International Journal of Computing and Digital Systems
ISSN (2210-142X)
Int. J. Com. Dig. Sys. 16, No.1 (Sep-24)

http://dx.doi.org/10.12785/ijcds/160186

Emotion Infused Rumour Detection Model using LSTM

Osheen Sharma[1,2], Monika Sethi[2] and Sachin Ahuja[3]

[1]*Post Graduate Department of Information Technology,Goswami Ganesh Dutta Sanatan Dharma College, Chandigarh, India*
[2]*Deepartment of Computer Science Engineering,Chitkara University Institute of Engineering Technology, Chitkara University, Punjab, India*
[3]*University Institue of Engineering, Chandigarh University, Punjab, India*

Received 7 Apr. 2024, Revised 21 May 2024, Accepted 22 May 2024, Published 1 Sep. 2024

Abstract: Twitter now 'X', is a highly favored platform for sharing brief messages, known as tweets, read and shared among users at a rapid pace. Hence, the dissemination of information occurs quickly within the community of users in network. Twitter's unregulated environment provides a suitable platform for individuals to share and circulate unverified information; this propagation of rumours can greatly affect society. The detection of rumour accurately on Twitter from tweets is a crucial task. In this study, we suggested an Emotion Infused Rumour Detection model based on an LSTM model that employs tweet text and twenty-one distinct linguistic, user, post, and network features to classify between rumour and non-rumour tweets. comparison of the proposed model i.e. Emotion Infused Detection model using LSTM was done with two different deep learning models to check the achieved outcomes. The findings of the evaluations exhibit the supremacy of the deep learning-based model for identifying rumours. The suggested Emotion Infused Rumour Detection model, which uses an LSTM model, earned an F1-score of 0.91 in identifying rumour and non-rumour tweets, outperforming the state-of-the-art findings. The suggested approach can lessen the influence of rumours on society, prevent loss of life and money, and increase users' confidence in social media platforms. The model proposed has the potential to promptly and accurately recognize tweets containing rumours, aiding in the prevention of the spread of misinformation.

Keywords: LSTM, Rumour Detection, Twitter, RNN, Deep Learning

1. INTRODUCTION

For the last many yonks social media has been a major attraction. Platforms like Twitter and Facebook has been an integral part of our daily existence and have become entrenched in our daily routines and are now considered indispensable. [1], [2], [3], [4], [5], [6],. Twitter, in particular, is a highly favored platform for sharing brief messages, known as tweets, which are limited to 280 characters. These messages are read and shared (or retweeted) among users at a rapid pace, making it a powerful tool for information dissemination. Hence, the dissemination of information occurs quickly within the community of users in network. In fact, many breaking news stories are first reported on Twitter before being picked up by traditional media outlets [7]. The data on Twitter has proven to be valuable for various purposes, such as disaster management [7], [8], [9], [10], [11], [12], predicting locations [10], managing customer relationships [6], [13], [14], [15], monitoring antisocial behavior [16], tracking government policies [17], and monitoring traffic [18], among others.

Twitter, re-branded and being referred to as 'X' since 2022, lacks advanced filtering and moderation systems that can verify the accurateness of posted content, leading to the dissemination of rumour easily [19], [20], [21], spam [22], biased sentiment [23], and other forms of inappropriate behavior. "Rumours are unverified and instrumentally relevant information in circulation that arises in contexts of ambiguity, danger, or potential threat" [24]. Twitter's unregulated environment provides a suitable platform for individuals to share and circulate unverified information.

The propagation of rumours can greatly affect society as they can misguide public perception or opinion, disrupt social harmony, erode citizens' confidence in the government, reduce the government's credibility, and pose a major risk to social constancy [25], [26], [27], [28]. For instance, in 2018, a report by The Indian Express revealed multiple cases of mob lynching resulting in fatalities of 27 people across nine different states. These incidents were resultants of rumours of child-kidnapping that spread throughout the areas. The states affected included Jharkhand (7 deaths), Tamil Nadu (1 death), Karnataka (1 death), Telangana (1 death), Assam (2 deaths), West Bengal (2 deaths), Chhattisgarh (1 death), Tripura (3 deaths), and Maharashtra (9 deaths) [29]. The report also stated that in the span of three months, 20 individuals were killed in mob lynchings across the country. All of the victims caught in the midst

E-mail address: osheensharma12@gmail.com, monika.sethi@chitkara.edu.in, ed.engineering@cumail.in

of people incited by rumours were innocent individuals. This kind of misinformation caused widespread panic in the US in 2013 when a prominent news agency tweeted false information. The tweets were made claiming that there had been an explosion at the White House, resulting in the injury of President Barack Obama. This news was believed by millions until White House Press Secretary Jay Carney clarified that the President was unharmed [30]. The rumour was finally put to rest when the Associated Press announced that their account was hacked. Fortunately, this false news was quickly debunked, unlike a similar incident during the Boston Marathon where a rumour about Obama's injury caused significant time and resources for the US government to rectify.

To enhance the trustworthiness of social networks and lessen the adverse effects by inaccurate as well as deceptive data known as rumours, it is vital to swiftly detect and manage the dissemination of rumour content on social media platforms. Utilizing an automated system for detecting rumours can effectively debunk them in their early stages, thus reducing their propagation and minimizing the resulting harm effects [19][21][31]. Detecting rumored data circulated over social media is a difficult job that requires extensive research [21][31][32][33].

A rating system for the news sources was also explained by researchers, according to which the low rating of source affects the credibility for anonymous sources [34]. According to a Kim et al. study in 2019, it was proposed that attributing a numerical value to the origin (source) could serve as an effective method for combatting fake news [35]. By using statistical method, it was found that for early detection of rumours user and linguistic features should be used as some of the features such as structural features and temporal can help in differentiating and identifying rumours and non-rumours. But it was also stated that these features are not accessible in the initial stage of propagation [36]. Some researchers explained about the presence of some manually extracted features also in addition to the ones explained till now for rumour identification. The features extracted by one researcher were textual and user based whereas other researchers extracted features from linguistic based features and from tweet depicting the characteristics of tweet in terms of support, denial, questioning or a regular tweet [37][38][39]. In 2018, Chen et al. in order to identify rumours, extracted features from tweets based on text using attention mechanism of deep learning model [40].

The objective of this work is to propose and implement a rumour detection model namely (EIRD i.e. Emotion Infused Rumour Detection) based on LSTM which has the ability to identify rumours along with tagging the malicious accounts and the tweets as positive and negative by finding the emotion of the tweet as well. The implementation of the model is explained in this paper, the explanation includes layer to layer description of the architecture. The work utilized the RumourEval2019 dataset explained in the methodology section. In this paper, we have extended the feature set by considering textual features which were extracted using model of deep learning and twenty-one manually extracted features from tweets based on linguistic, user, post and network to form a set of hybrid features. The deep learning model used to extract features automatically from tweet texts is Long Short-Term Memory (LSTM) network. Additionally, we propose an emotionally infused rumour detection model using LSTM network that incorporates this feature set (hybrid) to categorize tweets as either rumours or non-rumours.

The main contributions of our work are:

- Extracting and combining features from tweets to build a set of hybrid feature to accurately categorize those features, and for extracting features text and user characteristics from tweets we utilized both automatic extractions using deep learning and manual extraction;
- Introducing an emotionally infused rumour detection (EIRD) model using LSTM network for classifying tweets.
- Comparing the proposed model with two other models to evaluate its performance.

The article is organized in a way that the research related is discussed in 2nd section of paper; the used methodology is outlined in 3rd section; the 4th section presents the results obtained from the experiment; and the discussion over the outcomes of the experiments is done in 5th section. Lastly, 6th section concludes the paper.

2. RELATED WORK

The process of verifying the accuracy of content on social media is a complex undertaking. Some studies have focused on utilizing deep and machine learning methods to extract important characteristics from social media posts in order to detect rumours, while others have concentrated on the individuals responsible for spreading rumours through the network. We will provide a concise overview of several potential approaches in this section, that have been suggested in this field.

In 2011, Castillo et al. focused on posts related to trendy topics, for this they formed a classifier to classify the credibility of posts, using features based on content, user, topic and propagation [41]. Similarly, three types of features based on content, network and memes which are microblog specific were used by a researcher in 2011 to identify rumours. They also helped identifying the users who were in support and were part in propagation of the rumour [42]. The base of research by Liang et al. in 2015 was the eleven features extracted from messages based on linguistic and user features which they used to formulate a machine learning model to detect rumours on Sina Weibo [26].

On one side a study by Suchita Jain et al. formed two categories of accounts one of general public accounts and the other of verified News channels for real-time rumour detection on Twitter. Their analysis was based on sentiment and semantic approach, which stated the low reliability of provided information by news channel account as compared to public accounts [43]. On the other side, the study by Mao et al. in 2016 used sentiment orientation along with shallow statistical features and deep features which possessed the detection accuracy 3.9 and F1 score 4.6. Their study stated the effectiveness of sentiment orientation in detecting rumour [44].

Likewise, Sivasangari et al. calculated strength and sentiment category for textual data using rule-based heuristics approach. The researcher introduced VADER sentiment analysis to obtain the sentiment lexicon score for scraped dataset for distinguishing between rumour and genuine content [45].

SVM (Support Vector Machines) is another approach used for rumour detection along with sentiment analysis. Li (2016) created a hybrid kernel SVM (SHSVM) classifier that is based on sentiment analysis. This classifier uses an emotions dictionary to analyze the sentiment trends of comments on social networks [46]. In 2017 Qiao Zhang et al. for rumour detection utilized implicit features with shallow features and employed SVM and Random Forest to classify these features [47]. On the other hand, in 2019 Ajao et al. proposed a novel sentiment-aware algorithm for the detection of fake news, asserting that emotional words provided was advantageous in sentiment-aware rumour detection, outperforming state-of-the-art algorithms [48].

In 2017, Ma et al. studied the circulation of microblog posts to gather valuable insights on the evolution of the actual message over time. A propagation tree based on kernel was utilized to identify key patterns and differentiate rumours from the initial microblog post [49]. In a similar vein in 2017, Liu et al. took on the challenge of identifying rumours by focusing on how they spread and analyzed different aspects of the messages on Sina Weibo. Various features such as content, user, time and message structure were extracted which were used along with SVM classifier to accurately classify messages as rumour or non-rumour [50]. A researcher in 2017 used CNN for detecting tweet stance and veracity. They employed previously trained word embedding – GloVe for converting into vector from text data [51].

In the year 2017, Zubiaga et al. presented a state-of-the-art approach, using previously identified factors from related posts to determine tweet as rumour [52]. Some of the researchers applied multiple approaches and compared the results of each to identify the best suitable approach based on the results obtained. Zhiwei Jin et al. focused on specific political event – U.S. presidential election 2016, and for detection proposed an algorithm using multiple word matching methods such as TF-IDF, BM25, Word2Vec and Doc2Vec. They verified the rumours using the verified articles relating to the election candidates – Hillary Clinton and Donald Trump [53]. Bhutani et al. considered three datasets and text processing techniques. The results were compared on the basis of accuracy obtained by them. There prime focus was on utilizing sentiments to improve the accuracy on fake news detection [54].

Some researchers also tried developing some specialized tools for the task, like a FAkeNewsTracker tool was proposed by Shu et al. in 2019 which learning solutions and considered linguistic and social engagement features which enhanced the performance of the model, it was based on deep learning tool. Also, visual representation of the results was done using the tool for better understanding [55]. Similarly, Yan Zhang et al. experimented by hiding layers of autoencoder to check its effect on performance. They used different yet recent Weibo sets and applied multiple self-adapting thresholds for calculation. There work was restricted to Sina Weibo- microblogging site in China [56]. In the same manner Ghanem et al. introduced EIN model- a LSTM model based on unusual emotion patterns present in rumour tweets as according to their theory emotions have a vital role in gathering user interest towards rumours and this can be an essential feature in identifying rumours [57].

Clustering is yet another approach among the diverse approaches used by researchers for rumour detection. Using this approach, clusters were ranked based on their similarity of containing uncertain factual claims. For this the researcher gathered tweets related to inquiry and tweets not related to inquiry, then they used regular expressions in different sets to find the similarity index in order to form clusters [58]. Similarly, another work using cluster-based approach focused on political tweets relating to Hillary Clinton and Barak Obama, the then presidential candidates, of the month August 2015 and September 2015 respectively. Suspicious accounts were tagged on the basis of their history of posting rumourous news over Twitter. The researcher combined and compared the results of different parameter combinations but concluded that labelling manually the rumour clusters is tedious task and due to the presence of multiple parameters and combinations to test there is no particular approach of finding the best combination [59]. Another approach used by a researcher was J48 classification algorithm to achieve the best accuracy. They applied Weka classification tool and proposed an algorithm that identified rumour as well as the propagation source from the tweets related to London-Riots 2011 [60].

In 2016 by Zubiaga et al. studies the behaviour of tweets in terms of support, denial and propagation mechanism. They suggested the need for developing model for real-time rumour veracity detection based on machine learning [61]. Another researcher applied Hawkes process to train the stance classifier using Twitter dataset considering textual and temporary data from tweets [62]. For retrieving tweets related to rumours another researcher used TLV- Tweet Latent Vector features and applied semantic similarity generating 100-dimensional vector for features [63]. A heuristic algorithm was proposed by a researcher to find the source node over the network utilizing hitting time statistics of the surrogate random walking method. They considered various networks and compared the results which stated that the results provided were better than the traditional centrality-based heuristics [64]. RNN network used by Mao et al. to learn features automatically and identify the rumour veracity applying semantic information [39].

The study by Oh et al. in 2018 was primarily focused on the acceptance and consequences of rumours in crisis time, according to them the individuals who were closely connected were more prone to rumours [65]. Similarly, another study by Mondal et al. was focused on proposing an early-stage detection model following a disaster. They utilized a probabilistic model and used prominent characteristics which were propagating rumor from Chennai flood of 2015 [39].

In their 2018 study, Chen et al. tried experimenting by combining RNN with autoencoder to learn individual user behaviour and used self-adapting threshold to evaluate the model effectiveness by considering the errors obtained from different Weibo users [66]. Another RNN model based on GRU was proposed by Rath et al. for identifying rumour spreaders. For input features they used user embedding which were extracted using re-weighted retweet network [67]. A hybrid model for rumour identification with the amalgamation of LSTM and CNN was proposed by a researcher which concluded with the statement that deep neural network model can possess good accuracy even with limited data for training [68].

Various previous studies have emphasized on the significance of user feature in rumour detection [26][41][61]. Our model integrated hybrid features and tweet content and suggested an Attention based LSTM model for rumour detection. We used word embedding to get a better understanding of the hidden meanings of tweet text and classify tweets as rumours and non-rumours with more accuracy. In studies regarding rumour detection in the year 2018 and 2019, a thorough examination of techniques was carried out by the researchers. They stated that multiple features extracted from linguistic data were actually the base of many previous researches [69][31]. The results of these systems heavily relied on the usefulness of feature extraction. To address this, recent studies [19][51][68][70] have proposed deep learning models to overcome the limitations of manually crafted features in identifying rumours. Some researchers developed BiLSTM-CNN, to categorize tweets as either rumours or non-rumours [70]. They used Pheme dataset which is a publicly available dataset to attain a state-of-the-art [62]. Liu et al. (2019) utilized an LSTM network to identify rumours by analyzing changes in the spread of contents, spreaders, and diffusion structure [71]. Summarization of previous studies and works related to rumour stance classification is done in Table I.

A. Limitations of Related Work

Most of the researchers while detecting rumours focused on a particular topic or rumour that was spread on a social networking site and not on general information or any general rumour. For example, some tried detecting rumour related to US presidential elections, few only considered tweets related to London riots of 2011, or simply on rumours related to politics. Their area of research was very small, so was the dataset they used. The focus of their research was very specific and not on general information or any general rumour [47][53][60] Few only considered the verified accounts on social media to detect rumours [43][55]. Their approach was not that effective as they did not try to find out any malicious account. They just verified the information from verified accounts and tagged them as rumour or a non-rumour. Some only considered rumours already identified by some rumour detection websites (Politifact, Snopes, etc.) and showed the results based on that data. In most of the researches real time data was not considered whereas a dataset consisting of a very limited number of tweets/posts were considered. The accuracy provided with such research might not be trustworthy [48] [53] [55]. The real test of rumour detection model can only be done if real time data is considered. Most of the researchers have either analyzed the information to tag them as rumour and non-rumour or they have marked the information as positive and negative [45][53]. But there are various motives and sentiments behind a rumour such as Political intent, financial profit, or the rush to convey the information without any verification, etc... Early detection of rumour: Most researchers have done rumour detection on a specific dataset after the rumour was widespread [48][53][54]. Due to the fast circulation of information, the delay in identification and correction can cause greater damage. More the delay, the more will be the damage.

3. METHODOLOGY

A series of in-depth experiments was conducted to determine the veracity of rumours using traditional deep learning models. We used three distinct models: (i) Long-Short Term Memory (LSTM), (ii) Emotionally-Infused Model (EIN), and (iii) Emotion Infused Rumor Detection model using LSTM. By comparing the outcomes, we were able to identify the most effective model and feature set. This section defines the process that was followed for solving problems identified by the research and to complete the objectives of the research. Figure 1 illustrates the detailed methodology followed to complete each objective in order to attain the overall goal of the research starting from exploring and evaluating the existing models in literature survey to obtain the techniques and base of the research to proposing the model and implementing it to achieve the best outcomes by comparing the results with the existing models based on LSTM.

TABLE I. Present State of Art

Authors	Platform	Approach	Dataset	Features	Result
Sahana V P et. al. (2015)	Twitter	J48 decision tree Classifier	Tweets from Twitter	Linguistic, Network	Accuracy: .937
Liang et al. (2015)	Sina Weibo	Multiple classifiers	Microblog data from Sina Weibo	Linguistic, User	F1 score: 0.55 - 0.86
Mao et al. (2016)	Sina Weibo	Sentiment and Semantic Analysis	Ma-Weibo	Shallow statistical features and sentiment orientation deep features	F1 score: 0.39 - 0.46
Lukasik et al. (2016)	Twitter	Hawkes Processes	Ottawa shooting; Ferguson riots; Charlie Hebdo; Sydney siege	Temporal and textual information	Accuracy: 0.677 - 0.729
Chang et al. (2016)	Twitter	Clustering	Clinton and Obama datasets	Linguistic, User, Visual	F1 score: 0.83 -0.86
Suchita Jain et al. (2017)	Twitter	Sentiment and Semantic Analysis	Twitter	Content, User	Accuracy: 0.6078
QiaoZhang et al. (2017)	Sina Weibo	SVM	Weibo	Content-based, user-based, content-user based features	Precision: 0.71; Recall Rate: 0.63
ZhiweiJin et al. (2017)	Twitter	TF-IDF and BM25, Word2Vec and Doc2Vec, Lexicon matching	Twitter and snopes.com	Textual	Precision: 0.947
Yan Zhang et al. (2017)	Sina Weibo	Autoencoder	Weibo	Linguistic, User, Post	Accuracy: 0.88 F1 score: 0.82
Chen et al. (2018b)	Sina Weibo	RNN and Autoencoders	Weibo and comments	Content and Network	F1 score: 0.89
Ajao et al. (2019)	Twitter	SVM	PHEME	Linguistic, User, Post	Accuracy: 0.86
Ghanem et al. (2019)	Twitter	EIN	Newsarticle and Twitter	Linguistic	Accuracy: 0.9635 F1 score – 0.9633
Shu et al. (2019)	Twitter	Social Article Fusion (SAF) model	FakeNewsNet	Linguistic, User, Post, Network	Accuracy: 0.543 – 0.684 F1 score: 0.555 – 0.731
Abdullah Alsaeedi et al., 2020	Twitter	Word Embedding with CNN classifier	PHEME	Linguistic	Accuracy: 0.87
Neetu rani et al., 2021 90	Twitter	Word embedding (GloVe) with CNN+ BiLSTM	Kaggle	Contextual	Accuracy: 0.
Tokpa et al., 2023	PolitiFact and Wikipedia	DeepCnnBilstm and DeepCnnLstm	ISOT and FA-KES	Spatial and contextual features	Accuracy: 0.52 – 0.54 Precision: 0.53 – 0.55
Hemza Loucif, 2024	Twitter	Principal Component Analysis (PCA) with (CNN)	HoneyPot	Content and user-based; hybrid techniques	Accuracy: 0.95

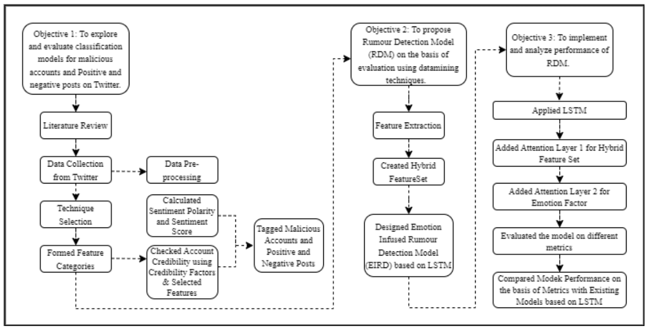


Figure 1. Comprehensive Research Methodology for Achieving the Objectives

A. Data Description

The EIN model used two datasets one NewsArticles including satire, hoaxes, propaganda with max length of words restricted to 300 words and second Twitter built by collecting limited number of tweets from 32 twitter accounts and merge it with the list of tweets obtained from annotated suspicious Twitter accounts based on public resources [57] whereas The LSTM model undergoes validation using the Pheme dataset, which is publicly accessible and consists of tweets pertaining to 5 distinct occasions i.e. Charlie Hebdo, Ferguson, German wings crash, Ottawa Shooting and Sydney Siege [61]. The validity of the proposed model has been confirmed through the utilization of the openly accessible RumourEval2019 dataset shown in Table III that includes rumour and non-rumour tweets and replies on the rumour tweets and non-rumour tweets with four classes- support, deny, query and comment. The RumourEval2019 test set contains 56 threads regarding tweets on natural disasters from Twitter and 25 thread set of Reddit data.

B. Feature Extractions

We collected a set of twenty-one different features from the tweets, including linguistic, user, post, and network aspects which were used for training and testing of the model [72]. The features based on linguistic characteristics are (i) presence of any question, (ii) tweets with words of supportive nature, (iii) tweets with words of denial nature, (iv) Tweet length - word count, (v) quoted words, and (vi) existence of links and graphs; User based features are (vii) account credibility- verified accounts or not, (viii) profile picture, (ix) account creation date, (x) tweets having URL and media, (xi) numbers of followers and followings, (xii) Number of hashtags, (xiii) tweet status count, and (xiv) Retweet count; Post based features are (xv) Emoticons, (xvi) Time series of post – to detect behaviour of post, (xvii) Vocabulary, and (xviii) Opinion – supporting, denying or querying; Network based features are (xix) User engagement, (xx) Influential account, (xxi) Association.

Table IV contains a full account of the features. These features may have varying variances that could overshadow other features when training classifiers. In order to best represent the features for the classifiers, the data is standardized. Each feature is standardized separately, with a mean of 0 and a SD (standard deviation) of 1. The features considered on the basis of results are then utilized by the model to categorize tweets as either rumours or non-rumours. The comprehensive outcomes of the model are shown in the 4th section of paper.

Figure 2. Most common emotions included in rumours.

C. Emotion Factor

Researches in the past show that emotions have prominent impact on rumour detection. Many incidents have been quoted in the past which shows the major reason behind the spreading of a rumour was to hamper the emotions of an individual or group [29]. Also, any rumour that spreads possess some emotion i.e., either negative or positive. It helps in detecting the veracity of rumour on the basis of emotions that the tweet possess. Emotions can be depicted on the basis of the words and phrases used in the tweet. Figure 2 shows the most common emotions in the tweets termed as rumours. To tag any tweet as Positive/ Negative we need to find sentiment polarity of the tweet and the sentiment score. Sentiment polarity means user's sentimental segmentation of the comments whether they are Positive / negative or neutral in respect to the thought entity. SP=(positive, p¿n; negative,p¡n) equation 1

Where, p is number of positive words n is number of negative words Polarity value in equation1 is increased by 1 with each positive word and decreased by 1 with each negative word. Sentiment score (SS) will be calculated using the sentiment dictionary. From equation1

SS= p/((p+n)) equation 2

The sentiment polarity obtained from equation 1 and sentiment score calculated from equation 2 that will be used to depict the positive and negative nature of the post.

Emoticons are good elements to identify sentiment. According to emoticon dictionary convert emoticons to words and analyze the meaning of words to find positive and negative emotions of the user and calculate the emotion quotient of tweet.

Process for calculating the emotion-quotient:

- Count the number of emoticons.

TABLE III. RumourEval2019

RumourEval2019	
support	1184 (14 percent)
deny	606 (7 percent)
query	608 (7 percent)
comment	6176 (72 percent)
Total	8574

TABLE IV. Feature Categories and types considered with description

Type	Feature
Linguistic	• Question existence • Tweet with supportive words • Tweet with words of denial nature • Tweet length – word count • Quoted words • Links and graphs
User	• Account credibility i.e Verified/ Unverified user • Profile picture • Account creation date • URL and media in tweet • Number of followers and followings • Number of hashtags • Status count of tweets • Retweet count of tweet
Post	• Emoticons • Time series of post – to detect behaviour of post • Vocabulary • Opinion – supporting, denying or querying
Network	• User engagement • Influential account • Association

- Match with emoticon dictionary.
- Convert the emoticons into words.
- Identify sentiment from words.

Whenever some information is posted over a social media platform, the information is instantly grasped by the social media and the circulation of that information starts at a great pace. There could be multiple reasons for the propagation of rumours. Figure 3 highlights some of those reasons: In the process of rumour generation, parameter selection is an important step as rumour generation process is dependent on parameters. These parameters are user specified. Users define the set of parameters and rumour is generated accordingly to fit those parameters.

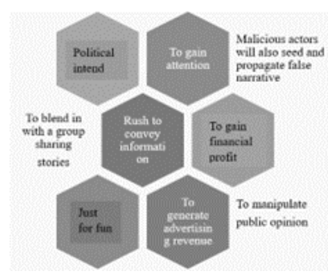


Figure 3. Reasons for spreading rumours

D. Word Embedding

Using a deep learning model, the contextual details of tweet content can be efficiently maintained while also eliminating the need for manually created features. For converting tweet into a dimension of fixed vector, word embedding technique is used. In word embedding, each tweet is represented as a n-dimensional dense vector given to deep learning models. It also manages to retain the semantic relationship among words. WE are like huge nexus of words that have semantically similar words forming clusters built with the help of a complex algo that establishes semantic relationships among words based on their usage in millions of sentences.

$$
\begin{array}{cc}
 & \begin{array}{ccccc} W_1 & W_2 & W_3 & \dots & W_m \end{array} \\
T_i = & \begin{bmatrix} e_{11} & e_{21} & e_{31} & \cdots & e_{m1} \\ e_{12} & e_{22} & e_{32} & \cdots & e_{m2} \\ e_{13} & e_{23} & e_{33} & \cdots & e_{m3} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ e_{1k} & e_{2k} & e_{3k} & \cdots & e_{mk} \end{bmatrix}
\end{array}
$$

Figure 4. embedded Tweet matrix

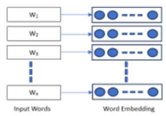


Figure 5. embedding with input words

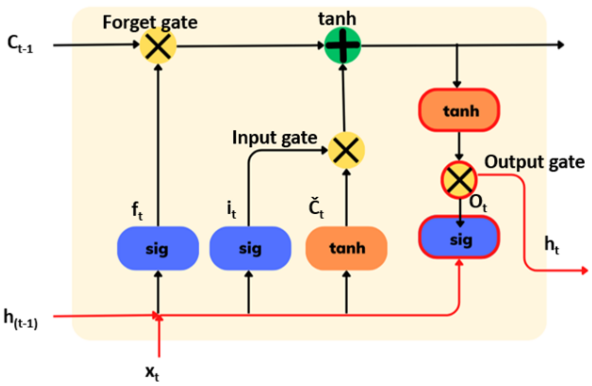


Figure 6. LSTM Layer Implementation

Matrix Ti represents the complete embedded tweet as shown in Figure 5 and Figure 5: The matrix of an embedded tweet, Ti, contains m words and will be padded if necessary. The purpose of padding is to ensure that all tweets are the same length. The vector [em1em2.....emk] represents the embedding for the word Wm, while k represents the dimension of the embedding. For this study, m is set at thirty-two, meaning that tweets having thirty-two words or more will be shortened, and tweets having word length not more than thirty-two will be padded to reach a length of thirty-two words.

E. Long-Short Term Memory (LSTM)

Specifically designed to receive these numerical representations as input, the LSTM model predicts whether the input contains rumours. Within the LSTM, gates are crucial part in determining relevant information for this prediction. The model undergoes training using labelled data, where tweets or tweet sequences are categorized as either "rumour" or "non-rumour." Through this training process, the LSTM learns to adjust its gate activations and cell state, enabling accurate predictions based on the provided data. Post-training, the LSTM model can then be applied to new Twitter data for real-time rumour detection. Figure 6 shows the layer-by-layer implementation of LSTM along with the gates placement and use to obtain the final output. The text of tweet is processed by a LSTM model [73] that consists of two layers, with the first layer having 200 dimensions hidden state vector and the second layer having 100 dimensions hidden state vector. The model is capable of training for 150 epochs using both rumour and non-rumour tweets. The features are extracted from the output of the second layer, which has a hidden state vector of 100 dimensions. For further processing these twenty-one linguistic, user, post and network features are concatenated with 100-dimensional feature map to build a 121-dimensional hybrid feature set.

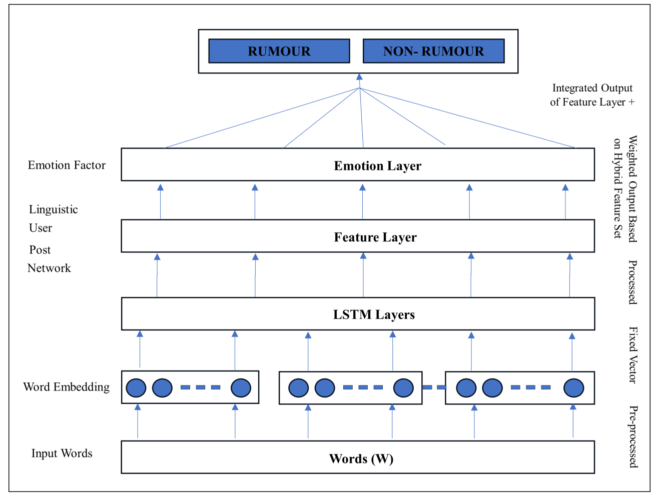


Figure 7. Layered architecture of Emotion Infused Rumour Detection model using LSTM

F. Emotion Infused Rumour Detection (EIRD) model using LSTM

As demonstrated in 2016 by Yang et al. about the success of various attention-based methods in NLP motivated us to develop a model to identify rumour tweets similar to those [74]. The model focuses on efficiently learning unique textual features. The attention layer is responsible for understanding the importance of each element in the sequence of input and then combining them to extract important data. The attention-based mechanism in detail is explained by Vaswani et al. [75]. In Emotion Infused Rumour Detection model using LSTM, after the second layer i.e. of LSTM, two attention layers are added- one for features with hybrid set of features while second is for emotion factor. Then twenty-one features - linguistic, user, post and network are concatenated to the first layer output which is the first attention layer and emotion feature is added to the second output layer which is the second attention layer. Figure 7 illustrates the organized illustration of the Emotion Infused Rumour Detection model using LSTM. Ultimately, the combined feature map is utilized to categorize tweets as rumours or non-rumours.

4. RESULTS

The results of all the three models considered are discussed and are being compared in this section of paper – EIN, LSTM and Emotion Infused Rumour Detection model using LSTM. The Emotion Infused Rumour Detection model using LSTM was evaluated through a 5-fold cross-validation method to assess its performance. The reason behind selecting 5-fold cross-validation is due to the dataset containing 8574 tweets, both rumour and non-rumour. As the number of folds increases, the size of the testing data decreases significantly.

A. *Metrics for Evaluation*

The evaluation of the models is done using the following parameters:

Precision (P) – The ratio of correctly predicted rumour tweets to the total number of predictions.

P =TP/(TP+FP) equation 3

Recall (R) – The ratio of correctly predicted tweets about rumours to the total number of tweets actually about rumours.

R=TP/(TP+FN) equation 4

F1 score (F1): harmonic mean between Precision and Recall.

F1=2 x (P x R)/(P+R) equation 5

The F1 score provides a well-balanced assessment that takes into account both Precision and Recall.

Accuracy (A): The proportion of TP and TN in relation to the entire dataset.

A=(TP+TN)/(TP+FP+FN+TN) equation 6

TP: count of rumour predictions as Rumour. FP: count of tweets which are not rumours predicted as Rumour. TN: count of non-rumour predictions as Non-Rumour. FN: count of rumour predictions as Non-Rumour.

B. *Parameter Configuration*

For the implementation of the Emotion Infused Rumour Detection model using LSTM the parameters considered were two LSTM layers, two dense layers emphasizing on hybrid feature set and emotion factor with 256 and 128 dimensions of hidden state vector. The model used two neurons (dense). The experiment was done with various tuning parameters such as time lags, optimizers, learning rates, epochs, hidden units, and batch sizes to optimize the model. Initially, the focus was on identifying the ideal time lags when using the Adam optimizer, a learning rate of 0.01, 150 epochs, 256 and 128 hidden units, and a batch size of 100. Subsequently, the impact was explored of different optimizers on the optimal time lags. Then, comparison of the model's performance across different learning rates was done to determine the most suitable rate. Then sequentially the optimal number of epochs and hidden units were determined. Finally, a comparative analysis was conducted to identify the optimal batch size, taking into consideration memory constraints. Through systematic testing of various parameter combinations, which ultimately identified the model best suited for the dataset.).

TABLE V. Hyper parameter settings for each model

Parameters	LSTM	EIN	EIRD model using LSTM
Number of layers	LSTM-2, Dense-1	LSTM-1, Dense-2	LSTM-2, Dense-2
Dimension of hidden state vector	200,100	120,60	256,128
Number of neurons (Dense)	2	2	2
Activation	Softmax	Relu	Softmax
Optimizer	Adam	Adam	Adam
Batch size	100	32	100
Epochs	150	120	150

The proposed EIRD model utilized dual LSTM layers with 256 and 128-dimensional hidden state vectors for the initial and subsequent LSTM layers, respectively, to extract tweet text features. The model underwent training using both rumour and non-rumour tweets for 150 epochs. Following this, the output of the second layer, featuring a 128-dimensional hidden state vector, was retained as features. These 128-dimensional features were then combined with twenty-one linguistic, post, network, and user features to form a 149-dimensional hybrid feature set for additional processing.

As mentioned earlier, the performance of the proposed model was measured on the basis of comparative analyses considering three models based on DL and LSTM – LSTM, EIN and the proposed model EIRD. Table V shows the details of the model's configuration and hyper-parameters.

C. *Results*

We utilized the attention technique combined with LSTM to construct our model, employing tweet texts and incorporating twenty-one linguistic, post, network, and user features for LSTM while integrating attention and hybrid features into the model (EIRD) using LSTM. To evaluate the performance of the model, we conducted 5-fold cross validation. The outcomes of the deep learning-based model are presented in Table VI . Figure 8 of box-and-whisker plot demonstrates the comparison of F1-scores between the LSTM model without hybrid features and the LSTM model with hybrid features, as determined through 5-fold cross validation. Figure 9 shows the confusion matrix for fold-5 in case of the proposed LSTM model with hybrid features i.e. Emotion Infused Rumour Detection model using LSTM.

TABLE VI. Results of EIRD model on the basis of evaluation metrics

Models	Precision	Recall	F1 score	Accuracy
EIRD	90.50	91.72	91.70	89.02

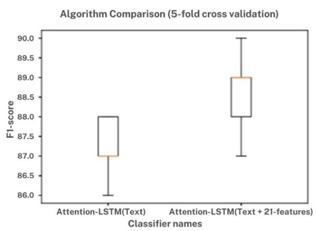


Figure 8. box-whisker showing 5-fold cross validation comparison between LSTM model without using Hybrid feature set and EIRD model using Hybrid feature set.

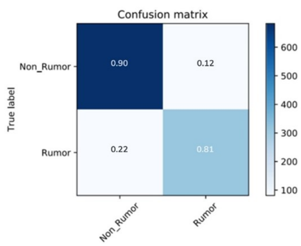


Figure 9. Confusion matrix for 5-fold

TABLE VII. Performance of EIRD Model on the Basis of Hybrid Feature Set

Models	Precision	Recall	F1 score	Accuracy
EIRD (without hybrid feature set)	90.00	89.00	89.00	86.00
EIRD (with hybrid feature set)	90.50	91.72	91.70	89.02

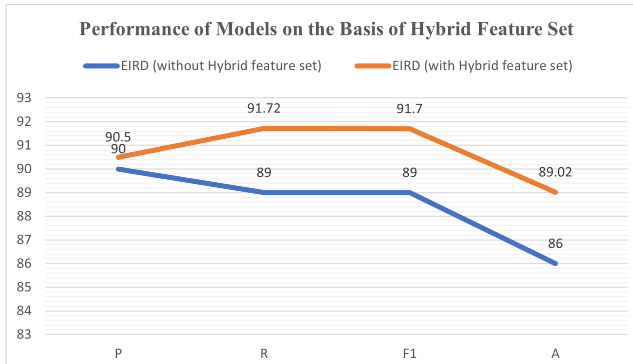


Figure 10. Performance of Models on the Basis of Hybrid Feature Set

The proposed model considered twenty-one manually extracted features. For comparative analysis of the proposed model, an experiment was conducted without using the hybrid feature set. This was done to show the difference between the outcomes obtained from the model considering hybrid feature set and without considering hybrid feature set.

The work has already explained the role of feature extraction and selection in rumour detection. The results shown in Table VII justifies the statement. In this comparative analysis the proposed model was first evaluated without using the hybrid feature set and then using the hybrid feature set then the results were compared. As shown in results it can be seen that the evaluation values significantly increased in the case when the hybrid features were considered. The accuracy obtained by the model not considering hybrid feature set was 86 percent whereas the accuracy obtained by the model considering hybrid feature set was 89.02 percent which shows an increase of 3.02 percent. Figure **??** is the graphical representation of the results obtained from this experiment of considering the hybrid feature set and without considering the feature set. The graph provides more clarity to the experiment claiming the role of feature extraction and of creating a hybrid feature set of twenty-one manually extracted features. The precision obtained in both cases were 90 and 90.5 percentage respectively. The recall value obtained in first case where no hybrid feature set was considered was 89 percent whereas in second case in which the hybrid features were considered obtained 91.72 percent of recall rate.

The next comparison was done on the basis of emotion factor as emotion is one of the major factors in the proposed model EIRD. Table V highlights the hyper parameters used. These parameters were used for EIN model considered for comparison with the proposed model to evaluate its performance on all scales considered. The results of both models are considering only linguistic feature and focused on emotion factor, no other feature category was considered for this experiment for comparative analysis. Table VIII illustrates the performance of the model using the hyper parameters selected for EIN model which was focused on emotions only. Only Linguistic feature was considered for the analysis.

Figure 11 depicts the graphical representation of the performance of the model on the basis of emotion factor.

TABLE VIII. Results Obtained for the Comparative Analysis of EIN and EIRD

Models	Precision	Recall	F1 score	Accuracy
EIN	95.74	96.97	96.35	96.31
EIRD	96.50	97.82	97.73	97.70

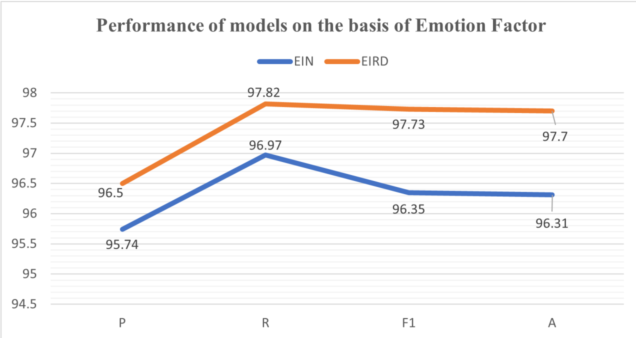


Figure 11. Performance of Models on the Basis of Hybrid Feature Set

As explained the model focuses on emotion factor being the emotion infused model. A significant increase in the values of evaluation metrics was experienced in this experiment as shown in Figure 6.8. As shown the value of accuracy obtained by EIN model was 96.31 percent whereas the accuracy percentage obtained by EIRD model was 97.70 considering only the linguistic features and emotion factor using the parameters selected for EIN.

5. DISCUSSION

The main discovery of the present study is that a model using LSTM with attention mechanism outperforms all other existing models in distinguishing tweets as either rumours or non-rumours. The twenty-one features extracted form post, network, linguistic, and user features manually contributed in classifying rumourous tweets, depicted in Table IV. Out of the DL models, Emotion Infused Rumour Detection model using LSTM was found to be particularly effective general rumour detection model considering two attention layers and multiple features, as it yielded higher accuracy compared to LSTM model, as shown in the Table IX.

The Table IX illustrates the outcomes of the two models discussed in research with all the evaluation metrics. The normal LSTM model provided accuracy of 88 percent with F1 score of 91. The table shows an increase in the accuracy in the model developed i.e., Emotion Infused Rumour Detection model using LSTM over LSTM model, the proposed model used more features and two attention layers which increased the model's efficiency. The model possesses the higher accuracy of 89.02 percent and F1 score of 91.70 which makes it most appropriate for detecting rumours in Twitter considering two attention layers and multiple features. The major short coming of previous models is the results are based on considering only single or combination of features considering dataset having tweets related to one or two incidents whereas the model proposed is a general model considering twenty-one linguistic, user, post and network features which makes it more useful and efficient. The graph in Figure 12 depicts the difference in the metrics values of LSTM and Emotion Infused Rumour Detection model using LSTM. The graph shows by using

TABLE IX. Performance of EIRD Model on the Basis of Hybrid Feature Set

Models	Precision	Recall	F1 score	Accuracy
LSTM	90.00	91.00	91.00	88.00
EIRD (with hybrid feature set)	90.50	91.72	91.70	89.02

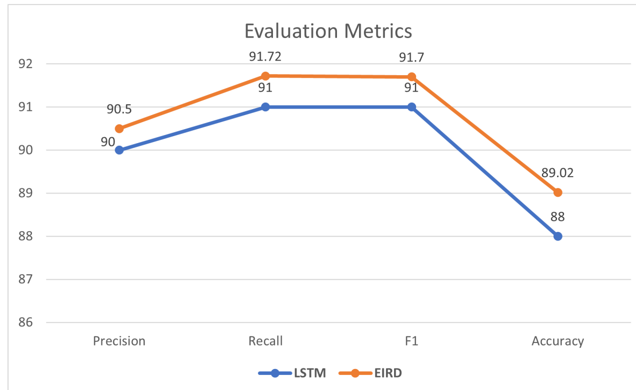


Figure 12. Comparison of Metrics Values of LSTM and Emotion Infused Rumour Detection Model Using LSTM

additional attention layer in the model it provided better results.

Our research aligns with the studies conducted previously in which it was observed that there is a strong positive association between a user's preexisting beliefs and the believability of an article [34][35]. Our analysis also found that an increase in the count of supportive or denial words in a tweet, which reflects the user's beliefs, leads to a higher accuracy in classification. Additionally, features related to the user's account, such as verification status, account age, and tweet count, were found to improve the accuracy of classification, which is similar to the approach proposed by one of the researchers [34]. While they both used Facebook and news sources, which provided a source rating, our study focused on Twitter. The rating of source can be obtained by factors such as inception of user account, tweet count, followers count and followees, and retweet count. Table IV shows breakdown of these features in detail.

A. *Theoretical Contributions*

The focal theoretical advancements of this study are the utilization of LSTM model using attention mechanism to classify tweets as rumours and non-rumours. This attention mechanism effectively captures the text that exhibits rumour-like behaviour. The attention layer, which is a sequential neural network component, concentrates on the words specifically in the existing in a particular class of tweets. Our research incorporates two attention layers to detect the words relevant for a class specifically.

TABLE X. Comparison of results of existing models with the proposed model

Authors	Technique/ Approach used	Features	Accuracy (percent)
Sahana v p et. al. (2015)	J48 decision tree Classifier using Synthetically generated training set [37]	Linguistic, Network	93.7
Yan zhang et al. (2017)	Auto encoder (Artificial Neural Network) [33]	Linguistic, User, Post	88
Bhutani et al. (2017)	TF-IDF and BM25, Word2Vec and Doc2Vec, Lexicon matching [34]	Linguistic, User, Post	79.9
Chang et al. (2018)	Clustering [38]	Linguistic, Visual, User	80
Sivasangari et al. (2018)	Sentiment and Semantic analysis [27]	Linguistic, User, Post	90
Ghanem et al. (2019)	Emotionally-Infused LSTM neural network (EIN)[35]	Linguistic	96.31
Ajao et al. (2019)	Support vector machine [30]	Linguistic, User, Post	89
Shu et al. (2019)	Social Article Fusion (SAF) model [36]	Linguistic, User, Post, Network	74.20
Proposed model	Emotion Infused Rumour Detection (EIRD) model using LSTM	Linguistic, User, Post, Network	89.02

Additionally, we also introduced a feature set hybrid in nature by extracting manually the features - linguistic, post, network and user from tweets, along with textual features from LSTM models. Previous studies have shown that using limited features in machine learning classifiers had limited success, as indicated in Table X. However, LSTM models with automatic text feature extraction also reached a limit, shown in Table VI. The previous study of some researchers resulted in precision scores of 0.83 and 0.86, which depicted that the combination of LSTM (BiLSTM) and CNN in hybrid deep learning models [68][70]. However, our attention model outperformed these models by achieving a precision score of 0.90, thanks to the incorporation of hybridization. This confirms the effectiveness of hybridization in rumour detection.

B. Implications for Practice

Emotion Infused Rumour Detection model using LSTM model had the ability to promptly and accurately recognize tweets containing rumours, aiding in the prevention of the spread of misinformation. By doing so, the system being proposed was able to mitigate the negative effect of rumours on society, as well as promoting trust in social media platforms. One practical application of this system was its potential to be developed as a smartphone app that categorizes tweets as either rumours or non-rumours. The existing system has some limitations, it only takes into account the text and user features of a tweet for this study. Other elements of a tweet, including pictures, graphics, audio, visuals, GIFs, memes, and links, could also assist in identifying tweets containing rumours. Another restriction of the model is that tweets in English language only are used to validate the model. This model may not yield the same results for other languages and multilingual tweets, which are prevalent in many non-English speaking countries. In the future, the existing system could be expanded to better align with the design science research guidelines [76][77].

C. Limitations and Future Research

The social media or microblogging platforms play a crucial role in the propagation of information irrespective of its accuracy and verification [78].Detection at an early stage plays crucial role in managing the situation effectively [79]. LSTM networks are a kind of RNN that include long short-term memory cells, allowing the RNN to recall the previous output for a longer time period [80][81].A drawback of the current research is that the model has only been tested and confirmed using data from only Twitter. This could potentially limit its usefulness on other social media platforms. Additionally, the focus has solely been on linguistic, user, post and network features. While text is the primary medium for spreading rumours, there are other elements that contribute to a rumour, such as images, videos, and emoticons. Therefore, a comprehensive rumour detection system should also incorporate these features. Furthermore, the proposed work is limited by its dependence on language, as it has only been trained and authenticated using English language tweets. This model may not yield the same results for other languages and multilingual tweets, which are prevalent in many non-English speaking countries. In this research in future, gathering of data considering various platforms of social media can be utilized to effectively verify the findings and make them more applicable to the model. Additionally, features like URL, emoticons, gifs, images, and videos may also be incorporated with the text. The model suggested is a supervised one, necessitating a substantial amount of labelled data for adequate training and authentication. As a future possibility, unsupervised models and GAN could be created to decrease or eliminate the necessity for a labelled dataset.

6. Conclusion

The detection of rumour accurately on Twitter from tweets is a crucial task. This research compares and implements the performance of various models based on deep learning to recognize rumour tweets at an early stage. To create a feature set hybrid in nature, twenty-one features are extracted from tweet related to linguistic, post, network and user, while using the LSTM model, extracted 100 features from text. The models are learned with the created feature set having hybrid nature. The improvised algorithm based on population is utilized for selecting the features in optimum number from the created feature set, which resulted in a reduction of over 20percent in total features. The findings of the experiment demonstrate the superiority of the deep learning-based model for identifying rumours. The proposed Emotion Infused Rumour Detection model using LSTM, with attention mechanism and created feature set hybrid in nature surpasses all prevailing models reaching F1-score of more than 0.91. Multiple experiments were carried out to get the best possible results of the model and also to evaluate and compare the results at different levels. In order to evaluate the performance of the model two other models were also considered so as to perform a complete comparative analysis on the basis of evaluation metrics. All the models were based on LSTM. Different parameters were considered for each experiment and the results were compared with the outcomes of the models used for comparison. The experiments included model-based LSTM using and without using hybrid feature set and on the basis of emotion factor. In conclusion, the model that used two LSTM layers, two dense layers emphasized on hybrid feature set and emotion factor with 256 and 128 dimensions of hidden state vector; that used two neurons (dense); softmax activation; Adam optimizer at learning rate of 0.001 using batch size 100 obtained the best outcome at 150 epochs. The main outcome of this work was that the EIRD model using LSTM based on attention mechanism outperforms all other existing models in distinguishing tweets as either rumours or non-rumours.

References

[1] Y. K. Dwivedi, K. K. Kapoor, and H. Chen, "Social media marketing and advertising," *The Marketing Review*, vol. 15, no. 3, pp. 289–309, 2015.

[2] A. Kumar and N. C. Rathore, "Relationship strength based access control in online social networks," in *Proceedings of First International Conference on Information and Communication Technology for Intelligent Systems:*, 2016, vol. 2, pp. 197–206.

[3] A. A. Alalwan, N. P. Rana, Y. K. Dwivedi, and R. Algharabat, "Social media in marketing: A review and analysis of the existing literature," *Telematics and informatics*, vol. 34, no. 7, pp. 1177–1190, 2017.

[4] M. A. A. Alryalat, N. P. Rana, G. P. Sahu, Y. K. Dwivedi, and M. Tajvidi, "Use of social media in citizen-centric electronic government services: A literature analysis," *International Journal of Electronic Government Research (IJEGR)*, vol. 13, no. 3, pp. 55–79, 2017.

[5] K. Tamilmani, N. Rana, M. Alryalat, W. Alkuwaiter, and Y. Dwivedi, "Social media research in the context of emerging markets: an analysis of literature published in senior scholars basket of is journals," *Journal of Advances in Management Research*, vol. 15, no. 2, pp. 115–129, 2018.

[6] M. A. Shareef, B. Mukerji, Y. K. Dwivedi, N. P. Rana, and R. Islam, "Social media marketing: Comparative effect of advertisement sources," *Journal of Retailing and Consumer Services*, vol. 46, pp. 58–69, 2019.

[7] J. P. Singh, Y. K. Dwivedi, N. P. Rana, A. Kumar, and K. K. Kapoor, "Event classification and location prediction from tweets during disasters," *Annals of Operations Research*, vol. 283, no. 32., 2019.

[8] A. Kumar, J. P. Singh, and N. P. Rana, "Authenticity of geo-location and place name in tweets," in *Proceedings of 23rd Americas Conference on Information Systems (AMCIS)*, 2017.

[9] A. Kumar and J. P. Singh, "Location reference identification from tweets during emergencies: A deep learning approach," *International journal of disaster risk reduction*, vol. 33, pp. 365–375, 2019.

[10] A. Kumar, J. P. Singh, Y. K. Dwivedi, and N. P. Rana, *A deep multi-modal neural network for informative Twitter content classification during emergencies.* Annals of Operations Research, 2022.

[11] B. Abedin and A. Babar, "Institutional vs," *non-institutional use of social media during emergency response: A case of twitter in 2014 Australian bush fire. Information Systems Frontiers*, vol. 20, pp. 729–740, 2018.

[12] S. Ghosh, K. Ghosh, D. Ganguly, T. Chakraborty, G. J. Jones, M. F. Moens, and M. Imran, "Exploitation of social media for emergency relief and preparedness: Recent research and trends," *Information Systems Frontiers*, vol. 20, pp. 901–907, 2018.

[13] K. K. Kapoor, K. Tamilmani, N. P. Rana, P. Patil, Y. K. Dwivedi, and S. Nerur, "Advances in social media research: Past, present and future," *Information Systems Frontiers*, vol. 20, pp. 531–558, 2018.

[14] H. Kizgin, A. Jamal, B. L. Dey, and N. P. Rana, "The impact of social media on consumers acculturation and purchase intentions," *Information Systems Frontiers*, vol. 20, pp. 503–514, 2018.

[15] A. M. Baabdullah, N. P. Rana, A. A. Alalwan, R. Algharabat, H. Kizgin, and G. A. Al-Weshah, "Toward a conceptual model for examining the role of social media on social customer relationship management (SCRM) system," in *Smart Working, Living and Organising: IFIP WG 8*, 2018, vol. 6, pp. 102–109.

[16] O. Oh, M. Agrawal, and H. R. Rao, "Information control and terrorism: Tracking the Mumbai terrorist attack through twitter," *Information Systems Frontiers*, vol. 13, pp. 33–43, 2011.

[17] P. Singh, Y. K. Dwivedi, K. S. Kahlon, R. S. Sawhney, A. A. Alalwan, and N. P. Rana, *Smart monitoring and controlling of government policies using social media and cloud computing*. Information Systems Frontiers, 123, 2019.

[18] S. Vallejos, D. G. Alonso, B. Caimmi, L. Berdun, M. G. Armentano, and Soria, "Mining social networks to detect traffic incidents." *Information systems frontiers*, vol. 23, no. 1, pp. 115–134, 2021.

[19] J. Ma, W. Gao, P. Mitra, S. Kwon, B. J. Jansen, K. F. Wong, and M. Cha, "Detecting rumors from microblogs with recurrent neural networks," in *Proceedings of the 25th International Joint Conference on Artificial Intelligence (IJCAI 2016)*, 2016, pp. 3818–3824.

[20] T. Mondal, P. Pramanik, I. Bhattacharya, N. Boral, and S. Ghosh, "Analysis and early detection of rumors in a post disaster scenario," *Information Systems Frontiers*, vol. 20, pp. 961–979, 2018.

[21] J. P. Singh, N. P. Rana, and Y. K. Dwivedi, "Rumour veracity estimation with deep learning for Twitter," in *Dwivedi*, A. Y., B. E., E. R., and J., Eds. Cham: Springer International Publishing, 2019, p. 351363.

[22] R. Aswani, A. K. Kar, V. Ilavarasan, and P., "Detection of spammers in twitter marketing: a hybrid approach using social media analytics and bio inspired computing," *Information Systems Frontiers*, vol. 20, pp. 515–530, 2018.

[23] K. S. Smith, R. McCreadie, C. Macdonald, and I. Ounis, "Regional sentiment bias in social media reporting during crises," *Information Systems Frontiers*, vol. 20, pp. 1013–1025, 2018.

[24] N. DiFonzo and P. Bordia, *Defining rumour*. Rumour psychology : Social and organizational approaches, 1134, 2007.

[25] H. Huang, "A war of (mis) information: The political effects of rumors and rumor rebuttals in an authoritarian country," *British Journal of Political Science*, vol. 47, no. 2, pp. 283–311, 2017.

[26] G. Liang, W. He, C. Xu, L. Chen, and J. Zeng, "Rumor identification in microblogging systems based on users behavior," *IEEE Transactions on Computational Social Systems*, vol. 2, no. 3, pp. 99–108, 2015.

[27] M. L. Khan and I. K. Idris, "Recognise misinformation and verify before sharing: a reasoned action and information literacy perspective," *Behaviour Information Technology*, vol. 38, no. 12, pp. 1194–1212, 2019.

[28] J. Lee, M. Agrawal, and H. R. Rao, "Message diffusion through social network service: The case of rumor and non-rumor related tweets during Boston bombing 2013," *Information Systems Frontiers*, vol. 17, pp. 997–1005, 2015.

[29] R. R, S. A, K. S, J. S, J. TA, J. A, P. P, and G. D, *2018 Murderous mob 9 states, 27 killings*, 2018.

[30] P. Domm, *False Rumour of Explosion at White House Causes Stocks to Briefly Plunge; AP Confirms Its twitter Feed Was Hacked*. Retrieved on April23, 2013.

[31] P. Meel and D. K. Vishwakarma, "Fake news, rumor, information pollution in social media and web: A contemporary survey of state-of-the-arts, challenges and opportunities." *Expert Systems with Applications*, vol. 153, p. 112986., 2020.

[32] M. G. Lozano, J. Brynielsson, U. Franke, M. Rosell, E. Tjrnhammar, S. Varga, and V. Vlassov, "Veracity assessment of online data." *Decision Support Systems*, vol. 129, p. 113132., 2020.

[33] E. Serrano, C. A. Iglesias, and M. Garijo, "A survey of twitter rumor spreading simulations," 2015.

[34] A. Kim, P. L. Moravec, and A. R. Dennis, "Combating fake news on social media with source ratings: The effects of user and expert reputation ratings," *Journal of Management Information Systems*, vol. 36, no. 3, pp. 931–968, 2019.

[35] A. Kim and A. R. Dennis, "Says who? The effects of presentation format and source rating on fake news in social media," *Mis quarterly*, vol. 43, no. 3, pp. 1025–1039, 2019.

[36] S. Kwon, M. Cha, and K. Jung, "Rumor detection over varying time windows," *PloS one*, vol. 12, no. 1, 2017.

[37] J. Ma, W. Gao, Z. Wei, Y. Lu, and K. F. Wong, *Detect rumors using time series of social context information on microblogging websites*. Proceedings of the 24th ACM international on conference on information and knowledge management, 2015.

[38] L. Derczynski, K. Bontcheva, M. Liakata, R. Procter, G. W. S. Hoi, and A. Zubiaga, "Semeval-2017 Task 8: RumourEval: Determining rumour veracity and support for rumours," *arXiv preprint arXiv:1704*, p. 05972., 2017.

[39] O. Enayet and S. R. El-Beltagy, "Niletmrg at semeval-2017 task 8: Determining rumour and veracity support for rumours on twitter," *Proceedings of the 11th international workshop on semantic evaluation*, pp. 470–474, 2017.

[40] T. Chen, X. Li, H. Yin, and J. Zhang, "Call attention to rumors: Deep attention based recurrent neural networks for early rumor detection," in *Trends and Applications in Knowledge Discovery and Data Mining: PAKDD 2018 Workshops, BDASC, BDM, ML4Cyber, PAISI, DaMEMO, Melbourne, VIC, Australia, June 3*, 2018, pp. 40–52.

[41] C. Castillo, M. Mendoza, and B. Poblete, *Information credibility on twitter*. Proceedings of the 20th international conference on World wide web, 2011.

[42] V. Qazvinian, E. Rosengren, D. Radev, and Q. Mei, *Rumor has it: Identifying misinformation in microblogs*. Proceedings of the 2011 conference on empirical methods in natural language processing, 2011.

[43] S. Jain, V. Sharma, and R. Kaushal, *Towards automated real-time detection of misinformation on Twitter*. international conference on advances in computing, communications and informatics (ICACCI), 2016.

[44] E. Mao, G. Chen, X. Liu, and B. Wang, "Research on detecting micro-blog rumors based on deep features and ensemble classifier," *Application Research of Computers*, vol. 33, no. 11, pp. 3369–3373, 2016.

[45] V. Sivasangari, A. K. Mohan, K. Suthendran, and M. Sethumadhavan, "Isolating rumors using sentiment analysis," *Journal of Cyber Security and Mobility*, vol. 7, no. 1, pp. 181–200, 2018.

[46] L. WY., *Research on Microblog Rumours Detection Pattern Based on Sentiment Analysis*. Chongqing University, 2016.

[47] Q. Zhang, S. Zhang, J. Dong, J. Xiong, and X. Cheng, *Automatic Detection of Rumor on Social Network, In: Springer International Publishing Switzerland 2015, Springer*, 2017.

[48] O. Ajao, D. Bhowmik, and S. Zargari, "Sentiment aware fake news detection on online social networks." *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 2507–2511, 2019.

[49] J. Ma, W. Gao, and K. F. Wong, "Detect Rumors in Microblog Posts Using Propagation Structure via Kernel Learning," in *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume*, 2017, vol. 1.

[50] Y. Liu, X. Jin, H. Shen, and X. Cheng, "Do rumors diffuse differently from non-rumors? a systematically empirical analysis

in sina weibo for rumor identification," in *Advances in Knowledge Discovery and Data Mining: 21st Pacific-Asia Conference, PAKDD 2017, Jeju, South Korea, May 23-26*, 2017, pp. 407–420.

[51] Y. C. Chen, Z. Y. Liu, and H. Y. Kao, "Ikm at semeval-2017 task 8: Convolutional neural networks for stance detection and rumor verification," *Proceedings of the 11th international workshop on semantic evaluation*, pp. 465–469, 2017.

[52] A. Zubiaga, M. Liakata, and R. Procter, "Exploiting context for rumour detection in social media," in *Social Informatics: 9th International Conference, SocInfo 2017, Oxford, UK, September 13-15*, 2017, pp. 109–123.

[53] Z. Jin, J. Cao, H. Guo, Y. Zhang, Y. Wang, and J. Luo, "Detection and analysis of 2016 US presidential election related rumors on twitter," in *Social, Cultural, and Behavioral Modeling: 10th International Conference, SBP-BRiMS 2017, Washington, DC, USA, July 5-8*, 2017, pp. 14–24.

[54] B. Bhutani, N. Rastogi, P. Sehgal, and A. Purwar, *Fake news detection using sentiment analysis.)*. twelfth international conference on contemporary computing (IC3), 2019.

[55] K. Shu, D. Mahudeswaran, and H. Liu, "Fakenewstracker: a tool for fake news collection, detection, and visualization," *Computational and Mathematical Organization Theory*, vol. 25, pp. 60–71, 2019.

[56] Y. Zhang, W. Chen, C. K. Yeo, C. T. Lau, and B. S. Lee, *Detecting rumors on online social networks using multi-layer autoencoder)*. IEEE technology engineering management conference (TEMSCON), 2017.

[57] B. Ghanem, P. Rosso, and F. Rangel, "An emotional analysis of false information in social media and news articles." *ACM Transactions on Internet Technology (TOIT)*, vol. 20, no. 2, pp. 1–18, 2020.

[58] Z. Zhao, P. Resnick, and Q. Mei, *Enquiring minds: Early detection of rumors in social media from enquiry posts*. Proceedings of the 24th international conference on world wide web, 2015.

[59] C. Chang, Y. Zhang, C. Szabo, and Q. Z. Sheng, "Extreme user and political rumor detection on twitter," 2016.

[60] V. P. Sahana, A. R. Pias, R. Shastri, and S. Mandloi, *Automatic detection of rumoured tweets and finding its origin*. International Conference on Computing and Network Communications (CoCoNet), 2015.

[61] A. Zubiaga, M. Liakata, R. Procter, W. S. Hoi, G., and P. Tolmie, "Analysing how people orient to and spread rumours in social media by looking at conversational threads," *PloS one*, vol. 11, no. 3, 2016.

[62] M. Lukasik, P. K. Srijith, D. Vu, K. Bontcheva, A. Zubiaga, and T. Cohn, *Hawkes processes for continuous time sequence classification: an application to rumour stance classification in twitter*. Association for Computational Linguistics, 2016.

[63] S. Hamidian and M. Diab, *Rumor identification and belief investigation on twitter*. Proceedings of the 7th Workshop on computational approaches to subjectivity, sentiment and social media analysis, 2016.

[64] A. Jain, V. Borkar, and D. Garg, "Fast rumor source identification via random walks," *Social Network Analysis and Mining*, vol. 6, pp. 1–13, 2016.

[65] O. Oh, P. Gupta, M. Agrawal, and H. R. Rao, "Ict mediated rumor beliefs and resulting user actions during a community crisis," *Government Information Quarterly*, vol. 35, no. 2, pp. 243–258, 2018.

[66] W. Chen, Y. Zhang, C. K. Yeo, C. T. Lau, and B. S. Lee, "Unsupervised rumor detection based on users behaviors using neural networks," *Pattern Recognition Letters*, vol. 105, pp. 226–233, 2018.

[67] B. RATH, W. GAO, J. MA, and J. SRIVASTAVA, "From retweet to believability: Utilizing trust to identify rumor spreaders on Twitter," in *Proceedings of the IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2017)*, 2017, pp. 179–186.

[68] O. Ajao, D. Bhowmik, and S. Zargari, *Fake news identification on twitter with hybrid cnn and rnn models*. Proceedings of the 9th international conference on social media and society, 2018.

[69] A. Zubiaga, A. Aker, K. Bontcheva, M. Liakata, and R. Procter, "Detection and resolution of rumours in social media: A survey," *Acm Computing Surveys (Csur)*, vol. 51, no. 2, pp. 1–36, 2018.

[70] M. Z. Asghar, A. Habib, A. Habib, A. Khan, R. Ali, and A. Khattak, "Exploring deep neural networks for rumor detection." *Journal of Ambient Intelligence and Humanized Computing*, vol. 12, pp. 4315–4333, 2021.

[71] Y. Liu, X. Jin, and H. Shen, "Towards early identification of online rumors based on long short-term memory networks," *Information Processing Management*, vol. 56, no. 4, pp. 1457–1467, 2019.

[72] S. O. and A. S., *Feature Extraction in Rumour Detection Model: Classification Significance.* Design Engineering, 2021.

[73] O. Sharma and S. Ahuja, "Novel framework for rumor detection using emotionally infused LSTM." *ECS Transactions*, vol. 107, no. 1, 2022.

[74] Z. Yang, D. Yang, C. Dyer, X. He, A. Smola, and E. Hovy, *Hierarchical attention networks for document classification*. Proceedings of the 2016 conference of the North American chapter of the association for computational linguistics: human language technologies, 2016.

[75] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, and . P. Gomez, A. N., "Attention is all you need." *Advances in neural information processing systems*, vol. 30., 2017.

[76] S. T. March and G. F. Smith, "Design and natural science research on information technology," *Decision support systems*, vol. 15, no. 4, pp. 251–266, 1995.

[77] R. Baskerville, A. Baiyere, S. Gregor, A. Hevner, and M. Rossi, "Design science research contributions: Finding a balance between artifact and theory," *Journal of the Association for Information Systems*, vol. 19, no. 5, 2018.

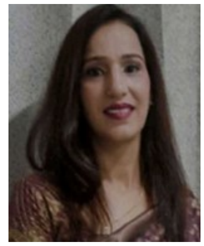
[78] P. Dhiman, A. Kaur, and A. Bonkra, *Fake Information Detection Using Deep Learning Methods: A Survey*. 2023 International Conference on Artificial Intelligence and Smart Communication (AISC), Greater Noida, India, 2023.

[79] A. K, K. A, G. Y, and H. Y, "Mir MS and Soomro AB (2023) Deep learning models/techniques for COVID-19 detection: a survey," *Front. Appl. Math. Stat*, vol. 9, p. 1303714, 2023.

[80] N. B. . A. Singla, "Deep learning based wheat crop yield prediction model in punjab region of north india," *Applied Artificial Intelligence*, vol. 35.

[81] A. K. R. Samrat, A. Bonkra and P. Dhiman, *Deep Learning Techniques for Sentiment Analysis: A Comparative Study)*. T6th International Conference on Contemporary Computing and Informatics (IC3I), 2023.

Ms. Osheen Sharma is currently working as Assistant Professor in Department of Information Technology, GGDSD College, Chandigarh. She has teaching experience of 8.5 years and has published paper in reputed international journals and conference which are also available online. Her main research work focuses on deep learning.



Dr. Monika Sethi is presently serving as Associate Professor in Department of Computer Science and Engineering Chitkara University, Punjab. She has more than a decade of extensive experience in the field of teaching and have published her work in various international journals and conferences.

Dr Sachin Ahuja is presently serving as the Executive Director of Engineering responsible for overseeing and leading the Engineering department previously. He has more than 2 decades of extensive experience and expertise in Engineering, Project Management, and Technical Leadership.