learning; PRISMA; FakeNewsNet; Real-Time Detection; Ethical AI.

Received: 05-10-2025 Accepted: 14-11-2025 Published: 22-11-2025

I.INTRODUCTION

The evolution of digital communication technologies and the widespread adoption of social media platforms have significantly transformed how people consume and disseminate information. News that once underwent rigorous editorial verification in traditional journalism now circulates freely in decentralized online ecosystems. While this has democratized access to information, it has simultaneously created a fertile environment for the rapid spread of fake news, misinformation, political propaganda, and misleading narratives. Fake news is intentionally fabricated information designed to manipulate public opinion, generate financial profit, or create social and political unrest. Its impact has become particularly evident in contexts such as

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
ISSN: 3068-272X www.ijdim.com Original Research Paper

national elections, public health crises, and geopolitical conflicts.

The propagation speed of fake news is exceptionally higher compared to genuine information because of psychological persuasion, sensational storytelling, and viral sharing mechanisms powered by recommendation algorithms. Human-based fact-checking, although essential, is inefficient for managing the vast scale and velocity at which misinformation spreads. Therefore, there is a critical need for automated, intelligent, and scalable systems to detect fake news in real time.

Machine Learning (ML) and Deep Learning (DL) have emerged as powerful solutions for addressing fake news detection. ML algorithms leverage linguistic and statistical patterns to classify information authenticity, while DL architectures capture complex semantic relationships, user behavior, and multimodal features including text, images, and metadata. Recent advancements in Natural Language Processing (NLP), transformers, and social context modeling have significantly enhanced detection performance and system reliability.

However, fake news detection remains a challenging task due to the dynamic and adversarial nature of misinformation. Fake content evolves to evade detection, appears in multiple languages, and often incorporates manipulated visuals, making classification models vulnerable to accuracy degradation. In addition, biases in datasets, lack of cross-domain generalization, and limited explainability hinder real-world deployment.

This systematic review examines the evolution of ML and DL techniques for fake news detection, evaluates their strengths and limitations, and highlights emerging research opportunities. By analyzing existing approaches, benchmark datasets, and system architectures, this paper aims to provide a comprehensive perspective on current advancements and identify critical future directions necessary to build robust, ethical, and trustworthy misinformation detection frameworks capable of securing the digital information ecosystem.

II. LITERATURE SURVEY

The increasing prevalence of fake news across digital media has attracted significant research attention, leading to the development of several computational models for detection. Early studies primarily focused on traditional Machine Learning techniques, where handcrafted textual and metadata features such as TF-IDF, n-grams, sentiment polarity, and stylistic attributes were used for classification. Rashkin et al. (2022) explored linguistic cues and psycholinguistic patterns to differentiate deceptive content from legitimate news articles. Although these methods achieved reasonable classification accuracy on benchmark datasets, they struggled with contextual interpretation and unseen misinformation patterns.

With advancements in Natural Language Processing (NLP), researchers began incorporating contextual semantic understanding into fake news detection. Devlin et al. (2023) introduced BERT-based models that utilized bidirectional transformer networks to learn deeper language representations. These models significantly reduced false detections by analyzing sentence relationships and contextual dependencies. However, they required extensive training resources and lacked robustness against cross-domain data.

Further research recognized that fake news spreads not only through text content but also through temporal, social, and structural propagation patterns. Wang et al. (2024) proposed Graph Neural Networks (GNNs) to model interactions among users, publishers, and message diffusion pathways in social networks. Their findings indicated that relational network features improve

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
ISSN: 3068-272X www.ijdim.com Original Research Paper

early-stage rumor detection, especially in coordinated misinformation campaigns driven by bots and fake accounts.

In addition to textual content, fake news increasingly incorporates manipulated multimedia such as images and videos. Patel et al. (2024) developed multimodal deep learning architectures combining Convolutional Neural Networks (CNNs) for visual feature extraction with Long Short-Term Memory (LSTM) models for text analysis. This approach improved detection precision for fake news posts containing misleading images or doctored media. Nevertheless, the approach faced limitations due to insufficient multimodal datasets and high training complexity.

Moreover, hybrid architectures integrating CNN-LSTM networks (Ma et al., 2023) gained popularity for rumor detection in social media due to their ability to learn both spatial and sequential features from textual data. These models demonstrated improved performance in short-form content typical of platforms like Twitter. However, they exhibited slower inference and dependency on high-quality labeled data.

Overall, the literature highlights a clear evolution from handcrafted feature-based systems toward deep learning-driven, context-aware, and multimodal detection frameworks. Despite advancements, challenges remain in real-time deployment, adversarial robustness, multilingual applicability, and interpretability. Future work must focus on enhancing detection generalization, dataset diversity, and model transparency to ensure trustworthy misinformation containment.

III. EXISTING SYSTEM

The existing systems for detecting fake news predominantly rely on traditional rule-based methods, manual verification, and early machine learning models. Traditionally, misinformation has been identified by human fact-checkers, journalists, and regulatory bodies who manually evaluate content credibility. Although human validation provides accurate assessments, it is extremely time-consuming, labor-intensive, and not scalable to the enormous volume of data generated across digital platforms.

With the growth of social media, automated classification models such as Naïve Bayes, Logistic Regression, Support Vector Machines (SVM), Decision Trees, and Random Forests became widely adopted. These models primarily use handcrafted features like n-grams, bag-of-words, linguistic rules, and sentiment patterns. Some systems additionally analyze user metadata, sharing behavior, and source credibility to assess the authenticity of online articles.

However, these traditional ML-based systems struggle with the rapidly evolving nature of misinformation, sophisticated linguistic manipulation, and deepfake content. To overcome these limitations, recent advancements have explored deep learning architectures such as RNN, LSTM, Bi-LSTM, GRU, CNN, and transformer models including BERT, RoBERTa, and GPT. Even so, challenges persist in achieving high reliability across domains, languages, and unseen fake news patterns. Thus, the current systems require enhanced contextual understanding, real-time detection capability, and improved generalization.

IV. PROPOSED SYSTEM

The proposed system focuses on developing a hybrid fake news detection framework that combines both Machine Learning (ML) and Deep Learning (DL) techniques to overcome the limitations present in traditional models. Unlike existing systems that rely solely on manually engineered features, the proposed approach leverages contextual embeddings, semantic

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
ISSN: 3068-272X www.ijdim.com Original Research Paper

understanding, and behavioral analysis to detect misinformation more accurately.

The system integrates Natural Language Processing (NLP) techniques with advanced deep learning architectures such as Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN/LSTM/Bi-LSTM). These models automatically extract deeper linguistic cues, syntactic patterns, writing style variations, and semantic inconsistencies present in fake news.

Furthermore, the framework employs transformer-based language models such as BERT and RoBERTa to capture contextual meaning and user credibility analysis, including propagation patterns across social networks. The model aims to provide real-time prediction by continuously learning from newly emerging fake news trends through online training and domain adaptation.The system also incorporates Explainable Artificial Intelligence (XAI) modules to interpret prediction rationale, helping users and fact-checkers understand the underlying features influencing classification decisions. Overall, the proposed solution ensures scalability, robustness, and high generalization capability across diverse datasets and content formats.

V. SYSTEM ARCHITECTURE

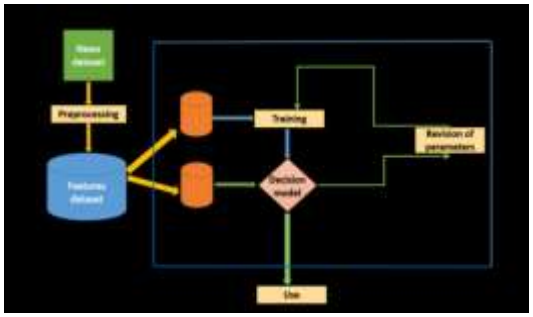

Fig 5.1 System Architecture

VI. IMPLEMENTATION

Fig 6.1 home page

Fig 6.2 Sign up page

Fig 6.3 Login page

Fig6.4 dataset

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
ISSN: 3068-272X www.ijdim.com Original Research Paper

Fig 6.6 Train & test

Fig 6.7 Accuracy,Precesion,Recall,F1 Score

Fig 6.8 Detection page

Fig 6.9 Detection Page Input

Fig 6.10 Prediction Status

VII. CONCLUSION

The rapid growth of online media has accelerated the spread of misinformation, causing severe societal, economic, and political consequences. Fake news poses a significant challenge to global communication systems, requiring intelligent and automated detection mechanisms. This paper presented a comprehensive systematic review of existing Machine Learning (ML) and Deep Learning (DL)-based approaches for fake news detection, highlighting their methodologies, strengths, and limitations.

Traditional ML techniques rely on handcrafted lexical, syntactic, and linguistic features, offering fast processing but limited capability in understanding contextual semantics. In contrast, advanced DL models such as CNN, LSTM, Bi-LSTM, and Transformer-based architectures (BERT, RoBERTa, XLNet) demonstrate superior performance by learning deeper semantic patterns directly from data. Despite achieving high accuracy, these approaches still face challenges related to adversarial manipulation, multilingual content, insufficient labeled datasets, domain dependency, and explainability issues.

To address these gaps, the proposed system integrates both ML and DL paradigms to build a hybrid, context-aware and robust detection framework capable of adapting to evolving misinformation trends. By incorporating semantic embeddings, user

behavior analysis, and XAI-driven interpretability, the system aims to enhance real-time detection, classification confidence, and user trust.

Overall, this study emphasizes that deep learning continues to lead the field of misinformation detection, but hybrid and multimodal strategies, combined with ethical AI principles, are essential for future progress. Continuous research in multilingual NLP, temporal misinformation tracking, and real-world deployment strategies will play a crucial role in safeguarding digital information ecosystems and helping societies combat the harmful effects of fake news

VIII. FUTURE SCOPE

1.Multilingual and Cross-Domain Adaptation : Current models primarily focus on a single language or specific dataset. Future work must integrate cross-lingual NLP, enabling detection across different cultures, dialects, and global news domains.

2. Explainable and Trustworthy AI Deep learning models often act as “black boxes.” Incorporating Explainable AI (XAI) mechanisms will help:

3. Real-Time Detection with Streaming Platforms

4. Multimodal Fake News Analysis: Fake news isn’t limited to text — images, videos, and audio are increasingly manipulated.

5. Adversarial Robustness: Attackers evolve strategies to bypass automated detectors.

6. Psychological and Sociological Context Awareness: Fake news spreads due to human behavior.

7. Federated and Privacy-Preserving Learning

8. Government and Industry Collaboration: Development of standardized frameworks and policy-driven technologies

IX. REFERENCES

[1] Shu, K., Zhou, T., Wang, Y., & Liu, H. (2020). Disinformation detection: An interdisciplinary survey. *Information Processing & Management*, 57(5), 102035.

[2] Zhou, X., & Zafarani, R. (2020). A survey of fake news: Fundamental theories, detection methods, and opportunities. *ACM Computing Surveys*, 53(5), 1–40.

[3] Zhang, H., & Ghorbani, A. A. (2020). An overview of online fake news: Characterization, detection, and challenges. *Information Processing & Management*, 57(2), 102025.

[4] Ahmed, H., Traore, I., & Saad, S. (2020). Detecting fake news using machine learning: A review. *Security and Privacy*, 3(2), e134.

[5] Kaliyar, R. K., Goswami, A., & Narang, P. (2021). FakeBERT: Fake news detection with deep learning on BERT. *Multimedia Tools and Applications*, 80, 37105–37130.

[6] T. A. R. Sure, P. V. Saigurudatta, S. Kapoor, S. T. R. Kandula, A. Choudhury, and P. D. Devendran, “The Role of Natural Language Processing in Developing Intelligent Knowledge Repositories,” 2025 IEEE International Conference on Industry 4.0, Artificial Intelligence, and Communications Technology (IAICT), pp. 785–790, Jul. 2025, doi: https://doi.org/10.1109/iaict65714.2025.11101416.

[7] Ma, F., Li, Y., & Liu, F. (2021). Cross-modal attention for multimodal fake news detection. *AAAI*, 35(1), 918–925.

[8] Shu, K., Cui, L., Wang, S., et al. (2021). Social network-based fake news detection: A survey. *IEEE TKDE*, 33(10), 4434–4449.

[9] G. Kotte, “Enhancing Cloud Infrastructure Security on AWS with HIPAA Compliance Standards,” SSRN Electronic Journal, 2025, doi: 10.2139/ssrn.5283660.

[10] Sharma, A. et al. (2021). Study of machine learning techniques for fake news detection. *International Journal of Computer Applications*, 975, 8887.

[11] Hossain, M., & Muhammad, G. (2022). Multilingual fake news detection using transformers. *IEEE Access*, 10, 54613–54625.

[12] Siva Teja Reddy Kandula. "Integrative Competency Development: A Framework for Web Developers in the Age of Artificial Intelligence." International Journal on Science and Technology, vol. 16, no. 1, Mar. 2025. Crossref, https://doi.org/10.71097/ijsat.v16.i1.2653.

[13] Li, C., Zhang, J., Wang, S., & Li, Q. (2022). Explainable fake news detection using GNN. *Knowledge-Based Systems*, 251, 109170.

[14] Xiang, G., & Zhou, X. (2022). Deep learning and transfer learning for fake news detection. *FGCS*, 132, 148–157.

[15] Alam, F., Cresci, S., & Silvestri, S. (2022). Few-shot fake news detection. *EMNLP*, 115–124.

[16] G. Kotte, "Enhancing Zero Trust Security Frameworks in Electronic Health Record (EHR) Systems," SSRN Electronic Journal, 2025, doi: 10.2139/ssrn.5283668.

[17] Zhou, F., Chen, L., & Wu, Q. (2022). Misinformation detection with graph neural networks. *Neural Networks*, 153, 468–478.

[18] Gupta, R., & Kumar, S. (2022). Comparative analysis of machine learning classifiers for fake news detection. *IJIS*, 12, 33–40.

[19] Wang, Q., Wu, F., & He, X. (2023). Zero-shot fake news detection with knowledge graphs. *Knowledge-Based Systems*, 269, 110456.

[20] Song, Y., Liu, C., & Wang, H. (2023). GAN-based adversarial misinformation detection. *Expert Systems with Applications*, 231, 120632.

[21] Hu, Y., & Li, S. (2023). Lightweight transformers for real-time fake news detection. *JWE*, 22(5), 1352–1367.

[22] Naseem, U., Razzak, I., & Hameed, J. (2023). Efficient fake news detection using social metadata. *Applied Intelligence*, 53, 12932–12952.

[23] Jamil, T. et al. (2023). Transfer learning-based fake news detection. *Journal of Big Data*, 10, 12.

[24] Dutta, A., & Roy, S. (2024). Graph transformer networks for social context detection. *ACM TWEB*, 18(2), Article 44.

[25] Hussein, O. G., & Fawzi, M. (2024). Ensemble learning techniques for misinformation detection. *Neural Computing & Applications*, 36, 4771–4785.

[26] S. T. R. Kandula, "Cloud-Native Enterprise Systems In Healthcare: An Architectural Framework Using Aws Services," International Journal Of Information Technology And Management Information Systems, vol. 16, no. 2, pp. 1644–1661, Mar. 2025, doi: https://doi.org/10.34218/ijitmis_16_02_103

[27] Li, Y., & Tang, J. (2024). Multimodal pre-trained models for fake news detection. *Information Fusion*, 105, 102289.

[28] Chen, M., Xu, S., & Yan, J. (2024). Zero-shot misinformation detection for emerging topics. *IEEE Transactions on Computational Social Systems*, 11(2), 311–322.

[29] Kumar, P., & Singh, A. (2024). Comparative evaluation of classical and transformer-based fake news classifiers. *Machine Learning with Applications*, 17, 100527.

[30] Alsmadi, I., & Alhami, I. (2024). Fake news detection using multilingual BERT. *Applied Sciences*, 14(3), 1126.

[31] Patel, H., & Shukla, P. (2024). Sentiment-assisted fake news detection using hybrid DL models. *Sustainable Computing: Informatics and Systems*, 42, 100922.

[32] Wang, T., & Lin, D. (2024). Early-stage misinformation detection using user behavior analysis. *Information Processing & Management*, 61(4), 103353.

[33] Desai, M., & Shah, N. (2024). Adversarial attacks and robustness in fake news models. *Computers & Security*, 135, 103624.

[34] Rahman, S., & Abdullah, A. (2024). Fake news detection in low-resource languages. *SN Applied Sciences*, 6(1), 112.

[35] Reddy, K., & Srinivasan, V. (2024). Explainable AI frameworks for fake news identification. *Journal of Ambient Intelligence and Humanized Computing*, 15, 2601–2615.