International Journal of Computational Research in Science and Technology
E-ISSN: 3108-0871, Volume 2 Issue 5
Sep-Oct- 2026, PP 01-6

AI-Based Fake News Detection Using Machine Learning: A Review Of Multilingual Detection, Automated Fact-Checking And Messaging-Platform Misinformation

Pawar Tanushri Sarjeet[1] Prof. M. R. Kumbhoje[2] Prof. M. C. Kolse[3] Prof. S. K. Said[4]
Department of Artificial Intelligence and Data Science
Jaihind College of Engineering
Pune, Maharashtra, India

Abstract—Misinformation spreads not only through websites and social networks but also through messaging platforms and in many languages, which is especially relevant in India. This paper reviews AI-based fake news detection using machine learning with emphasis on three aspects that are less covered by English-only text classifiers: multilingual detection, automated fact-checking and misinformation on messaging platforms. It discusses the social science of fake news and the characteristic features of false articles, such as sensational titles and simpler language, and reviews machine learning classifiers trained on labelled datasets. Automated fact-checking pipelines that identify check-worthy claims, retrieve evidence and verify claims against it are examined, together with cross-lingual transformer models and resources for Indian languages. Studies of misinformation on messaging platforms and pandemic-related fake news datasets are discussed. A five-layer architecture comprising multilingual ingestion, preprocessing, classification, evidence-based verification and user-facing explanation is proposed, along with research gaps, objectives, a comparison of approaches and challenges such as low-resource languages, encrypted messaging, evolving narratives and freedom-of-expression concerns. The review concludes that combining machine learning classification with evidence-based fact-checking provides more trustworthy detection.

Index Terms—Fake news detection, machine learning, fact-checking, multilingual NLP, Indian languages, misinformation, claim verification, messaging platforms

I. Introduction

1.1 Motivation

In India, a forwarded message in Marathi, Hindi or English can reach thousands of people within hours, often through family and community groups where trust is high and verification is rare. Misinformation about health, politics or communal events has caused real harm. Lazer *et al.* [1] described the science of fake news and called for interdisciplinary research on its spread and remedies.Most detection research uses English news articles and social media posts. Effective solutions for India must handle many languages, informal and code-mixed text, and content shared on messaging platforms. They should also explain their decisions with evidence rather than simply labelling content.

This paper reviews machine learning approaches to fake news detection from these perspectives and proposes a multilingual, evidence-based architecture.

1.2 Objectives of the Review

- **O1:** To review machine learning classifiers for fake news detection.
- **O2:** To examine automated fact-checking pipelines.

International Journal of Computational Research in Science and Technology
E-ISSN: 3108-0871, Volume 2 Issue 5
Sep-Oct- 2026, PP 01-6

- **O3:** To study multilingual and messaging-platform misinformation.
- **O4:** To propose a multilingual, evidence-based detection architecture.
- **O5:** To discuss challenges and responsible deployment.

II. Overview of Key Concepts

2.1 Detection Approaches

Fig. 1 organizes approaches into style-based classification, knowledge-based fact-checking, and context-based analysis of sources and spread.

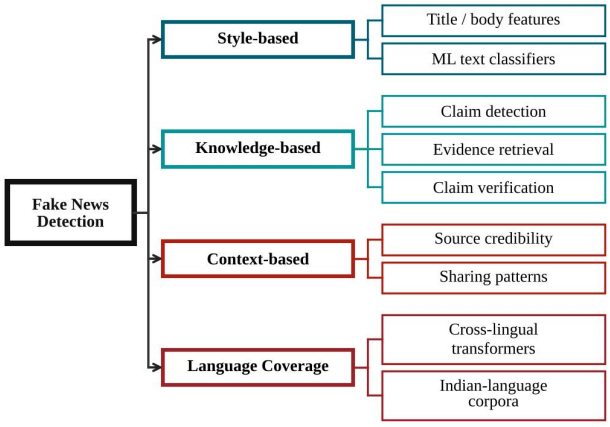


Fig. 1. Approaches to fake news detection.

2.2 Automated Fact-Checking

Fact-checking pipelines first detect claims worth checking, then retrieve relevant evidence and finally decide whether the evidence supports, refutes or is insufficient for the claim, as formalized in the FEVER task [2].

III. Review of Existing Work

Horne and Adali [3] analysed fake and real news and found that fake news titles are longer and contain more proper nouns, while body text is simpler and more repetitive, features useful for classification. Wang [4] released the LIAR dataset of fact-checked statements. Kumar and Shah [5] surveyed false information on the web and social media, covering its characteristics, impact and detection.

Hassan *et al.* [6] built ClaimBuster, which scores sentences by how check-worthy they are, helping fact-checkers prioritize. Thorne *et al.* [2] released FEVER, a dataset for verifying claims against Wikipedia evidence, which established evidence-based verification as a research task.

For languages, Conneau *et al.* [7] trained XLM-R, a cross-lingual transformer that performs well across many languages including Indian ones, and Kakwani *et al.* [8] provided Indian-language corpora and models. Patwa *et al.* [9] released a COVID-19 fake news dataset, reflecting the surge of health misinformation. Resende *et al.* [10] studied misinformation

International Journal of Computational Research in Science and Technology
E-ISSN: 3108-0871, Volume 2 Issue 5
Sep-Oct- 2026, PP 01-6

dissemination in public WhatsApp groups and proposed countermeasures, showing the importance of messaging platforms. Table I summarizes these works.

Sr. No.	Ref.	Year	Methodology	Findings	Research Gap
1	[3]	2017	Stylistic analysis of fake news	Distinct title and body features	English, small corpus
2	[4]	2017	LIAR dataset	Fact-checked statements benchmark	Short claims only
3	[6]	2017	ClaimBuster	Check-worthy claim detection	English political text
4	[1]	2018	Science of fake news	Interdisciplinary research agenda	Perspective article
5	[2]	2018	FEVER verification dataset	Evidence-based claim verification	Wikipedia evidence
6	[10]	2019	WhatsApp misinformation study	Dissemination patterns and countermeasures	Public groups only
7	[7]	2020	XLM-R cross-lingual model	Strong multilingual performance	Large model
8	[9]	2021	COVID-19 fake news dataset	Health misinformation benchmark	English only

TABLE I Summary of Literature Survey

IV. Identified Research Gaps

Labelled fake news data in Indian languages and code-mixed text are scarce.

End-to-end encrypted messaging limits platform-side detection.

Classifiers trained on past events struggle with new narratives.

Many systems give a label without supporting evidence.

Over-blocking risks suppressing legitimate speech.

V. System Architecture

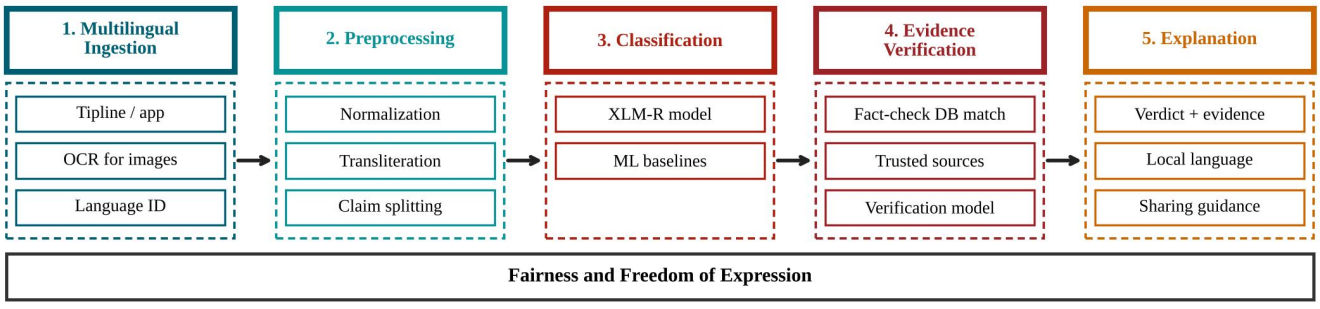


Fig. 2 shows the proposed architecture.

Fig. 2. Proposed multilingual, evidence-based fake news detection architecture.

International Journal of Computational Research in Science and Technology
E-ISSN: 3108-0871, Volume 2 Issue 5
Sep-Oct- 2026, PP 01-6

5.1 Multilingual Ingestion Layer

Users forward suspicious messages or links to a tipline or app; text is extracted from images when needed, and the language is identified.

5.2 Preprocessing Layer

Text is normalized, transliterated if written in Latin script and split into claims.

5.3 Classification Layer

A cross-lingual transformer such as XLM-R and baseline machine learning models score the likelihood that the content is misleading.

5.4 Evidence Verification Layer

Check-worthy claims are matched against fact-check databases and trusted sources; a verification model judges support or refutation.

5.5 Explanation Layer

Users receive a verdict with evidence links in their language and guidance on how to verify before sharing.

VI. Implementation Methodology

Fig. 3 describes the workflow. A claim c is verified against retrieved evidence E by a model estimating

$$P(\text{label} \mid c, E), \text{label} \in \{\text{supports, refutes, not enough info}\} \quad (1)$$

and the final verdict combines the classifier score with the verification result. Performance is evaluated with macro-F1 across languages, so that low-resource languages are not hidden by English performance.

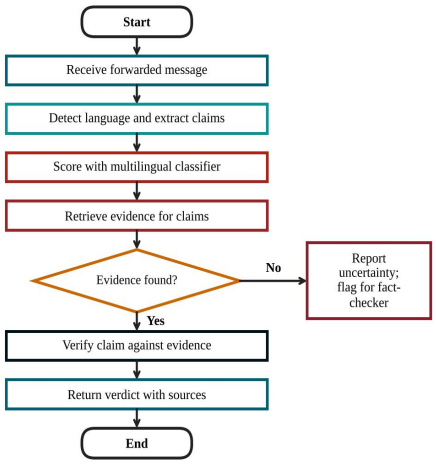


Fig. 3. Multilingual fake news verification workflow.

VII. Performance Comparison

Table II compares approaches. Style-based classifiers are fast but can be fooled, fact-checking provides evidence but needs reliable sources, and multilingual transformers extend coverage across languages.

Approach	Explainability	Language Coverage	Speed
Style-based ML	Low	Language-specific	Very fast
Monolingual BERT	Low–Medium	One language	Medium
Cross-lingual XLM-R	Low–Medium	Many languages	Medium
Evidence-based fact-checking	High	Depends on sources	Slower
Hybrid (proposed)	High	Many languages	Medium

TABLE II Comparison of Detection Approaches

VIII. Open Issues and Future Directions

Table III summarizes challenges. Future directions include voice-note and image misinformation detection, community-driven fact-checking tiplines, detection of AI-generated misinformation and media literacy programmes.

Challenge	Mitigation
Low-resource languages	Cross-lingual transfer; new datasets
Encrypted messaging	User-initiated tiplines
New narratives	Evidence retrieval over fixed labels
Over-blocking	Labels and context, not removal
AI-generated content	Provenance and detection tools

TABLE III Challenges and Mitigation

IX. Conclusion

Detecting fake news in a multilingual society requires more than an English text classifier. The reviewed literature shows that machine learning can identify stylistic signals of fake news, that automated fact-checking adds evidence and transparency, and that cross-lingual models extend coverage to Indian languages. Combining these approaches in a user-facing system that explains its verdicts can help people stop misinformation before they share it.

9.1 Acknowledgment

The authors express their sincere gratitude to the Department of Artificial Intelligence and Data Science, Jaihind College of Engineering, Pune, for providing the facilities, guidance and encouragement needed to carry out this work. The authors also thank the faculty members and colleagues whose valuable suggestions helped improve the quality of this review.

References

1. D. M. J. Lazer et al., “The science of fake news,” Science, vol. 359, no. 6380, pp. 1094–1096, 2018.
2. J. Thorne, A. Vlachos, C. Christodoulopoulos, and A. Mittal, “FEVER: A large-scale dataset for fact extraction and VERification,” in Proc. NAACL-HLT, 2018, pp. 809–819.

3. B. D. Horne and S. Adali, “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. ICWSM Workshops, 2017.
4. W. Y. Wang, ““Liar, liar pants on fire”: A new benchmark dataset for fake news detection,” in Proc. 55th Annu. Meeting ACL, 2017, pp. 422–426.
5. S. Kumar and N. Shah, “False information on web and social media: A survey,” arXiv:1804.08559, 2018.
6. N. Hassan, F. Arslan, C. Li, and M. Tremayne, “Toward automated fact-checking: Detecting check-worthy factual claims by ClaimBuster,” in Proc. 23rd ACM SIGKDD, 2017, pp. 1803–1812.
7. A. Conneau et al., “Unsupervised cross-lingual representation learning at scale,” in Proc. 58th Annu. Meeting ACL, 2020, pp. 8440–8451.
8. D. Kakwani et al., “IndicNLPSuite: Monolingual corpora, evaluation benchmarks and pre-trained multilingual language models for Indian languages,” in Findings EMNLP, 2020, pp. 4948–4961.
9. P. Patwa et al., “Fighting an infodemic: COVID-19 fake news dataset,” in Proc. CONSTRAINT Workshop (AAAI), 2021, pp. 21–29.
10. G. Resende et al., “(Mis)information dissemination in WhatsApp: Gathering, analyzing and countermeasures,” in Proc. Web Conf. (WWW), 2019, pp. 818–828.