Fake Media Detection Using NLP and Blockchain Framework

Vaibhav Sharma[1], GD Makkar[2], Pradeep Semwal[3], Harish Chandra Sharma[4], Archana Kero[5], Minit Arora[6*]

[1, 2, 3, 4, 6] School of Engineering & Technology, Shri Guru Ram Rai University, Dehradun, Uttarakhand-248001, India
[5] CSE Department, Shivalik College of Engineering, Dehradun, Uttarakhand-248197, India
Email: vsdeveloper10@gmail.com[1], gdmakkar@gmail.com[2], psemwal2222@gmail.com[3], hcs19@yahoo.com[4], archanakero@gmail.com[5], minitarora@gmail.com[6]
*Corresponding Author

Received: 24.04.2024 Revised : 26.05.2024 Accepted: 28.05.2024

ABSTRACT
The spread of false information on social media presents serious risks to the accuracy of information on a variety of platforms. Conventional detection techniques frequently find it difficult to keep up with the increasing complexity of fake content. This study presents a novel method of detecting false media by fusing Blockchain-based frameworks with Natural Language Processing (NLP) approaches. The suggested method analyzes and finds misleading patterns in text, photos, and videos by using natural language processing (NLP) techniques. In order to provide a decentralized and safe verification system, Blockchain technology is also used to guarantee the immutability and transparency of the detection process. Through this integration of technology, the framework provides a reliable and scalable way to identify and stop the spread of false information. In order to improve the precision, reliability, and traceability of material in digital ecosystems, this research shows how NLP and Blockchain can be integrated.

Keywords: Blockchain, NLP, platforms, technology, immutability.

INTRODUCTION

An important and novel strategy to counter the pervasive spread of misinformation in the current digital era is the identification of erroneous information using unsupervised models. Public discourse, democracy, and even public safety are at risk from the spread of false or misleading content, which has become a major worry with the growth of internet platforms and social media. Unsupervised models for fake news detection do not require pre-labeled training data; instead, they rely on the natural patterns and traits of textual data to distinguish between real news and bogus content. Through the application of methods like natural language processing, clustering, and anomaly detection, these models aim to automatically identify misleading stories and possibly hazardous data, offering a proactive and scalable response to the widespread problem of distinct fake news.

Enabling computers to understand text and spoken words in a way that is comparable to human comprehension is the main goal of the software engineering discipline known as "natural language processing" (NLP). Facts, artificial intelligence, and sophisticated learning models are combined with computational etymological rule-based human language display in NLP. These developments let computers to fully "comprehend" human language as message or auditory information, including the expectations and points of view of the essayist or speaker. Neural Language Processing (NLP) powers computer programs that translate text between languages, respond to spoken requests, and rapidly and even continuously summarize massive amounts of material. Digital assistants, speech-to-message transcription software, chatbots for customer service, voice-activated GPS devices, and other retail conveniences all employ natural language processing (NLP). On the other hand, NLP also plays a major role in large-scale commercial tactics that enhance important company processes, encourage worker productivity, and streamline operations.

Data can be kept in a way that makes fraud, hacking, and system alterations difficult or impossible when using a block chain. In its most basic form, a block chain is a computer network that duplicates and distributes a digital ledger of all network transactions. Every new transaction that happens on the block

chain is copied into each participant's ledger, and each block in the chain is made up of many transactions. Distributed ledger technology (DLT) is the term used to describe a decentralized database that is managed by several users. An unchangeable transactional record that has undergone cryptographic authentication and is shared by all network users is called a block chain. It is a continuously growing database. Every record refers to earlier transactions and has a time stamp. Access-right holders may utilize this data to go back in time to any point in the past associated with a transactional event that they own. Among the more generalized ideas of networked ledgers is a block chain.

Information deemed inaccurate or misleading but presented as news is known as fake news. A common goal of false news is to destroy someone or something's reputation or to generate generations of income from advertising. The term "fake news" was coined in the 1890s, a time when spectacular newspaper reports were commonplace, despite the fact that incorrect information has always been circulated throughout history. The word, for which there is no precise definition, is commonly used to refer to all incorrect information. It has also been used by well-known individuals to characterize any bad news about them. Intentionally distributing false information is known as disinformation, and hostile foreign entities frequently create and disseminate it, particularly during election seasons. Fake news can include articles with sensationalist or clickbait headlines that don't have any supporting information, as well as satirical pieces that are mistaken for the real thing. Researchers are starting to use the more objective and informative term "information disorder" to describe the multitude of fake news sources.

LITERATURE REVIEW

A tremendous amount of data is generated every second by different applications, users, and devices due to the rise in data traffic caused by the quick development of communication technologies and smart gadgets. Due to resource limitations, there is a growing need for concepts that may be used to analyze data changes over time. These solutions are known as concept drifts. Ahmad Abbasi [1] et al. describe in their work a novel method called ElStream that combines both real and artificial data to detect concept drifts using ensemble and traditional machine learning approaches. ElStream makes decisions by using the majority voting technique, which limits the number of votes to the best classifier. ElStream outperforms earlier state-of-the-art research and traditional machine learning algorithms in terms of accuracy. Experimental study demonstrates that the ensemble learning approach gives consistent performance for both simulated and real-world datasets. Due to its potential to offer priceless insights and advantages including cost savings, quicker decision-making, and innovation in new products across numerous industries, big data has attracted a lot of interest in the previous 10 years. Analysis is made more difficult by the fact that this data is frequently available in continuous streams. Big data is so complicated that typical data analysis methods are no longer useful. Alternatively, systems may recognize patterns and learn without explicit programming thanks to machine learning techniques.

Today's globe faces a serious problem with fake news [2], partly because social media is so widely used. Making sure information appears on social media is authentic requires checking that it originates from reliable sources. Still, it's difficult to gauge how serious and authentic online news can be. We provide a FNU-BiCNN model in this work that preprocesses data using NLTK features including stop words and stem words. After that, we use LSTM, batch normalization, and dense to compute the TF-IDF and use the WORDNET Lemmatize to select features. The datasets are trained using a Bi-LSTM with ARIMA and CNN, and then classified using a variety of machine learning methods. This model creates an ensemble approach to simultaneously learn news story depictions, authors, and titles by calculating believability scores from textual data. A Voting ensemble classifier is employed to attain higher accuracy, and its performance is evaluated against other machine learning techniques, including SVM, DT, RF, KNN, and Naive Bayes. With an accuracy of 99.99%, the voting ensemble classifier demonstrated the maximum performance, according to our findings. We evaluate classifiers based on their accuracy, recall, and F1-Score to determine their effectiveness. Malicious URLs have been detected by combining a trust computational model with a set of URL-based attributes. A 95% accuracy rate in malicious URL identification is attained by evaluating tweets' credibility with the use of Bayesian learning and Dempster-Shafer theory.

According to Chang Li [3] et al.'s article, online arguments can offer insightful information about a range of viewpoints. But comprehending the positions taken in these discussions is a challenging endeavor that necessitates modeling both the textual content and the user activities during the conversation. Existing methodologies adopt a collective classification approach, ignoring the connections among various debate issues. We propose to approach this issue as a representation learning problem in this work, and jointly embed the authors and the text according to their interactions. We analyze various methods for embedding structural information and assess our model using the Internet Argumentation Corpus. The experimental results show a considerable improvement over previous competitor models with our model.

The political discourse has been shaped by social media platforms in the last few years. Users can voice their opinions and converse with those who have opposing perspectives in online debate forums. Gaining knowledge from the way users interact on these platforms might help us better understand the political discourse, how people argue, and how the public feels generally about policy problems. A lot of study has been done to try to represent the relationships between writers and their content while capturing the debate structure because of how important it is to comprehend dispute conversation. One way to forecast similar or different views connected with their material is to look at how users engage with one another. This can reveal either agreement or disagreement. Since users typically stick to their positions during debates, resulting in stance agreement across all postings, the choice of stance can also be considered a user-level decision. Regretfully, posture categorization continues to be a difficult issue in spite of these study initiatives.

The idea of grouping reports—a crucial component in the domains of data and software engineering—was introduced by Umar Mohammed Abatcha [4] et al. in their article. This is the process of precisely classifying archives into designated groups, which is regarded as an essential way to arrange data. The number of reports has been rising steadily along with the development of technology and personal computers. As a result, it is crucial to organize these archives according to their content. Text classification is a widely used technique for classifying text into different groups. It consists of several steps that can be taken in different ways. Choosing the right approach for each category is essential to improving text processing performance. The difficult process of classifying archives according to their content is essential to the work of researchers and data specialists. It is essential to many applications, such as the planning, arranging, arranging, and effective management of massive amounts of data. This is especially crucial for bloggers, publishers, news organizations, and anyone working with large internal information libraries in an organization. Record sorting and grouping facilitates the easy execution of duties in that area. Records, both physical and digital, are growing both in quantity and magnitude.

This work introduces a novel feature selection technique and applies it to an actual data collection, as suggested by Aparna Kumari [5]et al. The recommended method specifically creates attribute subsets according to two standards: (1) distinct attributes with strong discrimination (classification) power; and (2) the attributes in the subset that complement one another by incorrectly identifying different classes. Using data from a confusion matrix, the approach assesses each attribute separately. The fundamental goal of classification issues is to achieve good classification accuracy, but finding the qualities with the highest separation power is equally interesting. Furthermore, feature selection has a significant impact on the classification process when dealing with big data sets, like brain MRI scans. This is mainly because data becomes increasingly sparse as the number of attributes rises, requiring a notably higher amount of training data to adequately represent such a broad domain. This leads to the underrepresentation of high-dimensional data sets, a phenomenon known as "the curse of dimensionality" in the literature. For example, the domain delineated by the corners (0,0) and (1,1) can be sufficiently covered by a 2attribute data set containing 10 samples. Ten points won't adequately cover the enlarged area, though, if the domain to be taught is a 3-D cube bounded by the corners (0,0,0) and (1,1,1). As such, minimizing the quantity of attributes in a classification issue is a well-researched topic.

People rely a lot on social media to consume and share news, which leads to the widespread spread of both real and false tales. Misinformation across a wide range of social media platforms has serious ramifications for society. The difficulty of differentiating between various types of incorrect information is a significant obstacle to efficiently detecting fake news on Twitter. Researchers have made headway in addressing this problem by concentrating on techniques for spotting bogus news. Four criteria for recognizing false news make up the FNC-1 dataset, which will be used in this study. Machine learning and big data technology (Spark) will be used to assess and contrast the most recent methods for identifying false news. The study's methodology entails building a stacked ensemble model using a decentralized Spark cluster. The suggested stacked ensemble classification model is applied following feature extraction using N-grams, Hashing TF-IDF, and count vectorizer.

PROPOSED SYSTEM

The methods of false news detection suggested are processing, reinforcement learning, and block chain technology. Gathering a sizable dataset of news stories with metadata, including author, date, and source, is part of the system. In order to clean and tokenize the text, the gathered data would be pre-processed using NLP methods. We would extract attributes like sentence length, readability, and word frequencies from the preprocessed data. To differentiate between genuine and fraudulent news, an RL agent would be trained using the features that were extracted. Correctly classifying news as false would earn the agent points, while misclassifying factual news as false would result in punishment. With its extracted features, the agent may be trained to determine if a new news story is true or false.

Blockchain technology and natural language processing can be used to recognize and detect fraudulent media, which is one possible way to stop the spread of misinformation and fake news. Examining the structure of news articles—which includes components like the headline, introduction, body, and conclusion—is a feasible strategy for handling this problem. It is possible to identify trends in news content that may indicate the presence of false media by closely examining how news is organized. To analyze news item content and spot trends that can point to the existence of false media, one approach is to use natural language processing, a branch of artificial intelligence that focuses on the interaction between computers and human language. NLP techniques, for example, can be used to evaluate the language used in news items and spot anomalies that might point to the existence of false media.

Blockchain and natural language processing-based approaches to false media identification can be further improved using data authentication mechanisms. In order to identify false media, it is essential to confirm the authenticity and integrity of the content being analyzed. Using digital signatures, which verify the news article's source, is one efficient way to provide data authentication. Cryptographic algorithms are used to create these digital signatures, which are used to confirm the accuracy of the data. The tamper-proof quality of the digital signature is guaranteed by appending it to the news story and preserving it on a blockchain, making verification simple. Inconsistencies in the data can also be found using machine learning methods. Training these algorithms, for example, can be used to detect linguistic discrepancies between the headline and substance of a bogus news story. Then, these discrepancies can be reported as possibly phony media.

In a Proof of Authority (PoA) system, a set of trustworthy validators is assigned to verify blockchain transactions. Most often, these validators are respectable companies or people with a solid reputation for integrity and honesty. Their duty is to add news articles on the blockchain after confirming their veracity. A system that is resistant to malicious actor attacks and capable of identifying false media can be created thanks to PoA. It is less probable that the validators will commit fraud or work with other validators to corrupt the system because they are reliable and have a reputation to uphold. Techniques for natural language processing can be applied to analyze news article language and spot possible instances of false media. The analysis's findings can subsequently be shown to the validators for approval.

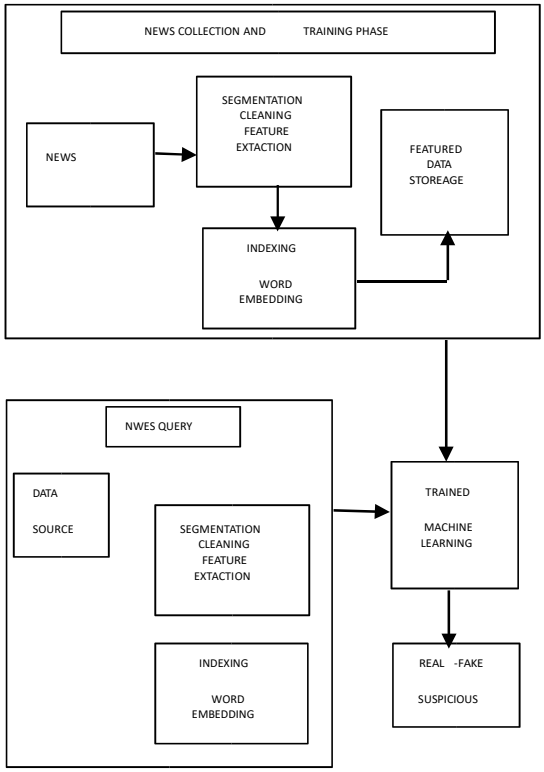

Figure 1. Block diagram

Detecting and countering the spread of bogus media can be effectively achieved by the application of block chain and natural language processing techniques.

Table 1. Comparison table

algorithm	accuracy	precision	recall	f1 score
NLP	89.67	88.78	86.18	87.46
RL	93.75	92.86	94.67	93.76
block chain	94.43	92.68	94.18	93.43

Anything from photos or videos to news stories that are purposefully designed to mislead the audience is considered fake media. To evaluate news articles' wording and spot possible instances of false media, natural language processing techniques might be utilized. If there is a difference between the headline and the body of the story, for example, NLP can identify discrepancies in the language used in the news item. The sentiment of the article can also be analysed using NLP, which can also spot any bias or false information. A tamper-proof and secure system for storing and authenticating news articles can be established with the use of block chain technology. It is simple to verify the integrity of news articles by assigning each one a distinct digital signature that is kept on the block chain. Establishing a system that is resistant to fraud and manipulation through the use of a block chain makes it possible to guarantee the authenticity of the data being analyzed.
Analyze the news article data beforehand. Tokenization, stop word removal, and text cleaning are all part of this process. Utilize the pre-processed data to extract characteristics. Readability, sentence length, and word frequency are some of the parameters that must be extracted for this. Apply the retrieved features to the deep learning model's training.

RESULT ANALYSIS

The efficacy of the suggested approach in identifying false news can be assessed using a variety of measures, including precision, recall, and F1 score. While accuracy measures the ratio of real positives to all projected positives, recall measures the ratio of genuine positives to all actual positives. Due to the weighted average nature of the F1 score, a higher score denotes superior performance.

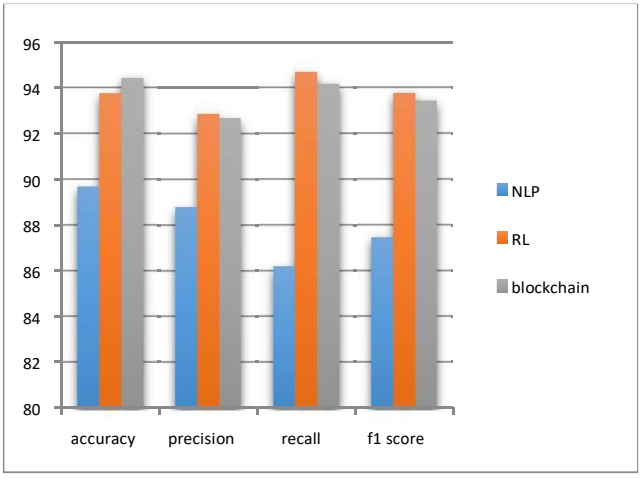

Figure 2. Comparison graph

The effectiveness of the suggested system can be assessed by contrasting its predictions with a labeled dataset of real and fake news stories. After that, the framework's expectations can be examined to determine its accuracy, review, and F1 score. It is also possible to assess the system's effectiveness by contrasting it with other state-of-the-art false news detection techniques. The suggested system's overall efficacy in identifying false news is influenced by the caliber of the dataset, the efficiency of the NLP techniques used to preprocess the data, the RL agent's architecture, and the precision of the block chain technology used to secure the data. In-depth testing and analysis are needed to assess the system's effectiveness and pinpoint areas in need of improvement. The feasibility of the project and the possibility that the system will benefit the organization are examined in the preliminary investigation. Testing the technical, operational, and financial viability of adding new modules and fixing an outdated operating system is the primary goal of the feasibility study. If resources are limitless and time is infinite, then any system can work.

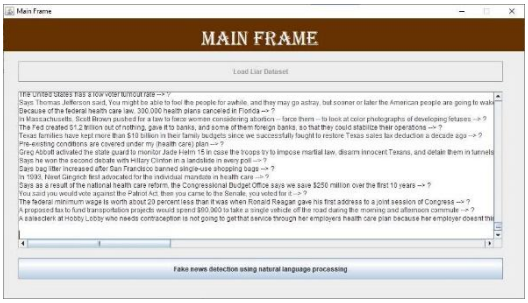

Figure 3. Main Frame

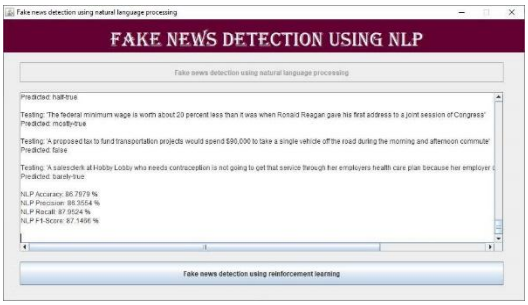

Figure 4. NLP

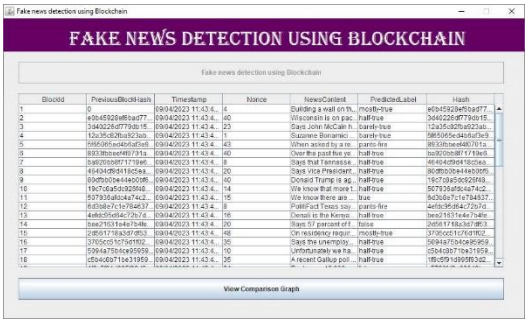

Figure 5. Blockchain

CONCLUSION

In conclusion, since spreading false information can have serious consequences in the modern world, identifying fake news is an important task. A promising answer to this problem is the proposed method for identifying bogus news, which makes use of blockchain technology, natural language processing, and reinforcement learning. An RL agent can be taught to recognize patterns that distinguish between genuine and fake news by using NLP techniques to pre-process and extract features from news items. It is also difficult for someone to alter the data covertly since blockchain technology ensures the authenticity and integrity of the data that has been analyzed. Basically, this approach that is being suggested might be extremely important in preventing the spread of fake news and encouraging the distribution of true information.

Future studies could look into ways to improve the suggested method even more in the area of false news identification. As more features might be looked into to strengthen the RL agent's ability to distinguish between real and fake news, the feature extraction procedure offers a chance for possible development. Further promising approaches to improving the system's overall performance are advanced natural language processing (NLP) techniques, such as deep learning models.

REFERENCES

[1] Augenstein, T. Rocktäschel, A. Vlachos, and K. Bontcheva conducted a study on stance detection using bidirectional conditional encoding in their paper published in 2020 on arXiv:1606.05464.

[2] M. Taulé, M. A. Martí, F. M. Rangel, P. Rosso, C. Bosco, and V. Patti provided an overview of the task on stance and gender detection in tweets related to Catalan independence at IberEval 2017. This work was presented in the proceedings of the 2nd Workshop on Evaluating Human Language Technologies for Iberian Languages (CEURWS), volume 1881, in 2017.

[3] M. Lai, A. T. Cignarella, D. I. Hernández Farías, C. Bosco, V. Patti, and P. Rosso focused on multilingual stance detection in social media political debates. Their research was published in the journal Computational Speech and Language in September 2020, with the article number 101075.
[4] B. Riedel, I. Augenstein, G. P. Spithourakis, and S. Riedel proposed a baseline approach for the Fake News Challenge stance detection task. Their work, which presented a simple yet effective method, was published in May 2018.
[5] C. Dulhanty, J. L. Deglint, I. B. Daya, and A. Wong explored automatic disinformation assessment through deep bidirectional transformer language models for stance detection in their paper published in 2019.
[6] S. Ochoa, G. D. Mello, L. A. Silva, A. J. Gomes, A. M. R. Fernandes, and V. R. Q. Leithardt published "FakeChain: A blockchain architecture to ensure trust in social media networks" in Proc. Int. Conf. Qual. Inf. Commun. Technol. Algarve, Portugal: Springer, 2019, pp. 105–118.
[7] Y. Wang, W. Yang, F. Ma, J. Xu, B. Zhong, Q. Deng, and J. Gao published Weak supervision for fake news detection via reinforcement learning in Proc. AAAI Conf. Artif. Intell., vol. 34, 2020, pp. 516–523.
[8] Chokshi and R. Mathew's "Deep learning and natural language processing for fake news detection: Research." January 2021, SSRN Electronic Journal. [Online]. Available at papers.ssrn.com/sol3/papers.cfm with abstract id=3769884
[9] J. A. Vijay, H. A. Basha, and J. A. Nehru, "A Dynamic Technique for Identifying the False News Using Random Forest Classifier and NLP," Computational Methods and Data Engineering, 2021, Springer, pp. 331-341.
[10] In a paired textual input schema, "Deep learning for fake news identification," By D. Mouratidis, M. Nikiforos, and K. L. Kermanidis, published in Computation, vol. 9, no. 2, p. 20 in February 2021.