iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** IV **Month of publication:** April 2026

DOI: https://doi.org/10.22214/ijraset.2026.80330

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue IV Apr 2026- Available at www.ijraset.com

A Real-Time Multilingual Fake News Detection

Ojaswi Bangre[1], Tanvir Malnas[2], Bhuwan Kanampalliwar[3], Ritesh Meshram[4], Tushar Pawar[5], Mahendra Patil[6], Sushama Telrandhe[7]

[1,2,3,4,5]*U.G. Student, Department of Computer Science and Engineering, Guru Nanak Institute of Engineering and Technology, Nagpur, Maharashtra, India.*

[6]*Professor, Department of Computer Science and Engineering, Guru Nanak Institute of Engineering and Technology, Nagpur, Maharashtra, India.*

[7]*Associate professor, Department of Computer Science and Engineering, Guru Nanak Institute of Engineering and Technology, Nagpur, Maharashtra, India.*

Abstract: ***The rapid expansion of information across the internet has made it increasingly difficult for users to differentiate between credible and misleading news. Conventional approaches to fake news detection typically rely on rule-based techniques or supervised machine learning models. While these methods can perform effectively on specific datasets, they often struggle to generalize across diverse contexts and are limited in their ability to identify newly emerging misinformation, particularly in multilingual environments.***

The proposed system addresses these challenges by analyzing multiple content modalities, including textual data, URLs, and images, to better understand the underlying narrative of news items. It verifies claims by comparing them against reliable and authoritative sources during content processing. In cases where no prior fact-check exists, the system employs advanced AI-driven reasoning mechanisms to assess the credibility of the information. Recognized for its capability to identify and mitigate misinformation, the framework is adaptable across various languages and digital platforms.

This paper presents a detailed overview of the system's architecture, operational workflow, implementation strategy, and evaluation results. It highlights how the integration of large language models with robust fact-verification techniques significantly enhances the accuracy and reliability of fake news detection.

Keywords: Fake News Detection, Artificial Intelligence, Fact Verification, Multilingual Processing, Misinformation, Large Language Models, Gemini API

I. INTRODUCTION

The proliferation of digital platforms has made misinformation one of the most critical challenges in the modern information landscape. False or misleading content can spread rapidly across social media, influencing public perception, shaping political discourse, and in some cases triggering social unrest or panic. Such widespread dissemination of inaccurate information can lead to public confusion and contribute to broader societal and political instability.

Existing fake news detection systems are largely based on machine learning models trained on static datasets. While these approaches perform effectively when identifying familiar patterns, they often struggle to adapt to evolving and previously unseen forms of misinformation. Furthermore, linguistic diversity presents an additional challenge. Many current systems are primarily designed for English, limiting their effectiveness in multilingual regions such as India, where regional languages are widely used and underrepresented in detection frameworks.

To address these limitations, this study proposes a real-time, AI-driven fake news detection system capable of processing and analyzing content across multiple languages. The system leverages advanced language understanding capabilities through the Gemini API to interpret context, translate text when required, and evaluate claims more effectively. Additionally, it incorporates external fact-checking services, such as Google Fact Check Tools, to validate information against credible sources and enhance detection reliability.

II. RELATED WORK

A substantial body of research has explored misinformation detection using traditional machine learning, deep learning, and linguistic analysis techniques. Early approaches relied on classical algorithms such as Naïve Bayes, Support Vector Machines (SVM), and Decision Trees, which primarily perform binary or multi-class classification tasks. While these methods can identify known patterns of misinformation within specific contexts, they often fail to generalize when confronted with novel or evolving forms of false information.

©IJRASET: All Rights are Reserved

The introduction of transformer-based architectures, including BERT, RoBERTa, and XLNet, significantly enhanced the ability of systems to capture contextual relationships within text. These models improved the semantic understanding of content and achieved better performance compared to traditional techniques. However, they remain heavily dependent on large, labeled datasets for training, which limits their adaptability to real-time changes in misinformation trends. Additionally, their multilingual capabilities are often insufficient for effectively handling content across diverse regional languages, particularly in linguistically rich regions such as India.

Recent advancements in large language models (LLMs) have opened new avenues for reasoning-based misinformation detection. Models such as Gemini enable deeper semantic interpretation and support multiple languages, including Hindi, Marathi, Bengali, and Tamil. Despite these advancements, there is still a lack of comprehensive systems that effectively integrate multilingual reasoning capabilities of LLMs with real-time fact-checking mechanisms through external APIs.

This gap highlights the need for a more robust and adaptive approach. The system proposed in this paper addresses this limitation by combining multilingual LLM-based reasoning with real-time fact verification, thereby providing a more practical and scalable solution for accurate fake news detection in dynamic and diverse information environments.

III. PROPOSED SYSTEM ARCHITECTURE

The proposed system is designed using a modular architecture capable of handling multiple input formats, including plain text, URLs, and image-based content processed through Optical Character Recognition (OCR). The workflow begins with an input processing layer, which standardizes incoming data by performing noise reduction and automatic language detection to ensure consistency across diverse inputs. A central component of the architecture is the Gemini API module, which performs advanced semantic analysis of the content. It interprets the underlying meaning, extracts relevant factual claims, and translates them into English when necessary to facilitate accurate verification across different languages.

Following claim extraction, the system employs a dual verification strategy. The first approach involves an external fact-verification module that queries reliable sources such as Google Fact Check Tools and NewsAPI to identify existing fact-checks, corresponding verdicts, and supporting references. In situations where no prior fact-check is available, the system activates an AI-based reasoning engine that evaluates the credibility of the claim using contextual understanding and logical inference.

The final stage consists of a classification and reporting layer, where the system aggregates all available evidence to categorize the content into classes such as *True*, *False*, or *Misleading*. It also generates a clear, user-friendly explanation of the decision, ensuring that the output is accessible and, when required, translated into the user's preferred language.

Following claim extraction, the system adopts a dual verification strategy to assess the credibility of the information. The first approach involves an external fact-verification module that queries reliable sources such as Google Fact Check Tools and NewsAPI to retrieve existing fact-checks, including verified verdicts and supporting source links.

In cases where no prior fact-check is available, the system activates an AI-driven reasoning mechanism that independently evaluates the claim's credibility based on contextual understanding and logical consistency.

The final stage of the process is the classification and reporting layer, which consolidates the gathered evidence to categorize the news into classes such as *True*, *False*, or *Misleading*. Additionally, the system provides a clear and concise explanation of the decision, ensuring transparency and interpretability for the end user.

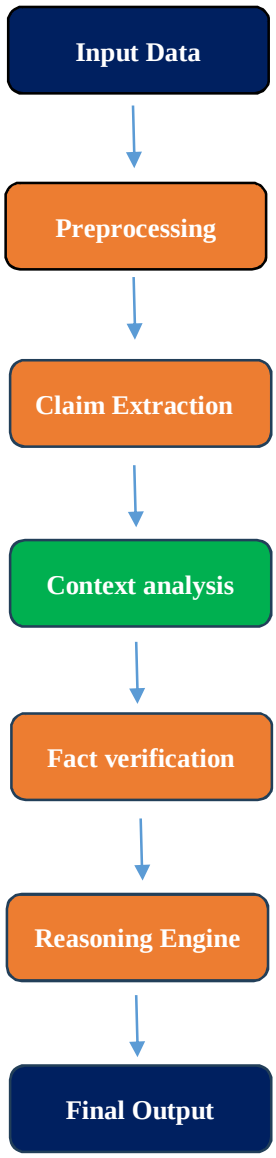
IV. IMPLEMENTATION AND METHODOLOGY

The backend of the proposed system is implemented using Node.js, which efficiently manages API interactions and communication with external services. The processing pipeline begins with input preprocessing, where the system performs language detection, noise removal, and, when required, translation to ensure uniformity in the input data.

Once the input is standardized, the extracted content is forwarded to the Gemini API for semantic analysis and claim extraction. The identified claims are then cross-verified against established fact-checking sources. For this purpose, the system integrates with trusted platforms such as Google Fact Check Tools, NewsAPI, and other reputable organizations that provide verified information through RSS feeds. If a matching fact-check is found, the system retrieves and returns the corresponding verified verdict along with supporting references. In scenarios where no existing verification is available, the system employs an advanced reasoning phase using the Gemini model. During this stage, the claim is analyzed for logical consistency, contextual alignment, and linguistic patterns commonly associated with misleading or false information. Based on this evaluation, the system assigns a credibility label along with a confidence score.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

The final output presented to the user includes the classification result, supporting evidence, a detailed explanation of the reasoning process, and translated content where necessary, ensuring both transparency and accessibility.



Implementation and Methodology Diagram

V. EXPECTED RESULTS AND DISCUSSION

A diverse set of materials—including statements created by users, confirmed examples of misinformation, posts in multiple languages, and authentic news reports—was used to evaluate the system. The results showed that it performed reliably and delivered accurate outputs. Even without relying on external fact-checking sources, the system demonstrated strong capability in identifying claims within text. The reasoning component provided clear and detailed explanations, helping users understand why a claim was labeled as true or false. One of its key strengths is its support for over 100 languages, which makes it especially well-suited for linguistically diverse regions such as India. Overall, the system proved to be highly useful.

There are, however, certain limitations to consider. Because the system depends on external APIs, temporary outages or slow response times can affect its accuracy. Additionally, statements that are satirical or lack clear context may sometimes be misclassified. Variations between different versions of the Gemini model can also lead to slight inconsistencies in the results. Despite these challenges, the system's overall performance indicates that the hybrid approach is reliable, adaptable, and well-suited for practical use in multilingual settings.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

VI. CONCLUSION AND FUTURE WORK

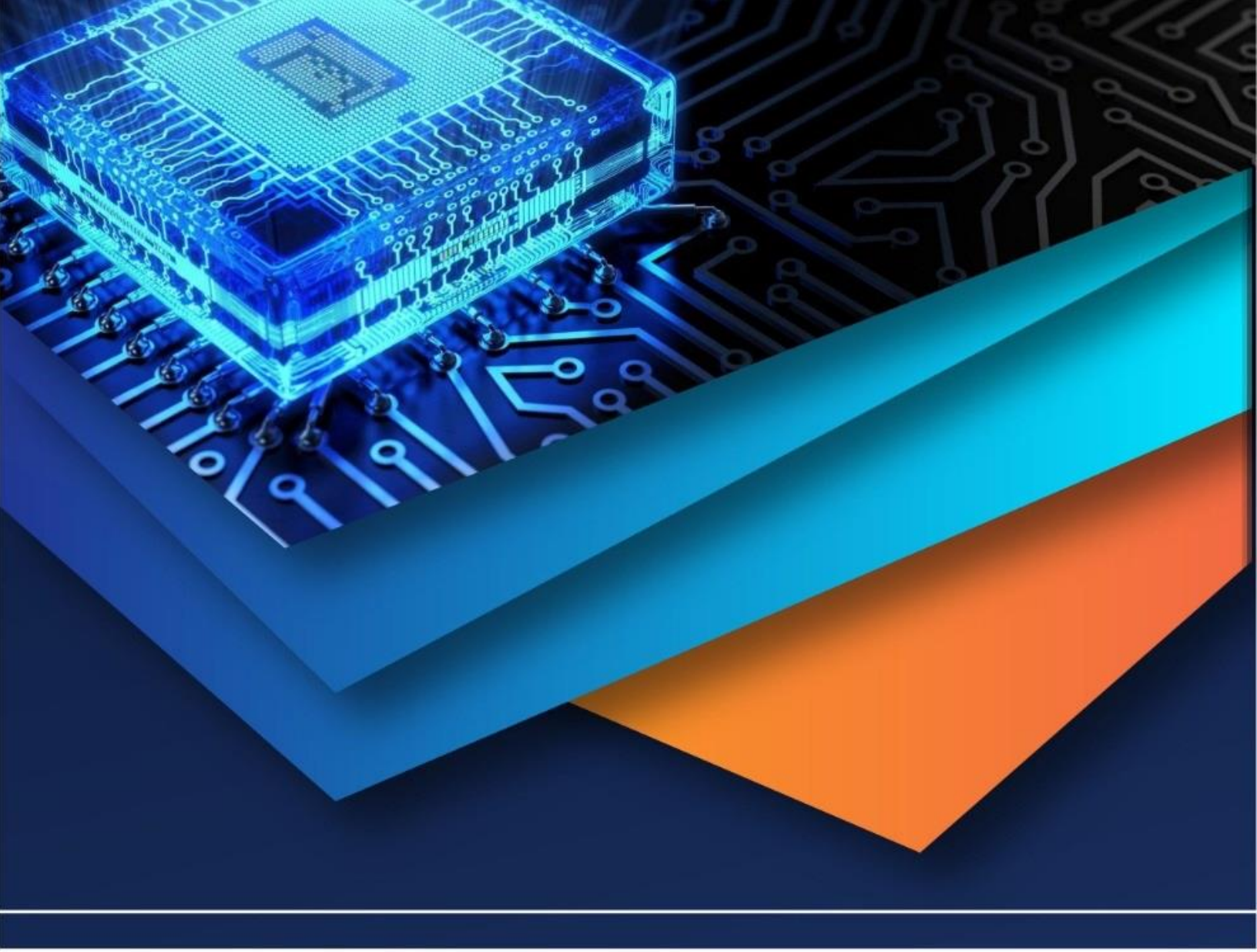
This study demonstrates that integrating LLM-based reasoning with real-time fact-checking sources offers an effective and scalable method for detecting misinformation. With features such as multilingual understanding, claim identification, contextual evaluation, and evidence verification, the system remains highly adaptable to evolving forms of disinformation.

Several enhancements could further strengthen the system. These include developing a browser extension for instant fact-checking, enabling integration with messaging platforms like WhatsApp, adding capabilities to detect deepfake images, and introducing voice-enabled fact-checking for hands-free interaction.

Overall, the system stands out for its adaptability and multilingual capabilities, making it particularly valuable in regions with linguistic diversity. Its ability to explain decisions further enhances user confidence and usability in real-world scenarios.

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 Explorations Newsletter, vol. 19, no. 1, pp. 22–36, 2017.
[2] X. Zhou and R. Zafarani, "A Survey of Fake News: Fundamental Theories, Detection Methods, and Opportunities," ACM Computing Surveys, vol. 53, no. 5, pp. 1–40, 2020.
[3] W. Y. Wang, "'Liar, Liar Pants on Fire': A New Benchmark Dataset for Fake News Detection," in Proc. 55th Annual Meeting of the Association for Computational Linguistics (ACL), 2017.
[4] A. Vaswani et al., "Attention Is All You Need," in Proc. 31st Conference on Neural Information Processing Systems (NeurIPS), 2017.
[5] Google, "Gemini API Documentation," 2024. [Online]. Available: https://ai.google.dev
[6] Google, "Fact Check Tools API," 2023. [Online]. Available: https://developers.google.com/fact-check
[7] BOOMLive, "Fact-Check Articles Database." [Online]. Available: https://www.boomlive.in
[8] AltNews, "Fact-Checking Reports." [Online]. Available: https://www.altnews.in
[9] NewsAPI, "News Data Aggregation API Documentation." [Online]. Available: https://newsapi.org
[10] S. Kumar and N. Shah, "False Information on Web and Social Media: A Survey," arXiv preprint, arXiv:1804.08559, 2018.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)