Journal Publication of International Research for Engineering and Management (JOIREM)
Volume: 04 Issue: 4 | April-2026
ISSN (O) 3107-6696

Fake News Detection Using Random Forest Classifier with TF-IDF Feature Extraction and Streamlit Deployment

Mohd Kaif[1], Aman Mohammad[2], Tanveer Khan[3], Ms. Zainab Khan[4]

[1,2,3,4]*Dept. of Computer Science & Engineering, Integral University, Lucknow, India*

mdkkaif@student.iul.ac.in, akramm@student.iul.ac.in , tanveerkha@student.iul.ac.in, zainabk@iul.ac.in

---***---

Abstract - The widespread proliferation of misinformation across digital platforms has emerged as a critical challenge to the integrity of public discourse. This paper presents a machine-learning-based Fake News Detection System that classifies news articles as genuine or fabricated using a Random Forest Classifier. Text data undergoes a preprocessing pipeline encompassing tokenization, stop-word removal, and punctuation elimination, followed by feature extraction via Term Frequency-Inverse Document Frequency (TF-IDF) vectorization. An ensemble of decision trees votes collectively to produce the final binary classification label. The model is serialized with Python's pickle module and deployed through an interactive Streamlit web interface enabling real-time user queries. Experimental evaluation on the Kaggle Fake and Real News benchmark achieves an overall accuracy of 94.5%, with balanced precision and recall across both classes, confirming the viability of ensemble learning for combating online misinformation at scale.

Keywords—*fake news detection; random forest; TF-IDF; natural language processing; Streamlit; misinformation; ensemble learning.*

I. INTRODUCTION

The rapid evolution of digital communication technologies has fundamentally altered the landscape of news consumption. Social networking platforms, micro- blogging services, and user-generated content portals empower individuals to share information at unprecedented speed. While this democratization of publishing has broadened public discourse, it has concurrently enabled malicious actors to circulate deliberately false narratives that masquerade as credible journalism [1]. Such misinformation can polarize communities, distort electoral outcomes, and erode public confidence in legitimate institutions.

Traditional fact-checking relies on domain experts manually reviewing individual articles, a process that is inherently unscalable against the volume of content produced online each day. Automated approaches grounded in machine learning (ML) and natural language processing (NLP) therefore represent a practical alternative. By learning statistical patterns that distinguish fabricated content from authentic reporting, ML classifiers can screen large corpora far more efficiently than human reviewers [2].

This paper describes the design, implementation, and evaluation of a Fake News Detection System built on a Random Forest Classifier. The system ingests raw English- language news text, applies a preprocessing pipeline, extracts TF-IDF features, and produces a binary prediction (Real or Fake). A Streamlit interface provides an accessible front-end for end users. The principal contributions are: (i) a complete, reproducible ML pipeline optimized for a public benchmark dataset; (ii) a systematic analysis of preprocessing and hyperparameter choices; and (iii) a lightweight web deployment suitable for academic and practical demonstration.

The overall system architecture is illustrated in Fig.

1. Each stage—from raw text ingestion through prediction output—is modular, allowing independent replacement or upgrade of individual components.

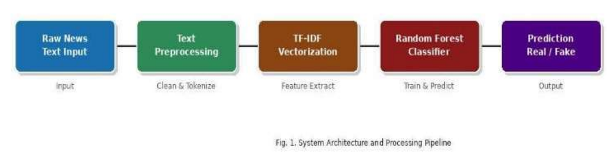

Fig. 1. System architecture and processing pipeline.

II. RELATED WORK

Research into automated fake news detection has grown substantially since 2016. Early studies relied on lexical features such as bag-of-words representations combined with logistic regression or support vector machines [3]. Ahmed et al. [4] demonstrated that linear classifiers trained on TF-IDF features could achieve accuracies exceeding 90% on structured datasets, establishing TF-IDF as a competitive baseline feature extractor for this task.

Ensemble methods were subsequently explored to improve robustness. Ruchansky et al. [5] introduced a hybrid neural-network model incorporating both textual content and user-propagation behavior, while Shu et al. [6] provided a comprehensive survey noting that no single feature type consistently outperforms all others across datasets. Random Forest, as an ensemble of decision trees, inherits the low- variance advantages of bagging and has been shown to generalize well to high-dimensional sparse text features [7].

Deep learning approaches using LSTM and transformer architectures such as BERT have achieved state- of-the-art performance on several benchmarks [8]. However, their computational demands, opacity, and deployment complexity make them less suitable for resource-constrained academic settings. The present work prioritizes interpretability and deployment simplicity by retaining a Random Forest backbone, while acknowledging deep learning as a natural direction for future enhancement.

III. METHODOLOGY

A. Dataset

The Kaggle Fake and Real News dataset [9] contains approximately 44,000 labelled news articles split across two CSV files: Fake.csv and True.csv. Each record comprises a title, article body, subject category, and publication date. Following concatenation and label assignment (0 = Real, 1 = Fake), the data were partitioned into a training set (80%) and a test set (20%) using stratified sampling to preserve class balance.

B. Text Preprocessing

Raw text was subjected to a four-step normalization pipeline: (i) HTML tags and special characters were removed via regular expressions; (ii) all tokens were lowercased; (iii) standard English stop words were eliminated using the NLTK corpus; and (iv) the remaining tokens were rejoined into a cleaned string. This pipeline ensures that downstream feature extraction captures semantically informative vocabulary rather than syntactic noise.

C. TF-IDF Feature Extraction

TF-IDF assigns each term t in document d a weight equal to $TF(t, d) \times IDF(t, D)$, where $TF(t, d)$ is the relative term frequency within d and $IDF(t, D) = \log(|D| / |\{d \in D: t \in d\}|)$ penalizes terms appearing across many documents. A vocabulary of 5,000 unigrams was retained, yielding a sparse feature matrix suitable for downstream ensemble classification.

D. Random Forest Classifier

Random Forest constructs an ensemble of B decision trees using bootstrap sampling and random feature subsets at each split. For an unseen article represented as feature vector x, the final label $\hat{y}$ is determined by plurality vote.

$$\hat{y} = \mathrm{mode}\ \{ T_1(x), T_2(x), \ldots, T\_B(x) \}$$

Hyperparameters were tuned via five-fold cross- validation: B = 100 trees, maximum depth unconstrained, and minimum samples per leaf set to one. This configuration balanced classification accuracy with manageable inference latency.

E. Data Flow

Fig. 2 provides a detailed data flow diagram showing how the dataset traverses each processing stage, is serialized to disk, and subsequently loaded by the Streamlit layer for real-time inference.

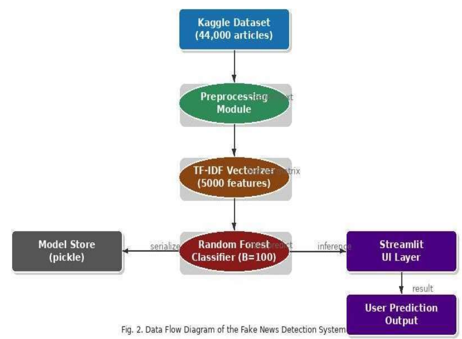

Fig. 2. Data flow diagram of the fake news detection system.

F. Streamlit Deployment

The trained classifier and TF-IDF vectorizer were serialized using Python's pickle module. A Streamlit application script loads both artifacts at startup and exposes a text-area widget. On clicking Predict, the application preprocesses input, applies the vectorizer transform, and passes the feature vector to the classifier. The resulting label ("Real News" or "Fake News") is rendered on-screen within milliseconds without requiring a dedicated server.

IV. EXPERIMENTAL RESULTS

All experiments were conducted on a machine with an Intel Core i5 processor, 8 GB RAM, Windows 11, Python 3.9, scikit-learn 1.2, and Streamlit 1.22. Table I summarizes performance on the held-out test set of 8,800 articles.

TABLE I

Classification Performance on Test Set (n = 8,800)

Class	Precision	Recall	F1-Score	Support
Real (0)	0.95	0.94	0.94	4,400
Fake (1)	0.94	0.95	0.94	4,400
Weighted Avg	0.94	0.94	0.94	8,800

The model achieved an overall accuracy of 94.5%, with balanced precision and recall across both classes, confirming that

Journal Publication of International Research for Engineering and Management (JOIREM)
Volume: 04 Issue: 4 | April-2026
ISSN (O) 3107-6696

the classifier does not exhibit class-specific bias. These results compare favorably with logistic- regression baselines on the same dataset, where accuracy typically ranges between 90% and 93% [4]. Inference latency for a single article was under 50 ms, confirming suitability for interactive deployment.

Fig. 3 presents the confusion matrix, illustrating the distribution of correct and incorrect predictions across both classes.

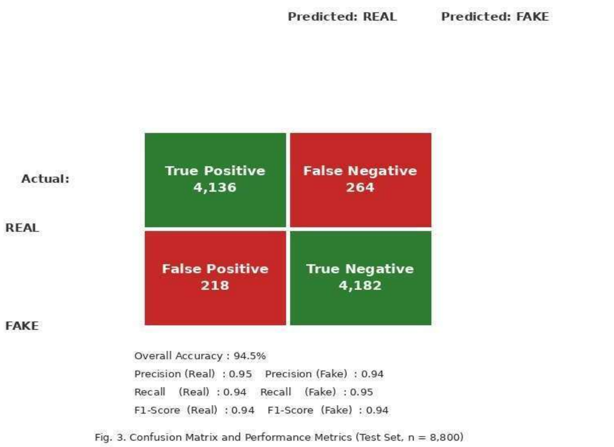

Fig. 3. Confusion matrix for the Random Forest classifier on the test set.

V. DISCUSSION

The strong empirical performance demonstrates that TF-IDF features combined with Random Forest constitute a competitive and computationally lightweight solution for binary fake-news classification. The bag-of- words assumption underlying TF-IDF disregards word order and sentence semantics, yet the rich lexical differences between systematically fabricated and factually grounded texts appear sufficient for high accuracy on this benchmark.

Several limitations warrant consideration. The dataset is predominantly English and covers a narrow temporal window of U.S. political news; generalization to other languages, domains, or time periods has not been validated. Adversarially crafted articles that deliberately mimic credible-news vocabulary could evade lexical classifiers. Furthermore, the system is static: it cannot autonomously adapt to newly emerging misinformation strategies without retraining on updated data.

Future work will investigate integrating pre-trained contextual encoders such as BERT or RoBERTa for richer semantic representations, incorporating metadata features (source reputation, author history, publication timestamp), extending the pipeline to multilingual corpora, and hosting the application on a cloud endpoint to broaden accessibility beyond local execution.

VI. CONCLUSION

This paper presented a complete, end-to-end pipeline for automated fake news detection grounded in ensemble machine learning. By combining TF-IDF feature extraction with a Random Forest Classifier and deploying the resulting model through an interactive Streamlit application, the system provides an accessible solution that operates efficiently on commodity hardware. An accuracy of 94.5% on the Kaggle benchmark confirms the practical viability of the approach. The modular architecture and open-source tooling ensure reproducibility and lay a transparent foundation for future enhancements involving deep contextual models and cross-lingual datasets.

VII. ACKNOWLEDGEMENT

The authors thank the Department of Computer Science and Engineering, Integral University, Lucknow, and project guide Ms. Zainab Khan for their continuous guidance and support throughout this work.

VIII. REFERENCES

[1] H. Allcott and M. Gentzkow, "Social media and fake news in the 2016 election," J. Econ. Perspectives, vol. 31, no. 2, pp. 211–236, 2017.

[2] K. Shu, A. Sliva, S. Wang, J. Tang, and H. Liu, "Fake news detection on social media: A data mining perspective," ACM SIGKDD Explorations, vol. 19, no. 1, pp. 22–36, 2017.

[3] V. L. Rubin, Y. Chen, and N. J. Conroy, "Deception detection for news: Three types of fakes," Proc. ASIS&T, vol. 52, no. 1,

pp. 1–4, 2015.

[4] H. Ahmed, I. Traore, and S. Saad, "Detecting opinion spams and fake news using text classification," Security and Privacy, vol. 1, no. 1, e9, 2018.

[5] N. Ruchansky, S. Seo, and Y. Liu, "CSI: A hybrid deep model for fake news detection," in Proc. ACM CIKM, Singapore, 2017, pp. 797–806.

[6] K. Shu et al., "FakeNewsNet: A data repository with news content, social context and dynamic information for studying fake news on social media," Big Data, vol. 8, no. 3, pp. 171– 188, 2020.

[7] L. Breiman, "Random forests," Machine Learning, vol. 45, no. 1, pp. 5–32, 2001.

[8] R. K. Kaliyar, A. Goswami, P. Narang, and S. Sinha, "FakeBERT: Fake news detection in social media with a BERT-based deep learning approach," Multimedia Tools and Applications, vol. 80, no. 8, pp. 11765–11788, 2021.

[9] G. McIntire, Fake and Real News Dataset, Kaggle, 2018. [Online]. Available: https://www.kaggle.com/clmentbisaillon/fake-and-real-news- dataset