International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

COPY RIGHT

IJIEMR Transactions, online available on 07th Sept 2023. Link :http://www.ijiemr.org/downloads.php?vol=Volume-12&issue=Issue 09

10.48047/IJIEMR/V12/ISSUE 09/10

Title **DETECTING BOGUS NEWS USING ARTIFICIAL INTELLIGENCE, MACHINE LEARNING AND BLOCKCHAIN**

Volume 12, ISSUE 09, Pages: 92-100

Paper Authors **Dr.Venkata Kishore Kumar Rejeti, V.Venkata Veera Anjaneyulu, T.Hemanth Eswar Baba, Y.Sateesh, P. Kartheek Babu**

USE THIS BARCODE TO ACCESS YOUR ONLINE PAPER

To Secure Your Paper As Per UGC Guidelines We Are Providing A Electronic Bar Code

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

DETECTING BOGUS NEWS USING ARTIFICIAL INTELLIGENCE, MACHINE LEARNING AND BLOCKCHAIN

[1]Dr.Venkata Kishore Kumar Rejeti, [2]V.Venkata Veera Anjaneyulu, [3]T.Hemanth Eswar Baba, [4]Y.Sateesh, [5]P. Kartheek Babu

[1]Professor, Department of CSE, KKR & KSR Institute of Technology and Sciences
1sheela.softinfo@gmail.com

[2,3,4,5] B.Tech Students, Department of CSE, KKR & KSR Institute of Technology and Sciences
2velpurianji2002@gmail.com, 3thbaba2001@gmail.com, 4yengantisateesh@gmail.com, 5posanikartheekbabu@gmail.com

ABSTRACT: Given the recent developments and advancements in the software engineering field, the interest-based entertainment network is the one of the most important aspects of human existence. This environment has established itself as a popular form for exchanging information and news on all topics as well as daily reports, which is the major period for information and news on all topics as well as daily reports, which is the major period for information collecting and transmission. There are a variety of advantages to this environment, but from another angle, there are a lot of false data and information that lead readers and clients astray while they are looking for the information they need. One of the major problems with this approach us the lack of reliable data and true new insight regarding internet entertainment data. There is a main disadvantage is that the platform is also been used for producing the fake news and users are being miss guided by the information. This all has come with the lack of trustable platform and trustable information prevailing in the other platforms.

To combat this problem, we have created an integrated framework for various blockchain and Natural language processing (NLP) components that applies AI techniques to recognize fake news and better anticipate fake client records and postings. This methodology uses the support learning approach. The decentralized blockchain structure was used, which provides the framework of computerized contents authority verification, to work on this stage with regard to security. More specifically, the goal of this framework is to promote a secure environment for spotting and identifying fake news in online entertainment companies.

KEYWORDS: Fake news, Blockchain, NLP, Reinforcement learning, AI, Security, Media .

I. INTRODUCTION

Fighting false news begins as an unusual problem in informal organizations at the application layer of information and data utilization and then grows into significant and challenging issue in the headway that arises in the political, financial, and discretionary domains. The publishing of false statistics highlights the organization's asset's perpetual inefficiency. We gathered the online entertainment materials from Face book and Twitter, two popular social media sites for exchanging information with a big audience and spreading a tone of daily news and messages on numerous subjects. As a result of the easily accessible support and a contextual inquiry, it also contains the material in its totality and authenticity.

Fig 1: social media

Fig 2: Social Media Network

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

This work uses block chain, NLP, and AI approaches to validate fraudulent users and data. The proposed framework is a defence tactic against the idea of fraudulent information extraction fusing with game elements, to be more precise. Support learning is the learning-based computation that raises the quality of the framework in light of the supplied data. In the case that the data is faulty, the framework bans utilizing comparable data in order to reduce the fraudulent and incorrect data rating. Utilizing Normal Language Processing (NLP) for a thorough text analysis and developing a false news counteraction framework rather than a geographic one in light of the shared components.

Fig 3: Flow of NLP

Setting up financial sources and putting the power convention verification into practice. This interaction is necessary for the system to find bogus client data and records. use the Support Learning method to predict the framework's rate of learning and eliminate false records. In an effort to prevent the transmission of erroneous information, it is necessary to identify the connections between the contents, distinguish between the relative importance, and analyse the organization of the transmitted material.

The types of media in the society have been shown in Table-1, to know about their role and example of each type of media. The domain we are choosing depends individually on the type of media we are choosing to detect the fake news in that platform. These media platforms need to be registered in the blockchain network to get their identification with a unique ID and then their published content will be verified by the Natural language processing (NLP) which is a machine learning technique used for feature extraction from the data given and also use Reinforcement learning to take feedback from every operation the algorithm as gone through.

S.No	Types of media	Role of media	Examples
1	News media	It takes mass media elements and uses them to deliver news to a target audience.	BBC, CNN
2	Social media	It is a digital tool, such as website or an app that allows users to create & share content with public.	Facebook, Twitter, Instagram, Pinterest
3	Web media	It is a form of media that uses audio, text, visuals on web. It is owned by an individual.	Blog Posts, Internet, Podcasts.

Table-1: Types of media

The scope of the fake news in the different areas of the society are taken into consideration while creating this integrated system for detection of fake media and also fake news. They are given as:

Effects on society
Effects on political area
Effects on Individual
Effects on media organizations

The effects mentioned above have widely increased when a fake content or fake news is being circulated in the society. The above Fig.4 show that the prevailing of the fake news in different media platforms in India. According to this Times Now has misreported 8 times after performing the fact checks in the year 2021. The above diagram shows the misreporting of the news by the known news outlets in the year 2021 and the information has been gathered from the website thewire.in, which publishes the headlines and news from day to day.

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

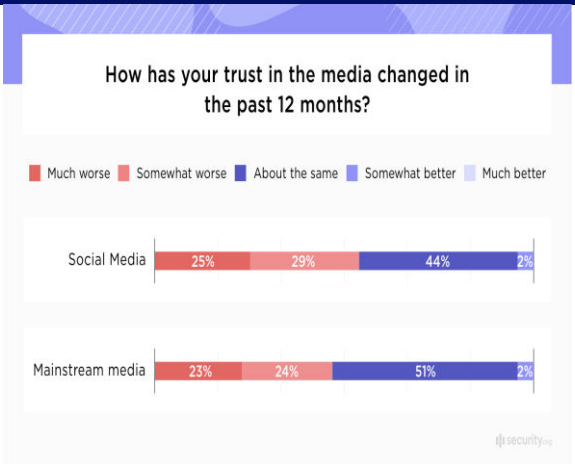

Fig-5: Change of media in past 12 months

Though many people rely on social and news media to get information, public opinion of both sources is dropping. Fig-5 shows how has the news media changed, One-half of Americans hold the media in lower regard than they did just twelve months ago, and one-quarter say their trust has grown "much worse" over that time. Only two percent of adults said their trust improved.

Name of news media	Fake news casted	Published date	Effected by	Demerits
Times of India	Death of Sushant Singh, Fans started an online movement to get him justice alleging that he was coerced to take his life.	June 14, 2020	The false news spreads that, he was murdered.	Shows that the individuals are involved in the case.
The Hindu	Published a study of Filo-virus linked to Corona virus.	October, 2019	Public was panicked due to fear that the virus is spreading.	Fear has taken them before corona did.
TV9 Bharatvarsh	Asked people to give up eating tomatoes claiming an unnamed virus affecting the fruit in Maharashtra.	October, 2019	The society has been affected due to unavailability of the tomatoes.	The tomato farmers have lost their profit which they have invested.
Aaj Tak, Times Now	Aired video of 1962 war memorial as "proof of Galwan".	June, 2020	Soldiers who died during 1962 war & their graves are shown, which panicked the country.	Provoking the clash between India & China.

Table-2: News media platforms and its effects

The table-2 shows how different news channels, which have produced the fake news at different times in year and their effects on the society. The effects may lead to any consequence such as an Inflation or even recession in a country or may lead to political outrage among the parties.

FAKE MEDIA

Fake news is information that is false or misleading yet is reported as news. The destruction of someone or something's reputation or the generations of advertising revenue are frequent objectives of false information has always been shared throughout history, the term "fake news" was first used in the 1890s, a time when spectacular newspaper tales were common. The term, which has no specific definition, is frequently used to describe all false information. High-profile people have also used it describe any negative news that pertains to them. [17] Disinformation is also the deliberate spread of misleading information, and it is commonly produced and spread by hostile foreign actors, especially during election seasons. Stories with sensationalist or click bait headlines without any underlying material are some examples of fake news, are satirical articles that are misconstrued as the genuine thing. Due to the variety of false news sources, researches are beginning to adopt the term "information disorder" since it is more objective and informative.

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

II. LITERATURE REVIEW

According to Arian Balouchestani [1], he used Blockchain with the help of decentralized application and smart contracts can provide a platform in which fake news can be detected through the public participation. [1] Publishing news anonymously and news validation and evaluation by the users.
[2] Adnan Qayyum and Junaid Qadir, in this paper We can tackle the fake news by the blockchain in the new post-Truth world and check the spread [2]. A network analysis approach has been developed for detection and mitigation of fake news. Blockchain by using smart contracts and public keys are mapped to the media organizations. Depending on the proposed paper "Fake news detection in social media using Blockchain" by Shovon Paul & Jubair Islam Joy [3] with a Noval machine learning counterfeit news location technique which has news substance. Breadth first search algorithm is used to explore the dataset. Zonyin shae & Jeffrey J.P Tsai [4] proposed a trusted mechanism built with Blockchain in a distributed manner with the aim of rebuilding the trust relationship. Artificial intelligence is also used for prediction of fake news in this proposed system.

By produced paper of Xishuang Dong & Uboho Victo [5], they have developed a semi-supervised learning technique to detect the fake news which also uses a conventional neural network (CNN). Through this the fast propagation of the illegitimate news can be stopped. According to Aditya Chokshi [6], who published on "Deep learning and natural language processing for fake news detection" where N-grams language models are being used for processing of data using a bunch of input features for indexing the data. LSTMs are being used in the hidden layer. As provided by Nicollas R.de Oliveira [7] in his paper "Identifying fake news on social media platforms based on Natural language processing" gives us a way to understand how the fake news can be collected and training on the obtained data is done to get the classification model to predict the fake news.

Depending on the proposed paper "Detecting COVID-19 Fake news using Deep-Learning" by Anmol Tukrel and Avalon Wolfe [8]. The author builds a baseline LSTM model as it was the best performing model to store and detect the fake news from the given dataset which contains 300 real news articles and 300 fake news articles. The paper published by Giuseppe Sansonnetti & Fabio Gasparetti [9] on "Unreliable users' detection in social media" helps in detecting the fake users who are producing or publishing the unreliable content on social platforms.

This article gives a better perspective to identify who are publishing the fake news rather than also finding the fake news. Mohammad Mahyoob & Jeehan Algaraady

III. ARCHITECTURE DIAGRAM

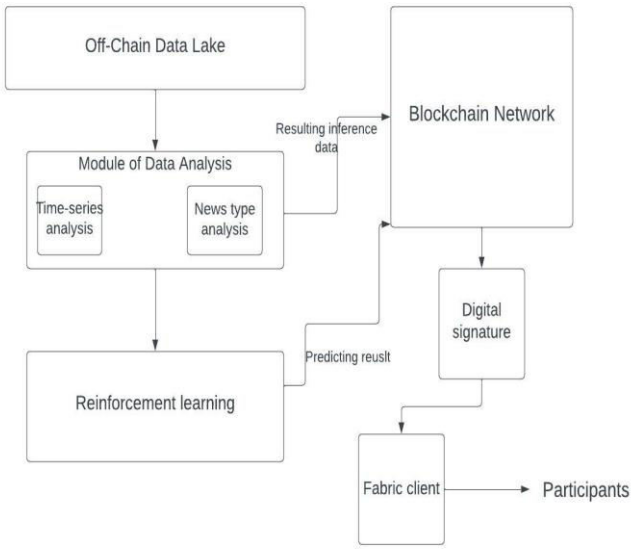

Fig-8: Architectural diagram of Proposed system

The Fig-8 shows the flow of the data from the initial state of entering into the application to the final output of detecting the news.
The individual components as in the Fig-8 working is as follows:

1) Off-chain data Lake:

It contains the details related to the user, News sharing schedule, History of news records and provides data analysis. The user is allowed to register in blockchain where that person is provided with a Unique ID and Publisher ID. The user can publish the news and also validate the news published by others in the supply chain. This also contains a record of the already published news in the history to access whenever needed. The data is then

transferred to the next module where the data is analysed.

2) Module for data analysis:

This module will analyse the obtained data from the user by using different techniques such as Time-series Analysis, News Type Analysis where Time series analysis is a specific way of analysing a sequence of data points collected over an interval of time and analyst record data points at consistent intervals over a set period of time rather than recording the data points intermittently or randomly. News Type analysis gathers the information about the headline of the news article and include that headline into the hypothesis we are building, this also take the voice of the article, tone and rhetoric structure of the data we have taken. It also examines the structure of the news article to see how much of a personal opinion is included in this article.

3) Analysis Module:

This is the module where the analysis and prediction depending on the news we have gathered from the user and analysed data taken from the data analysis module. The Reinforcement learning model is being used to detect and output the prediction of the fake news and real news. This module also contains Natural language processing, where we can identify the text and convert it into the machine understandable language, instead it also gathers the context of the text we have taken as input and make it to computer understandable which is very useful to train the model for better prediction.

4) Blockchain network

This network is a decentralized network where no central authority takes control. This network provides to store data in a distributed ledger. This takes input from the module of data analysis and analysis module such as Inference data and prediction result from the all the module and store them. This network contains smart contracts and peer machines. Digital signature method is used for securing the data in a blockchain network where only authenticated persons only can access the network.

IV. EXISTING SYSTEM

Recent technological advancements and the employment of programmes in daily life have created a problem as a result of the broadcasting of bad and needless situations on social media. One of the social media sites mentioned above, Twitter, has a sizable user base. Every day, millions of tweets with a wide range of terminology and themes are shared. The blockchain technology and machine learning are crucial in the battle against the spread of misleading information. There are several deep learning algorithms for detection that are based on vast amounts of data. Convolution neural networks (CNN) were also examined for the classification of text and images. The combination of three components—the fake news detector, reinforcement learning, and annotator—was proposed as a framework for false news identification. By choosing high-quality samples and eliminating the weak labels from the material, this procedure has been used to detect fake news.

V. PROPOSED SYSTEM

The proposed model is an integrates system of various technologies and computer science tools used for detecting, identifying and then finally giving the output to the individuals whether the news published is fake or real. The detection of fake news utilizing a mix of Natural Language Processing, Reinforcement Learning, and Blockchain are being used and these have been explained in length in this section. The system's two key components are its learning phases and its suggested data security strategy.

The data security strategy which is also a key component is provided by the Blockchain which has a wide range of applications in to developing a secure platform which cannot be modified by any individual. This is a con for the developing system and providing security at its best.

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

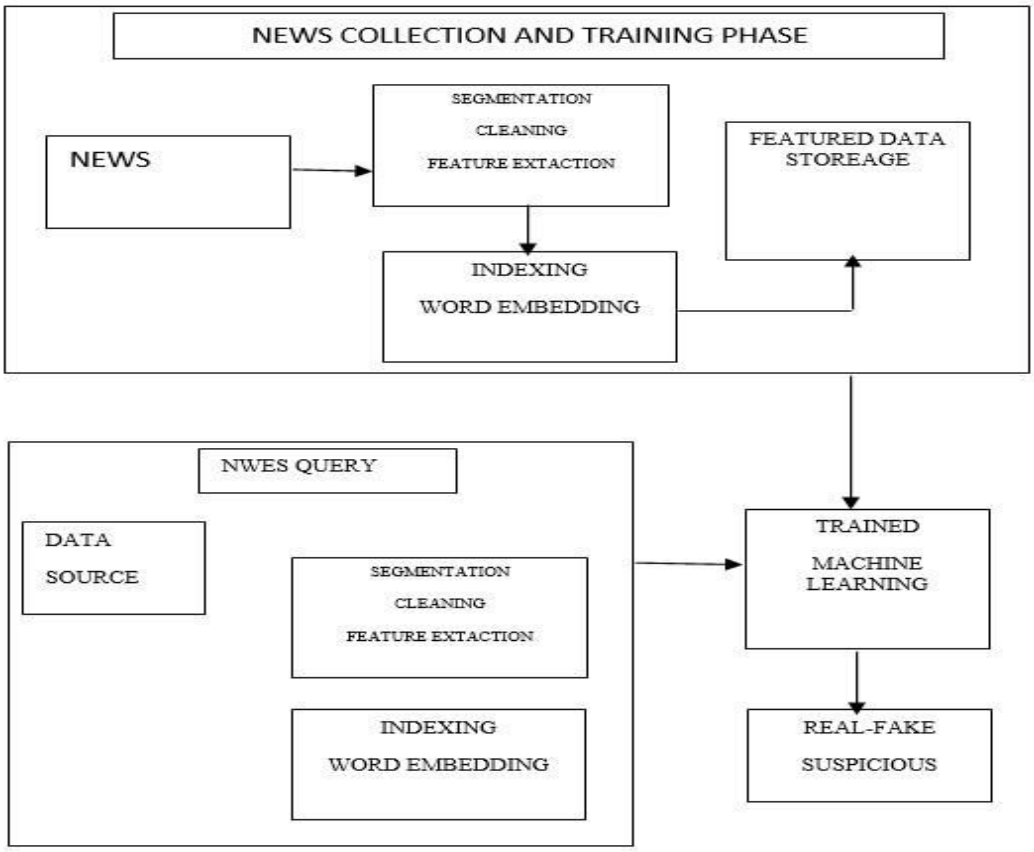

Fig-10: Detection of Fake news using NLP

FAKE NEWS DETECTION USING NATURAL LANGUAGE PROCESSING

Data cleaning, segmentation, stop words, feature extraction, word indexing, and embedding are important steps in the data preparation process. NLP is one among them. Before beginning any processing on the data, data preparation is responsible for cleaning it. The data is then transformed into vectors and saved into a database after feature extraction in the next phase. A query is sent to the data source on a recurring basis using the data that has been collected for feature extraction to look for the relevant news on the Internet. The feature extraction module may group the news into a list based on distance and query after comparing its contents to the news into a list based on the distance and query after comparing its contents to the news storage to determine how similar they are.

FAKE NEWS DETECTION USING REINFORCEMENT LEARNING

Deep reinforcement learning combines deep learning with reinforcement learning to make decisions from unstructured data. Most of the news that circulates on social media has to be verified in order to be believed. The key justification for include this algorithm in the proposed system is its learning-based element, which is a step in the right direction toward enhancing the efficacy of fake news identification. Markov

Decision Process is used to create problems (MDP). In order to take action and get the prize for the state that follows, each timestamp agent must be physically present in that state. The agent chooses a learning approach based on growing incentive returns that are assembled into a news list.

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

FAKE NEWS DETECTION USING BLOCKCHAIN.

The blockchain idea was initially popularized by the Bit coin platform since every node may participate in the network and be a miner. Proof-of-Work (PoW), Proof-of-Authority (PoA), and Proof-of-Stake are a few examples of this (PoS). One of the newest additions to the Byzantine Fault Tolerant (BFT) algorithm is PoA. The chosen party proposes the next algorithmic block and is in charge of mining. PoA offers a speedier communications technique and performs better than BFT. The network assumptions that underlie PoA implementation disclose what is most advantageous for system preservation. The correspondence identity is connected to the platform's validator identification, such as the credibility scorer. The scenario offered for detecting fake news on a blockchain is based on PoA with a high transaction rate. The credibility score of the consensus was applied.

A crucial step in the organization of news is the identification of the organizations with the ability to register and have them take part in transaction and publishing validation. As a news organization, we don't rely on the dependability and legitimacy of news sources in the actual world to grow the network. Some organizations, like CNN and the BBC, require an application before they may register on the blockchain. Certain information, such precise figures or documents, is needed for data authenticity and news organization accreditation. We will check for correctness before granting permission to use this content on TV, in newspapers, or on the radio. the next level of integrating news registrations into a blockchain uses smart contract rules to verify node authentication. A record of the entire process is stored on the blockchain. A crucial component of the suggested system is Proof-of-Authority (PoA), a consensus technique for the identification of fake news. The following news outlet may request publishing once you upload the news and seek node authentication.

The publication has an impact on where the credibility score is placed. During this phase, some nodes adopt the role of validators to vouch for the accuracy of the news and the transaction. More specifically, the transaction enters the validation step when the news is submitted for verification. Each participant is given a credibility score based on their behaviour during this process. Each member typically has a primary key and credit score. The person with the greatest primary credibility score will be awarded the individual ID after registering with the blockchain network.

PERFORMANCE EVALUATION OF BLOCKCHAIN FRAMEWORK

In this part, the performance of the blockchain network is assessed. To assess this system's efficiency, we used the PBFT and RAFT consensus algorithms, PBFT provides Byzantine fault tolerance, whereas Raft only supports crash fault tolerance. For nodes of the bounded type, the PBFT offers Byzantine fault tolerance (BFT) with minimal latency and high transaction rates. The PBFT method has issues with poor scalability and huge transmission records for
consensus messages. The blockchain network being utilized has permissions. The primary justification for utilizing PBFT in a secure environment is that it is message-based. Because of its high throughput and low latency, Proof of Authority is a popular technique in permission blockchain. Use of the RAFT approach results in low TL and high TT.

VI. RESULT

The output of this model gives us an ultimate result whether the given news is bogus or fake in the real world. This helps in reduction of illegitimate content in the society and spreading of this bogus content. The result of this model is evaluated based on the accuracy, precision, recall and F-score of the model.

Precision is defined as follows:

$$\text{Precision} = TP / (TP+FP)$$

Recall is calculated as the ratio between the numbers of positive samples correctly classified as positive to the total number of positive samples.

$$\text{Recall} = TP/ (TP+FN)$$

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

Accuracy is one metric for evaluating classification models. It is the fraction of predictions the model got right or correct.
F-score is the harmonic mean of a system's precision and recall values.

F-score = 2 * [(precision * recall) / (precision + recall)]
If the F-score is greater than 0.9 then the classification if very good.
As we have the model, we have calculated the different parameters of the different algorithms.

Algorithm	Accuracy	Precision	Recall	F1-score
NLP	89.21	91.97	87.71	87.32
RL	90.18	90.14	92.4	91.34
Blockchain	93.74	91.5	94.4	92.97

Table-3: Analysis of algorithms

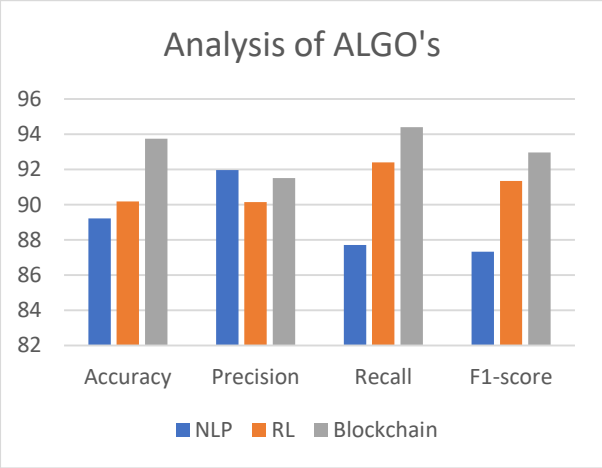

Fig-11: Algorithm analysis of Proposed system

The above Fig-11 is a graph that shows the graphical representation of the data given in the above table-3 which has the data that evaluates the performance of the proposed model. The Y-axis contains the data related to the performance of each model and X-axis contains the tuples i.e., Accuracy, Precision, Recall, F1-score. Each colour shows different algorithms that are being used in the proposed system.

One skew aspect of the data values in the acquired dataset is the time series frequency of the shared news. To prevent skewers and produce a consistent range of data values, data normalization is crucial. Among other methods for normalizing data, decimal scaling normalization, Z-score normalization, and min-max normalization are used. The above data is normalized using the min-max normalization and this method scales the characteristics between zero and one in the order to normalize them and provide data features in uniform mode.

VII. CONCLUSION

The spread of fake news, which arises from a lack of faith in the veracity of the material shared on social media, is one of the most well-known study issues in modern technology. In this post, we discussed how to use blockchain technology and machine learning techniques to create a trust-based architecture for news that is shared online. We used a learning-based method called reinforcement learning to construct a trustworthy decision-making architecture that is suitable for the Proof-of-Authority protocol. We combined it with a unique consensus algorithm, blockchain architecture, and smart contracts. In this process, social media is crucial. It is crucial to look into and enhance the Proof-of-Authority protocol and user validation since the shared information platform promotes erroneous information.

VIII. REFERENCES

[1] A. Balouchestani, M. Mahdavi, Y. Hallaj, and D. Javdani, ''SANUB: A new method for sharing and analyzing news using blockchain,'' in Proc. 16th Int. Conf. Inf. Secur. Cryptol. (ISCISC), Aug. 2019, pp. 139–143.
[2] A. Qayyum, J. Qadir, M. U. Janjua, and F. Sher, ''Using blockchain to rein in the new post-truth world and check the spread of fake news,'' IT Prof., vol. 21, no. 4, pp. 16–24, Jul. 2019.
[3] S. Paul, J. I. Joy, S. Sarker, H. Shakib, S. Ahmed, and A. K. Das, ''Fake news detection in social media using blockchain,'' in Proc. 7th Int. Conf. Smart Comput. Commun. (ICSCC), Jun. 2019, pp. 1–5.
[4] Z. Shae and J. Tsai, ''AI blockchain platform for trusting news,'' in Proc. IEEE 39th Int. Conf. Distrib. Comput. Syst. (ICDCS), Jul. 2019, pp. 1610–1619

[5] 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

[6] A. Chokshi and R. Mathew, ‘‘Deep learning and natural language processing for fake news detection: A survey,’’ SSRN Electron. J., Jan. 2021. [Online]. Available: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3769884

[7]] 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

[8] 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

[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.