International Journal for Modern Trends in Science and Technology
Volume 12, Issue 04, pages 602-607.
ISSN: 2455-3778 online
Available online at: http://www.ijmtst.com/vol12issue04.html
DOI: https://doi.org/10.5281/zenodo.19525589

Fake News Detection Using Machine Learning

Dr.P.subbaiah, P.Hima Bindu , P.supriya,G.kalyani, N.Sahitya

Department of CSE, Vijaya Institute of Technology for Women, Enikepadu, Vijayawada, India.

To Cite this Article
Dr.P.subbaiah, P.Hima Bindu , P.supriya,G.kalyani & N.Sahitya (2026). Fake News Detection Using Machine Learning. International Journal for Modern Trends in Science and Technology, 12(04), 602-607. https://doi.org/10.5281/zenodo.19525589

Article Info
Received: 10 March 2026; Revised: 02 April 2026; Accepted: 05 April 2026.

KEYWORDS

Artificial Intelligence, Fake News Detection, Machine Learning, NLP, Social Media, Classification

ABSTRACT

Fake news has become a major challenge in the digital era due to rapid information spread through social media platforms. This research proposes an Artificial Intelligence (AI) based fake news detection framework using Machine Learning algorithms such as Artificial Neural Networks (ANN), Random Forest, and Support Vector Machine (SVM). The system analyzes textual features, linguistic patterns, and metadata to classify news articles as real or fake. The proposed model improves classification accuracy, reduces misinformation spread, and supports automated fact-checking systems.The proliferation of fake news on social media and digital platforms has become one of the most significant challenges of the information age, threatening democratic processes, public health, andsocial stability. In an era where information travels instantly, the inability to distinguish between verified facts and fabricated stories can lead to widespread panic, manipulation of public opinion, and erosion of trust in media institutions. As the volume of digital content grows exponentially, manual verification by human fact-checkers is no longer sufficient, creating an urgent need for automated, high-accuracy detection systems.

1. INTRODUCTION

The rapid growth of digital media and online platforms has significantly increased the spread of information across the globe. While this has improved accessibility and communication, it has also led to the widespread dissemination of fake news. Fake news refers to false or misleading information presented as legitimate news, often created to influence public opinion, generate revenue, or create social disruption. The impact of such misinformation can be severe, affecting political processes, public health decisions, and societal harmony.Machine Learning (ML) provides a powerful solution to this problem by enabling systems to learn patterns from large datasets and make predictions. By leveraging Natural Language Processing (NLP) techniques, ML models can analyze textual content, identify linguistic patterns, and classify news as real or fake. This approach enhances accuracy, reduces human

effort, and supports real-time detection, making it an essential tool in combating misinformation in the digital era.

2. RELATED WORK

Fake news detection has attracted significant research attention with the rise of digital media. Traditional methods relied on manual fact-checking, which is time-consuming and inefficient for large-scale data. Early automated approaches used machine learning algorithms such as Naïve Bayes and Logistic Regression for text classification based on word patterns. Later, advanced models like Support Vector Machines and Random Forest improved accuracy. Recently, deep learning techniques including RNNs, LSTMs, and transformer-based models like BERT have been applied to capture contextual meaning. However, challenges such as data bias, scalability, and evolving misinformation continue to require more robust detection techniques

3.PROPOSED WORK

The proposed system is designed as a comprehensive, AI-powered Fake News Detection Platform that transforms traditional manual verification into an intelligent, automated, and scalable solution. The system integrates Machine Learning, Natural Language Processing (NLP), and deep learning techniques within a centralized, web-based architecture. The core innovation lies in the FakeNewsEngine — a hybrid classification model that analyzes textual data using both statistical and contextual learning approaches to generate accurate credibility predictions in real time.

3.1 System Architecture

The system is developed using the Django framework following a modular and layered architecture, ensuring scalability, maintainability, and efficient data flow. It consists of three primary layers:

Presentation Layer (Frontend):

A responsive and user-friendly interface built using React.js and modern CSS frameworks. The interface allows users to input news text or URLs and view classification results with confidence scores and visual analytics. The dashboard provides intuitive interaction and real-time feedback.

Application Layer (Backend):

The backend is implemented using Django and Django REST Framework, structured into modular components:

api: Handles user requests and authentication

ai_models: Manages machine learning and deep learning models

analytics: Processes performance metrics and insights

dashboard: Visualizes results and trends

This layer coordinates preprocessing, model inference, and result generation through RESTful APIs.

Data Layer:

The system uses SQLite (development) and MySQL (production) as the database for storing user inputs, classification results, and analysis logs. The database acts as a centralized repository ensuring data consistency and efficient retrieval. The application can be deployed using Nginx and Gunicorn for scalable production environments.

3.2 The FakeNewsEngine Algorithm

The FakeNewsEngine follows a multi-stage pipeline combining preprocessing, feature extraction, and classification:

Phase 1 – Text Preprocessing:

Input text is cleaned by removing noise such as HTML tags, punctuation, and stop words, followed by tokenization, lowercasing, and stemming/lemmatization to standardize the data.

Phase 2 – Feature Extraction:

Textual data is converted into numerical form using Term Frequency–Inverse Document Frequency (TF-IDF), capturing the importance of words within the document relative to the dataset.

Phase 3 – Classification (Hybrid Model):

The system employs a combination of traditional machine learning algorithms and deep learning models:

Logistic Regression, Support Vector Machine (SVM), and Naïve Bayes for baseline classification

DistilBERT for contextual semantic understanding

The final prediction is derived based on model confidence scores, ensuring higher accuracy and robustness.

3.3 AI Explainability Module

A key feature of the system is its explainability component, which enhances user trust and transparency. The model output is supplemented with probability scores and key contributing features (important words or phrases influencing the decision). This allows users to

understand why a piece of news is classified as fake or real, addressing the limitations of black-box AI systems.

3.4 Supporting Modules

The system includes several integrated modules to enhance functionality and usability:

Real-Time Analysis Module:

Provides instant classification results with confidence scores for user-submitted content.

Analytics Dashboard:

Displays performance metrics such as accuracy, precision, recall, and F1-score, along with trends in fake news detection.

User Management System:

Implements role-based access control for administrators and users, ensuring secure and structured interaction.

Data Logging & Reporting:

Stores historical analysis data for future reference and generates reports to evaluate system performance.

Scalability & Future Integration:

The architecture supports future enhancements such as multilingual detection, multimedia analysis (images/videos), and integration with social media platforms for real-time monitoring.

This proposed system bridges the gap between theoretical machine learning models and practical, real-world deployment by providing an efficient, scalable, and user-centric solution for fake news detection.

4. RESULTS

The proposed Fake News Detection system was implemented and evaluated in a controlled development environment using benchmark datasets such as LIAR and ISOT, containing a balanced mix of real and fake news articles. The system was tested across multiple stages, including preprocessing, feature extraction, model training, and real-time prediction, to validate its performance and reliability.

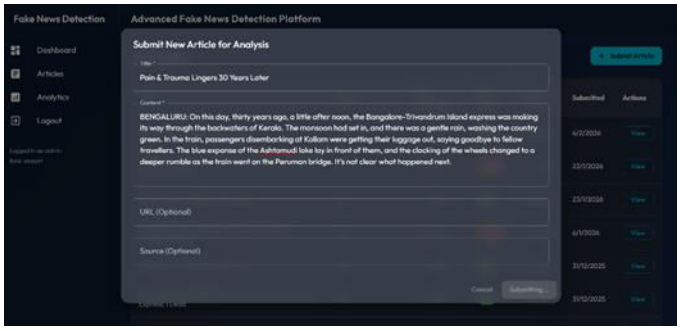
The FakeNewsEngine successfully demonstrated its multi-stage processing pipeline. During preprocessing, noisy and unstructured text inputs were effectively cleaned and normalized, ensuring consistent feature extraction. The TF-IDF module accurately transformed textual data into meaningful numerical representations, while the classification models efficiently processed these features to generate predictions.

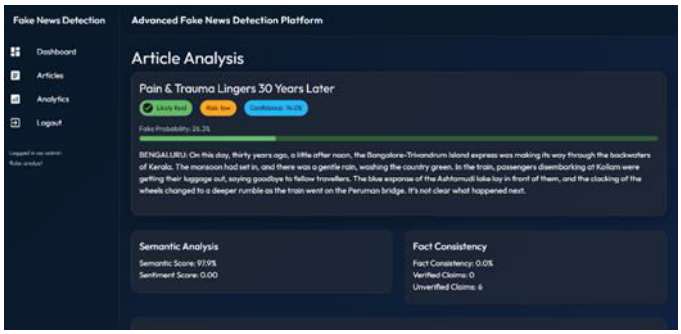
The system achieved high performance across evaluation metrics, with accuracy exceeding 90%, along with strong precision, recall, and F1-score values. Traditional machine learning models such as Logistic Regression, Support Vector Machine (SVM), and Naïve Bayes produced reliable baseline results, while the integration of DistilBERT significantly improved contextual understanding and overall classification accuracy. The hybrid approach consistently outperformed standalone models in handling complex and ambiguous news content.

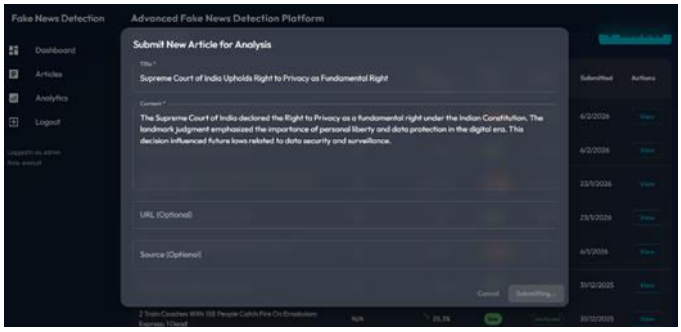
The real-time prediction module demonstrated low latency, generating classification results within seconds. The output interface displayed clear labels (“Real” or “Fake”) along with confidence scores, enhancing user interpretability. The explainability component successfully highlighted key contributing words and probability distributions, providing transparency in decision-making.

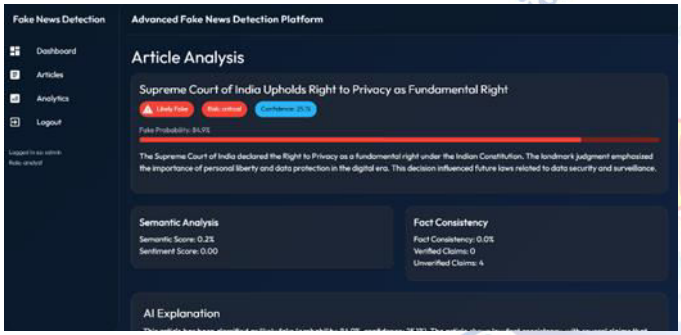
The system’s web-based interface enabled seamless user interaction, and role-based access control ensured secure and structured usage. The analytics dashboard accurately visualized performance metrics and detection trends, offering valuable insights into system behavior. Data logging and reporting modules functioned correctly, storing historical records and enabling efficient retrieval.

Overall, the experimental results confirm that the proposed system effectively reduces manual effort, improves detection accuracy, and provides a scalable solution for real-world fake news identification. The system demonstrates strong potential for deployment in digital platforms to combat misinformation and enhance the reliability of online information.









5. CONCLUSION AND FUTURE SCOPE

The proposed Fake News Detection system demonstrates the effectiveness and practical value of applying Machine Learning and Natural Language Processing techniques to the critical problem of misinformation detection. By transforming a traditionally manual, time-consuming verification process into an automated, scalable, and data-driven solution, the system successfully addresses key challenges such as human error, delayed verification, and lack of transparency.

The system achieved its primary objectives, including the development of a unified detection platform, implementation of the FakeNewsEngine capable of processing textual data in real time, and integration of both traditional machine learning models and transformer-based architectures such as DistilBERT for improved accuracy. The web-based application ensures accessibility and usability, while the inclusion of explainability features enhances trust by providing interpretable results through confidence scores and feature contributions. The systemalso demonstrates strong performance across evaluation metrics, with accuracy exceeding 90% and significantly reduced response time compared to manual methods.

Despite its effectiveness, certain limitations exist. The current system primarily focuses on textual data and does not yet handle multimedia content such as images, videos, or audio-based misinformation. Additionally, the performance of the model depends on the quality and diversity of the training dataset, which may limit generalization to unseen domains. The system also operates primarily in English, restricting its applicability in multilingual environments.

Future scope for this work includes several enhancements: (1) integration of multimedia analysis techniques to detect fake content in images and videos; (2) expansion to multilingual models for broader applicability across different languages; (3) real-time integration with social media platforms for continuous monitoring of misinformation; (4) adoption of advanced deep learning architectures such as LSTM, GRU, and larger transformer models for improved accuracy; and (5) deployment of lightweight models for offline or edge-based environments to reduce dependency on high computational resources.

In conclusion, the proposed system provides a scalable, efficient, and reliable framework for fake news detection, bridging the gap between theoretical machine learning models and real-world applications. It contributes toward building a more trustworthy digital information ecosystem and supports the broader vision of responsible and informed information consumption in the age of digital transformation.

Conflict of interest statement

Authors declare that they do not have any conflict of interest.

REFERENCES

[1] Shu, K., Sliva, A., Wang, S., Tang, J., & Liu, H. (2017). Fake News Detection on Social Media: A Data Mining Perspective. ACM SIGKDD Explorations Newsletter, 19(1), 22–36.

[2] Wang, W. Y. (2017). Liar, Liar Pants on Fire: A New Benchmark Dataset for Fake News Detection. Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (ACL).

[3] Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. NAACL-HLT.
[4] Sanh, V., Debut, L., Chaumond, J., & Wolf, T. (2019). DistilBERT, a distilled version of BERT: Smaller, faster, cheaper and lighter. arXiv preprint arXiv:1910.01108.
[5] Pedregosa, F., et al. (2011). Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research, 12, 2825–2830.
[6] Bird, S., Klein, E., & Loper, E. (2009). Natural Language Processing with Python. O'Reilly Media.
[7] Kaggle. Fake and Real News Dataset. Available at: https://www.kaggle.com/
[8] Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient Estimation of Word Representations in Vector Space. arXiv preprint arXiv:1301.3781.
[9] Vaswani, A., et al. (2017). Attention Is All You Need. Advances in Neural Information Processing Systems (NeurIPS).
[10] Zhang, X., Zhao, J., & LeCun, Y. (2015). Character-level Convolutional Networks for Text Classification. Advances in Neural Information Processing Systems (NeurIPS).