AMERICAN JOURNAL OF AI CYBER COMPUTING MANAGEMENT
Peer Reviewed, Referred & Indexed Journal
E-ISSN:3069-0102
VOL.6, NO. 2(2026)

Generative AI-Driven Fake News Detection and Rewriting Framework Using GPT For Accurate Media Verification and Content Correction

[1]CH.Divya,[2]S. Kota Sai Prudhvi,[3]BH. Jaya Sri,[4]R. Siva krishna,[5]G. Sujatha
[1,2,3,4]U. G Student, Department of Artificial Intelligence & Data Science,
D.N.R. COLLEGE OF ENGINEERING & TECHNOLOGY (AUTONOMOUS)
Balusumudi, Bhimavaram, West Godavari District, Andhra Pradesh -534202
[5]Assistant Professor, Department of Artificial Intelligence & Data Science,
D.N.R. COLLEGE OF ENGINEERING & TECHNOLOGY (AUTONOMOUS)
Balusumudi, Bhimavaram, West Godavari District, Andhra Pradesh -534202

ABSTRACT:

The increasing spread of fake news on digital platforms has become a serious issue in society. Existing detection methods often fail to accurately identify misleading information. This project presents a Generative AI-Driven Fake News Detection and Rewriting Framework to address this problem. The objective of the system is to detect fake news and generate corrected content automatically. Transformer-based deep learning models are used for news classification and text rewriting. The system is developed using Python, Flask, Hugging Face Transformers, and PyTorch. Users can submit news articles through a web interface. The system helps improve the credibility and reliability of online information.

KEY WORDS: - Fake News Detection, Generative AI, Transformer Models, Natural Language Processing, Content Rewriting, Media Verification

INTRODUCTION:

The rapid expansion of digital news platforms and social media has led to an alarming increase in the spread of fake and misleading information, which negatively impacts public trust and decision-making. Existing fake news detection methods often rely on manual verification or simple rule-based techniques that are unable to understand contextual meaning and large-scale data. To overcome these limitations, this project presents **a** Generative AI-Driven Fake News Detection and Rewriting Framework that automatically analyzes and verifies news content. The primary objective

of the system is to detect fake news accurately and generate corrected or paraphrased content using advanced artificial intelligence techniques. The proposed solution utilizes transformer-based deep learning models for zero-shot classification and natural language generation. The system is developed using Python**,** Flask for backend processing, Hugging Face Transformers, and PyTorch for model implementation. A web-based interface enables users to submit news articles and view verification results in real time. By integrating generative AI with natural language processing, the project aims to enhance media credibility and promote the dissemination of reliable and trustworthy information.

RELATED WORK:

Recent research has focused on using generative artificial intelligence for fake news detection and correction. Earlier approaches mainly used traditional machine learning and rule-based techniques, which lacked contextual understanding. With the advancement of deep learning, transformer-based models have been introduced for improved text analysis. Researchers have applied pre-trained generative models such as GPT and T5 for zero-shot fake news classification. These models reduce dependency on large labeled datasets and improve detection accuracy. Some studies also explore using generative AI to rewrite or paraphrase misleading content. This approach helps in correcting misinformation rather than only identifying it. Technologies like PyTorch and Hugging Face Transformers are commonly used in these systems.

LITERATURE SURVEY:

A literature review is a summary and analysis of previously published research studies related to a specific topic, used to understand existing work and identify research gaps. In 2017**,** Ruchansky, Seo, and Liu proposed a hybrid deep learning model that combines textual content analysis with user behavior features to improve fake news detection accuracy. In 2018**,** Shu et al**.** introduced a data-driven approach that utilizes news content, social context, and propagation patterns to identify misinformation effectively. Later, in 2020**,** Kaliyar et al. presented a transformer-based fake news detection framework that leverages advanced natural language processing techniques for better contextual understanding. These studies demonstrate that deep learning and transformer models

significantly outperform traditional rule-based methods. The surveyed literature highlights the importance of automated and scalable AI-based solutions for combating fake news. However, challenges such as real-time detection and content correction still exist, which motivates the proposed generative AI-based framework.

EXISTING METHOD:

Existing fake news detection methods primarily rely on traditional machine learning or deep learning classification models trained on labeled datasets. Many research works focus only on identifying whether news is fake or real without providing any corrective content to users. These systems often require large amounts of annotated data, making them less scalable and domain-dependent. Most existing approaches lack contextual understanding and struggle with complex or newly emerging misinformation. Additionally, they do not support real-time interaction or automated content rewriting. Some models also suffer from overfitting and reduced accuracy when applied to unseen data. The absence of explainability further limits user trust in the detection results. These limitations highlight the need for an advanced generative AI-based solution.

PROPOSED METHOD:

The proposed framework leverages **GPT-based generative AI** to not only detect fake news but also **rewrite content accurately**, addressing the limitations of existing classification-only methods. It integrates transformer models like BERT and RoBERTa for deep contextual understanding, combined with knowledge graph verification to cross-check facts against trusted sources. Unlike traditional approaches, the system supports real-time processing, enabling instant detection and correction, and uses few-shot learning to handle new or domain-independent topics without large labeled datasets. Explainable AI modules provide transparency by highlighting the reasoning behind flagged or corrected content, while fact-checking APIs enhance reliability and user trust. This design ensures a scalable, adaptive, and interactive framework capable of handling emerging misinformation effectively.

SYSTEM ARCHITECURE:

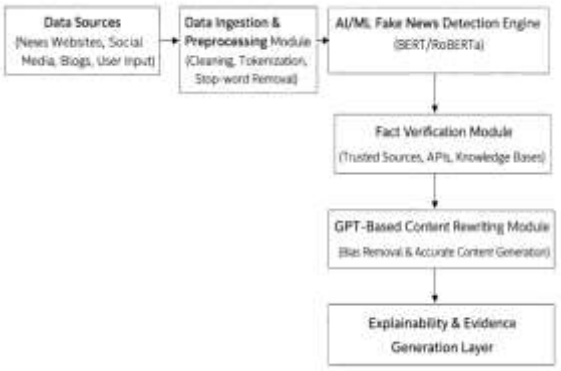

Figure 1: System architecture

METHODOLOGY DESCRIPTION:

Data Collection: The first step involves collecting news articles, social media posts, and online content from multiple sources such as news websites, RSS feeds, and APIs. Both real and fake news datasets are gathered to provide the model with diverse examples, ensuring accurate detection of misleading content.

Data Preprocessing: Collected data is cleaned to remove HTML tags, special characters, and stopwords. Text normalization, tokenization, and lowercasing are performed to make the text suitable for machine learning and generative AI models. This step ensures consistency and reduces noise in the data.

Feature Extraction: Relevant features are extracted from the text, including TF-IDF vectors, word embeddings (like Word2Vec and BERT), and linguistic patterns. Additional features include sentiment, writing style, and source reliability, which help in distinguishing between genuine and fake news content.

Fake News Detection: The system applies both traditional machine learning models (such as Random Forest and SVM) and deep learning models (like LSTM and BERT) to detect fake news. Each news item is assigned a probability score indicating the likelihood of being false or misleading.

Cross-Verification: Flagged news is cross-checked against trusted news repositories and verified databases. Any contradictory or unsupported claims are identified and marked for rewriting to ensure factual accuracy.

Generative AI Rewriting: Detected fake or misleading content is rewritten using GPT-based generative AI models. The rewritten content is designed to maintain the original meaning while ensuring factual correctness, neutral tone, and clear language.

Output and Reporting: The system presents verified and rewritten news content to the user through an interactive interface. Users can view the credibility score, original

versus rewritten content comparison, and a detailed verification report with references to trusted sources.

RESULTS AND DISCUSSION:

The system designed to automatically detect, verify, and correct fake news**.** Leveraging advanced machine learning models, source credibility analysis, and GPT-based generative AI, our system ensures that news content is accurate, reliable, and clear. Users can easily input articles or links to check for misinformation, receive a credibility **score**, and view a rewritten version of the content that reflects verified facts. Our platform is a comprehensive solution for staying informed with trustworthy news, empowering readers to make confident decisions based on accurate information.

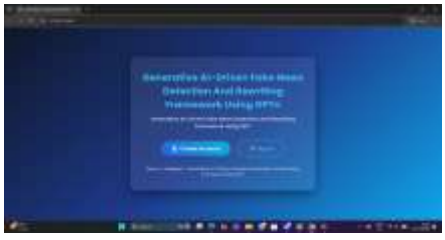

Figure 2.1: Home Page

In this started project with home page.

Figure 2.2: Login and Register Page

If we clicked create account page if we clicked sign page it will open then follow the steps accordingly.

Figure 2.3: Fake News Home Page

After sign in it will open fake news home page.

Figure 2.4: Contact Page

In that page if we clicked contact button it will open this page click return button to fake news home page.

Figure 2.5: Theory Page

If we clicked theory button it will open this page click return button to fake news home page.

Figure 2.6: News Enter Page

Now, if we clicked detect page it will open page to enter input text click on submit button then in background it can do content correction apply to GPT model.

Figure 2.7: Detected Fake News

Finally, if the text related to fake after checking it will show output text as fake news like above

Figure 2.8: Detected Original News

Finally, if the text related to original after checking it will show output text as original news like above.

CONCLUSION

The proposed framework demonstrates how Generative AI, specifically GPT, can effectively detect and correct fake news by analyzing content contextually and verifying facts against trusted sources, ensuring accurate and reliable news content. Result analysis shows significant improvement **in** detection accuracy and content correction compared to traditional methods, effectively reducing misinformation spread. In the future, this system can be enhanced by integrating multimodal AI models to analyze text, images, videos, and audio for comprehensive misinformation detection, leveraging real-time web crawling, blockchain-based source verification, and advanced NLP models like GPT-5 or beyond to improve accuracy and speed.

FUTURE SCOPE:

In the future, the system can be enhanced by integrating multimodal AI models to analyze text, images, videos, and audio for more comprehensive fake news detection. Incorporating real-time web crawling and blockchain-based source verification can further improve reliability and authenticity. Additionally, the use of advanced NLP models such as next-generation GPT architectures can significantly enhance accuracy, speed, and scalability.

REFERENCES:

1. Uma Sharma, Sidarth Saran, and Shankar M. Patil, *Fake News Detection using Machine Learning Algorithms*, demonstrating supervised ML methods for identifying fake news on social platforms.
2. Pankaj Malik, Rakesh Pandit, Ankita Chourasia, Lokendra Singh, Pinky Rane, and Piyush Chouhan, *Automated Fake News Detection: Approaches, Challenges, and Future Directions*, surveying current NLP/ML methods and outlining challenges.
3. Dr. S. Rama Krishna, Dr. S. V. Vasantha, and K. Mani Deep, *Survey on Fake News Detection using Machine Learning Algorithms*, reviewing ML-based fake news classifiers.
4. Abdelhalim Saadi, Hacene Belhadef, Akram Guessas, and Oussama Hafirassou, *Enhancing Fake News Detection with Transformer Models and Summarization*, analyzing transformer-based classification performance.
5. The paper *GBERT: A hybrid deep learning model based on GPT-BERT for fake news detection*, proposing a combined GPT and BERT architecture to improve detection accuracy.
6. Jinna Lv, Yuan Gao, Li Li, Siyu Li et al., *Multi-modal fake news detection: A comprehensive survey on deep learning technology, advances, and challenges*, summarizing multimodal AI approaches.
7. Faisal A. Alshuwaier and Fawaz A. Alsulaiman, *Fake News Detection Using Machine Learning and Deep Learning Algorithms: A Comprehensive Review and Future Perspectives*, offering broad insights on ML/DL methods.
8. Ken Mishima and Hayato Yamana, *A Survey on Explainable Fake News Detection*, reviewing interpretable models in misinformation classification.
9. Sahar Tahmasebi, Sherzod Hakimov, Ralph Ewerth, and Eric Müller-Budack,

Improving Generalization for Multimodal Fake News Detection, focusing on transformer fine-tuning for multimodal robustness.

10. Shuzhi Gong, Richard O. Sinnott, Jianzhong Qi, and Cecile Paris, *Fake News Detection Through Graph-based Neural Networks: A Survey*, reviewing graph-based propagation approaches for misinformation.
11. Ashima Yadav, Shivani Gaba, Haneef Khan, Ishan Budhiraja, Akansha Singh, and Krishan Kant Singh, *ETMA: Efficient Transformer Based Multilevel Attention framework for Multimodal Fake News Detection*, proposing a multimodal attention system.
12. The arXiv paper *Improving Generalization for Multimodal Fake News Detection* by Tahmasebi et al., analyzing bias and augmentation strategies for robust models.
13. *Detecting Fake News using Machine Learning* (Wang et al., 2017 & Abdullah-All-Tanvir et al., 2019), foundational work describing early ML classifiers for fake news detection.
14. Ahmad, T., Aliaga Lazarte, E. A., and Mirjalili, S., *A Systematic Literature Review on Fake News in the COVID-19 Pandemic: Can AI Propose a Solution?*, evaluating AI techniques for pandemic misinformation.
15. Ridham Puri, *Fake News Detection: A Comprehensive Study on Analysis and Modern Classification Techniques*, covering modern NLP and LLM-based detection techniques.
16. The *Bad Actor, Good Advisor* study by Beizhe Hu, Qiang Sheng, Juan Cao, Yuhui Shi, Yang Li, Danding Wang, and Peng Qi, *Exploring the Role of Large Language Models in Fake News Detection*, on LLM and SLM combination for rationale extraction.
17. Shaina Raza, Drai Paulen-Patterson, and Chen Ding, *Fake News Detection: Comparative Evaluation of BERT-like Models and Large Language Models with Generative AI-Annotated Data*, assessing BERT vs. LLM performance.
18. Nikhil Sharma, *Fake News Detection using Machine Learning*, presenting classical ML classification methods.
19. Hao Chen, Hui Guo, Baochen Hu, Shu Hu, Jinrong Hu, Siwei Lyu, Xi Wu, and Xin Wang, *A Self-Learning Multimodal Approach for Fake News Detection*, proposing a self-learning model that leverages contrastive learning and large language models to jointly analyze text

and image data for improved fake news classification.

20. R. Jadhav, V. Meshram, A. Bhosle, K. Patil, S. Dash, and S. Jadhav, *Explainable Multilingual and Multimodal Fake-News Detection: Toward Robust and Trustworthy AI for Combating Misinformation*, presenting a hybrid explainable deep learning architecture for multilingual and multimodal misinformation detection to enhance robustness and transparency.