Fake News Detection Using Artificial Intelligence: A Performance Comparison

[1]URAVAKONDA VASIF, [2]DR.S. BHARATH BHUSHAN
[1]Student, [2]Professor
School of Computer Science and Engineering,
VIT-AP University, Amaravati, Andhra Pradesh, 522241, India

***Abstract*—** Nowadays, spreading fake news is a major problem. It leads to misleading people on various platforms, like newspapers and social media. Ensuring the credibility of news is a great challenge. Traditional cross-checking of facts in the news consumes more time and requires various resources. This research paper helps to ensure the truth and credibility of news by classifying text into fake or real news. The current dataset contains 21,417 articles labeled as true news and 23,481 articles labeled as fake news. A total of seven classifiers, namely, Logistic Regression (LR), Decision Tree (DT), Gradient Boosting (GB), Random Forest (RF), and XGBoost (XGB), along with deep learning models, Long-Short Term Memory (LSTM), and a hybrid CNN+LSTM model. Accuracy, precision, F1-score, sensitivity, specificity, and AUROC were considered for evaluation. This approach helps to determine which model performs well on the dataset with high evaluation metrics. XGBoost outperformed the other classifiers with a testing accuracy of 0.9970, a precision of 0.9959, an F1-score of 0.9969, a sensitivity of 0.9978, a specificity of 0.9963, and an AUROC of 0.9971.

***Index Terms*—** Fake News Detection, Artificial Intelligence (AI), Machine Learning (ML), Natural Language Processing (NLP), XGBoost, Performance Evaluation

I. INTRODUCTION

Social media has become a dominant platform for information exchange, allowing users to gain and share news quickly. However, the speedy spread of fake news and misinformation has become a critical concern, significantly affecting public perception and trust in legitimate news sources. Fake news circulates widely online, misleading users, shaping political discourse, influencing public opinion, and causing reputational damage through false allegations. The widespread dissemination of fake news reduces the credibility of trustworthy news sources, making the engineering of an effective fake news detection system an urgent necessity. AI has demonstrated remarkable progress in handling complex textual data, making it a promising outcome for automated fake news detection. Conventional fact-checking methods, such as manual verification and rule-based techniques, are time-consuming and struggle to keep up with the huge volume of data generated daily. Machine learning (ML) and NLP offer scalable, data-driven solutions that improve the efficiency and accuracy of fake news classification. However, detecting phony news remains challenging, as deceptive content often mimics the language and format of authentic news articles, making it difficult for keyword-based detection methods to differentiate between Real and false information. The rise of AI-generated fake news further complicates the detection process, necessitating more advanced methodologies. AI's impact extends beyond fake news detection, significantly advancing fields like healthcare [1-3], where deep learning models aid in disease diagnosis and predictive analytics. The emergence of Explainable AI (XAI) [4] has enhanced trust in AI-driven decisions by making them more transparent, while Federated Learning [5] enables decentralized model training, preserving data privacy. At the same time, Generative AI has revolutionized content creation but also poses new challenges by facilitating AI-generated misinformation. The emergence of Large Language Models (LLMs) [6], like GPT and BERT, has further improved NLP capabilities, making fake news detection systems more effective. These enhancements emphasize the need for continuous research in ethical AI development and misinformation mitigation to address the evolving landscape of digital deception. To address these challenges, researchers have increasingly adopted sophisticated deep learning techniques and NLP to improve the precision of unreal news identification. These models can recognize subtle language structure, analyze relevant relationships within text, and capture semantic meaning, enabling more accurate classification. This research focuses on enhancing fake news detection by leveraging AI-based methods. Our study evaluates these models based on key evaluation metrics, including accuracy, precision, F1-score, sensitivity, specificity, and AUROC, to determine the most effective approach for false news classification. The rest of the paper is arranged as follows: Section 2 (Related Works) discusses existing research on false news detection. Section 3 (Methodology) describes the experimental workflow, including data preprocessing and model implementation. Section 4 (Results and Discussion) shows the experimental findings and analyzes the efficiency of different models. Finally, Section 5 (Conclusion) briefs the key findings and outlines possible steps for future research.

II. RELATED WORKS

The increasing concern over the spread of misinformation has led to crucial improvements in automated fake news detection, enabling accurate and efficient classification of news articles. Prior research has explored several techniques, including deep learning, ML, and NLP, to enhance detection accuracy. [7] The research addresses critical ethical issues and best practices while providing a thorough overview of the current state of available datasets. It checks the evolution of false news detection models, considering the availability of diverse datasets and potential future directions for advancing detection technologies. By offering insights into dataset properties, challenges, biases, and ethical considerations, this survey supports researchers in developing more effective and resilient detection models. [8] The purpose of this research is to check the veracity of news content, including texts, images, and videos, by summarizing intentional creation and heteromorphic transmission. It analyzes deceptive intent, news propagation patterns, and user reactions to enhance false news detection. The study addresses the significant issue of fake news propagation on online platforms and its negative impact on countries, societies, and individuals. [9] The review of varied research efforts on fake news detection approaches highlights the main issues related to datasets, data fusion, and feature representation. It investigates the Effect of multiple factors on detection accuracy, such as dataset limitations, underfitting/overfitting, image-based features, feature vector representation, machine learning models, and data fusion techniques. By addressing these issues, the review provides valuable insights for experts and scholars working in the field. [10] The study offers a detailed overview of definitions, datasets, tools, and methods used for false news detection while highlighting the need for effective debunking tools to ensure the dissemination of genuine information. It reviews fake news detection by defining key terms like misinformation and disinformation and categorizing detection methods into content-based approaches (textual and visual analysis) and social context-based approaches (propagation and publisher credibility). By addressing the difficulties posed by the widespread dissemination of misinformation, this research highlights the importance of effective detection tools for maintaining social harmony and provides a detailed analysis of existing methods and datasets. [11] The investigation into the use of machine learning classifiers for automatically detecting unreal news has become increasingly prevalent on online platforms. It reviews the collection and analysis of research papers from various databases based on predefined inclusion and exclusion criteria. By exploring automated systems to combat the spread of fake news, this research addresses the growing impact of misinformation on individuals and organizations. [12] The research identifies and categorizes the key approaches presently available for identifying fake news, illustrating them with relevant examples and discussing main challenges. It classifies the existing approaches into five main sections: language-based approaches, domain-independent approaches, machine learning approaches, knowledge-based approaches, and hybrid approaches. The research highlights the influence of misinformation on public decision-making and emphasizes the necessity of a hybrid approach that combines human wisdom with digital tools for more effective detection. Reference [13] reviews NLP solutions for fake news detection, analyzing their alignment with NLP tasks while discussing key assumptions and challenges. It compares task formulations, datasets, and NLP solutions for unreal news detection, showcasing technical challenges and machine learning approaches used to address them. The research underscores the need for automated fake news detection and distinguishes it from fact-checking, rumor detection, stance detection, and sentiment analysis, while also emphasizing their unique challenges. [14] The review of the detection of false news on online platforms includes a description based on psychology and public theories, as well as current algorithms from a data mining viewpoint. It explores fake news detection through characterization (psychological and social aspects) and detection (data mining), categorizing methods into News Content and Social Context Models, covering feature extraction, model construction, datasets, and metrics. By handling the growing influence of fake news on society, the research highlights detection challenges on social media while reviewing existing methods, their limitations, and potential future directions to improve detection and mitigation.

III. METHODOLOGY

data collection, data preprocessing, data splitting, feature extraction, model training, and visualization as depicted in Fig. 1. The dataset (https://www.kaggle.com/code/therealsampat/fake news-detection) consists of two files - Fake.csv and True.csv, which contain 21,417 articles titled as true news and 23,481 articles titled as fake news. The articles are assigned labels: 0 for Fake News and 1 for True News. Initially, both datasets are combined and shuffled. The data is pre-processed for a better understanding of the models, which includes various steps: text cleaning, lowercasing, removing URLs, special characters, digits, and punctuation, expanding contractions, removing stop words, lemmatizing words, and conducting spelling correction. The dataset is then randomly separated into training (70%), validation (15%), and test (15%) sets. For feature extraction, TF-IDF vectorization is employed to transform text into numerical vectors appropriate for traditional machine learning models, while tokenization and padding alter text into sequences for deep learning models using Tokenizer and Padding Sequences. Model selection and training include both machine learning and natural language processing algorithms, along with one hybrid model. The algorithms used are LR, DT, GB, RF, XGB, LSTM, and a hybrid model (CNN+LSTM). Deep learning classifiers were trained up to 200 epochs with hyperparameters optimized for better performance.

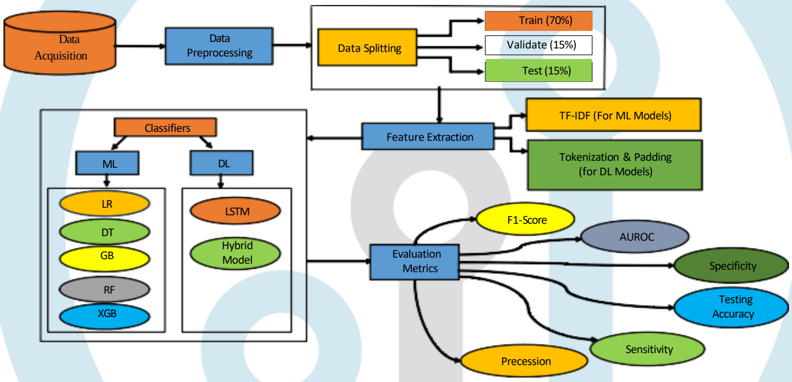

Figure 1. Experimental Flow

IV. RESULTS AND DISCUSSION

Table 1 summarizes the performance evaluation of the different models and ranks them from highest to lowest accuracy. Overview of Model Performance With a 99.70% accuracy, 99.78% sensitivity, 99.63% specificity, and an AUROC of 99.71%, XGB outperformed all other tested models (Figure 6, 13). With a well separated curve showing a substantial distinction between fake and authentic news, XGB's AUROC graph (Figure 13) demonstrates its powerful classification capacity. Following closely behind were RF and DT, which achieved 99.60% and 99.61% accuracy, respectively, with well-balanced specificity and high precision of 99.62% and 99.44%. Their excellent classification ability is demonstrated by the AUROC graphs of DT (Figure 3) and RF (Figure 5). With 99.72% sensitivity and 99.57% accuracy, GB likewise demonstrated comparable performance (Figure 4, 11). The accuracy of the deep learning models, LSTM and Hybrid CNN+LSTM, was 98.86% and 98.94%, respectively. When compared to tree-based models, their AUROC curves (Figures 7, 8) show a good but marginally worse classification performance. On the other hand, LR performed the worst, with 98.57% accuracy, 98.04% precision, and 98.22% specificity. LR's AUROC graph (Figure 2) shows that it has a lower capacity for classification than the other models. The confusion matrices for each model are shown in Figure 9–15 for additional model performance analysis. With 3,584 positive instances and 3,181 negative cases accurately identified, XGB (Figure 13) demonstrated its excellent prediction abilities. Despite its strong performance, LSTM (Figure 14) and hybrid CNN+LSTM (Figure 15) displayed somewhat greater rates of misclassification, which dataset size restrictions may have brought on. The greatest number of misclassifications occurred in LR (Figure 9), highlighting its shortcomings in identifying intricate patterns of bogus news. On comparing with the state-of-the-art models, the suggested XGB model performance was much better than the existing models, as shown in Table 2. It can be observed from the comparison that XGB is the best model for detecting fake news since it performs better than any other model in terms of accuracy, AUROC, precision, and sensitivity. A visual comparison of model accuracies is shown in Figure 16, which makes it evident that LR fared the worst and XGB the best, followed by DT, RF, and GB. In conclusion, tree-based models (XGB, DT, RF, and GB) demonstrated the best accuracy and AUROC scores, making them the most dependable for detecting bogus news. Due to dataset size limitations, deep learning models (LSTM, CNN+LSTM) showed promise but were somewhat less successful. The least successful model was logistic regression (LR), which had trouble identifying intricate patterns in bogus news. The outcomes confirm that

XGB is the most reliable and accurate model, surpassing both deep learning and conventional machine learning techniques.

TABLE 1. Evaluation of models

Model	**Accuracy %**	**Sensitivity %**	**Specificity %**	**Precision %**	**F1 Score %**	**AUROC %**
LR	98.57	98.96	98.22	98.04	98.96	98.50
DT	99.61	99.56	**99.63**	**99.62**	99.56	99.59
GB	99.57	99.72	99.44	99.37	99.72	99.55
RF	99.60	99.72	99.49	99.44	99.72	99.58
XGB	99.70	**99.78**	**99.63**	**99.62**	**99.78**	**99.69**
LSTM	98.86	98.86	98.87	98.75	98.84	98.79
CNN+LSTM	98.84	98.68	98.99	98.87	98.68	98.78

TABLE 2. State-of-the-Art Comparison

Ref	**Model**	**Accuracy%**	**Sensitivity %**	**Specificity %**	**Precision %**	**F1-Score**	**AUROC %**
[15]	BERT	99	98	-	99	99	-
[16]	TI-CNN	-	92.77	-	92.20	92.10	-
[17]	BERT	98	-	-	-	-	-
[18]	BETO+LSTM	80	-	-	-	-	-
[19]	BOOST	86.5	-	-	-	86.5	-
[20]	BERT	88.063	-	-	-	-	-
[21]	DNN	94.31	-	-	-	-	-
Proposed	XGB	99.70	99.78	99.63	99.62	99.78	99.69
	Superior model	XGB	XGB	XGB	XGB	XGB	XGB

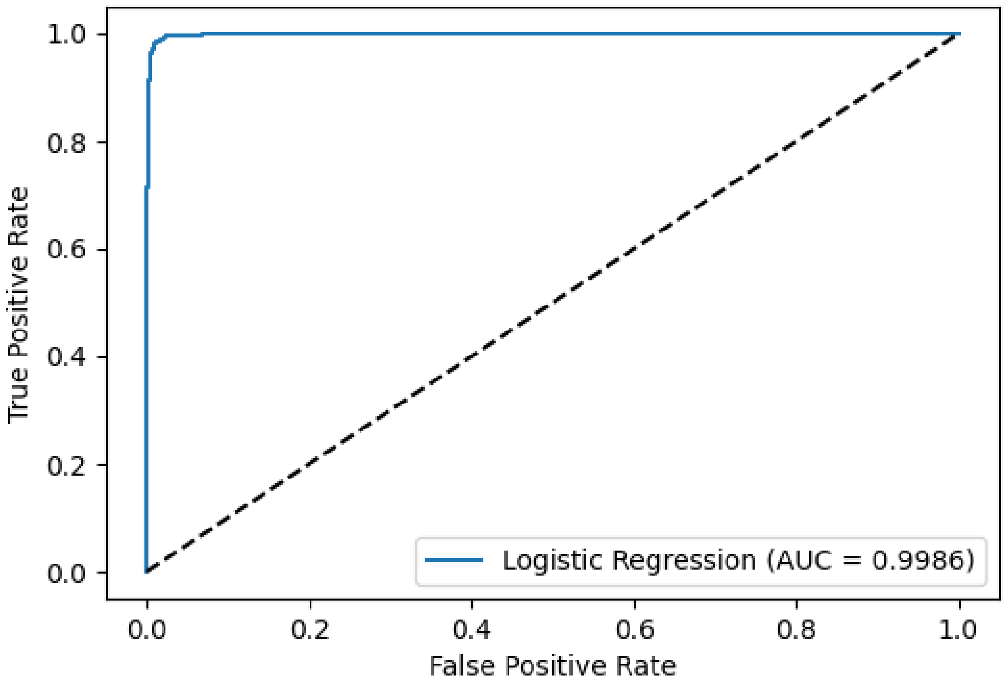

Figure 2. AUROC of LR

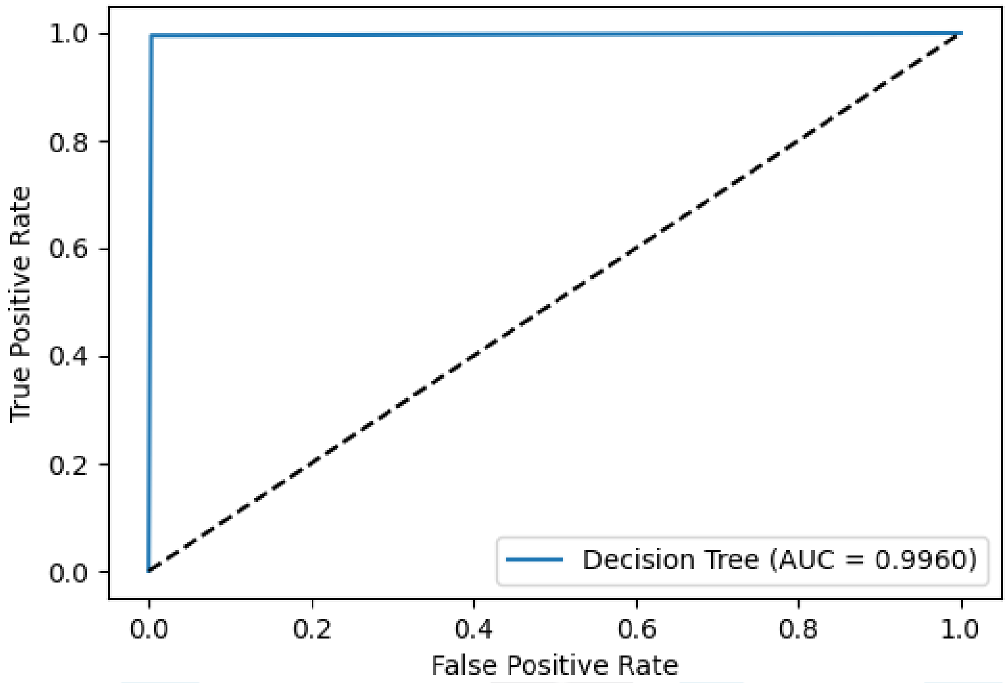

Figure 3. AUROC of DT

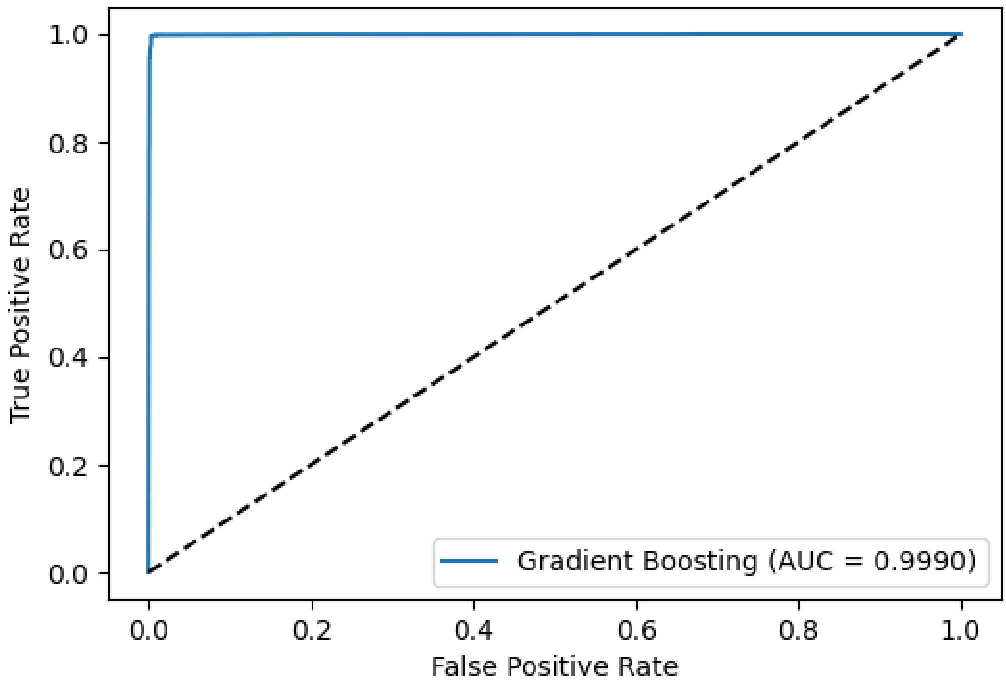

Figure 4. AUROC of GB

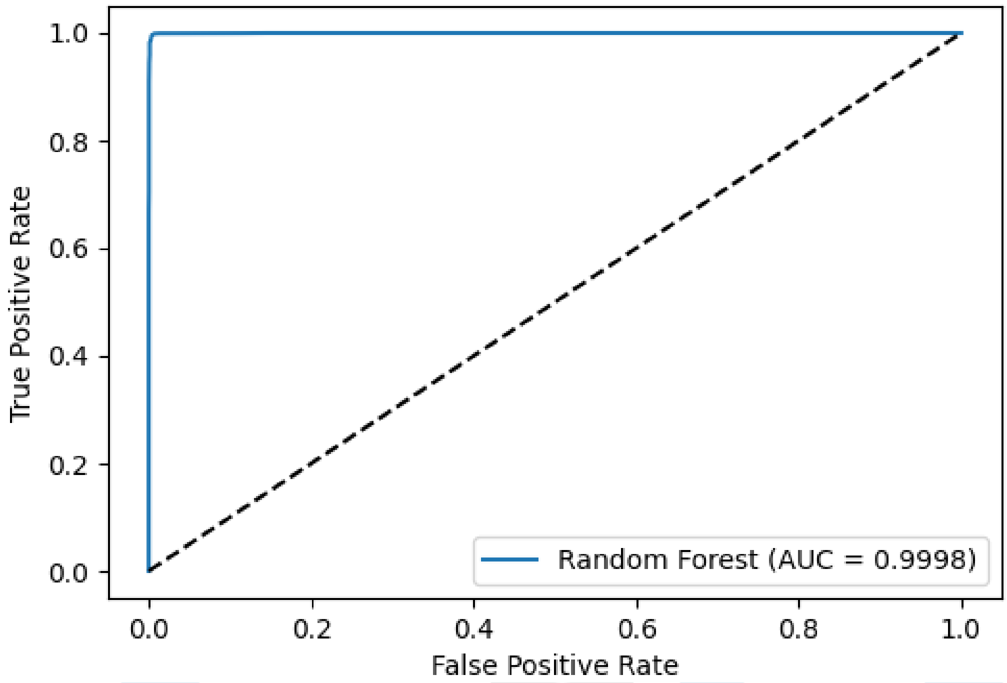

Figure 5. AUROC of RF

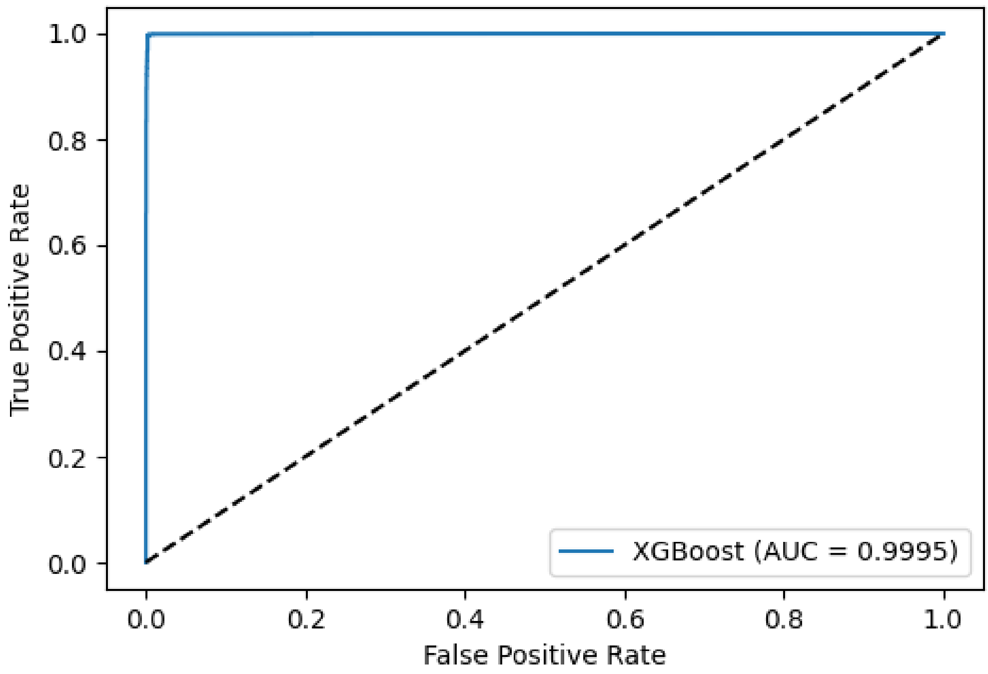

Figure 6. AUROC of XGBoost

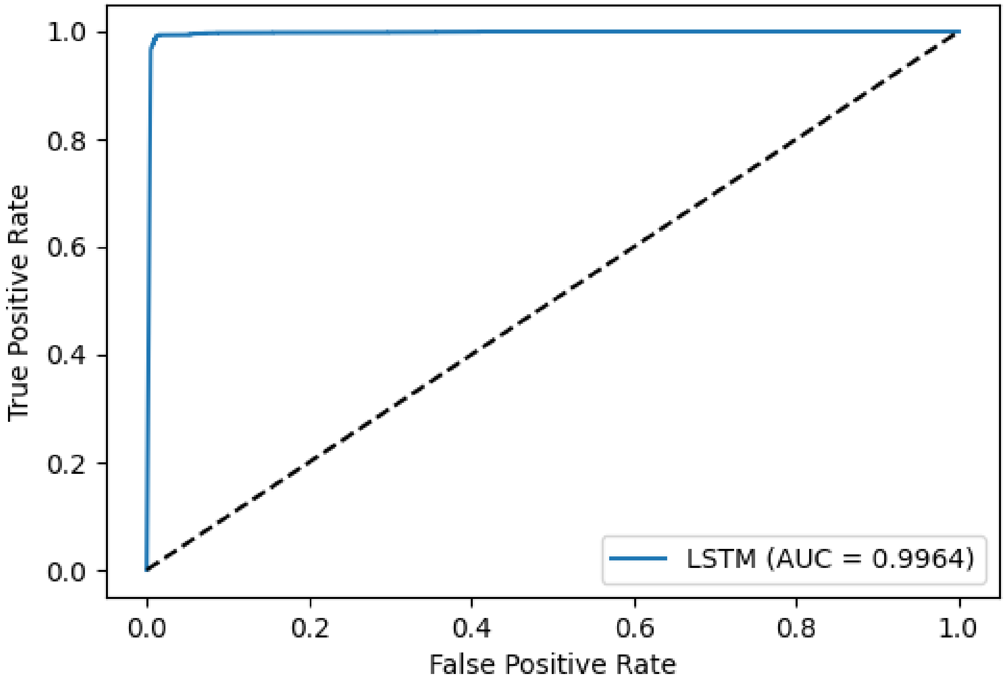

Figure 7. AUROC of LSTM

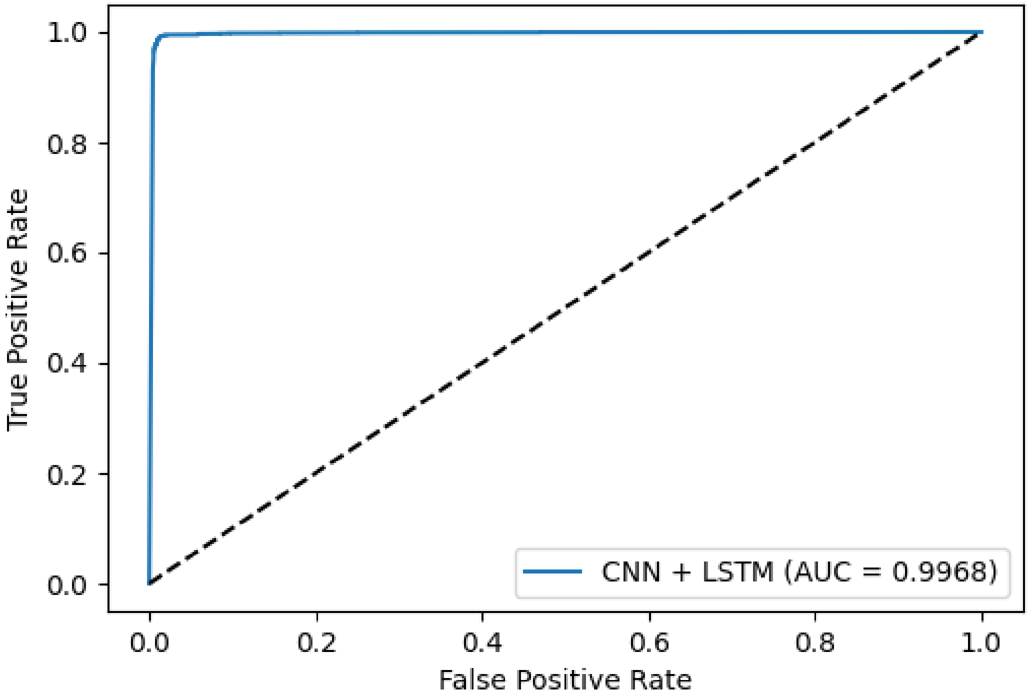

Figure 8. AUROC of CNN + LSTM

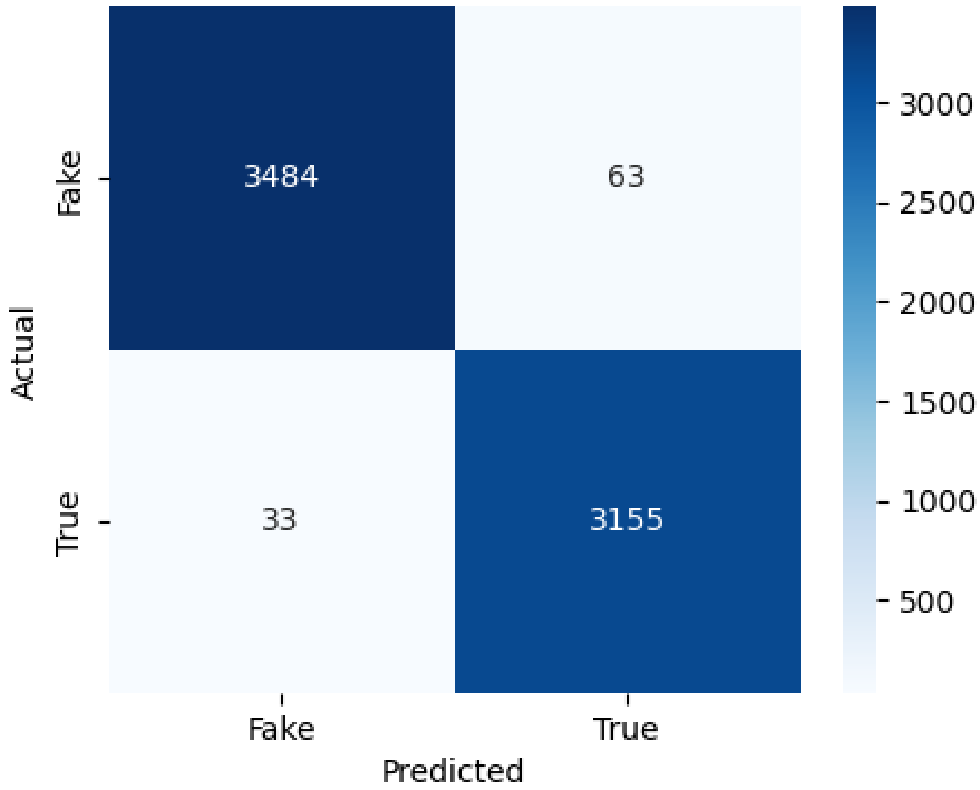

Figure 9. Confusion of LR

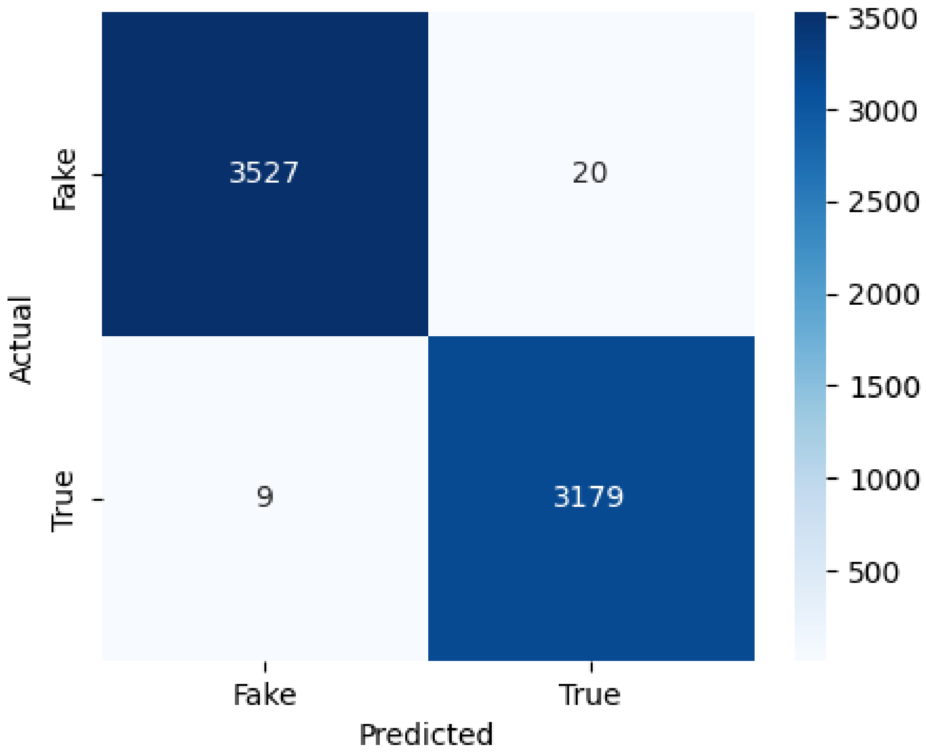

Figure 10. Confusion of DT

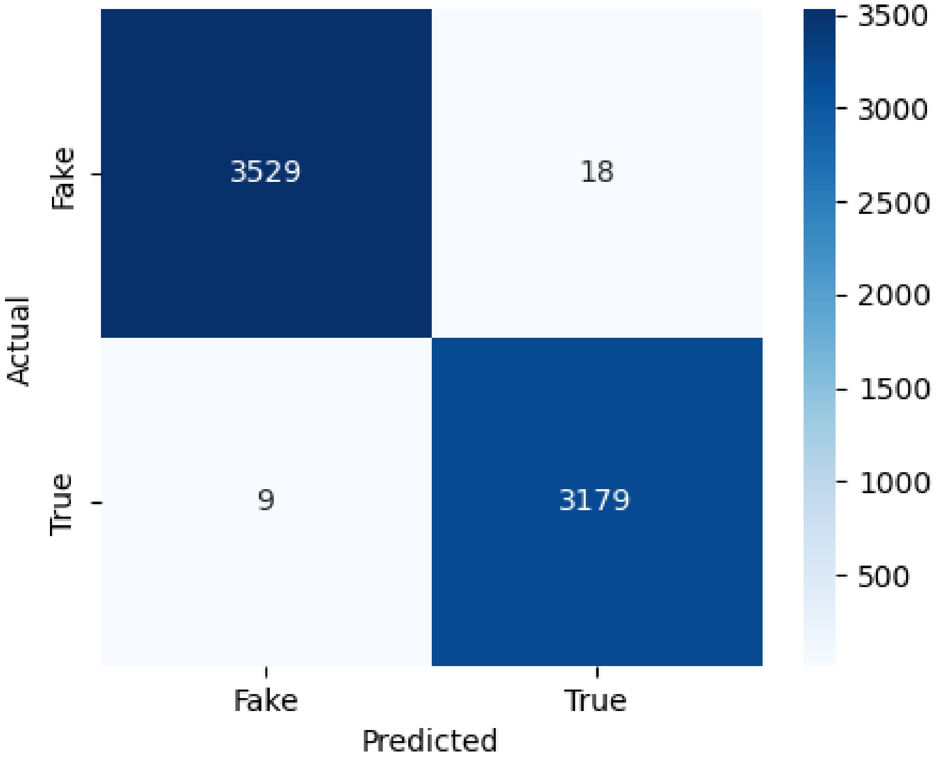

Figure 11. Confusion of GB

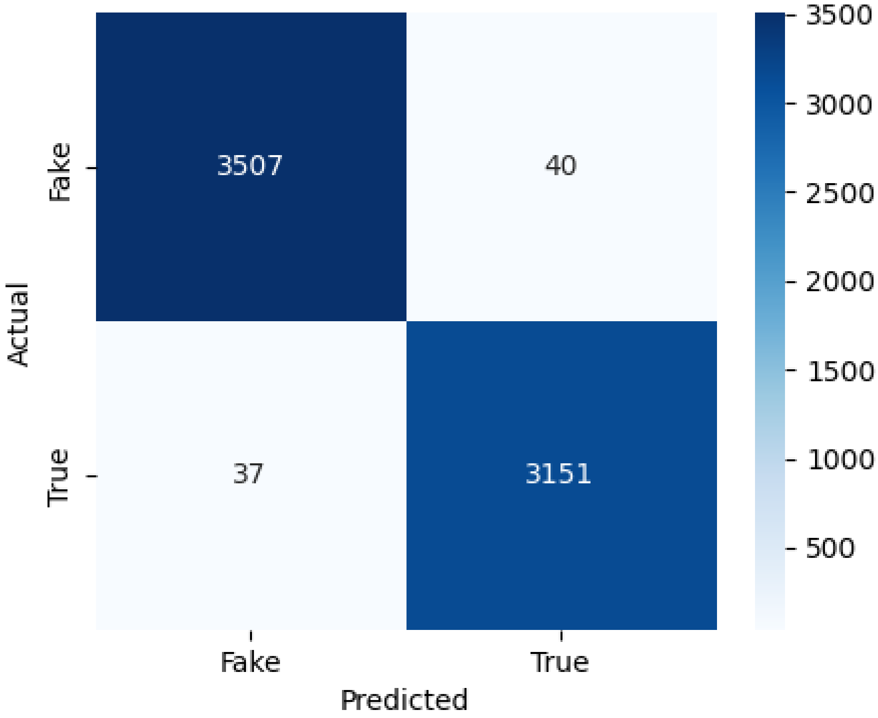

Figure 12. Confusion of RF

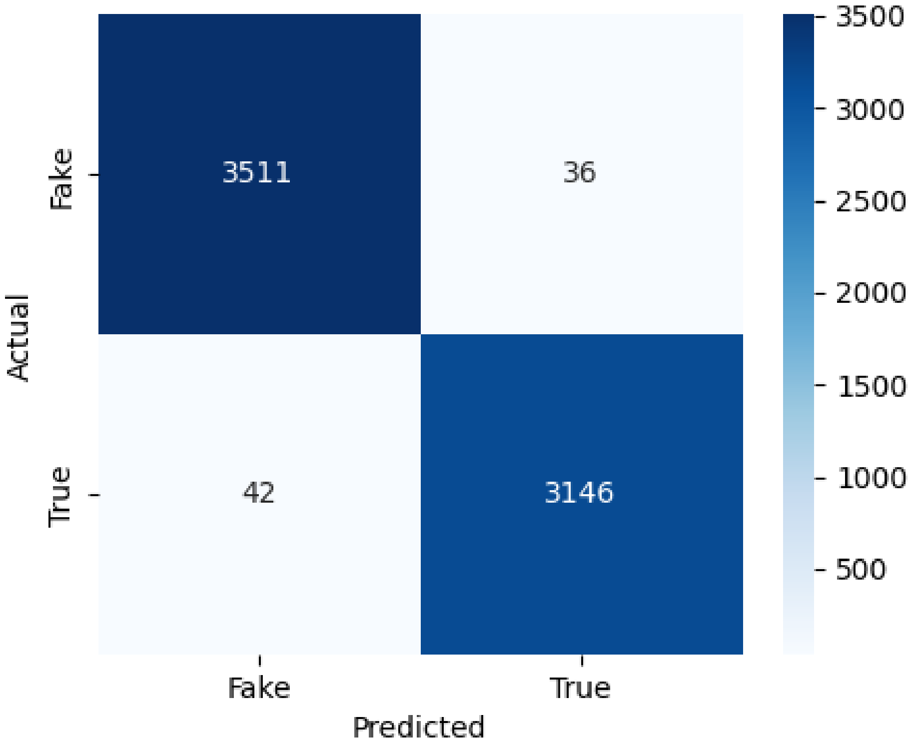

Figure 13. Confusion of XGB

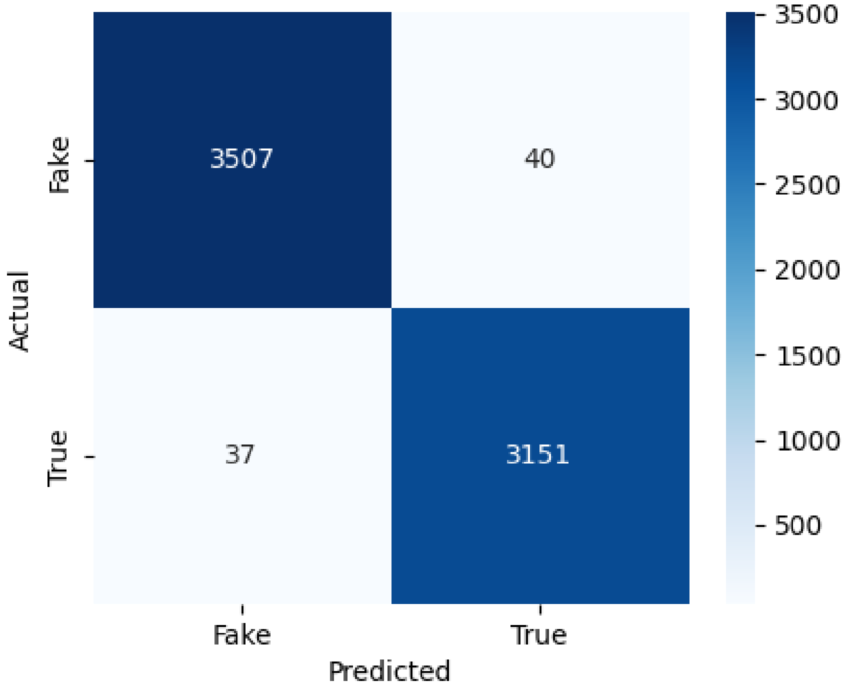

Figure 14. Confusion of LSTM

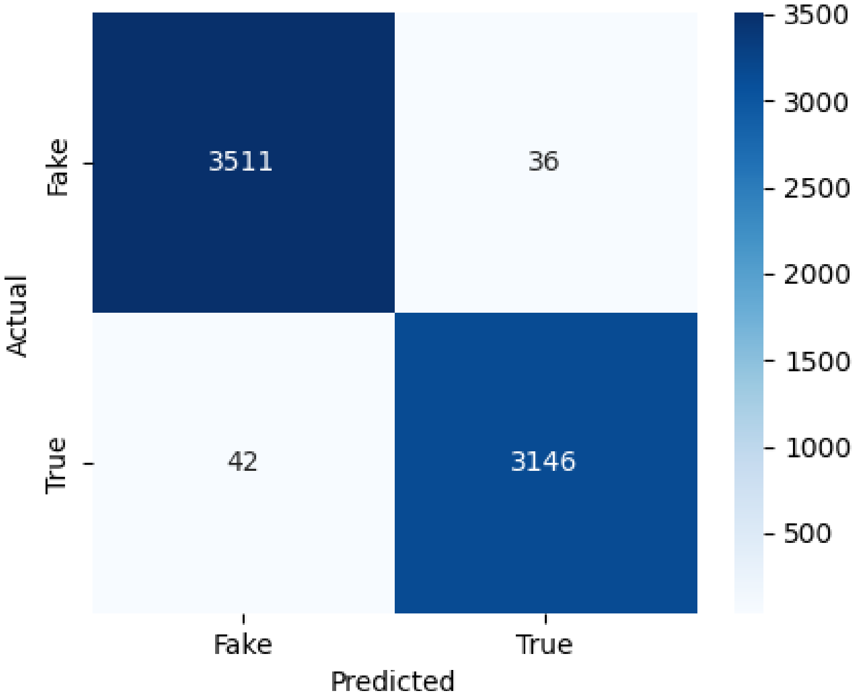

Figure 15. Confusion of CNN+LSTM

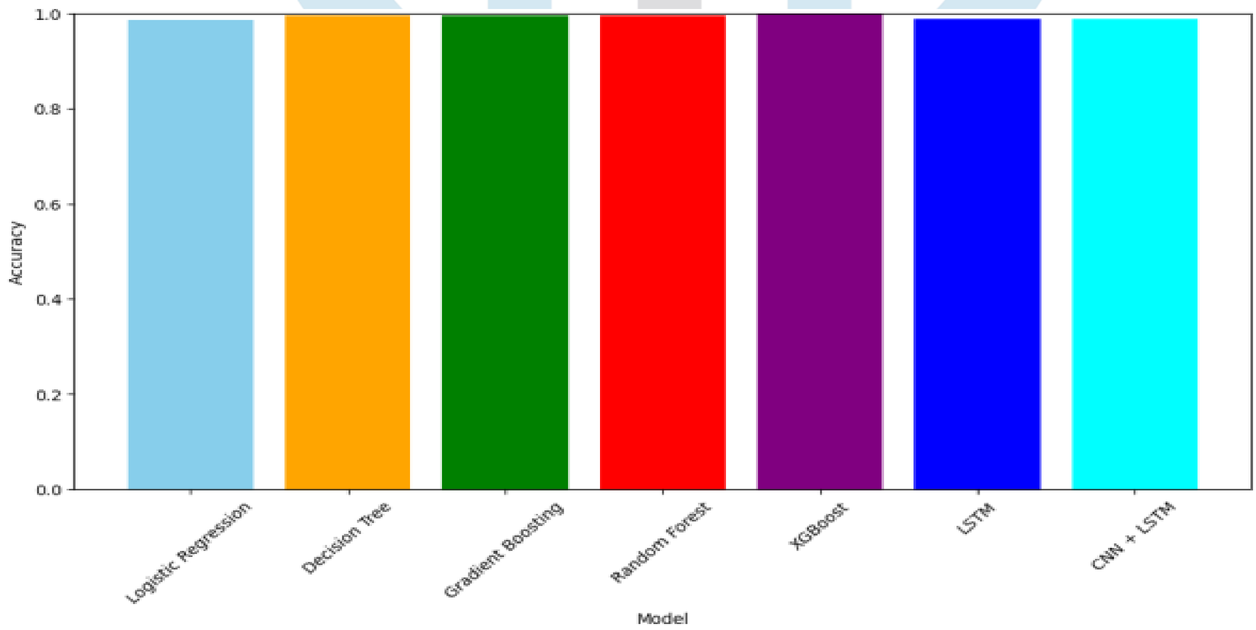

Figure 16. Model Accuracy Comparison

V. CONCLUSION

This study explores machine learning and deep learning approaches for fake news detection. Traditional machine learning models, such as LR, DT, GB, RF, and XGB, were employed, along with deep learning models, including LSTM and a hybrid CNN+LSTM architecture. To enhance model performance, the dataset underwent extensive preprocessing, including text normalization, spell correction using SymSpell, stopword removal, and lemmatization. Experimental results demonstrated that XGB performed well compared to the machine learning and deep learning models. The evaluation metrics, including confusion matrices and ROC curves, indicated that deep learning models effectively captured complex textual patterns, leading to superior classification performance.

VI. FUTURE WORK

The outcomes of this experiment are encouraging; additionally, some fields need more research and work. Using a more advanced NLP Technique Might lead to an increase in accuracy by capturing contextual data with meaning and using transformer-based architectures like BERT, RoBERTa, or XLNet. Utilizing different kinds of data, including Photos, text, and videos, to understand fake news could increase the ability of fake news detection. Expanding a dataset is an important step as it will help to improve the resilience and generalizability of the model by adding more diverse and larger datasets. Explainable AI (XAI) methods provide a better choice of models, and automated detection of fake news can become more trustworthy and transparent. Creating a real-time and scalable false news detection model that helps to identify fake news on news websites, online platforms, and social media sites. To understand the generation of fake news detection in the current digital age, more efficient and reliable false news detection systems can be made by pursuing these future research avenues.

VII. ACKNOWLEDGMENT

This study was conducted with the support of VIT-AP University, Amaravati, Andhra Pradesh. We extend our gratitude to our mentors and faculty members for their guidance and valuable insights throughout the research. We appreciate ACM SIGCHI letting us make changes to the templates they developed.

VIII. REFERENCES

[1] S. Sanjeev, G. Sai Ponnekanti and G. Pradeep Reddy, "Advanced Healthcare System using Artificial Intelligence," 2021 11th International Conference on Cloud Computing, Data Science & Engineering (Confluence), Noida, India, 2021, pp. 76-81, doi: 10.1109/Confluence51648.2021.9377084.

[2] Pradeep Reddy G, Rohan D, Kumar YVP, Prakash KP, Srikanth M. Artificial Intelligence-Based Effective Detection of Parkinson's Disease Using Voice Measurements. Engineering Proceedings. 2024; 82(1):28. https://doi.org/10.3390/ecsa-11-20481.

[3] Rohan, D., Reddy, G.P., Kumar, Y.V.P. et al. An extensive experimental analysis for heart disease prediction using artificial intelligence techniques. Sci Rep 15, 6132 (2025). https://doi.org/10.1038/s41598-025-90530-1.

[4] G. P. Reddy and Y. V. P. Kumar, "Explainable AI (XAI): Explained," 2023 IEEE Open Conference of Electrical, Electronic and Information Sciences (eStream), Vilnius, Lithuania, 2023, pp.1-6, doi: 10.1109/eStream59056.2023.10134984.

[5] G. P. Reddy and Y. V. Pavan Kumar, "A Beginner's Guide to Federated Learning," 2023 Intelligent Methods, Systems, and Applications (IMSA), Giza, Egypt, 2023, pp. 557-562, doi: 10.1109/IMSA58542.2023.10217383.

[6] G. P. Reddy, Y. V. Pavan Kumar, and K. P. Prakash, "Hallucinations in Large Language Models (LLMs)," 2024 IEEE Open Conference of Electrical, Electronic and Information Sciences (eStream), Vilnius, Lithuania, 2024, pp.1-6,doi: 10.1109/eStream61684.2024.10542617.

[7] Kuntur, Soveatin, Anna Wr'oblewska, Marcin Paprzycki and Maria Ganzha. "Fake News Detection: It's All in the Data!" ArXiv abs/2407.02122 (2024): n. pag.

[8] Bo Hu, Zhendong Mao, Yongdong Zhang, An overview of fake news detection: From a new perspective, Fundamental Research, Volume 5, Issue 1, 2025, Pages 332-346, ISSN2667-3258, https://doi.org/10.1016/j.fmre.2024.01.017.

[9] Suhaib Kh Hamed, Mohd Juzaiddin Ab Aziz, Mohd Ridzwan Yaakub, A review of fake news detection approaches: A critical analysis of relevant studies and highlighting key challenges associated with the dataset, feature representation, and data fusion, Heliyon, Volume 9, Issue 10, 2023, e20382, ISSN 2405-8440, https://doi.org/10.1016/j.heliyon.2023.e20382.

[10] Hangloo, Sakshini, and Bhavna Arora. "Fake News Detection Tools and Methods - A Review." ArXiv abs/2112.11185 (2021): n. pag.

[11] Ahmed, Alim Al Ayub & Aljarbouh, Ayman & Choi, Myung. (2021). Detecting Fake News using Machine Learning: A Systematic Literature Review. Psychology (Savannah, Ga.). .58.1932-1939. 10.17762/pae.v58i1.1046.

[12] de Beer D, Matthee M. Approaches to Identify Fake News: A Systematic Literature Review Integrated Science in Digital Age 2020. 2020 May;136:13-22. PMCID: PMC7250114.

[13] Oshikawa, Ray & Qian, Jing & Wang, William. (2018). A Survey on Natural Language Processing for Fake News Detection. 10.48550/arXiv.1811.00770.

[14] Shu, Kai & Sliva, Amy & Wang, Suhang & Tang, Jiliang & Liu, Huan. (2017). Fake News Detection on Social Media: A Data Mining Perspective. ACM SIGKDD Explorations Newsletter. 19. 10.1145/3137597.3137600.

[15] Cavus, N., Goksu, M. & Oktekin, B. Real-time fake news detection in online social networks: FANDC Cloud-based system. Sci Rep 14, 25954 (2024). https://doi.org/10.1038/s41598-024-76102-9.

[16] Yang, Yang & Zheng, Lei & Jiawei, Zhang & Cui, Qingcai & Li, Zhoujun & Yu, Philip. (2018). TI-CNN: Convolutional Neural Networks for Fake News Detection. 10.48550/arXiv.1806.00749.

[17] Prachi, Noshin Nirvana & Habibullah, Md & Rafi, Emanul Haque & Alam, Evan & Khan, Riasat. (2022). Detection of Fake News Using Machine Learning and Natural Language Processing Algorithms. Journal of Advances in Information Technology. 13. 10.12720/jait.13.6.652-661.

[18 Martínez-Gallego, Kevin, et al. "Fake News Detection in Spanish Using Deep Learning Techniques." ArXiv abs/2110.06461 (2021): n. pag.

[19] Xinyi Zhou, Atishay Jain, Vir V. Phoha, and Reza Zafarani. 2020. Fake News Early Detection: A Theory-Driven Model. Digital Threats 1, 2, Article 12 (June 2020), 25 pages. https://doi.org/10.1145/3377478.

[20] Yang, Kai-Chou et al. "Fake News Detection as Natural Language Inference." ArXiv abs/1907.07347 (2019): n. pag.

[21] Ahluwalia, Simrat & Lohia, Nibrat & Thota, Ashwin. (2018). Fake News Detection: A Deep Learning Approach. 10.