IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 9, Issue 5, May 2021

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 7.488

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 5, May 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0905176 |

Rumour Detection on Twitter Sites Using LSTM

Shantanu Shinde, Sahil Tagunde, Ganesh bankar, Prof.Prakash D Kshirsagar

Department of Computer Engineering, Zeal College of Engineering and Research, Narhe, Pune, India

ABSTRACT: Microblogging platforms like Twitter serves as an ideal place for fabricating and diffusing rumours. These rumours disseminate at a rapid pace and lead to grievous social issues. Rumours can cause social panic and adversely affect people and organizations. Therefore, detecting rumours accurately, quickly and automatically is important. Existing models detect rumours on social media sites by employing feature extraction which is time-consuming, biased and labour-intensive. Even though, recent studies use machine learning-based methods for automatic rumour detection by extracting features of rumour contents (e.g., people's opinions, questions, etc.) and static spreading processes, early detection of rumour remains a challenge. The proposed system employs a deep learning model, Long Short-Term Memory (LSTM) in collaboration with pooling function of Convolutional Neural Network (CNN) to quickly detect rumour. LSTM networks makes predictions based on time series data. It also overcomes problems faced while training traditional Recurrent Neural Network (RNN), like, exploding and vanishing gradient. The dynamic changes of forwarding contents are taken into consideration using LSTM based model.

KEYWORDS: Rumour identification, Twitter, Long Short-Term Memory, Microblogging sites, Early detection.

I. INTRODUCTION

Social media has become one of the major platforms which facilitate people and organizations to create or share information, career interests, ideas, news and much more. Social media is blooming, particularly microblogging sites like Twitter, Tumblr, and SinaWeibo are admired widely because these sites provides fast propagation and acquisition of information. Nowadays, social media is easily available which helps us to connect people across the globe. Twitter has become one of the popular social networking site on which users interact and post messages called "tweets". These tweets comprises of user's opinion, photos, videos, article links and quotes. Twitter gained popularity as a microblogging service because of its 280-character messages called tweets. These short message service provided by Twitter became extensively popular amongst people. These microblogging sites generate a large quantity of multimedia content which is crucial in many important applications. In spite of Twitter being a great platform for sharing and gaining information, it is vulnerable to quick dissemination of rumours which can cause damage to the society.

A rumour is a statement which is circulated without confirming the facts [1]. Rumours arise in the context of ambiguity or when the situation is not clear to people [2]. Rumours hence are harmful force that affects people and organizations [3]. For example, a rumoured tweet saying two explosions in White House and Barack Obama is injured was released from the official Twitter account of the Associated Press (AP) which was hacked on April 23, 2013.

II. LITERATURE SURVEY

Ma et al. [4] was first to apply RNN for finding rumours. It was noticed that an event comprises of the original message or tweet and some other messages like comments related to the event and reposts of the messages and this created a continuous stream of messages. Therefore, to extract the features efficiently from these messages, they were batched into time intervals with variable length, and contemplate these as a unit. The proposed model evaluates

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 5, May 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0905176 |

the RNN-based model using three recurrent units namely GRU, LSTM and tanh. GRU as well as LSTM shows potential to remarkably capture the long-term dependencies of messages and the performance of the detection model is high in the initial stage.

Chen et al. [5] proposed a model which uses deep-attention for early identification of rumours. The model uses RNN for identifying rumours. The model CallAtRumours(Call Attention to Rumours) detects rumours by learning temporary hidden characteristics and representations from the messages. Initially the model batches the continuous streams of messages into variable length units of time series and then soft-attention mechanism is incorporated into recurrence to draw out distinct features and shun duplicity.

Weiling Chen, Chai Kiat Yeo, Chiew Tong Lau, Yan Zhang , Bu Sung Lee [6] proposed a detection model which uses combination of RNN and Autoencoder. The user's behaviour is seen to vary while commenting on a rumour message and on a genuine message, this is reported with the use of comment-based features. RNN helps in analyzing the features which change over time. The time-dependent and time-independent properties are merged and passed to an autoencoder for identifying rumour.

Jing Ma, Wei Gao, Kam-Fai Wong [7] proposed a model namely RvNN, which is a type of tree structured neural network that links content's semantics and dissemination hints. Various classes of rumours are classified using neural networks to learn the distinct features of tweets.

OluwaseunAjao, DeepayanBhowmik, ShahrzadZargari [8] provides a structure that identifies rumoured messages and further classifies them by employing combination of LSTM and CNN models on Twitter. CNN provides a pooling function which assists in circumventing over-fitting and lowers the cost of detection by lowering the dimensionality. The model used texts and images to learn in an unsupervised manner and classifies the tweet accordingly.

N. Ruchansky et al. [9] proposed a model that combined the properties like the message, the response and the source of message for accurate results and predictions. The behavior of the spreaders was monitored to detect rumours efficiently. Using the motivation of the properties mentioned above, a CSI model was proposed which consisted modules namely, capture, score and integrate.

III. PROPOSED SYSTEM

A rumour is a doubtful and unverified statement, which on investigation, can come out as a legitimate fact or an actual misinformation (rumour). The system aims in debunking actual misinformation. The misinformation can be deliberately fabricated to mislead public or it can emerge out of a misunderstanding. Whatever maybe the cause of the diffusion of misinformation, it ultimately harms society emotionally and financially. To avoid confusion regarding the rumour coming out as a legitimate fact or an actual misinformation, the later sections follows the convention of referring a legitimate fact as "non-rumour" and an actual misinformation as "rumour".

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 5, May 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0905176 |

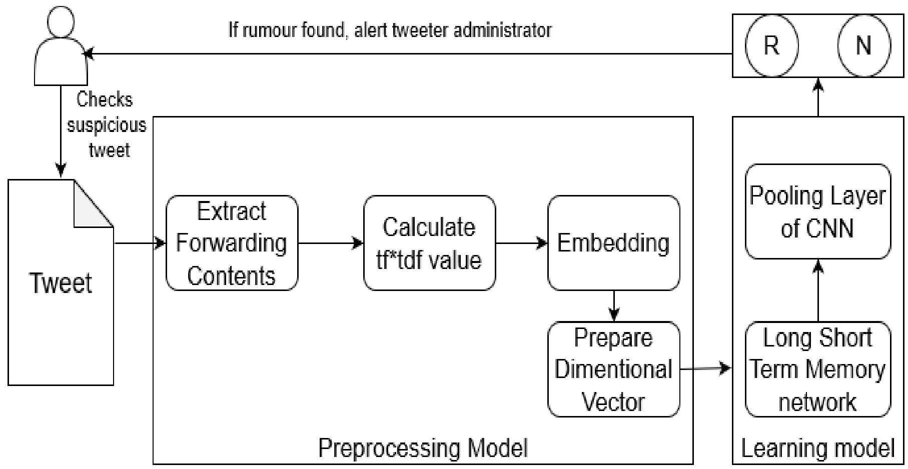

Figure 1: Block Diagram of Rumour Detection System

Figure 1 displays the block diagram for proposed rumour detection model. If the administrator comes across a dubious tweet, then this tweet can be fed to the model for inspection of its credibility. Using the information gained by extracting forwarding contents of the tweet, the results are produced. The model is divided into two phases namely, the processing phase and the learning phase.

Preprocessing phase: It is observed that the tweet's content gets altered over the course of time. In this phase, extraction of the key features of the forwarding content is the major goal. To do the needful, first, the common and unimportant words in the tweet are given less weightage by calculating the *tf*idf* values. These values help in determining the stop words which are then eliminated. Next is the embedding process which is prominently used for the purpose of transforming a word to its vector equivalent. The neural network like LSTM requires a numeric data for processing. As the LSTM cannot process words directly, embedding plays a major role in mapping the words to their respective vector values. To perform this mapping, a shallow two-layered neural network called Word2Vec model is used to create word embedding for representing words as numeric values. A dimension D_c is created from the vector of vocabulary and the *tf*idf* values and fed to the LSTM in the next phase.

Learning phase: this phase comprises of the LSTM and the pooling function of CNN. LSTM aids in capturing key features that will help distinguish rumours and non-rumours. LSTM overcomes the drawback of traditional neural networks as it captures the long-term dependencies of the tweets. LSTM comprises of gates and memory cell state s_t which controls the flow of information. The output h_t of the LSTM unit can be computed using the below equations [6].

$$i_t = \sigma(U_i h_{t-1} + W_i x_t + V_i s_{t-1} + b_i) \quad (1)$$
$$f_t = \sigma(U_f h_{t-1} + W_f x_t + s_{t-1} V_f + b_f) \quad (2)$$
$$\bar{s}_t = tanh(U_s h_{t-1} + W_s x_t + b_c) \quad (3)$$
$$s_t = f_t x_{t-1} + i_t \bar{s}_t \quad (4)$$
$$o_t = \sigma(U_o h_{t-1} + W_o x_t + V_o s_t + b_o) \quad (5)$$
$$h_t = o_t tanh(s_t) \quad (6)$$

To downsample the input matrix and to efficiently capture the hidden clues, pooling function is used. Pooling function proves beneficial acquiring influential local information. Lastly, the output of model is displayed as "rumour" or "non-rumour".

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 5, May 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0905176 |

IV. RESULTS AND EVALUATION

Proposed model is implemented using python on Anaconda platform. TensorFlow and Keras are math libraries used for numeric calculation. Google Word2Vec is a tool is used for word embedding. Tweepy library is used to access Twitter API. Using the tweet text and the time of post of tweet, results are obtained.

Table 1 shows the summary of input provided by the model which shows total number of rumour and non-rumour events taken into consideration, it also shows total number of tweets posted and maximum and minimum number of tweets posted for an event.

Rumour	484
Non – Rumour	193
Total Post	349897
Maximum Post	15196
Minimum Post	6

Table 1: Summary of input

Table 2 shows the test accuracy for 677 events with 484 rumour events 193 non-rumour events. Along with accuracy; precision, recall and F1 score are some of the important model metrics used to calculate the goodness of the model. These metrices are calculated using True Positive (TP), False Positive (FP), True Negative (TN), False Negative (FN).

Results

Accuracy	0.85
Precision	0.88
Recall	0.93
F1 Score	0.91

Table 2: Results of Model

Figure 2 shows the graph of training accuracy versus epochs. This can be used to visualize the increase in accuracy of model during training in proportion to increase in epoch.

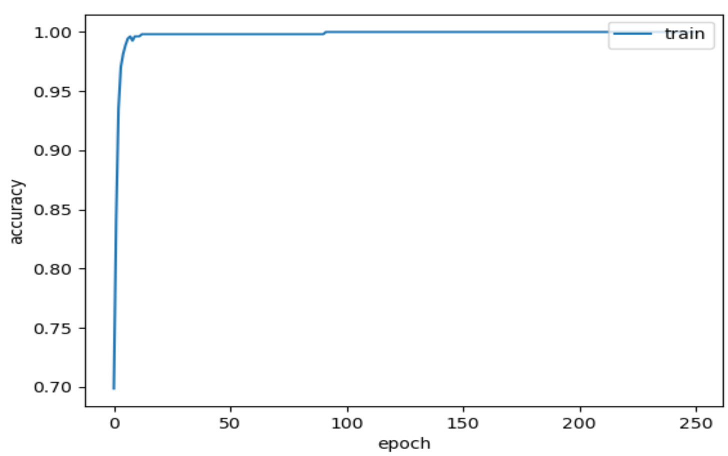

Figure 2: Graph of Train Accuracy vs. Epoch

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 5, May 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0905176 |

This can be used to visualize the training loss of model gradually decreases as number of epochs are increased.

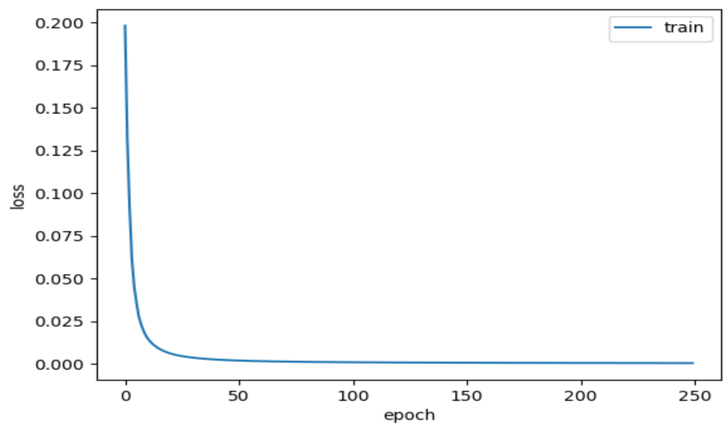

Figure 3: Graph of Train Loss vs. Epoch

Figure 3 gives comparison of different methods used for rumour detection on twitter like DT-Rank, attRNN, DTC, RFC, CallAtRumour, LSTM and HAS-BLSTM.

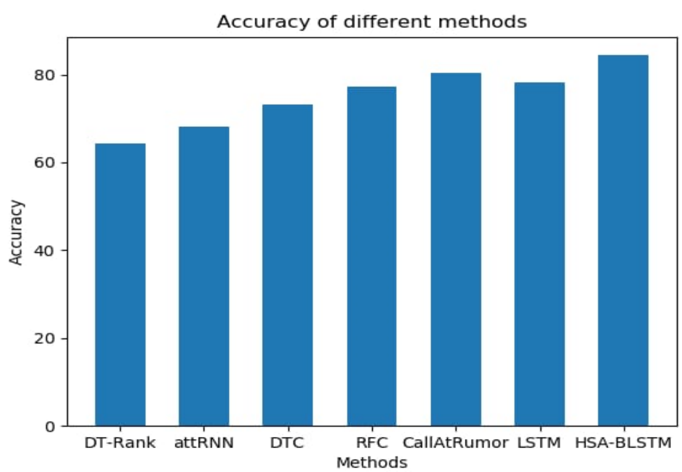

Figure 4: Comparison of Accuracy of Different Methods

V. GRAPHICS USER INTERFACE OF THE SYSTEM

Rumoured tweet – Spraying or introducing bleach or another disinfectant into your body will protect you against #corona

The administrator enters the keyword related to event and number of tweets to be displayed.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 5, May 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0905176 |

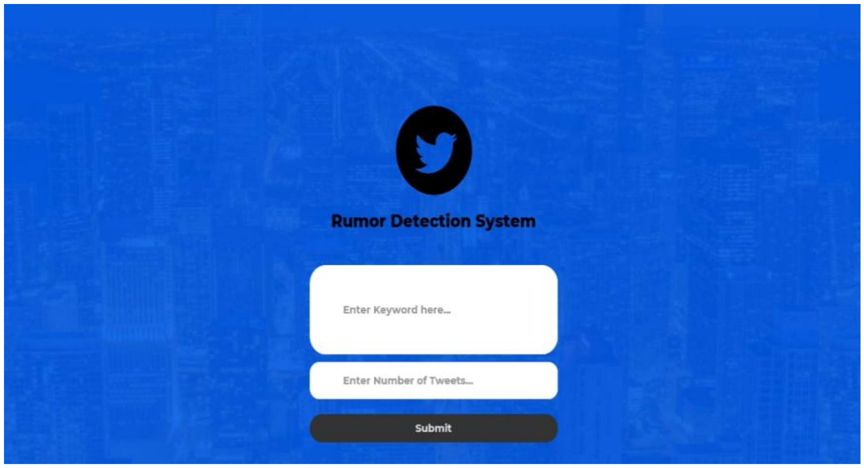

Figure 5: Interface to check rumour

After keyword is entered by user the corresponding tweets are retrieved and displayed on screen.

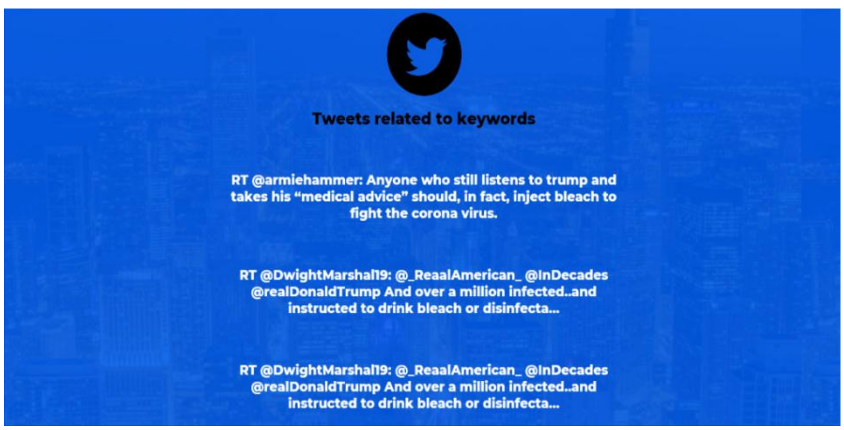

Figure 6: Tweets retrieved related to keywords

A graph of number of tweets posted per day is displayed to get an idea of propogation of rumour.

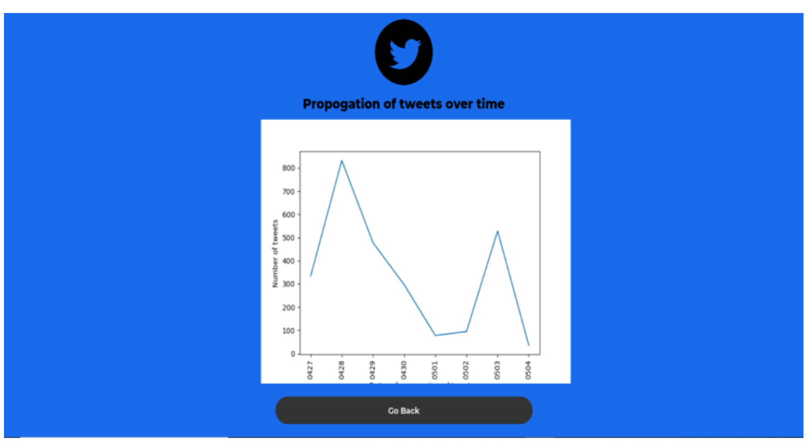

Figure 7: Interface to view graph

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 5, May 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0905176 |

Rumour detection

The model alerts the administrator whether the given event is rumour or not so that the administrator can take action when event is rumour.

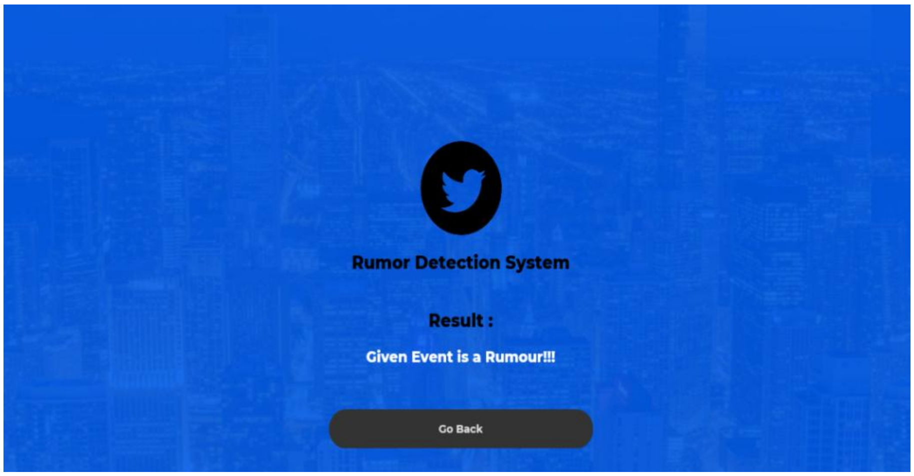

Figure 8: Result obtained for the tweet

VI. CONCLUSION

Currently, some models use ML based techniques to detect rumours on social media sites. These models employ feature extraction which is time-consuming and biased. This system assists user to successfully detect rumoured tweets using deep learning based method. LSTM networks and pooling function of CNN are used in this model to accurately debunk the suspicious tweet as rumour or non-rumour. The system considers the forwarding contents of the user for detection purpose. Using the max pooling function, cost and dimensionality of the model is reduced and performance is improved. The accuracy obtained using this model is 85.29% for total of 677 events. The efficiency of the model can be further improved by adding numerous hidden layers to the LSTM.

REFERENCES

[1] G. Allport and L. Postman, “The psychology of rumour”, Rinehart & Winston, 1947.
[2] N. DiFonzo, R. Rosnow and P. Bordia, “Reining in rumours”, Organizational Dynamics, 1994, 23(1):47–62p.
[3] N. DiFonzo and P. Bordia, “Rumour psychology: Social and organizational approaches”, American Psychological Association, 2007.
[4] J. Ma, B. J. Jansen, W. Gao, M. Cha, P. Mitra, K. Wong and S. Kwon, “ Detecting rumors from microblogs with recurrent neural networks”, in Proceedings of IJCAI, 2016.
[5] T. Chen, L. Wu, X. Li, Jun Zhang, Hongzhi Yin and Yang Wang, “Call attention to rumours: Deep attention based recurrent neural networks for early rumour detection”, 2018, 40-52p.
[6] W. Chen, Y. Zhang, C. Yeo, C. Lau, B. Sung Lee. “Unsupervised rumour detection based on user’s behaviours using neural networks”, Journal paper vol.105, April 2018, 226-233p.
[7] J. Ma, K.Wong and W. Gao “Rumor Detection on Twitter with Tree-structured Recursive Neural Network’’ in Proc. 56th Annu. Meeting Assoc. Comput. Linguistics, vol. 1, 2018.
[8] O. Ajao, D. Bhowmik, and S. Zargari, “Fake news identification on Twitter with hybrid CNN and RNN models”, in Proc. 9th Int. Conf. Social MediaSoc., 2018, 226-230p.
[9] N. Ruchansky, S. Seo, and Y. Liu, “CSI: A hybrid deep model for fake news detection” in Proceedings of the 2017 ACM on Conference on Information and Knowledge Management.
[10] T. Lan, C. Li, J. Li “Mining Semantic Variation in Time Series for Rumor Detection via Recurrent Neural Networks”, International Conference.