Journal of Ambient Intelligence and Humanized Computing (2024) 15:435–449
https://doi.org/10.1007/s12652-022-03900-2

ORIGINAL RESEARCH

Blockchain-based rumor detection approach for COVID-19

Poonam Rani[1] · **Vibha Jain**[1] · **Jyoti Shokeen**[2] · **Arnav Balyan**[1]

Received: 29 August 2021 / Accepted: 4 May 2022 / Published online: 20 May 2022

Abstract
The ubiquity of handheld devices and easy access to the Internet help users get easy and quick updates from social media. Generally, people share information with their friends and groups without inspecting the posts' veracity, which causes false information propagation in the network. Moreover, detecting false news and rumors in such a massive load of unstructured information is a very tedious task. Results, many literature papers explored different machine learning and deep learning approaches to detect the presence of rumors on social media networks. Although detection of misleading news and rumors is not sufficient, therefore, we have proposed a model for the detection and prevention of transmitted rumors in this paper. In this paper, we use blockchain technology to verify the credibility of information and design a framework with four layers: network layer, blockchain layer, machine layer, and device layer, to prevent the propagation of rumors in the network. We also use deep learning techniques to identify the anomalies in the network. The Bi-directional Long Short Term Memory (Bi-LSTM) model is used to prevent the introduction of new rumors by continuously monitoring incoming messages in the network. The experimental results demonstrate that the proposed Bi-LSTM model outperforms state-of-the-art machine learning methods and recent baseline work. Performance is compared over different metrics such as accuracy, precision, recall, f1-score, and specificity. Experiment results show that our Bi-LSTM model outperforms all the other approaches and achieved 99.63 % accuracy. Additionally, the probability of incorrect detection is significantly low with only 0.13% false positive.

Keywords COVID-19 · Blockchain · LSTM · Rumor

1 Introduction

The unforeseeable COVID-19 pandemic situation led the world to a whole bunch of new challenges. Government and non-government organizations have taken several steps to overcome the problem, such as social distancing and partial or complete lockdown. Millions of people around the globe are forced to stay indoors by adjusting their current lifestyle and continuing their work from home. During the COVID-19 pandemic, millions of people have gone online for entertainment, education, etc. Consequently, according to preliminary statistics, total internet usage has surged between 50% and 70%, compared to pre-lockdown scenarios, whereas some areas also witnessed a 100% surge in internet usage (Pandey and Pal 2020). The pandemic time has helped more people come online because it has pushed humans to seek education and entertainment on the internet (Subudhi and Palai 2020).

On the one hand, online life improves our living style by maintaining social distancing. On the other hand, it brings new risks to human life. Rumors are elements deeply implanted in human communication and interaction. A rumor is defined as the piece of information that is either false or unverified and disseminates rapidly. In other words, a rumor is a fast-checkable and controversial statement. A rumor can become extremely dangerous when exchanging parties are unaware of the credibility of the information. False news often spreads quickly through social media sites

✉ Poonam Rani
poonam.rani@nsut.ac.in

Vibha Jain
vibha.jain.cs19@nsut.ac.in

Jyoti Shokeen
jyotishokeen12@gmail.com

Arnav Balyan
arnavb.co18@nsut.ac.in

[1] Department of Computer Engineering, Netaji Subhas University of Technology, Delhi, India

[2] Department of Computer Science and Engineering, University Institute of Engineering and Technology, Maharshi Dayanand University, Rohtak, Haryana, India

such as Twitter, YouTube, and Facebook because social networks are the platforms that are easily accessible by people these days. Social networks play an important role in our lives in all aspects. Social networks are highly complex and uncertain. Moreover, the information floating in the social networks is highly dynamic, i.e., information changes from node to node (Rani et al. 2019b, 2021e). The rapid change in information from node to node introduces fake information or rumor spreading in networks. Sometimes the information and its meaning get changed during information propagation. The gradual change of information in social networks is one of the potential challenging tasks that need to be tackled efficiently with accuracy. In this paper, we address this issue by integrating deep learning and blockchain technologies. We have covered and tackled the potential issues of social networks in our papers (Rani et al. 2018, 2019a, c, 2021a). Therefore, recognizing the origin of false news can help in reducing the damage further. Widespread rumors can form public opinion in countries (Jin et al. 2017), impact financial markets (Oberlechner and Hocking 2004), cause panic and even break out wars (Del Vicario et al. 2016). In the pandemic, misinformation about "cures" for COVID-19 and rumors about best practices have proven dangerous and even led to multiple deaths. Pertaining to the pandemic, multitude of misinformation and conspiracy theories have been discovered, like the virus being non-existent and being a mere media facade, it being a bio-warfare weapon developed in China (Uchejeso et al. 2021), the virus being non-contagious (Toshida and Jagruti 2020) or incorrect medicines. There has also been misinformation saying wearing masks increases the spread of the virus by activating it.

Some bad actors on the Internet tend to spread rumors, and there is no one to ratify online content. Misinformation and rumors are spread quickly on the Internet due to the high connectivity. Fake news subjects innocent and naive people to a great deal of danger. The effect of rumors/fake news on a person can be both immediate and long-lasting. In most cases, rumors can rile up one's emotions and even change a person's mental state. Depending on the strength of one's feelings, the story, and the reaction it gave someone, it can stick in one's head, even after one finds out it's false. Rumors have proven to cause long-term physical and mental health issues, including but not limited to clinical depression, post-traumatic stress disorder, panic attacks, guilt, and even suicide.

It is essential to detect and stop the spread of false news disguised as rumors. Machine learning is a field of study which involves making decisions and classifying results based on learning and identifying patterns from data with minimal human intervention. Powerful machine learning algorithms clubbed with correct data sources, and the proper simulation can help construct models (Al-Asadi and Tasdemır 2022) Such models can be scaled to detect and stop the spread of unverified news, rumors, and misinformation in real-world applications suffering from this problem, like social media networks. Machine learning has been a prevalent and paramount tool for classification and decision-making of all kinds. The role of machine learning in rumor/false news detection has received significant attention in recent years (Manzoor et al. 2019). This problem has been approached through social network analysis, data mining, and natural language processing. Literature indicates that classic machine learning models like decision trees and Support Vector Machine (SVM) have been used to get good accuracy for fake news classification (Thakur et al. 2018; Pathak et al. 2020). Researchers have also used convolutional neural network (CNN) and LSTM to attempt to classify rumors and legitimate information in deep learning. However, the scope of this work in the domain of deep learning remains limited.

In this paper, we propose a social media network equipped with blockchain and robust deep learning models for swiftly verifying the credibility of information and design measures to stop the propagation of misinformation and rumors in the network. Deep learning models help find anomalies in information, whereas blockchain aims to find the source of that anomaly to minimize the spread of rumors in the future. The major contributions of our paper are as follows:

- First, we have considered a biased graph-based social media network with N number of nodes in which M nodes are malicious and can spread rumors to the network.
- A deep learning-based Bi-LSTM model is introduced to find anomalies in generated information. Designed social media network is integrated with blockchain to verify the credibility of each transmitted information and find the source of rumor to prevent the growth in future.
- We compare the simulation results of the proposed work with other state-of-art machine learning and recent baseline works to measure the effectiveness of the proposed model.
- Performance is compared over different metrics such as accuracy, precision, recall, f1-score, and specificity. Experiment results show that our Bi-LSTM model outperforms all the other approaches and achieved 99.63 % accuracy. Additionally, the probability of incorrect detection is significantly low with only 0.13 % false positive. r

The rest of the paper is divided into the following sections. Section 2 includes the survey of related works. Next, Sect. 3 describes some preliminaries used in the paper. Section 4 explains the working of the proposed model, including the architecture of the designed social media network, the working mechanism of blockchain, and the implementation of several machine learning algorithms for rumor detection.

Section 5 presents the description of dataset, hyperparameters tuning, performance metrics and experimental results. Section 6 presents the comparison of the proposed work with the existing approaches. At last, we conclude our work in Sect. 7.

2 Related works

Automatic recognition of rumors and blocking their propagation are challenging tasks in designing trusted social media sites. Several techniques have been devised in the literature to detect and block the misleading information on social media sites automatically. However, only a few of them used blockchain to resolve the security issues of information on social media. Most of the existing techniques treat rumor detection as a classification problem. A comprehensive review of different approaches for source detection of rumors is given by Shelke and Attar (2019). Li et al. (2019) discussed various machine learning-based techniques for rumor detection in social media. They also provided a list of datasets used by different rumor detection techniques. They divided the approaches of rumor detection into content-based, user information-based, propagation path-based, and network-based approaches. The content-based approaches include textual content, or visual content or both. The authors also discussed some future research areas in rumor detection. Recently, Al-Asadi and Tasdemır (2022) investigates and reviews the use of various artificial intelligent tools in fake news detection. Their study concluded that machine learning and deep learning are the widely-adopted tools for classifying fake news.

Jin et al. (2017) attempted to gain insights from the online behavior of rumors in American politics. They analyzed rumors from 8 million tweets of supporters of two presidential candidates, namely, Donald Trump and Hillary Clinton. The authors gathered a variety of datasets to evaluate the rumors. The reason for evaluating rumors was to answer the main concerns about rumors in the 2016 US presidential elections. The significant points of concern were determining what rumors were posted, who posted such rumors, when the rumors were posted and which side of the candidate's followers posted the most rumors. For this, they proposed a text-matching approach and matched it against verified rumor articles.

Diffusion models are the widely used models for source detection of rumors in social networks. These models mine the data and answer the questions about where and when the piece of information originated in the network and how fast it disseminates (Guille et al. 2013). The epidemic model is a model that explains information spread in the network. There exist four phases in an epidemic model: Susceptible-Infected (SI), Susceptible-Infected-Susceptible (SIS), Susceptible-Infected-Recovered (SIR) and Susceptible-Infected-Recovered-Susceptible (SIRS). Zhou et al. (2019) employ the Susceptible (S) - Exposed (E) - Infected (I) - Recovered (R) model to detect the actual source of rumor. They take the snapshots of the network to observe the simulation of state transitions of nodes in the network. It is assumed that each node exists in any of the above four states in the network. Initially, all nodes are in S state and only one node turns into I state, and such a node of I state is called a rumor node. The infected node sends the rumors to susceptible nodes. The susceptible nodes will turn into exposed nodes with some probability. If any exposed node trusts the rumor message, then the exposed node will turn into an infected node with some probability. If the exposed node does not trust the rumor, it will change its state to the recovered state.

An increasing usage of blockchain as a secure system in networks can be seen in recent years (Rani et al. 2020, 2021d; Khanna et al. 2021). Shae and Tsai (2019) harnessed the idea of using blockchain in solving the forged news crisis. It is challenging to assess the credibility of information on social media sites due to the fast proliferation of information on social media. Gao and Gao (2020) proposed a blockchain-based model for rumor prevention in social media networks. They used the voting rule and rumor detection system to identify the false messages in the network. Based on the influence of nodes in the network, this model divides the nodes into important and less important nodes in social media networks. However, the authors did not mention the consensus mechanism followed in designing the model. Recently, Połap et al. (2020) used blockchain and federated learning to propose an Internet of Medical Things (IoMT) architecture. They used CNN as a classifier to process x-ray images of patients. Połap et al. (2021) also uses blockchain technology and threaded federated learning in proposing an IoMT architecture. Their architecture, which is based on multi-agent system, uses blockchain to share and protect patients' private data. Further, Dibaei et al. (2021) attempted to integrate machine learning and blockchain technologies to secure vehicular networks.

Yang et al. (2018) proposed a unified fake news detection model to analyze the image and textual news data. They implemented CNN model on the image and textual data on the above dataset and obtained an accuracy of 92.10%. If there are m news articles, then the authors have represented the textual and image data as a set of tuples $A = \{A_i^I, A_i^T\}_i^m$ where i is any news article, T represents the text data, and I represents the image data. They denote the real news by [1, 0] label and fake news by [0, 1] label. They extract a set of latent features from the tuple-set.The objective function of their fake news detection model was to construct a model $f : \{X_i^I, X_i^T\}_i^m \in X \rightarrow Y$, where X_i^I and X_i^T represent the latent features for images and text data, respectively that

are extracted to derive the labels from news articles. However, the use of LSTM could have trained the model much faster. Further, Tida et al. (2022) applied the Bidirectional Encoder Representations from Transformers (BERT) model using the transfer learning approach on three datasets and achieved the highest accuracy of 97%. O'Brien et al. (2018) uncovered the black box-based approach of deep learning by capturing the input words significant for classification. They studied a fake news detector that achieved an accuracy of 93.50%. Other studies (Ghanem et al. 2018; Singh et al. 2017; Ahmed 2017; Ruchansky et al. 2017) have also used techniques like linguistic analysis, n-gram models, and word embeddings to classify fake news.

Alsaeedi and Al-Sarem (2020) devised a CNN-based deep learning model for rumor detection on Twitter. However, their model might not be capable of detecting rumors online. Recently, Raza and Ding (2022) also proposed a fake news detection model intending to resolve the issue of unavailability of labelled data in the training detection model. They also used the transformer approach similar to Tida et al. (2022) in which the encoder is used to learn the representations of fake news, and the decoder predicts the future labels. However, their work has certain limitations, namely, domain-level error analysis, use of ground-truth labels, weak supervision of deep neural networks, and less quantity of user profile data. A summary of the related works with their limitations is given in Table 1.

3 Preliminaries

In this section, we first explain the social media network with mathematical representation. Next, we define the concept of blockchain and the consensus algorithm to used in the proposed work.

3.1 Social media network

With the evolution of the Internet, new social media networks have popped up. Social media networks play a vital role in humans communication and human interaction. The information sharing in social media is often informal and shared on a personal level with other users (Osatuyi 2013). Researchers have tried to analyze user behavior and the nature of these networks, which have led to exciting patterns. Social networks also have acted as an interface for people with similar interests to interact and share information. They are the platforms to exchange opinions, knowledge, and recommendations that might not be true often. In research, this has given birth to "trust," which acts as a degree of confidence a user/node has in its neighbors about their genuine information. Literature has used this factor and simulated social networks for various reasons (Shokeen et al. 2021; Shokeen and Rana 2021). This is based on the assumption that nodes are connected to their neighbors based on similar

Table 1 Summary of related works

Authors	Main work	Technique	Parameters	Limitations
Jin et al. (2017)	Online rumor behavior	Text-based	Textual data	No comparison with previous studies
Yang et al. (2018)	CNN	Image and textual data	Slow	
O'Brien et al. (2018)	Fake news detection	Deep learning, CNN	Language patterns	No comparison with previous studies
Shae and Tsai (2019)	Fake news detection and prevention	AI, Blockchain	Smart contracts, News rooms	No implementation performed
Gao and Gao (2020)	Rumor prevention	Voting rule, Blockchain	Private blockchain	No consensus mechanism
Alsaeedi and Al-Sarem (2020)	Rumor detection	CNN-based, deep learning	Layers	Not capable of detecting rumors online
Polap et al. (2020)	Internet of Medical Things	Blockchain, CNN	Federated learning, Medical data	No consensus mechanism provided
Polap et al. (2021)	Internet of Medical Things	Blockchain, Machine learning	Threaded federated learning, Multi-agent systems, Security	Internet connection problem
Dibaei et al. (2021)	Vehicular networks	Blockchain, Machine learning	Security	–
Tida et al. (2022)	Fake news detection	BERT, Transfer learning	Encoders, decoders	–
Raza and Ding (2022)	Fake news detection	Transformer approach	Encoders, decoders, social context, news content	Domain level error analysis, use of ground-truth labels, weak supervision of deep neural networks, and less quantity of user profile data

interests. Information is then propagated using the trust factor (Shokeen and Rana 2021).

We consider a graph-based network $GN = (N, C)$ where N are the nodes and C are the connections connecting the nodes. Each node has a degree D associated with it, which denotes the number of connections the node has with its neighbors. The nodes follow the degree distribution scheme as given in Eq. (1) as follows:

$$F(k) = e^{-kc} * \left(\left(kc^k\right)/k!\right) \tag{1}$$

where F is a function of k denoting the probability of the n^{th} node having k degree. It is assumed that the nature of all nodes is unbiased, i.e., the nodes may or may not try to transmit false information over the network. Likewise, a message can belong to two categories: correct message and false message (rumor). The network follows an epidemic routing protocol. When two neighbor nodes interact, the node checks missing information from its counterpart. It transmits the missing data to the other nodes and stores the complete information in the blockchain as a transaction. This process repeats for all neighbors of a host node in a single transmission attempt until the message reaches almost all the other nodes.

3.2 System dynamics and metrics

The N node network is initialized by M malicious nodes, which intend to introduce rumors to the network and spread them subsequently. The remaining $(N - M)$ nodes are unbiased and may or may not propagate a message that they receive. These nodes send messages having real information. We consider the network state after t time units have elapsed. As time elapses, rumors already present and malicious nodes in the network should decrease. We call the state of the network at tth timestamp as $GN(t) = (N, C)$. Metrics like the percentage of rumors remaining in the network, percentage of malicious nodes, etc., are representative enough to assess and evaluate the system's efficiency.

3.3 Blockchain

A blockchain is a set of interconnected blocks in a linear order which holds the complete history of previous transactions and node interactions. The information in blockchain is immutable in nature due to a block's property of interconnection through hash values. The primary usage of blockchain is to ensure that the information circulating in the network is immutable and correct. This is dealt with using machine learning and judicious usage of the network resources. Figure 1 illustrates a typical blockchain structure. We consider a block to be a collection of the following entities: parent block hash, data hash, timestamp, version, and nonce.

The use of blockchain to make data secure and valid in decentralized networks has been investigated in the literature. The design and architecture of blockchain ensure that information cannot be modified and the data remain valid. Blockchain has also been combined with the Internet of Things (IoT) applications (Rani et al. 2020, 2021d). Another use case of blockchain has been in supply chain management (Rani et al. 2021b). Further, blockchain is popular in cyber security fields, preventing identity thefts and securing online transactions (Rani et al. 2021c).

Consensus algorithms aim to verify and validate transactions before storing them permanently on the distributed ledger. We use properties of two widely used consensus algorithms, Proof of Work (PoW) (Vukolić 2015) and Proof of Stake (PoS) (Li et al. 2017). In PoW, nodes have to perform a computationally expensive mathematical task to

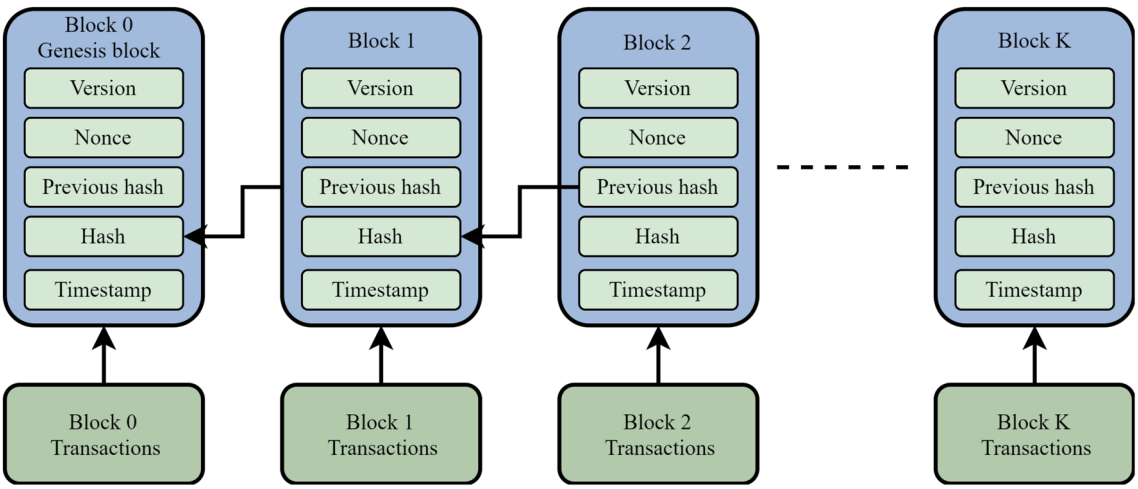

Fig. 1 Blockchain

establish their credibility. However, this becomes impractical in closed networks using private blockchain due to its heavy computational requirements. In PoS, a node aims to stake its peer's economic share within the network. However, this bypasses the expensive computation and raises the issue of lack of randomization. The proposed approach uses a joint PoW-PoS strategy to keep the heavy computations in check while maintaining enough randomization for security.

A malicious node is bound to make wrong transactions and transmit fake messages repeatedly. Therefore, to protect the system's security and resources, each node has a credibility rating, which is regulated based on the transaction's legitimacy. This rating is adjusted using Eq. (2) where $\lambda \in [0, 1]$ is the credibility constant, and $\omega \in [1, 1]$ is a penalizing factor used to increase or decrease the credibility rating of any node.

$$C(X) = \lambda(C(X))old + (1 - \lambda)\omega \tag{2}$$

If this credibility rating goes below a predetermined threshold (0.25 in our case), then the node is not allowed to transmit messages over the network. On initializing the network, we set each node's rating to 0.4.

4 Proposed work

This section explains the details of the proposed mechanism for rumor identification and source detection. The main aim of the paper is to verify the credibility of digital content and fake news and prevent any scams that might be running on the network via machine learning and blockchain. We consider a graph-based social network where multiple connected nodes can transmit information to one another. Further, we exploit blockchain to ensure the credibility of nodes in the network and perform a comparative analysis of various machine learning models for building an optimal model for rumor detection and news verification.

The misinformation about COVID-19 has proven extremely dangerous, ranging from the widespread misinformation about injecting bleach to kill the virus to fake news suggesting that wearing a mask increases the risk of "activating" the virus. Given the devastating impact and loss of life from fake news itself, it is imperative to find a solution that can swiftly verify the credibility of information. The distributed peer-to-peer nature of the network allows for a vast scan of digital content. It provides users with a degree of confidence about the credibility and trustability of the content they receive. Further, using a multi-layer architecture including a social media network, advanced artificial intelligence models and blockchain results in a highly robust and efficient solution for detecting and stopping the propagation of rumors and fake news.

The system consists of the following layers stacked up as illustrated in Fig. 2. As the data passes through each layer, it gets verified and validated at each step. Each layer acts as a filter, validating the transmission from different security points. The system consists of a machine learning layer (MLL) for rumor detection, a blockchain layer for immutable messages, and a network layer to act as medium for transmission and connection. The designed system follows the following infrastructure: When a node introduces a message in the network, MLL ensures that misinformation does not enter the network from the source node itself. This is done using a Bi-LSTM trained on a large news dataset containing both real news and rumors. Further, blockchain ensures that information is not edited/changed by any party while in transit. Ultimately, it ensures that true information enters the network first, and this true nature is maintained via blockchain until it reaches the destination.

Figure 3 illustrates the proposed architecture. A message initially gets introduced to the network using the device layer. Next, the machine learning layer is used to check the legitimacy of the message. Once approved, the message gets converted to a block within the blockchain layer. Finally, it propagates throughout the network using the network layer.

4.1 Pre-processing

The dataset used consists of around 50,000 documents (consisting of raw text data of various news articles). The data is labeled and consists of two classes (real and fake). Each class consists of more than 22,000 news text articles. We follow the following steps to create effective classifiers, as shown in Fig. 4. Pre-processing is a crucial part of any machine learning method by changing the data and making it machine-readable. We pre-process the dataset by removing

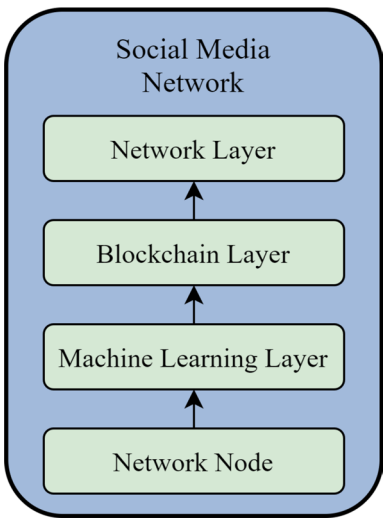

Fig. 2 Message flow in network

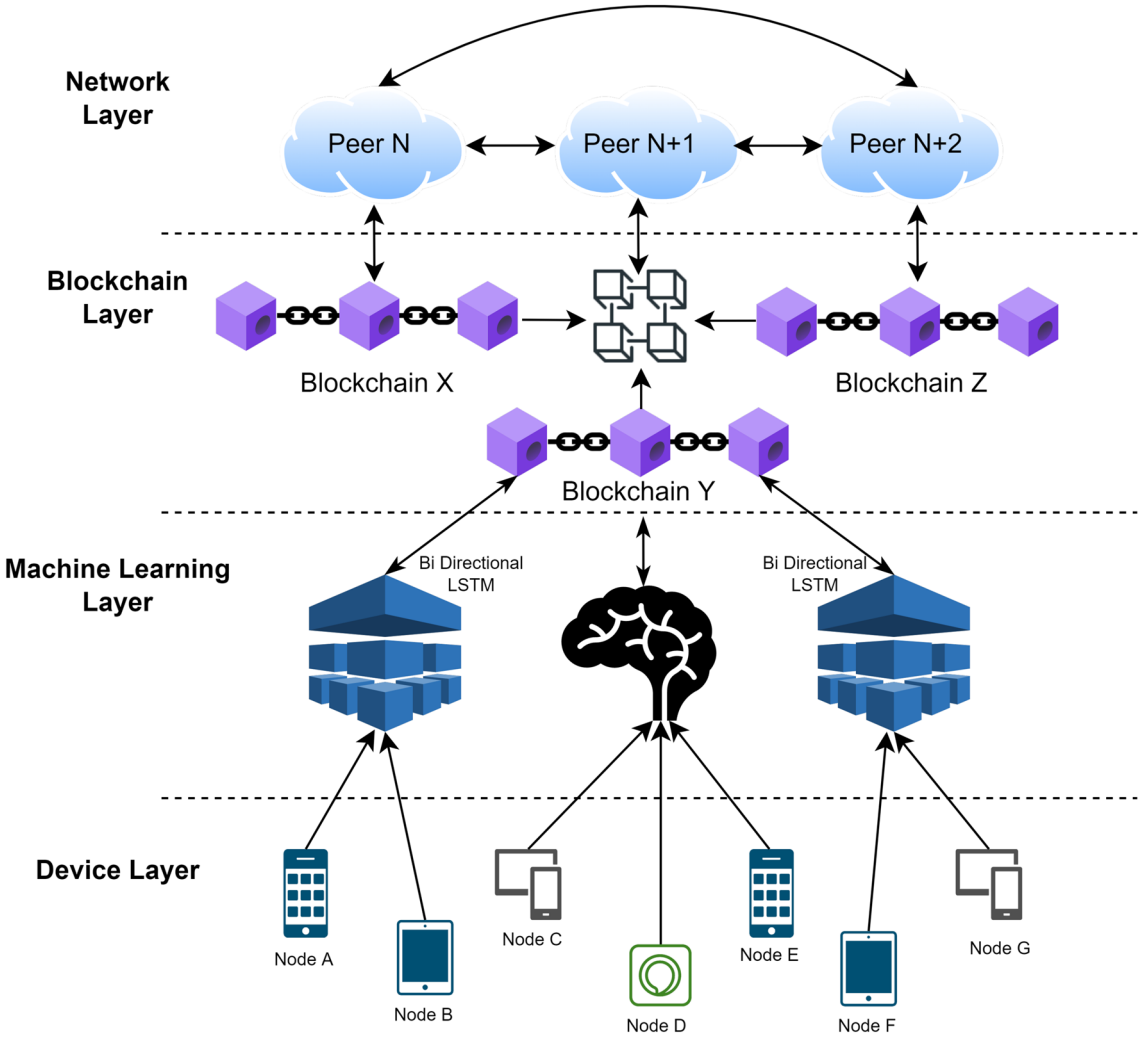

Fig. 3 System architecture

stop words and punctuations and applying stemming and lemmatization. Stemming is removing a part of a word, i.e., reducing the word to its stem word, whereas the word reduces to a dictionary word on lemmatization. We then select latent features for the classification of data.

The first step of preprocessing is stopword removal. We create a stopword list with some of the most common stopwords and the most redundant words in our training data that do not typically appear in default stopword lists. After this, we remove all stopwords from the data. However, the amount of data after this step is still huge and consists of words that can be reduced further.

Typically, natural language data has multiple worlds with similar semantics but different syntax that increase inconsistencies and features for models to process. To address this problem, we use stemming. Stemming is the procedure of all words with the same stem to a single standard form. For instance, the following words (consign, consigned, consignment, consigning) get reduced to their stem word-consign. Stemming reduces the data and features to an optimal amount. Once the data is cleaned, it is fed into a machine learning model for feature extraction.

We convert the preprocessed data into features that our model will train. We use a count vector featurizer to convert words into meaningful features. The idea is to maintain

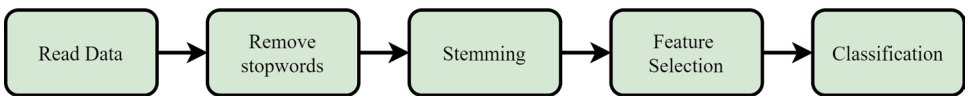

Fig. 4 Steps for classification

a large vector having dimensions equal to the size of our vocabulary. These vectors hold the frequency of each word that exists in the document. They estimate each word's impact in a document/dataset. Finally, we use these features and feed them into supervised models, which give meaningful predictions. To accomplish this, we split the data into 75% training data and 25% testing data.

4.2 LSTM model

LSTM is a type of recurrent neural network (RNN) primarily used to predict time series/sequential data. The problem with traditional RNN is the embedded short-term memory. In RNN, it is hard to persist in the long sequence from the earlier stage to eventually the last step. RNN also suffers from the vanish gradient problem during backpropagation. Therefore, layers associated with the small gradient do not update properly. LSTM is designed to overcome such issues of RNN by using different gates with hidden layers.

LSTM models have various advantages over conventional networks such as CNN and feed-forward neural networks. The unique 3 gate structure of LSTM gives them the ability to learn from time-series data, as the feedback allows for past inputs to leave a footprint on the model. It makes LSTM model suitable for text classification tasks where temporal information needs to be preserved. To preserve temporal data, we use a Bi-LSTM model as the given text information can flow in both directions (forward and backward). This property makes the proposed model highly suitable for text classification tasks like fake news detection.

Different types of LSTMs are employed in various applications like weather forecasts, natural language processing, computer vision, etc. Rani et al. (2022) used LSTM model to propose a social pandemic model for sentiment analysis on COVID-19 dataset. However, in the proposed work, we use a variant of the LSTM model, that is, Bi-LSTM to improve the system's overall accuracy. In Bi-LSTM models, two LSTM models are trained simultaneously. The first model is trained on the same input sequence. On the other hand, the second LSTM model uses the reverse of the input sequence for training. This technique improves the model efficiency significantly as the amount of information increases in the network. Figure 5 shows the structure of our designed Bi-LSTM model in which x_i and y_i indicate the input and output, respectively.

The classification model is used to reduce the loss function. Eq. (3) defines the loss function as follows:

$$H_p(q) = -\frac{1}{N}\sum_{i=1}^{N} y_i \cdot log(p(y_i)) + (1 - y_i) \cdot log(1 - p(y_i)) \tag{3}$$

where y is the class for the input, $p(y)$ denotes the predicted probability and $H_p(q)$ denotes the loss.

5 Experimental results and analysis

In this section, we prelude the experiments and evaluation as follows: dataset description, hyperparameter tunung, performance metrics, and results and discussion.

5.1 Dataset description

The dataset (Ahmed et al. 2018) consists of real-world news sources and was gathered by crawling various news and information websites. Truthful articles were obtained by crawling articles from the famous news website-Reuters. Whereas, fake news articles were obtained from various sources which were flagged for fake news by PolitiFact and

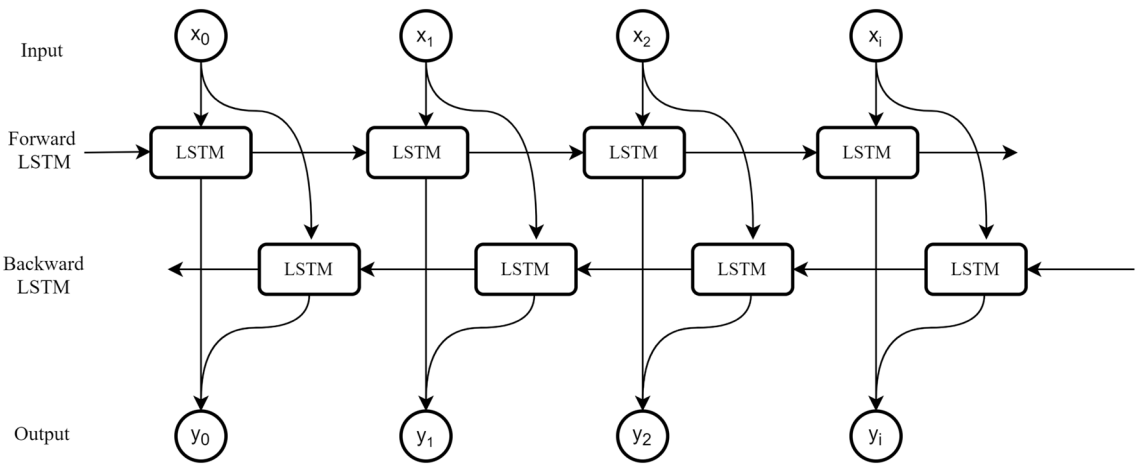

Fig. 5 LSTM model

Wikipedia. PolitiFact is a website used for fact checking and was used to determine the validity of a given claim. The authors performed data augmentation on the above dataset to create a more robust and adaptable model. Data augmentation was implemented by synonyms replacement, sentence shuffling and random swap operations.

5.2 Hyperparameter tuning

The procedure of choosing hyperparameters is a vital aspect of neural network model. Hyperparameters are the variables that need to set before enforcing any deep learning model to any dataset. The optimal values of hyperparameters are context-dependent and task-specific. Simulation parameters used in implementing the proposed model for achieving the accurate classification are enlisted in Table 2. Keras tuner was used to tune the model hyperparameters. Adam optimizer is utilized with a binary cross-entropy loss function. The model is run with a batch size of 16 and for 8 epochs. Further, a 20% validation split is used for fine-tuning the hyperparameters.

5.3 Performance metrics

The performance of a classification model is measured by counting the correct and incorrect predictions, whereas a confusion matrix provides the best pictorial representation of these terms. The performance metrics used for experimental results are accuracy, precision, recall, F1-score, and specificity. Accuracy is a metric to measure the correct prediction news out of all the classification instances present in the dataset. Precision counts the proportion of how much news is actually rumored out of all the detected rumors. Recall or sensitivity counts the successfully detected rumors out of the rumors presented in the dataset initially. They are computed as follows:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{4}$$

Table 2 Parameters used for the proposed model

Parameters	Information
Number of layers	6
Activation function	Relu
Dropout rate	30%
Optimizer	Adam
Learning rate	0.001
Loss function	Binary crossentropy
Batch size	16
Number of Epochs	8
Validation split	20%

$$Precision = \frac{TP}{TP + FP} \tag{5}$$

$$Recall = \frac{TP}{TP + FN} \tag{6}$$

where TP, TN, FP, and FN represent true positive, true negative, false positive, and false positive, respectively.

F1-score considers the impact of recall and precision and provides an aggregated value as follows:

$$F1 - score = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{7}$$

Specificity counts the measure of news that is incorrectly predicted as a rumor out of all the correct information. Specificity provides the exact opposite measure of recall. The formula for the computation of specificity is given in Eq. (8):

$$Specificity = \frac{TN}{TN + FP} \tag{8}$$

5.4 Results and discussion

This section presents the experimental results performed to measure the effectiveness of the proposed approach with the state-of-the-art models. A static network is considered for simulation where multiple connected nodes are deployed in a fixed location. We consider a social network with 50 nodes placed at predefined coordinates which have a property associated with them (malicious or genuine). Nevertheless, a malicious node has 50% probability of creating/spreading or illegitimately modifying a message in transmission, while a genuine node always initiates legitimate messages. The simulation is run for 2000 seconds to allow message transmission and assess system performance. Figure 6 depicts the simulation of the network in a 100×100 meter square area. The default simulation settings are taken to implement different classification models.

The confusion matrix and classification reports of the models are given in Figs. 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 and21. Figures 7 and 8 are the confusion matrix and classification report of random forest model, respectively. Figures 9 and 10 are the confusion matrix and classification report of Ada Boost models, respectively. Figures 11 and 12 are the confusion matrix and classification report of logistic regression, respectively. Similarly, Figs. 13 and 14 are the confusion matrix and classification report of KNN model. Figures 15 and 16 are the confusion matrix and classification report of Bernoulli Naive Bayes model. Figures 17 and 18 are the confusion matrix and classification report of Gradient Boost. Figures 19 and 20 are the confusion matrix and classification report of extra tree model. The confusion matrix of the proposed Bi-LSTM-based model is

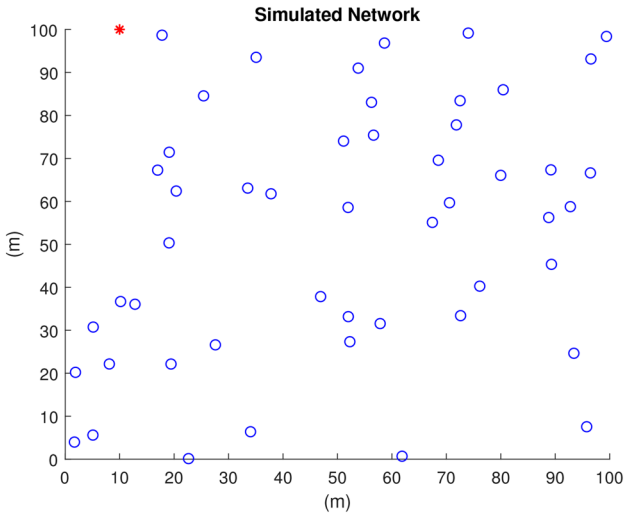

Fig. 6 A simulation of the network in a 100 × 100 meter square area

presented in Fig. 21. The accuracy and loss of the proposed model are depicted in Figures 22 and 23, respectively. It is clear from the results that the proposed Bi-LSTM model works better than other baseline methods on different performance metrics, with the best-provided accuracy of 99%. After the LSTM-based model, Ada-boost gives better accuracy, whereas *k*-nearest neighbor has the worst performance. Figure 24 shows the graph of block size versus network traffic.

Figure 22 depicts LSTM training and validation accuracy with increasing epochs. Figure 23 depicts LSTM Training and Validation Loss with increasing epochs. From these figures, we conclude that proposed LSTM models work better and give high accuracy under training compared to validation.

Figure 24 shows the effect of changing messages per block on the network traffic. We see an exponential change in traffic upon increasing messages per block. This line follows the perfect exponential curve in ideal conditions but will vary in a practical scenario.

6 Comparative analysis with existing results

We compare the proposed model with some baseline classification methods to validate the performance of our model. The classification methods used for comparison are random forest, Ada boost, logistic regression, k-nearest neighbor, Bernoulli naive Bayes, gradient boost, and extra tree classifier. Table 3 shows the comparative performance analysis of different machine learning algorithms and the proposed Bi-LSTM model under different training and testing data splits, that is, 70–30% and 60–40%.

Table 3 Comparison of proposed approach with different methods

Measure	Algorithm															
	Random Forest		Ada Boost		Logistic regression		k-nearest Neighbour		Bernoulli Naive Bayes		Gradient boost		Extra tree		Bi-LSTM	
	60–40%	70–30%	60–40%	70–30%	60–40%	70–30%	60–40%	70–30%	60–40%	70–30%	60–40%	70–30%	60–40%	70–30%	60–40%	70–30%
TP	10377	5827	10379	5835	10417	5842	8836	4868	10140	5663	10449	5829	10268	5759	4559	4647
TN	5615	5316	9661	5343	9576	5345	6992	3957	9351	5176	9657	5338	9542	5242	4234	4318
FP	4066	37	20	10	105	8	2689	1396	330	177	24	15	139	71	100	12
FN	147	45	145	37	107	30	1688	1004	384	209	75	43	256	113	87	3
Accuracy	0.79	0.98	0.99	0.99	0.98	0.97	0.78	0.78	0.96	0.96	0.99	0.99	0.98	0.98	0.97	0.99
Precision	0.71	0.99	0.99	0.99	0.99	0.99	0.76	0.77	0.96	0.96	0.99	0.99	0.98	0.98	0.97	0.99
Recall	0.98	0.99	0.98	0.99	0.98	0.99	0.83	0.82	0.96	0.96	0.99	0.99	0.97	0.98	0.98	0.99
F1-score	0.83	0.99	0.99	0.99	0.98	0.99	0.80	0.80	0.96	0.96	0.99	0.99	0.98	0.98	0.97	0.99
Specificity	0.38	0.90	0.92	0.91	0.91	0.91	0.60	0.63	0.89	0.88	0.92	0.91	0.91	0.89	0.90	0.92

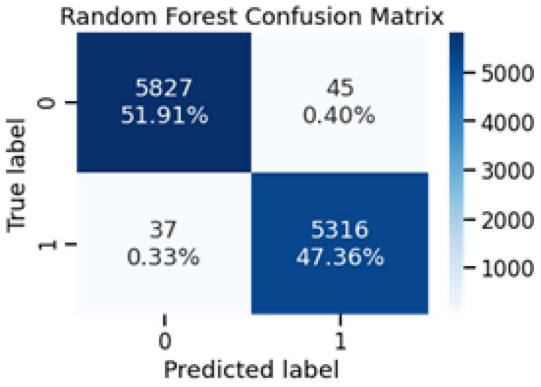

Fig. 7 Random Forest confusion matrix

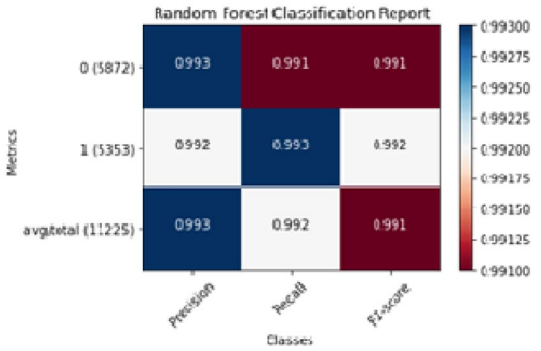

Fig. 8 Random Forest classification report

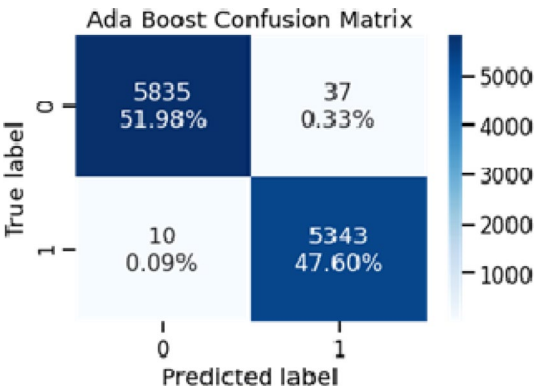

Fig. 9 Ada Boost confusion matrix

To validate the performance of the proposed methodology, we have also conducted a comparative analysis considering various deep learning model found in previous model. Table 4 summarizes the accuracy obtained by different recent work using distinguish methodology. From the results, its is evident that the proposed work achieved high accuracy with 99.63% and outperformed the other literature.

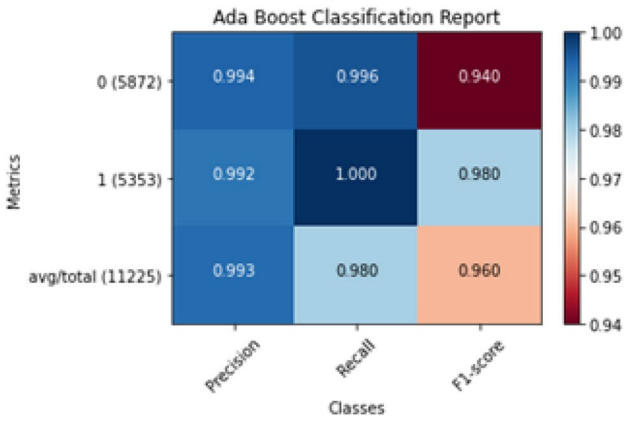

Fig. 10 Ada Boost classification report

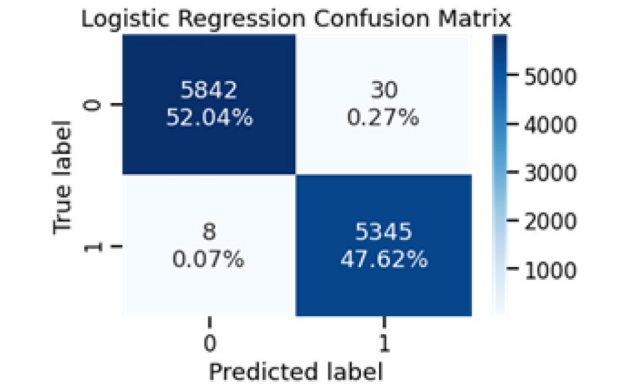

Fig. 11 Logistic Regression confusion matrix

Fig. 12 Logistic Regression classification report

7 Conclusion and future work

Detecting rumors on social media is a very challenging task. Therefore, various machine learning and deep learning-based approaches have been proposed to solve the

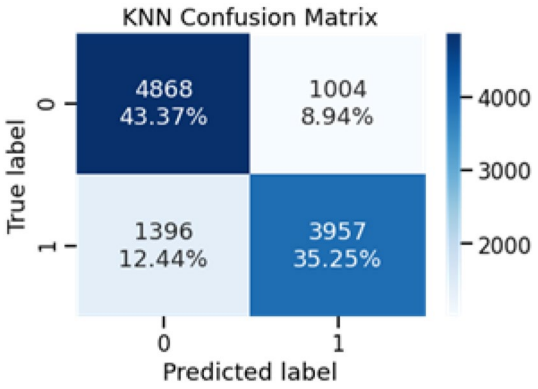

Fig. 13 KNN confusion matrix

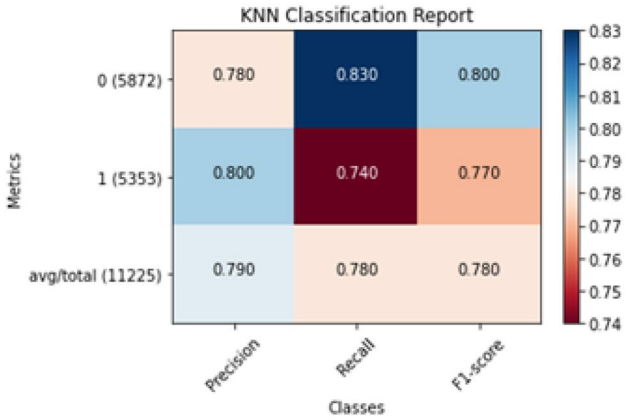

Fig. 14 KNN classification report

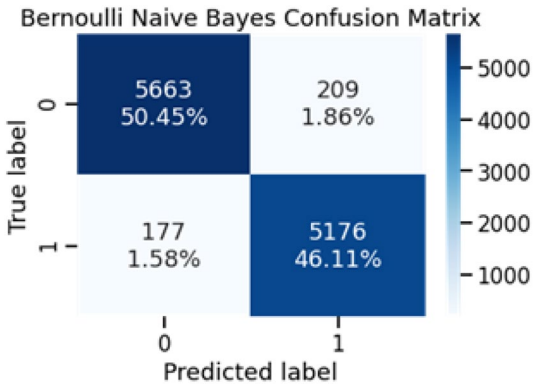

Fig. 15 Bernoulli Naive Bayes confusion matrix

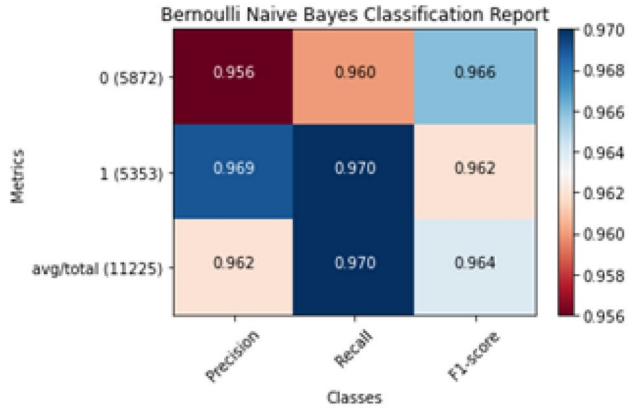

Fig. 16 Bernoulli Naive Bayes classification report

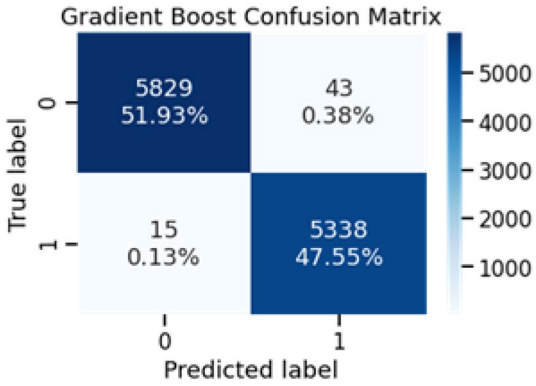

Fig. 17 Gradient Boost confusion matrix

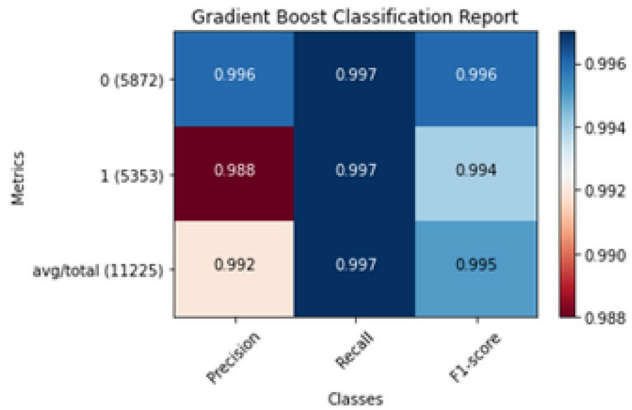

Fig. 18 Gradient Boost classification report

aforementioned problem. Results show that the deep learning models have more accuracy and significant improvements than the state-of-art machine learning algorithms. Current literature is limited by providing reactive measures to detect rumors. We have modeled an integrated blockchain and deep learning approach for rumors detection and prevention in the proposed work, giving both reactive and proactive solutions. The introduction of new rumors or false news to the system is detected using a Bi-LSTM model that continuously monitors incoming messages and detects false messages.

On the other hand, blockchain ensures the legitimacy of current information circulating in the network. This is done

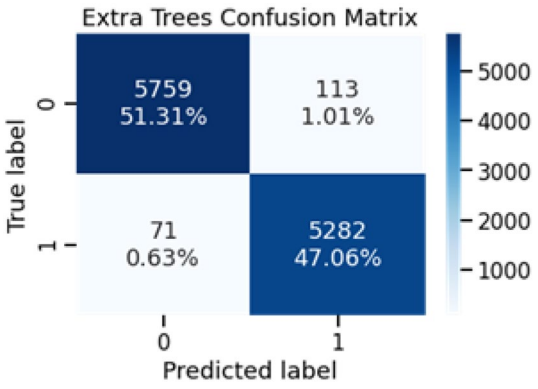

Fig. 19 Extra Tree confusion matrix

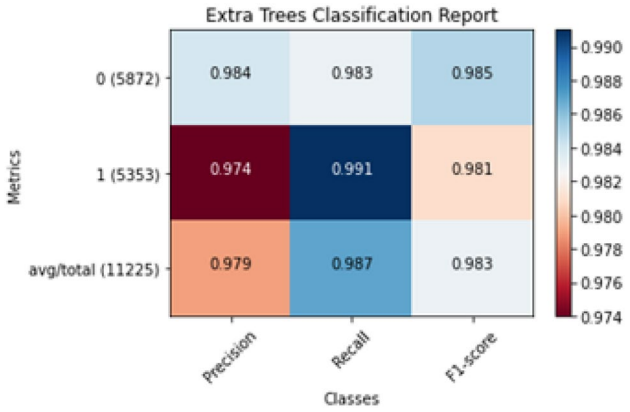

Fig. 20 Extra Tree classification report

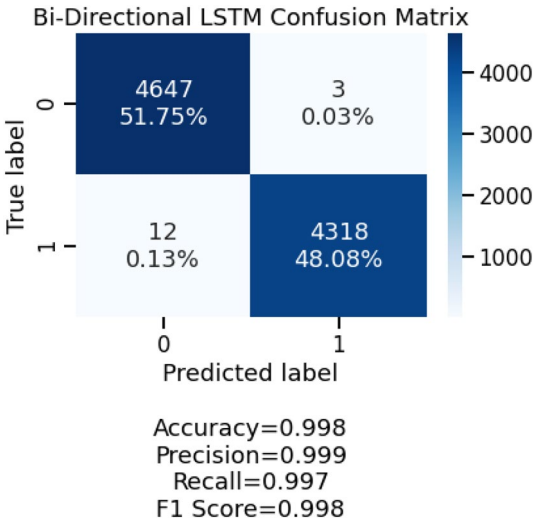

Fig. 21 Bi-LSTM confusion matrix

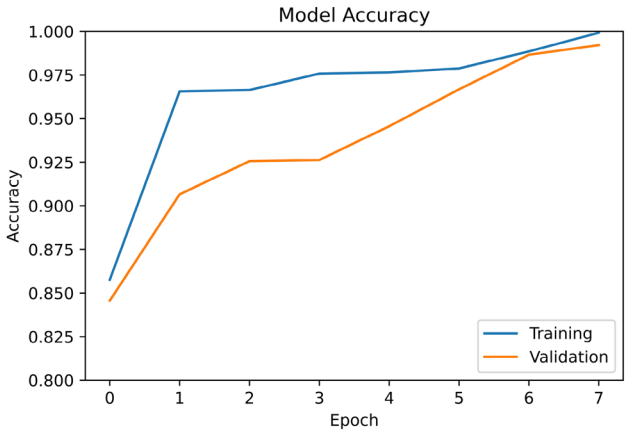

Fig. 22 Accuracy under LSTM model training vs validation

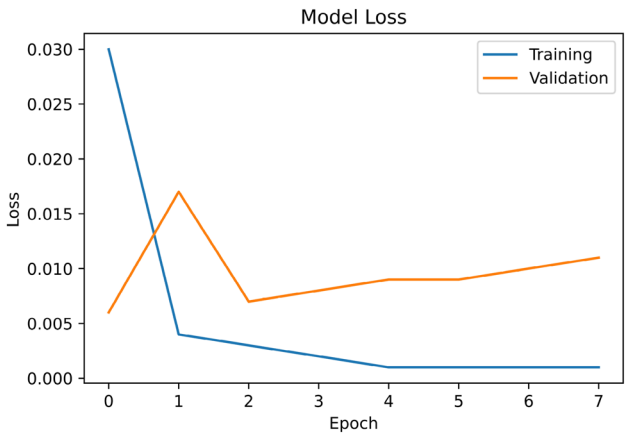

Fig. 23 Loss under LSTM model training vs validation

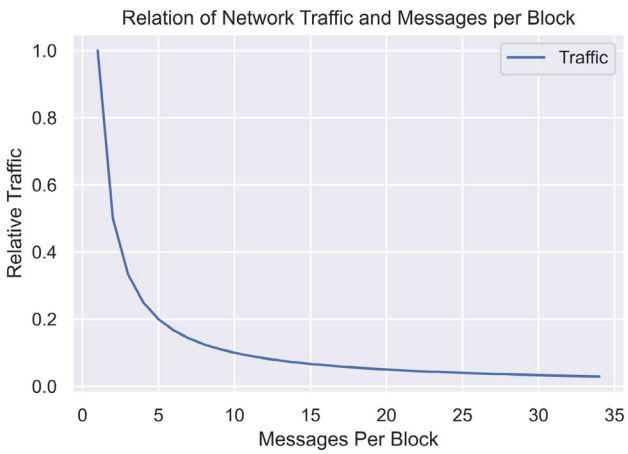

Fig. 24 Block size vs. network traffic

by using joint PoW-PoS consensus algorithms. The network load and resource protection were guaranteed by blacklisting malicious nodes identified by a credibility score. Overall, the system successfully ensured that all nodes got access to only accurate and credible information. To measure the

Table 4 Comparison of results with recent works

Paper	Methodology	Accuracy (%)
Yang et al. (2018)	CNN	92.10
O'Brien et al. (2018)	Deep neural network (black-box approach)	93.50
Ghanem et al. (2018)	Word embeddings and n-gram	48.80
Singh et al. (2017)	Linguistic analysis and word count	87.00
Ahmed (2017)	Uni-gram model	89.00
Ruchansky et al. (2017)	Hybrid model	89.20
Tida et al. (2022)	Hybrid BERT	97.00
Proposed model	Bi-LSTM	99.63

effectiveness of the proposed work, we have conducted various extensive simulations on the dataset of around 50,000 documents. Results are compared with recent work that comprises both machine learning and deep learning techniques. The proposed model outperforms the existing baseline work and predicts the rumors with 99.63% accuracy. Using the designed model, the probability of incorrect detection is significantly lower with only 0.13% false positive.

The proposed work has a few intriguing limitations that we will cover in our future work. In the future, a more extensive dataset can be involved consisting of multiple languages for the geographical-independent model with significantly improved accuracy. There is also scope for scaling the social media network to a real-world scenario and implementing it for small-scale organizations. In the current work, only a binary dataset is considered, with only two classes predicting whether the information is rumor or not. For the future, we are planning to incorporate a hybrid model with multi label-datasets with a less complex deep learning model.

To compare the accuracy scores of the proposed model with baseline models, T-test (paired) hypothesis testing was used. Accuracies obtained from benchmark CNN and BERT based models were compared with the proposed model. Hypothesis: H0: The accuracies of the proposed model did not improve compared to the benchmark H1: The accuracies of the proposed model improved compared to the benchmark The T-test statistic was found to be 4.441490, while the p-value was found to be 0.003378. Since the p-value (0.003378)< alpha (0.05). The null hypothesis H0 is rejected. It can be concluded that the accuracies of the proposed model have improved compared to the benchmark.

Data availability statement The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

Declarations

Conflict of interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

References

Ahmed H (2017) Detecting opinion spam and fake news using n-gram analysis and semantic similarity. PhD thesis

Ahmed H, Traore I, Saad S (2018) Detecting opinion spams and fake news using text classification. Secur Privacy 1(1):e9

Al-Asadi MA, Tasdemır S (2022) Predict the value of football players using FIFA video game data and machine learning techniques. IEEE Access 10:22631–22645

Alsaeedi A, Al-Sarem M (2020) Detecting rumors on social media based on a CNN deep learning technique. Arab J Sci Eng 45(12):10813–10844

Del Vicario M, Bessi A, Zollo F, Petroni F, Scala A, Caldarelli G, Stanley HE, Quattrociocchi W (2016) The spreading of misinformation online. Proc Natl Acad Sci 113(3):554–559

Dibaei M, Zheng X, Xia Y, Xu X, Jolfaei A, Bashir AK, Tariq U, Yu D, Vasilakos AV (2021) Investigating the prospect of leveraging blockchain and machine learning to secure vehicular networks: A survey. IEEE Trans Intell Transp Syst 23(2):683–700

Gao H, Gao T (2020) Prevention of rumor spreading based on blockchain. In: 2020 IEEE 20th international conference on communication technology (ICCT), IEEE, pp 1174–1178

Ghanem B, Rosso P, Rangel F (2018) Stance detection in fake news a combined feature representation. In: Proceedings of the first workshop on fact extraction and VERification (FEVER), pp 66–71

Guille A, Hacid H, Favre C, Zighed DA (2013) Information diffusion in online social networks: a survey. ACM SIGMOD Rec 42(2):17–28

Jin Z, Cao J, Guo H, Zhang Y, Wang Y, Luo J (2017) Detection and analysis of 2016 US presidential election related rumors on twitter. In: International conference on social computing, behavioral-cultural modeling and prediction and behavior representation in modeling and simulation, Springer, pp 14–24

Khanna A, Rani P, Sheikh TH, Gupta D, Kansal V, Rodrigues JJ (2021) Blockchain-based security enhancement and spectrum sensing in cognitive radio network. Wirel Person Commun 2021:1–23

Li Q, Zhang Q, Si L, Liu Y (2019) Rumor detection on social media: datasets, methods and opportunities. arXiv:1911.07199

Li W, Andreina S, Bohli JM, Karame G (2017) Securing proof-of-stake blockchain protocols. In: Data privacy management, cryptocurrencies and blockchain technology. Springer, pp 297–315

Manzoor SI, Singla J, et al (2019) Fake news detection using machine learning approaches: a systematic review. In: 2019 3rd international conference on trends in electronics and informatics (ICOEI), IEEE, pp 230–234

Oberlechner T, Hocking S (2004) Information sources, news, and rumors in financial markets: insights into the foreign exchange market. J Econ Psychol 25(3):407–424

O'Brien N, Latessa S, Evangelopoulos G, Boix X (2018) The language of fake news: Opening the black-box of deep learning based

detectors. In: 32nd conference on neural information processing systems, center for brains, minds and machines (CBMM)
Osatuyi B (2013) Information sharing on social media sites. Comput Hum Behav 29(6):2622–2631
Pandey N, Pal A et al (2020) Impact of digital surge during covid-19 pandemic: a viewpoint on research and practice. Int J Inf Manage 55:102171
Pathak AR, Mahajan A, Singh K, Patil A, Nair A (2020) Analysis of techniques for rumor detection in social media. Procedia Comput Sci 167:2286–2296
Połap D, Srivastava G, Jolfaei A, Parizi RM (2020) Blockchain technology and neural networks for the internet of medical things. In: IEEE INFOCOM 2020-IEEE conference on computer communications workshops (INFOCOM WKSHPS), IEEE, pp 508–513
Połap D, Srivastava G, Yu K (2021) Agent architecture of an intelligent medical system based on federated learning and blockchain technology. J Inf Secur Appl 58:102748
Rani P, Bhatia M, Tayal D (2018) Qualitative SNA methodology. In: 2018 5th international conference on computing for sustainable global development), IEEE, pp 4223–4228
Rani P, Bhatia M, Tayal D (2019a) A comparative study of qualitative and quantitative sna. In: 2019 6th international conference on computing for sustainable global development (INDIACom), IEEE, pp 500–504
Rani P, Bhatia M, Tayal D (2019b) A soft computing-based approach to group relationship analysis using weighted arithmetic and geometric mean. In: International conference on innovative computing and communications, Springer, pp 171–178
Rani P, Tayal DK, Bhatia M (2019c) SNA using user experience. In: 2019 international conference on machine learning, big data, cloud and parallel computing (COMITCon), IEEE, pp 125–128
Rani P, Balyan A, Jain V, Sangwan D, Singh PP, Shokeen J (2020) A probabilistic routing-based secure approach for opportunistic IoT network using blockchain. In: 2020 IEEE 17th India council international conference (INDICON), IEEE, pp 1–7
Rani P, Bhatia M, Tayal DK (2021a) Conical SNA using fuzzy k-medoids based on user experience. Int J Electr Eng Educ 2021:0020720920988490
Rani P, Jain V, Joshi M, Khandelwal M, Rao S (2021b) A secured supply chain network for route optimization and product traceability using blockchain in internet of things. In: Data analytics and management, Springer, pp 634–647
Rani P, Shokeen J, Agarwal A, Bhatghare A, Majithia A, Malhotra J (2021c) Credit card fraud detection using blockchain and simulated k-means algorithm. In: International conference on innovative computing and communication (ICICC). Springer, Berlin
Rani P, Singh PP, Balyan A, Shokeen J, Jain V, Sangwan D (2021d) A secure epidemic routing using blockchain in opportunistic internet of things. In: Data Analytics and Management. Springer, Berlin, pp 101–110
Rani P, Tayal DK, Bhatia M (2021e) Sociocentric SNA on fuzzy graph social network model. Soft Comput 2021:5
Rani P, Shokeen J, Majithia A, Agarwal A, Bhatghare A, Malhotra J (2022) Designing an LSTM and genetic algorithm-based sentiment analysis model for covid-19. In: Proceedings of data analytics and management. Springer, Berlin, pp 209–216
Raza S, Ding C (2022) Fake news detection based on news content and social contexts: a transformer-based approach. Int J Data Sci Anal 2022:1–28
Ruchansky N, Seo S, Liu Y (2017) Csi: a hybrid deep model for fake news detection. In: Proceedings of the 2017 ACM on conference on information and knowledge management, pp 797–806
Shae Z, Tsai J (2019) AI blockchain platform for trusting news. In: 2019 IEEE 39th international conference on distributed computing systems (ICDCS), IEEE, pp 1610–1619
Shelke S, Attar V (2019) Source detection of rumor in social network—a review. Online Soc Netw Media 9:30–42
Shokeen J, Rana C (2021) A trust and semantic based approach for social recommendation. J Ambient Intell Hum Comput 2021:1–15
Shokeen J, Rana C, Rani P (2021) A trust-based approach to extract social relationships for recommendation. In: Data analytics and management. Springer, Berlin, pp 51–58
Singh V, Dasgupta R, Sonagra D, Raman K, Ghosh I (2017) Automated fake news detection using linguistic analysis and machine learning. In: International conference on social computing, behavioral-cultural modeling, & prediction and behavior representation in modeling and simulation (SBP-BRiMS), pp 1–3
Subudhi RN, Palai DP (2020) Impact of internet use during COVID lockdown. J Human Soc Sci Res 2:59–66
Thakur HK, Gupta A, Bhardwaj A, Verma D (2018) Rumor detection on twitter using a supervised machine learning framework. Int J Inf Retrieval Res (IJIRR) 8(3):1–13
Tida VS, Hsu D, Hei D et al (2022) Unified fake news detection using transfer learning of bidirectional encoder representation from transformers model. arXiv:2202.01907
Toshida T, Jagruti C (2020) COVID-19-rumours and facts in media. Int J Res Pharmaceut Sci Spec Issue 11:171–174
Uchejeso O, Etukudoh N, Chukwudimma O, Ogechukwu E, Amadi K et al (2021) Coronavirus: the biological threat of our time. J Curr Emerg Med Rep 1(1):1–7
Vukolić M (2015) The quest for scalable blockchain fabric: Proof-of-work vs. BFT replication. In: International workshop on open problems in network security. Springer, Berlin, pp 112–125
Yang Y, Zheng L, Zhang J, Cui Q, Li Z, Yu PS (2018) TI-CNN: convolutional neural networks for fake news detection. arXiv:1806.00749
Zhou Y, Wu C, Zhu Q, Xiang Y, Loke SW (2019) Rumor source detection in networks based on the SEIR model. IEEE Access 7:45240–45258