

A Comparative Analysis of Machine Learning and Deep Learning Approaches to Enhanced Fake News Detection

Wesam Ahmed

Department of Information Technology, Faculty of Computers and Artificial Intelligence, Hurghada University, Hurghada, Egypt
wesamahmed929@yahoo.com

Arshad Ali

Faculty of Computer and Information Systems, Islamic University of Madinah, Al Madinah Al Munawarah, Saudi Arabia
dr.aali7pk@gmail.com (corresponding author)

Gahangir Hossain

Anuradha and Vikas Sinha Department of Data Science, University of North Texas, Denton, TX, USA
Gahangir.Hossain@unt.edu

Mohammad Husain

Faculty of Computer and Information Systems, Islamic University of Madinah, Al Madinah Al Munawarah, Saudi Arabia
dr.husain@iu.edu.sa

Amena Mahmoud

Computer Science Department, Faculty of Computers and Information, Kafr Elsheikh University, Kafr El Sheikh, Egypt
Amena_mahmoud@fci.kfs.edu.eg

Received: 18 December 2025 | Revised: 16 January 2026 | Accepted: 28 January 2026

Licensed under a CC-BY 4.0 license | Copyright (c) by the authors | DOI: <https://doi.org/10.48084/etasr.17036>

ABSTRACT

Fake news influences major aspects of society. To mitigate the dangers of misinformation, manual fact-checking is frequently implemented. Original fake news propagators target innocent individuals to disseminate misleading information. The government and society must first determine the pattern of false news dissemination to address this sequence of events. However, since manually fact-checking the substantial volume of newly generated data is insufficient, machine learning algorithms can be used to detect fake news on different social media platforms. This study applies and evaluates four machine learning algorithms and three deep learning models in detecting fake news. The area under the receiver operating characteristic curve (AUC), precision, F1-score, recall, and accuracy were used to determine the best algorithm for classifying fake news. Logistic Regression (LR) and Long Short-Term Memory (LSTM) both achieved the highest accuracy of 92.97%. This work elucidates the capacity of machine learning and deep learning models to detect fake news.

Keywords-fake news; machine learning; deep learning; LSTM; NLP; logistic regression

I. INTRODUCTION

Social networks have become the most important source of information, changing the way people access and engage with information [1, 2]. Freedom of speech and fast access to

information on social media make fake news immensely popular, especially among young people. People worldwide use these sites to acquire news about everything, and often do not even think about whether the news is trustworthy or not [3]. Misinformation, or simply fake information that can be verified

as false, is a threat to the very essence of democracy, as it undermines the public's trust in government and influences crucial areas such as elections, the economy, and people's stance on wars to a great extent [4, 5]. Previous studies have revealed that fake news has a larger audience on social media than true news, with instances being observable on the top social media platforms [6, 7]. Fake news is everywhere on social media, undermining trust in online information, making it more vital than ever to come up with effective ways to deal with this issue. As the amount of data continues to grow, it becomes increasingly vital to quickly and effectively obtain and process relevant information.

For this purpose, computational linguistic tools play an important role. Using Artificial Intelligence (AI) approaches becomes critical, as they provide better ways to detect false information [8]. AI is a must for false news detection because it observes the minor components of language and context in a way that human censors might overlook [9, 10]. Research interest in automatic fake news identification is increasing, and several research approaches take advantage of AI and Natural Language Processing (NLP) [11]. This study investigates misinformation categorization using Machine Learning (ML) and Deep Learning (DL) methods on a comprehensive fake news detection dataset, contributing the following:

1. Evaluates the strengths, shortcomings, and applicability of ML and DL models using a dataset for reliable and scalable misinformation detection in digital ecosystems.
2. Evaluates the models using different performance metrics to assess model correctness and resilience from all angles. Logistic Regression (LR) achieved the greatest accuracy of 92.97%, showing how computationally efficient algorithms can discover fake news.
3. Evaluates several ML and DL models using a standardized fake news detection dataset to develop rigorous performance standards. These criteria enable future research to compare and create enhanced misinformation-prevention methods, increasing confidence and dependability in online information systems.

II. RELATED WORK

In social media contexts, recent research has increasingly focused on the development of effective methods to identify fake news. In [12], a deep learning model (Bi-LSTM-GRU) employed the LIAR dataset to examine news, achieving an F-score of 0.914, a precision of 0.913, a recall of 0.916, and an accuracy of 0.898. In [13], ML and DL models were tested on a dataset with Pakistani news, with the LSTM with Glove achieving the best results. In [14], actual tweets were collected and processed, and four ML algorithms were tested on the detection of fake news, with SVM achieving the highest performance. In [15], an ensemble classification model demonstrated superior accuracy compared to other models, but testing accuracy on the LIAR dataset was only 44.15%.

III. METHODOLOGY

This study investigated the identification of fake news using different models. Figure 1 shows the method's architecture, comprising data preparation, feature extraction, hyperparameter tuning of the ML models, and model training. This study selected Logistic Regression (LR), SVM, Naive Bayes (NB), and XGBoost as representative and widely used ML algorithms. For DL, LSTM and RNN were chosen due to their ability to model sequential text data, while MLP was included as a standard feedforward baseline.

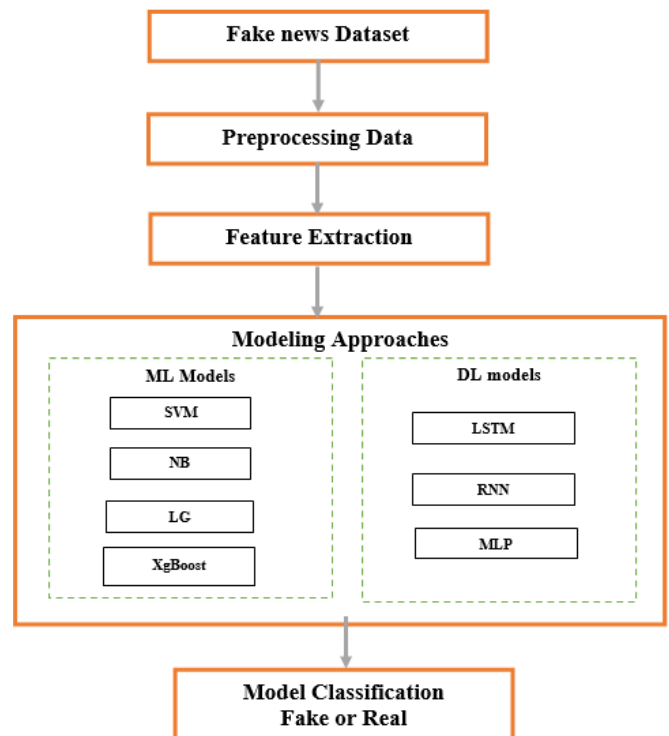


Fig. 1. The framework of this study.

A. Dataset Description and Preprocessing

This study examined fake news detection as a binary classification problem, using the dataset in [16], comprising 12520 rows with news articles from the New York Times website and fake news scraped from the AFP Fact Check website, and two features.

TABLE I. SOME EXAMPLES OF THE DATASET

Text	Label
help kid wear mask sisodias mask photoop Arvind kejriwal ji that s call advertisement person hindila	1
bearbite boost intelligence increase iq upper extreme	1
crackdown pearl square hundred riot police officer rain tear gas percussion grenade thousand demonst	0

In the preprocessing phase, all the text was changed to lowercase. The text was broken down into its parts, with stop words and punctuation deleted, and each word lemmatized (Lemmatization). This preprocessing pipeline aimed to improve the quality and relevance of the input data.

TABLE III. HYPERPARAMETERS FOR DL MODELS

Model	Parameters
MLP	Batch size = 32, Epochs = 20
RNN	Epochs = 50, Learning rate = 1e-4, Optimizer = Adam, Batch size = 128
LSTM	Epochs = 50, Optimizer = Adam, Learning rate = 1e-4, Batch size = 64

TABLE IV. PERFORMANCE RESULTS FOR ALL MODELS

Model	Accuracy (%)	Recall (%)	Precision (%)	F1-score (%)	AUC (%)
LR	92.97	95.50	92.56	94.01	97.98
XGBoost	91.01	95.09	89.92	92.43	97.04
SVM	92.37	95.22	91.86	93.51	97.18
NB	91.45	91.70	93.38	92.53	97.18
LSTM	92.97	95.22	92.78	93.99	97.85
MLP	91.33	93.28	91.00	92.32	97.38
RNN	92.17	93.95	92.24	93.09	96.00

As shown in Table IV, all models performed strongly on the dataset, achieving accuracies greater than 91%. The LSTM and LR models had the best accuracy (92.97%), followed closely by SVM (92.37%) and RNN (92.17%). This means that both ML and DL algorithms work well on fake news detection, with just a little difference in total accuracy (91.01–92.97%). The models show very good performance at finding true positives (fake news) and lowering false positives. This is shown by their high recall (91.70–95.50%), precision (89.92–93.38%), F1-score (92.32–94.01%), and AUC (96.00–97.98%). The high AUC values, especially for LR (97.98%) and LSTM (97.85%), show that these models are quite reliable across all classification thresholds. The F1-score is an important measure for unbalanced classification since it balances precision and recall. LSTM (93.99%) and RNN (93.09%) were close behind, while LR (94.01%) once again achieved the best performance. The AUC values for all models were above 96%, suggesting that all models are good at recognizing the difference between real and fake news in this dataset. The AUC of LR was the highest (97.98%), which means it is even more dependable, while LSTM (97.85%) and SVM (97.18%) were quite close behind.

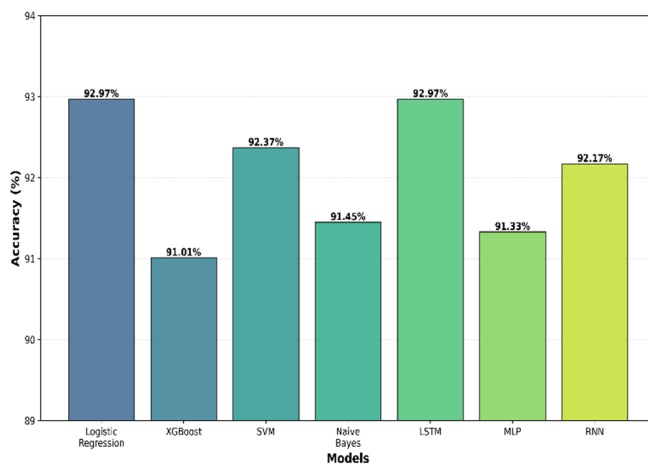


Fig. 3. Accuracy comparison.

A significant conclusion is that LR, a relatively simple linear classifier, surpassed or equaled more complex deep learning architectures like LSTM and RNN on this dataset, as seen in Figure 3. LR is considered the more pragmatic choice because of its simplicity, lower computational cost, and ease of deployment, making it suitable for real-world applications. However, it should be noted that this is due to the characteristics and the quantity of this dataset.

V. LIMITATIONS OF THE STUDY

The evaluation results were based on one dataset, which could restrict their applicability to other datasets or real-world situations that have different data distributions and characteristics. This study focused on the predictive performance of the models, but no thorough examination was performed regarding the interpretability of the models or the computational efficiency. It is worth noting that the use of a Count vectorizer with (1,2)-grams, while simple and effective, may have favored the performance of linear models such as LR and SVM. Future work could explore more sophisticated embeddings, such as TF-IDF or word2vec, to potentially enhance model performance."

VI. CONCLUSION

This investigation conducted a thorough assessment of seven DL and ML models in identifying fake news utilizing a standardized dataset. This problem was investigated using a multi-metric assessment approach. The results showed that both LR and LSTM achieved the highest accuracy of 92.97%, indicating that they were the most effective and efficient for identifying fake news. More specifically, LR is a more pragmatic choice due to its simplicity, lower computational cost, and ease of deployment, making it suitable for real-world applications. LR also allows inspection of feature coefficients, enabling transparency and explainability. This characteristic is particularly valuable in real-world and sensitive applications, such as fake news detection, where understanding model decisions is important. These results are valuable to businesses, including online news sites and social media companies, as they can be used to determine how to prevent the dissemination of fake news and contribute to the development of trust in digital ecosystems. Additionally, the performance benchmarks used in this study can serve as a basis for future research on automated fake news detection.

A potential future direction is the development of Large Language Models (LLMs) that are specifically designed to detect deceptive, nuanced information. Such future approaches can leverage the sophisticated natural language comprehension of these models to enhance their ability to combat fake news and manage a greater volume of it. Future research could also explore larger and more diverse datasets, as well as advanced DL architectures, including transformer-based models such as BERT and RoBERTa, to further improve the robustness and accuracy of fake news detection systems.

REFERENCES

- [1] X. Zhou and R. Zafarani, "A Survey of Fake News: Fundamental Theories, Detection Methods, and Opportunities," *ACM Computing Surveys*, vol. 53, no. 5, June 2020, Art. no. 109, <https://doi.org/10.1145/3395046>.
- [2] A. Dhankhar, A. Prakash, and S. Juneja, "Feature Extraction from Text using Grasshopper optimization algorithm for identifying hate speech," in *2023 International Conference on Advances in Computation, Communication and Information Technology (ICAICIT)*, Nov. 2023, pp. 494–498, <https://doi.org/10.1109/ICAICIT60255.2023.10466147>.
- [3] H. Thakar and B. Bhatt, "Fake news detection: recent trends and challenges," *Social Network Analysis and Mining*, vol. 14, no. 1, Aug. 2024, Art. no. 176, <https://doi.org/10.1007/s13278-024-01344-4>.
- [4] A. Dhankhar, A. Prakash, S. Juneja, and S. Prakash, "A Survey on Multi-Modal Hate Speech Detection," in *2023 IEEE 11th Region 10 Humanitarian Technology Conference (R10-HTC)*, Oct. 2023, pp. 225–230, <https://doi.org/10.1109/R10-HTC57504.2023.10461793>.
- [5] U. Sharma and J. Singh, "A comprehensive overview of fake news detection on social networks," *Social Network Analysis and Mining*, vol. 14, no. 1, June 2024, Art. no. 120, <https://doi.org/10.1007/s13278-024-01280-3>.
- [6] A. D. Radhi, H. A. H. A. Naffakh, A. I. Fuqdan, B. A. Hakim, and B. Al-Attar, "A comprehensive review of machine learning-based models for fake news detection," *BIO Web of Conferences*, vol. 97, 2024, Art. no. 00123, <https://doi.org/10.1051/bioconf/20249700123>.
- [7] E. F. Ayetiran and Ö. Özgöbek, "A Review of Deep Learning Techniques for Multimodal Fake News and Harmful Languages Detection," *IEEE Access*, vol. 12, pp. 76133–76153, 2024, <https://doi.org/10.1109/ACCESS.2024.3406258>.
- [8] I. Saha and S. Puthran, "Fake News Detection: A Comprehensive Review and a Novel Framework," in *2024 OITS International Conference on Information Technology (OCIT)*, Dec. 2024, pp. 463–469, <https://doi.org/10.1109/OCIT65031.2024.00087>.
- [9] S. I. Manzoor, J. Singla, and Nikita, "Fake News Detection Using Machine Learning approaches: A systematic Review," in *2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI)*, Apr. 2019, pp. 230–234, <https://doi.org/10.1109/ICOEI.2019.8862770>.
- [10] N. Capuano, G. Fenza, V. Loia, and F. D. Nota, "Content-Based Fake News Detection With Machine and Deep Learning: a Systematic Review," *Neurocomputing*, vol. 530, pp. 91–103, Apr. 2023, <https://doi.org/10.1016/j.neucom.2023.02.005>.
- [11] X. Zhang and A. A. Ghorbani, "An overview of online fake news: Characterization, detection, and discussion," *Information Processing & Management*, vol. 57, no. 2, Mar. 2020, Art. no. 102025, <https://doi.org/10.1016/j.ipm.2019.03.004>.
- [12] N. Aslam, I. Ullah Khan, F. S. Alotaibi, L. A. Aldaej, and A. K. Aldubaikil, "Fake Detect: A Deep Learning Ensemble Model for Fake News Detection," *Complexity*, vol. 2021, no. 1, 2021, Art. no. 5557784, <https://doi.org/10.1155/2021/5557784>.
- [13] A. Kishwar and A. Zafar, "Fake news detection on Pakistani news using machine learning and deep learning," *Expert Systems with Applications*, vol. 211, Jan. 2023, Art. no. 118558, <https://doi.org/10.1016/j.eswa.2022.118558>.
- [14] M. Sudhakar and K. P. Kaliyamurthi, "Detection of fake news from social media using support vector machine learning algorithms," *Measurement: Sensors*, vol. 32, Apr. 2024, Art. no. 101028, <https://doi.org/10.1016/j.measen.2024.101028>.
- [15] S. Hakak, M. Alazab, S. Khan, T. R. Gadekallu, P. K. R. Maddikunta, and W. Z. Khan, "An ensemble machine learning approach through effective feature extraction to classify fake news," *Future Generation Computer Systems*, vol. 117, pp. 47–58, Apr. 2021, <https://doi.org/10.1016/j.future.2020.11.022>.
- [16] "Fake News Detection Dataset." Kaggle, [Online]. Available: <https://www.kaggle.com/datasets/abdellahelfedg/fake-news-detection-dataset>.
- [17] J. Alghamdi, Y. Lin, and S. Luo, "A Comparative Study of Machine Learning and Deep Learning Techniques for Fake News Detection," *Information*, vol. 13, no. 12, Dec. 2022, <https://doi.org/10.3390/info13120576>.
- [18] S. Kumar and B. Arora, "A Review of Fake News Detection Using Machine Learning Techniques," in *2021 Second International Conference on Electronics and Sustainable Communication Systems (ICESC)*, Aug. 2021, pp. 1–8, <https://doi.org/10.1109/ICESC51422.2021.9532796>.
- [19] A. A. A. Ahmed, A. Aljabouh, P. K. Donepudi, and M. S. Choi, "Detecting Fake News Using Machine Learning: A Systematic Literature Review," *arXiv*, Feb. 08, 2021, <https://doi.org/10.48550/arXiv.2102.04458>.
- [20] M. Lahby, S. Aqil, W. M. S. Yafooz, and Y. Abakarim, "Online Fake News Detection Using Machine Learning Techniques: A Systematic Mapping Study," in *Combating Fake News with Computational Intelligence Techniques*, M. Lahby, A. S. K. Pathan, Y. Maleh, and W. M. S. Yafooz, Eds. Springer International Publishing, 2022, pp. 3–37.
- [21] R. C. Thompson, S. Joseph, and T. T. Adeliyi, "A Systematic Literature Review and Meta-Analysis of Studies on Online Fake News Detection," *Information*, vol. 13, no. 11, Nov. 2022, <https://doi.org/10.3390/info13110527>.
- [22] O. Bashaddadh, N. Omar, M. Mohd, and M. Nor Akmal Khalid, "Machine Learning and Deep Learning Approaches for Fake News Detection: A Systematic Review of Techniques, Challenges, and Advancements," *IEEE Access*, vol. 13, pp. 90433–90466, 2025, <https://doi.org/10.1109/ACCESS.2025.3572051>.
- [23] D. Choudhury and T. Acharjee, "A novel approach to fake news detection in social networks using genetic algorithm applying machine learning classifiers," *Multimedia Tools and Applications*, vol. 82, no. 6, pp. 9029–9045, Mar. 2023, <https://doi.org/10.1007/s11042-022-12788-1>.
- [24] W. H. Bangyal et al., "Detection of Fake News Text Classification on COVID-19 Using Deep Learning Approaches," *Computational and Mathematical Methods in Medicine*, vol. 2021, no. 1, 2021, Art. no. 5514220, <https://doi.org/10.1155/2021/5514220>.
- [25] I. Q. Abduljaleel and I. H. Ali, "Deep Learning and Fusion Mechanism-based Multimodal Fake News Detection Methodologies: A Review," *Engineering, Technology & Applied Science Research*, vol. 14, no. 4, pp. 15665–15675, Aug. 2024, <https://doi.org/10.48084/etasr.7907>.
- [26] R. Jehad and S. A. Yousif, "Classification of fake news using multi-layer perceptron," *AIP Conference Proceedings*, vol. 2334, no. 1, Mar. 2021, Art. no. 070004, <https://doi.org/10.1063/5.0042264>.
- [27] B. Xie and Q. Li, "Detecting fake news by RNN-based gatekeeping behavior model on social networks," *Expert Systems with Applications*, vol. 231, Nov. 2023, Art. no. 120716, <https://doi.org/10.1016/j.eswa.2023.120716>.
- [28] P. Bahad, P. Saxena, and R. Kamal, "Fake News Detection using Bi-directional LSTM-Recurrent Neural Network," *Procedia Computer Science*, vol. 165, pp. 74–82, Jan. 2019, <https://doi.org/10.1016/j.procs.2020.01.072>.
- [29] A. Jarrahi and L. Safari, "Evaluating the effectiveness of publishers' features in fake news detection on social media," *Multimedia Tools and Applications*, vol. 82, no. 2, pp. 2913–2939, Jan. 2023, <https://doi.org/10.1007/s11042-022-12668-8>.
- [30] A. D'Ulizia, M. C. Caschera, F. Ferri, and P. Grifoni, "Fake news detection: a survey of evaluation datasets," *PeerJ Computer Science*, vol. 7, June 2021, Art. no. e518, <https://doi.org/10.7717/peerj-cs.518>.