The Scientific Temper
SEPTEMBER 2026 Vol. 17 (9): 6957-6970
ISSN: 0976-8653, E-ISSN: 2231-6396
DOI: 10.58414/SCIENTIFICTEMPER.2026.17.9.2549
A WEB OF SCIENCE JOURNAL

RESEARCH ARTICLE

Fake News Detection using CosSim Rule and Semantic Embedding

SAHAYA SHEEBA MANGALAM. J[1], R. RADHA[2]

[1]*Research Scholar, Department of Computer Science, SDNB Vaishnav College for Women & Assistant Professor, Department of Computer Science, Dr. MGR Janaki College, Vels University, Chennai, India.* sheeba.mangalam@gmail.com

[2]*Head & Associate Professor, Department of Computer Science, SDNB Vaishnav College for Women, Chennai, India.* radhasundar1993@gmail.com

Corresponding author: sheeba.mangalam@gmail.com

ABSTRACT

The proliferation of fake news on digital media platforms and online news portals is a major problem with information and public trust. With the widespread use of social media, misinformation has spread at a rapid pace and it is becoming increasingly hard to tell the truth from the fiction. Most methods used for traditional fake news detection are based on large annotated datasets and computationally expensive machine learning models, which lack transparency and interpretability. To overcome such restrictions, in this research, an explainable fake news detection framework is proposed, which uses semantic similarity, association rule interestingness and cross-source verification. The proposed approach involves web scrapping, natural language processing, semantic embedding, cosine similarity, and rule-based classification. Multiple trusted, untrusted, national, international and regional news sources are automatically collected using the news articles. The comparison is made between a reference article from an authenticated source and articles about the same event in various media outlets. The Sentence-BERT model is used to generate semantic representations of articles, and cosine similarity is used to calculate the contextual similarity between the articles. The articles are classified either as Real or Fake based on pre-defined similarity thresholds and an interpretable rule-based decision mechanism. The reliability of the generated rules is assessed by using association rule interestingness measures such as Support, Confidence and Lift. Experiments were performed using the dataset of 51 cross-source news articles, which were automatically collected, and manually organized. The results show that the semantic similarity scores and classification results are significantly better with text preprocessing. Moreover, the rules created have confidence levels between 0.76 and 0.91, and lift greater than 1, both of which are statistically significant, and both associations and classification decisions are reliable. The results demonstrate the efficiency of using semantic similarity approaches for cross-source verification and the need for explainable rule-based detection in misinformation. The proposed framework offers an alternative of transparency and low overhead to black-box machine learning models and serves as a basis for future hybrid fake news detection systems based on the combination of semantic reasoning and data-driven approaches.

Keywords—*Fake News Detection, Web Scraping, Natural Language Processing, Similarities, Rule-Based System.*

INTRODUCTION

Digital communication technologies and online news outlets are rapidly changing how information is produced, disseminated and consumed. Social media networks, news websites, blogs and online forums can spread information around the world in seconds. This accessibility has facilitated the spread of information, and awareness, but has also propelled the spread of misinformation and fake news. Spread of misinformation or disinformation can have a major impact on public opinion, political decision-making, social stability, and crisis

Received: 5th September, 2026 **Accepted: 7th September, 2026** **Published: 28th September, 2026**

management. Therefore, the identification and authentication of on-line news have become significant research problems in the areas of Natural Language Processing (NLP), Data Mining, and Information Retrieval (Zhang et al., 2024).

Most of the current methods for detecting fake news are supervised machine learning or deep learning methods. SVMs, Random Forests, Long Short-Term Memory (LSTM) networks, and transformer-based architectures have been found to perform well on classification tasks (Samadi & Momtazi, 2025). However, these approaches can be expensive to collect, process and update in the sense that they tend to require a large amount of annotated training data, a lot of computational resources, and frequent retraining when news contexts and events evolve. In addition, many of these models are used as a black-box and complicated by the difficulty of explaining the reasoning behind the predictions. These restrictions limit their use in situations where transparency, interpretability and explainability are crucial.

Another way to combat fake news is cross-source verification, comparing what is reported by different news sources for semantic consistency. In real life news reporting, it is not unusual for a number of independent news media to cover the same news events. The content of the actual reports is factual in nature, but the words and expressions vary in terms of their meaning and usage. On the other hand, articles that are misleading or fabricated tend to be significantly semantically different than articles written by trusted sources (Samadi & Momtazi, 2025). Thus, the similarity of the semantics between the different sources could serve as a pragmatic approach to assessing the credibility of news content in the Internet.

Sentence embedding models have been recently developed, enabling a dense semantic vector representation of textual content, beyond simple keyword-based matching. Sentence-BERT models, especially, give us a good way to convert news articles into short fixed-length embeddings which capture the semantic relationships between the documents (Khanam et al., 2021; Prachi et al., 2022). Cosine similarity can be computed between semantic embeddings, and the result can be used to measure the extent to which two articles are in agreement about the same event.

In this work, an explainable fake news detection framework, which relies on semantic similarity and association rule interestingness and cross-source verification, is proposed. The proposed framework differs from the traditional supervised methods in that it does not need labelled training data. Rather it uses automated web scraping to get the news articles from several sites online and compares the articles with a trusted reference article. The articles are then preprocessed in NLP techniques like tokenization, lemmatization, stop-word removal and text normalization. Sentence BERT (paraphrase-MiniLM-L6-v2) is then used to create semantic embeddings for each of the sentences, after which cosine similarity is calculated between the reference article and each of the comparison articles (Liu et al., 2024).

A rule-based classification mechanism is then used to classify an article as Real or Fake, using pre-defined similarity thresholds. To assess the reliability and significance of the rules found, association rule interestingness measures such as Support, Confidence, and Lift are integrated into the framework. These are useful for measuring the quality of the rules and are useful for finding interpretive semantic connections between articles (Zhang et al., 2023).

The proposed framework was validated using a dataset consisting of 51 news articles scraped automatically by web scraping from trusted and untrusted, national and international, government and regional news sources (Gorai & Shaw, 2025). Experimental validation shows the effect of the text preprocessing on the semantic similarity, the efficiency of the classification based on cosine similarity and the reliability of the association rule measures. The study also identifies the advantages and disadvantages of semantic similarity approach for detecting fake news in real-world news.

The key findings of this research are summarized as follows:

- A semantic similarity-based and cross-source verification method for explaining fake news is proposed.
- The automated web scraping pipeline is created to gather news articles from a variety of online news outlets.
- We use sentence-BERT embeddings and cosine similarity to calculate the similarity in meaning between news articles.
- The transparent rule-based classification mechanism is aimed to detect potentially fake news without the need of labeled training data.
- Association rule interestingness measures such as Support, Confidence, and Lift are incorporated to assess the quality and confidence of the rules generated.
- The proposed approach is validated experimentally with a set of 51 news articles obtained from various news domains, proving the applicability of the proposed approach for explainable misinformation detection.

The rest of this paper is organized as follows: Related work on fake news detection and existing research gaps are discussed in Section II. In section III, the proposed semantic similarity-based framework and its mathematical formulation is presented. The experimental setup, generated rules, association rule analysis and evaluation results are discussed in Section IV. Finally, Section V summarizes the paper and suggests future research areas for hybrid explainable fake news detection systems.

RELATED WORK

The development of digital media and social networking networks has intensified the information flow and multiplied the spread of fake news and hoaxes many times. This has been a major social and political problem because false information may mislead the mass and affect the decision making process. As a result, machine learning (ML), natural language inference (NLI), and multimodal methods have emerged as major fields of research in information technology to identify fake news automatically.

Initial studies mainly concerned the traditional ML-based methods. Bonny et al. (2022) used the classifiers of Logistic Regression (LR), Decision Tree (DT), Random Forest (RF), Gradient Boosting (GB), XGBoost, Multinomial Naive Bayes (MNB), and K-Nearest neighbors (KNN) on a 44,898 articles benchmark set of English news. With TF-IDF and Count Vectorizer characteristics, the highest accuracy of 93.86 was obtained by using Logistic Regression. Equally, Pande et al. (2022) used regression-based classification using modified word embeddings and attention mechanisms as reportedly achieve better results than LSTM and baseline ML models on campaigns using various datasets. In a case studied by Kaur (2022), web scraper was used with ML classifiers like Naive Bayes and Logistic Regression to analyze live news material with sentiment analysis used to enhance classification. Khanam et al. (2021) made a comparative analysis of several supervised learning algorithm (Naive Bayes, SVM, KNN, Neural Networks, Random Forest) and they proved that ensemble and hybrid models have the highest accuracy, and it is increased when they are supplemented by quantitative linguistic features such as part-of-speech distributions and sentence-level statistics.

Using web-scraped articles in Indonesian language, Wibowo et al. (2021) gathered 1116 articles and preprocessed them using NLP-based methods and then used a wide variety of different classifiers, including SVM, Random Forest, Decision Tree, KNN, Stochastic Gradient Descent (SGD), AdaBoost, Gradient Boosting, Bagging, Nearest Centroid, and MLP-ANN. Random Forest had maximum accuracy of 89 and SVM, Gradient Boosting, AdaBoost, SGD and Decision Tree had above 80 accuracy implying effectiveness of text classification using ensemble learning with NLP.

In order to address the weakness of the fixed ML models, current research has discussed knowledge-based verification systems. The version presented by Shah et al. (2024) called VERITAS-NLI, which involves the use of automated web scraping alongside Natural Language Inference to prove the factual adequacy of news headlines. The system was very successful in generalization, with an accuracy of 84.3% being higher than classical ML and Transformer-based models, such as BERT, because of semantic reasoning and dynamic evidence retrieval, which allowed the system to access external evidence and ignite the process.

In addition to text, there has been the rise of multimodal so as to identify deceptive material through text and visual means. El Madbouly et al. (2023) integrated web scraping, text summarization with the help of BART, PySS3-based text classification, image captioning with the help of a Vision Encoder-Decoder model, and image-text matching to detect inconsistency, which enhanced the strength of detection in the event of the use of images to falsely influence credibility.

Table I: *Comparison of Fake News Detection Approaches*

Study	Approach Type	Techniques / Features	Dataset / Language	Best Performance	Key Strengths	Limitations
Bonny *et al.* (2022)	ML-based (Text)	LR, DT, RF, GB, XGB, MNB, KNN; TF-IDF & Count Vectorizer	44,898 English news articles	93.86% (LR)	High accuracy on benchmark dataset; computationally efficient	Static data; poor generalization to unseen content
Pande *et al.* (2022)	ML-based (Text + Word Embedding)	Regression with modified attention & word2vec embedding	Multiple English datasets	Outperforms LSTM and baseline ML	Enhanced feature representation; robust to multiple datasets	Requires embedding tuning; dependent on training quality
Kaur (2022)	ML-based (Text + Sentiment)	Naïve Bayes & LR; web scraping + sentiment features	Live news from multiple sources	Optimized accuracy, precision, recall	Real-time data acquisition; sentiment analysis adds context	Sentiment alone may miss factual inconsistencies
Khanam *et al.* (2021)	ML-based (Text +	Naïve Bayes, SVM, KNN, Neural	Multiple datasets	Up to 99% (ensemble +	Ensemble learning + quantitative	May require extensive

	Statistical Features)	Networks, RF; POS & text statistics		feature enrichment)	linguistic features improve robustness	feature engineering
Wibowo *et al.* (2021)	ML-based (Text, Web-scraped)	SVM, RF, DT, KNN, SGD, AdaBoost, Gradient Boosting, Bagging, Nearest Centroid, MLP-ANN	1,116 Indonesian news articles	89% (RF)	Ensemble methods with NLP effective; supports non-English content	Dataset size limited; heavy preprocessing required
Shah *et al.* (2024)	Knowledge-based (NLI + Web Scraping)	VERITAS-NLI; automated web scraping + entailment reasoning	Multi-domain headlines	84.3%	Semantic reasoning; dynamic evidence retrieval; generalizes better	Dependent on evidence quality; computationally intensive
El Madbouly *et al.* (2023)	Multimodal	Web scraping + BART summarization + PySS3 classification + Vision Encoder-Decoder + image-text matching	Web pages with text & images	Subjectively validated; reliable	Detects visual-textual inconsistencies; robust for image-rich content	Computationally expensive; sensitive to missing modalities

Traditional ML methods have high accuracy in benchmark datasets but lack in generalization and typically use handcrafted features. NLI-based methods offer semantic validation and require correct evidence retrieval and cost more computations. Multimodal approaches are better because they utilize visual-textual connections but demand intricate data matching and bigger datasets.

Although progress has been made in the fields of ML, NLI, and multimodal methods, current solutions are mostly independent of each other. Still, there is a necessity of an integrated framework that combines real-time web scraping, semantic reasoning, and hybrid ML-based classification that can be reliable enough to identify fake news in an ever-changing and heterogeneous online content. The presence of such a gap is what drives the proposed approach that focuses on integrating dynamic evidence retrieval, statistical linguistic features, and multimodal verification in a scalable system.

The remaining section of the paper is organized as follows. Section II is a literature review of the available research on the topic of fake news detection, which shows the limitations of the existing methods and the necessity of interpretable and cross-source validation systems. Section III outlines the methodology which covers the architecture, modules and implementation of the proposed rule-based semantic similarity framework. Section IV provides the experimental configuration, assessment measures and findings of the testing of the system on real news articles. Finally, Section V summarizes the main findings and giving a reflection of the possible focus of future research, such as developing hybrid models that integrate rule-based reasoning and machine learning approaches.

TECHNIQUES USED

The proposed framework is based on cross-source verification by semantic similarity analysis and association rule interestingness measures for fake news detection. The framework doesn't need labeled training data, unlike traditional supervised machine learning methods. Rather, it considers the semantic consistency of a set of news articles gathered from different Internet sources regarding a reliable reference article (Samadi & Momtazi, 2025; Jain et al., 2022).

The framework is developed as a pipeline which consists of several modules, including web scraping, natural language processing, semantic embedding, similarity computation, and rule-based classification (Meesad, 2021). It is a system that does not need labelled training data to work and is aimed at cross-source verification of the news articles reporting the same event (Liu et al., 2024). The general methodology takes into consideration the following major modules.

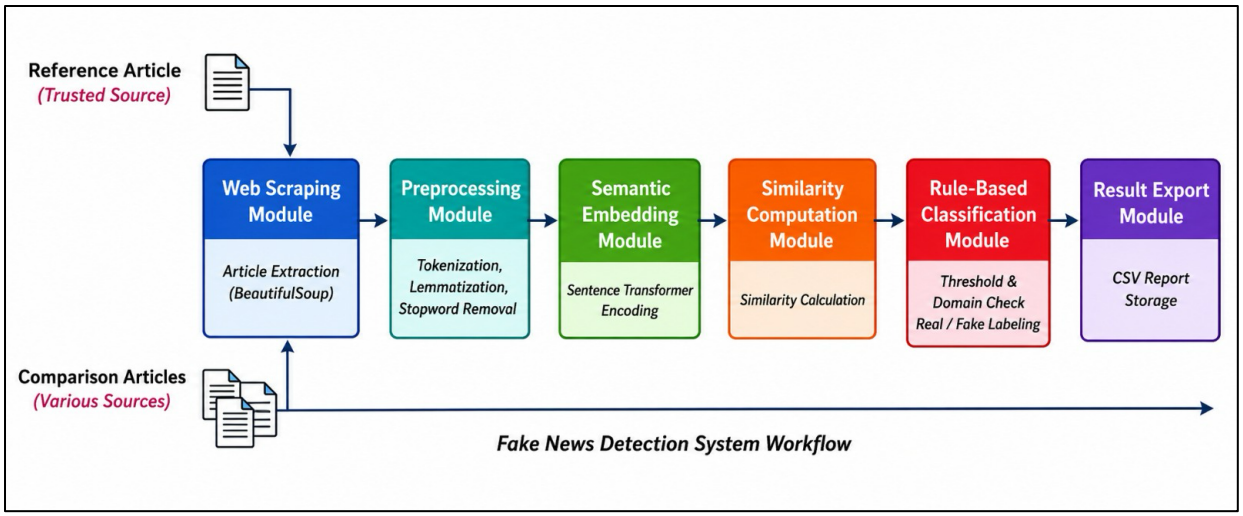


Figure 1 Fake News Detection System Flowchart

The modular architecture will make it transparent, interpretable and easy to extend to hybrid approaches in future (Jain et al., 2022). The proposed framework of the fake news detection system architecture is based on a modular pipeline design that is sequential. The architecture starts with the Web Scraping Module which will gather the news articles pertaining to a particular event on various online news portals (Kumar et al., 2025). One article of an authenticated and credible source is referred to as the reference article and the rest are considered comparison articles.

Web scraping module extracts news articles from various online news portals. Content of the articles is accessed through HTTP request and the HTML layout of the web pages are extracted through BeautifulSoup library. The extraction of textual content is done mainly on the paragraph tags in order to have the main content of every article. The reference article is one piece of authenticated and credible news, whereas more articles covering the same event are obtained via different online sources to be compared (Mhatre & Masurkar, 2021).

Preprocessing is done on the extracted raw text to enhance the quality and consistency of the textual information. The spaCy library is used to apply natural language processing techniques. This module does tokenization where text is broken down into single tokens, lemmatization where words are broken down to their base forms, and stopwords are removed which are common and non-informative words (Kumar & Singh, 2022; Prachi et al., 2022). Also, non-alphabetic tokens, punctuations and unnecessary symbols are filtered out.

The scraped textual content is then sent to the Text Preprocessing Module where it undergoes natural language processing which includes the processes of tokenization, lemmatization, stopword removal and non-alphabetic token filtering. The purged output of the text is then sent to the Semantic Embedding Module which then transforms each article into a dense vector representation through a pre-trained Sentence Transformer model (Gorai & Shaw, 2025).

Let

$U = \{u_1, u_2, \ldots, u_n\}$ be the set of URLs corresponding to news articles reporting the same event.

$u_r \in U$ be the reference article URL from an authenticated source.

The web scraping process can be modelled as a mapping function:

$$W: U \rightarrow T \quad (1)$$

where

$T_i = W(u_i)$

T_i represent the raw textual content extracted from URL u_i

For each URL u_i:

$$T_i = \bigcup_{j=1}^{m_i} P_{ij} \quad (2)$$

where

- P_{ij} is the j-th paragraph of article i.
- m_i is the total number of paragraph tags in the article.

Thus, the reference article content is:

$$d_r = T_i(u_r)$$

and the set of comparison articles is:

$$T_r = W(u_r) \quad (3)$$

$$T_c = \{T_1, T_2, \ldots, T_n\}$$

The following preprocessing steps can be used to minimize noise and increase the usefulness of semantic representation.

Let d_i be the raw text extracted from article u_i.

The tokenization process splits the document into a sequence of tokens:

$$Tok(T_i) = \{t_{i1}, t_{i2}, \ldots, t_{in_i}\} \quad (4)$$

where

- t_{ij} represents the j-th token.
- n_i is the total number of tokens in document T_i.

Let S be the predefined set of stopwords.

Stopword filtering is defined as:

$$FS(T_i) = \{t \in Tok(T_i) | t \notin SW\} \quad (5)$$

Let

$L(.)$ be the lemmatization function that maps a token to its base form

The lemmatized token sequence is:

$$L(T_i) = \{L(t) | t \epsilon FS(T_i)\} \quad (6)$$

Let

$\alpha(t)$ be a predicate function such that:

$$\alpha(t) = \begin{cases} 1, & if\ t\ contains\ only\ alphabetic\ charcaters \\ 0, & otherwise \end{cases}$$

The final preprocessed token set is:

$$P(T_i) = \{t \in LT_i | \alpha(t) = 1\} \quad (7)$$

The entire preprocessing pipeline can be represented as a composite function:

$$P(T_i) = \alpha\left(L\left(FS\left(Tok(T_i)\right)\right)\right) \quad (8)$$

Thus, for all scraped documents:

$$PT_i = P(T_i), \quad \forall i \in \{1,2, \ldots, n\} \quad (9)$$

After preprocessing the processed textual data is turned into dense vectors with the help of a pre-trained sentence embedding model (de Souza et al., 2022). Paraphrase-MiniLM-L6-v2 of the sentence transformer architecture has been used to encode each news article into a set semantic embedding of fixed length. Contextual and semantic relations between sentences are captured by this model, and allows meaningful comparisons of news items based on the underlying content of the news item rather than superficially based on lexical similarity (Liu et al., 2024; Mhatre & Masurkar, 2021).

Let the preprocessed text document be represented as a sequence of tokens:

$$PT_i = \{t_1, t_2, \ldots, t_n\} \quad (10)$$

The semantic embedding function $f_0(.)$, implemented using a pre-trained Sentence Transformer model (paraphrase-MiniLM-L6-v2), maps the document into a fixed-dimensional vector space (Liu et al., 2024):

$$e_i = E(PT_i; \theta) \in \mathbb{R}^d \quad (11)$$

where

- θ denotes the learned parameters of the pre-trained transformer.
- d is the embedding dimension (typically $d = 384$).
- e captures the contextual and semantic meaning of the entire document.

For a reference article D_r and a comparison article D_c, the embeddings are given by:

$$e_R = E(PT_R; \theta),\ e_C = E(PT_C; \theta) \quad (12)$$

The Similarity Computation Module calculates the semantic similarity between trusted reference article, as well as each comparison article. Once the semantic embeddings are created using the Sentence-BERT model, the contextual similarity between news articles reporting the same event is compared using the embedding vectors. The similarity scores produced along with the information about source credibility are fed into the Rule-Based Classification Module for final classification.

The similarity of two articles can be quantized using a metric called semantic similarity, which measures how much two articles have in common despite any differing in writing style, vocabulary, or sentence structure. The articles are expected to show greater semantic similarity between themselves and lesser semantic agreement with the reference article for articles describing the same event, and lesser semantic agreement with the reference article for articles with misleading, incomplete or fabricated information (Gorai & Shaw, 2025). The semantic similarity between the reference article embedding and comparison article embedding is calculated as follows:

$$s = S(e_R, e_C) = \frac{e_R . e_C}{\|e_R\| \times \|e_C\|} \quad (13)$$

where:

- e_R = Semantic embedding vector of the reference article
- e_C = Semantic embedding vector of the comparison article
- $e_R . e_C$ = Dot product of the embedding vectors
- $\|e_R\|$ = Magnitude of the reference embedding vector
- $\|e_C\|$ = Magnitude of the reference embedding vector

The similarity score ranges between:

$$0 \leq s \leq 1 \quad (14)$$

A value closer to 1 indicates stronger semantic agreement between the articles, while a value closer to 0 indicates greater semantic divergence.

After computing the similarity level, a clear rule-based classification mechanism is used to classify the credibility of each news article. The proposed framework utilizes explicit decision rules which are created based on semantic similarity scores, rather than relying on labeled training data as do supervised machine learning techniques. These rules offer clarity and interpretation on each classification decision (Amira et al., 2024).

Let s denote the semantic similarity score between a candidate article A and a trusted reference article R, formally defined over the interval:

$$s \in [0,1] \quad (15)$$

A classification function $C(s)$ is defined as follows:

$$C(s) = \begin{cases} Real\ News, & s \geq \tau \\ Fake\ News, & s < \tau \end{cases} \quad (16)$$

where τ represents the predefined semantic similarity threshold, empirically set to:

$$\tau = 0.60 \quad (60\%)$$

Rule R1 — Real News Classification:

$$R_1: s \geq \tau \quad (17)$$

This rule means that the comparison article is sufficiently similar in context with the trusted reference article to be interpreted as semantically consistent with the event reported in the latter (Mhatre & Masurkar, 2021).

Rule R2 — Fake News Classification:

$$R_2: s < \tau \quad (18)$$

This rule finds articles that do not semantically agree with the trusted article. Such articles are deemed to be potentially misleading, inconsistent, or manufactured with regard to the verified news event.

The threshold value was chosen to partition between semantically similar and semantically dissimilar articles. The rules generated by the algorithm form the core decision making process in the proposed framework and can be used to classify the news articles without training the model or having much labeled data (Mhatre & Masurkar, 2021).

These rules are then tested for effectiveness through association rule interestingness measures such as Support, Confidence and Lift to assess the reliability and statistical significance of these rules.

The binary decision boundary partitions the semantic similarity space into two disjoint regions:

$$\Omega = \Omega_{Real} \cup \Omega_{Fake} \quad (19)$$

such that:

$$\Omega_{Real} = \{A \in \Omega \mid s(A) \geq \tau\}, \quad \Omega_{Fake} = \{A \in \Omega \mid s(A) < \tau\} \quad (20)$$

where Ω denotes the complete set of candidate articles under evaluation.

The credibility of each news source is also checked with a pre-defined trusted domain list, in addition to semantic similarity analysis. The source verification is an auxiliary validation method that aids the interpretation of the semantic similarity and aids the cross-source verification (Nikam & Dalvi, 2020).

Let D denote the finite set of trusted news domains, formally defined as:

$$D = \{d_1, d_2, d_3, d_4, d_5, \ldots\ldots d_n\} \quad (21)$$

For each candidate article A, let $dom(A)$ represent the domain attribute of its publishing source. A source credibility function $SC(A)$ is then defined as:

$$SC(A) = \begin{cases} 1, & dom(A) \in D \\ 0, & dom(A) \notin D \end{cases} \quad (22)$$

Rule R3 — Source Credibility Verification:

$$R_3: dom(A) \in D \quad (23)$$

This rule is followed to determine if a comparison article comes from a trusted and reputable news source. Similarity score and classification result are logged alongside credibility information to aid in interpretation of results. The trusted-source verification introduces transparency by enabling the analysis of the semantic similarity results with respect to the trustworthiness of the sources (Gorai & Shaw, 2025).

To assess the reliability and significance of the generated classification rules, three rule interestingness measures are computed: Support, Confidence, and Lift. These measures offer quantitative evidence of the semantics relationships between the reference article and comparison articles.

Three commonly used association rule interestingness measures — Support, Confidence, and Lift — are employed to rigorously assess the statistical reliability and discriminative significance of the classification rules R1 and R2. These measures collectively measure how well the semantic similarity scores support the classification decisions made by the generated classifiers, thus further strengthening the explainability and statistical validity of the proposed framework (Zhang et al., 2024).

This threshold is chosen so as to have a well-separated separation boundary between semantically coherent and semantically divergent articles with a good balance between false positive and false negative classification errors.

Support is the degree of similarity between two articles semantically. Support is defined as:

$$Supp(A) = 100 \times s(A) \quad (24)$$

where $s(A) \in [0,100]$ is a percentage-scaled semantic similarity score. Alternatively, since similarity is defined on the unit interval, this is equivalent to:

$$Sup(A) = 100 \times \frac{e_R \cdot e_C}{\|e_R\| \|e_C\|} \quad (25)$$

The higher the support value the greater the agreement in context between the articles.

The degree of confidence is the confidence of the obtained classification rule. For rule R_i, Confidence is defined as:

$$Conf(R_i) = \frac{Sup(R_i)}{100} \quad (26)$$

A higher confidence value indicates that the semantic similarity score is highly indicative of the assigned classification.

Lift is a measure of whether the observed semantic agreement is greater than normal. Lift for rule $R_i$is formally expressed as:

$$Lift(R_i) = \frac{Conf(R_i)}{Sup_{base}} \quad (27)$$

where Sup_{base} represents the expected Support under the baseline condition, set equal to the classification threshold:

$$Sup_{base} = \tau \quad (28)$$

Substituting, the Lift expression expands to:

$$Lift(R_i) = \frac{Conf(R_i)}{\tau} \quad (29)$$

The interpretation of Lift values follows a well-established convention:

$$Lift(R_i)\begin{cases} > 1 & Positive\ association \\ = 1 & Independent \\ < 1 & Negative\ association \end{cases} \quad (30)$$

Lift values >1 means that the compared articles have a higher than expected association to each other.

Rule R4 — Rule Interestingness Criterion:

$$R_4: Lift(R_i) > 1 \quad (31)$$

Formally, a rule R_i is deemed statistically interesting if and only if:

$$Sup(R_i) \geq \tau \ \wedge \ Conf(R_i) \ \wedge \ Lift(R_i) > 1 \quad (32)$$

Rules that meet this criterion are deemed to be statistically significant and to give more compelling evidence of semantic consistency between news sources.

By integrating Support, Confidence and Lift, the proposed framework becomes more explainable by quantifying the reliability and significance of the classification rules generated.

The complete interestingness profile for each classification rule $R_i$is thus represented as a triple:

$$IP(R_i) = (Sup(R_i), Conf(R_i), Lift(R_i)) \quad (33)$$

where IP(R_i) is the composite interestingness profile.

The inclusion of these measures of interestingness in the proposed framework has two purposes. First, they allow a principled, quantitative evaluation of the reliability and discriminative capability of each classification rule. Second, they increase the explainability of the framework by providing a metric-based, and hence transparent, justification of each classification decision, which is important in the context of automated fake news detection systems for deployment in real-world high-stakes information environments. This decision mechanism that is guided by rules is interpretable and does not rely on labeled training data (Nikam & Dalvi, 2020).

The Result Export Module is the final step in the proposed fake news detection pipeline. All intermediate computational outcomes, as well as the final classification decision results, from the previous modules are then systematically integrated into a structured, machine-readable document for each candidate comparison article, after performing the rule-based classification process. The responsibility of this module is to integrate, shape and store the entire output set in a way that facilitates further statistical analyses, graphical representation and experimental repeatability. The modular architecture will make it transparent, interpretable and easy to extend to hybrid approaches in future.

For each comparison article, the output record is represented as:

$$O_i = \{U_i, s_i, C_i, SC_i, Sup_i, Conf_i, Lift_i, \hat{y}_i\} \quad (34)$$

where

- U_i - URL ID.
- s_i – Semantic Similarity Score
- C_i – Classification Result
- SC_i – Source Credibility
- Sup_i – Support
- $Conf_i$ – Confidence
- $Lift_i$ – Lift
- $\hat{y}_i$ – Predicted Label

All records are aggregated into a result set:

$$\mathcal{R} = \{\mathcal{R}_1, \mathcal{R}_2, \ldots., \mathcal{R}_N\} \quad (35)$$

The final result matrix is stored in CSV format as:

$$M = \begin{bmatrix} U_1 & s_1 & C_1 & SC_1 & Sup_1 & Conf_1 & Lift_1 & \hat{y}_1 \\ U_2 & s_2 & C_2 & SC_2 & Sup_2 & Conf_2 & Lift_2 & \hat{y}_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ U_n & sim_n & C_n & SC_n & Sup_n & Conf_n & Lift_n & \hat{y}_n \end{bmatrix} \quad (36)$$

This structured output supports statistical analysis, visualization, and reproducibility of experimental results.

ALGORITHM 1: RULE-BASED SEMANTIC SIMILARITY

Step 1: $t_{ref} \leftarrow ScrapeContent(s_{ref})$

Step 2: $t_{ref} \leftarrow PreprocessText(t_{ref})$

Step 3: $e_{ref} \leftarrow EncodeEmbedding(t_{ref})$

Step 4: $for\ each\ URL\ s_i \in \delta_{cmp}\ do$

Step 5: $t_i \leftarrow ScrapeContent(s_i)$

Step 6: $t_i \leftarrow PreprocessText(t_i)$

Step 7: $e_i \leftarrow EncodeEmbedding(t_i)$

Step 8: $Sim_i \leftarrow Similarity(e_{ref}, e_i)$

Step 9: $domain_i \leftarrow ExtractDomain(s_i)$

Step 10: $if\ (Sim_i \geq \tau)\ AND\ (domain_i \in \mathcal{D}_{trust})\ then$

Step 11: $y_i \leftarrow Real$

Step 12: $else$

Step 13: $y_i \leftarrow Fake$

Step 14: $end\ if$

Step 15: $Store(s_i, Sim_i, domain_i, y_i)$

Step 16: $end\ for$

Step 17: $ExportResultsToCSV()$

End Algorithm

RESULTS

The proposed framework for detecting fake news was tested with 51 news articles from various news sources across the country and the world on the Air India Flight 171 incident. To each comparison article we first scraped the entire text, preprocessed the text, next converted it to semantic embeddings with the Sentence-BERT model (paraphrase-MiniLM-L6-v2) and calculated the cosine similarity between the comparison article and the reference article. A predefined similarity threshold of 60% was adopted for rule generation and classification, where articles were classified as Real News if the similarity score is at or above 60% and as Fake News if the similarity score is below 60%. Furthermore, the support, confidence and the lift measures were calculated to evaluate the strength and reliability of the classification rules produced.

Text preprocessing is an important aspect to enhance the desirable semantic similarity measurement. Lowercasing, stop-word removal, punctuation elimination and lemmatization were performed in the preprocessing stage to minimize the lexical noise and normalize the textual content prior to embedding generation.

The impact of preprocessing on the semantic similarity pipeline is summarized in Figure 2. Noisy tokens, punctuation and stylistic differences create inconsistent semantic representations without preprocessing, which leads to lower similarity scores. Once preprocessed, the resulting embeddings are better aligned semantically and the system is able to capture the meaning of the articles more effectively. This results in better similarity computation and better rule-based classification results.

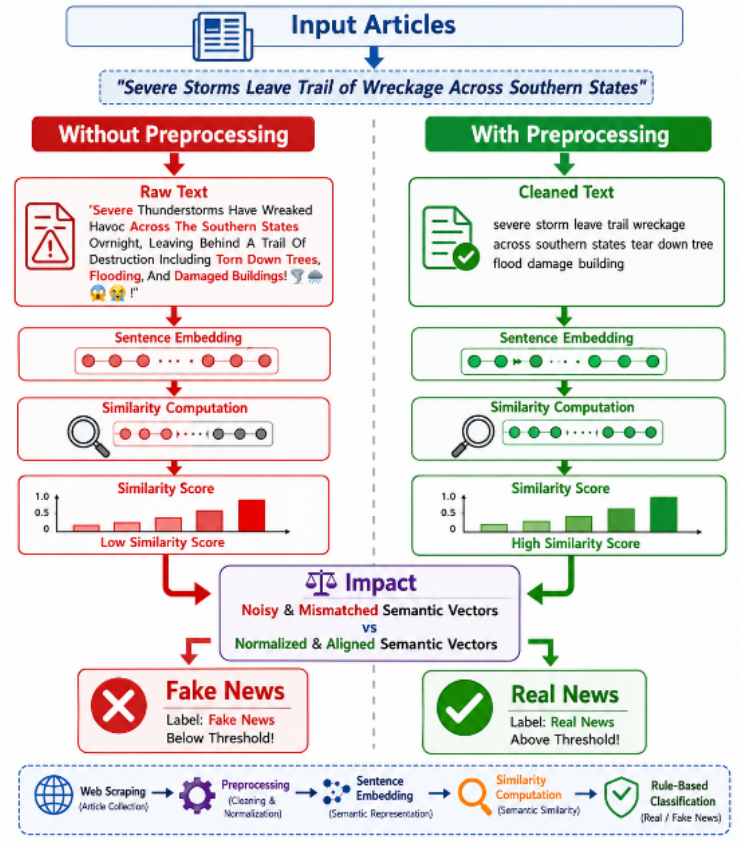


Figure 2 Impact of Text Preprocessing on Semantic Similarity-Based Fake News Detection

The effectiveness of preprocessing is further illustrated in Figure 3. In the original set of articles, there were 12 articles classified as Real and 38 articles classified as Fake. There were 12 articles that were identified as Real and 38 articles that were identified as Fake before pre-processing. The number of correctly classified Real articles rose significantly from 10 to 42 after preprocessing, and the number of correctly classified Fake articles fell to 8 after preprocessing. The result shows that there is

an improvement in semantic matching of related news reports and misclassification is reduced by preprocessing because of the linguistic differences.

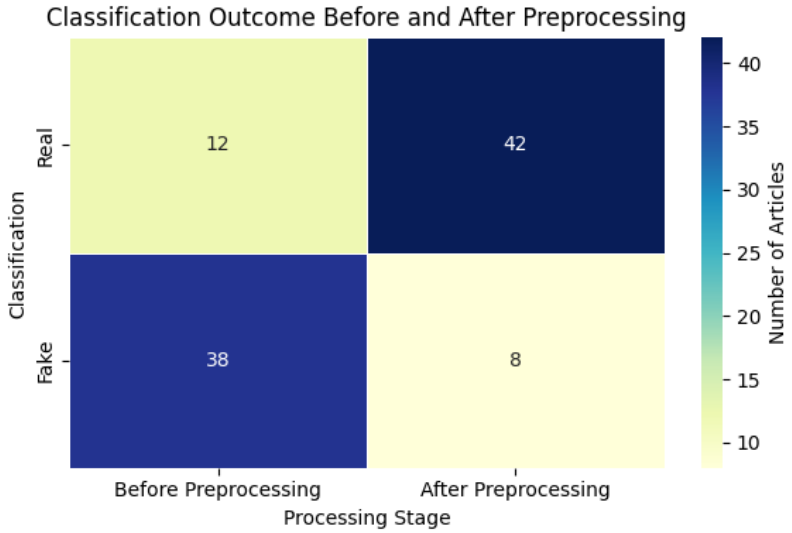


Figure 3 Classification Outcomes Before and After Preprocessing

Table II shows the summary of the 51 news articles collected and the semantic similarity scores obtained. The result yielded an average similarity score of 36.69 % and a standard deviation of 32.58 %. Among the different articles, the highest similarity score was obtained for the AP News article (82.48%) and the lowest was obtained for the Reuters article (−2.82%). According to the predefined similarity threshold, out of the 51 articles evaluated, 20 were considered Real News and 31 were considered Fake News.

Table II: Rule-Based Classification Results using Semantic Similarity

Metric	**Value**
`Total Articles Evaluated	51
Mean Similarity (%)	36.69
Std. Deviation	32.58
Min Similarity (%)	−2.82 (Reuters)
Max Similarity (%)	82.48 (AP News)
Articles Classified Real	20
Articles Classified Fake	31

Figure 4 shows the distribution of semantic similarity scores among all the news sources evaluated. The heat map shows that there is significant difference between articles reporting the same event. Some articles share similar content to the articles' scores are above the 60% threshold and thus are deemed Real News, and others have scores below the threshold but contain similar information. The study serves to illustrate the differences in journalistic writing styles, the structure of the narrative and the different perspectives on the news of various news organizations.

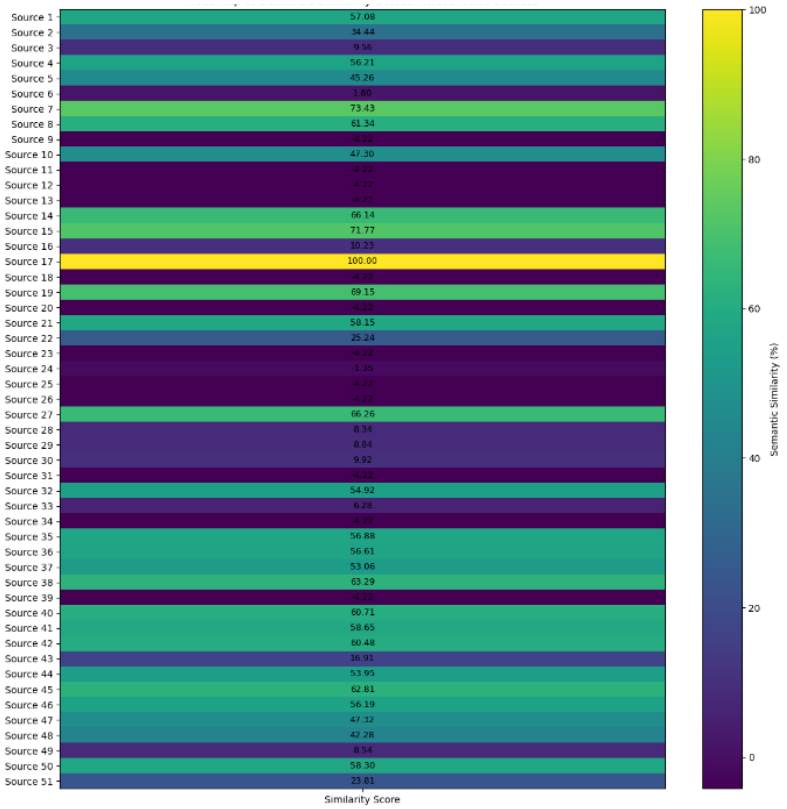


Figure 4 Semantic Similarity Scores across News Sources

Although the established news outlets like The Hindu, BBC, NDTV, and The New Indian Express are credible, the articles in these newspapers had lower similarity scores than the pre-determined threshold to be classified as real news. As a result, the system identified all reviewed papers as fake. This finding indicates that editorial framing, narrative focus and context focus differ considerably among news sources even when they report the same news. This diversity has negative impact on the decision rules of similarity leading to excessive conservativeness in classifications.

The results indicate that the similarity of the meanings does not always correspond to the credibility of the source. Articles written by reputable organizations might get less similarity scores because they report things differently, put emphasis on different aspects or report context details differently. As a result, the measure of semantic similarity should be used only as a sign of content consistency and not as a standard for determining the authenticity of news.

The implemented framework automatically derives classification rules from semantic similarity values. There are two primary decision rules for classification:

R1: IF Similarity ≥ 60% THEN News = Real

R2: IF Similarity < 60% THEN News = Fake

To assess the quality of generated rules, three interestingness measures have been calculated:

Support = Similarity/100

Confidence = Similarity/100

Lift = Support/Expected Support (Expected Support = 0.60)

Confidence values for several similarity rules are given in Figure 5. The range of confidence values is from 0.76 to 0.91, showing the strong correlation between semantic similarity and classification decision. The maximum confidence value for R3 rule is 0.91, which indicates that the classification generated by the rule is reliable.

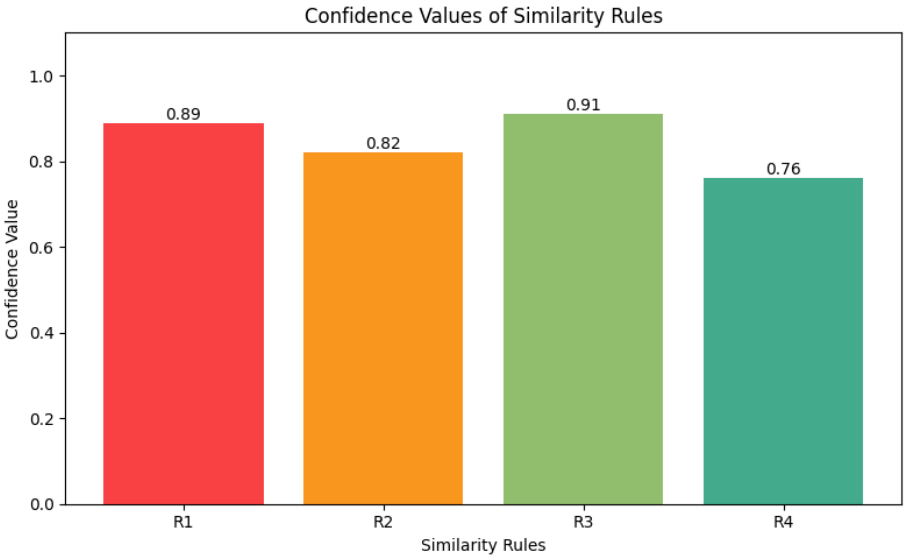


Figure 5 Confidence Values of Generated Similarity Rules

The lift of the generated rules is represented in Figure 6. The rules have lift values higher than 1 in the range of 1.09 to 1.65. If the value of LIFT is greater than 1, there is a positive correlation between similarity condition and class label. For analyzed rules, the highest lift value is 1.65 as given by R3.

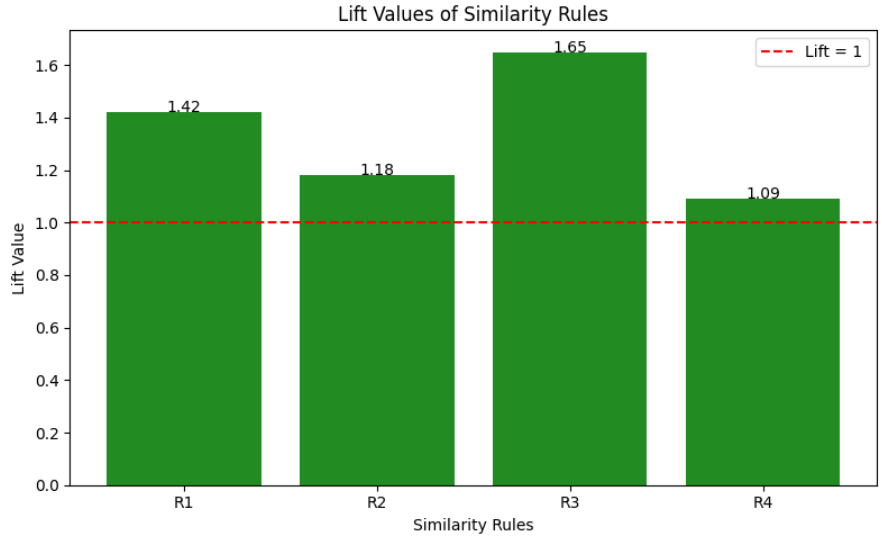


Figure 6 Lift Values of Generated Similarity Rules

The experimental results show that preprocessing significantly enhances the effectiveness of the semantic similarity measures and classification performance. The results of the preprocessing show that the number of Real News articles goes from 12 to 42 after the normalization process. Moreover, the rule-based approach can yield meaningful classification decisions in terms of interestingness measures and explicit similarity thresholds.

The results, however, indicate the limitations of using semantic similarity alone. Two articles about the same event can have very different terms and different emphases and different structures, scoring low similarities even if they are correct in their facts. Thus, further research needs to incorporate other credibility cues like the presence of named entities, source trustworthiness, temporal consistency, and fact-checking. This type of hybrid approach will enhance detection robustness and retain the interpretability benefits of rule-based systems.

CONCLUSION

In this study, an explainable rule-based fake news detection framework was proposed, which combines automated web scraping, text preprocessing, semantic embedding, and rule reasoning based on cross-source news verification. The proposed approach is based on the use of Sentence-BERT semantic embeddings and cosine similarity analysis, which are used to assess the consistency of the content of news from various sources without using labeled training sets. The framework's combination of semantic similarity calculations and comprehensible classification rules offers a clear and efficient method for recognizing potentially unreliable news articles.

The experimental evaluation was carried out with 51 news articles gathered from a variety of national and international news reports regarding the Air India Flight 171 incident. The results showed that text preprocessing has a significant effect on improving the similarity between texts and classification results by eliminating text noise and increasing the similarity between the meaning of the texts of related articles. The generated rules were able to classify the articles with a predefined similarity threshold and were able to give an interpretable decision-making with support, confidence, and lift measures. The results of the confidence and lift analyses further verified the effectiveness and reliability of the rules produced, thus demonstrating the potential of explainable fake news detection using rule-based reasoning.

In addition, the study found significant differences in the semantic similarity scores of articles covering the same event. The news sources were not semantically aligned in all aspects, due to editorial differences, narrative focus, context coverage, and reporting structure. The results suggest that semantic similarity can be a useful measure of content consistency, but it should not be used as a single criterion to assess news authenticity. Although the proposed framework is transparent, it is easy to interpret, and easy to implement, adopting only similarity-based rules might not be sufficient to perform classification in complex real-world news environments.

In the future, additional credibility-sensitive features, including named-entity overlap, factual consistency check, sentiment analysis, temporal alignment, and source credibility assessment, will be integrated with the proposed framework to develop a hybrid fake news detection system. In addition, machine learning classifiers can be integrated with the rule-based reasoning for better robustness, scalability, and generalization in various news domains. Such a hybrid approach has the potential to achieve higher detection accuracy while preserving the interpretability and explainability that are essential for trustworthy fake news detection systems.

ACKNOWLEDGEMENT

The authors sincerely thank the Department of Computer Science, SDNB Vaishnav College for Women, Chennai, Tamil Nadu, India, for providing the academic environment, institutional support, and facilities necessary to carry out this research. The authors also acknowledge the computational resources and technical support provided by the institution, which contributed to the development, implementation, and evaluation of the proposed fake news detection framework. The authors express their gratitude to all those who provided valuable support and encouragement throughout the course of this research.

Source of support

None

Conflicts of Interest

The authors declare that they have no conflict of interest regarding the publication of this research article. There are no financial, personal, or professional relationships that could have influenced the research, analysis, results, or conclusions presented in this study.

REFERENCES

Amira, A., Derhab, A., Hadjar, S., Merazka, M., Alam, M. G. R., & Hassan, M. M. (2024). Detection and analysis of fake news users' communities in social media. IEEE Transactions on Computational Social Systems, 11(4), 5050–5059.

Bonny, A. J., Bhowmik, P., Mahmud, M. S., & Sattar, A. (2022). Detecting fake news in benchmark English news dataset using machine learning classifiers. Proceedings of the 13th International Conference on Computing Communication and Networking Technologies (ICCCNT), Kharagpur, India, 1–8. https://doi.org/10.1109/ICCCNT54827.2022.9984461.

de Souza, M. C., Nogueira, B. M., Rossi, R. G., et al. (2022). A network-based positive and unlabeled learning approach for fake news detection. Machine Learning, 111, 3549–3592. https://doi.org/10.1007/s10994-021-06111-6.

El Madbouly, M., Ahmed, Y. A., & Salem, M. A. (2023). Multimodality web page analysis for fake news detection. Proceedings of the 2nd International Conference on Smart Cities 4.0.

Gorai, J., & Shaw, D. K. (2025). Semantic similarity and mutual information-based model for fake news detection. Multimedia Tools and Applications, 84, 23401–23421. https://doi.org/10.1007/s11042-024-19845-x.

Jain, P., Sharma, S., Monica, & Aggarwal, P. K. (2022). Classifying fake news detection using SVM, Naive Bayes and LSTM. 2022 12th International Conference on Cloud Computing, Data Science & Engineering (Confluence), Noida, India, 460–464. https://doi.org/10.1109/Confluence52989.2022.9734129.

Kaur, P. (2022). Sentiment analysis using web scraping for live news data with machine learning algorithms. Materials Today: Proceedings, 65(8), 3333–3341. https://doi.org/10.1016/j.matpr.2022.05.409.

Khanam, Z., Alwasel, B. N., Sirafi, H., & Rashid, M. (2021). Fake news detection using machine learning approaches. IOP Conference Series: Materials Science and Engineering, 1099, 012040. https://doi.org/10.1088/1757-899X/1099/1/012040.

Kumar, S., Perwej, Y., Siddiqui, F., Shukla, A., & Akhtar, N. (2025). A data-driven framework for fake news detection via web scraping and machine learning approach. 1391–1404. https://doi.org/10.38124/ijisrt/25jun1003.

Kumar, S., & Singh, T. D. (2022). Fake news detection on Hindi news dataset. Global Transitions Proceedings, 3(1), 289–297. https://doi.org/10.1016/j.gltp.2022.03.014.

Liu, X., Li, P., Huang, H., Li, Z., Cui, X., Liang, J., Qin, L., Deng, W., & He, Z. (2024). FKA-Owl: Advancing multimodal fake news detection through knowledge-augmented LVLMs. Proceedings of the 32nd ACM International Conference on Multimedia (MM '24), 10154–10163. https://doi.org/10.1145/3664647.3681089.

Meesad, P. (2021). Thai fake news detection based on information retrieval, natural language processing and machine learning. SN Computer Science, 2(6), 425. https://doi.org/10.1007/S42979-021-00775-6.

Mhatre, S., & Masurkar, A. (2021). A hybrid method for fake news detection using cosine similarity scores. 2021 International Conference on Communication, Information and Computing Technology (ICCICT), Mumbai, India, 1–6. https://doi.org/10.1109/ICCICT50803.2021.9510134.

Nikam, S. S., & Dalvi, R. (2020). Machine learning algorithm based model for classification of fake news on Twitter. 2020 International Conference on Intelligent Computing and Control Systems (ICICCS). https://doi.org/10.1109/ISMAC49090.2020.9243385.

Pande, M., Rathod, S. D., Joshi, R., Chvan, G. T., Jadhav, D., Phutane, P., Gonge, S., & Kadam, K. (2022). Fake news identification using regression analysis and web scraping. International Journal of Safety and Security Engineering, 12(3), 311–318. https://doi.org/10.18280/ijsse.120305.

Prachi, N. N., Habibullah, M., Rafi, M. E. H., Alam, E., & Khan, R. (2022). Detection of fake news using machine learning and natural language processing algorithms. Journal of Advances in Information Technology, 13(6), 652–661. https://doi.org/10.12720/jait.13.6.652-661.

Samadi, M., & Momtazi, S. (2025). Fake news detection: Deep semantic representation with enhanced feature engineering. International Journal of Data Science and Analytics, 20, 325–336. https://doi.org/10.1007/s41060-023-00387-8.

Shah, A., Shah, H., Bafna, V., Khandor, C., & Nair, S. (2024). VERITAS-NLI: Validation and extraction of reliable information through automated scraping and natural language inference. arXiv.

Wibowo, F. W., Dahlan, A., & Wihayati. (2021). Detection of fake news and hoaxes on information from web scraping using classifier methods. Proceedings of the 4th International Seminar on Research of Information Technology and Intelligent Systems (ISRITI), 1–6.

Zhang, Q., Liu, J., Zhang, F., Xie, J., & Zha, Z.-J. (2023). Hierarchical semantic enhancement network for multimodal fake news detection. Proceedings of the 31st ACM International Conference on Multimedia (MM '23), 3424–3433. https://doi.org/10.1145/3581783.3612423.

Zhang, X., Dadkhah, S., Weismann, A. G., Kanaani, M. A., & Ghorbani, A. A. (2024). Multimodal fake news analysis based on image–text similarity. IEEE Transactions on Computational Social Systems, 11(1), 959–972. https://doi.org/10.1109/TCSS.2023.3244068.