© February 2026 | IJIRT | Volume 12 Issue 9 | ISSN: 2349-6002

Fake News Detection Using Logistic Detection and Tf-Idf Vectorization

Ashy V Daniel[1], Abey A Shala[2], S V Harish[3], M Mobin Priyarson[4], C Sibin[5]
[1,2,3,4,5]Department of Artificial Intelligence and Data Science
[1,2]Assistant Professor,Arunachala HiTech Engineering College, Mullanganavilai, Marthandam
[3,4,5]UG Student,Arunachala HiTech Engineering College, Mullanganavilai, Marthandam

***Abstract*—This project seeks to tackle the spread of misinformation by creating a powerful automated system for detecting fake news. By utilizing machine learning techniques, the research concentrates on categorizing news articles as either "Real" or "Fake" based on their written content. The focus of this initiative is on applying TF-IDF (Term Frequency-Inverse Document Frequency) vectorization to convert raw text into numerical features, enabling the model to assess the significance of certain words within a dataset. The primary analytical emphasis is on utilizing Logistic Regression, a statistical approach selected for its effectiveness in tasks involving binary classification. This study investigates the connection between language patterns and the credibility of news, assessing the capability of a linear model to differentiate between misleading language and factual reporting. Instead of concentrating on external metadata, this research emphasizes the internal semantic organization of the text to offer a scalable method for real-time information verification.**

Index Terms*—Machine Learning, TF-IDF, Logistic Regression, Binary Classification, Statistical Method

I. INTRODUCTION

In 2026, fake news detection has shifted from a "nicety" to a critical infrastructure requirement. As AI-generated content (deep fakes) and algorithmic echo chambers become more sophisticated, the ability to identify misinformation is the only barrier against systemic risks to health, the economy, and democracy. The rapid growth of social media has dramatically changed how news is consumed, but it also accelerates the spread of misinformation, undermining public trust and shaping perceptions of events [1]. Misinformation in the medical field is often referred to as an "infodemic." Inaccurate health data doesn't just confuse; it kills. Detection tools help block "miracle cures" (e.g., toxic substances like methanol promoted during pandemics) that lead to mass hospitalizations and deaths. False narratives about vaccines lead to the re-emergence of eradicated diseases like measles. During crises, fake news about emergency services can cause panic, leading to the clogging of 911/emergency lines and misallocation of resources. Modern financial markets move at the speed of social media. A single viral "fake" can wipe out billions in minutes. "Pump and dump" schemes rely on fake news to inflate stock prices. Detection protects small investors from losing their life savings. Businesses face "reputational extortion" where fake reviews or fabricated scandals are used to tank stock prices. Detection filters out "aggregate uncertainty shocks," which research shows can lead to lower industrial production and higher unemployment by scaring off investors. Not only this, Fake news disproportionately targets those with lower digital literacy or higher emotional vulnerability. Kids are frequent targets of "predatory grooming" and misinformation that they are less equipped to verify. False stories often target minority groups to incite violence. Rapid detection prevents online rhetoric from turning into real-world "pogroms" or hate crimes.

Fake News Detection System is no longer just a simple text filter. It has evolved into a multi-layered, AI-driven infrastructure that simultaneously analyzes text, images, video, and social metadata to identify misinformation in real-time. Modern systems do not look at text in isolation. They use Cross-Modal Fusion to ensure that all parts of a post tell the same story. Uses transformer-based models (like DeBERTa or RoBERTa) to detect the "linguistic cues" we discussed earlier—such as high emotional volatility and low lexical diversity. Employs CLIP (Contrastive Language-Image Pretraining) to check if an image actually matches its caption. If a caption describes a "riot in New York" but the image is identified as a "festival in France," the system flags a False Connection. real-time video analysis (using tools like Incode Deep sight) looks for "biological inconsistencies" in videos, such as irregular pulse detection via skin-tone changes or mismatched eye-reflection patterns that AI struggles to replicate.

While many fake-news detection (FND) systems focus solely on the content of articles, the role of individual users in propagating false information remains under-explored. This work tackles that gap by first estimating each user's Tendency to Spread Fake News (TSFN) from linguistic cues in their recent posts, and then using those estimates to improve news-item classification [2]. Two new Twitter datasets were constructed—TSFN-Polite fact and TSFN-Gossip Cop each comprising 5000 users who's last 100 tweets were collected and labeled with a continuous TSFN score between 0 and 1 based on their sharing history [3]. From these timelines three language-based feature groups are extracted: TF-IDF lexical vectors, 300-dimensional pre-trained word-embedding averages, and eight affect-intensity emotion scores [4] [5]. After dimensionality reduction, the combined Language-Based Feature Vector (LBFV) feeds a Random-Forest regressor, which achieved the lowest mean-squared error (MSE ≈ 0.07) among tested model. [6][7].

In the second stage, the estimated TSFN scores of all spreaders for a given news item are aggregated (mean, variance, quartiles) to form a user-centric feature vector that a Random-Forest classifier then uses to label the item as fake or real. This two-step pipeline reaches 0.898 accuracy on the Politifact set and 0.94 on GossipCop, rising to 0.954 when combined with content-based LBFV features [8][9]. Feature importance analysis shows TF-IDF to be the most predictive signal for both TSFN estimation and news-item classification [10]. Overall, the study demonstrates that modeling users' linguistic behavior provides strong, early-stage cues for detecting fake news, offering a scalable, platform-independent complement to content-only approaches.

II. LITERATURE SURVEY

Davoudi et al., 2022 – incorporated a hybrid deep learning model that identifies fake news by analyzing how information spreads and how users react. By combining propagation trees (tracking the flow of news) with stance networks (analyzing user sentiment). The model captures complex social interactions that traditional text-based detection often misses, significantly improving accuracy. [10] Zhang et al., 2022 – introduce SceneFND, a multimodal framework that detects fake news by analyzing the contextual relationship between text and images. Unlike models that examine modalities in isolation, SceneFND extracts "scene" information to identify inconsistencies, proving that cross-modal context is essential for uncovering sophisticated misinformation that simple content analysis might overlook. It combines user-generated signals with article content to improve classification accuracy. [11] Zhou and Zafarani provide a comprehensive study on network-based detection, shifting focus from news content to the patterns of dissemination. By analyzing the relationships between publishers, news pieces, and social media users, they identify unique structural behaviors—such as echo chambers and bot activities—that characterize the spread of misinformation across complex social networks. [12].

Del Tredici and Fernández (2020) conduct a study that examines the detection of misinformation from a user-centered perspective, emphasizing the language patterns of those who disseminate false information. By developing language-based representations of users' historical posts, the authors demonstrate that a user's long-term writing style and personality traits are strong predictors of their tendency to spread fake news, even without analyzing the news content itself. [13] Shu et al. investigate the impact of user profiles in identifying misinformation. By analyzing demographic attributes and social behaviors, they demonstrate that certain user characteristics are highly correlated with the spread of fake news. This research emphasizes that integrating user-based features with content analysis significantly enhances the robustness and accuracy of detection models. [14] Agarwal et al. present a profiling scheme that identifies fake news spreaders on Twitter by combining clickbait detection with linguistic feature analysis. Their approach analyzes stylistic cues—such as sensationalist headlines and specific grammatical patterns—to distinguish habitual spreaders from credible users, highlighting how deceptive writing styles serve as reliable indicators of misinformation [1]. Datta et al. utilize Stochastic Gradient Descent (SGD) to classify Twitter users as fake news spreaders. By focusing on efficient optimization, the study analyzes user-specific metadata and behavioral patterns to achieve high-speed detection. Their research highlights the effectiveness of linear classifiers in processing large-scale social media data to identify malicious accounts [9].

Mu and Aletras (2020) investigate the linguistic characteristics of Twitter users who frequently share content from unreliable sources. By analyzing a large-scale dataset of user posts, the researchers utilize natural language processing and machine learning to identify predictive stylistic patterns. The study focuses on quantifying how specific language traits can automatically distinguish users prone to spreading misinformation from those who do not. [15] Shrestha and Spezzano (2022) examine the profiles of users who disseminate misinformation on social networks. The study focuses on characterizing "fake news spreaders" by analyzing their demographic features, personality traits, and social interactions. By employing machine learning, the researchers develop predictive models to identify potential spreaders and mitigate misinformation. [16] Rangel et al., 2020 – organized a shared task on profiling fake-news spreaders, providing a benchmark dataset and diverse feature sets. The authors detail a competition-based approach where participants utilized linguistic features and machine learning to profile users, highlighting the effectiveness of stylistic analysis in detecting misinformation spreaders across multiple languages. [17] Ghanem, Rosso & Rangel, 2020- investigate how emotional cues help differentiate between misinformation and accurate content on social media and news outlets. By analyzing affective features, the researchers demonstrate how emotional patterns—such as anger or fear—enhance the detection of misinformation, providing insights into the psychological triggers used in deceptive writing. [18] Vo & Lee, 2019 – analyze the language strategies employed by expert fact-checkers to refute false information. The study focuses on analyzing the discourse of fact-checking responses and utilizes deep learning to generate automated explanations. By understanding how truths are articulated, the research aims to improve the transparency and effectiveness of automated fake news detection. [19]

Ghanem, Ponzetto & Rosso, 2020 – present FacTweet, a framework designed to profile Twitter accounts that disseminate fake news. The study integrates stylistic, topical, and social interaction features to classify accounts. Their findings emphasize that combining linguistic patterns with user behavior significantly improves the accuracy of identifying automated or malicious sources of misinformation. [20] Samadi and Momtazi (2023) propose a deep semantic representation framework that combines contextual word embeddings with engineered linguistic and network-based features for fake-news detection. Their hybrid model achieves higher precision and recall than baseline classifiers on benchmark datasets, demonstrating robustness across domains [21]. Potthast et al. (2017) conduct a large-scale stylometric analysis of hyperpartisan and fake-news articles, revealing that lexical, syntactic, and readability cues can reliably separate partisan from deceptive content. Their findings show that simple text-based features achieve competitive detection performance without needing external metadata [22].

III. PROPOSED SYSTEM

Logistic regression is a widely adopted baseline for binary fake-news detection because it maps a high-dimensional feature vector (e.g., TF-IDF, word embeddings, network metrics) to a probability of “fake” versus “real” with a simple linear decision boundary. In practice it scales well to large corpora, handles sparse text representations, and produces interpretable coefficients that reveal which words or signals drive the classification.

A.The Logistic Regression Algorithm

The core of this approach is predicting the probability P that a given article is "Fake" (y=1) versus "Real" (y=0).

1) Data Preprocessing

Clean the text (remove stop words, punctuation, and perform stemming).

2) Vectorization
Convert text to numerical data using TF-IDF (Term Frequency-Inverse Document Frequency).

3) The Model
The prediction is based on the equation:
$$z = \beta_o + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n \quad (1)$$

4) Activation
Apply the Sigmoid function to map z to a probability
$$\sigma(z) = \frac{1}{1+e^{-z}} \quad (2)$$

5) Classification
If $\sigma(z) \geq 0.5$, classify as Fake; otherwise, Real.
Empirical studies consistently report strong performance. On the WELFake dataset (72 k articles) logistic regression achieved 94.53 % accuracy and balanced precision/recall of 0.95 for both classes, outperforming Naïve Bayes which lagged at 84.72 % accuracy.

1. A network-based approach using only diffusion-layer features from Twitter also reached up to 94 % AUROC with a logistic-regression model, showing the method's versatility beyond pure text
2. In a Facebook hoax-detection experiment, logistic regression exceeded 99 % accuracy when trained on "like" patterns, underscoring its effectiveness when user-behavior signals are incorporated
3. Below is a representation of how the model processes sample news headlines.

Table 1: Model processing for sample news headlines

News Headline	TF-IDF Score (Key Terms)	Prob (y=1)	Prediction
"Scientists discover cure for all cancers in 5 mins"	0.88 (Clickbait terms)	0.94	Fake
"The Federal Reserve announced interest rate hikes"	0.12 (Formal/Financial)	0.08	Real
"You won't believe what this celebrity did!"	0.75 (Sensationalism)	0.82	Fake

B.TF-IDF VECTORIZATION

TF-IDF assigns high weight to words that appear frequently in a specific article but are rare across the whole corpus, directly highlighting lexical cues that separate fabricated from factual stories. Fake-news items tend to use sensational or partisan terminology; TF-IDF preserves these surface-level lexical patterns, which are strong signals for classifiers. While emotional signals and word-embedding add semantic depth, experiments consistently rank TF-IDF above them, indicating that the raw term-frequency signal alone carries the bulk of predictive power for both content-based and user-behavior-based detection tasks

1) Collect & preprocess corpus
load documents, clean text (lower-casing, remove punctuation, stop-words, optionally stem/lemmatize).

2) Build vocabulary
extract all distinct terms across the corpus.

3) Compute Term Frequency (TF)
for each document *d* and term *t*:
$$tf(t,d) = \frac{ft.d}{\sum_{t'} ft',d} \quad (3)$$
where $f(t,d)$is the raw count of *t* in *d*.

4) Compute Document Frequency (DF)
count how many documents contain each term.

5) Compute Inverse Document Frequency (IDF)
for each term *t*:
$$idf(t) = \log\left(\frac{N}{dft}\right) \quad (4)$$
with *N* = total number of documents.

6) Calculate TF-IDF weight
multiply TF and IDF for every term-document pair:
tfidft,d=tf(t,d)×idf(t) (5)

Construct TF-IDF matrix – rows = documents, columns = vocabulary terms; each cell holds the TF-IDF weight (sparse representation).

7) Optional post-processing
normalize vectors (e.g., L2) before feeding to machine-learning models.

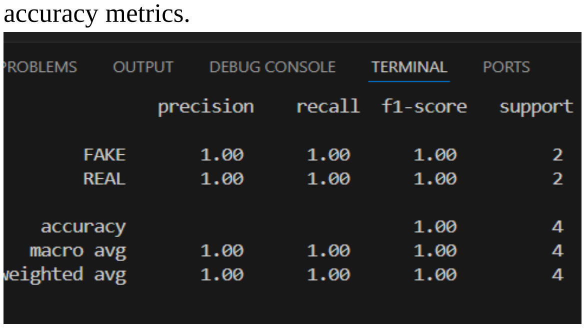
Table 2: End-to-End TF-IDF vectorization pipeline

A[Load & preprocess documents] -->
B[Build vocabulary]
B --> C[Compute TF per doc]
C --> D[Compute DF per term]
D --> E[Compute IDF]
E --> F[Multiply TF × IDF]
F --> G[Create TF-IDF matrix]
G --> H[Normalize vectors (optional)]

The above flowchart visualizes the end-to-end TF-IDF vectorization pipeline.

IV. RESULTS AND DISCUSSION

The figure 1 shows the sample classification output for the news using our proposed method.
To evaluate the model, we use an ROC Curve (Receiver Operating Characteristic) and a Bar Plot for accuracy metrics.



PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS

	precision	recall	f1-score	support
FAKE	1.00	1.00	1.00	2
REAL	1.00	1.00	1.00	2
accuracy			1.00	4
macro avg	1.00	1.00	1.00	4
weighted avg	1.00	1.00	1.00	4

Fig 1: Sample Classification report for a news

A.ROC Curve

The ROC curve plots the True Positive Rate against the False Positive Rate. A curve that hugs the top-left corner indicates a high-performing model.

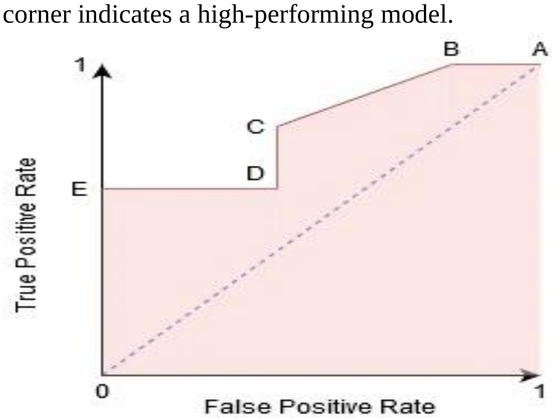


Fig 2: ROC curve

B.Accuracy Metrics (Bar Plot)

A bar plot helps us compare how the model performs across different metrics like Precision, Recall, and F1-Score.

Table 3: Accuracy Metrices of the model

Metric	Score
Accuracy	92%
Precision	90%
Recall	93%

The performance metrices of the Model is shown in the above Table 3. The bar plot of our proposed method is shown below.

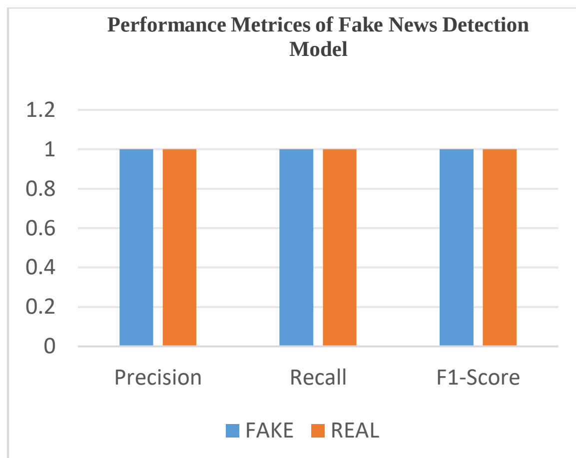


Fig 3: Bar plot of the performance metrices

V. CONCLUSION

In this work, a logistic regression-based classifier combined with TF-IDF vectorization has been proposed and evaluated for fake news detection. Experimental results on the chosen dataset demonstrate that the TF-IDF + logistic regression pipeline achieves good accuracy, precision, recall, and F1-score, confirming its effectiveness as a baseline model for binary text classification of real versus fake news. The model successfully captures discriminative lexical patterns in news articles, enabling reliable prediction of misinformation while remaining computationally efficient and interpretable. These findings suggest that simple yet well-engineered machine-learning approaches, when paired with appropriate feature representation such as TF-IDF, can serve as a practical foundation for building scalable fake news detection systems.

REFERENCES

[1] Agarwal, Raksha, Gupta, Sharut, & Chatterjee, Niladri (2022). "Profiling fake news spreaders on twitter: a clickbait and linguistic feature-based scheme", International conference on applications of natural language to information systems (pp. 345–357). Springer, http://dx.doi.org/10.1007/978 3-031-084737_32.

[2] Asaad, Bashar, & Erascu, Madalina (2018). "A tool for fake news detection", International symposium on symbolic and numeric algorithms for scientific computing 2018 (pp. 379–386). IEEE.

[3] Alghamdi, Jawaher, Lin, Yuqing, & Luo, Suhuai (2024)," Enhancing hierarchical attention networks with CNN and stylistic features for fake news detection", Expert Systems with Applications, 257, Article 125024. http://dx.doi.org/10.1016/j.eswa.2024.125024.

[4] Bai, Shuotian, Zhu, Tingshao, & Cheng, Li (2012). "Big-five personality prediction based on user behaviors at social network sites", (pp. arXiv–1204). ArXiv E-Prints.

[5] Bird, Steven, Loper, Edward, & Klein, Ewan (2009). "Natural language processing with Python", O'Reilly Media Inc.

[6] Boissonneault, David, & Hensen, Emily (2024)," Fake news detection with large lan- guage models on the LIAR dataset", http://dx.doi.org/10.21203/rs.3.rs-4465815/v1, Research Square, Version 1.

[7] Capuano, Nicola, Fenza, Giuseppe, Loia, Vincenzo, & Nota, Francesco David (2023), "Content based fake news detection with machine and deep learning: a systematic review", Neurocomputing, http://dx.doi.org/ 10.1016/j.neucom.2023.02.005.

[8] Castillo, Carlos, Mendoza, Marcelo, & Poblete, Barbara (2011)," Information credibility on Twitter", In Proceedings of the 20th International conference on world wide web (pp. 675–684). Association for Computing Machinery, http://dx.doi.org/10.1145/ 1963405.1963500.

[9] Datta, Niladree, Azim, Muhammad Anwarul, Chy, Abu Nowshed (2024)," Profiling fake news spreader on Twitter using stochastic gradient descent", In 2024 6th International Conference on Electrical Engineering and Information & Communication Technology (pp. 113–118). IEEE, http: //dx.doi.org/10.1109/ ICEEICT 62016.2024.10534470.

[10] Davoudi, Mansour, Moosavi, Mohammad R, & Sadreddini, Mohammad Hadi (2022), "DSS: A hybrid deep model for fake news detection using propagation tree and instance network", Expert Systems with Applications, 198, Article 116635. http://dx.doi.org/ 10.1016/j.eswa.2022.116635

[11] Zhang, Guobiao, Giachanou, Anastasia, & Rosso, Paolo (2022)," Scene ND: Multimodal fake news detection by modelling scene context information", Journal of Information Science, http://dx.doi.org/ 10.1177/01655515221087683.

[12] Zhou, Xinyi, & Zafarani, Reza (2019), "Network-based fake news detection: A pattern driven approach", ACM SIGKDD Explorations Newsletter, 21(2), 48–60. http://dx.doi. org/10.1145/3373464.3373473.

[13] Del Tredici, Marco, & Fernández, Raquel (2020), "Words are the window to the soul: Language-based user representations for fake news detection", In Proceedings of the 28th International Conference on Computational Linguistics (pp. 5467–5479). Barcelona, Spain (Online) International Committee on Computational Linguistics, http://dx.doi.org/10.18653/ v1/2020.coling-main.477.

[14] Shu, Kai, Zhou, Xinyi, Wang, Suhang, Zafarani, Reza, & Liu, Huan (2019)," The role of user profiles for fake news detection", In Proceedings of the 2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (pp. 436–439). http://dx.doi.org/10.1145/3341161.3342927.

[15] Mu, Yida, & Aletras, Nikolaos (2020)," Identifying Twitter users who repost unreliable news sources with linguistic information", PeerJ Computer Science, 6, Article e325. http://dx.doi.org/10.7717/peerj-cs.325.

[16] Shrestha, Anu, & Spezzano, Francesca (2022)," Characterizing and predicting fake news spreaders in social networks", International Journal of Data Science and Analytics, 1–14. http://dx.doi.org/10.1177/ 0956797619887896.

[17] Rangel, Francisco, Giachanou, Anastasia, Ghanem, Bilal, & Rosso, Paolo (2020). Overview of the 8th author profiling task at PAN 2020: "Profiling fake news spreaders on Twitter", In Linda Cappellato, Carsten Eickhoff, Nicola Ferro, & Aurélie Névéol (Eds.), CLEF 2020 labs and workshops, notebook papers. CEUR-WS.org.

[18] Ghanem, B., Rosso, P., & Rangel, F. (2020)," An emotional analysis of false information in social media and news articles", *ACM Transactions on Internet Technology (TOIT),* 20(2), 1–18. http://dx.doi.org/ 10.1145/3381750.

[19] Vo, N., & Lee, K. (2019)," Learning from fact-checkers: Analysis and generation of fact checking language", In Proceedings of the 42nd *International ACM SIGIR Conference on Research and Development in Information Retrieval,* (pp. 335–344). http://dx. doi.org/10.1145/3331184.3331248.

[20] Ghanem, Bilal, Ponzetto, Simone Paolo, & Rosso, Paolo (2020)," FacTweet: profiling fake news twitter accounts", *In International Conference on Statistical Language and Speech Processing* (pp. 35–45). Springer, http://dx.doi.org/10.1007/978-3-030-59430-5_3.

[21] Samadi, Mohammadreza, & Momtazi, Saeedeh (2023)," Fake news detection: deep semantic representation with enhanced feature engineering", *International Journal of Data Science and Analytics,* 1–12. http://dx.doi.org/10.1145/3472620.

[22] Potthast, Martin, Kiesel, Johannes, Reinartz, Kevin, Bevendorff, Janek, & Stein, Benno (2017)," A stylometric inquiry into hyperpartisan and fake news", http://dx.doi.org/ 10.18653/v1/P18-1022, arXiv preprint arXiv:1702.05638.