

Study of Various Technique of Rumor Detection in Twitter: A Comparative Analysis

Mr. Raj Kumar Gupta
Associate Professor, Department of CSE
Madhyanchal Professional University, Bhopal, India

Abstract - In social media, the veracity of information has long been a problem that affects society and companies. Online social media becomes the stage for rumor starters since they look for social networks that are open to infinite dissemination. As a result, this false information does enormous harm to people, businesses, the government, etc. Previous research that examines the linguistic and temporal aspects of rumors suggests that there is plenty of time for rumors to spread. Social media rumors have the potential to cause fear in the public and harm to particular people. Thus, it is essential to implement automated rumors identification and to ban users who propagate rumors on a regular basis. There may also be rumors in this material. Rumors and false information spread fear across the community, exacerbating the negative impacts on companies, government agencies, groups, and individual citizens. Given the negative effects of false news, it is crucial to identify rumors early on, confirm their accuracy, and stop them from spreading. This system's primary goal is to create a social network free from the propagation of false information. The social network based on Twitter was able to anticipate the mechanism of false news dissemination. This system uses Graph Convolutional Network and Recurrent Neural Network for false news identification. Experiments using two real-world datasets demonstrate the superiority of this approach above the state-of-the-art in rumor detection.

Keywords- Discrimination, Inequality, Equality, Gender, Unfair gender treatments.

How to cite this article: Mr. Raj Kumar Gupta, "Study of Various Technique of Rumor Detection in Twitter: A Comparative Analysis" Published in International Journal of Scientific Modern Research and Technology (IJSMT), ISSN: 2582-8150, Volume-12, Issue-3, Number-1, September 2023, pp.1-5, URL: <https://www.ijsmrt.com/wp-content/uploads/2023/10/IJSMT-23120301.pdf>

Copyright © 2023 by author (s) and International Journal of Scientific Modern Research and Technology Journal. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0) (<http://creativecommons.org/licenses/by/4.0/>)



IJSMT-23120301

I. INTRODUCTION

A popular definition of a rumor is a statement that is purposefully untrue or whose veracity cannot be established. These malevolent or fictitious rumors proliferate on social media and will do enormous harm to both people and society. People from all areas of life should now be concerned about the propagation of false news and its increasing power. On occasion, it is

claimed that the spread of false information through social media, notably microblogs like Twitter, has affected election outcomes. Fake news is becoming more and more prevalent worldwide, impacting everything from politics to education and the financial sector on a regular basis. This is becoming a constant source of worry for both residents and policymakers. The majority of rumor detection techniques in use today are derived from statistical machine learning.

The success of neural network models in other tasks has recently inspired several models that use neural networks for rumor detection.

Social media's widespread usage has made it simple for consumers to get information. Sites for microblogging, like Twitter, provide a wealth of news. A Pew Research Center study found that 57% of respondents believe much of the news they encounter on social media is false. People may become unstable if they are exposed to fake news about any political, corporate, or governmental entity, celebrity, or natural occurrence. The necessity for this study arose from the widespread and quick dissemination of information and the lack of methods to guarantee its accuracy. An unconfirmed assertion of facts that is in circulation and useful for some purpose is called a rumor. Rumors are common and, when spread widely by the public, have the power to seriously harm people, institutions, and the government. False reports going viral about people that blame them for their misdeeds might spark hate campaigns and damage their reputation in the long run.

In the first quarter of 2016, there were an average of 310 million users of the microblogging service (Statista - The statistics site), and an average of 6000 tweets are caught globally each second (Internet live stats). Identifying a rumor inside a database this size is difficult. In order to address this issue, researchers have started examining the many aspects of rumors on social media. In order to verify the topic dependability, one of the approaches uses a baseline approach that is framed by looking for relevant material in other significant media (Hashimoto, Kuboyama & Shirota, 2011). Using temporal features is another approach. In this case, a periodic time series model that takes into account both daily and external shock cycles is presented, and the model shows that rumors probably fluctuate over time (Sejeong, Meeyoung, Kyomin, Wei & Yajun, 2013).

An alternative method that incorporates different social background information concerning rumors is time series modeling (Ma, Gao, Wei, Lu & Wong, 2015). Here, a technique for identifying the decentralized phenomena is suggested, which eliminates the main issues as seen in the case that follows. When it comes to sensitive matters that lie behind rumors, people are easily misled. As a result,

maintaining integrity becomes very important. Take the Kolkata riots (Scroll.in, 2015) or stories about Hindus leaving a hamlet. These rumors need a fast detection procedure, which is absent from the previously listed techniques. There are obviously two obstacles in the way: the first is the massive amount of social media data, for which this method includes a filter mechanism; the second is the discovery of a characteristic that doesn't give viewers enough time to respond before being revealed. The dual strategy centers on using a supervised machine learning model to identify rumors. Because specialists explain rumors as statements conveying unsubstantiated facts rather than gauging their truthfulness. Consequently, this serves as the foundation for the first classification model that filters tweets that are factual or educational. Consequently, the audience's first assessments of any material shared on social media are based on these reactions, and these response tweets aid in the early detection of rumors. Finally, taking into account the language feature of rumor response tweets, questioned threads are identified as rumor because they emphasize material that is deemed incomplete or raises doubts; other threads are classified as non-rumor information.

The following are some reasons why rumor detection is relevant:

- Automated rumor detection offers several benefits over manual detection.
 - There isn't much time to share content on social media.
- Therefore, the likelihood that these rumors may affect a person or a community will grow if it takes longer to find them on social media.
- It's critical to identify and disprove rumors as soon as possible in order to reduce the negative impacts of rumors.
 - Online fact-checking resources have a lengthy wait time and limited thematic coverage.

Obstacles Social networking on autopilot There are various challenges in detecting rumors due to the abundance of social multimedia data that is growing. Furthermore, it is very difficult to interpret and

comprehend internet data due to its non-structural, fragmentary, and noisy character.

Propagation with a heterogeneous structure is one of the main issues. When rumors circulate on social media, users have the option to talk about them and leave comments. Characterizing user behaviors on the network of propagation is crucial to recognizing online rumors, since these behaviors contribute to the formation of the underlying social environment for rumors.

II. LITERATURE REVIEW

Zubiaga et al., 2018 - The conventional technique and the artificial intelligence approach are the two major strategies utilized to disprove false information. The conventional approach defines the essential characteristics prior to identification by manually analyzing text using statistics.

Through the analysis of user traits, rumor distribution pathways, and language, Castillo et al. (2011) offered a significant number of features for rumor identification. Some researchers provide distinct feature sets from many angles.

Liu et al., 2015: As artificial intelligence has advanced, several researchers have tried to use deep learning to identify rumors.

Ma et al. (2016) learned the abstract language of rumors using an RNN model. enhanced the accuracy of rumor detection by combining a deep learning model with feature engineering to create a hierarchical social attention model. The hardest aspect of rumor detection is early detection. Many efforts have been made to identify rumors early on.

Wang et al. (2017) examined key elements of rumor dissemination and put forward a probabilistic model. discovered that linguistic and user characteristics might be valuable markers for rumor detection. For early rumor identification, machine learning and deep learning have been used in addition to conventional techniques.

A hybrid SVM classifier based on graph kernels and capable of capturing high-order propagation patterns

was presented by Wu et al. in 2015. created a method centered on looking for query terms that performed well.

A heterogeneous network with a 61% accurate rate for early rumor detection was presented by Zhang et al. in 2018. an RNN network with structures for gentle attention. constructed a neural network framework with the ability to choose discriminative features, develop a rumor classifier, and infer rumor types. Furthermore, Nguyen et al. (2017) described a method that used a time series and convolutional neural networks to learn the hidden representations of every tweet. We have created a multiloss hierarchical BiLSTM model with an attenuation factor that excels in early rumor detection and has good accuracy, drawing inspiration from both multitask learning and hierarchical structures.

Divya Patel, a In 2017, The detection method used in this study is based on sentiment categorization. Results of a comparison of many supervised learning methods for rumor identification are presented in the study. The drawback is that a rumor can only be identified using one feature—sentiment polarity. Their model has an accuracy range of 60–70% depending on the methodology they use.

2018's Rosa Sicilia In order to find an efficient mix of classifiers and features, the research investigates the selection of various characteristics utilizing various classification techniques for Rumor detection. It attempts to examine which characteristics are the most informative and whether or not the recently added features have any impact on the categorization process.

Anand Kumar 2018. This paper proposes a two-step supervised machine-learning framework that first filters rumors before examining their language characteristics. This approach trains numerous classification algorithms with an accuracy greater than 81.079% in an effort to automate filtering. Lastly, rumors are found by applying textual features to the filtered data.

Kaushal Rishabh In 2016. This study presents an algorithm that uses tweets from verified news outlets as a basis to identify rumors on Twitter. The algorithm operates on the idea that Twitter accounts belonging to

Verified News Channels provide reliable information in contrast to user accounts that are unsophisticated and unverified. The study suggests a method for using sentiment and semantic analysis to extract rumors from the data. The model's accuracy falls between 60 and 76%.

2018 saw K. Suthendran. Rather of utilizing Twitter API, the model obtains data using Twitter Scraper. Sentiment analysis is employed for that vast amount of scraped data to distinguish between genuine and fake content. Positive and negative hashtags are used to divide the text's hashtag input.

2019 Oluwaseun Ajao. The research makes the hypothesis that there is a connection between the mood of words posted online and rumors. They take into account the feelings associated with the keywords and words when determining the relevance and use of such terms within the corpus. Finding the most relevant words and ideas within a text corpus is made possible by topic models. It is calculated to find the emotion ratio of negative to positive words.

III. CONCLUSIONS

As the use of social media has grown, so has research into the creation of tools for rumor detection and verification. These tools allow professional practitioners as well as regular users to obtain news and facts instantly, but there is a negative side effect: the spread of unverified information. This research has presented a novel approach to the development of the two primary components, rumor detection and rumor veracity verification, and has reviewed works published in the scientific literature toward the development of rumor detection systems. Significant improvement over previously established rumor identification systems is achieved by combining the veracity verification model with the rumor detection model utilizing real-time data. This online application would function as a cutting-edge platform for several industries, including politics, the government, business, health, and the film industry, to access trending tweets about the relevant topic, analyze sentiment, and find any rumors that may be spreading. This would assist them in stopping the dissemination of incorrect information, misleading instructions,

deliberate defamation, damage to their reputation, dissemination of private information, etc.

REFERENCES

- [1] Dr. Dinesh B. Vaghela¹, Divya M. Patel², "Rumor Detection with Twitter and News Channel Data Using Sentiment Analysis and Classification" in International Journal of Advance Engineering and Research Development Volume 5, Issue 02, February -2018.
- [2] V. Sivasangari, Ashok Kumar Mohan, K. Suthendran, M. Sethumadhavan, "Isolating Rumors Using Sentiment Analysis" in Journal of Cyber Security and Mobility, Vol. 7 1, 181–200. River Publishers, 12 June 2018.
- [3] Oluwaseun Ajao¹, Deepayan Bhowmik² and Shahrzad Zargari, "Sentiment Aware Fake News Detection on Online Social Networks" in ICASSP, IEEE 2019.
- [4] Jin, Z., Cao, J., Guo, H., Zhang, Y., Wang, Y., & Luo, J. (2017, July). Detection and analysis of 2016 us presidential election related rumors on twitter. In International conference on social computing, behavioral-cultural modeling and prediction and behavior representation in modeling and simulation (pp. 14-24). Springer, Cham.
- [5] Krishnan, S., & Chen, M. (2018, July). Identifying tweets with fake news. In 2018 IEEE International Conference on Information Reuse and Integration (IRI) (pp. 460-464). IEEE.
- [6] Kumar, A., Singh, J. P., Dwivedi, Y. K., & Rana, N. P. (2020). A deep multi-modal neural network for informative Twitter content classification during emergencies. *Annals of Operations Research*, 1-32.
- [7] Kwon, S., Cha, M., Jung, K., Chen, W., & Wang, Y. (2013, December). Prominent features of rumor propagation in online social media. In 2013 IEEE 13th International Conference on Data Mining (pp. 1103-1108)
- [8] Neal U. Almost 10 percent Of Twitter Is Spam; 2015. Accessed: 2016-05-12.

<http://www.fastcompany.com/3044485/almost10-of-twitter-is-spam>.

neural networks for early rumor detection, in Pacific-Asia Conference on Knowledge Discovery and Data Mining. Springer, 2018, pp. 4052.

[9] Kalbitzer J, Mell T, BERPohl F, Rapp MA, Heinz A. Twitter psychosis: a rare variation or a distinct syndrome? *The Journal of nervous and mental disease*. 2014

[10] Yang C, Harkreader R, Gu G. Empirical evaluation and new design for fighting evolving Twitter spammers. *IEEE Transactions on Information Forensics and Security*. 2013; 8(8):1280±1293. <https://doi.org/10.1109/TIFS.2013.2267732>.

[11] Kwon, S., Cha, M., & Jung, K. 2017. Rumor Detection over Varying Time Windows. *PLOS ONE*, 12(1), 1- 19. <https://doi.org/10.1371/journal.pone.01683>

[12] Liu, X., Nourbakhsh, A., Li, Q., Fang, R., & Shah, S. 2015. Real-time Rumor Debunking on Twitter. *Proceedings of the 24th ACM International on Conference on Information and Knowledge Management - CIKM '15*, 1867–1870. <https://doi.org/10.1145/2806416.2806651>

[13] Lukasik, M., Cohn, T., & Bontcheva, K. 2015. Classifying Tweet Level Judgements of Rumors in Social Media. *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, (September), 2590–2595. <https://doi.org/10.18653/v1/D15-1311>.

[14] Y.-C. Chen, Z.-Y. Liu, and H.-Y. Kao, Ikm at semeval2017 task8: Convolutional neural networks for stance detection and rumor verification, in *Proceedings of the 11th International Workshop on Semantic Evaluation (SemEval-2017)*, 2017, pp. 465469.

[15] C. Song, C. Tu, C. Yang, Z. Liu, and M. Sun, Ced: Credible early detection of social media rumors, *arXiv preprint arXiv:1811.04175*, 2018

[16] 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*. ACM, 2017, pp. 797806.

[17] T. Chen, X. Li, H. Yin, and J. Zhang, Call attention to rumors: Deep attention based recurrent