



International Journal of Innovative Research in Science Engineering and Technology (IJIRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 9.935

Volume 15, Issue 6, June 2026

Hybrid Evidence – Grounded AI for Indian News Verification

Deepak Kumar.V

Student, Department of Computer Science and Engineering, Sri Venkateswara College of Engineering and Technology,
Thiruvallur, Tamil Nadu, India

ABSTRACT: The rapid growth of digital media has significantly accelerated the spread of misinformation, making the timely verification of news a critical challenge. Existing fake news detection systems often rely solely on machine learning models or static datasets, limiting their ability to validate claims against real-world evidence and effectively process multilingual content. This paper presents a Hybrid Evidence-Grounded AI System for verifying the authenticity of Indian news by combining large language model (LLM)-based reasoning with real-time evidence retrieval. The proposed framework utilizes the Google Gemini API to perform comprehensive textual analysis and cross-validates extracted claims using articles retrieved from trusted Indian news sources, including The Times of India, NDTV, and The Hindu, through NewsAPI, GDELT, and Google News RSS feeds. To address India's linguistic diversity, the system supports multiple Indian languages by translating regional-language inputs into English for consistent semantic evaluation. The verification process generates confidence scores, evidence-backed explanations, and source references, enabling transparent and interpretable decision-making. By integrating natural language understanding, multilingual processing, real-time evidence retrieval, and structured data storage, the proposed system offers a scalable and practical approach for combating misinformation in India's multilingual digital ecosystem.

KEYWORDS: Fake News Detection, Evidence-Grounded Artificial Intelligence, Google Gemini API, Real-Time Evidence Retrieval, Misinformation Detection, News Verification, Cross-Source Validation, Indian News Ecosystem.

I. INTRODUCTION

The rapid growth of digital communication platforms has transformed the way news and information are created, shared, and consumed. While online news portals and social media have made information more accessible, they have also accelerated the spread of fake news and misinformation. False information can mislead the public, influence political and social decisions, create panic, and undermine trust in credible media sources. Consequently, the automatic detection and verification of fake news has become an important research challenge.

Traditional fact-checking methods rely on manual verification by experts using multiple trusted sources. Although accurate, these methods are time-consuming and cannot cope with the massive volume of digital content generated every day. Recent advances in Artificial Intelligence (AI), Natural Language Processing (NLP), and Large Language Models (LLMs) have enabled automated fake news detection systems. However, many existing approaches focus primarily on text classification using static datasets and do not validate news claims against real-time evidence. As a result, their performance is often limited when dealing with emerging events or newly published information.

The challenge is even greater in India due to its multilingual environment and diverse digital news ecosystem. News is published in English as well as several regional languages, while many existing verification systems are designed mainly for English-language content and do not effectively utilize trusted Indian news sources. This limitation reduces their ability to identify misinformation that spreads through regional media platforms.

To address these challenges, this paper proposes a Hybrid Evidence-Grounded AI System for Indian news verification. The proposed system combines the reasoning capabilities of the Google Gemini API with real-time evidence retrieved from trusted Indian news sources such as The Times of India, NDTV, and The Hindu using NewsAPI, GDELT, and Google News RSS feeds. The framework also supports multiple Indian languages by translating regional-language inputs into English before semantic analysis. The final verification result is presented with a confidence score and evidence-based explanation, enabling transparent and reliable news verification. By integrating AI-based reasoning,

multilingual processing, and live evidence retrieval, the proposed system provides a scalable approach for combating misinformation in India's digital ecosystem.

The remainder of this paper is organized as follows. Section II reviews the related work on fake news detection and evidence - based verification. Section III presents the proposed methodology and system architecture. Section IV discusses the implementation and experimental results. Finally concludes the paper and outlines future research..

OBJECTIVES OF THE STUDY

The primary objective of this study is to develop a hybrid, evidence-grounded artificial intelligence system for verifying the authenticity of Indian news articles using both AI-based reasoning and real-time evidence retrieval. The specific objectives are:

1. To develop a hybrid news verification framework that combines Large Language Model (LLM)-based reasoning with real-time evidence collection.
2. To utilize the Google Gemini API for semantic analysis and contextual understanding of news claims.
3. To retrieve and analyze supporting evidence from trusted Indian news sources through NewsAPI, GDELT, and Google News RSS.
4. To support multilingual news verification by translating regional Indian language content into English for consistent analysis.
5. To generate confidence scores and evidence-based explanations to improve the transparency and interpretability of verification results.
6. To provide a user-friendly interface for submitting news text or URLs and displaying verification reports.
7. To maintain a structured database of verification results for future reference, auditing, and system improvement.

II. REVIEW OF RELATED LITERATURE

The increasing spread of misinformation across digital platforms has led to significant research on automated fake news detection. Early approaches primarily relied on traditional machine learning algorithms such as Naïve Bayes, Support Vector Machines (SVM), Random Forest, and Logistic Regression, where handcrafted linguistic and statistical features were used to classify news articles as genuine or fake. Although these methods achieved satisfactory performance on benchmark datasets, they often struggled to generalize to unseen events and rapidly evolving news topics.

With the advancement of deep learning, researchers introduced neural network-based models, including Long Short-Term Memory (LSTM) networks, and Transformer-based architectures. These models automatically learn semantic representations from text and generally outperform traditional machine learning techniques. contextual reasoning, and text generation, making them promising tools for fake news detection and claim verification.

Several studies have emphasized that content-based classification alone is insufficient for reliable fact-checking because the authenticity of a claim often depends on external evidence. Consequently, evidence-based verification systems have emerged, where news claims are validated against trusted sources such as news articles, knowledge graphs, and fact-checking databases. These approaches improve transparency by providing supporting or contradicting evidence rather than relying solely on classification scores. However, many existing systems depend on static datasets or international news repositories. The challenge becomes more significant in the Indian context due to linguistic diversity and the availability of news in multiple regional languages. Most existing fake news detection models are designed primarily for English-language content and provide limited support for multilingual verification. Furthermore, relatively few systems integrate AI-based reasoning with live evidence retrieval from trusted Indian news portals, reducing their effectiveness in identifying misinformation circulating within the Indian digital ecosystem.

To address these limitations, the proposed work presents a Hybrid Evidence-Grounded AI System that combines the reasoning capabilities of the Google Gemini API with real-time evidence retrieval from trusted Indian news sources, including The Times of India, NDTV, and The Hindu, using NewsAPI, GDELT, and Google News RSS feeds. By integrating multilingual processing, semantic analysis, and evidence-grounded verification, the proposed framework aims to provide more accurate, explainable, and scalable fake news detection for India's multilingual information environment.

The reviewed literature demonstrates that artificial intelligence and natural language processing have significantly improved the automation of fake news detection. However, existing methods often suffer from limitations such as dependence on historical datasets, insufficient multilingual support, and the absence of real-time evidence verification. Based on these observations, the proposed Hybrid Evidence-Grounded AI System integrates Large Language Model (LLM)-based reasoning with live evidence retrieval from trusted Indian news portals to provide accurate, transparent, and explainable news verification. This approach addresses the identified research gaps and offers a scalable solution for combating misinformation in the Indian digital ecosystem.

III. METHODOLOGY

A. The proposed Hybrid Evidence-Grounded AI System follows a structured pipeline that processes user input, extracts meaningful claims, retrieves supporting evidence, and generates a final verification result with confidence scoring. The system is designed to integrate Large Language Model (LLM) reasoning with real-time information retrieval to ensure accurate and explainable news verification.

B. The process begins when the user provides either a news text as input. The system first performs language detection to identify whether the input is in English or any regional Indian language. If a non-English language is detected, the text is translated into English using a translation module to ensure uniform processing and semantic consistency across different languages.

C. After preprocessing, the translated or original English text is passed to the Google Gemini API for AI-based verification. The model performs semantic analysis of the content, extracts key factual claims, and evaluates their contextual meaning. This stage enables the system to understand the intent and factual structure of the input news.

D. In parallel, the system performs live evidence retrieval using APIs such as NewsAPI, GDELT, and Google News RSS. Relevant articles from trusted news sources, including *The Times of India*, *NDTV*, and *The Hindu*, are collected and filtered based on relevance to the extracted claims. These sources provide supporting or contradicting evidence for verification.

E. The AI-generated interpretation and retrieved evidence are then combined in a fusion layer, where consistency between claims and external sources is analyzed. Based on this comparison, the system computes a confidence score and determines the final verification outcome, categorizing the news as likely true, false, or uncertain. This ensures that the decision is not solely dependent on the language model but is grounded in real-world evidence.

F. The operational flow of the system can be summarized as follows: the process starts with user input (text), followed by language detection and translation (if required), then AI-based verification using the Gemini API, parallel retrieval of live evidence from news APIs, fusion of AI reasoning with retrieved data for confidence scoring, and finally display of the result along with report generation.

IV. RESULTS AND DISCUSSION

A. The proposed Hybrid Evidence-Grounded AI System was evaluated based on its ability to correctly verify news claims using a combination of Large Language Model (LLM) reasoning and real-time evidence retrieval. The system was tested using a diverse set of news inputs, including factual statements, misleading claims, and ambiguous articles collected from online news sources and social media platforms.

B. The results indicate that the system is capable of effectively distinguishing between true, false, and uncertain news statements by analyzing contextual meaning and comparing it with evidence retrieved from trusted news portals. The integration of Google Gemini API enabled accurate semantic understanding of complex news text, while the use of real-time data sources such as NewsAPI, GDELT, and Google News RSS improved factual validation by providing up-to-date supporting or contradicting information.

C. A key strength of the system is its ability to generate confidence scores along with explanations for each decision. This improves transparency and helps users understand the reasoning behind the classification outcome. In cases where sufficient evidence was available, the system demonstrated high confidence in classification. However, for emerging or

less-covered news topics, the confidence score was comparatively lower, indicating uncertainty due to limited external verification data.

D. The inclusion of multilingual support also improved the system's applicability in the Indian context. Inputs in regional languages were successfully translated into English before processing, ensuring consistent analysis across different linguistic inputs. This feature significantly enhances the usability of the system in a diverse multilingual environment.

E. Despite its effectiveness, certain limitations were observed. The system's performance depends heavily on the availability and reliability of external news sources. In some cases, delays in API responses or lack of relevant articles affected the completeness of evidence retrieval. Additionally, while the Gemini API provides strong reasoning capabilities, it may still produce inconsistent interpretations for highly complex or sarcastic content.

F. Overall, the results demonstrate that combining LLM-based reasoning with real-time evidence retrieval significantly improves the accuracy and interpretability of fake news detection systems. The proposed approach provides a scalable and practical solution for combating misinformation in real-world scenarios, particularly within the Indian digital news ecosystem.

VI CONCLUSION

A. This paper presented a Hybrid Evidence-Grounded AI System for the verification of Indian news, designed to address the growing challenge of misinformation in digital media. The proposed approach integrates Large Language Model (LLM)-based reasoning using the Google Gemini API with real-time evidence retrieval from trusted news sources such as The Times of India, NDTV, and The Hindu through APIs including NewsAPI, GDELT, and Google News RSS. The system also incorporates multilingual support to handle regional Indian languages by translating input text into English before analysis..

B. The results demonstrate that combining AI-based semantic understanding with live evidence retrieval significantly improves the reliability, transparency, and interpretability of fake news detection. The system is capable of generating confidence scores and evidence-based explanations, making the verification process more understandable and trustworthy for end users. This hybrid approach reduces the limitations of traditional classification-based models that rely only on static datasets.

C. Overall, the proposed system provides a scalable and practical framework for combating misinformation and strengthening trust in digital news ecosystems, particularly in a multilingual and diverse country like India.

VI. ACKNOWLEDGEMENTS

A. I would like to express my sincere gratitude to all those who supported me throughout the development of this project.

B. I am deeply thankful to my guide and faculty members for their valuable guidance, continuous encouragement, and insightful suggestions that helped me successfully complete this work. Their expertise and support played a crucial role in shaping the direction of this project.

REFERENCES

1. J. Zhang, C. Wang, and H. Li, "Fake News Detection: A Survey of Methods, Challenges, and Opportunities," *IEEE Access*, vol. 9, pp. 12032–12050, 2021
2. A. Sharma and R. Singh, "Natural Language Processing Approaches for Fake News Detection," **International Journal of Computer Applications**, vol. 182, no. 8, pp. 1–6, 2019
3. K. Shu, S. Wang, and H. Liu, "Beyond News Contents: The Role of Social Context for Fake News Detection," in **Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining**, pp. 1137–1146, 2017.
4. V. Gupta and P. Kumar, "Multilingual Fake News Detection Using Deep Learning Techniques," **Springer Journal of Information Processing**, vol. 15, no. 3, pp. 245–256, , 2022.



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  ijirset@gmail.com



www.ijirset.com

Scan to save the contact details