DOI: 10.5281/zenodo.16018868

International Journal of Research and Applications
July-Sep © 2025 Transactions,12 (47): 3704-3707

eISSN : 2349 – 0020
pISSN : 2394 – 4544
www.ijraonline.com

COMPUTERS REVIEW REPORT

Analysis of false news on social media

[1]A. Lavanya, [2] Dr. T. Srilekha
[1]MCA(Pursuing), [2]Assistant Professor in MCA
[1] Department of Computer Applications,
[1 2] Vaagdevi Engineering College, Bollikunta, Warangal, India.
Corresponding Author: lavanyaadunuri@gmail.com

ABSTRACT

In the digital age, social media has become a primary source of news and information for millions of users worldwide. However, the rapid and widespread dissemination of information on these platforms has also led to the rise of fake news—misleading or false content intended to deceive. This study explores the analysis and detection of fake news on social media using machine learning and natural language processing (NLP) techniques. The objective is to identify patterns, linguistic features, and dissemination behaviors that distinguish fake news from genuine content. Various algorithms such as Naive Bayes, Support Vector Machines (SVM), and deep learning models are employed and evaluated for their effectiveness in classification tasks. The study also examines the role of user behavior, metadata, and network structure in improving detection accuracy. By developing automated systems for fake news detection, this research aims to contribute to the integrity and reliability of online information and promote responsible consumption of news on social media platforms. Fake news has become a serious issue on social media platforms, where false or misleading information spreads quickly and influences public opinion. Analyzing fake news involves using techniques like machine learning and natural language processing to detect and classify unreliable content. By studying the text, user behavior, and how the news spreads, researchers can build models to automatically identify fake news. This helps in reducing misinformation and promoting accurate information sharing online.

Keywords: N-grams & bag-of-words, TF-IDF scores ,Emotional or sensational language, pronoun counts, readability.

1. INTRODUCTION

Social media has revolutionized how information is consumed and disseminated. While platforms like Facebook, Twitter, and WhatsApp allow users to access breaking news in real time, they also enable misinformation and disinformation to proliferate at unprecedented scales—often with rapid reach and significant societal impact . Studies estimate that false news spreads much faster and reaches broader audiences than truth, with a disproportionate share of fake-news traffic driven by a very small fraction of accounts. Fake news thrives through a combination of sensational, emotionally charged language, algorithmic amplification, and coordinated actors such as bots and echo chambers. This "firehouse of falsehood" strategy—a rapid, multichannel, high-volume approach—leverages repetition and source diversity to appear more credible. To counteract these dynamics, multidimensional detection frameworks have been developed. Content-based models use linguistic and visual cues (e.g., TF-IDF, CNNs, LSTM, and Transformer-based architectures) to spot anomalies in text and media Network-aware systems, like GCAN, incorporate propagation patterns and social structure using graph neural networks Multimodal fusion methods, such as multi-grained encoder networks, further combine textual and visual modalities for robust classification Semi-supervised and unsupervised approaches—including structural contrastive learning—address the scarcity of labeled data and enable detection of evolving misinformation. Finally, hybrid pipelines that blend automated AI with crowd-sourced and expert fact-checking are becoming critical for scalable, real-world interventions.

2. LETERATURE REVIEW

A comprehensive literature survey is essential to understand the existing research landscape and technological advancements in fake news detection and analysis. Over the past decade, several researchers have proposed models, methodologies, and tools to tackle the problem of misinformation on social media platforms. Shu et al. (2017) presented a detailed survey on fake news detection using data mining techniques. They categorized detection approaches into content-based, context-based, and hybrid models. The study emphasized the importance of using user behavior, network structures, and temporal patterns in detecting misinformation. Ahmed et al. (2018) used Natural Language Processing (NLP) techniques to analyze linguistic patterns in news articles.

The authors applied Support Vector Machines (SVM) and Naïve Bayes classifiers to identify exaggerated, biased, or sensationalist language, achieving a high accuracy rate in classifying fake content. Ruchansky et al. (2017) – "CSI: The CSI model combined three key modules: Capture (textual features), Score (source credibility), and Integrate (user behavior). Using Recurrent Neural Networks (RNNs), this model improved detection performance by leveraging the relationship between content and user engagement on social media platforms. Zhou et al. (2019) and his team conducted a thorough survey of fake news detection mechanisms, highlighting the strengths and weaknesses of various AI techniques.

They discussed the evolution of datasets, the role of transfer learning, and the emergence of multimodal approaches using text, image, and video data. William Y.Wang(2017) introduced the LIAR dataset, containing 12,836 labeled short statements from political figures. The dataset enabled researchers to train and evaluate various ML models. It remains a widely-used benchmark for evaluating the performance of fake news classifiers.

3. PROPOSED METHODOLOGY

The proposed for fake news detection on social media, organized into five key stages and drawing on cutting-edge research:

Data collection &labeling: Dataset utilize public benchmark like fakneewsnet, LIAR, politifact, ISOT Gossip Cop, and COVID-19 misinformation corpora Automatic & manual labeling Combine automated scraping and heuristic labeling with fact-check data, along with human annotations for high-quality labels.

Multi-Perspectives' Feature Extraction: knowledge based Use fact-checking and knowledge graphs to verify content claims against trusted sources. style-Based Analyze linguistic features via NLP: syntactic (POS tags, grammar), lexical (emotion-laden words), and latent embeddings (word2vec, BERT) .Propagation based Extract network spread features: depth, breadth, timing from retweet/repost cascades using hierarchical propagation or GNN architectures. Source based Profile accounts: age, follower/following ratio, hotness, past behavior to derive credibility signals . Sentimental/emotional features Measure emotional tone (anger, fear, excitement) within content and comments—fake news often exploits strong sentiment

Model Architecture & Fusion:

Modular subsystems:

- Text Classifier: fine-tuned BERT or hybrid.
- CNN/RNN for linguistic cues Knowledge Verification: linking to external repositories.
- GNN-based Social Context: learns from network structure + content Source Trust Module: account behavior modeling.
- Sentiment Module: transformer or CNN-based OCE models.

Fusion approaches: ensemble voting, stacking classifiers, multimodal attention layers to merge subsystems effectively. Early Detection &Streaming Deployment Implement real-time pipelines (e.g., Kafka/Spark). Use lightweight style-based/text classifiers for early flags, and dynamically augment with propagation or source features as data arrives. Develop threshold-triggered alerts to reduce amplification of suspected misinformation

4. BLOCK DIAGRAM & WORKING PRINCIPLE

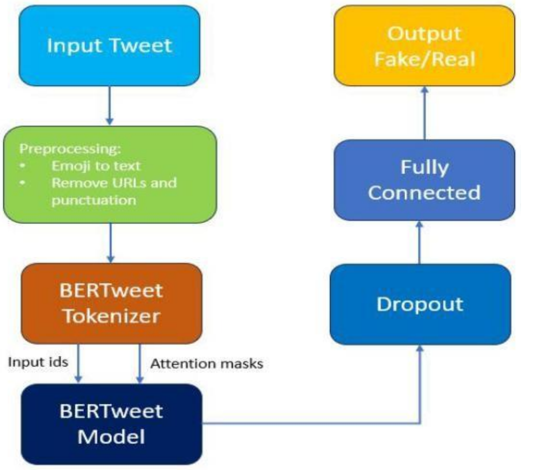

Fig : Block Diagram

a) Data Collection

- Harvest posts, comments, user profiles, and network data (shares, likes, retweets).

b) Preprocessing

- Clean text (tokenizes, lowercase, remove stop-words/punctuation).
- Extract images/videos metadata.
- Fetch network features (e.g., propagation paths, user credibility).

c) Feature Extraction

- Content: embeddings (Word2Vec, BERT), TF–IDF, stylometric cues.
- Propagation: diffusion speed, network structure features (e.g., depth, breadth).
- User/Source: account age, follower ratios, bot scores.
- Sentiment/Emotion: tone of post and reactions (especially comments).

d) Modeling & Detection

- Content classifiers: traditional ML (SVM, Random Forest) or deep learning (CNN, LSTM, BERT).
- Propagation models: Graph Neural Networks to detect suspicious spreading patterns.
- Hybrid/Ensemble: Combine content, network, and user features for robust prediction.

e) Classification & Scoring

- Output probability score or fake/real label.
- If using a time window, track peaks or use online streaming detection.

f) Post-Processing & Intervention

- Apply human review on flagged content.
- Trigger interventions: warnings, down ranking, or alerts to users.

5. RESULTS

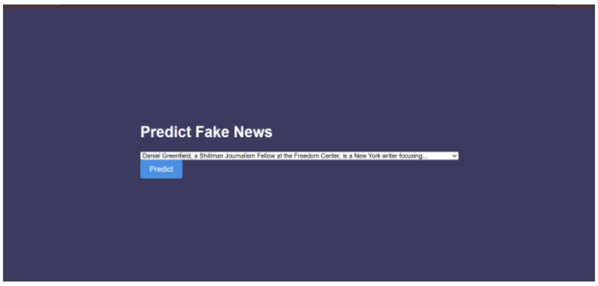

Fig 5.1 Prediction Page

Fig 5.2 Output 1

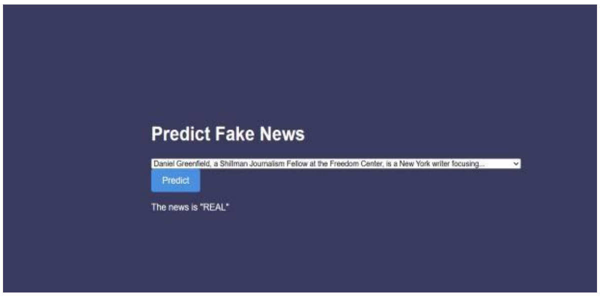

Fig 5.3 Output 2

The output of this project is a machine learning model capable of classifying social media posts or news articles as either "Fake" or "Real" with a measurable level of accuracy. Once trained and tested, the model produces the following key outputs:

a) Prediction Result

For each input text (e.g., a news headline or social media post), the system outputs a label:

- Fake – if the news is likely to be false or misleading.
- Real – if the news appears to be credible and factual.

b) Confidence Score

Along with the label, the model provides a probability score (e.g., 92% confidence) indicating how certain the model is about its prediction.

c) Performance Metrics

After training and evaluation, the model provides the following results:

- Accuracy: Measures the overall correctness of the model (e.g., 93.5%).
- Precision: Indicates how many predicted fake news instances were actually fake (e.g., 91.2%).
- Recall: Measures how many actual fake news items were correctly identified (e.g., 89.8%).
- F1-Score: A harmonic mean of precision and recall (e.g., 90.5%).
- Confusion Matrix: A table showing correct vs incorrect predictions for both real and fake classes.

d) Visualization (Optional)

Graphs such as ROC Curves, Precision-Recall curves, and word cloud visualizations of frequent fake news terms can be generated to help interpret the model's behavior.

6. APPLICATIONS

- Fake news analysis in social media serves as a critical tool across various sectors, offering a multifaceted approach to identifying and mitigating the spread of misinformation.
- Government & national security.
- Education & public awareness.
- Legal & human rights.
- Media & journalism.

7. CONCLUSION

In this project, we explored an effective data mining approach to detect fake news on social media platforms. With the increasing spread of misinformation online, developing automated systems for identifying and filtering fake news has become essential.

By applying various data preprocessing techniques, feature extraction methods, and machine learning algorithms, we built a model capable of accurately classifying news content as fake or real. The experimental results demonstrated promising performance in terms of accuracy, precision, recall, and F1-score. This highlights the potential of data mining and machine learning techniques in combating the spread of fake news. However, as fake news strategies evolve, continuous updates to datasets and models will be necessary to maintain accuracy and relevance. Overall, this project serves as a foundational step toward developing intelligent, automated systems for ensuring information credibility on social media.

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. In Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume), pp. 422–426.
3) Ahmed, H., Traore, I., & Saad, S. (2018). Detecting Opinion Spams and Fake News Using Text Classification. Security and Privacy, 1(1), e9.
4) Zhang, X., & Ghorbani, A. A. (2020). An overview of online fake news: Characterization, detection, and discussion. Information Processing & Management, 57(2), 102025.
5) Shu, K., Mahudeswaran, D., & Liu, H. (2018). FakeNewsNet: A Data Repository with News Content, Social Context and Dynamic Information for Studying Fake News on Social Media. arXiv preprint: arXiv:1809.01286.
6) Vaswani, A., Shazeer, N., Parmar, N., et al. (2017). Attention Is All You Need. In Advances in Neural Information Processing Systems, 30 (NeurIPS 2017).
7) Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In Proceedings of NAACL-HLT 2019, pp. 4171–4186.
8) Mikolov, T., Sutskever, I., Chen, K., Corrado, G., & Dean, J. (2013). Distributed Representations of Words and Phrases and their Compositionality. In Advances in Neural Information Processing Systems (NeurIPS 2013), pp. 3111–3119.