This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

SaTYa: Trusted Bi-LSTM-Based Fake News Classification Scheme for Smart Community

Pronaya Bhattacharya, Shivani Bharatbhai Patel, Rajesh Gupta, *Student Member, IEEE*, Sudeep Tanwar, *Senior Member, IEEE*, and Joel J. P. C. Rodrigues, *Fellow, IEEE*

***Abstract*—This article proposes a SaTya scheme that leverages a blockchain (BC)-based deep learning (DL)-assisted classifier model that forms a trusted chronology in fake news classification. The news collected from newspapers, social handles, and e-mails are web-scrapped, prepossessed, and sent to a proposed Q-global vector for word representations (Q-GloVe) model that captures the fine-grained linguistic semantics in the data. Based on the Q-GloVe output, the data are trained through a proposed bi-directional long short-term memory (Bi-LSTM) model, and the news is classified as real-or-fake news. This reduces the vanishing gradient problem, which optimizes the weights of the model and reduces bias. Once the news is classified, it is stored as a transaction, and the news stakeholders can execute smart contracts (SCs) and trace the news origin. However, only verified trusted news sources are added to the BC network, ensuring credibility in the system. For security evaluation, we propose the associated cost of the Bi-LSTM classifier and propose vulnerability analysis through the smart check tool for potential vulnerabilities. The scheme is compared against discourse-structure analysis, linguistic natural language framework, and entity-based recognition for different performance metrics. The scheme achieves an accuracy of 99.55% compared to 93.62% against discourse structure analysis. Also, it shows an average improvement of 18.76% against other approaches, which indicates its viability against fake-classifier-based models.**

***Index Terms*—Bi-directional long short-term memory (Bi-LSTM), blockchain (BC), deep learning (DL), fake news identification.**

I. Introduction

In RECENT years, there has been an exponential growth in social network platforms. Anyone can broadcast any message, or news (i.e., verified or unverified) on the social media network, which can be further spread or viral without any formal verification [1]. This causes numerous news posts to get flooded on the social network platforms without any credibility. This poses a huge obstacle for various social network sites such as Facebook, Twitter, WhatsApp, and LinkedIn to classify the massive number of news posts as real or fake. Furthermore, the widespread of fake news has negatively impacted the social community and can adversely affect the mental well-being of its people in numerous ways. The studies have used several techniques to assess the veracity of news circulated on social media. Machine learning (ML) techniques like random forest classifiers, support vector machines, neural networks, and many others have been used to check the credibility of posted news content.

Some studies have also explored the linguistic analysis using natural language processing (NLP) for the same [2]. In NLP systems, the high-dimensional news article sources are converted to minimized word space vector, and different approaches are employed that exploit the contextual properties of word forms in the sentence [3]. Other approaches normally employ embedding word vector (WV) to multi-scale latent features and use bidirectional encoder representations for extraction of word embedding from news sources [4]. However, in most approaches, the semantic interpretation is not conserved, and thus the operational complexity increases. In news classification models, the news posts are required to be analyzed and classified in real-time to reveal the underlying pattern for the identification of fake news. To fulfill this, the use of ML approaches alone cannot satisfy the need of the hour.

Deep learning (DL) algorithms like long short-term memory (LSTM), bi-directional LSTM, and many others can be utilized to pioneer the model to capture and extract the features from the data to grab the hidden implications of news content over time [5]. Then, adding the extra hidden layers to the DL model can further improve its performance. The model contains sensitive training and testing data, with various security and privacy concerns such as data modification. This causes inaccurate results in classifying news as fake or real. However, even after news classification, the purpose of elimination of fake news generation is not achieved, as the chronology of the news generation is unclear.

To combat digital deception, a distributed ledger technology (DLT), i.e., blockchain (BC) [6] can be used to ensure the security, privacy, and trust in a decentralized peer-to-peer (P2P) network [7], [8]. A transaction, i.e., data, cannot tamper once it is distributed, accepted, and validated by the network consensus algorithm [9]. The features like authentication, efficiency, and secure data storage can be achieved using DLT,

Manuscript received August 2, 2021; revised October 27, 2021; accepted November 29, 2021. This work was supported in part by FCT/MCTES through national funds and when applicable co-funded EU funds under Project UIDB/50008/2020 and in part by the Brazilian National Council for Scientific and Technological Development (CNPq) under Grant 313036/2020-9. *(Corresponding author: Sudeep Tanwar.)*

Pronaya Bhattacharya, Shivani Bharatbhai Patel, Rajesh Gupta, and Sudeep Tanwar are with the Department of Computer Science and Engineering, Institute of Technology, Nirma University, Ahmedabad, Gota, Gujarat 382481, India (e-mail: pronoya.bhattacharya@nirmauni.ac.in; 17bce117@nirmauni.ac.in; 18ftvphde31@nirmauni.ac.in; sudeep.tanwar@nirmauni.ac.in).

Joel J. P. C. Rodrigues is with the Senac Faculty of Ceará, Fortaleza 60160-194, Brazil, and also with the Instituto de Telecomunicações, 6201-001 Covilhã, Portugal (e-mail: joeljr@ieee.org).

Digital Object Identifier 10.1109/TCSS.2021.3131945

and the role of a third party for trust establishment can be eliminated using the digital SC [10]. DLT helps to stop the distribution of fake news as soon as it is identified by the DL model [11]. Various researchers across the globe have given their contributions to the field of fake news identification in social media platforms.

A. Existing Schemes

In literature, researchers globally have proposed solutions to predict news popularity [12] as well as address the classification of fake news circulation over the social platform through ML and DL approaches, and have proposed trusted fake news origin classification via BC. For example, Vishwakarma *et al.* [13] analyzed the accuracy of the information on various social media platforms by exploring it on the web and checked its credibility by comparing it against the top 15 Google search results. The real parameter is calculated subsequently, compared against the threshold to classify the information as fake or real. Then, Dey *et al.* [14] presented the linguistic analysis on tweets and applied the k-nearest neighbors' algorithm for news post classification. Zhang *et. al* [15] proposed an analytics-driven scheme by using a double-layered approach for news classification. Xu *et al.* [16] analyzed the domain reputation and content characterizations of fake and real news, which provides key insights to detect fake news on social media. Then, the authors in [17] extracted the relationship between various news texts using principal component analysis (PCA) and convolutional neural network (CNN).

Later, a deep CNN with multiple hidden layers has been proposed by Kaliyar *et al.* [18]. They addressed the problem of selecting an optimal depth of CNNs for fake news detection. Then, Traylor *et al.* [19] used a quoted attribution in the Bayesian network system as a key feature to identify the fake news. Furthermore, Ishida and Kuraya [20] proposed a bottom-up approach using a dynamic relational network in which each node evaluates the other nodes and cross-validates the information to identify the fake news. In line with the DL models, Bahad *et al.* [21] presented a model which is based on Bi-directional LSTM-recurrent neural network (RNN) on a balanced and imbalanced high dimensional news dataset. Then, Ozbay and Alatas [22] proposed a text-mining and supervised artificial intelligence (AI) model to detect fake news and conclude the best performance metrics that have been achieved using a decision tree, ZeroR, CV Parameter Selection, and weighted instances handler wrapper algorithms. Later, Nyow and Chua [23] presented the mechanisms of identifying the significant tweets' attributes and application architecture to automate the classification of online news. Ahn and Jeong [24] utilized the Google-based bidirectional encoder representations from transformers, [25] an NLP technique to pre-train its model. Using this model, fine-tuning matches the dataset detected by fake Korean news, and an AUROC score of 83.8% is achieved. Similarly, Shae and Tsai [26] proposed an integrated AI-BC platform to provide journalists with BC crowd-sourced and AI validated factual data on emerging news.

B. Novelty

Most of the aforementioned schemes are heavily tilted toward classifying fake news based on DL, NLP, and supervised AI models. However, once a piece of news is classified as real or fake news, we require a trust-based DLT mechanism that can chronologically trace the fake news generation to curtail further fake news sources from that origin. In such cases, BC is preferred, as the blocks are chronologically linked to store the information, and the records are transparent and immutable. Through BC, the origin address of the fake news generator is traced, and the address is publicly visible to other news stakeholders in the BC network.

Thus, to address the research gap, the authors proposed integration of BC and DL-based scheme to classify fake news and store the information as a trusted ledger. Moreover, the schemes incorporate both NLP-based Q-GloVe and Bi-LSTM frameworks that preserve the semantic interpretation of textual form before classification. The owner of the news source is traced through SC that executes as part of the BC framework. The scheme considers a public BC so that records are transparent to all stakeholders. To ensure high scalability, we convert the news articles as textual information using the Q-GloVe method, and the input is fed to the Bi-LSTM model that generates a news classification score, which is stored as a short external 32-byte hash address in the main BC. Thus, a large number of news sources are verified in real-time in the proposed ecosystem.

C. Research Contributions

Following are the contributions of this article.

1) To propose a Q-GloVe methodology that creates word embedding vectors of the collected data in a matrix form and presents the co-relation of the probe word against the stored vectors.
2) To design a Bi-LSTM model that classifies the news as fake or valid with probability p. It accepts the input as training, validation, and testing datasets, which is based on the output of Q-GloVe method.
3) To develop SCs for tracing the news's source origin and comparing p against a specific threshold. If the news source is genuine, only the result is stored in the BC network. It ensures trust in the ecosystem.

D. Article Layout

The article layout is as follows. Section II presents the system model and the problem formulation of the proposed scheme. Section III presents the details of the proposed scheme and discusses the proposed Q-GloVe method, the Bi-LSTM model, and the discussion of proposed SC execution and transaction storage in the BC network. Section IV discusses the performance evaluation, and finally Section V concludes the article.

II. SaTYa: System Model and Problem Formulation

In this section, we first propose the details of a three-layered DL and BC integrated scheme SaTYa for fake news classification and identification for smart community interactions. Next,

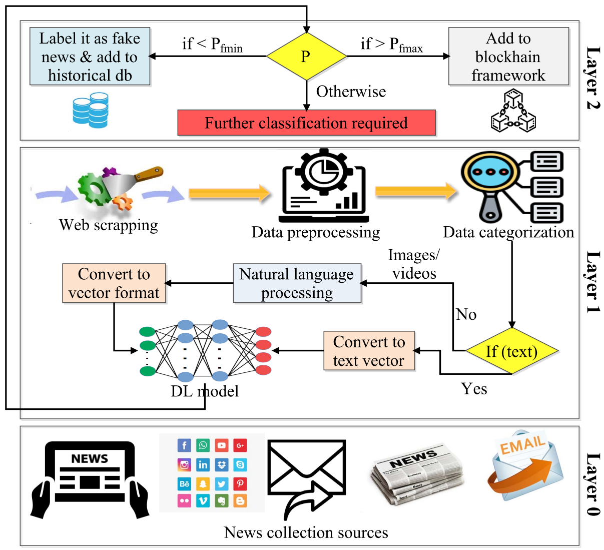


Fig. 1. SaTYa: The system model.

we present the details of the problem formulation that handles the collected data to perform WV embedding, and then the model is trained through the Bi-LSTM model. Finally, the details are presented as follows.

A. System Model

We present the details of the system model. It consists of an entity-set $E \in \{E_{\mathrm{NW}}, E_{\mathrm{SM}}, E_{\mathrm{MF}}, E_{\mathrm{EM}}\}$, where E_{NW}, E_{SM}, E_{MF}, and E_{EM} represent the data forwarded through persons via news websites (media), social media handles, message forward through chat communities, and e-mail forward, respectively. The details of the proposed layered stack are shown in Fig. 1. The stack consists of three layers, with the movement of data from Layer 0 to Layer 2. In the scheme, Layer 0 is named as data collection layer and is responsible for the collection of raw data from heterogeneous sources as $\mathrm{DC} = \{D_1, D_2, \ldots, D_n\}$, where DC denotes the collection of raw data to Layer 1. DC is mined through web-scrapper W to extract meaningful patterns that decrease the size of DC, which also reduces the computational cost of the proposed scheme. The mined data, called data forward DF, has a trivial condition, $\mathrm{DF} < \mathrm{DC}$.

DF is now forwarded to Layer 1, where the mined data are pre-processed categorically as $\mathrm{DB} = \{D_t, D_i, D_a, D_v\}$, where D_t, D_i, D_a, and D_v denotes the text, image, audio and video, respectively. A mapping M is defined as $M : \mathrm{DC} \to \mathrm{DB}$ to denote the association. Based on M, DB is converted to WV forms as $\mathrm{WV} = \{V_1, V_2, \ldots, V_k\}$, based on a proposed NLP method Q-GloVe. WV is now streamed as bits $\{b_0, b_1, \ldots, b_l\}$, with condition $l > k$. WV is now fed into a Bi-LSTM model through input gate I_t because using Bi-LSTM model ensures additional training of the model by traversing the data in both the orders: 1) left to right and 2) right to left. It generates a classification score (CS), and propagates CF to Layer 2, denoted as transaction layer.

At Layer 2, a fake news probability P_f is computed and is compared with minimum and maximum threshold $P_{f_{\min}}$ and $P_{f_{\max}}$, set as 0.35 and 0.65, respectively. In case $P_f < P_{f_{\min}}$, the news is labeled as fake news and is discarded. In case $P_f > P_{f_{\max}}$, the news is verified as valid news, and the transaction is appended to BC. In the case of mid-range, the classification fails and WV is further processed back at Layer 1. Thus, only verified transactions classified through Bi-LSTM are stored as transparent and immutable ledger in BC.

B. Problem Formulation

As discussed in Section II-A, the proposed scheme, SaTYa collects the data (News) from different sources such as news websites, E_{NW}, social media handles, E_{SM}, message or text forward, E_{MF}, and emails, E_{EM} at Layer 0. The collected data is then segregated and classified based on its format: text D_t, image D_i, audio D_a and video D_v and stored in the database, DB, as

$$\mathrm{DB} \leftarrow \{D_t, D_i, D_a, D_v\}. \tag{1}$$

The collected data are then pre-processed in three steps: data cleaning, data transformation, and data reduction. In data cleaning, the missing values are handled by either ignoring the tuples or adding the missing values and noisy data. Three techniques are applied to fill missing values, namely, the binning method, regression, and clustering. In the binning scheme, we prefer the equal width binning technique over frequency binning, as the frequency technique does not accurately minimize the overfitting problem for small datasets. Contrary, in the equal width binning, we consider equal-sized bins defined over a minimum bin size $B_{\min}$, and maximum bin size $B_{\max}$. The window range, W_{bin} is then defined over the number of created bins B_n as follows:

$$W_{\mathrm{bin}} = \frac{B_{\max} = B\min}{B_n}. \tag{2}$$

Based on the value of W_{bin}, the equal bin-width is defined as follows:

$$\mathrm{BW} = \{(B_{\min} + W_{\mathrm{bin}}), (B_{\min} + 2W_{\mathrm{bin}}), \ldots, (B_{\min} + q\,W_{\mathrm{bin}})\} \tag{3}$$

where q is the number of equal-bin widths defined. In the regression analysis, we use the independent vector variables V_i, and dependent scalar variables, D_i, error term e_i, and unknown parameter τ. D_i is now presented as $f(V_i, \tau)$ as follows:

$$D_i = f(V_i, \tau) + ei \tag{4}$$

subject to minimize the value of the constraint C_1 : $\min(\sum_i (D_i - f(V_i, \tau))^2$. Next, we apply the k-means cluster approach, by initially selecting k-centroids, where k depends on DB. Any point P is assigned to the closest centroid, to form the cluster C. The points P are iterated until all the points are arranged in some specific cluster C. Once the missing values are filled, DB is transformed into a suitable form for mining process. Once mining process is completed, we apply the PCA on DB to find effective dimension d out of n-dimensional space, where $d < n$. To formulate the same, we consider a given set of data vectors V_i, where $i \in \{1, n\}$, and form d orthonormal axes. The first principal component PC_1, is chosen to have maximum variance. We assume that data values $\{v_1, v_2, \ldots, v_n\}$ are presented into columnar form in a $n \times d$ matrix V. Each column $C(V)$ corresponds to n-dimensional observation space, and a total of d observations are chosen. PC_1 is defined on linear combination of V, with weights $w = [w_1, w_2, \ldots, w_n]$, presented as follows:

$$\mathrm{PC}_1 = w^T V$$
$$\mathrm{var}(\mathrm{PC}_1) = \mathrm{var}(w^T V) = w^T S w \tag{5}$$

where S denotes the covariance on V. In order to increase the value of $\mathrm{var}(\mathrm{PC}_1)$, we maximize $w^T S w$, where w is constrained to unit length vector, i.e., $C_2 : \max(w^T S w) = 1$. To solve the optimization problem constraint C_2, we consider Lagrangian multiplier β_1 as follows:

$$L(w, \beta_1) = w^T S w - \beta_1(w^T w - 1). \tag{6}$$

Differentiating (6) w.r.t. w, we obtain n equations as follows:

$$S w = \beta_1 w. \tag{7}$$

Next, we permute both sides by w^T, and obtain

$$w^T S w = \beta_1 w = \beta_1 \tag{8}$$

where $\mathrm{var}(\mathrm{PC}_1)$ is maximized only if β_1 is the largest eigenvector of S. Based on $\mathrm{var}(\mathrm{PC}_1)$, the reduction space is at d-dimensional space, which makes the data storage efficient. From DB, we present the reduced $\mathrm{DB}_{\mathrm{red}}$ through PCA. In $\mathrm{DB}_{\mathrm{red}}$, we take $\{D_v, D_a, D_i\}$ files, and convert them to D_t through the Markov model [27]. For D_v, we extract the frame sequences $\{f_1, f_2, \ldots, f_t\}$ defined over a time interval t as still D_i, and store them in frame buffer F_v for processing. Once D_t is generate, it is analyzed through the proposed Bi-LSTM model, as presented in Section III-B.

Vector representations for D_t words are obtained through the proposed Q-GloVe scheme that presents an unsupervised learning approach that converts the word to vector form. Through the Q-GloVe algorithm, we present an unsupervised learning approach that generates high-dimensional news articles are generated and is then pre-trained through Q-GloVe word embeddings. The details of the proposed Q-GloVe scheme are presented in Section III-A. The proposed technique optimizes the traditional GloVe algorithm, where random weights are loaded. Instead, in Q-GloVe, we uploaded weights from the embedding layer, which reduces the overall operational complexity. The technique applies to global aggregated co-occurrence statistics across all the words in the news article corpus.

Once the vectors are generated through Q-GloVe, they are presented as inputs to the Bi-LSTM model, and the model is trained for fake-news classification. To optimize the performance of the classification model, we add multiple stacked LSTM layers in the model, which allows the model to gain a high-level understanding of the presented input sequences. Each layer in the stacked LSTM model presented a word hierarchy that receives the hidden state information from the previous layer as input.

From the output, the metadata M of the processed information is extracted and stored in the BC network. As the meta-information is stored in the BC network, it becomes immutable and transparent to all users. This allows trusted fake-news classification for all stakeholders and is also publically verifiable. Also, meta-information transactions are small in size, and thus more transactions are added per block, and the scalability of the BC network is preserved. In case of updates in the Bi-LSTM model, a new meta-information block is appended in BC, and thus at any time instant, the true state of the fake-news classification is preserved. Fig. 2 presents the operational flow sequence of the proposed scheme.

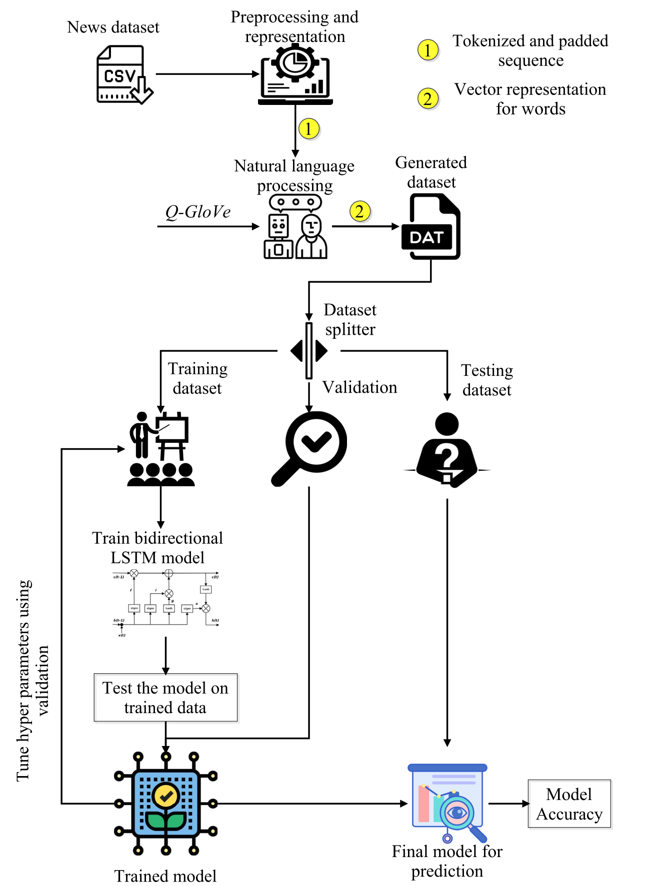


Fig. 2. SaTYa: The operational flow sequence.

III. SaTYa: The Proposed Scheme

In this section, we propose the scheme, SaTYa for fake news classification. We first present the details of the Q-GloVe methodology, and then present the proposed Bi-LSTM model. The details are presented as follows.

A. Q-GloVe Methodology

As described in Section II-B, we present the Q-GloVe methodology that creates word embedding vectors in matrix form. To formulate the same, we consider a word p in any ith text D_t and tabulate the frequency of occurrence of word q in the context for word p. The matrix of word q co-occurrence with p is denoted by N^i_{pq}, where N^i is the sum of all occurrences of correlated words and is shown as follows:

$$N^i_{pq} = \sum_r N^i_{rk}. \tag{9}$$

Here $\sum_r N^i_{rk}$ represents the number of times word k occurs in the context of word r. We define the conditional probability P^i_{pq} that represents the probability of occurrence of word q with p and is represented as follows:

$$P^i_{pq} = P^i(p|q) = \frac{N^i_{pq}}{N^i_p} \tag{10}$$

where, P^i_{pq} is the relative probability of occurrence of event N^i_{pq} in relevance to N^i_p. Based on the formulation, the ratio of co-occurrence probabilities is presented as follows:

$$\Psi\left(W^i_p, W^i_q, \bar{W}^i_r\right) = \frac{P^i_{pr}}{P^i_{qr}} \tag{11}$$

where Ψ denotes a function that determines the co-relation of probe word with the word W^i_p and W^i_q. $W^i \in V_M$ are WVs from ith article from D_t, and $\bar{W}^i_r$ is a probe word in which with reference the co-relation would be determined, and P^i_{qr} is the probability of co-occurrence between word W^i_q and W^i_r.

With respect to the probe word, the given ratio can be of any value determining the towardness of co-relation. If the ratio is greater than $\bar{W}^i_r$, it is more co-related to W^i_{qr}. So, the ratio provides us with the relation between three words.

The model can be developed for function Ψ, taking into mind the concept of linearity for embedding WV W^i as

$$\Psi\left(\left(W^i_p - W^i_q\right)^T \bar{W}^i_r\right) = \frac{P^i_{pr}}{P^i_{qr}}. \tag{12}$$

This relation is also described as the difference and similarity of word embeddings for the variables in Ψ. Also, the relation is symmetrical, and hence (12) can be presented as follows:

$$\Psi\left(\left(W^i_p - W^i_q\right)^T \bar{W}^i_r\right) = \frac{\Psi\left({W^i_p}^T \bar{W}^i_r\right)}{\Psi\left({W^i_q}^T \bar{W}^i_r\right)}. \tag{13}$$

Over the period of time T, all the WV embeddings have maintained linearity, and to satisfy the condition, $\Psi(x)$ is defined as an exponential function as follows:

$$\Psi(x) = \exp(x). \tag{14}$$

From (12) to (13), we obtain as follows:

$$\Psi\left({W^i_p}^T \bar{W}^i_r\right) = P^i_{pr} = \frac{N^i_{pr}}{N^i_p}. \tag{15}$$

Since $\Psi(x) = \exp(x)$

$${W^i_p}^T \bar{W}^i_r = \log\left(P^i_{pr}\right) = \log\left(N^i_{rk}\right) - \log\left(N^i_p\right) \tag{16}$$

$${W^i_p}^T \bar{W}^i_r + b^i_p + \left(\bar{b}^i_r\right) = \log\left(N^i_{rk}\right) \tag{17}$$

where $\log(N^i_p)$ is then converted into constant term as N^i_p is not dependent on N^i_{rk}, and hence the term can be converted into two bias constant r and p, described by b to maintain the symmetrical requirements. Here W and b are defined as embedding vector matrices. Therefore, the long occurrence can be counted by the dot product of two embedding matrices. Based on this, in Q-GloVe methodology, we compute the mean square root to find the error between predicted and the original value of co-occurrence counts. Due to the unusual occurrence frequency of every pair, weight is assigned to re-adjust the cost of each word pair. Thus, we introduce a weighing function $f(N^i_{pq})$ into the cost function in the model as follows:

$$J = \sum_{p,q=1}^{V} f\left(N^i_{pq}\right)\left({W^i_p}^T \bar{W}^i_r + b^i_p + \bar{b}^i_r - \log\left(N^i_{rk}\right)\right)^2 \tag{18}$$

where V is the vocabulary size, the properties of the weighting function can be defined as follows.

1) $f(0) = 0$. If the function can be defined as continuous, the conversion of $x \rightarrow 0$ should be fast enough to converge according to limiting factor defined by
$$\lim_{x\rightarrow 0} f(x)\log^2 x.$$
2) The word that has rare occurrence, and thus should not be over-weighed so that $f(x)$ should be non-decreasing.
3) Frequent overweight items should not be over-weighed so that $f(x)$ should be relatively small for large x.

Based on the defined properties of weighting function, the function f is defined as follows:

$$f(x) = \begin{cases} (x/x_{\max})^a, & \text{if } x < x_{\max} \\ 1, & \text{otherwise.} \end{cases} \tag{19}$$

The final representations formalize the significant linear substructures of the WV space, $\mathrm{DB_{ve}}$.

$$\mathrm{DB_{ve}} \xleftarrow{\text{GloVe}} \mathrm{DB} \tag{20}$$

where $\mathrm{DB_{ve}}$ is then split into training, validation, and testing datasets. The tuples in training are used to train the model according to Section III-B. Once we obtain the results, the information is stored in the BC network, so prediction results do not tamper. The details of the Bi-LSTM model and BC network specifications are presented as follows.

B. Bi-Directional LSTM Model

In traditional neural networks, the feed input and predicted output are independent of each other. In the fake news classification problem, we have to perform the prediction based on previous training samples. Therefore, there is a requirement to retain the previous information through hidden layers. Thus, we prefer LSTM over the RNN model, storing information in hidden layers.

In SaTYa, we prefer a bi-directional LSTM model that allows access to both the backward and forward information about the sequence at every iteration step [28]. In general, we emulated the feed-forward linear sequence through the forward layer and the back-propagation mechanism to move from data at time $t + k$ to time t. This allows the scheme to identify the semantic context with more accuracy and allows optimization of linear weights if required. Thus, one layer presents the information as forwarding propagation, i.e., from the past to the future, and the other presents the backward propagation layer from future to past. The bi-directional propagation layers are required to understand the semantic context in a better fashion, and the information can also be preserved for future access and alterations if required. However, on the downside, to maintain the bi-directional access, the LSTM

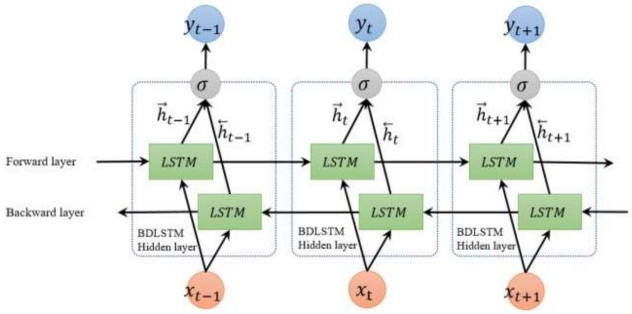


Fig. 3. Bi-directional LSTM design used in the DL model.

network requires twice the memory space to manage the weight and bias parameters, compared to the uni-directional LSTM architecture. Fig. 3 presents the details of the model.

In the Bi-LSTM network, the model has loops that take in sequential inputs. The independent activation functions are converted into dependent activation by assigning weights and bias to all the layers, and thus the complexity of increased parameters is reduced. Each previous output is memorized as input in the next hidden layer. For the same, in the Bi-LSTM network, every LSTM cell has a memory unit, denoted as $\text{ME}_t \in V_M$ at a particular time t instead of a simple RNN unit. ME_t works with V_M, where M denotes the vector dimension. The model has the basic three gates that handle the flow of data in the cell, namely, the input Gate $I_t \in V_M$ for inputs, the forget gate $F_t \in V_M$, and the output Gate $O_t \in V_M$ that conditionally displays the output from the memory unit.

For information transfer, two more activation functions are defined, the sigmoid function σ, and the hyperbolic tanh function. σ is defined over the normalized range [0, 1]. When the value of σ is 0, it discards the whole information of the gate, and if the value is 1, it allows the information to pass through the gate. To solve the problem of gradient descent and allow the information to be stored between cells, we define the tanh function. A WV space W is taken at the input gate I_t for the model's training. WV represents a restructured WV space of news article, as described in Section III-A, and it is then normalized into the cells of the Bi-LSTM model through the σ function of I_t. The dataset DB_v is passed through a dataset splitter which splits it into three parts, the training dataset, the validation dataset, and the testing dataset. A weight matrix is assigned as a hyper-parameter to optimize for better accuracy between I_t and hidden layers, defined as $\text{WM}_{I\to H}$, and subsequently between the hidden layers and output gates, defined as $\text{WM}_{H\to O}$

$$\begin{aligned} I_t &= \delta(\text{WM}_{H\to O}) \\ i_t &= f(\text{WM}_{I\to H}\cdot W + Qi_{t-1}) \end{aligned} \tag{21}$$

where I_t defines the hidden layer as a function of the previous input state, i.e., I_{t-1}, that denotes the values of the inputs from the $t-1$ layer. The output gate O_t is fully connected, and hidden layer I_t is recurrent due to weight matrix $W_{I\to H}$. The process is then iterated for z times and is denoted as follows:

$$\begin{aligned} I_t &= \delta(\text{WM}_{H\to O}) \\ I_t &= \delta\text{WM}_{H\to O}\cdot i_t \\ I_t^1 &= \delta(\text{WM}_{H\to O}\cdot(f(W_{I\to H}\cdot W + Q\cdot i_{t-2})) \\ I_t^2 &= \delta(\text{WM}_{H\to O}\cdot(f(\text{WM}_{I\to H}\cdot W + Q\cdot i_{t-3}))) \\ &\vdots \\ I_t^z &= \delta(\text{WM}_{H\to O}\cdot(f(\text{WM}_{I\to H}\cdot W + Q\cdot i_{t-z})) \end{aligned} \tag{22}$$

After the z iterations, the LSTM component processes the information from I_{t-1}, and sets the forget gate component as follows:

$$F_t = \sigma(Q_f\cdot[I_{t-1}\cdot W + b_f). \tag{23}$$

The above process represents the in-depth analysis of the LSTM hidden layer process. To ensure higher semantic training and context information, the Bi-LSTM model hidden layer is then adopted as follows:

$$I_t = \overrightarrow{I_t} \oplus |\overleftarrow{I_t}| \tag{24}$$

where $\overrightarrow{I_t}$ and $\overleftarrow{I_t}$ denote the hidden state forward and backward movements defined at a given time t. These two hidden layers can be defined as follows:

$$\overrightarrow{I}_t = f(\overrightarrow{Q}x_t + \overrightarrow{V}\overrightarrow{I}_{t-1} + \overrightarrow{b}) \tag{25}$$

and

$$\overleftarrow{I}_t = f(\overleftarrow{Q}x_t + \overleftarrow{V}\overleftarrow{I}_{t+1} + \overleftarrow{b}). \tag{26}$$

The final classification result is produced through the (25) and (26), that produces $\hat{y}$ as follows:

$$\hat{y} = g(Uh_t + c) = g(\text{WM}[\overrightarrow{I}_t;\ \overleftarrow{I}_t] + c). \tag{27}$$

Algorithm 1 presents the entire process of the proposed Bi-LSTM model for fake news classification. In the proposed algorithm, we consider input WV of length k. Lines 1–8 first normalize the input between min and max and divide the vector space into the train, test, and validation. Then lines 10–25 presents the gate functionalities of the model to classify the news data. Finally, lines 26–34 compute the fake news probability, denoted as P_f. The model is trained to output $\hat{y}$, and subsequently generate P_f, the probability value that determines whether the news is classified as fake or real news.

C. Smart Contract Execution

Once the output is obtained through $\hat{y}$, based on the probability P_f, we check the value of minimum threshold $P_{f_{\min}}$, and maximum threshold $P_{f_{\max}}$, to classify whether the news is fake or not. In case $P_f < P_{f_{\min}}$, we classify the news as fake, and the news is not added to the BC network. In case $P_f > P_{f_{\max}}$, the news is verified as valid, and then added to BC. If P_f is mid-range between $P_{f_{\min}}$ and $P_{f_{\max}}$, the classification has failed and is processed again at Layer 1. After the generation of P_f in Section III-B, the SC execution phase begins. In this layer, the SC decides whether the classified news is added to the BC scheme or not.

Once the transaction is added to BC, Algorithm 2 presents the methodology behind the execution of our SC for the classification of news. In Algorithm 2, the first sub-procedure generate the fixed-length address of news from varied sources and form the mapping of the news with the probability P_f.

Algorithm 1 SaTYa: Bi-Directional LSTM Model for News Classification

Input: WV
Output: Probability of the news being fake, PF

procedure SATYA_BILSTM_MODEL
 for $k \leftarrow 1$ to n **do**
  $A[k] \leftarrow WV_i$
  $NM \leftarrow Normalize(A[k], min, max)$
  $D[k] \leftarrow Initialize\_vector(V^M, A[k])$
  $DS_{tr} \leftarrow Split\_Train(DS[: e], A[k])$
  $DS_{va} \leftarrow Split\_Validation(DS[e+1 : f], A[k])$
  $DS_{te} \leftarrow Split\_Test(DS[f+1 :], A[k])$
 end for
 for $t \leftarrow 1$ to z **do**
  $R_t \leftarrow DS_{tr}$
  $I_t \leftarrow \delta(WM_{H \rightarrow O})$
  $i_t \leftarrow f(WM_{I \rightarrow H}.W + Qi_{t-1})$
  $F_t \leftarrow \sigma(Q_f.R_t + b_f)$
  $m \leftarrow Forget\_gate\_mapping(R_t, 0, 1)$
  if $(m == 0)$ **then**
   print ***"Fake News"***
   $M_s \rightarrow Update\_Dataset(w_t, \bar{C}_t, \alpha)$ as defined in eq. (27)
   $C_s \rightarrow Delete\_News(C_t, \alpha)$
   $O_t \rightarrow \sigma(Q_O.[M_t, C_t, \alpha] + b_O)$
   $w_t^O \rightarrow softmax(O_t)$
   $c = \sum_{t=1}^{l} w_t^O.h_t$
   $y = Sigmoid(O_t + c)$
  else if $(m == 1)$ **then**
   print ***"Not Fake News"***
   $M_s \rightarrow Update\_Dataset(w_t, \bar{C}_t, \alpha)$ as defined in eq. (27)
   $C_s \rightarrow Publish\_News(C_t, \alpha)$
   $O_t \rightarrow \sigma(Q_O.[M_t, C_t, \alpha] + b_O)$
   $wm_t^O \rightarrow softmax(O_t)$
   $c = \sum_{t=1}^{l} w_t^O.h_t$
   $y = Sigmoid(O_t + c)$
  else
   print ***"Testing Model is not up to date, Print -1 and exit"***
  end if
 end for
end procedure

Algorithm 2 SaTYa: Smart Contracts

Input: News in following formats: $\{E_{NW}||E_{SM}||E_{MF}||E_{EM}\}$
Output: Notification with probability, P_f

1: **procedure** SATYA_ENTER_STRING($\{E_{NW}||E_{SM}||E_{MF}||E_{EM}\}$)
2: $A_{ddr} \leftarrow GENERATE\_ADDRESS(\{E_{NW}||E_{SM}||E_{MF}||E_{EM}\})$
3: $P_f = procedure\ SaTYa\_BILSTM\_MODEL(A_{ddr})$
4: *Mapping*: $A_{ddr} \rightarrow P_f$
5: $return A_{ddr}$
6: **end procedure**

1: **procedure** SATYA_GET_RESULT(A_{ddr})
2: $P_f = Mapping[A_{ddr}]$
3: **if** $(P_f >= P_{fmax})$ **then**
4:   print **"True news."**
5: **else if** $(P_f <= P_{fmin})$ **then**
6:   print **"False news."**
7: **else**
8:   print **"Testing model is not up to date."**
9: **end if**
10: **end procedure**

TABLE I
SIMULATION PARAMETERS FOR BI-LSTM MODEL

Simulation Parameters	
No. of nodes	256
No. of validators	24
Total news articles	45000
No. of Sources	78
Epochs	1000
Max_Seq_Length	144
Batch_size	32
dim_length	16
Optimizer	Adam
Loss	Mean Squared Error
Dropout	0.322
Neuron_activation	128
Activation Function	Sigmoid

Lines 1–5 form the mapping of SC with the generated P_f. Once A_{ddr} is generated, we compare it with threshold probability P_{th}, depicted in the second sub-procedure. News is taken in as input in one of the given formats (E_{NW}, E_{EM}, E_{MF}, E_{EM}) and an address is generated for it. The news is fed to the Bi-LSTM model that generates P_f, which is mapped correspondingly with the address, A_{ddr}.

IV. PERFORMANCE EVALUATION

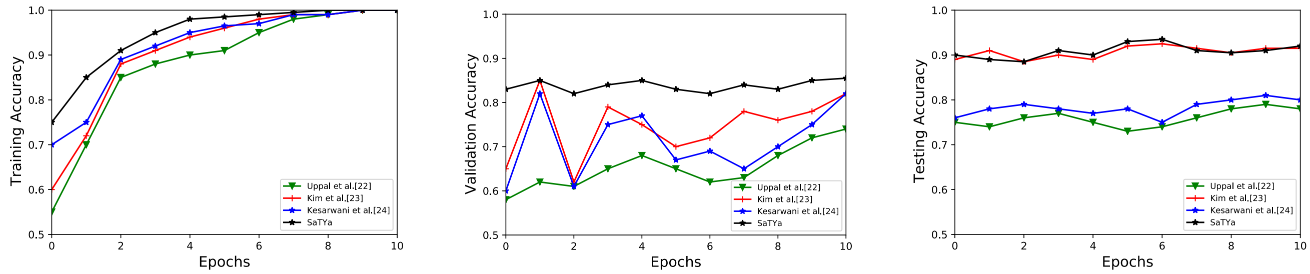
In this section, we present the performance evaluation of our proposed framework, SaTYa. For the evaluation purpose, we have prepared our real-time dataset consisting of 45 000 records from various sources like Google News, Media websites, messages, and emails (in the form of tweets and posts). We have fed them to our training and testing model. The target is to classify the news as fake or real news. The data was scraped using Beautiful Soup, Selenium, and Scrapy. Several hyper-parameters are given in Table I along with their values, like the number of nodes, validators, sources, and many more. This table can be used to map with the proposed model SaTYa for thorough understanding.

Once the dataset was collected, the primary goal was to clear, filter, and pre-process the dataset. For the same, we have focused on cleaning and filtering unnecessary entries. We focused on stemming and lemmatization, removing stop words, hyper-links, redundant and unnecessary words, and sentences. This was followed by normalization for transformation into a canonical form. Through the above process, a more consistent dataset was obtained, and anomalies were minimized. The cleaned dataset contains a total of 38 729 entries, where fake news constituted 17 903 rows, and 20 826 rows were true news. At last, we also removed the null entries from the dataset to improve our accuracy. Lastly, we also removed the NULL entries from our dataset. This eventually resulted in higher efficiency and better accuracy. Parts of speech tagging, shallow parsing, and named entity recognition were employed to derive useful information.

A. Simulation Results

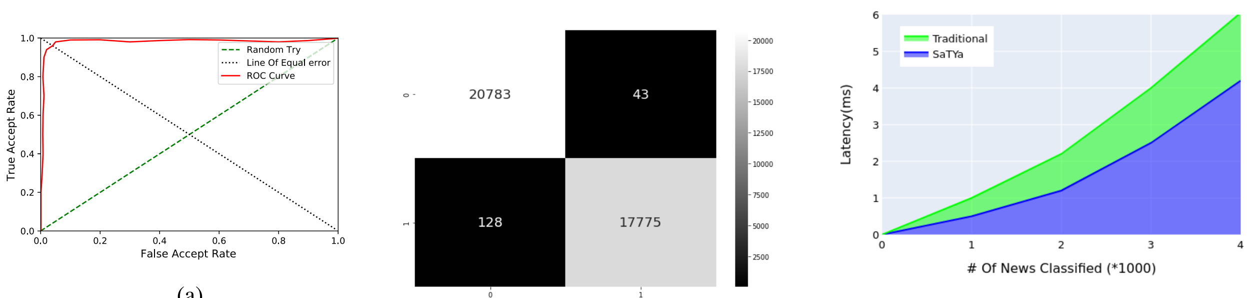
In the simulation results, we consider the results of the Bi-LSTM classifier model that generates a value P_f that differentiates the news as fake or real news from the considered dataset. The obtained results are presented in Table II.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.



(a) (b) (c)

Fig. 4. Comparative analysis of Training, validation, and testing accuracy. (a) Comparative analysis of accuracy for training dataset. (b) Comparative analysis of accuracy for validation dataset. (c) Comparative analysis of accuracy for testing dataset.



(a) (b) (c)

Fig. 5. Receiver operating characteristic (ROC) curve and area under the curve (AUC), confusion matrix, and scalability of mining latency of classified news in BC of the proposed scheme. (a) Analysis of ROC and AUC. (b) Confusion matrix. (c) Scalability of mined transactions.

TABLE II

PERFORMANCE METRICS BASED ON CONFUSION MATRIX VALUES IN SATYA

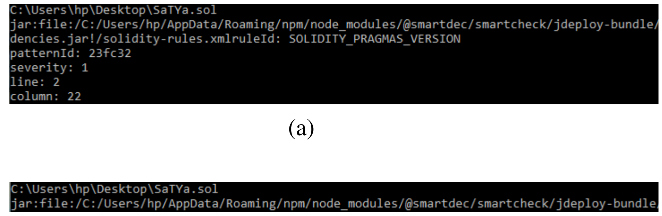
Performance metrics	
Accuracy	.9955
Error rate	.02500
Precision value	.9938
True Negative Rate	.94339
False Positive Rate	.05661
Recall	.9979
F-measure	.9958

1) Bi-LSTM Model-Performance Analysis: We present the performance analysis plots of our proposed approach against Uppal *et al.* [29], Kim and Jeong [30], and Kesarwani *et al.* [31] in terms of training, validation, and testing accuracy of dataset. In Uppal *et al.* [29], a discourse-structure analysis is performed to differentiate between fake and real news, and the baseline model achieved 74% accuracy. Kim and Jeong [30] presented a semantic-rule-based bidirectional multi-perspective matching for natural language sentence framework. Still, the model suffered from the limitation of the longer length of the input sentence. They overcame the limitation through entity-based matching and obtained an accuracy of 0.9362 on the test dataset, with an AUC value of 0.782. In Kesarwani *et al.* [31], K-nearest neighbor classification model is employed, and an accuracy of 79% is achieved.

In the proposed framework SaTYa, a Bi-LSTM model is preferred. Due to the movement in both forward and backward directions, the method eliminates the limitations of the vanishing gradient problem. Thus a higher accuracy is obtained in comparison to the previous state-of-the-art models. The results of the same are demonstrated in Fig. 4. In Fig. 4(a), we present the details of the training accuracy of our proposed model against the aforementioned schemes. It is evident that at six epochs, the training accuracy of the model is close to 0.9878, compared to 0.9655 in Kim *et al.*, 0.9434 in Kesarwani *et al.*, and 0.9343 in Uppal *et al.* Similarly, we obtain a higher validation accuracy in the proposed approach, as evident in Fig. 4(b). The proposed approach has a consistent validation accuracy of >0.8 throughout. Fig. 4(c) presents the testing accuracy of the proposed scheme. The propose scheme has a low error rate of 0.02500, with a precision of 0.9938, and an accuracy of 0.9955, that is higher than 0.9362 in Kim *et al.*, 0.791 in Kesarwani *et al.*, and 0.74 in Uppal *et al.* As the vanishing gradient problem is resolved, the Bi-LSTM model converges faster than the conventional schemes, and the learning rate is improved.

2) Analysis of Operating Characteristics and Mining Latency of Added Transactions in BC: Next, we present the results of the different operating characteristics of the proposed scheme and the mining latency of added verified transactions in BC. Fig. 5 presents the details. In Fig. 5(a), we represent the performance of the classification model presented at all the different classification thresholds. The receiver-operating characteristics (ROC) curve indicates that the proposed scheme has more true positive hits in identifying fake news than false positives. As the Bi-LSTM updates the probability value P_f at

Fig. 6. SaTYa: Functionalities of Smart Contract. (a) Add scheme. (b) Get scheme. (c) Add scheme. (d) Get scheme.



(b)

Fig. 7. Security analysis of the proposed SaTYa Smart contract with and without vulnerabilities. (a) Smart contract of SaTYa scheme with vulnerabilities. (b) Smart contract of SaTYa scheme without vulnerabilities.

minimal latency due to bi-directional movements, it results in a higher area under the curve (AUC), which provides us with the collective measure of performance at all the different classification thresholds. The value obtained is ≈0.987. Fig. 5(b) presents the details of the confusion matrix, and an accuracy of 0.9955 is achieved of the proposed framework. The number of false negatives is 43, and false positives are 128. The achieved precision value is 0.9938, recall value of 0.9979, and we obtain an F-score of 0.9958. Finally, we present the analysis of the mining latency of the added verified news transactions in BC to emphasize the scalability of the proposed scheme. Fig. 5(c) shows the mining latency against traditional approaches by authors in [29]–[31]. The proposed approach has higher scalability than the model used through the Bi-LSTM, which stores the value of P_f in itself, and thus reduces the processing time. As the number of transactions increases, more transactions are appended in the same unit time, and therefore more news can be accommodated in the BC network. Thus, in real-time, more users are serviced through the proposed scheme.

B. Functionalities of SC

In this subsection, we present the details of the various SC functionalities. Fig. 6 presents the details. The process starts with the execution of SC for verified news functionalities with news stakeholders, where the stored valid information is accessed through the input address. Fig. 6(a) specifies the address details where the information access address is entered, based on the outputs of the Bi-LSTM model. In Fig. 6(b) in the event of true classification, the data is appended in the BC network and added to other nodes through proposed consensus. It forms a trust among all the stakeholders about the authenticity of the classified information. Fig. 6(c) presents the details from the classification model that signifies that the entered news is true, and Fig. 6(d) signifies that the entered news from the classification model is a piece of fake news. Fake news is disposed, and not added to the BC network.

C. Cost Evaluation

For cost analysis, we present the bytes exchange cost, which mainly depends on the cost of the Bi-LSTM model, and the block proposal during SC deployment. As depicted in Algorithm 1, we consider the sequence length k, and we run it for n epochs. The WVs are stored in cell-array $A[k]$, and the feed-forward model run sequentially for z times. The LSTM input I_t involves three weights, each of 1 bit. Thus, the cost involved 3 bits. At the forget gate F_t, there is a σ operation of 1 bit, and the state transitions would involve the sequence length z. The Forget_gate_mapping would be 1 bit. The storage requirement would depend on weights w_i. Thus the storage cost would be $z \cdot W_i$. If we consider $w = 2$, and we max_sequence_length of 144, the cost involved would be 288 bits. As per the values, the involved cost is ≈37 bytes for LSTM feed-forward proposal. In the back-propagation model, a feedback unit is designed for parameter weight optimization and would involve the same operations. Thus, the cost is ≈74 bytes.

Next, we evaluate the cost of Algorithm 2. At input, we collect news from different sources and form the string concatenation of q bits. Next, we generate a 32-byte hash output of q and is passed as input to the Bi-LSTM model, which involves a cost of $z \cdot w_i$. The mapping $A_{\mathrm{ddr}} \rightarrow P_f$ would be 1 bit. In the second sub-procedure, the mapping is fetched, and the comparison would exchange 1 bit. For block proposal, the metadata involves a timestamp of 32 bits and the previous hash value of 32 bytes. The cryptographic nonce is 64 bytes, and the merkle_root address is 32 bytes. Thus, the total involved cost is ≈162 bytes. Thus, the overall communication cost is ≈236 bytes.

D. Security Evaluation

In this subsection, the security analysis of SCs has been performed using the smartcheck security tool for the proposed SaTYa scheme. We have implemented the solidity source code in the Remix IDE Ethereum platform to scrutinize the security-related issues in the proposed SaTYa scheme. Fig. 7(b) shows the validation of the source code of the SaTYa scheme using a smart check tool that does not contain any vulnerability or severity. But, Fig. 7(a) shows that the smart check tool yields one severity in the source code by default [32]. Therefore, we have eliminated that particular severity as depicted in Fig. 7(b) leading to a secure SaTYa scheme without any vulnerability.

V. Conclusion

In modern smart communities, fake news circulation has grown exponentially due to the rise of user interactions among social network platforms. Thus, the authenticity and credibility of the news source are required to be verified. Recently, researchers have proposed various ML and DL techniques to classify the news as real or fake. In a similar direction, the authors propose a scheme, named SaTYa, that integrates the Bi-LSTM model and BC network to classify the news as real or fake. Once the news is classified, the results are appended in the BC network, which ensures that an adversary does not alter the classification of the news. Moreover, the authors propose the usage of SC that fetches the news source origin based on the stored hash address value. This allows users to judge the news credibility, the source of the news, and verify the authenticity of the posted feeds on social platforms. The scheme is compared against the semantic-rule-based approach, bi-directional multi-perspective matching, and K-nearest neighbor classification models. The presented Bi-LSTM model assures movement of data in both forward and backward propagation paths, which optimizes the training weights in less time, and bias is minimized. The obtained results indicate higher accuracy than conventional frameworks and allow meta-classified information to be stored as transactions in BC ledgers, which reduces the mining latency and increases the scalability of the BC network.

In the future, the authors would like to improve the hyper tuning of the model's regularization parameters, which would reduce the model complexity and improve the loss functions. This ensures faster classification and convergence, further improving the model response against real-time fed data as inputs to the model.

References

[1] S. Leonardi, G. Rizzo, and M. Morisio, "Automated classification of fake news spreaders to break the misinformation chain," *Information*, vol. 12, no. 6, p. 248, Jun. 2021.

[2] 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.

[3] K. L. Tan, C. Poo Lee, and K. M. Lim, "FN-Net: A deep convolutional neural network for fake news detection," in *Proc. 9th Int. Conf. Inf. Commun. Technol. (ICoICT)*, Yogyakarta, IN, USA, Aug. 2021, pp. 331–336.

[4] M. Choudhary, S. S. Chouhan, E. S. Pilli, and S. K. Vipparthi, "BerConvoNet: A deep learning framework for fake news classification," *Appl. Soft Comput.*, vol. 110, Oct. 2021, Art. no. 107614.

[5] P. N. Sureshbhai, P. Bhattacharya, and S. Tanwar, "KaRuNa: A blockchain-based sentiment analysis framework for fraud cryptocurrency schemes," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC Workshops)*, Dublin, Ireland, Jun. 2020, pp. 1–6.

[6] S. Paul, J. I. Joy, S. Sarker, A. A. 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)*, Sarawak, Malaysia, 2019, pp. 1–5.

[7] P. Fraga-Lamas and T. M. Fernandez-Carames, "Fake news, disinformation, and deepfakes: Leveraging distributed ledger technologies and blockchain to combat digital deception and counterfeit reality," *IT Prof.*, vol. 22, no. 2, pp. 53–59, Mar. 2020.

[8] S. K. Singh, P. K. Sharma, Y. Pan, and J. H. Park, "BIIoVT: Blockchain-based secure storage architecture for intelligent Internet of Vehicular Things," *IEEE Consum. Electron. Mag.*, early access, Jun. 24, 2021, doi: 10.1109/MCE.2021.3089992.

[9] U. Bodkhe, D. Mehta, S. Tanwar, P. Bhattacharya, P. K. Singh, and W.-C. Hong, "A survey on decentralized consensus mechanisms for cyber physical systems," *IEEE Access*, vol. 8, pp. 54371–54401, 2020.

[10] S. K. Singh, A. E. Azzaoui, T. W. Kim, Y. Pan, and J. H. Park, "DeepBlockScheme: A deep learning-based blockchain driven scheme for secure smart city," *Hum.-Centric Comput. Inf. Sci.*, vol. 11, p. 12, Mar. 2021.

[11] S. B. Patel, P. Bhattacharya, S. Tanwar, and N. Kumar, "KiRTi: A blockchain-based credit recommender system for financial institutions," *IEEE Trans. Netw. Sci. Eng.*, vol. 8, no. 2, pp. 1044–1054, Apr. 2021.

[12] K. Akyol and B. Sen, "Modeling and predicting of news popularity in social media sources," *Comput., Mater. Continua*, vol. 61, no. 1, pp. 69–80, 2019.

[13] D. K. Vishwakarma, D. Varshney, and A. Yadav, "Detection and veracity analysis of fake news via scrapping and authenticating the web search," *Cognit. Syst. Res.*, vol. 58, pp. 217–229, Dec. 2019.

[14] A. Dey, R. Z. Rafi, S. Hasan Parash, S. K. Arko, and A. Chakrabarty, "Fake news pattern recognition using linguistic analysis," in *Proc. Joint 7th Int. Conf. Informat., Electron. Vis. (ICIEV)*, Kitakyushu, Japna, Jun. 2018, pp. 305–309.

[15] C. Zhang, A. Gupta, C. Kauten, A. V. Deokar, and X. Qin, "Detecting fake news for reducing misinformation risks using analytics approaches," *Eur. J. Oper. Res.*, vol. 279, no. 3, pp. 1036–1052, Dec. 2019.

[16] K. Xu, F. Wang, H. Wang, and B. Yang, "Detecting fake news over online social media via domain reputations and content understanding," *Tsinghua Sci. Technol.*, vol. 25, no. 1, pp. 20–27, Feb. 2020.

[17] L. Yin, X. Meng, J. Li, and J. Sun, "Relation extraction for massive news texts," *Comput., Mater. Continua*, vol. 60, no. 1, pp. 275–285, 2019.

[18] R. K. Kaliyar, A. Goswami, P. Narang, and S. Sinha, "FNDNet—A deep convolutional neural network for fake news detection," *Cognit. Syst. Res.*, vol. 61, pp. 32–44, Jun. 2020.

[19] T. Traylor, J. Straub, and N. Snell, "Classifying fake news articles using natural language processing to identify in-article attribution as a supervised learning estimator," in *Proc. IEEE 13th Int. Conf. Semantic Comput. (ICSC)*, Newport Beach, CA, USA, Jan. 2019, pp. 445–449.

[20] Y. Ishida and S. Kuraya, "Fake news and its credibility evaluation by dynamic relational networks: A bottom up approach," *Proc. Comput. Sci.*, vol. 126, pp. 2228–2237, Jan. 2018.

[21] P. Bahad, P. Saxena, and R. Kamal, "Fake news detection using bi-directional LSTM-recurrent neural network," *Proc. Comput. Sci.*, vol. 165, pp. 74–82, Jan. 2019.

[22] F. A. Ozbay and B. Alatas, "Fake news detection within online social media using supervised artificial intelligence algorithms," *Phys. A, Stat. Mech. Appl.*, vol. 540, Feb. 2020, Art. no. 123174.

[23] N. X. Nyow and H. N. Chua, "Detecting fake news with tweets-T properties," in *Proc. Conf. Appl., Inf. Netw. Secur. (AINS)*, Pulau Pinang, Malaysia, 2019, pp. 24–29.

[24] Y.-C. Ahn and C.-S. Jeong, "Natural language contents evaluation system for detecting fake news using deep learning," in *Proc. 16th Int. Joint Conf. Comput. Sci. Softw. Eng. (JCSSE)*, Chonburi, Thailand, Jul. 2019, pp. 289–292.

[25] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," in *Proc. NAACL-HLT*, 2019, pp. 1–16.

[26] Z. Shae and J. Tsai, "Ai blockchain platform for trusting news," in *Proc. IEEE 39th Int. Conf. Distrib. Comput. Syst. (ICDCS)*, Dallas, TX, USA, Dec. 2019, pp. 1610–1619.

[27] Y. He and X. Dong, "Real time speech recognition algorithm on embedded system based on continuous Markov model," *Microprocessors Microsyst.*, vol. 75, Jun. 2020, Art. no. 103058.

[28] L. Xiang, G. Guo, Q. Li, C. Zhu, J. Chen, and H. Ma, "Spam detection in reviews using LSTM-based multi-entity temporal features," *Intell. Autom. Soft Comput.*, vol. 26, no. 4, pp. 1375–1390, 2020.

[29] A. Uppal, V. Sachdeva, and S. Sharma, "Fake news detection using discourse segment structure analysis," in *Proc. 10th Int. Conf. Cloud Comput., Data Sci. Eng. (Confluence)*, Noida, India, Jan. 2020, pp. 751–756.

[30] K.-H. Kim and C.-S. Jeong, "Fake news detection system using article abstraction," in *Proc. 16th Int. Joint Conf. Comput. Sci. Softw. Eng. (JCSSE)*, Chonburi, Thailand, Jul. 2019, pp. 209–212.

[31] A. Kesarwani, S. S. Chauhan, and A. R. Nair, "Fake news detection on social media using K-nearest neighbor classifier," in *Proc. Int. Conf. Adv. Comput. Commun. Eng. (ICACCE)*, Las Vegas, NV, USA, Jun. 2020, pp. 1–4.

[32] A. López Vivar, A. T. Castedo, A. L. Sandoval Orozco, and L. J. García Villalba, "An analysis of smart contracts security threats alongside existing solutions," *Entropy*, vol. 22, no. 2, p. 203, Feb. 2020.