Detecting Fake News: A Hybrid Approach with Linguistic and Knowledge-Based Analysis

Mr. RAGHUVARAN E, M.E., Assistant Professor,Department of Computer Science and Engineering

Mr. M.MOHAMMED ALJAMIN, B.E, Student of Computer Science Engineering

Mr. D.RUTHESH, B.E, Student of Computer science and Engineering

St. Joseph College of Engineering, Sriperumbudur, Chennai.

Abstract

Integrates advanced machine learning techniques for comprehensive analysis of textual, visual, and contextual features, ensuring robust detection of fake news across diverse media formats. Employs dynamic learning mechanisms to continuously adapt and respond to evolving misinformation tactics, enabling real-time monitoring and swift identification of emerging patterns. Demonstrates exceptional precision and recall rates through extensive validation on varied datasets, offering reliable and trustworthy results crucial for safeguarding the integrity of digital information ecosystems. The proposed model consists of two phases: first, global features are extracted by TF-IDF, spatial features by a convolutional neural network (CNN), and temporal features by bi-directional long short-term memory (BiLSTM) simultaneously.

Key Terms

Social Media - Textual Analysis - Dynamic Learning - Natural Language Processing (NLP) - Data Mining - Automated System - Machine Learning Models - Fact-Checking - Information Verification - Real-time Monitoring - Validation

Introduction

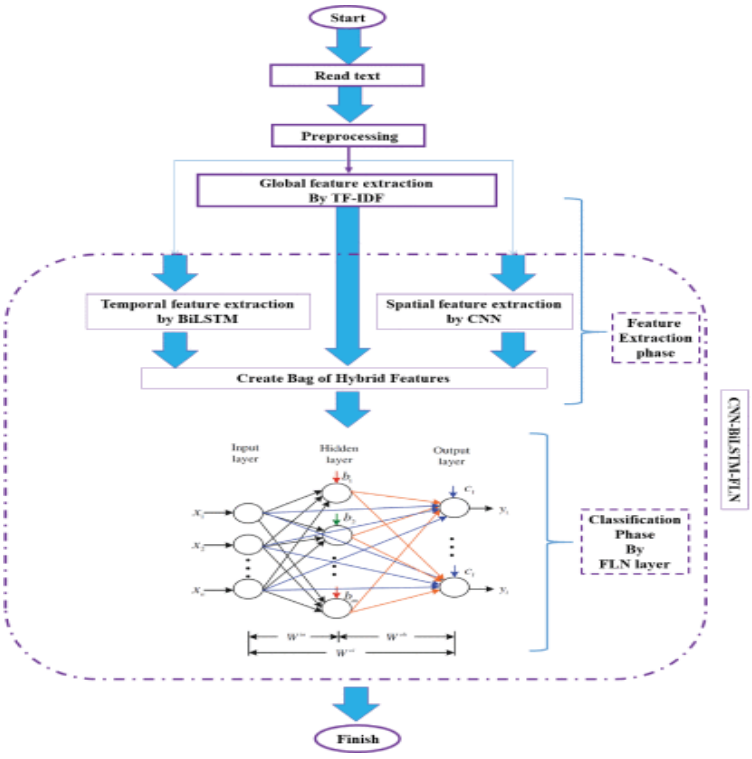
In today's interconnected digital age, the rapid dissemination of information has empowered individuals with unprecedented access to news and knowledge. However, this landscape is also fraught with challenges, particularly the proliferation of fake news and misinformation. The consequences of fake news can be far-reaching, leading to misinformation, polarization, and erosion of trust in media and institutions. To address this pressing issue, fake news detection systems have become a critical area of research and development. These systems leverage cutting-edge technologies such as machine learning, natural language processing (NLP), and data analytics to identify and combat misinformation across various digital platforms. In this context, this paper explores the key concepts, methodologies, and challenges associated with fake news detection systems. Additionally, it discusses the ethical considerations, such as bias detection and fairness, that are essential for developing trustworthy and effective fake news detection solutions. Through a comprehensive examination of the current state-of-the-art techniques and best practices, this paper aims to contribute to the ongoing discourse on combating fake news and promoting information integrity in the digital era.

Literature Survey

The paper “Topic-Aware Fake News Detection Based on Heterogeneous Graph” (2023) by Lijuan Sun and Hongbin Wang, provides the Heterogeneous graph which allows different types of edges or nodes to have different dimensions of features or attributes, it contains more comprehensive information and richer semantics.

The paper “Enhancing Fake News Detection by Multi-Feature Classification” (2023**)** by Ahmed Hashim, provides the architecture with the integration of CNN+BiLSTM and TF-IDF ithin the deep neural network architecture.using these three components at the same time gives a better view of the input text to the classifier, resulting in improved detection accuracy.

System Design



The system design for fake news detection begins with data collection from diverse and reliable sources, encompassing news articles, social media posts, and multimedia content. This data undergoes rigorous preprocessing to ensure consistency and relevance, including tasks such as text normalization, tokenization, and filtering out noise and irrelevant information. The goal of this stage is to prepare the data for analysis and feature extraction while maintaining its integrity and quality.

Next, the system focuses on feature extraction, where relevant features are derived from the preprocessed data. These features may include textual attributes like word frequencies and sentiment analysis scores, visual characteristics such as image metadata and visual cues, and contextual information like source credibility and publication history. By leveraging these diverse features, the system gains a comprehensive understanding of the data, enabling it to detect patterns and nuances indicative of fake news.

The core of the system lies in its machine learning models, which are trained using labeled datasets to distinguish between genuine and fake news articles. These models encompass a range of techniques, from traditional supervised classifiers like support vector machines and random forests to more advanced deep learning architectures like convolutional neural networks (CNNs)

and transformers. By continuously learning from the data and updating their predictive capabilities, these models form the backbone of the fake news detection system, providing accurate and reliable results in identifying misinformation.

Implementation

Snapshots

Conclusion

In conclusion, fake news detection systems are pivotal in addressing the rampant spread of misinformation in today's digital landscape. Through the integration of advanced technologies like machine learning, NLP, and multimodal analysis, these systems have made significant strides in identifying and combating fake news. However, there are still challenges to overcome, such as improving contextual understanding, mitigating biases, and enhancing transparency. Future enhancements in these areas, coupled with real-time monitoring capabilities and collaborative approaches, hold promise in strengthening the effectiveness and reliability of fake news detection systems. By continuing to innovate and collaborate across disciplines, we can work towards a more informed and trustworthy information ecosystem, safeguarding the integrity of public discourse and democratic processes.

Future Enhancemments

Future enhancements in fake news detection systems include advanced multimodal analysis to detect misinformation across various media formats, improved contextual understanding to identify nuanced forms of fake news, and real-time monitoring capabilities to swiftly counter emerging trends. Additionally, leveraging deep learning architectures, implementing explainable AI, and developing bias mitigation strategies are crucial for enhancing accuracy, transparency, and fairness in fake news detection. Integrating collaborative filtering algorithms, exploring blockchain technology for source verification, and incorporating human-in-the-loop systems are also key areas for further development in combating misinformation at scale.

References:

[1] H. Ahmed, I. Traore and S. Saad, "Detecting opinion spams and fake news using text classification", Secur. Privacy, vol. 1, no. 1, pp. 9, Jan. 2018.

[2] S. Vosoughi, D. Roy and S. Aral, "The spread of true and false news online", Science, vol. 359, no. 6380, pp. 1146-1151, Mar. 2018.

[3] V. L. Rubin, Y. Chen and N. K. Conroy, "Deception detection for news: Three types of fakes", Proc. Assoc. Inf. Sci. Technol., vol. 52, no. 1, pp. 1-4, Jan. 2015.

[4] G. Li, P. Niu, X. Duan and X. Zhang, "Fast learning network: A novel artificial neural network with a fast learning speed", Neural Comput. Appl., vol. 24, no. 7, pp. 1683-1695, Jun. 2014.

[5] W. Han and V. Mehta, "Fake news detection in social networks using machine learning and deep learning: Performance evaluation", Proc. IEEE Int. Conf. Ind. Internet (ICII), pp. 375-380, Nov. 2019.

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

[7] R. M. Vaibhav and E. Hovy, "Do sentence interactions matter? Leveraging sentence level representations for fake news classification", Proc. 13th Workshop Graph-Based Methods Natural Lang. Process., pp. 134-139, 2019.

[8] L. Hu, T. Yang, L. Zhang, W. Zhong, D. Tang, C. Shi, et al., "Compare to the knowledge: Graph neural fake news detection with external knowledge", Proc. 59th Annu. Meeting Assoc. Comput. Linguistics 11th Int. Joint Conf. Natural Lang. Process., pp. 754-763, 2021.

[9] A. Bondielli and F. Marcelloni, "A survey on fake news and rumour detection techniques", Inf. Sci., vol. 497, pp. 38-55, Sep. 2019.

[10] H. Wang, K. Tian, Z. Wu and L. Wang, "A short text classification method based on convolutional neural network and semantic extension", Int. J. Comput. Intell. Syst., vol. 14, no. 1, pp. 367-375, 2021.

Author 1

Mr.E.Raghuvaran M.E., Assistant Professor in the Department of Computer Science and Engineering at St. Joseph College of Engineering, Sriperumbudur, Chennai, Tamil Nadu. Currently, Pursuing Ph. D in Pondicherry University.

Author 2

Mr. M.Mohammed Aljamin B.E., Student of Computer Science and Engineering at St.Joseph College of Engineering, Sriperumbudur, Chennai, TamilNadu. I had attended many Workshops, Seminars in Python, Machine Learning and UI/UX designing.

Author 3

Mr. D. Ruthesh B.E., Student of Computer Science and Engineering at St.Joseph College of Engineering, Sriperumbudur, Chennai, Tamil Nadu. I had attended many Workshops and Seminars in the area of Python and Machine Learning.