WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Sentiment-Based News Veracity Detector: Distinguishing Fake News from Real News Using Emotion and Sentiment Cues

Prof. Swathi A[1], Yogesh Naik M[2], Sathvik P[3], Shashikiran[4], Sharath Kumar K R[5]

[1]Assistant Professor, Computer Science and Engineering, Dayananda Sagar Academy of Technology & Management, Bengaluru, India
swathianjanappa7777@gmail.com
[2]Student, 3rd Year, B.E Computer Science and Engineering, Dayananda Sagar Academy of Technology & Management, Bengaluru, India
[2]1dt22cs189@dsatm.edu.in, [3]1dt22cs137@dsatm.edu.in, [4]1dt22cs146@dsatm.edu.in, [5] 1dt22cs141@dsatm.edu.in

ABSTRACT-

The proliferation of fake news has become a critical issue in today's digital information age, influencing public opinion and disrupting societal harmony. Traditional methods of fake news detection primarily rely on linguistic features or source credibility, often failing to capture the subtle emotional and sentiment-driven patterns that differentiate real from deceptive content. This study proposes a web-based application that leverages sentiment and emotion cues extracted from textual content to accurately classify news as real or fake. By employing Natural Language Processing (NLP) techniques combined with machine learning classifiers, the system analyzes emotional tone and sentiment polarity to identify deceptive patterns. The platform provides real-time predictions, ensures transparency in classification, and offers an interactive dashboard for end-users and administrators. The system's modular design allows for easy scalability and integration with news platforms, aiming to enhance digital media literacy and curb the spread of misinformation.

Keywords- Fake News Detection, Sentiment Analysis, Emotion Recognition, NLP, Machine Learning, Web Application, Real-Time Classification, News Veracity.

I. INTRODUCTION

The rapid dissemination of information through digital platforms has revolutionized communication but has also facilitated the widespread circulation of misinformation and fake news. In a country like India, where social media serves as a major information source, fake news has the potential to influence public sentiment and decision-making on a large scale. Conventional fake news detection approaches often rely on checking factual correctness or evaluating the credibility of sources, but these methods face challenges such as delayed verification and limited adaptability to diverse content styles.

This paper introduces a novel sentiment-based approach to fake news detection, where emotional and sentiment patterns in text are used as key indicators of veracity. Fake news often exhibits exaggerated emotional tones or extreme sentiment polarities, which can be effectively captured using NLP tools. By integrating sentiment and emotion analysis into a machine learning pipeline, we aim to create a more nuanced and scalable solution for fake news detection.

The system is designed as a web-based application, providing an intuitive interface for users to input news articles or headlines and receive real-time feedback on their authenticity. It also includes an administrator dashboard for managing datasets, training models, and viewing analytical insights. This integrated approach enhances user trust, supports informed decision-making, and contributes to building a resilient information ecosystem.

II. LITERATURE REVIEW

2.1 Fake News Detection Techniques

Traditional approaches to fake news detection primarily rely on content-based analysis, source credibility checks, and fact-matching techniques. Early research applied machine learning models using linguistic features such as n-grams, POS tags, and readability scores to distinguish fake from real news articles [1]. While these methods detect stylistic patterns, they often fail to capture subtle manipulations and emotionally charged misinformation.

To improve detection accuracy, researchers have developed hybrid models that integrate user interaction data and social propagation behavior [2]. These models leverage metadata such as retweet patterns and user credibility, offering improved performance on social media platforms. However, their effectiveness is limited in environments lacking rich user interaction data or where such metadata is inaccessible.

2.2 Role of Sentiment and Emotion in Fake News

More recent studies emphasize the emotional tone and sentiment polarity of news content as significant indicators of veracity. Fake news often contains highly emotive language, especially emotions like fear, anger, or surprise, while real news tends to maintain a neutral or objective tone [3].

Basic sentiment analysis techniques classify text as positive, negative, or neutral, but they fall short in capturing the complexity of emotional manipulation. Advanced emotion detection models, using tools like the NRC Emotion Lexicon and deep learning-based classifiers, can detect a broader range of emotions—such as disgust, joy, trust, and sadness—making them more suitable for fake news detection [4]. These emotional signals enhance the model's ability to understand the intent behind the content, which is especially useful since fake news is designed to provoke reactions and go viral.

2.3 Advances in NLP for Fake News Detection

The introduction of transformer-based models, including BERT, RoBERTa, and DistilBERT, has significantly advanced the field of natural language understanding. These models use self-attention mechanisms to better capture semantic and contextual relationships within text [5][6]. When fine-tuned on domain-specific datasets, such as LIAR, FakeNewsNet, and CREDBANK, these models achieve state-of-the-art performance in fake news classification.

Researchers have also found success in combining these deep contextual embeddings with traditional classifiers like Logistic Regression, Support Vector Machines (SVM), and Random Forest, creating systems that balance accuracy with interpretability [7]. This approach allows for high-performance models that can still provide explainable predictions, which is crucial for real-world applications.

2.4 Hybrid Approaches and Multimodal Systems

Fake news detection is no longer limited to text alone. Multimodal approaches incorporate images, videos, and social context to enhance reliability. For example, studies show that pairing visual sentiment analysis with text-based cues significantly improves detection of manipulated media on social platforms [8].

Some systems also integrate knowledge graphs and fact-checking databases (e.g., Snopes, PolitiFact) to validate claims in real-time. These fact-aware systems boost accuracy but demand substantial computational and database resources, making them less suitable for lightweight or real-time applications [9].

Although this project focuses on text-based sentiment and emotion analysis, it lays the foundation for future expansion into multimodal or fact-integrated systems.

2.5 Limitations of Existing Systems and Research Gaps

Despite technological advancements, current systems face several practical challenges:

- Lack of interpretability: Deep learning models often act as black boxes, offering limited transparency into which features (e.g., emotion or sentiment) influenced their decisions.
- Scalability issues: Most models are not optimized for real-time detection on high-volume data streams.
- Language and cultural limitations: Existing datasets are primarily English-based, lacking multilingual and cross-cultural emotion annotations [10].
- Dataset limitations: Few public datasets contain emotion-labeled fake news, which hampers training of emotion-aware classifiers.
- End-user disconnect: Few systems offer interactive dashboards or transparent output explanations, making them less usable by non-technical stakeholders.

These gaps underline the need for systems like the one proposed in this project—combining emotion and sentiment analysis with explainability, multilingual support, and real-time scalability.

III. METHODOLOGY

The proposed system uses a multi-step approach to detect fake news by leveraging sentiment analysis, emotion detection, and machine learning techniques. The process begins with collecting diverse labeled datasets containing real and fake news articles, which are then cleaned and preprocessed using techniques like tokenization, stop word removal, and lemmatization to prepare the text for analysis. Sentiment and emotion cues are extracted from the text using lexicon-based tools and deep learning models, providing additional features that help distinguish between real and fake news. Various textual and semantic features—including TF-IDF, n-grams, word embeddings, and readability metrics—are combined with sentiment and emotion scores to create a robust feature set. Multiple machine learning models, including classical algorithms and deep learning architectures, are trained and evaluated

using stratified cross-validation and hyperparameter tuning. The system emphasizes interpretability through feature importance analysis and is deployed in a user-friendly web dashboard for practical application.

Key steps in the methodology include-

Data Collection & Preprocessing:

- Use publicly available datasets like LIAR, FakeNewsNet, and Kaggle datasets.
- Clean and normalize text via tokenization, stop word removal, and lemmatization.
- Balance datasets to avoid model bias.

Sentiment and Emotion Extraction:

- Apply sentiment analysis tools (VADER, TextBlob, transformer-based models).
- Detect emotions using NRC Emotion Lexicon and deep learning classifiers.
- Encode sentiment and emotion features numerically.

Feature Engineering:

- Extract textual features: TF-IDF, n-grams, POS tags, readability scores.
- Generate semantic embeddings: Word2Vec, GloVe, BERT.
- Integrate emotion and sentiment features into the feature set.

Model Training and Evaluation:

- Train models including Logistic Regression, Random Forest, LSTM, BERT, etc.
- Perform hyperparameter tuning using grid search or Bayesian optimization.
- Use stratified k-fold cross-validation for reliable performance estimation.
- Evaluate with metrics like accuracy, precision, recall, F1-score, and ROC-AUC.
- Use SHAP and LIME for model interpretability.

Deployment:

- Develop a web-based dashboard for user input and real-time classification.
- Implement backend with Flask/Django and frontend with React.js/HTML-CSS.
- Expose model via REST API for easy integration and scalability.

This approach ensures a comprehensive and scalable solution for accurate fake news detection that combines textual content analysis with emotional and sentiment insights.

IV. CONCLUSION

This project presents a novel approach to detecting fake news by leveraging sentiment and emotion cues through Natural Language Processing and machine learning techniques. By analyzing the underlying emotional tone of news content, the system enhances accuracy in identifying deceptive narratives compared to traditional keyword- or metadata-based methods. The web-based platform offers real-time news classification, a user-friendly interface, and an admin panel for monitoring and model updates. While the system shows promising results in distinguishing fake from real news, future improvements can focus on expanding multilingual support, refining emotion classification, and integrating real-time news feeds for continuous analysis. Ultimately, this solution contributes to combating misinformation and promoting a more informed digital society.

REFERENCES-

[1] K. Shu, A. Sliva, S. Wang, J. Tang, and H. Liu, “Fake News Detection on Social Media: A Data Mining Perspective,” *ACM SIGKDD Explor. Newsl.*, vol. 19, no. 1, pp. 22–36, 2017.

[2] H. Rashkin, E. Choi, J. Y. Jang, S. Volkova, and Y. Choi, “Truth of Varying Shades: Analyzing Language in Fake News and Political Fact-Checking,” in *Proc. EMNLP*, 2017, pp. 2931–2937.

[3] X. Zhou and R. Zafarani, “Fake News: A Survey of Research, Detection Methods and Opportunities,” *arXiv preprint arXiv:1812.00315*, 2018.

[4] S. M. Mohammad and P. D. Turney, “Crowdsourcing a Word–Emotion Association Lexicon,” *Comput. Intell.*, vol. 29, no. 3, pp. 436–465, 2013.

[5] A. Giachanou, E. A. Ríssola, B. Ghanem, F. Crestani, and M. Fernández, "Emotion Analysis for Fake News Detection," in *Proc. 42nd Int. ACM SIGIR Conf.*, 2019, pp. 1105–1108.

[6] S. Singhania, N. Fernandez, and S. Rao, "3HAN: A Deep Neural Network for Fake News Detection," *arXiv preprint arXiv:1705.10049*, 2017.

[7] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," in *Proc. NAACL-HLT*, 2019, pp. 4171–4186.

[8] B. D. Horne and S. Adalı, "This Just In: Fake News Packs a Lot in Title, Uses Simpler, Repetitive Content in Text Body, More Similar to Satire than Real News," in *Proc. AAAI Conf. Web Soc. Media*, 2017, pp. 759–766.

[9] F. Alam, S. Shaar, G. Da San Martino, and P. Nakov, "Fighting the COVID-19 Infodemic: Modeling the Perspective of Journalists, Fact-Checkers, Social Media Platforms, Policy Makers, and the Society," in *Proc. EMNLP*, 2021, pp. 81–96.

[10] S. Kumar, R. West, and J. Leskovec, "Disinformation on the Web: Impact, Characteristics, and Detection of Wikipedia Hoaxes," in *Proc. 25th Int. Conf. WWW*, 2016, pp. 591–602