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

http://dx.doi.org/10.12785/ijcds/160114

# Graph-Based Rumor Detection on Social Media Using Posts and Reactions

**Nareshkumar R[1], Nimala K[2], Sujatha R[3], Shakila Banu S[4], Sasikumar P[5] and Balamurugan P[6]**

[1,2,3]*Department of Networking and Communications, School of Computing, College of Engineering and Technology, SRM Institute of Science and Technology, Kattankulathur, Chennai, Tamilnadu, India*
[4]*Artificial Intelligence and Data Science Department, CARE College of Engineering, Trichy, Tamilnadu, India*
[5]*Department of Artificial Intelligence and Machine Learning, Sphoorthy Engineering College, Hyderabad, Telangana, India*
[6]*Department of Computer Science and Engineering, MLR Institute of Technology, Dundigal, Hyderabad, Telangana, India*
*Received 27 Sep. 2023, Revised 13 Apr. 2024, Accepted 16 Apr. 2024, Published 1 Jul. 2024*

**Abstract:** In this article, researchers deliver a novel method that makes use of graph-based contextual and semantic learning to detect rumors. Social media platforms are interconnected, so when an event occurs, similar news or user reactions with common interests are disseminated throughout the network. The presented research introduces an innovative graph-based method for identifying rumors on social media by analyzing both posts and reactions. Identifying and dealing with online rumors is an important and increasing difficulty. We use real-world social media data to create a solution based on data analysis. The process involves creating graphs, identifying bridge words, and selecting features. The proposed method shows better performance than the baselines, indicating its effectiveness in addressing this significant issue. The method that is being offered makes use of tweets and people's replies to them in order to comprehend the fundamental interaction patterns and make use of the textual and hidden information. The primary emphasis of this effort is developing a reliable graph-based analyzer that can identify rumors spread on social media. The modeling of textual data as a words co-occurrence graph results in the production of two prominent groups of significant words and bridge connection words. Using these words as building pieces, contextual patterns for rumor detection may be constructed and detected using node-level statistical measurements. The identification of unpleasant feelings and inquisitive components in the responses further enriches the contextual patterns. The recommended technique is assessed by means of the PHEME dataset, which is open to the public, and contrasted with a variety of baselines as well as our suggested approaches. The results of the experiments are encouraging, and the strategy that was suggested seems to be helpful for rumor identification on social media platforms online.

**Keywords:** Neural network, Rumors, Social media, NLP

## 1. INTRODUCTION

In the recent years, a variety of social media platforms, including Twitter, Facebook, Instagram, and others, have gained a lot of popularity due to the fact that they make the simple acquisition of information extremely easy and give a speedy platform for the exchange of information. Researchers have focused a significant amount of emphasis on the availability of unauthentic data on social media platforms, which have evolved into hubs for the dissemination of false information [1]. Because of the enormous damage that can be caused by fake news, there has been an increase in attention paid to the problem by scholars, journalists, politicians, and members of the general public. A difficult but crucial activity, rumor detection on social media seeks to discover and categories rumors or incorrect information that are being propagated across different social media platforms. Now that more people than ever are turning to social media to get their news and information, it's more important than ever to be able to spot false claims before they spread and cause widespread fear [2]. Research into rumor detection on social media focuses primarily on a variety of strategies and approaches that make use of techniques such as machine learning, natural language processing (NLP) [3], network analysis, and data mining. This is because these are the four primary areas in which rumors may be found. In the context of writing, "fake news" refers to articles that are created or published with the intention of misleading the general public and tarnishing the reputation of an organization, institution, or individual for the purpose of gaining financial or political gain.

The spread of false information is risky and may hinder our attempts to solve global concerns since it is the source of many of the problems and contributes to the distortion of

*E-mail Address:nr7061@srmist.edu.in, nimalak@srmist.edu.in, sujathar2@srmist.edu.in, fshakilabanu@care.ac.in, sasi@sphoorthyengg.ac.in and balamurugan.p@mlrit.ac.in*

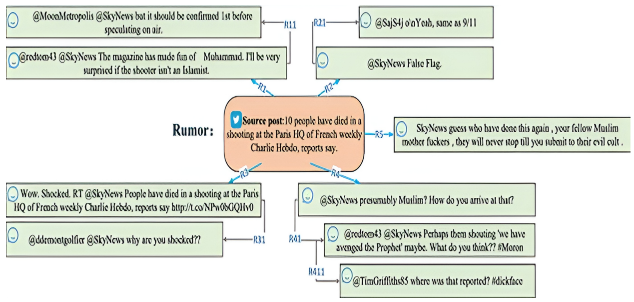


Figure 1. An example of rumor event on online website with a source post and related comments

others. The spread of false information operates similarly to that of a virus in that it preys on our frailties, our preferences, our prejudices, and our emotions. Misinformation has been dubbed a "infodemic" by the World Health Organization because it is growing at a higher rate than COVID-19, so undermining efforts to improve public health and warping good scientific guidance.

The propagation of rumors on social media platforms is rapid, they have a broad range of potential influencers, they have a high cost, and they are difficult to disprove manually.Rumor detection has been a significant focus of research with the rise in popularity of social networks. In a preliminary investigation, a group of regular expressions was established to identify signal posts that inquire about the truthfulness of a post. This approach's limitation is in its reliance on predefined regular expressions, which are unable to cover the wide range of writing styles found in postings. Early studies also concentrated on developing a supervised classifier by manually extracting characteristics from post metadata [4]. Those methods have a drawback because to their heavy dependence on feature engineering, which might be biased and require a lot of effort.

As a result, the detection of rumors has evolved into an exceptionally difficult research problem within the field of text classification, and as a result, it has received a lot of focus from both the academic community and the business world. In this piece of research, we describe a graph-based contextual learning representation for the purpose of identifying rumors on Twitter.

Utilizing a graph-based methodology, we were able to extract the contextual information included within the tweets. The graph-based representation [5] learning that our method employs, in addition to the identification of two frequent groups of words, makes it one of a kind in terms of its ability to extract tumorous patterns. In addition, the emotionally charged and intellectually stimulating phrases are thought to render patterns more general.

Several cataloging representations are trained on the initial PHEME dataset as well as a balanced distribution of the dataset that is freely accessible to the public. The benefit of the pattern-based approach over the word-based technique is that it is more representational and maintains the grammatical meaning, in contrast to the word-based method [6].

## 2. RELATED WORK

In recent years, the identification of rumors and false news has been one of the most researched topics of study. The rumor has been around for a long time, but its effects of misinformation and falsehoods spread via online social networks only become clear when major events take place and people are forced to depend on social media for news and information. For instance, the widespread dissemination of false news on social media had an impact on the presidential election in the United States in 2016. Therefore, automatically recognizing rumors is useful since it allows for the provision of early safeguards to reduce the harmful effect it has. Here is also an example of Figure 1 shows rumor event on social media.

In order to mitigate the harm that might result from spreading incorrect information on social media sites like Twitter, Christof et al. [14] this study proposed a unique probabilistic mixture model called the mixture marked Hawkes model (MMHM). The necessity for feature engineering was eliminated thanks to their framework, which formalized the self-exciting nature of true vs. false retweeting processes. The MMHM accurately inferred the veracity of internet rumors by explicitly modelling the spreading process. The approach was able to identify false rumors with a balanced accuracy of 64.97 % and an area under the curve (AUC) of 69.46 %, exceeding state-of-the-art baselines while being completely interpretable, according to an evaluation of 13,650 retweet cascades on Twitter. The method has real-world applications since it can identify fake information early on social media by using the process of spreading as an implicit quality indicator.

S. Y. Bae et al. [15] analysis of survey data from social media users in South Korea investigates the spread of political rumors through online social networks. Users' trust in political rumors is strongly correlated with their usage of social media as a news source, according to the study's authors. The significance of knowing how social media users digest false information is also highlighted. The study sheds light on the variables that lend credence to rumors and false information by analyzing network features. In sum, the research sheds new light on the dynamics and variables impacting the veracity of political rumors in online social networks.

A. E. Fard et al. [16] research was to investigate the feasibility of utilizing deep learning algorithms for sentiment analysis, with a special emphasis on rumor categorization. While sentiment categorization has been accomplished using machine learning techniques, the accuracy of deep learning methods is far higher. This paper highlights the usefulness of deep learning for sentiment analysis, a task that is not often associated with it. The deployed deep

TABLE I. Previous Study

| Study | Approach/Model Used | Contributions | Limitations# |
|---|---|---|---|
| Ma et al., 2018 [7] | Recursive Neural Networks | Utilized recursive neural networks for text feature extraction | Limited use of social network structure information |
| Khattar et al., 2019 [8] | Variational Autoencoder | Employed variational autoencoder for text and image feature extraction | May not fully utilize social network structure information |
| Liu and Wu, 2018 [9] | RNN and CNN | Modeled propagation paths of tweets as multivariate time series | Limited use of social network structure information |
| Yuan et al., 2019 [10] | Graph Attention Network (GAT) | Leveraged GAT for text and network structure feature extraction | Does not consider camouflage behaviors in the social network |
| Wang et al., 2018 [11] | Generative Adversarial Network (GAN) | Used GAN for obtaining non-event-specific features from texts and images | Limited focus on social network structure information |
| Ma et al., 2019 [12] | Generative Adversarial Network (GAN) | Employed GAN to generate uncertain or conflicting voices for learning features | Limited focus on social network structure information |
| Haq et al., [13] | Graph-based Perturbation | Proposes perturbing the network structure to learn stronger structure features. Involves graph adversarial attacks with consideration for heterogeneity and domain constraints in rumor detection | Focused on leveraging social network structure information |

learning model, built on top of a Recurrent Neural Network with Long Short-Term Memory, outperformed other state-of-the-art algorithms with an accuracy of 87.53%. Grid search CV and a pipeline are used to preprocess the data in order to optimize the classification model's parameters. The research demonstrates the power of deep learning for analyzing emotions and classifying rumors. The results of earlier research on the identification of rumors on social media are shown in Table I.

Thota et al. [17] detects rumors on social media, where misinformation spreads quickly. rumor categorization and prediction mostly use machine learning with hand-crafted textual characteristics. Dynamic data features are needed because to the large volume of data and inaccuracy of hand-crafted features. Graph Convolutional Networks (GCN) are used to graph rumor propagation trees and update node representations depending on rumor replies. Pattern matching algorithms quickly recreate comparable sub-graph patterns to discover rumors early. The model predicts rumors early and uses social network dynamics to outperform other methods.

Chen et al. [18] to investigate the validity of the rumor. The CNN model takes the tree structure that is generated using the source-response connection as its starting point, extracts the feature vector from the textual data that is located at each node and then provides this vector to neural networks. Recurrent Neural Networks, often known as RNN, are an additional type of neural network that is

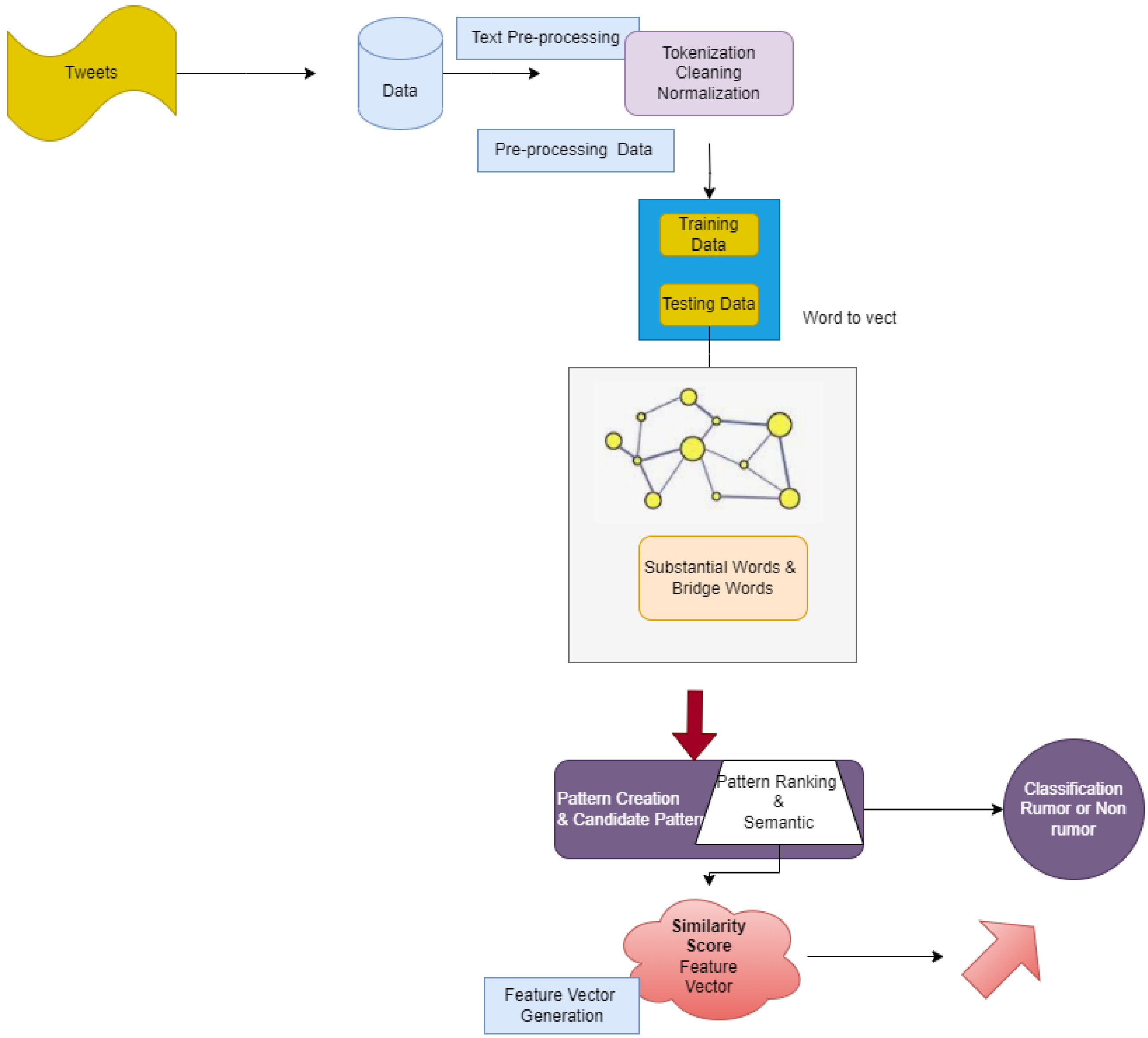


Figure 2. Proposed Architecture

frequently used in this industry.

## 3. METHODOLOGY

We approach the issue of rumor detection as a binary classification one, in line with other efforts that have been done before that are considered to be state-of-the-art. Figure 2 displayed a schematic illustration of the proposed architecture.

Distinguishing characteristics based on its spread patterns and other types of data. To describe the propagation behavior of a post in social networks, we recommend utilizing a graph structure to extract valuable insights about how the message spreads. Rumor identification challenge tries to estimate a function r represent of a tweet as either a rumor or a non-rumor when given a tweet from a Twitter dataset T called $t_i \in T$.

$$r(t_i) = \begin{cases} 1 & \text{If } t_i \text{ is a rumor} \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

### *A. Data Pre-processing*

A series of important steps are required to prepare a Twitter dataset. These steps include data collection, text cleaning to remove noise such as URLs and special characters, tokenization for word separation, lowercase conversion for uniformity, stopword removal, normalization [19] using techniques such as stemming, emoticon handling, duplicate

TABLE II. Hyperparametric

| Hyperparametric | Numerical value |
|---|---|
| $\mathrm{Num}_h eads$ | 6 |
| dropout | 0.2 |
| $\mathrm{Learning}_r ate$ | 5e-5 |
| Epsilon for adam optimizer | Le-8 |
| $\mathrm{Num}_t rain_e pochs$ | 30 |
| $\mathrm{Hidden}_s ize$ | 200 |
| $\mathrm{Batch}_s ize$ | 32 |
| Embedding $_d im$ | 300 |

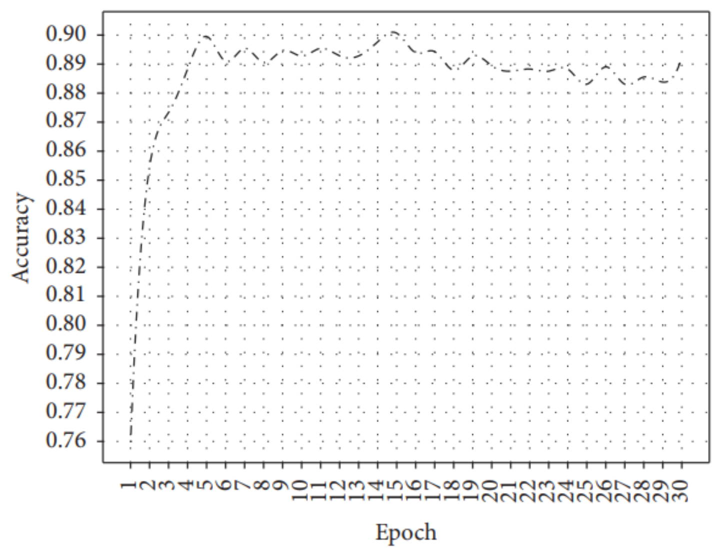


Figure 3. Evaluation accuracy

elimination, abbreviation expansion, sentiment analysis for emotional assessment, feature extraction, entity recognition and data labeling. These procedures, taken as a whole, transform the data into a format that is more organized and refined, making it more amenable to later analysis and applications of machine learning.

### *B. Tweet-Reactions Relationships*

The object of this subsection is to determine the degree to which the source tweets and responses are related. The responses are a description of the hidden information behind them. When asked about the veracity of a tweet, users' responses often include skepticism, correction, and verification. As a means of acquiring extra information and making tweets explicable in relation to rumors, their responses have been put alongside them.

For each particular tweet $t_i$ and the responses it receives

$$R(t_i) = \left\{r_1^{t_1}, r_2^{t_2}, \ldots, r_n^{t_i}\right\} \tag{2}$$

By computing the proportional difference between a tweet of its responses, which is referred to as equation 2, the repeats may be deleted.

$$SD_i = t_i \Delta \left\{r_n^{t_i}\right\}_1^n \tag{3}$$

The following are the regular differences that may be formalized between a tweet and reactions:

$$t_i \left\{r_n^{t_i}\right\}_1^n \tag{4}$$

In order to ensure that a tweet and its responses are as inclusive as possible, a sole input tweet is generated by merging a tweet and there optimized responses via the creation of a union of all of the symmetric differences between the two SDi

$$X = (t_i R_{ti}) \tag{5}$$

$$X = t_i + \cup_1^n SD_i^{t_i} \tag{6}$$

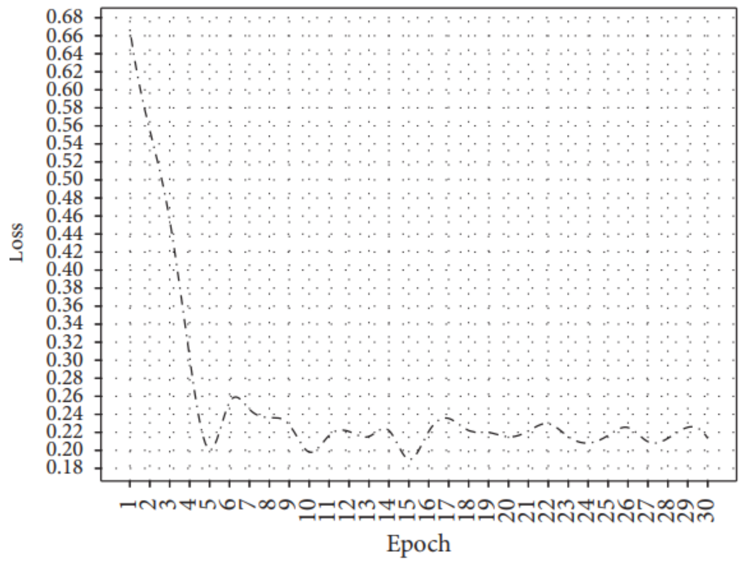


Figure 4. Evaluation Loss

TABLE III. Performance Results of Our Proposed Approach - Rumors

| Methods | Accuracy | Precision |
|---|---|---|
| Baseline | 0.547 | 0.580 |
| CNN | 0.652 | 0.654 |
| LSTM | 0.795 | 0.705 |
| RNN | 0.784 | 0.711 |
| GRU | 0.735 | 0.791 |
| BERT | 0.751 | 0.791 |
| BI GCN | 0.879 | 0.936 |
| Ours proposed approach | 0.884 | 0.936 |

### *C. Graph Generation and Bridge Words*

The graph-based technique efficiently captures linguistic variance [20] and contextual information in textual data with the least amount of subject expertise. Scene graphs are effective visual representations that break down pictures into its abstract semantic components, such as objects and their relationships, making it easier to understand visual information and make logical deductions. On the other hand, commonsense knowledge graphs are rich reservoirs that store information about how the world is organized and how various ideas relate to one another. By calculating node-level statistical metrics, the graph is utilized significant bridge connection words. Collecting words really important for creating rumors patterns is the main goal. In Table II, you will get a summary of the hyperparametric properties of all of the models.

## 4. RESULTS AND DISCUSSION

Throughout the working out of the approach, we decided to set up 30 rounds of training, and As the number of training cycles increased, we discovered that the predictions made by the model became more reliable.

Timely identification of rumors can greatly reduce the harm caused by their dissemination. Many researchers have used differential equations and numerical simulation approaches to analyze the speed of rumor spread and solutions for controlling it. Table III and IV contains an overview of all of the models' performances, from which it is possible to draw the conclusion that the accuracy and F1 value of the algorithm ar low. This is due to the fact that the traditional ML manual feature extraction is laborious, which is why, when it comes to rumor detection.

The accuracy value achieved by the system during the first four rounds was quite low; the rumor detection impact of this model was bad; but, during the fifth round, the accuracy value of the model grew significantly. In assumption, the number of disks training had a significant influence on the outcomes of the experiment.It reached a point of stability after a specific number of rounds had been completed. The model that was given in this study had excellent detection effects when it reached a stable state.

Figure 3 displayed a schematic illustration of the improvements in evaluation index accuracy. Figure 4 displayed

TABLE IV. Performance Results of Our Proposed Approach - Non Rumors

| Methods | Accuracy | Precision |
|---|---|---|
| Baseline | 0.577 | 0.548 |
| CNN | 0.663 | 0.652 |
| LSTM | 0.723 | 0.696 |
| RNN | 0.700 | 0.678 |
| GRU | 0.781 | 0.732 |
| BERT | 0.781 | 0.732 |
| BI GCN | 0.852 | 0.846 |
| Ours proposed approach | 0.855 | 0.851 |

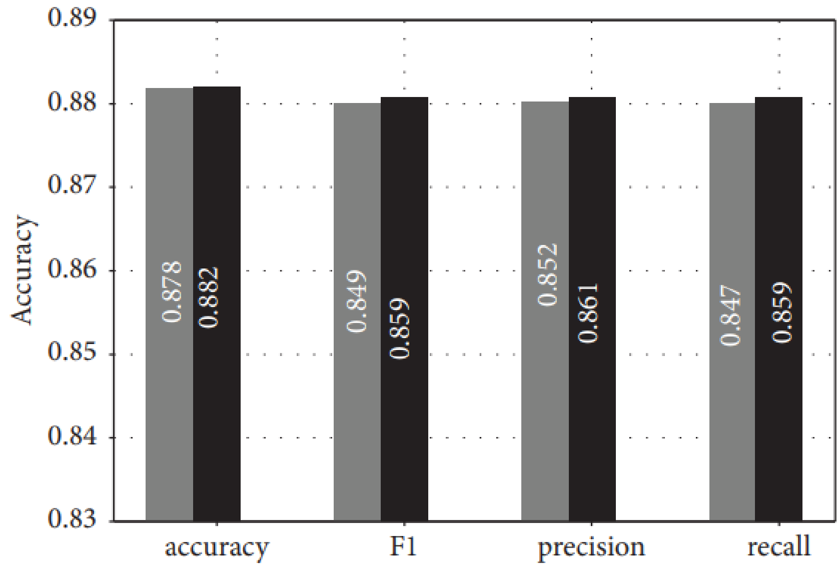


Figure 5. Proposed Model Results

a schematic depiction of the evaluation loss index's change in value. We perform our experimentation using a PHEME dataset. The dataset is comprised of a collection of tweets, together with the tweeters' replies (direct or nested), as well as information about contravention news events.

For instance, it could involve the position taken by a reply post in relation to its original post. If a post contains erroneous information, direct responses are expected to contradict it, whereas posts with accurate information will generally receive supportive responses. We create a propagation graph to study the relationship between the original post and its responses. When it comes to the process of generating and mining rumor features, the basic recurrent neural network model does not take into consideration the characteristics associated with significant spatial levels among objects. When compared to the approach of modeling baseline time series, the outcome of the method of modeling rumors based on graph and tree layout is superior. Figures 5 and 6 provide a graphical representation of the findings of the comparison analysis performed on the proposed.

The comparative findings are summarized in Table III and IV performance completed with the dataset that was originally used. It is clear that the suggested method beats any and all other methods for each of the four classification algorithms.

According to the F1 score, which is a machine learning assessment metric, the accuracy of a model is evaluated. Specifically, it combines the accuracy scores of a model with its recall ratings. This statistic determines the frequency with which a model generated an accurate prediction throughout the entirety of the dataset. When the dataset is class-balanced, or when each class in the dataset has an equal number of samples, then and only then can this metric be believed.

The reason for the constant performance of the suggested method is because the graph-based approach covers the relevant terms, whilst the word embedding technique covers the higher dimensional semantic space. This allows for the proposed approach to consistently perform well.

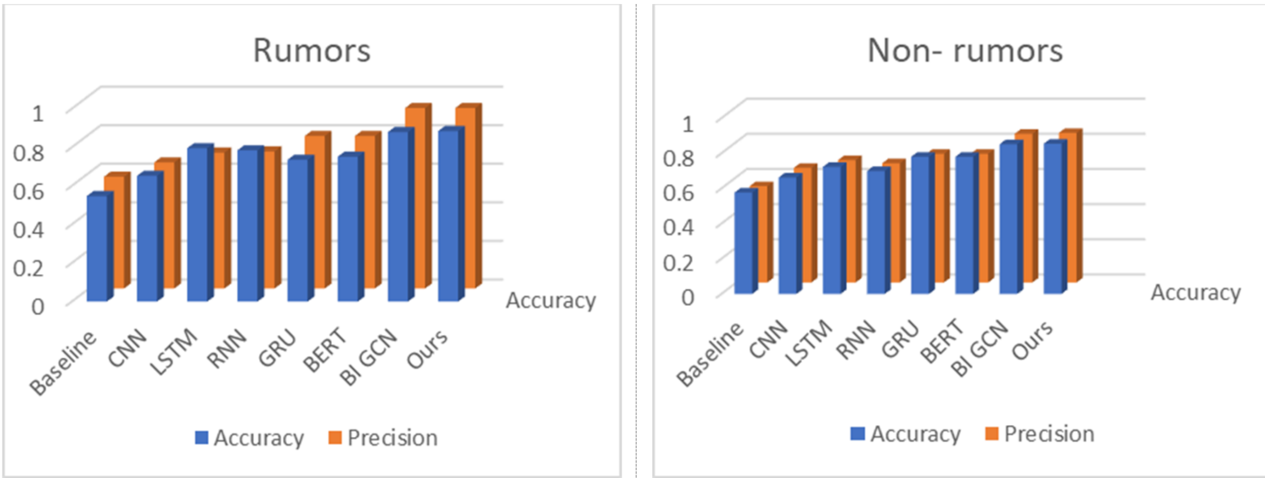


Figure 6. Comparison of Evaluation Accuracy

## 5. CONCLUSIONS

The current research introduced an innovative graph-based method for identifying rumors on social media by analyzing both posts and reactions. We tackled the crucial task of recognizing online rumors by utilizing actual social media data. The data-driven strategy outperformed the baselines, showcasing its effectiveness and potential for widespread adoption. In this investigation, we describe a framework for the identification of rumors that makes use of a graph-based method to make use of tweets and responses in order to identify patterns that are indicative of rumors. For the purpose of capturing contextual information. It is possible to determine the inquisitiveness, skepticism, sentimentality, and emotionalism of users of social media platforms by analyzing the writing styles shown in their responses to a tweet. The rumor detection model is trained and taken the emotional, skeptical, and rumor respectively. The results of the experiments showed that the suggested technique is superior to the existing approaches in terms of its performance on the identification of rumor detection.

## 6. LIMITATION AND FUTURE WORK

The recent study presented a novel graph-based approach to detect rumors on social media by examining both posts and reactions. We will recognize the subjectivity and biases that come with choosing features and explore automated feature learning solutions in future research. We will address the scalability issues of the graph-based approach when dealing with extensive datasets and investigate optimization techniques to enhance scalability. We will recognize the constraints of utilizing only one social media platform and emphasize the necessity for additional assessment on various platforms and datasets to determine the generalizability of the outcomes. We have enhanced the future work portion by detailing specific research directions to tackle the highlighted constraints and investigate further enhancements. Exploring automated feature learning approaches such as deep learning and natural language processing to extract features automatically, aiming to decrease bias and enhance performance. Exploring methods to enhance scalability: We will explore ways to optimize graph algorithms and utilize distributed processing approaches to efficiently manage larger datasets. We will conduct tests on several social media platforms and datasets to evaluate the generalizability of the approach and determine any necessary platform-specific modifications.

## References


[1] S. S. Samal, K. Pathak, K. Nimala, and R. Nareshkumar, "Sentiment analysis for depression using human expression," in *2023 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES)*, 2023, pp. 1–7.

[2] R. Nareshkumar and K. Nimala, "Interactive deep neural network for aspect-level sentiment analysis," in *2023 International Conference on Artificial Intelligence and Knowledge Discovery in Concurrent Engineering (ICECONF)*, 2023, pp. 1–8.

[3] L. J. Sailesh, V. K. Kumar, K. Nimala, and R. Nareshkumar, "Emotion detection in instagram social media platform," in *2023 International Conference on Artificial Intelligence and Knowledge Discovery in Concurrent Engineering (ICECONF)*, 2023, pp. 1–6.

[4] R. Nareshkumar and K. Nimala, "An exploration of intelligent deep learning models for fine grained aspect-based opinion mining," in *2022 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES)*, 2022, pp. 1–7.

[5] L. D. Xu, Y. Lu, and L. Li, "Embedding blockchain technology into iot for security: A survey," *IEEE Internet of Things Journal*, vol. 8, no. 13, p. 10452–10473, Jul. 2021. [Online]. Available: http://dx.doi.org/10.1109/jiot.2021.3060508

[6] R. Nareshkumar, G. Suseela, K. Nimala, and G. Niranjana, *Feasibility and Necessity of Affective Computing in Emotion Sensing of Drivers for Improved Road Safety*. IGI Global, Sep. 2022, p. 94–115. [Online]. Available: http://dx.doi.org/10.4018/978-1-6684-3843-5.ch007

[7] J. Ma, W. Gao, and K.-F. Wong, "Rumor detection on

twitter with tree-structured recursive neural networks," in *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. Association for Computational Linguistics, 2018.

[8] D. Khattar, J. S. Goud, M. Gupta, and V. Varma, "Mvae: Multimodal variational autoencoder for fake news detection," in *The World Wide Web Conference*, ser. WWW '19. ACM, May 2019.

[9] Y. Liu and Y.-F. Wu, "Early detection of fake news on social media through propagation path classification with recurrent and convolutional networks," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 32, no. 1, Apr. 2018. [Online]. Available: http://dx.doi.org/10.1609/aaai.v32i1.11268

[10] C. Yuan, Q. Ma, W. Zhou, J. Han, and S. Hu, "Jointly embedding the local and global relations of heterogeneous graph for rumor detection," in *2019 IEEE International Conference on Data Mining (ICDM)*. IEEE, Nov. 2019.

[11] ser. KDD '18. ACM, Jul. 2018.

[12] J. Ma, W. Gao, and K.-F. Wong, "Detect rumors on twitter by promoting information campaigns with generative adversarial learning," in *The World Wide Web Conference*, ser. WWW '19. ACM, May 2019.

[13] A. Haque and M. Abulaish, "An emotion-enriched and psycholinguistics features-based approach for rumor detection on online social media," in *Proceedings of the 11th International Workshop on Natural Language Processing for Social Media*. Association for Computational Linguistics, 2023.

[14] C. Naumzik and S. Feuerriegel, "Detecting false rumors from retweet dynamics on social media," in *Proceedings of the ACM Web Conference 2022*, ser. WWW '22. ACM, Apr. 2022.

[15] S. Y. Bae, "The social mediation of political rumors: Examining the dynamics in social media and belief in political rumors," *Journalism*, vol. 21, no. 10, p. 1522–1538, Sep. 2017. [Online]. Available: http://dx.doi.org/10.1177/1464884917722657

[16] A. E. Fard, M. Mohammadi, and B. v. de Walle, *Detecting Rumours in Disasters: An Imbalanced Learning Approach*. Springer International Publishing, 2020, p. 639–652. [Online]. Available: http://dx.doi.org/10.1007/978-3-030-50423-6_48

[17] N. R. Thota, X. Sun, and J. Dai, "Early rumor detection in social media based on graph convolutional networks," in *2023 International Conference on Computing, Networking and Communications (ICNC)*. IEEE, Feb. 2023.

[18] 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," in *Proceedings of the 11th International Workshop on Semantic Evaluation (SemEval-2017)*. Association for Computational Linguistics, 2017.

[19] R. Nareshkumar and K. Nimala, *An Enhanced BERT Model for Depression Detection on Social Media Posts*. Springer Nature Singapore, 2024, p. 53–64. [Online]. Available: http://dx.doi.org/10.1007/978-981-99-8479-4_5

[20] R. Sujatha and K. Nimala, "Harnessing attention-based graph recurrent neural networks for enhanced conversational flow prediction via conversational graph construction," *Journal of Information amp; Knowledge Management*, Mar. 2024. [Online]. Available: http://dx.doi.org/10.1142/s0219649224500382

**Nareshkumar R** received the M.E. and B.E. Computer Science and Engineering from Anna University, Chennai, India . He is pursuing Ph.D degree in Computer Science and Engineering, SRM Institute of Science and Technology, Kattankulathur, Chennai. His research includes machine learning, NLP, Text Mining, Deep Learning Analytics. He can be contacted at nr7061@srmist.edu.in.

**Nimala K** is an Associate Professor in Department of Networking and Communications in School of computing, SRM Institute of Science and Technology, Kattankulathur, Chennai, India. She obtained her Doctorate in Computer Science Engineering from SRM Institute of Science and Technology, Kattankulathur, Chennai, India in 2020. Her research interest includes Machine Learning, Deep Learning, Cloud Computing, Analytics etc. Her main focus of research includes Sentiment Analysis, Topic modelling, Predictive analytics and big data. She has published 30+ papers in many reputed journals and also published many book chapters. She has guided many UG and PG projects. She is a member of many Professional bodies like Institute of Electrical and Electronics (IEEE), Indian Society for Technical Education (ISTE), The Institution of Engineers India (IEI) etc. She can be contacted at nimalak@srmist.edu.in.

**Sujatha R** received the B.E degree in Computer Science and Engineering from the Bharathidasan University, Tiruchirappalli, in 2004 and M.E degree in Computer Science and Engineering from Anna University, Chennai, in 2014 and She is currently working as an Assistant Professor(Jr.G) in the Department of Networking and Communications and pursuing Ph.D degree in Computer Science and Engineering, SRM Institute of Science and Technology, Kattankulathur, Chengalpattu. Her research interests include Natural Language Processing and Data mining. She can be contacted at sujathar2@srmist.edu.in.

**Shakila Banu S** received the M.E and B.E (CSE) from Anna university, Chennai, India in 2015 and 2017 respectively. She is currently a Research Scholar at Department of Information and Communication Engineering, Anna university, Chennai, India. Her research includes Machine learning, Artificial Intelligence and Wireless communication, Data analytics. She can be contacted at fshakilabanu@care.ac.in.

**Sasikumar P** is an Associate Professor in Department of Artificial Intelligence and Machine Learning in Sphoorthy Engineering College , Hyderabad, India. His research interest includes Machine Learning, Deep Learning, Data Analytics etc. His main focus of research includes Sentiment Analysis, Predictive analytics and big data Live Streaming . He has published 15+ papers in many reputed journals ( Scopus, WoS, SCI). He has guided many UG and PG projects in India and Ethiopia . He is a Life time member of many Professional bodies like Indian Society for Technical Education (ISTE), Computer Science Teachers Association, USA. International Association of Engineers (IAENG) , Editorial Board member and reviewer for various Reputed Journals etc. He can be contacted at sasi@sphoorthyengg.ac.in.

**Balamurugan P** is an Associate Professor in department of computer science and engineering MLR Institute of Technology, Dundigal, Hyderabad, Telangana, India. balamurugan.p@mlrit.ac.in, He research interest includes Machine Learning, Deep Learning, Analytics etc. He main focus of research includes Sentiment Analysis, Predictive analytics and big data. he has guided many UG and PG projects. he is a member of many Professional bodies like Indian Society for Technical Education (ISTE), The Institution of Engineers, India.(IEI).etc. He can be contacted at balamurugan.p@mlrit.ac.in.