EasyChair Preprint

№ 10323

Fake News Detection Using Blockchain

Veetu Kaushik, Taslim Raza Ahmad, Jai Sanger and Renu Mishra

June 2, 2023

Fake News Detection Using Blockchain

Veetu Kaushik
Dept. of computer science engineering
Sharda University
Grater Noida, India
2019521381.veetu@ug.sharda.ac.in

Taslim Raza Ahmad
Dept. of computer science engineering
Sharda University Greater Noida, India
taslim932@gmail.com

Jai Sanger
Dept. Of computer science engineering
Sharda University Greater Noida, India
2019628544.jai@ug.sharda.ac.in

Renu Mishra
Dept. of computer science engineering
Sharda University Greater Noida, India
Renu.mishra@sharda.ac.in

Abstract : Social media networks have emerged as an essential component of human existence as a result of recent technological advancements in computing era. This climate has turned into a notable stage for trading data and news on various topics, as well as day to day reports, and it is the essential time frame for information assortment and transmission. This climate enjoys many benefits, however it likewise contains a ton of fake news and data that confounds clients and perusers about the data they require. The absence of real-time social media news and reliable information is a major concern for the system. We suggested a machine learning-based integrated solution that would better anticipate bogus user accounts and postings and identify false news for multiple blockchain components using natural language processing (NLP). The proposed methodology utilizes the Support Learning procedure and decentralized blockchain architecture, which gives an outline of computerized content power verification to improve the stage's security. The target of this framework is to offer a solid strategy for foreseeing and distinguishing fake news via social media stages.

Keywords : *Natural language processing , blockchain , fake media , reinforcement learning*

1. **INTRODUCTION** : “Fake news, disinformation and Misinformation is worse than an epidemic” is an statement of Marcia McNutt, president of the National Academy of Sciences of the United States. Since, everybody utilizes virtual entertainment to get to news content,The involvement of vindictive people in virtual entertainment means changing real news and spreading counterfeit news, which can hurt web-based entertainment distinction.This phenomenon, known as fake news, is facilitated by the large number of people using social media to share articles, posts, videos, and news. To address this problem, the quality of trust (QoT) is used to evaluate the accuracy of information. Machine learning text classification is used to improve the safety of online networks, by removing harmful and unwanted content to avoid policy issues and free up space on servers.Fake news is a type of misleading publicity, often used for political purposes. The language used in creating fake news is designed to manipulate people's emotions and spread false information. Identifying fake news involves assessing the veracity of shared information. As data becomes more unstructured and voluminous, there is a growing need for a reliable method of detecting fake news. Recent advancements in artificial intelligence, machine learning, and deep learning have made this task more feasible.Blockchain technology provides a reliable and decentralized platform for combating fake information

in digital data exchange. Each new block in a blockchain contains information from the previous block, ensuring the trustworthiness of stored data. This study focused on using blockchain, natural language processing, and AI to combat fake users and data on Facebook and Twitter. The proposed method includes a gamification approach and reinforcement learning to improve system quality and restrict the use of incorrect data in the future. This approach aims to prevent the extraction and spread of misleading information.This paper presents a framework to counteract misleading news using Natural Language Processing (NLP) for text analysis and establishing financial roots. The system utilizes authority protocol as evidence and employs reinforcement learning to identify and remove fake user accounts. The proposed strategy involves tracking the association between items, identifying common material, and stopping the spread of misleading information. The system also involves joining a permissioned blockchain network where only verified users can participate. The goal is to store news data in a secure and reliable platform.

In different sections of the paper ,we present the research problem and the existing challenges, discuss the state of the art in existing approaches for fake news detection.A Blockchain enabled and NLP based framework is presented as the core of the paper and it's analysis is done with major expected outcomes. At last we concluded and point out the future research directions in this domain.

2. LITERATURE SURVEY

With the advance utilization of social media applications in daily life, fake posting cause a mess in various platforms. This problem is identified and addressed by various authors to find a proper and suitable news on the topic searched. They all explored and found that, machine learning and blockchain can be combine together to overcome fake news sharing issue.

The review [1] takes a gander at the extreme changes in proficient political correspondences that are occurring in the computerized age. The heuristic possibilities of framework humanistic system, intuitive methodology, and hypothesis of informative activity can be generally integrated into the examinations on account of the multiparadigm approach's capability as the strategic establishment..

Author Name	Advantages	Disadvantages	Methodology
Nicollas R. [2]	computational expressive examination in light of NLP	It has 86% exactness and 94% accuracy	It Uses NLP and ML Calculations for calculating the exactness of the report.
Guanfeng Liu, [3]	idea of Quality of Trust (QoT) is used to show the level of reliability	better but not 100% accurate.	Monte Carlo-based estimate calculation.
Maria Nefeli Nikiforos, [4]	gives a spic and span informational index of 2,366 English- tweets	Covers only particular event not all and still in Research stage	Two ML models were utilized in preliminaries
Adline Rajasenah [5]	different exemplary ML and deep learning systems	better than the other wokrs but not 100% accurate.	This research is completely based on deep learning models.
Xishuang Dong [6]	This paper is also effective even with limited labeled data.	very limited labeled data is available, which is inadequate for the supervised model	proposed a deep semisupervised learning model of limited labeled data.
S. Kumar, [7]	achieved accuracy of 88.78%.	Tested 1356 instances so not a fully tested model.	Used(CNNs), (LSTMs), ensemble methods, and attention mechanisms.
P. Bahad, [8]	Bi-directional LSTM model	2 publicly available unstructured news	Bi-directional LSTM-recurrent neural network.

		articles datasets used	
G. Sansonetti, [9]	used two approaches online and offline analysis. online content	offline analysis is completely dependent on individual	Used deep learning for offline and online is done through classification

Correspondence innovation for "fake news" has the accompanying elements: rate with which these messages are spread, wide allure, and trouble finding .They have additionally started to "blockcade" political and correspondence between genuine political players. To ensure the veracity of the data that is circulated and the shortfall of sham disinformation, specialists should apply extra exertion in dissecting web reports.

In [2], to distinguish counterfeit news in virtual entertainment posts, they utilize a computational expressive examination in light of natural language processing and ML calculations are utilized. As far as identification quality, even with a layered decrease to one-6th of the quantity of unique elements, 86% exactness and 94% accuracy stick out. While our strategy can give a high certainty file to deciding if news is real or fake, it has a low above.

As a pert of various improving exercises in [3] ,they first present an original complex informal organization structure and a pristine idea of Quality of Trust (QoT) to show the ability to ensure a specific level of reliability in trust evaluation. The ideal social trust course decision is in this manner exhibited as a NP-Complete Multi-Constrained Optimal Path (MCOP) assurance issue. To resolve this issue, theypresent MONTE K, a proficient Monte Carlo-based estimate calculation. Our examination on a genuine world dataset of interpersonal organizations shows that our proposed calculation beats existing ones as far as proficiency as well as the nature of the social trust channels picked.

In [4], The quantity of individuals who utilize virtual entertainment and online data has soar as an immediate consequence of the quick improvement of organization administrations. These stages are utilized by pernicious records to spread misleading data and publicity across enormous client organizations. Thusly, a computerized framework for recognizing trickery and misleading data is required. This study gives a spic and span informational index of 2,366 English-language tweets about Hong Kong's occasions in August 2019 and a clear cut calculation for etymological and network-based misleading news recognition. Two ML models were utilized in preliminaries to approve our strategy, and it performed better compared to past examinations.

In [5] , Machine learning is especially important when it comes to fake social media posts because of the scope of ecological and social challenges. A false or true message that goes viral can be created by anyone. During general elections, political parties in nations like India use social media groups to spread fake messages across the nation. While some of the messages might be genuine, many of them are fakes. Social media groups can be used to send spam messages and fake messages with false information. Policymakers and private residents alike face a serious problem like this. Utilizing numerous AI draws near, specialists recognize spam interchanges and fake news. A subset of machine learning known as "deep learning" mimics the human brain in its ability to solve numerous challenging issues. As a matter of fact, deep learning techniques are right now utilized by society since they can effectively tackle text acknowledgment issues like recognizing fake news and spam. These models work rapidly on crude information by concentrating on

undeniable level qualities all alone. To recognize deceitful and spam correspondences, different exemplary ML approaches as well as deep learning systems are talked about in this article. An outline of a few examinations on the distinguishing proof of phony news utilizing conventional ML techniques and Deep Neurral Networks is likewise remembered for this review.

3. METHODOLOGY

Numerous algorithms are existing but unfortunately, all of which are based on traditional machine learning algorithms like SVM or Random Forest to prevent such fake news and lacks in authorization and security. To resolve the issue, we proposed a coordinated answer for different pieces of blockchain and regular language handling (NLP) to recognize misleading news and better expect counterfeit client records and postings utilizing ML draws near. Our methodology utilizes the Support Learning procedure.

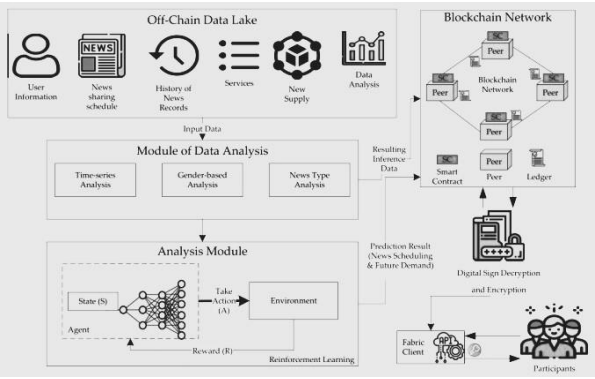


Fig 1 Proposed Architecture

As shown in the Figure , we proposed a decentralized blockchain design, which gives an outline of computerized content power confirmation, to improve the stage's security. Our goal is to offer a solid strategy for anticipating and distinguishing fake news via Block chain and NLP.Key benefits of this approach is that , Blockchain confirms all hashcodes prior to putting away new blocks, and on the off chance that all blocks are effectively checked, just the new block is put away; The term "proof of work" or "proof of authority" refers to this method. This innovation is intended to give a solid strategy to foreseeing and recognizing fake news via social media stages.

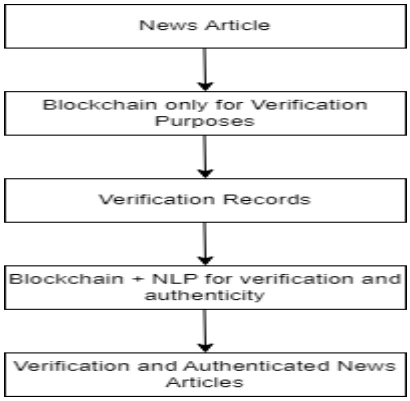
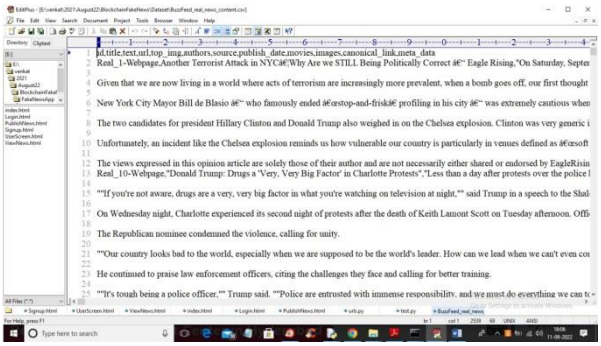


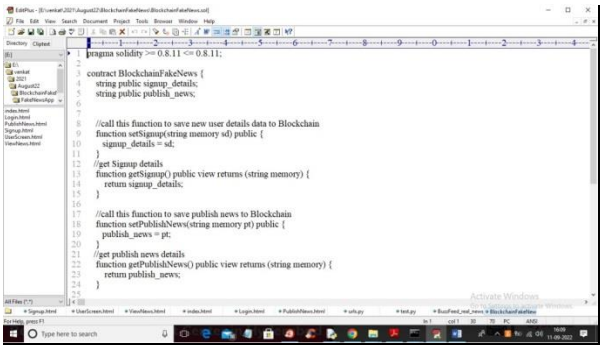
Fig.2 Detection with or without NLP

A blockchain is a conveyed record that is included a steadily extending rundown of passages (blocks) that are cryptographically hashed to safely connect them together. A timestamp, exchange information, and the cryptographic hash of the block that preceded it are totally remembered for each block. Exchange information is commonly portrayed as a Merkle tree, with information hubs being addressed by leaves. The timestamp shows that the exchange information were available when the block was made. They structure a chain (see connected list information structure), with each resulting block interfacing with the ones that preceded it on the grounds that each block contains data about the one that preceded it. Along these lines, blockchain exchanges can't be switched in that a solitary block's items can't be changed in that frame of mind without influencing every single ensuing block.

Blockchains are regularly overseen by a peer-to-peer (P2P) PC network for use as a public disseminated record. To add and check new exchange blocks, hubs on the whole stick to an agreement calculation convention. Because of the chance of blockchain forks, blockchain records might be viewed as secure by plan and act as a delineation of a circulated figuring framework with solid Byzantine adaptation to internal failure, despite the fact that they are not unalterable. A blockchain was underlying 2008 by an individual or gathering utilizing the name or nom de plume Nakamoto to act as the public circulated record for bitcoin digital currency exchanges. It depended on past work by Stuart Haber, W. Scott Stornetta, and David Bayer. Bitcoin was the principal computerized money to take care of the twofold spending issue without a focal server or believed power thanks to the execution of blockchain inside the cash. Blockchains and applications that can be perused by general society have arisen because of the bitcoin idea and are regularly utilized in digital forms of money. The blockchain may be viewed as a kind of portion rail. It has been proposed to utilize private blockchains for business purposes. Computerworld has named the offer of such private blockchains without a reasonable security design "snake oil." In any case, others have proposed that permissioned blockchains, whenever organized accurately, may really be more decentralized and, thus, more secure than permissionless blockchains.



In above screen first row contains dataset column names and remaining rows contains dataset values and by using above dataset we will train Reinforcement algorithm.To store or access data from Blockchain we need to develop SMART CONTRACT which will contains functions to STORE and READ data and below is the SMART CONTRACT for Fake News application.

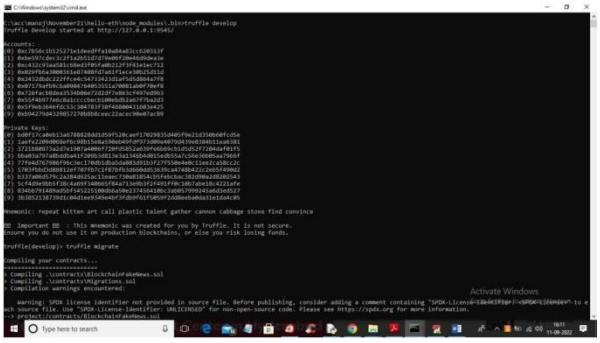


```
pragma solidity >= 0.8.11 <= 0.8.11;

contract BlockchainFakeNews {
    string public signup_details;
    string public publish_news;

    //call this function to save new user details data to Blockchain
    function setSignup(string memory sd) public {
      signup_details = sd;
    }
    //get Signup details
    function getSignup() public view returns (string memory) {
      return signup_details;
    }

    //call this function to save publish news to Blockchain
    function setPublishNews(string memory pt) public {
      publish_news = pt;
    }
    //get publish news details
    function getPublishNews() public view returns (string memory) {
        return publish_news;
    }
```

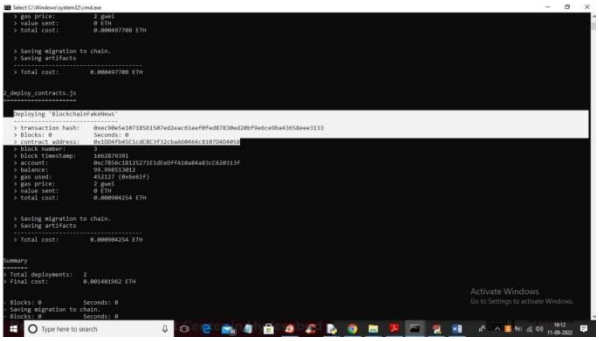
In above screen we have define smart contract functions to store USER & NEWS details and we need to deploy this contract in Blockchain server and for deployment we need to follow below steps

First go inside ‘hello-eth/node_modules/.bin’ folder and then double click on ‘runBlockchain.bat’ file to get

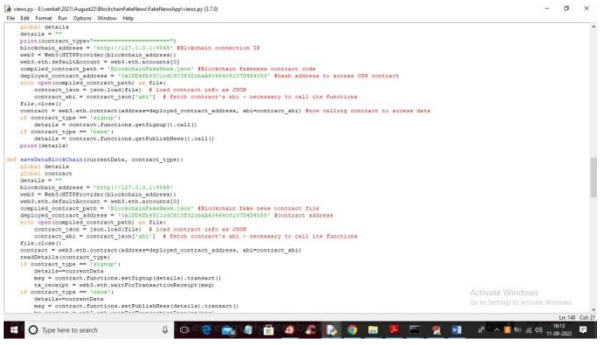
below screen



In above screen Blockchain generated Private keys and default account address and now type command as ‘truffle migrate’ and then press enter key to deploy contract and get below output.



In above screen in white colour text we can see Blockchain Fake News contract deployed and we got contract address also and this address we need to specify this address in python program to access above contract to store and read data from Blockchain.



In above screen read red colour comments to know about how to call Blockchain function to store and read data using Python program

4. ANALYSIS OF THE PROPOSED FRAMEWORK

Blockchain technology can be used to verify the authenticity and integrity of information by creating an immutable record of transactions. However, it alone may not be enough to detect fake news. Natural Language Processing (NLP), on the other hand, is a branch of artificial intelligence that deals with the processing and analysis of natural language data. NLP techniques can be used to analyze the content of news articles and determine their authenticity.

Using only blockchain for fake news detection would involve creating a blockchain-based system where news articles are stored and verified through cryptographic algorithms. While this would provide a reliable record of the publication of a news article, it would not necessarilydetermine the authenticity of its content.On the other hand, using blockchain with NLP would involve combining the two technologies to verify both the publication of news articles and the accuracy of their content. NLP algorithms could be used to analyze the language and sentiment of news articles to determine their authenticity, while blockchain could provide a secure and immutable record of the verification process.

In summary, while blockchain can provide a secure and reliable way to verify the publication of news articles, using NLP in combination with blockchain can help to determine the authenticity of their content.

5. CONCLUSION

Fake news spread is one of the most well-known concentrate on points in present day innovation, inferable from an absence of safety and trust in the veracity of data distributed via virtual entertainment. Here, we offered arrangements and fostered a trust-based engineering for online news sharing utilizing a blend of blockchain and ML techniques. We joined a blockchain structure, shrewd agreements, and a custom agreement calculation that is great for the Evidence of-Authority convention with the support learning approach, a learning-based calculation, to make a powerful dynamic design. In this cycle, virtual entertainment is pivotal. The Evidence of-Authority convention and client approval face a critical test given that the common data stage contains fake news.

REFERENCES

[1] V. P. Miletskiy, D. N. Cherezov, and E. V. Strogetskaya, ‘‘Transformations of professional political communications in the digital society (by the example of the fake news communication strategy),’’ in Proc. Commun. Strategies Digit. Soc. Workshop (ComSDS), 2019, pp. 121–124.

[2] N. R. de Oliveira, D. S. V. Medeiros, and D. M. F. Mattos, ‘‘A sensitive stylistic approach to identify fake news on social networking,’’ IEEE Signal Process. Lett., vol. 27, pp. 1250–1254, 2020.

[3] G. Liu, Y. Wang, and M. Orgun, ‘‘Optimal social trust path selection in complex social networks,’’ in Proc. AAAI Conf. Artif. Intell., vol. 24, 2010, pp. 1391–1398.

[4] M. N. Nikiforos, S. Vergis, A. Stylidou, N. Augoustis, K. L. Kermanidis, and M. Maragoudakis, ‘‘Fake news detection regarding the Hong Kong events from tweets,’’ in Proc. Int. Conf. Artif. Intell. Appl. Innov. Greece: Springer, 2020, pp. 177–186.

[5] A. R. Merryton and G. Augasta, ‘‘A survey on recent advances in machine learning techniques for fake news detection,’’ Test Eng. Manag, vol. 83, pp. 11572–11582, 2020.

[6] X. Dong, U. Victor, S. Chowdhury, and L. Qian, ‘‘Deep two-path semisupervised learning for fake news detection,’’ 2019, arXiv:1906.05659. [Online]. Available: http://arxiv.org/abs/1906.05659

[7] S. Kumar, R. Asthana, S. Upadhyay, N. Upreti, and M. Akbar, ‘‘Fake news detection using deep learning models: A novel approach,’’ Trans. Emerg. Telecommun. Technol., vol. 31, no. 2, p. e3767, Feb. 2020.

[8] P. Bahad, P. Saxena, and R. Kamal, ‘‘Fake news detection using bidirectional LSTM-recurrent neural network,’’ Procedia Comput. Sci., vol. 165, pp. 74–82, Jan. 2019.

[9] G. Sansonetti, F. Gasparetti, G. D’Aniello, and A. Micarelli, ‘‘Unreliable users detection in social media: Deep learning techniques for automatic detection,’’ IEEE Access, vol. 8, pp. 213154–213167, 2020.

[10] M. Mahyoob, J. Algaraady, and M. Alrahaili, ‘‘Linguistic-based detection of fake news in social media,’’ Int. J. English Linguistics, vol. 11, no. 1, p. 99, Nov. 2020.

[11] A. Koirala, ‘‘COVID-19 fake news classification using deep learning,’’ Tech. Rep., 2020. [Online].

Available: https://www.cs.ait.ac.th/xmlui/handle/123456789/981

[12] H. Gill and H. Rojas, ‘‘Chatting in a mobile chamber: Effects of instant messenger use on tolerance toward political misinformation among south Koreans,’’ Asian J. Commun., vol. 30, no. 6, pp. 470–493, Nov. 2020.

[13] J. L. Alves, L. Weitzel, P. Quaresma, C. E. Cardoso, and L. Cunha, ‘‘Brazilian presidential elections in the era of misinformation: A machine learning approach to analyse fake news,’’ in Proc. IberoamericanCongr. Pattern Recognit. Havana, Cuba: Springer, 2019, pp. 72–84.

[14] N. R. de Oliveira, P. S. Pisa, M. A. Lopez, D. S. V. de Medeiros, and D. M. F. Mattos, ‘‘Identifying fake news on social networks based on natural language processing: Trends and challenges,’’ Information, vol. 12, no. 1, p. 38, Jan. 2021.

[15] D. Mouratidis, M. N. Nikiforos, and K. L. Kermanidis, ‘‘Deep learning for fake news detection in a pairwise textual input schema,’’ Computation, vol. 9, no. 2, p. 20, Feb. 2021