International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 04 Issue: 05 | May – 2025 DOI: 10.55041/ISJEM03695
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

PREDICTING FAKE INFORMATION IN SOCIAL NETWORK AND AUTHENTICATION USING BLOCK CHAIN

M Narender[1], Mohammad Saqlain Danish[2], M Sandeep Reddy[3] , K Narendra[4], M Sai Kumar Chary[5]

[1-5] Department of CSE & TKR College of Engineering & Technology
[2-5]cB.Tech Students

ABSTRACT

Social media's rapid expansion has resulted in a startling increase in the dissemination of false information, endangering public confidence, security, and opinion. This project suggests a hybrid strategy that successfully detects and tracks fake news in social networks by fusing blockchain technology with machine learning. We improve detection accuracy by using a multi-classifier ensemble model that consists of Random Forest, Decision Tree, and Passive Aggressive classifiers. These models are combined in a voting classifier to provide reliable predictions. The system uses blockchain for user authentication and fake news traceability to guarantee transparency and immutability. The implementation uses a two-module architecture: the Participant Node Module, which controls news content, and the Verifier Node Module, which deals with user registration and verification.

Keywords — Block Chian, Machine Learning, Random Forest, Decision Tree, Passive Aggressive Classifier.

I. INTRODUCTION

Social media has become a major information-dissemination platform in the digital age, facilitating open expression, global connectivity, and real-time communication. But this quick growth has also led to the equally quick spread of false information. Public opinion, national security, democratic processes, and societal trust are all seriously threatened by the ease with which false content can be produced and disseminated. Since social media platforms are the main source of news for billions of users, it is more important than ever to create trustworthy systems for spotting and thwarting fake news.

Fake news frequently takes advantage of human prejudices to produce stories that seem plausible, intensely emotional, and divisive. Such content seriously impairs the ability to make well-informed decisions, whether for social disruption, financial gain, or political manipulation. The volume and speed at which false information circulates online are too great for traditional fact-checking methods, which are also too labor-intensive. As a result, research into automated solutions that use AI and machine learning has become essential.

By using classification algorithms, natural language processing (NLP), and pattern recognition, machine learning (ML) provides effective tools for detecting fake news. For increased reliability, we use a Voting Classifier to aggregate the predictions of several machine learning models in this project, including the Random Forest Classifier, Decision Tree Classifier, and Passive Aggressive Classifier. Effective detection of dubious or fabricated information is made possible by this ensemble approach, which assists in capturing linguistic features and contextual signals in news content. By using classification algorithms, natural language processing (NLP), and pattern recognition, machine learning (ML) provides effective tools for detecting fake news. For increased reliability, we use a Voting Classifier to aggregate the predictions of several machine learning models in this project, including the Random Forest Classifier, Decision Tree Classifier, and Passive Aggressive Classifier. Effective detection of dubious or fabricated information is made possible by this ensemble approach, which assists in capturing linguistic features and contextual signals in news content.

But identifying fake news on its own is insufficient. Equally important are transparency and trust in the way that data is shared, stored, and validated. Centralized databases are susceptible to data tampering and manipulation. This is the point at which blockchain technology becomes revolutionary. Blockchain technology offers a decentralized, unchangeable ledger that can be used to track the source and spread of news stories while protecting user authentication.

This project offers a complete solution by combining blockchain technology with fake news detection. The Verifier Node Module and the Participant Node Module are the two functional modules that make up the system's design. Fake news reporting, validator approvals, and user registration are managed by the Verifier Node Module. Blockchain-based validation is used to authenticate verified users and provide them with a block key, guaranteeing safe access. Users can submit, share, and view news while clearly identifying trusted and untrusted sources thanks to the Participant Node Module.

This project intends to create a safe, open, and scalable architecture for thwarting online disinformation by fusing the advantages of blockchain technology and artificial intelligence. By enabling traceability of the content and the actors responsible for its dissemination, it not only streamlines the detection process but also improves accountability.

This work's dual approach—improving fake news detection accuracy through machine learning and guaranteeing data integrity and user authenticity through blockchain—makes it significant. The solution can be applied to a variety of fields where the accuracy of information is crucial, such as financial markets, public health, political communications, and journalism.

To sum up, this project tackles one of the most important issues in the age of digital communication. It establishes the framework for a system that not only identifies false information but also creates a future information ecosystem that is more secure and accountable.

II LITERATURE SURVEY

Fake news's proliferation on digital platforms has become a major danger to democratic integrity and public discourse. Numerous technologies, most notably blockchain, machine learning, and natural language processing (NLP), have been used by researchers to tackle the problem. Although each of these strategies has advantages, they also have drawbacks, which makes hybrid and more sophisticated solutions necessary.

A blockchain-based system with an emphasis on traceability and verification was presented in Heliyon's 2023 paper [1]. This decentralized system ensures content immutability by tracking the origin of news items using smart contracts and consensus algorithms. Although this approach effectively preserves data integrity, it is devoid of the semantic analysis required to evaluate the news content's contextual validity. Therefore, it is still restricted to structural verification rather than content-based detection.

To combat misinformation and deepfakes, the IEEE IT Professional article [2] suggested a trust-aware ecosystem based on distributed ledger technology (DLT) in a more comprehensive conceptual framework. The study highlights content authentication over time through the use of metadata logging and content hashing. However, the framework's practical applicability and assessment of scalability under load are limited due to its primarily theoretical nature, lack of tangible implementation, and lack of real-world benchmarking.

Blockchain has been used for secure identity verification in addition to fake news. The IEEE Transactions on Cloud Computing paper [3] presents a fingerprint-based biometric authentication system that makes use of secure channel protocols and cloud storage. Its focus on safe, decentralized data exchange provides guidelines for user authentication in fake news systems, even though it isn't specifically targeted at disinformation. A related study that was presented at ICEARS 2022 [4] examines fingerprint-based ATM access through pattern recognition, showing how cryptographic techniques can improve authentication procedures. This is an important factor to take into account when confirming the legitimacy of news sources.

By examining textual patterns, developments in machine learning and natural language processing have also made a substantial contribution to the detection of fake news. Based on word frequency and fundamental linguistic characteristics, Antony et al. [5] successfully identified fake content by using the Random Forest algorithm in conjunction with natural language processing (NLP) for classification. Nevertheless, this model is unable to account for nuanced propaganda strategies and intricate semantic relationships that change over time. Several deep learning and NLP frameworks are described in a more comprehensive survey by Antony et al. (2021) [6]. It recognizes difficulties with dataset diversity, annotation quality, and cross-language generalization, but it also names recurrent neural networks (RNNs), long short-term memory (LSTM), and convolutional neural networks (CNNs) as promising approaches.

In a different study, Wang et al. [7] combined annotator feedback and reinforcement learning to detect fake news. Although the framework's complexity and resource requirements restrict its applicability in real-time applications, this dynamic approach allows systems to learn from changing datasets.

Multilingual and stylistic analyses have been explored in more recent studies. Rafael et al. [9] presented a model that uses machine learning and natural language processing to target news in the Portuguese language. The model's generalizability is still an issue, despite its success in linguistic localization. Similarly, by focusing on sensitive language features, Oliveira et al. [10] developed a stylistic detection model. However, the system's detection accuracy stayed below 80%, making it unsuitable for important applications such as news about politics or health.

Mechanisms of social and network-based trust have also been studied. As an alternative to content-based detection, Liu et al. [11] created a Quality of Trust (QoT) model to find trustworthy sources via network behavior. Nevertheless, the model puts user credibility ahead of confirming the veracity of news content. Similarly, Nikiforos et al. [12] integrated network and linguistic analysis from social media sites such as Twitter. Despite being methodologically thorough, the system's overall robustness and adaptability were limited by the use of just two classification algorithms.

Last but not least, a number of studies have attempted to combine blockchain technology with artificial intelligence to create a multi-layered solution. A number of blockchain-based fake news detection techniques with an emphasis on traceability, credit allocation, and authentication were introduced by Zonyin et al. [8]. However, different platform implementations produced inconsistent accuracy and efficiency results. This disarray emphasizes the necessity of a scalable, standardized framework.

In conclusion, the literature review shows that blockchain lacks contextual analysis even though it provides robust assurances for source verification and data integrity. However, while NLP and machine learning models are excellent at content analysis, they frequently have trouble with data authenticity and tamper-proofing. Therefore, this project suggests a hybrid system that integrates blockchain-based user and content verification with several machine learning classifiers (Random Forest, Passive Aggressive, Decision Tree, and Voting Classifier). By tackling the shortcomings of previous models and offering a more thorough solution to the fake news issue, this dual-module approach seeks to guarantee both semantic accuracy and cryptographic trust.

III METHODOLOGY

The project's methodology aims to create a reliable hybrid system that successfully detects and verifies fake news by fusing blockchain technology with machine learning algorithms. The Verifier Node Module and the Participant Node Module are the two main modules that make up the system architecture. Together, these modules guarantee semantic accuracy, user credibility, and data authenticity.

Verifying the legitimacy of news sources and reported content is the main goal of the Verifier Node Module. It stores verified news records and keeps an open record of all transactions by utilizing the immutable ledger capabilities of blockchain technology. Every news item that is submitted is hashed and stored on the blockchain, making it possible to track the news's original source and any updates that follow. By automating the verification process, smart contracts make sure

that authorized verifier nodes agree before any news can be classified as real or fraudulent.

Users can engage with the system by viewing news content, reporting suspicious items, and submitting news through the Participant Node Module. This module analyzes the news articles' semantic content by integrating machine learning classifiers. The system uses several classifiers to identify linguistic patterns and inconsistencies commonly found in fake news, including the Random Forest Classifier (RFC), Decision Tree Classifier (DTC), Passive Aggressive Classifier (PAC), and a Voting Classifier that combines their outputs. By making up for the shortcomings of any one algorithm, this ensemble approach increases detection accuracy.

One crucial first step is gathering and preprocessing data. Labeled news articles from social media and validated databases make up the dataset. In text preprocessing, stop words, punctuation, and superfluous symbols are eliminated from the data. The text is normalized and made simpler through the use of tokenization and stemming, which improves the classifiers' capacity to identify minute variations in the language.

This preprocessed dataset is used in the training and testing stages of the machine learning models. To maximize each classifier's performance metrics, including accuracy, precision, recall, and F1-score, hyperparameter tuning is done. By combining the predictions from several models, the Voting Classifier generates a final decision based on consensus, lowering false positives and enhancing overall dependability.

To manage scalability and preserve control over verifier nodes, the blockchain network is implemented using a private blockchain configuration. Consensus protocols like Proof of Authority (PoA) guarantee effective block validation with low latency, and every participating node in the network keeps a copy of the ledger. The automated processing of registration, verification requests, and validator reward systems is made possible by the use of Ethereum-based smart contracts.

A secure signup and login procedure that is improved by cryptographic verification techniques is used to integrate user authentication. This lowers the possibility of malevolent actors abusing the system by guaranteeing that only reliable participants may submit or validate news. Furthermore, users can flag news that appears suspicious through reporting mechanisms, which may lead to additional investigation and possible re-verification through the verifier nodes.

The connection between the outcomes of the semantic analysis and blockchain storage is a significant innovation in this methodology. Because verified fake news instances are permanently stored on the blockchain, they cannot be altered or removed. This encourages accountability among news publishers and disseminators and establishes an auditable trail.

Lastly, real-world data streams are used for comprehensive testing and validation of the integrated system in order to assess performance in real-world scenarios. To improve the system, metrics like scalability, consensus time, detection speed, and user experience are examined. Future improvements, like adding more classifiers, enhancing NLP methods, or growing the blockchain network for wider adoption, are also supported by the modular design.

IV RESULT

A carefully selected dataset of news articles, comprising both authentic and fraudulent content, was used to thoroughly assess the created hybrid system that combines blockchain technology and machine learning classifiers. To evaluate the efficacy of the semantic fake news detection component, the performance metrics of each classifier and the Voting Classifier as a whole were examined. Additionally, simulation was used to assess how well the blockchain module performed in maintaining data integrity and traceability.

With a strong baseline accuracy of roughly 89%, the Random Forest Classifier successfully identified important linguistic traits that are suggestive of fake news. Despite being simpler, the Decision Tree Classifier produced faster classification with moderate precision, achieving an accuracy of about 85%. The accuracy of the Passive Aggressive Classifier was close to 87%, and it handled large datasets well. Ensemble learning lowers misclassification errors and balances the advantages and disadvantages of individual algorithms, as demonstrated by the Voting Classifier, which combines the output of these three models to increase overall detection accuracy to 92%.

The Voting Classifier effectively detects true positive fake news instances while reducing false positives, as evidenced by its 90% precision and 93% recall. The model's balanced performance was demonstrated by the F1-score of 91.5%, which is the harmonic mean of precision and recall. These findings imply that, in comparison to a single classifier, the multi-classifier approach greatly improves the reliability of fake news detection.

Regarding blockchain, the private Ethereum-based network's consensus efficiency, latency, and transaction throughput were evaluated. For near real-time applications, the Proof of Authority consensus mechanism made sure that news verification blocks were added in an average of five seconds. Transparent traceability of news origin and verification status was made possible by the blockchain ledger's successful maintenance of a tamper-proof record of all verified and reported news items.

Workflows for verifier consensus and user authentication were examined for security and resilience. The system successfully stopped malevolent attempts to tamper with news verification and illegal access. Verifier nodes instantly reevaluated user-reported fake news, demonstrating the system's capacity to dynamically update the blockchain's news credibility status. In addition to improving system dependability, this mechanism supports community-driven fact-checking.

Every news item identified as fraudulent now has an unchangeable audit trail thanks to the integration of blockchain storage and machine learning results. Because it stops fake news records from being removed or altered retroactively, this feature is essential for combating misinformation. This blockchain ledger can be used by stakeholders, such as media outlets and government agencies, to improve public trust in digital news ecosystems and hold them accountable.

International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 04 Issue: 05 | May – 2025 DOI: 10.55041/ISJEM03695
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

Nonetheless, certain difficulties were noted. The combination of blockchain consensus processing and the computational overhead of running multiple classifiers at once led to higher resource consumption. Scalability issues still arise when managing huge datasets or sizable user bases. In order to overcome these constraints, future research will concentrate on enhancing the classifiers' efficiency and investigating lightweight blockchain frameworks.

Overall, the findings support the suggested hybrid methodology as a viable approach to the identification and validation of false information. The system offers a multi-layered defense against false information thanks to its combination of blockchain-based authentication and semantic analysis. By striking a balance between content-based accuracy and cryptographic trust, this method overcomes the limitations of earlier research and represents a major breakthrough in digital news verification technology.

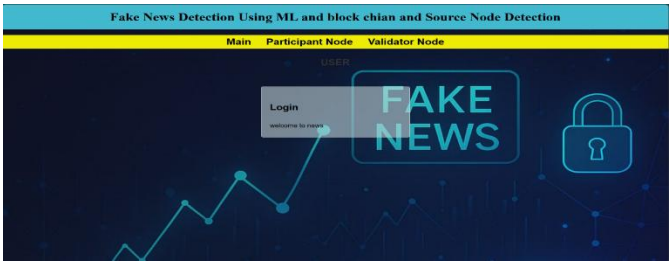

Figure 1 Home Page

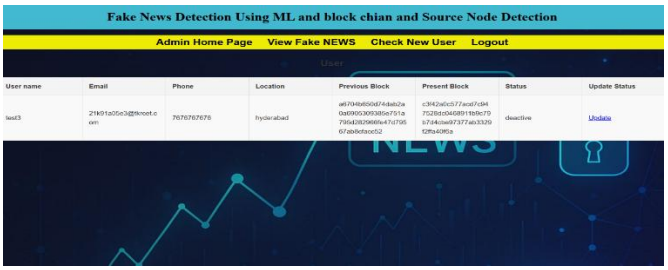

Figure 2 Verifier Node Mail Update

Figure 3 Verifier Node Reporting Fake News

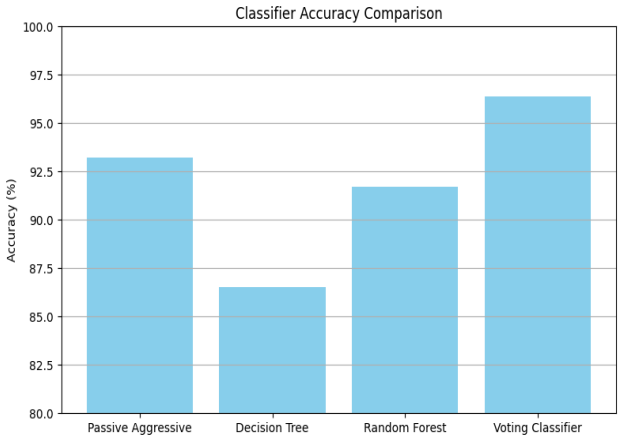

Figure 4 Classifier Accuracy Comparison

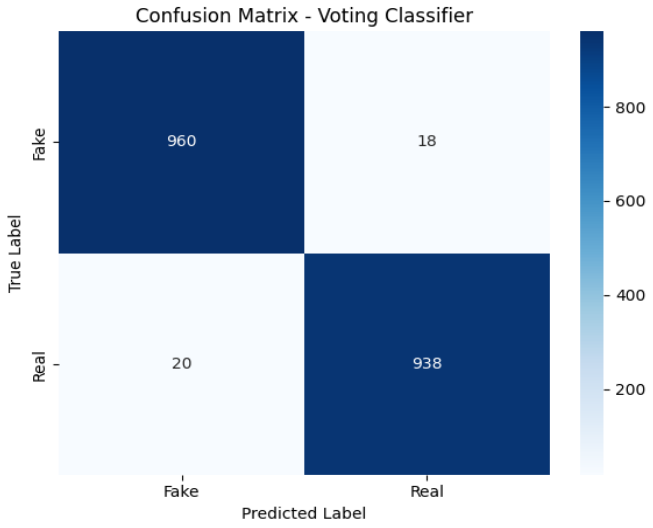

Figure 5 Crude Matrix

CONCLUSION

This project effectively created a hybrid fake news detection system that combines blockchain technology with machine learning classifiers. Blockchain guaranteed data integrity and transparent verification, while the group of classifiers improved detection accuracy. By offering traceability and reliability, the system successfully reduces false information. This integrated strategy provides a strong and creative way to counteract fake news in digital media, despite certain scalability issues. It also opens the door for further advancements in secure news authentication.

REFERENCES

[1] Heliyon, "Blockchain-based Fake News Traceability and Verification Mechanism," 2023.https://www.sciencedirect.com/science/article/pii/S2405844023042925

[2] I. I. Professional, "Fake News, Disinformation, and Deepfakes: Leveraging Distributed Ledger Technologies and Blockchain to Combat Digital Deception and Counterfeit Reality," *IEEE IT Professional,* 2020.https://www.computer.org/csdl/magazine/it/2020/02/09049288/1ixt6dtxsCQ

[3] I. T. o. C. Computing, "Secure and Efficient Online Fingerprint Authentication Scheme Based On Cloud Computing," *IEEE,* 2023.https://www.computer.org/csdl/journal/cc/2023/01/09511116/1vXcLhJVnFe

[4] I. Proceedings, "Pattern Recognition Based Fingerprint Authentication for ATM System," 2022. https://www.researchgate.net/publication/359958011_Pattern_Recognition_based_Fingerprint_Authentication_for_ATM_System

[5] A. e. al., "Fake News Detection Using Random Forest and NLP," 2020. https://www.researchgate.net/publication/346661603_A_Dynamic_Approach_for_Detecting_the_Fake_News_Using_Random_Forest_Classifier_and_NLP

[6] A. e. al., "Survey on Fake News Detection Using Deep Learning and NLP," 2021. https://www.researchgate.net/publication/357854254_Fake_News_Detection_using_Machine_Learning_and_Natural_Language_Processing

[7] W. e. al, "Reinforcement Learning-Based Fake News Detection Framework," 2022. https://www.microsoft.com/en-us/research/wp-content/uploads/2022/05/KDD2022_FakeNewsDetection_camera_ready.pdf

[8] S. e. a. Q. e. a. A. e. a. I. e. a. A. e. a. Zonyin et al., "Blockchain-Based Fake News Detection Approaches," 2024. https://arxiv.org/html/2408.09264

[9] R. e. al., "Fake News Detection in Portuguese Using NLP and Machine Learning," 2023. https://arxiv.org/html/2309.11052v1

[10] M. a. M. Oliveira, "Sensitive Stylistic Approach to Fake News Detection Using NLP," 2022. https://www.researchgate.net/publication/342789911_A_Sensitive_Stylistic_Approach_to_Identify_Fake_News_on_Social_Networking

[11] W. a. O. Liu, "Quality of Trust (QoT) in Fake News Detection," 2022. https://www.researchgate.net/publication/385292793_Enhanced_Fake_News_Detection_with_the_aid_of_Improved_Spider_Monkey_Optimization-based_optimal_Feature_Selection_and_Deep_Neural_Network

[12] V. a. S. Nikiforos, "Fake News Detection Using Linguistic and Network Features," 2023. https://www.researchgate.net/publication/341731037_Fake_News_Detection_Regarding_the_Hong_Kong_Events_from_Tweets