Multimedia Tools and Applications (2024) 83:82089–82106
https://doi.org/10.1007/s11042-024-18168-1

Detection on early dynamic rumor influence and propagation using biogeography-based optimization with deep learning approaches

R. Amutha[1]

Received: 31 July 2023 / Revised: 4 October 2023 / Accepted: 3 January 2024 /
Published online: 12 March 2024

Abstract
The far and wide distribution of innumerable rumors and fake news have been a serious threat to the truthfulness of microblogs. The earlier works have frequently aimed at remembering the earlier state with n consideration to the next context information. Also, a majority of the works before have made use of conventional feature representation approaches preceding a classifier. In this research, we evaluate the rumor detection problem by examining multiple Deep Learning approaches, with a focus on forward and backward direction analysis. The proposed technique incorporates Optimal Bidirectional Long Short-Term Memory and Convolutional Neural Network in order to correctly classify tweets as rumor or non-rumor. Then the Biogeography-based optimization (BBO) provides recommendations for fine-tuning the Bi-LSTM-CNN model's hyperparameters. According to the results of the experiments, the suggested technique is more precise than conventional methods, with an accuracy of 86.12%. The statistical analysis further demonstrates that the suggested model is much more successful than the appropriate alternatives.

Keywords Convolutional Neural Network · Dynamic Rumor Influence · Propagation Detection Model · Bidirectional Long Short-Term Memory and biogeography-based optimization

1 Introduction

The fast-paced evolution of social media sites and the exponential rise in social media content, user-created messages are very easy to get to a wide number of masses. This capability for quick and far-fetched distribution of information in social media puts forward unexpected problems in the commitment towards information quality and management. According to current National Crime Records Bureau (NCRB) statistics, incidents of spreading "false/fake news" and falsehoods, a felony under the Indian Penal Code,

✉ R. Amutha
amutharajan@gmail.com

[1] Assistant Professor, Department of Computer Science, PSG College of Arts & Science, Coimbatore, India

increased by over three times in 2020 compared to 2019 [1, 2]. The pandemic year has seen an increase of 214 percent in the number of instances of false news that were reported, which is much higher than the number of cases reported in 2019 (486) and in 2018 (280), the year in which the category was introduced for the first time. Many of the social media sites, who are concerned about information quality presently measure the information trustworthiness and identify rumors through manual means [3]. This method cannot be scaled economically or efficiently to handle the massive amounts of messages seen in today's typical social media environment.

The deficit experienced, regretfully enables the rumors to easily spread to a huge population quite rapidly, resulting in increased societal hazards and considerable harms. As a consequence, it is crucial and beneficial for society for automated rumors detection to be used in social media. Even though several existing studies highlighted on Twitter, in this research, the well-known Chinese social media site–Sina Weibo [4] is taken to be due to three main factors, the experimental social media setting. Sina Weibo is initially the most popular and prominent social media network in China, as well as the largest microblog service website. On Sina Weibo's homepage, a dedicated section contains all of the recognized rumors. About two messages are labeled as rumors every day, according to self-posted statistics from Sina Weibo. Before to being officially refuted, most of those rumors had already received hundreds of retweets. This large number of retweets considered is an indicant that the propagation of rumors is an important concern in Sina Weibo in spite of the site's earnest effort to deal with the incidents [5].

Using a person's record, the doctors can evaluate the body mass index (BMI) and risks to health with improved accuracy and precision. Opinions and perspectives on the real incidents can be shared very easily taking the aid of the Internet [6]. This study demonstrates the effect analysis, representation, comparison and comprehensive evaluation of the present condition of techniques, technologies, tools to filter the danger that information corruption imposes. But, early identification of rumors are few huge challenges, which still need to be addressed and require more exploration [7]. Now [8] this highly acclaimed study is examined from three perspectives in order to thoroughly sort the research situation of rumor detection from several aspects: Feature Selection, Model Structure, and Research Methodologies are the three main topics covered here. In terms of the selection of features, the methods may be divided into three categories: the content feature, the structure of the rumor's distribution, and the social feature. The key drawback of these methods is that they cannot get the dynamic interaction information of postings across a range of time periods since they are all reliant on the traditional feature-based models. Generally, dynamic distribution forms of rumor and non-rumor can be easily distinguished when compared to those of the final static forms.

The primary focus of the prevalent studies is on three wide classes of misinformation, this includes the identification of incorrect information, fake news, and rumours. In response to the above concerns, the following is provided: a detailed examination of automatic misinformation detection concerning incorrect information, rumours, fake news, spam, and disinformation. An effort is made to present a baseline study of misinformation detection (MID) in which deep learning (DL) is useful in the automatic processing of data and the construction of patterns to make judgments that execute the extraction of global features and produce superior outcomes [8]. Therefore, in this research, few efficient roles imparted by removing the misinformation on online Social Network using DL approaches is studied. Also, the effect, features, and detection of misinformation applying DL methods are focused. Hence, deep learning-based rumor detection is still a hot subject of research and requires extension in the future research work. In addition, the distribution

form yields important information, depicting the features of the distribution process. The propagation form is established by the postings' response correlation. Typically, a tree serves as the propagation form, with the root node being the tweet itself.

The primary contribution made by this research involves its introduction of an innovative according to a model of information dissemination for social media use cases with varied user characteristics. In this research, Optimal Bidirectional Long Short-Term Memory and Convolutional Neural Network in order to correctly classify tweets as rumor or non-rumor. Then the Biogeography-based optimization (BBO) provides recommendations for fine-tuning the Bi-LSTM-CNN model's hyperparameters. In a social media setting, the innovative information propagation model has the ability to define information dissemination patterns across diverse user subpopulations. One apparent application of the paradigm in the context of social media is the detection and differentiation of rumours and true messages. Using the Sina Weibo platform, a number of studies have been undertaken to examine the efficacy and advantages of using the revolutionary information propagation model for rumour identification.

The remaining sections of the research is structured as given: Section 2 provides a short review of the existing studies relevant to this research. The algorithmic technique for rumor identification in a social network (media) setting is addressed in Section 3 of the research. Section 4 shows the outcomes of the experiments carried out with the proposed technique for rumor detection is explained. At last, Section 5 provides the conclusion of this research.

2 Related work

In-depth research has been done on and presented in the literature on rumor detection in social media scenarios. Wang et al. [9] introduced by blocking a certain selection of nodes, a dynamic rumor influence minimization with user experience (DRIMUX) model may help lessen the impact that the rumor. On the basis of a real-world scenario, a dynamic Ising propagation model that takes into account both the rumor's worldwide popularity and individual interest is illustrated. But, practically, even if there are equal number of online friends for two users, their individual response to a rumor might be different due to their individual capabilities in manipulating the rumors.

Wu et al. [10] tried reducing the negative information like online rumors, which has gained massive focus. Virtually all of the rumor-blocking algorithms now in use have at least partially stemmed the spread of rumors, but their primary focus is on the network's microstructure, namely the pairwise associations or similarity between nodes, which were built entirely on social networks. A few networks may not want these methods' blocking effects because of the microstructure's limited data.

Hosni et al. [11] introduced the dynamic technique for rumour influence reduction in order to address these challenges and capitalise on their benefits (DARIM). The objective is to strike a balance between anti-rumor campaign strategies that limit rumour effect to the maximum degree possible and blocking nodes that prohibit rumour spread. In keeping with this, the survival theory is utilised to suggest a solution to a network inference problem that is constructed from a standpoint of network inference. But, due to the rumor distribution having the features of several propagation paths, rapid speed of distribution, extensive field of distribution and time-changing nature, improving the efficiency of the rumor blocking algorithm is a daunting task.

Hosni et al. [12] suggested combining the Truth Campaign Strategy (TCS) with the Nodes or Links Blocking Strategy (BNLS) to create this hybrid method. Its main objective is to determine the best combination of nodes for BNLS and TCS that will have the least negative impact on the system from rumors. Nevertheless, this method was created based on the full social network, emphasizing the network's microstructure, which includes pairwise correlations or similarity between nodes, which might have a blocking effect in certain networks owing to the problem of sparse data in the microstructure.

Ni et al. [13] recommended NP-Hard issues may be solved using the two-stage discrete gradient descent (TD-D) approach. In order to somewhat verify the suggested technique, a two-stage greedy (TG) algorithm is created. Finally, synthetic and authentic multi-network datasets are used to assess the effectiveness of the suggested strategies. Nevertheless, it ignores the impact of node correlation on the rumor-dissemination process.

Hu et al. [14] recommended an elaborate for the proactive and ongoing halting of the rumor spread in OSNs, the Hybrid Clustered Shuffled Frog-Leaping Algorithm-Particle Swarm Optimization (HCSFLA-PSO) method has been developed. First, a revolutionary trust technique of refuting rumors and an avant-garde depiction of trust degree are presented via the deconstruction of social interactions and the assessment of the degree of intimacy, self-reliability, and trustworthiness. But, it is still a crucial challenge to have an immediate and continuous blocking of the rumor distribution in OSNs, particularly in the times when the information technology has been rapidly developing.

Lin et al. [15] studied about a model, which extracts textual, distribution and structural information. The model includes three elements, which are Decoder, Encoder, and Detector. The encoder leverages a robust Graph Convolutional Network to take into consideration the input text and update the representation through propagation in order to learn the text and distribution data. The encoded form is then exploited by the subsequent decoder, which uses the AutoEncoder to gain information about the overall structure. But, it is a huge risk to social stability and safety of the public if the information is detected to be fake, misinforming or undesirable.

Yang et al. [16] investigated rumour detection using a graph attention capsule network (GACN) on dynamic propagation structures. To mine the rumor's deep-level characteristics, GACN consists of two components: a capsule-powered graph attention network that can encode static graphs into substructure classification capsules and a dynamic network architecture that can partition the rumour structure into multiple static graphs chronologically to capture dynamic interactive attributes during the development of a rumour. These two components are known collectively as the graph attention network. However, more study is necessary to identify the exact implications that each channel in the capsule discloses for illustrative purposes of the improvement in rumour detection.

Concerning the topic of rumour detection, Asghar et al. [17] investigated several Deep Learning models that lay a focus on considering the context of a text in both the backward and forward directions forward as well as backward. Using convolutional neural network and bidirectional long-short-term memory, the proposed system successfully divides tweets into rumour and non-rumor categories. These models fail, however, since they do not account for the characteristics of the structure of rumour dispersion.

Zojaji and Tork Ladani [18] demonstrated an innovative adaptive cost-sensitive loss function to learn Deep neural networks are used to balance out uneven stance data, which improves stance classifier performance in classes that aren't as prevalent. Essentially, the suggested loss function represents cross-entropy loss in a cost-sensitive approach. In contrast to the vast majority of previous cost-sensitive deep neural network models, the considered cost matrix is not explicitly stated: nonetheless, its adaptive tuning may be carried

out during the learning phase. The suggested model is not limited to the categorization of stances and is capable of empowering various classification techniques in the classification of imbalanced data.

Sailunaz et al. [19] examined a technique for the classification of Decision Tree and Logistic Regression were implemented on both tweet and user characteristics gathered from a common rumor dataset called "PHEME," to rank non-rumor and rumor tweets using a novel tweet and user feature ranking approach. For identifying the rumors, deep learning methods as well as supervised classification algorithms were used. However, rather than taking into account the global forwarding links between postings on social networks, the main goal of these algorithms is about understanding the serialization properties from a distribution standpoint.

Xu et al. [20] suggested using Hierarchically Aggregated Graph Neural Networks (HAGNN), a novel framework for Graph Neural Networks (GNN)-based rumor detection. The objective of this project is to acquire various granularities of high-level representations of text content and combine them with the pattern of rumour propagation. For learning the text-granularity representations of occurrences that are propagating, it applies a graph of rumor dispersion using a Graph Convolutional Network (GCN). Taking accuracy into consideration, the combination of an advanced text classification strategy with huge vectorization models may not be advantageous at this time.

Tembhurne et al. [21] presented a Mc-DNN model for the detection of fake news leverages the articles' headlines and articles' contents on altogether different channels. The Mc-DNN outperforms most of the fact-checking websites and deep models created by the various researchers. The performance is analyzed for Mc-DNN using the combination of CNN and RNN, CNN and GRU, CNN and LSTM, CNN and BiGRU, and CNN and BiLSTM. We found that the CNN and BiLSTM Mc-DNN is best for the task of fake news detection and achieve the highest accuracy of 99.23% and 94.68% on ISOT fake news dataset and FND, respectively.

Summary: Even though the above-mentioned approaches have gained the function of preventing the distribution of rumor data, most of them concentrate on the network's microstructure the paired correlation or similarity between nodes and ignore the topic content and distribution structure. This may lead to an undesirable blocking impact in few networks due to the data sparsity in microstructure, and in addition, still there is some improvement required in DL based learning process.

3 Proposed methodology

This study introduces an innovative method for the automated identification of rumours propagated in social media based on the fact that their dissemination patterns differ from those of truthful information. By accumulating user context-sensitive information from numerous profiles, the novel technique can forecast whether a message is a rumour or a genuine message and, if it is a genuine message, how likely it is to be retweeted. Using information propagation models proposed for rumour and credible message modes, the technique can determine whether a message is more likely to be a rumour or a credible one, based on an estimate of the likelihood that a network of information about the message will be disseminated. For the purpose of identifying rumours on social media, a neural model called OBi-LSTM-CNN is used in this study. Utilizes the text content and dynamic distribution patterns depicted in Fig. 1. The time at which each tweet was posted is used to

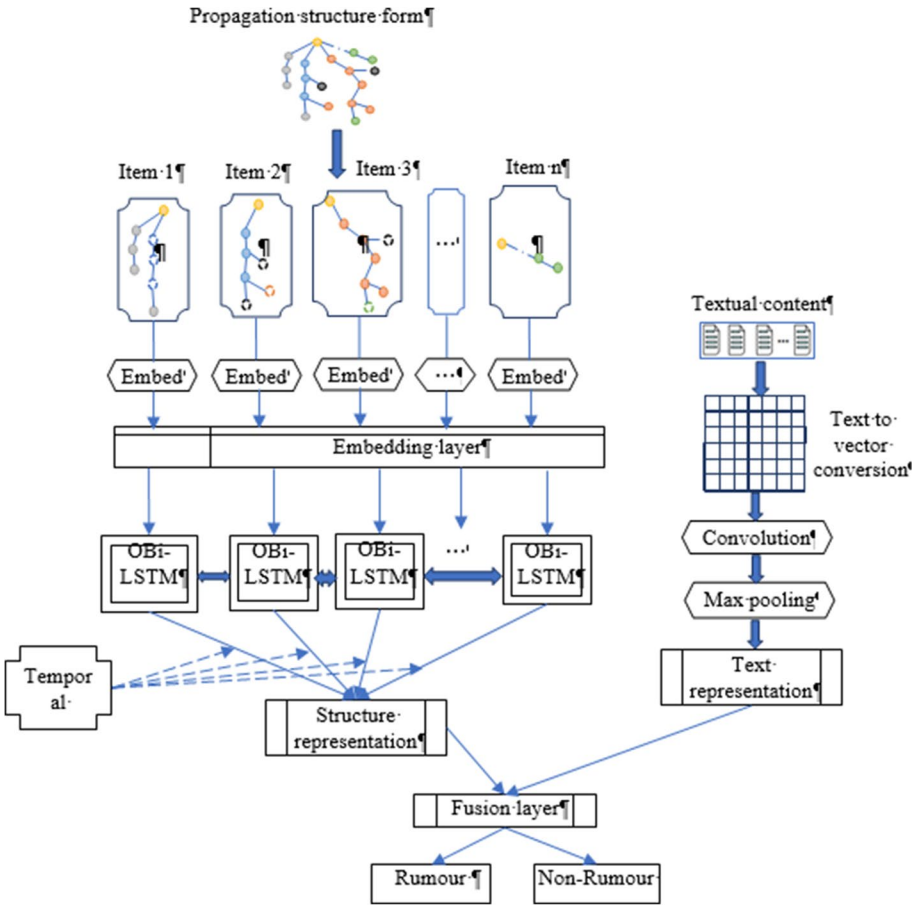

Fig. 1 An overview of the OBi-LSTM and CNN Framework

divide the propagation structure into smaller segments, and a Bi-LSTM model is employed to train a representation of the structure in order to extract dynamic information. In a subsequent phase, we will devise a temporal attention strategy to focus in on stages containing valuable structures. This model also contains propagation structure and content information for the purpose of identifying rumour sources. In research employing social media datasets, this model, which integrates dynamic propagation processes into the disinformation detection process, was found to be effective and reliable.

3.1 System model

In the proposed OBi-LSTM and CNN framework, the textual contents and propagation structures' initial representations are learnt using the structure network and content network correspondingly as illustrated in Fig. 1. The structure network employs the temporal attention strategy to highlight distribution stages having advantageous structures and divides the propagation structure according to the posting time. A content network assigns the postings' contents chronologically, and embedding vectors are trained for each content

containing semantic and sentiment data. Each tweet's publishing time is utilised to segment the propagation structure, which determines the distribution pattern for each occurrence. The tweets in each of the several time units that comprise the whole distribution period form a substructure. By encoding each substructure into a vector, the input of the OBi-LSTM is then created. Utilizing the temporal attention method to highlight time units with pertinent structures is proposed.

Propagation structure segmentation process: By segmenting the structure according to the posting time of each tweet, the dynamic propagation structure's change during the course of distribution is recorded. In particular, the total amount of time is broken up into a number of different time units, each of which has a certain time interval. The time units are ordered sequentially, represented as $\left[\mathfrak{tu}_1, \mathfrak{tu}_2, \ldots, \mathfrak{tu}_n\right]$, where $\mathfrak{tu}_i$ indicates the i-th time unit.

Propagation structure encoding and embedding process: Every time unit $\mathfrak{tu}$ has a sub-structure that the tweets published during this time unit form and every sub-structure is encoded and later embedded into a dense vector. It has to be observed that the optimization of the way structural features is represented during the training process can be achieved [22]. Particularly, the below three structural features are used in every time unit:

Feature 1: Every layer of the propagation tree's retweet-to-repost ratio represents the percentage of both actions j-th layer in the t-th time window to be $rr_{\mathfrak{tu},j}$.
Feature 2: The ratio of the post number of the j-th layer in the tu-th time unit and the percentage of post numbers in neighbouring layers, which both illustrate the characteristic of diffusion depth in the propagation process $\mathbb{N}_{\mathfrak{tu},j}$ and $\mathbb{N}_{(\mathfrak{tu}-1),j}$ is denoted as $\mathfrak{pn}_{\mathfrak{tu},j}$.
Feature 3: In the t-th time frame, the number of posts detected and the number of the j-th layer are given as follows $\mathfrak{np}_{\mathfrak{tu},j}$. After that, the information on the tu-th substructure is integrated as $x_t = embed\left(rr_{\mathfrak{tu}}, \mathfrak{pn}_{\mathfrak{tu}}, \mathfrak{np}_{\mathfrak{tu}}\right)$, where $rr_{\mathfrak{tu}}, \mathfrak{pn}_{\mathfrak{tu}}, \mathfrak{np}_{\mathfrak{tu}}$ refer to the lists consisting of of $rr_{\mathfrak{tu},j}, \mathfrak{pn}_{\mathfrak{tu},j}, \mathfrak{np}_{\mathfrak{tu},j}$ correspondingly.

Bi-LSTM learning for each sub-structure: As each one of the time units are structured chronologically, RNN is used for providing the structure. To effectively represent the context information of the whole structure-development process, Bi-LSTM is used for learning the representations for dynamic structures as depicted in Fig. 2. LSTM is robust when a general RNN technique is used; but, for estimating the hidden state, it employs a variety

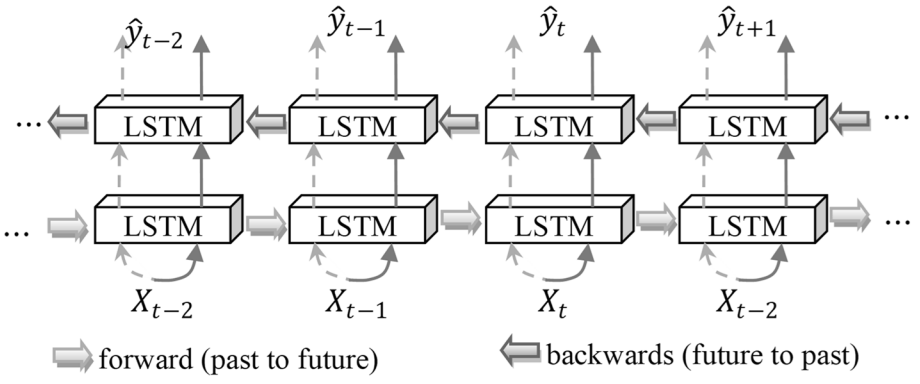

Fig. 2 The BiLSTM structure with three-time steps related

of approaches, which helps finding a solution to the problem of RNN and does not have the potential to address long-distance dependency. A group of identical memory units with three gates make up the LSTM method. The values of the various states of the LSTM word unit are illustrated using the text feature vector FV and the sample word.

Listed below are the individual estimate functions, where Σ refers to a sigmoid function and ⊙ represents the dot multiplication [23]. When the sigmoid function is provided with data from both the previous hidden state and the current input, it produces an output value that falls between [0,1]. Near to 0 implies the higher chances for its elimination, and the near to 1 implies the more chances for its retention. The F_{gt} specifies a forget gate:

$$Fgt_{\mathbb{t}} = \Sigma\left(W_{Fgt} * \left[hs_{\mathbb{t}-1}, x_{\mathbb{t}}\right] + b_{Fgt}\right) \tag{1}$$

The input gate $\mathrm{Igt}_{\mathbb{t}}$ is helpful in updating the cell state. The sigmoid function receives the previous layer's hidden state and current input in the first step. To decide what requires updating, the value between 0 and 1 is modified. 0 means unimportant, 1 means important. Second, the tanh function generates a new candidate value vector using the hidden state of the previous layer and the current input *cv*. Finally, sigmoid and tanh outputs are multiplied. The sigmoid output value determines whether information in the tanh output value is important and should be kept.

$$\mathrm{Igt}_{\mathbb{t}} = \Sigma\left(W_{Igt} * \left[hs_{\mathbb{t}-1}, x_{\mathbb{t}}\right] + b_{Igt}\right) \tag{2}$$

$$\widetilde{cv_{\mathbb{t}}} = \tanh s\left(W_{cv} * \left[hs_{\mathbb{t}-1}, x_{\mathbb{t}}\right] + b_{cv}\right) \tag{3}$$

The current state *Cst* is used for calculating the cell state. Initially, a point-by-point sequential multiplication of the forgetting vector and the cell state of the previous layer is performed. In the event that it is multiplied by a number that is somewhat close to zero, this indicates that this information has to be removed from the new cell state. Subsequently, this value is successively added to the input gate's output value, and the cell state is updated with the neural network's new knowledge. This is the new cell state.

$$Cst_{\mathbb{t}} = Fgt_{\mathbb{t}} \bullet Cst_{\mathbb{t}-1} + Igt_{\mathbb{t}} \times \widetilde{cv_{\mathbb{t}}} \tag{4}$$

The output gate *Ogt* is helpful in deciding the value of the upcoming hidden state. The data that was previously supplied is present in the concealed state. The newly acquired cell state is sent to the tanh function after being supplied to the sigmoid function together with the previous concealed state and current input. Finally, tanh and sigmoid outputs are multiplied to determine the concealed state's information. Then, the hidden state becomes the current cell's output, and the next time step receives the new cell state and hidden state.

$$Ogt_{\mathbb{t}} = \Sigma\left(W_{Ogt} * \left[hs_{\mathbb{t}-1}, x_{\mathbb{t}}\right] + b_{Ogt}\right) \tag{5}$$

$$hs_{\mathbb{t}-1} = Ogt_{\mathbb{t}} \bullet \tanh s\left(Cst_{\mathbb{t}}\right) \tag{6}$$

where W_{Fgt}, W_{Igt}, W_{cv} and W_{Ogt} refer to the weight of the LSTM. b_{Fgt}, b_{Igt}, b_{cv} and b_{Ogt} indicates the bias of the LSTM. $hs_{\mathbb{t}}$ stands for the hidden state in time $\mathbb{t}$. Σ indicates the activation function sigmoid, tan*hs* signifies the hyperbolic tangent function. LSTM data is inadequate. Future operations may benefit from this knowledge. Bidirectional LSTM layers are front and backward. This strategy uses a forward layer with sequence data and a backward layer with fresh sequence data. A similar output layer joined these layers. Moreover,

this model completely utilizes sequence content information. Time input is assumed $\mathbb{t}$ is given by the word embedding $w_{\mathbb{t}}$, at time $\mathbb{t}-1$, such that the result of forward hidden unit becomes $\overrightarrow{hs_{\mathbb{t}}}$, and simulation result of the backward hidden unit is given by $\overleftarrow{hs_{\mathbb{t}-1}}$, Finally, the outcomes obtained using backward and a concealed unit at time $\mathbb{t}$ are equivalent as seen below:

$$\overrightarrow{hs_{\mathbb{t}}} = \mathfrak{L}\left(w_{\mathbb{t}}, \overrightarrow{hs_{\mathbb{t}-1}}, cv_{\mathbb{t}-1}\right) \tag{7}$$

$$\overleftarrow{hs_{\mathbb{t}}} = \mathfrak{L}\left(w_{\mathbb{t}}, \overrightarrow{hs_{\mathbb{t}-1}}, cv_{\mathbb{t}-1}\right) \tag{8}$$

$\mathfrak{L}(\bullet)$ denotes the hidden layer job of the LSTM hidden layer. The forward output vector is $\overrightarrow{hs_{\mathbb{t}}} \in R^{1\times\mathfrak{H}}$, whereas the backward output vector is $\overleftarrow{hs_{\mathbb{t}}} \in\leftarrow R^{1\times\mathfrak{H}}$. Clearly, these vectors must be merged to produce the text feature. $\mathfrak{H}$ reveals the total number of cells in the concealed layer:

$$\mathfrak{H}_t = \overrightarrow{hs_{\mathbb{t}}} \parallel \overleftarrow{hs_{\mathbb{t}}} \leftarrow \tag{9}$$

BBO based Bi-LSTM Model (OBi-LSTM): The DE model is utilized to optimize the BiLSTM classification hyper-parameters for optimum classification outcomes. Batch size and hidden neuron number are employed here. From the initial randomly generated conclusion, the DE approach improves the sentiment classification model's accuracy until the stop requirement is fulfilled. The fitness function ($\mathfrak{FF}$) BiLSTM estimates rumour classification accuracy. BBO utilises crossover and mutation models but, unlike the Genetic Algorithm, has an explicit upgrading function (GA).

Biogeography-based optimisation (BBO) is a potent optimisation technique, particularly when coping with multidimensional and high-objective optimisation problems. This team describes the phenomena that permit the delineation of the BBO:

Initialization of Habitats: Habitats are locations of residence, reproduction, mutation and death of the species $\mathcal{S}$. The attribute is connected with a set of problem solutions, whose optimization is to be done.

Description of Suitable Index Variable (SIV): Variable is useful for describing ambient factors of species that exist in all environments, such as the availability of water, sunlight, sustenance, the extent of vegetation, etc. This variable corresponds to the set's independent variable when optimising a collection of solutions to a problem.

Description of Habitats Suitable Index (HSI): Finding a suitable habitat for a species is made easier with the aid of HSI. The SIV can determine the worth of this variable by using the objective function. This variable is useful for gauging how appealing various solutions are to the optimisation challenge at hand. The linear model of BBO, including the species' migratory pattern and habitats, is shown in Fig. 3.

In the BBO optimisation of a problem, there are two stages: migration and mutation. There are two distinct actions that can be taken during the migration phase: emigration and immigration. After the algorithm completes the migration process, the immigration SIV and emigration SIV are computed using the migration rate. For each SIV, the algorithm generates a random value and compares it to the mutation rate to determine whether or not the mutation operation should be executed. If the value is lower than the mutation rate, mutation occurs and the corresponding SIV is indiscriminately set to a new value. Other than that, the mutation will not occur. This phrase is used to describe both mutation and migration:

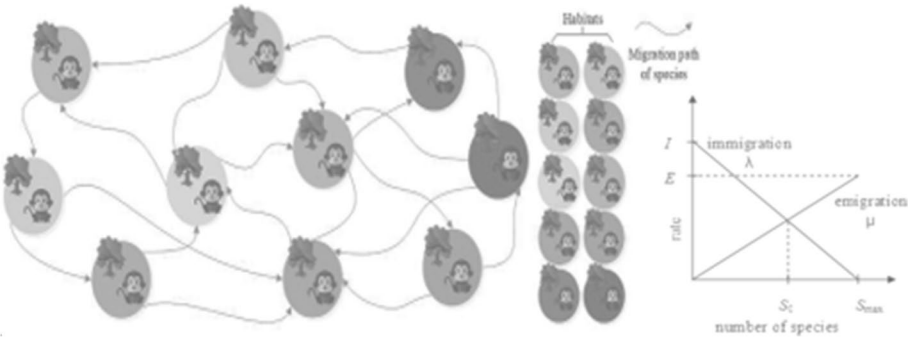

Fig. 3 Habitats and species migration path of BBO [24]

$$\mathfrak{I}_{\mathcal{S}} = \mathfrak{I}_{max}\left(1 - \frac{\mathcal{S}}{\mathcal{S}_{max}}\right) \tag{10}$$

$$\mathfrak{E}_{\mathcal{S}} = \frac{\mathfrak{E}_{max} \bullet \mathcal{S}}{\mathcal{S}_{max}} \tag{11}$$

$$\mathbb{N}_{\mathcal{S}} = \mathbb{N}_{max}\left(\frac{1 - \mathbb{P}_{\mathcal{S}}}{\mathbb{P}_{max}}\right) \tag{12}$$

If $\mathcal{S}$ is the species number, then $\mathfrak{I}_{\mathcal{S}}$ is the immigration rate; otherwise, $\mathfrak{I}_{max}$ is the maximum immigration rate. Number of species, denoted by SS. If the number of species is$\mathcal{S}$, then. $\mathfrak{E}_{\mathcal{S}}$ represents the emigration rate. E_max represents the maximum emigration rate. If $\mathcal{S}$ is the number of species, then $\mathbb{N}_{\mathcal{S}}$ is the mutation rate. Maximum mutation rate, denoted by $\mathbb{N}_{max}$. The probability of finding $\mathcal{S}$ species in each environment is denoted by the symbol$\mathbb{P}_{\mathcal{S}}$. Additionally, $\mathbb{P}_{max}$ represents the highest possible species probability. The tuning procedure for the hyperparameter in the solution that needs optimisation is improved by both migration and mutation in the BBO. The mutation is frequently carried out post migration. Simultaneously, elitism is presented for maintaining the optimal solution found during every iteration efficiently.

As per the abovementioned content, BBO offers the benefits of less number of parameters, plainness in operation and good speed of convergence. It is capable of solving the issue of making Bi-LSTM hyperparameters optimized. The BBO algorithm is used to determine the Bi-hyperparameters, LSTM's which include the batch size and the number of hidden layers. When BBO is used for optimizing the hyperparameters of Bi-LSTM, the respective correlation between CNN, and BBO variables is as given. SIV indicates the value of the hyperparameters requiring optimization. The objective function of BBO stands for the Bi-LSTM. HSI indicates the output of Bi-LSTM. Habitat refers to a bunch of solutions for all the hyperparameters requiring optimization. When the BBO is performed, by iterating continuously, the optimization operation is finished [25].

CNN: Next, to the thick layer, the CNN's output features are mixed with the Bi-output. LSTM's The concatenated feature also serves as the dense layer's input. It then moves into the flatten layer and dropout layer. Reduced redundancy and improved orthogonality between various characteristics at each layer are the direct effects of dropout layer, with

the sparse viewpoint of data representation also giving an accurate interpretation of these effects. The Flatten layer is helpful in "flattening" the input. This implies that a multidimensional input is flattened into one dimension. A CNN is helpful in extraction of features and particularly, one-dimensional convolution operation is used with filters having diverse sizes $\mathbb{w} \in R^{\mathfrak{H} \times k}$ for the extraction of several features. Dynamic k-max-pooling [26] is also used for getting the k-largest value of the feature map, represented as P^{k-max} which is capable of well-capturing the associations of long-range component present in the sequence. The j-th feature map is derived as

$$P_j^{k-max} = \mathfrak{L}\left(ReLU\langle w_{\mathbb{t}}, \widehat{X}\rangle_T + b\right) \tag{13}$$

where $\mathfrak{L}$ indicates the k-max pooling operation, $w_{\mathbb{t}}$ stands for the filter with h_j height, T signifies Trace/Frobenius inner product and b specifies the bias. The textual content structure TC is represented using the calculation

$$TC = \mathfrak{c}_s\left(P_1^{k-max}, P_2^{k-max}, \ldots, P_m^{k-max}\right) \tag{14}$$

where $\mathfrak{c}_s$ refers to the process of transformation and concatenation, and m represents the number of filters to be applied. In the instance of the fusion layer, the representations of the content network and the structure network are combined via a concatenate operation and a following completely connected layer. A softmax function is used, at long last, to calculate the chance that a certain event is only a rumor:

$$X = \mathfrak{c}\left(TC, \mathfrak{H}_t\right) \text{and} Y = \Sigma(f(X)) \tag{15}$$

where Y indicates the prediction result specifying if the incident is a rumor, f stands for the fully connected layer, Σ refers to the sigmoid function. TC indicates the representation of textual content network and cross-entropy is used as the loss function, i.e.

$$Loss = \sum_{i=1}^{m} Y_i \log \widehat{Y}_i + \left(1 - Y_i\right) \times \log\left(1 - \widehat{Y}_i\right) \tag{16}$$

where Y_i and $\widehat{Y}_i$ stands for the prediction outcome and true label of the i-th event correspondingly. Adam [26] works as an optimizer to accelerate convergence during training and a dropout [27] mechanism in the penultimate layer to prevent overfitting.

4 Experimental results and discussion

The experiment is implemented in MATLAB. The performance achieved with the five classifiers like Naïve Bayes, KNN, Random Forest, Adaboost and Enhanced Adaboost with PSO are assessed in terms of four classification metrics including F-Measure, Accuracy, Recall, and Precision are referred as bias metrics if imbalance occurs. Non-rumors to be negative class and rumor is regarded as positive class. The data utilized for this technical work is known as the PHEME dataset and can be accessed online by PHEME dataset (2016). The outcomes of imbalance and balance datasets like Charlie Hebdo, Ferguson, Germanwings Crash, Ottawa shooting, Sydneysiege incidents are analyzed.

This dataset consists of a set of Twitter non-rumours and rumours published when breaking news are portrayed with its True, veracity value, False or Unverified. Data is structured as given. Each event must have directory, made up of two subfolders, non-rumours and rumours.

These two directories contain tweet ID folders. Tweets are created based on the'source-tweet' directory of tweets evaluated, whereas the 'reactions' directory contains tweets responding to the original tweet. The distribution of rumours and non-rumours in the real dataset is depicted in Fig. 4.

For measuring the recital of different approaches on rumor detection, Accuracy metric is defined as below:

$$Accuracy = \frac{cp}{Tes} \tag{17}$$

where 'cp' indicates the sum of correctly predicted non-rumors and rumors in testing set, and 'Tes' indicates the size of testing set. Concurrently, the performance of those models are computed with recall, precision and F-measure metrics.

Measurement of rumors have more important, precision is used often, which is shown as given:

$$Precision(PR) = \frac{\#TP}{\#TP + \#FP} \tag{18}$$

where #TP indicates the sum of true positives, indicating the sum of accurately identified rumors, #FP stands for the sum of false positives, which signifies the number of incorrectly detected non-rumors as non-rumors.

Recall can be used for finding sensitivity:

$$Recall(RE) = \frac{\#TP}{\#TP + \#FN} \tag{19}$$

where #FN sum of false negatives, indicating the number of rumors identified as non-rumors.

F-measure is used for adding both recall and precision, computed as below:

$$F - measure = \frac{2.(PR.RE)}{(PR + RE)} \tag{20}$$

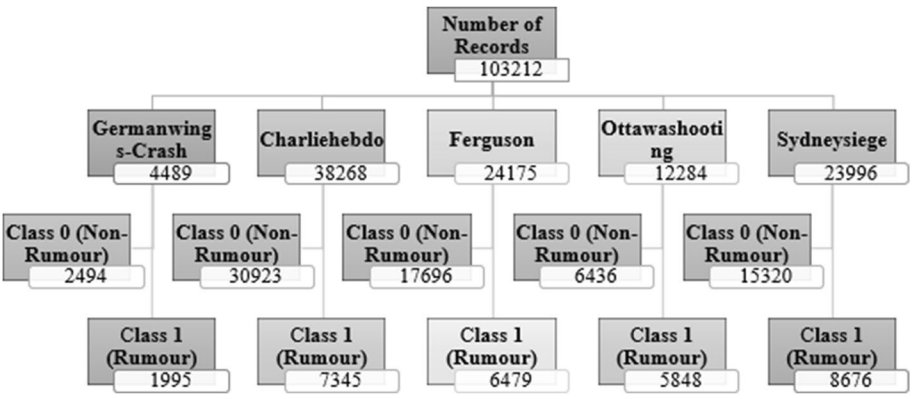

Fig. 4 SmartArt about PHEME Dataset

In this section, the comparison between the DRIMUX [9], HCSFLA-PSO [14], HAGNN [20] and proposed OBi-LSTM-CNN is validated in terms of recall, precision, accuracy, and f-measure.

4.1 Comparison results for germanwings-crash

Figure 5 illustrates the class imbalanced results of Germanwings-Crash dataset for the techniques like DRIMUX, HCSFLA-PSO, HAGNN and OBi-LSTM-CNN. Consequently, the proposed OBi-LSTM-CNN yields an improved accuracy, precision, recall and F-measure on Germanwings-Crash test sets yielding 74.86%, 65.80%, 73.81% and 71.64%. The performance of the structure network is superior to that of other methods since it employs just dynamic structural information and no extra textual information. Due to the incorporation of structure information, the performance of OBi-LSTM with CNN surpasses that of other approaches, demonstrating that the dynamic structure information provides unique evidence in rumour detection. The findings suggest that the client context-sensitive information transmission model proposed for rumour identification is both effective and efficient.

4.2 Comparison results for charliehebdo

Figure 6 illustrates the class imbalanced outputs of Charliehebdo dataset for the techniques like DRIMUX, HCSFLA-PSO, HAGNN and OBi-LSTM-CNN. Consequently, the proposed OBi-LSTM-CNN yields an improved accuracy, precision, recall and F-measure on Charliehebdotest sets yielding 92.10%, 71.11%, 93.92% and 80.92%. Also, the way the BBO algorithm is designed for optimizing the parameter of Bi-LSTM lead to better overall classification outcomes and the propagation structure is the same when the data is collected. Here, the performance textual content-based techniques, including DRIMUX, HCSFLA-PSO and HAGNN, is not up to the mark. The proposed model combines the

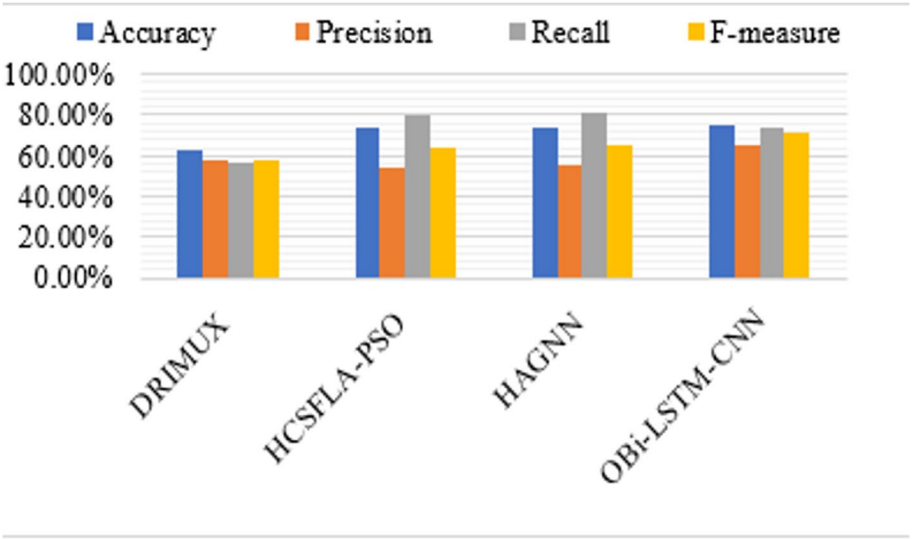

Fig. 5 Comparison Results for class imbalanced: Germanwings-Crash

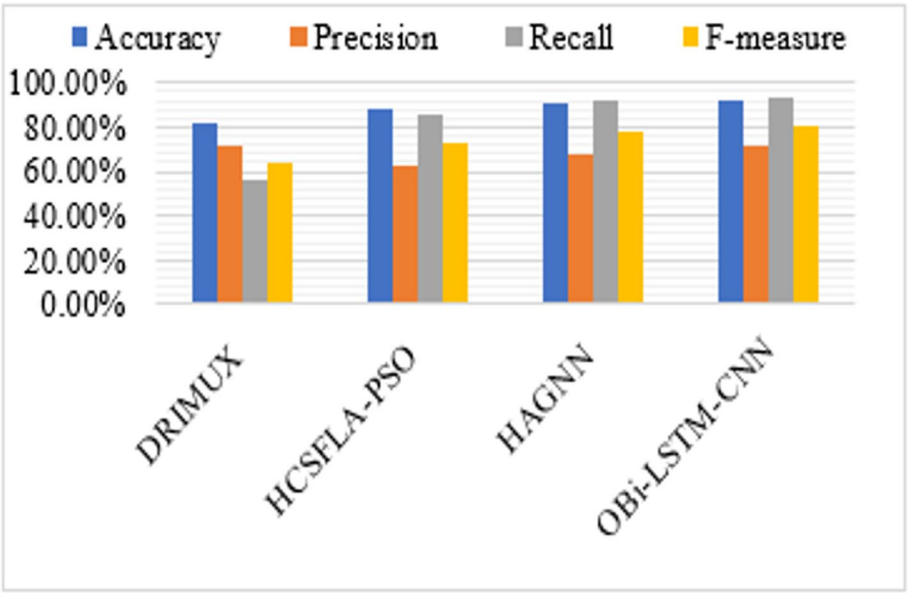

Fig. 6 Comparison Results for class imbalanced: Charliehebdo

structural information of reposts and content information, and therefore its performance excels other techniques.

4.3 Comparison results for ferguson

Figure 7 illustrates the class imbalanced results of Ferguson dataset for the techniques like DRIMUX, HCSFLA-PSO, HAGNN and OBi-LSTM-CNN. Consequently, the proposed OBi-LSTM-CNN yields a better accuracy, precision, recall and F-measure on Ferguson test sets achieving 92.10%, 71.11%, 93.92% and 80.92%. It is found from the results that

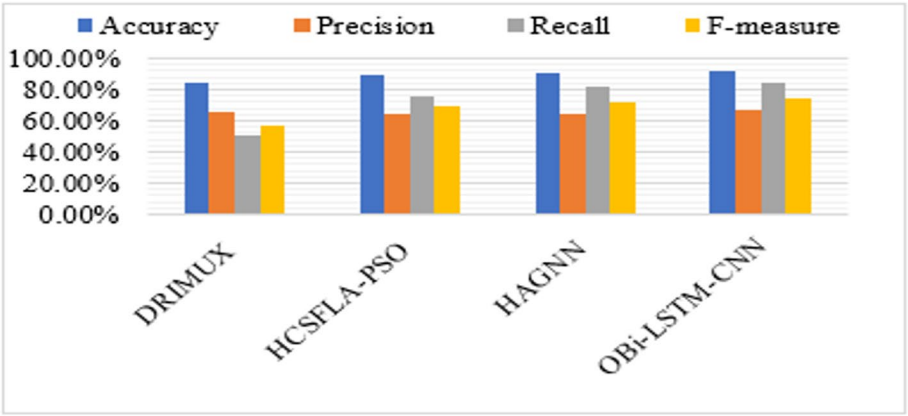

Fig. 7 Comparison Results for class imbalanced: Ferguson

the first level prediction was right, and that the combination of CNNs and OBi-LSTMs have the potential to make the best use of both the CNN's strength in identifying thec local patterns, and the OBi-LSTM's capability to exploit the text's sequence. But, the sequencing of the layers in the models have a critical function in their performance.

4.4 Comparison results for ottawashooting

Figure 8 illustrates the class imbalanced results of Ottawashooting dataset for the techniques including DRIMUX, HCSFLA-PSO, HAGNN and OBi-LSTM-CNN. Consequently, the proposed OBi-LSTM-CNN yields an improved accuracy, precision, recall and F-measure on Ottawashootingtest sets achieving 87.05%, 82.14%, 93.31% and 87.36%. it is understood from the results that it is no coincidence that very less variation is seen between the models. It has been observed that the initial convolutional layer of the OBi-LSTM-CNN does not lose a significant amount of the text's order or sequence information. As a consequence of this, the LSTM layer will only operate as a fully connected layer in the event that the sequence of the convolutional layer does in fact give any information. This model is found to successfully exploit the entire potential of the LSTM layer and therefore does not reach its peak.

4.5 Comparison results for sydneysiege

Figure 9 shows the class imbalanced results of Sydneysiege dataset for the techniques including DRIMUX, HCSFLA-PSO, HAGNN and OBi-LSTM-CNN. Consequently, the proposed OBi-LSTM-CNN yields an improved accuracy, precision, recall and F-measure on Sydneysiege test sets yielding 79.46%, 65.43%, 76.46% and 70.51%. Due to this, the performance of the proposed model is found to be remarkable due to its initial OBi-LSTM layer functioning like an encoder so that for each token in the input, an output token exists, containing the information not just of the actual token, but every other preceding tokens.

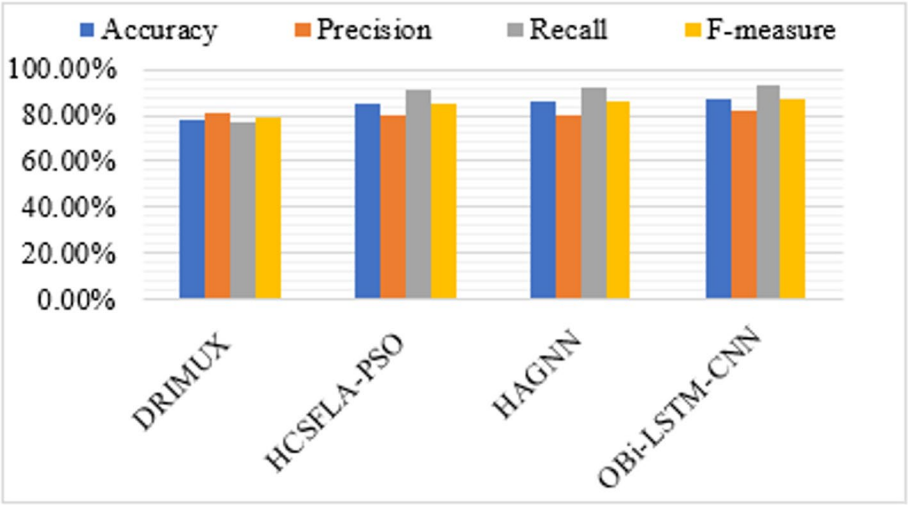

Fig. 8 Comparison Results for class imbalanced: Ottawashooting

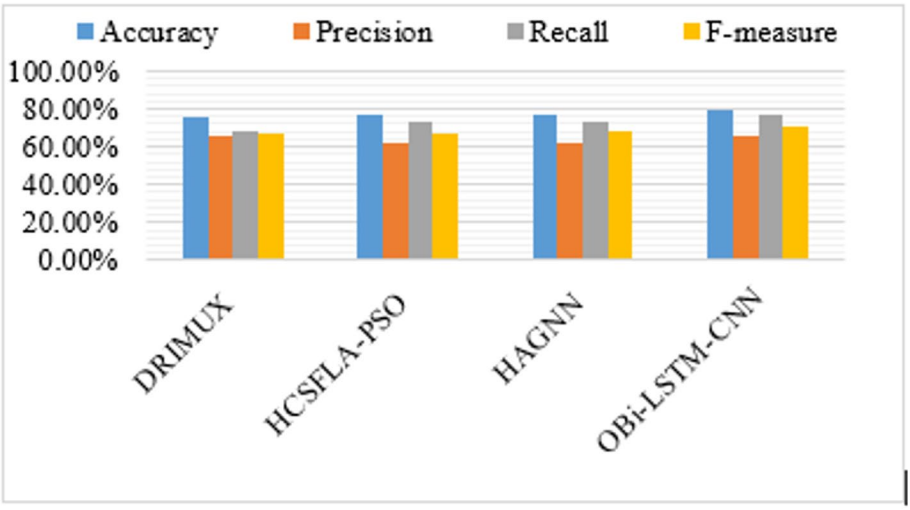

Fig. 9 Comparison Results for class imbalanced: Sydneysiege

Later, the CNN layer will get the local patterns applying this better representation of the actual input, so that the accuracy is improved.

5 Conclusion and future work

In this study, a combination model called OBi-LSTM-CNN that incorporates structural and textual data is generated using an attention-based model for the extraction of features from dynamic propagation structures. This model is not only capable of the meaningful extraction of local features of text by utilizing convolutional neural networks, but it also makes use of the Bi-LSTM to consider the global features of the text, in addition to giving complete consideration to the contextual semantic information of the words. This capability is in addition to the model's ability to meaningfully extract local features of text. This study adds an attention mechanism before CNN and Bi-LSTM models, dropout layer, flatten layer, and fully connected layer to increase performance. From the result analysis, the the proposed OBi-LSTM-CNN yields a better accuracy, precision, recall and F-measure on Ferguson test sets achieving 92.10%, 71.11%, 93.92% and 80.92%. The suggested model's classification accuracy is greater than the CNN and Bi-LSTM models. The findings show that the suggested model in this study outperforms the related models in classification accuracy and actively improves the model's accuracy and thoughtfulness in exploiting dynamic propagation structures. Especially, for sites such as Facebook, supporting elaborate user profiles, the technique would operate with more efficiency compared to another set of sites, like Twitter, where there is very less personal information. As a futuristic approach, it will be attempted to examine models yielding superior performance on social network. In this study, shallow CNN is employed, however in the future, deeper CNN will be fused with OBi-LSTM or other models to classify text sentiment. Second, Deep Learning will be used to increase the model's text sentiment analysis accuracy by combining the advantages of traditional Machine Learning or Sentiment Lexicon with Deep Learning.

Authors contributions All authors are approved for this work.

Funding No fund received for this project.

Data availability Not Applicable.

Declarations

Ethical approval and human participation No ethics approval is required.

Conflicts of interest The authors declare that they have no conflict of interest.

References

1. Bali A, Desai P (2019) Fake news and social media: Indian perspective. Media Watch 10(3):737–750
2. Yang F, Liu Y, Yu X, Yang M (2012) Automatic detection of rumor on sina weibo. In Proceedings of the ACM SIGKDD workshop on mining data semantics, p 1–7 https://www.semanticscholar.org/paper/Automatic-detection-of-rumor-on-Sina-Weibo-Yang-Liu/96b17f09f3afcc9238c1789208eeb016bc2a84d8
3. Vosoughi S, Mohsenvand MN, Roy D (2017) Rumor gauge: Predicting the veracity of rumors on Twitter. ACM transactions on knowledge discovery from data (TKDD) 11(4):1–36
4. Sun S, Liu H, He J, Du X (2013) Detecting event rumors on sina weibo automatically. In: Asia-Pacific web conference. Springer, Berlin, Heidelberg, pp 120–131 https://link.springer.com/chapter/10.1007/978-3-642-37401-2_14
5. Meel P, Vishwakarma DK (2020) Fake news, rumor, information pollution in social media and web: A contemporary survey of state-of-the-arts, challenges and opportunities. Expert Syst Appl 153:112986
6. Ahsan M, Kumari M, Sharma TP (2019) Rumors detection, verification and controlling mechanisms in online social networks: A survey. Online Social Networks and Media 14:100050
7. Tan L, Wang G, Jia F, Lian X (2022) Research status of deep learning methods for rumor detection. Multimed Tools Appl 82(2):2941–2982
8. Islam MR, Liu S, Wang X, Xu G (2020) Deep learning for misinformation detection on online social networks: a survey and new perspectives. Soc Netw Anal Min 10(1):1–20
9. Wang B, Chen G, Fu L, Song L, Wang X (2017) Drimux: Dynamic rumor influence minimization with user experience in social networks. IEEE Trans Knowl Data Eng 29(10):2168–2181
10. Wu Q, Zhao X, Zhou L, Wang Y, Yang Y (2019) Minimizing the influence of dynamic rumors based on community structure. Int J Crowd Sci 3(3):303–314
11. Hosni AIE, Li K, Ahmad S (2019) DARIM: Dynamic approach for rumor influence minimization in online social networks. In: International conference on neural information processing. Springer, Cham, pp 619–630. https://www.researchgate.net/publication/335910204_DARIM_Dynamic_Approach_for_Rumor_Influence_Minimization_in_Online_Social_Networks
12. Hosni AIE, Hafiani KA, Chenoui A, Beghdad Bey K (2022) Hybrid approach for rumor influence minimization in dynamic multilayer online social networks. In: International conference on computing systems and applications. Springer, Cham, pp 275–285. https://link.springer.com/chapter/10.1007/978-3-031-12097-8_24
13. Ni P, Zhu J, Wang G (2023) Misinformation influence minimization by entity protection on multi-social networks. Appl Intell 53(6):6401–6420
14. Hu X, Xiong X, Wu Y, Shi M, Wei P, Ma C (2023) A Hybrid Clustered SFLA-PSO algorithm for optimizing the timely and real-time rumor refutations in Online Social Networks. Expert Syst Appl 212:118638
15. Lin H, Zhang X, Fu X (2020) A graph convolutional encoder and decoder model for rumor detection. In: 2020 IEEE 7th International conference on data science and advanced analytics (DSAA). IEEE, pp 300–306. https://www.researchgate.net/publication/344712414_A_Graph_Convolutional_Encoder_and_Decoder_Model_for_Rumor_Detection
16. Yang P, Leng J, Zhao G, Li W, Fang H (2023) Rumor detection driven by graph attention capsule network on dynamic propagation structures. J Supercomput 79(5):5201–5222
17. Asghar MZ, Habib A, Habib A, Khan A, Ali R, Khattak A (2021) Exploring deep neural networks for rumor detection. J Ambient Intell Humaniz Comput 12(4):4315–4333

18. Zojaji Z, Tork Ladani B (2022) Adaptive cost-sensitive stance classification model for rumor detection in social networks. Soc Netw Anal Min 12(1):1–17
19. Sailunaz K, Kawash J, Alhajj R (2022) Tweet and user validation with supervised feature ranking and rumor classification. Multimed Tools Appl 81(22):31907–31927
20. Xu S, Liu X, Ma K, Dong F, Riskhan B, Xiang S, Bing C (2023) Rumor detection on social media using hierarchically aggregated feature via graph neural networks. Appl Intell 53(3):3136–3149
21. Tembhurne JV, Almin MM, Diwan T (2022) Mc-DNN: Fake news detection using multi-channel deep neural networks. Int J Semantic Web Inf Syst (IJSWIS) 18(1):1–20
22. Liu Y, Xu S (2016) Detecting rumors through modeling information propagation networks in a social media environment. IEEE Trans Comput Soc Syst 3(2):46–62
23. Guo H, Cao J, Zhang Y, Guo J, Li J (2018) Rumor detection with hierarchical social attention network. In: Proceedings of the 27th ACM international conference on information and knowledge management, p 943–951. https://dl.acm.org/doi/10.1145/3269206.3271709
24. Rahmati SHA, Zandieh M (2012) A new biogeography-based optimization (BBO) algorithm for the flexible job shop scheduling problem. Int J Adv Manuf Technol 58(9):1115–1129
25. Grefenstette E, Blunsom P (2014) A convolutional neural network for modelling sentences. In: The 52nd Annual Meeting of the Association for Computational Linguistics, Baltimore, Maryland. https://aclanthology.org/P14-1062.pdf
26. Wang S, Kong Q, Wang Y, Wang L (2019) Enhancing rumor detection in social media using dynamic propagation structures. In: 2019 IEEE International Conference on Intelligence and Security Informatics (ISI). IEEE, pp 41–46. https://arxiv.org/abs/2008.12154
27. Li Q, Zhang Q, Si L (2019) Rumor detection by exploiting user credibility information, attention and multi-task learning. In: Proceedings of the 57th annual meeting of the association for computational linguistics, pp 1173–1179. https://aclanthology.org/P19-1113/