International Journal of Computing and Artificial Intelligence

E-ISSN: 2707-658X
P-ISSN: 2707-6571
Impact Factor (RJIF): 5.57
https://www.computersciencejournals.com/ijcai/
IJCAI 2026; 7(1): 01-04
Received: 02-11-2025
Accepted: 05-12-2025

Priya Sharma
Integral University, Department of Computer Application, Lucknow, Uttar Pradesh, India

Mohd Waris Khan
Integral University, Department of Computer Application, Lucknow, Uttar Pradesh, India

Corresponding Author:
Mohd Waris Khan
Integral University, Department of Computer Application, Lucknow, Uttar Pradesh, India

Comparative analysis in fake news detection using machine learning techniques

Priya Sharma and Mohd Waris Khan

DOI: https://www.doi.org/10.33545/27076571.2026.v7.i1a.233

Abstract
The rapid proliferation of misinformation on social media has made fake news detection a critical challenge in the digital era. Although recent deep learning-based methods have demonstrated high performance, classical and hybrid machine learning approaches remain highly relevant, particularly in resource-constrained environments. This study presents a comparative analysis of machine learning-based fake news detection approaches reported between 2020 and 2025 that achieve classification accuracies below 92%, with a focus on identifying their strengths and limitations. Building on this analysis, a hybrid classification framework combining Logistic Regression, Random Forest, and XGBoost is proposed. The proposed system achieves an accuracy of 96.96% on the experimental dataset. Furthermore, the factors contributing to the superior performance of the proposed approach are analyzed, its limitations are discussed, directions for future research are outlined, and strategies for mitigating the challenges of fake news detection are subsequently discussed.

Keywords: Fake news, machine learning techniques, support vector machine, deep learning, logistic regression, CNN

1. Introduction

Social media's widespread use has contributed to the spread of fake news, which is a major threat to politics, public health, and public discourse. To assist moderators, fact-checkers, and users in navigating an increasingly loud information environment, automatic detection of such misinformation is crucial. Different machine learning models have been created by researchers, including ensemble or hybrid systems, deep neural networks (LSTM, CNN), and traditional classifiers (SVM, Random Forest, etc).

In spite of many modern pipelines are successful, not all systems report exceptionally high accuracy; instead, they frequently employ less computation, data, or simpler features and are easier to understand or use in limited environments. Furthermore, a comparison of these approaches aids in highlighting the trade-offs between generalizability, complexity, and performance.

In this paper, machine learning-based studies published between 2020 and 2025 are systematically reviewed based on their reported accuracy, and their underlying techniques, advantages, and limitations are comparatively analyzed. These approaches are then contrasted with the proposed high-accuracy hybrid system combining Logistic Regression, Random Forest, and XGBoost, which achieves an accuracy of 96.96% on the dataset.

The remainder of this paper is structured as follows: Section 2 presents a comparative analysis of selected fake-news detection studies published between 2020 and 2025. Section 3 provides a detailed results analysis, including comparisons with prior works as well as a discussion of caveats and risks. Future research directions are outlined in Section 4. Section 5 discusses strategies to mitigate the challenges associated with fake-news detection. Finally, the conclusions are presented in the last section.

2. Comparative Analysis

Table 1 presents a comparative summary of recent (2020-2025) studies on machine learning based fake news detection.

Table 1: Comparative analysis of selected fake-news detection works (2020-2025)

S.N.	Paper (Year)	Technology used	Advantages	Disadvantages	Accuracy (%)
1	Velivela & Kumari (2022) -*Detection of Fake News using Machine Learning Models* [1].	Naïve Bayes, SVM, others, TF-IDF	Very simple setup; interpretable; fast to train	Feature engineering limited; shallow models; may not generalize well	86%
2	Dev et al. (2024) - *Hybrid RFSVM: Hybridization of SVM and Random Forest* [2].	SVM + Random Forest hybrid	Combines linear/separable (SVM) with ensemble robustness (RF)	Complexity increases; risk of overfitting; feature scaling needed	~88%
3	Janssen (2023) -*Comparative analysis of ML algorithms for fake news* [3].	Deep learning (Bi-RNN with LSTM) + classical model	Explores several models; shows deep models' potential	Deep models may overfit; needs more hyperparameter tuning	91% (bidirectional RNN with 4 LSTM layers)
4	IJCRT (2024) - *Fake News Detection Using Deep Learning (LSTM)* [4].	LSTM, also SVM / RF / LR baselines	Captures sequential dependencies; better recall	Requires more data; slower training; may have vanishing gradients	88% (LSTM)
5	Al-Obaidi et al. (2024) -*Automated Fake News Detection System* [5].	Gradient Boosting, Decision Tree, LR, SVM	Several baseline methods; lightweight	Low accuracy; limited feature richness; perhaps small / noisy dataset	82.24% (Gradient Boosting)
6	Ilyas et al. (2024) - *Fake News Detection on Social Media Using Ensemble (CNN + LSTM etc.)* [6].	Ensemble of CNN, LSTM, SVM, RF	Multimodel ensemble; captures feature variety	Deep + classical mix is complex; CNN+LSTM gave poor accuracy	88% (CNN + LSTM ensemble)
7	Al-Tarawneh et al. (2025) - Towards Accurate Fake News Detection: Evaluating Machine Learning Approaches and Feature Selection Strategies [7].	Decision Tree, SVM, MLP, RF, XGBoost	Evaluates many models; shows trade-offs; use of ensemble	Some models underperform; classical models limited by feature set	91.006% (Decision Tree)
8	Saini & Khatarkar (2023) - *A Review on Fake News Detection using Machine Learning* [8].	Survey of algorithms (NB, SVM, RF, LSTM, CNN)	Broad coverage; identifies trends; low barrier to entry	Not all experimental; reported some ML accuracies from literature	~85% (as reported for Random Forest / NB etc.)
9	Hamed & Singh (2023) - 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 [9].	Survey / meta-analysis	Highlights challenges, domain shift, interpretability	Not a new model; no original experiments; broad rather than deep	Reported typical ML accuracy between 70-80%
10	Hoy & Koulouri (2025) - An exploration of features to improve the generalisability of fake news detection models [10].	Logistic Regression, SVM, Decision Tree, RF, Gradient Boosting	Focus on generalisability; interpretability; novel features	Cross-dataset drop in accuracy; token-models less robust	75% cross-dataset accuracy on Facebook URLs

3. Results Analysis

3.1 Methodology: LR + RF + XGBoost

To enhance predictive performance, a hybrid ensemble framework is employed that leverages the complementary strengths of multiple machine learning classifiers [11, 12]. The proposed approach integrates Logistic Regression, Random Forest, and XGBoost, achieving an accuracy of 96.96% on the dataset. This ensemble is designed to balance linear and non-linear feature interactions while maintaining high levels of accuracy, precision, recall, and model interpretability. By combining these classifiers, the framework effectively captures diverse data patterns and improves overall classification robustness. The operational workflow of the proposed framework is described in the following stages:

A. Data Collection

- Real-world fake and real news articles are included in the benchmark dataset FakeNewsNet.
- These databases offer a fair portrayal of news pertaining to politics and entertainment.

B. Data Cleaning

- Prior to feature extraction, data cleaning guarantees consistent and high-quality input.
- The preprocessing procedures listed below are used:
- Convert all text to lowercase for uniformity.
- Remove URLs and non-alphabetical characters (e.g., numbers, special symbols).
- Collapse multiple spaces into single spaces.
- Filter out stop words to reduce noise and focus on meaningful tokens.
- After cleaning, TF-IDF (Term Frequency-Inverse Document Frequency) is used to vectorise the text into numerical representation.

C. Resampling (SMOTE)

- There is frequently a class imbalance in the dataset (e.g., less fake news than true news or vice versa).
- The SMOTE (Synthetic Minority Oversampling Technique) algorithm is used to solve this.
- Text is first converted into TF-IDF vectors because SMOTE only processes numerical input. To ensure a balanced class distribution for training, SMOTE then creates synthetic minority class samples.

D. Hybrid Model (Ensemble)

- For robustness, a hybrid ensemble is built rather than depending on a single classifier.
- The ensemble combines:

- **Logistic Regression (LR):** Favorable for interpretability and linear separability.
- **Random Forest (RF):** Minimizes overfitting and manages non-linear interactions.
- **XGBoost:** An effective gradient boosting method that recognizes intricate patterns.

- To increase detection accuracy, the combined model makes use of each classifier's advantages.

E. Evaluation & Visualization

Several metrics are used to evaluate the hybrid model's performance:

- **Accuracy:** Overall accuracy of forecasts.
- **Precision, Recall, and F1 score:** Crucial in situations involving class imbalance.
- **ROC AUC and PR AUC:** Analyse the discriminative capacity of the model.

Techniques for visualization include:

- **Confusion Matrix:** Displays the distribution of accurate and inaccurate classifications.
- **ROC Curve:** shows the trade-off between the false positive rate and the true positive rate.
- **Precision-Recall Curve:** demonstrates performance on datasets that are unbalanced.

3.2 Comparison with Prior Works

In comparison with the studies reviewed above, the proposed hybrid model demonstrates substantially improved performance, as summarized below:

- Compared to the majority of classical or hybrid ML models in the literature, the 96.96% is far higher. For example, Velivela & Kumari (2022) reported only ~86% [1].
- Even some more intricate or group techniques, like Dev's RF + SVM hybrid, achieved about 88% [2].
- In certain studies, deep learning models like LSTM (2024) only achieved about 88% [4].
- Al-Tarawneh (2025) achieved about 91% with decision trees and simpler models [5].
- Although they have superior domain robustness, models optimized for generalizability (Hoy & Koulouri, 2025) have lower cross-dataset accuracy (~75%) [10].

As a result, the proposed model outperforms many existing approaches, suggesting the following:

- **Effective feature engineering and ensemble synergy:** Combining LR (which manages linear separability), RF (which captures nonlinear and high-order interactions), and XGBoost (strong gradient boosting) may provide complementing strengths that previous works' smaller or single models were unable to completely utilize.
- **Good dataset quality or alignment:** The ensemble model can effectively capture meaningful patterns when the dataset is clean, well-balanced, and closely aligned with the data distributions on which the individual models were trained.
- **Rigorous validation:** The system demonstrates stability and reduced risk of overfitting when the reported accuracy is obtained on a properly held-out test set using appropriate cross-validation, early stopping, and systematic hyperparameter tuning.

3.3 Caveats and Risks

Despite the strong performance, several potential concerns should be considered.

- **Overfitting risk:** Accuracy close to 97% may indicate overfitting, particularly in the presence of data leakage or if the test set is not fully independent of the training data. Therefore, it is essential to report additional evaluation metrics such as precision, recall, F1-score, and confusion matrices alongside accuracy to provide a more comprehensive assessment of model performance.
- **Generalizability:** The extent to which the proposed model generalizes to unseen data, domain shifts, or diverse news categories (e.g., political, health, and social news) remains uncertain. This limitation contrasts with the work of Hoy and Koulouri (2025), which explicitly evaluates cross-dataset performance and reports a significant decline in accuracy, highlighting the challenges of model transferability.
- **Explainability:** Compared to single, straightforward models, ensemble models can be more difficult to understand. Explainability and interpretability the ability to explain why a piece of news is deemed fake may be necessary for practical deployment.
- **Scalability and efficiency**: Three models (LR, RF, and XGBoost) may increase computational costs, memory overhead, and latency during training and deployment, making them unsuitable for real-time or limited systems.

4. Future Work

Based on the promising results of the proposed approach and the identified gaps in existing literature, several directions for future research are suggested.

4.1 Cross-domain and cross-dataset evaluation

- To comprehensively assess the generalizability of the proposed model, future studies should evaluate its performance across multiple datasets representing diverse domains, such as political news, health-related misinformation, COVID-19-related content, and other domain-specific fake news.
- Additionally, manually fact-checked external validation datasets should be employed to strengthen the reliability of the evaluation, including benchmark resources such as the Facebook URLs dataset used by Hoy and Koulouri (2025).

4.2 Feature explainability and interpretability

- To determine which features or ensemble components (LR, RF, and XGBoost) have the greatest influence on decisions, use explanation tools (SHAP, LIME).
- To determine whether bias exists, report class-wise metrics (precision, recall, F1) and confusion matrices (e.g., false negatives/positives).

4.3 Model compression and optimization

- Examine whether performance can be maintained while inference time is shortened using a distilled or lighter ensemble (e.g., trimming RF or boosting).
- Investigate dimensionality reduction and feature selection to streamline the process with little loss.

4.4 Adversarial robustness

- Assess the robustness of the proposed model against adversarial attacks, including paraphrasing, antagonistic commentary, and intentional text manipulations, such as those generated by large language models.
- To enhance resistance to such manipulations, robust learning strategies, adversarial training, and data augmentation techniques can be incorporated.

4.5 Multimodal and propagation-based features

- The proposed system can be extended beyond textual analysis by integrating multimodal information, including social graph signals (e.g., propagation and diffusion patterns), metadata (such as source credibility and author information), and visual content (images or videos, where available).
- Additionally, temporal and early-detection capabilities should be explored to determine whether propagation-based features can enable the early identification of fake news before widespread dissemination.

5. Mitigate the Challenges in Fake News Detection

- Overfitting prevention can be achieved through cross-validation to ensure the model isn't overfitting to a single dataset. Regularization techniques like L1/L2 regularization can be applied to logistic regression and decision trees to avoid complexity. Early stopping during model training can also prevent overfitting, particularly with ensemble models like XGBoost.
- To enhance generalizability, testing the model across various domains (e.g., political news, health misinformation) is crucial. Cross-domain testing ensures the model performs effectively across diverse types of fake news. Domain adaptation techniques or transfer learning can also help fine-tune the model for specific fake news categories.
- Model explainability can be improved by using interpretability tools such as SHAP or LIME to make it easier to understand how decisions are made. Identifying the most influential features using feature importance ranking can ensure the system operates transparently and predictably.
- To defend against adversarial attacks, adversarial training can introduce manipulated or paraphrased content during model training. Data augmentation techniques such as synonym replacement or back-translation can help the model handle such adversarial examples.
- Improving scalability and computational efficiency can be achieved through model optimization techniques like distillation or pruning, which reduce the model size while maintaining or improving accuracy. Smarter feature selection can reduce dimensionality, eliminating irrelevant features and improving training efficiency.
- To enhance fake news detection, integrating social media signals such as user interactions, shares, and sentiment analysis, along with analyzing multimedia content like images, videos, and audio, can provide more context and improve accuracy.

6. Conclusion

This study compared machine learning-based fake news detection systems published between 2020 and 2025 that reported accuracies below 92%. Despite the use of classical classifiers and simple ensemble methods, their performance was often limited by factors such as data quality, feature representation, and overfitting. In contrast, the proposed hybrid ensemble combining Logistic Regression, Random Forest, and XGBoost achieved an accuracy of 96.96%, demonstrating that carefully designed classical and ensemble approaches can outperform many simpler or less optimized systems.

Nonetheless, comprehensive evaluation, including cross-domain testing, interpretability analysis, and robustness assessment, is crucial to assess the model's practical utility. For real-world deployment, future research should focus on improving generalization, explainability, scalability, and resilience to adversarial attacks.

References

1. Lakshmi VD, Kumari CS. Detection of fake news using machine learning models. International Journal of Computer Applications. 2022 Jan; 183 (47):22-27. doi:10.5120/ijca2022921874.
2. Dev DG, Bhatnagar V. Hybrid RFSVM: hybridization of SVM and random forest models for detection of fake news. Algorithms. 2024;17(10):459. doi:10.3390/a17100459.
3. Janssen J. Comparative analysis of machine learning algorithms for fake news detection. 2023.
4. SK, Allada VR, Asif M, Sabeel UR, Shariq MB. Fake news detection using deep learning (LSTM). International Journal of Creative Research Thoughts. 2024;12(5):2320-2882.
5. Al Obaidi SA, Çağlıkantar T. Automated fake news detection system. Iraqi Journal for Computer Science and Mathematics. 2024;5(4). doi:10.52866/27887421.1200.
6. Ilyas MA, Rehman A, Abbas A, Kim D, Naseem MT, Min Allah N. Fake news detection on social media using ensemble classifiers. Computers, Materials and Continua. 2024;81(3):4525-4549. doi:10.32604/cmc.2024.056291.
7. Al-Tarawneh MAB, Al-Khresheh A, Al-Irr O, Kulaglic A, Danach K, Kanj H, et al. Towards accurate fake news detection: evaluating ensemble and classical methods. European Journal of Pure and Applied Mathematics. 2025; 18(2): 6087. doi:10.29020/nybg.ejpam.v18i2.6087.
8. Saini P, Khatarkar V. A review on fake news detection using machine learning. Smart Moves Journal Ijoscience. 2023. doi:10.24113/ijoscience.v9i2.511.
9. Hamed SK, Ab Aziz MJ, Yaakub MR. 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. 2023;9(10):e20382. doi:10.1016/j.heliyon.2023.e20382.
10. Hoy N, Koulouri T. An exploration of features to improve the generalisability of fake news detection models. Expert Systems with Applications. 2025;275:126949. doi:10.1016/j.eswa.2025.126949.
11. Parveen N, Khan MW. Proposed algorithm and models for sentiment analysis and opinion mining using web data. Nanotechnology Perceptions. 2024;20(6):3900-3910.
12. Mishra A, Khan MH, Khan W, Khan MZ, Srivastava NK. A comparative study on data mining approach using machine learning techniques: prediction perspective. In: Pervasive healthcare. Berlin/Heidelberg: Springer; 2022. p. 153-165.