



Research Article

A Comparative Analytical Study of Machine Learning Paradigms for Fake News Detection

Anjali Pal ^{1*},  Amit Jain ²,  Ronak Duggar ³

¹⁻³ School of Computer Science & Engineering, Geeta University, Panipat, India

Corresponding Author: *Anjali Pal

DOI: <https://doi.org/10.5281/zenodo.20033313>

Abstract

The quantity of fake news has increased recently due to the quick development of digital media, which explains why automated detection is more relevant than ever. The study conducts a comparative analytical investigation of four machine learning frameworks—classical machine learning, deep learning, transformer-based models, and multimodal models—for the purpose of detecting fake news. Rather than proposing a new model, the study evaluates existing paradigms in terms of performance, interpretability, scalability, and computational complexity. The analysis indicates that advanced models like transformers are more effective in giving a better understanding of the context, but need more computational efforts and have lower interpretability. The research concludes that fake news detection must also be treated as a multi-dimensional system design issue and not necessarily aiming at classification accuracy only.

Manuscript Information

- ISSN No: 2583-7397
- Received: 10-04-2026
- Accepted: 25-04-2026
- Published: 05-05-2026
- IJCRM:5(3); 2026: 22-26
- ©2026, All Rights Reserved
- Plagiarism Checked: Yes
- Peer Review Process: Yes

How to Cite this Article

Pal A, Jain A, Duggar R. A Comparative Analytical Study of Machine Learning Paradigms for Fake News Detection. Int J Contemp Res Multidiscip. 2026;5(3):22-26.

Access this Article Online



www.multiarticlesjournal.com

KEYWORDS: Fake News Detection, Deep Learning, Misinformation Detection, Multimodal Analysis, Machine Learning, Transformer Models.

I. INTRODUCTION

The fast increase in digital media platforms has reshaped information dissemination and consumption, however, also contributed to the dissemination of false information. Nowadays, false information might be mistaken for authentic content, making it difficult to distinguish between the two. To address this problem, machine learning techniques have been widely applied, including automatic fake news detectors. Early algorithms were built on traditional models, such as Support Vector Machines and Naive Bayes, which rely on textual features rather than context.

Deep learning models enhanced both feature representation and feature detection, but increased the level of computational complexity [11], [12].

Transformer-based models have recently been used with higher contextual abilities, but with greater computation and low interpretability [2], [3], [14]. There is also recent research that examines the multimodal methods that combine both textual and visual data to boost the accuracy of detection [5],[7]. Nevertheless, a lot of the available techniques rely on domain-related data, which constrain their application in other situations [15].

II. LITERATURE REVIEW

Initial research on fake news detection models was heavily text-based, classic machine learning models that performed reasonably but were not context sensitive [8], [10]. These plans were good and founded on the superficial trends of linguistics. With the further development of deep learning, feature representations within models, i.e., CNNs and LSTMs, kept on enhancing hierarchical and sequential information, leading to better performance in the detection of features [11], [12]. More computational rates and lesser interpretability were linked to those models, however. Another development in fake news detection was based on transformer models that applied their contextual dependencies as attention to improve their performance compared to the former models [2], [3]. Even though these models are useful, they require significant computing power, and their scalability is an issue [14].

The most recent research has been on multimodal, which involves textual and visual elements to enhance detection accuracy [5]-[7]. Whereas these techniques offer more contextual insight, they make the systems more complex and difficult to manage in terms of big data.

Table 1: Overview of Literature Review

Study	Model	Contributions	Limitations
1	Cross-feature ambiguity learning	Enhances contextual representation using shared feature relationships	Increased model complexity
2	FakeBERT (Transformer)	Combines style analysis and credibility verification for improved detection	High computational cost
3	Triple-attention mechanism Systematic review of detection algorithms	Improves contextual feature extraction in social media data Gives a summary of ML and DL techniques for identifying bogus news.	Limited cross-domain evaluation Lack of a unified evaluation framework
4	Multimodal feature fusion	Combines textual and visual features to improve fake news detection accuracy	Higher computational complexity
5	Multimodal contrastive learning	Improves multimodal representation using contrastive learning	Requires large-scale data
6	SVM-based detection	Efficient supervised learning with high classification accuracy	Requires manual feature engineering
7	ML vs DL comparative model	Shows deep learning improves performance over classical methods	Higher training complexity
8	ML & DL systematic review	Summarises content-based detection approaches	Lack of standardised benchmarks
9	Benchmark ML study	Compares the performance of ML models across datasets	Limited cross-domain generalisation

The key developments and constraints of the current research on the topic of false news identification are summarised in Table 1. It points out that although new models are more likely to enhance performance, they tend to complicate and become less understandable.

A. Research Gap

Despite the fact that many works propose machine learning and deep learning models in order to detect fake news, most of the studies are primarily focused on improving the accuracy of the classification. Expressiveness in other areas has been given little attention, such as the scalability, interpretability and computational efficiency of these models. Additionally, the existing technologies rely mostly on the domain-specific datasets, particularly the political news corpora and therefore lack generality when applied to a broad spectrum of domains such as health, technology and multilingual contents. The other

The challenge is that no standard evaluation structures exist where a fair comparison may be made in different approaches to detection.

Therefore, machine learning paradigms must somehow be presented in contrast to the strengths and weaknesses and their application with the practical use to understand their applicability to the context of fake news detecting systems.

III. RESEARCH METHODOLOGY

A. Problem Definition

Fake news detection is defined as a binary classification problem where a news piece is classified as either phoney or real. This paradigm is formulated to capture contextual and linguistic patterns, and the study does not propose a new model, but makes comparisons between existing machine learning paradigms. The most important thing is to get context patterns

on the one hand and to be interpretable and scalable on the other hand.

B. Dataset Description

Table 2: Dataset Overview

Category	Description
Total Articles	44,897
Real News Articles	21,416
Fake News Articles	23,480
Features	Title, Text, Subject, Date

The publicly available Kaggle dataset presented in Table 2 is used in this work; it contains 44,898 labelled news articles, and its real and fake classes are also almost balanced, which contributes to reducing model bias and providing a reliable evaluation. The dataset features title, full text, subject, and date, where the textual content makes inference possible to extract semantic, syntactic, and contextual patterns and might be used in both traditional and modern machine learning methods of classifying fake news.

C. Evaluation Metrics

Accuracy, precision, recall, and F1-score are the metrics most frequently used to assess the models' performance. The measurements are frequently employed to assess a model's performance, particularly in classification-related issues [11], [12].

$$F1 = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$$

Among them, the F1-score is particularly helpful since it strikes a compromise between recall and precision. Recall is used to demonstrate how many of the actual positive cases that are actually anticipated as positive are also predicted as positive, while precision is used to demonstrate how many of the cases that the model projected to be positive are actually true. Because the F1-score considers both of these factors, it is more useful in situations where the dataset is lopsided, given that both false positives and false negatives are taken into account.

IV. Analysis

This section will compare various machine learning paradigms for false news detection while keeping some important characteristics of such a system, including robustness, scalability, interpretability, computational complexity, and performance. Conventional machine learning models, such as Support Vector Machines and Naive Bayes, are very interpretable and inexpensive to compute. Nevertheless, they are dependent on human qualities and are unable to learn more profound contextual correlations of textual material [8], [10]. Because deep learning models develop sequential and hierarchical text representations, they are utilised to enhance detection results. These models, however effective, are more computationally consuming, and their interpretability is limited [11], [12]. Transformer-based models offer better contextual information with attention mechanisms, which are highly accurate when it comes to fake news detection tasks [2], [3].

But their complexity is very high with input and hence cannot be used in large-scale real-time applications [14].

Table 3: The Comparative Paradigm Analysis

Paradigm	Models	Complexity	Scalability
Classical ML	SVM, Naive Bayes	Low	High
Deep Learning	CNN, LSTM	Medium	Moderate
Transformers	BERT, FakeBERT	High	Limited
Multimodal	Text + Image	Very High	Moderate

Fig 1: Comparative computational complexity of machine learning paradigms

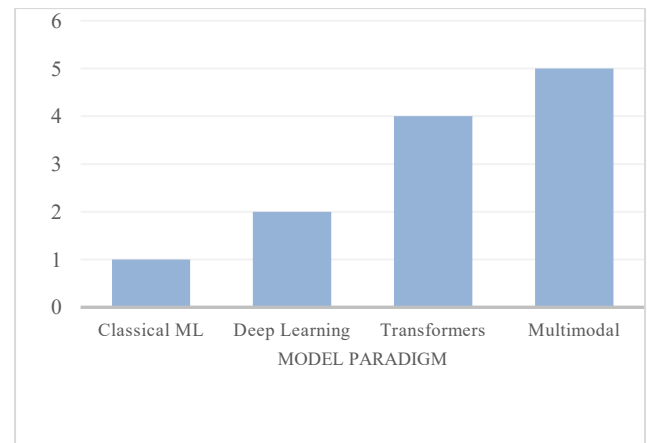


Table 3 compares the top machine learning paradigms on a range of key dimensions, including complexity and scalability, showing that model complexity can make the system better and therefore improved, but also compromises scalability and computing efficiency. The above diagram illustrates the growing complexity of many computations at various paradigms (Fig. 1). Classical models are by far less complicated and multimodal, and transformer-based models require much more computational power.

V. Conceptual Framework

As the review above demonstrates, the optimal machine learning paradigm does not exist, depending on all aspects of evaluation. The two methods entail compromises of performance, interpretability, computational complexity and scalability. As such, it is not enough to choose a detection model regardless of its classification accuracy [4], [14].

To resolve this weakness, a conceptual analytical system proposal is suggested to be used in providing a systematic and context fiction-sensitive selection of fake news detection methods. The structure brings together important elements, like the data representation, the model properties, and the deployment limitations, to allow setting the decision-making to be balanced.

A. Framework Overview

The framework is divided into three major elements:

1. Representation Layer: focuses on data representation that comprises statistical data representation methods (e.g. TF-IDF) and deep learning, transformer-based representations [2], [11].

2. Decision Layer: The decision layer is composed of such classification models as classical machine learning models, deep neural networks, and transformers.

3. Deployment Layer: It is a layer that considers the use of practical constraints, such as within the real world, in practice, such as cost of computation, scalability and interpretability are practical constraints [4], [14].

B. Multi-Dimensional Evaluation

A multi-dimensional evaluation method is used to rank different paradigms on the basis of different factors according to performance, interpretability, complexity of computations, scalability, robustness or adaptability.

Table 4: Multi-Dimensional Comparative Matrix

Dimension	Classical ML	Deep Learning	Transformers	Multimodal
Performance	Moderate	High	Very High	Very High
Interpretability	High	Low	Extremely Low	Extremely Low
Complexity	Extremely low	Moderate	Extremely High	Very High
Scalability	High	Moderate	Limited	Moderate
Robustness	Moderate	Data-dependent	Context-sensitive	Platform-dependent
Adaptability	Low	Moderate	Moderate	Moderate

The table presents various machine learning paradigms in comparative terms, which allows one to easily visualise the trade-offs between performance, interpretability, complexity and scalability.

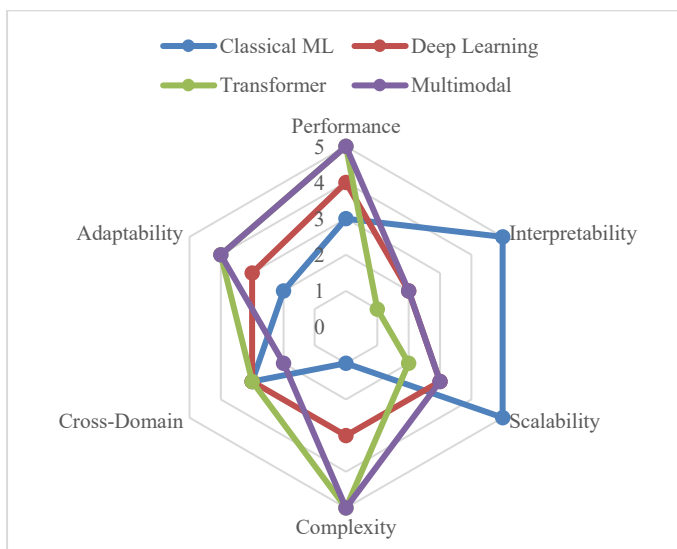


Fig. 2. Comparison of Fake News Detection Paradigms in Multi-Dimension

As illustrated in the diagram above, transformer and multimodal models are more efficient and scalable (classical models) but less interpretable and more computationally expensive. Consequently, only practical deployment needs must be used to select a model that has been practised, not just based on accuracy.

VI. DISCUSSION

The findings of the present paper indicate that the issue of identifying false news should be considered a systemic problem, and not as a task of classification. Other more advanced frameworks, such as deep learning and transformer style, are more precise, but also generate computational and model interpretability problems [2], [11], [14]. Such limitations may interfere with their real-life application in large and time-constrained settings. Classical machine learning models are not particularly effective in the contextual information operation, yet it is also useful, given their efficiency, scalability, and transparency [8], [10]. This renders them to be applied in areas where computer power and interpretability are important. The reliance on specific datasets based on a domain is another valuable observation as well. Most models - even the ones that excel in benchmark datasets - cannot be generalised between domains (health, finance, or multilingual content) [15]. Through this, it becomes clear that there is a need to strengthen and make detection systems more versatile. Moreover, the growing adoption of multimodal solutions enhances the detection performance at the cost of further complications and increased data specifications [5],[7]. With the constant transformation of misinformation, the detection systems will also be required to keep pace with the new trends and forms of content. On the whole, the findings indicate that it is necessary to focus on finding the balance of performance, interpretability, scalability, and practicality in fake news detectors in their design.

VII. CONCLUSION

In the present paper, a comparative analytical study of machine learning paradigms for fake news detection methodology, including classical machine learning, deep learning, transformer-based models, and multimodal models, was presented. In place of the model development, the study was able to perform the analyses of these paradigms along major factors that included performance, interpretability, scalability, and computational complexity. It is found that more sophisticated models are more successful in understanding the context and are more accurate, but have enormous demands on calculation processes and are less understandable. Unlike modern machine learning models, classical machine learning models are efficient, scalable, and more transparent, and thus can be deployed in resource-absorbing settings. The paper shows that the detection of fake news is not a problem of accuracy optimisation but also a multi-dimensional system design. The choice of the model is to be based on the conditions of application, resources, and restrictions during the deployment. The future research must be based on advancing cross-domain ability, better interpretability of advanced models, and working on adaptive systems that could address changing patterns of misinformation.

REFERENCES

1. Cui S, Duan K, Hao Z, Shinnou H. Fake news detection based on shared information and cross-feature ambiguity learning. Eng Appl Artif Intell. 2026;163:112841.
2. Tanaja RN, Johnny, Rafif MA, Santoso Gunawan AA. Fake news detection using machine learning: Integrating

- FakeBERT classification, style analysis, and credibility verification. *Procedia Comput Sci.* 2025;269:1067-1076.
3. Jin H, Wang P. Fake news detection on social media using a triple-attention mechanism optimised by an advanced tailor optimisation algorithm. *Egypt Inform J.* 2025;32:100815.
 4. Dal Forno AJ, Richetti GP, Knaesel VH. Fake news detection algorithms: A systematic literature review. *Data Knowl Eng.* 2025;158:102441.
 5. Yan F, Zhang M, Wei B, Ren K, Jiang W. FMC: Multimodal fake news detection based on multi-granularity feature fusion and contrastive learning. *Alex Eng J.* 2024;109:376-393.
 6. Shen X, Huang M, Hu Z, Cai S, Zhou T. Multimodal fake news detection with contrastive learning and optimal transport. *Front Comput Sci.* 2024;6:1473457.
 7. Chen H, et al. A self-learning multimodal approach for fake news detection. *Front Artif Intell.* 2025;8:1665798.
 8. Sudhakar M, Kaliyamurthi KP. Detection of fake news from social media using support vector machine learning algorithms. *Meas Sens.* 2024;32:101028.
 9. Park M, Chai S. Constructing a user-centred fake news detection model using classification algorithms in machine learning techniques. *IEEE Access.* 2023;11:71517-71527.
 10. Arowolo MO, Misra S, Ogundokun RO. A machine learning technique for the detection of social media fake news. *Int J Semant Web Inf Syst.* 2023;19(1):1-25.
 11. Kishwar A, Zafar A. Fake news detection on Pakistani news using machine learning and deep learning. *Expert Syst Appl.* 2023;211:118558.
 12. Saeed A, Solami EA. Fake news detection using machine learning and deep learning methods. *Comput Mater Contin.* 2023;77(2):2079-2096.
 13. Reddy J, Mundra S, Mundra A. Ensembling deep learning models for fake news classification. *Procedia Comput Sci.* 2024;235:2766-2774.
 14. Capuano N, Fenza G, Loia V, Nota FD. Content-based fake news detection with machine learning and deep learning: A systematic review. *Neurocomputing.* 2023;530:91-103.
 15. Khan JY, Khondaker MTI, Afroz S, Uddin G, Iqbal A. A benchmark study of machine learning models for online fake news detection. *Mach Learn Appl.* 2021;4:100032.

Creative Commons (CC) License

This article is an open-access article distributed under the terms and conditions of the Creative Commons Attribution–Non-Commercial–No Derivatives 4.0 International (CC BY-NC-ND 4.0) license. This license permits sharing and redistribution of the article in any medium or format for non-commercial purposes only, provided that appropriate credit is given to the original author(s) and source. No modifications, adaptations, or derivative works are permitted under this license.

About the corresponding author



Anjali Pal is associated with the School of Computer Science & Engineering at Geeta University, Panipat, India. Her academic interests include artificial intelligence, data science, and software development. She is committed to research and innovation, contributing to emerging technologies through analytical skills, technical expertise, and a dedication to advancing computing and engineering fields.