



INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

PROGRESSIVE SAGACIOUSNESS OF RUMORS IN ONLINE SOCIAL NETWORK USING ARTIFICIAL NEURAL NETWORK

¹Ayesha Taranum, ²Thejaswini K M, ³Thushara N ⁴Vinutha M

¹Assistant Professor ^{2,3,4}Student

Department of Information Science and Engineering

^{1,2,3,4}GSSS Institute of Engineering and Technology for Women, Karnataka, India.

Abstract: Rumors can spread among thousands of users immediately without verification and can cause serious damages. Rumor detection, allows refuting misinformation before it causes harm. Our project aims in detecting the rumors by using the Bi-Directional Graph Convolutional Networks (Bi-GCN) in Artificial Neural Network. Analyzing contradicting claims and graphs propagation patterns allows accurate rumor identification once they have sufficiently spread.

Index Terms -- Bi-GCN, networks, Artificial Neural Network, graphs, patterns.

I. INTRODUCTION

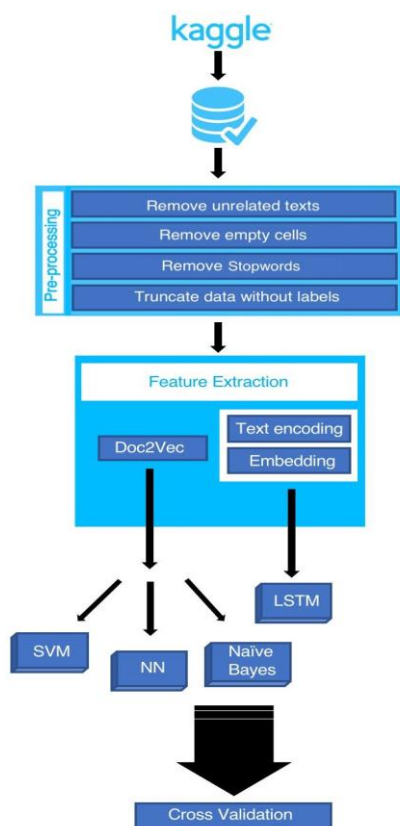
Recent uproar of fake news and misinformation on social media platforms has sparked the interest in the scientific community to automatically detect and refute them. The most popular research task to counteract misinformation, Rumor detection, requires repeated signals to reach adequate detection accurate. Consequently, rumor detection recognizes rumors only when they have started spreading and causing harm. We introduce a new task called “rumor prediction” that assesses the possibility of a document arriving from a social media stream becoming a rumor in the future. Note that rumor prediction differentiates itself from rumor detection through instant decision making. This allows refuting misinformation before it spreads and causes harm. Our approach to rumor prediction harnesses content-based features in combination with novelty-based features and pseudo feedback. Our experiments show that we are able to accurately predict, whether a document is a rumor or not. Additionally, we show how rumor prediction can significantly improve the accuracy of rumor detection systems.

A. ARTIFICIAL NEURAL NETWORKS (ANN)

Artificial Neural Network (ANN) uses the processing of the brain as a basis to develop algorithms that can be used to model complex patterns and prediction problems. ANN is a machine learning algorithm used for classification, regression and clustering problems. It is the building block of deep neural networks. It’s majorly used to learn complex non-linear hypothesis when the data-set is too large where we will end up with too many features. ANN uses different layers of mathematical processing. It has several units organized in several layers. A single unit is called a neuron. The input units in the input layer receives various information from the outside world as inputs.

The main objective of the project is to

- The aim is to develop a real time application for society and to identify Fake messages or rumors.
- The objective is to build a system which avoids damage to the society by predicting rumors.
- To approach the problem only based on content, by extracting each word from the post or article then match those with a dictionary or bag of words for classification.
- To build a system, which improve accuracy of prediction whether an information is rumors or not, to develop real time application which is useful for the society.



II. LITERATURE REVIEW

Before you begin to format your paper, first write and save the content as a separate text file. Keep your text and graphic files separate until after the text has been formatted and styled. Do not use hard tabs, and limit use of hard returns to only one return at the end of a paragraph. Do not add any kind of pagination anywhere in the paper Do not number text heads—the template will do that for you. Finally, complete content and organizational editing before formatting. Please take note of the following items when proof reading spelling and grammar returns of the shares and estimated betas.[1] Many rumor detection methods have been proposed in recent years, and most of them are based on machine learning (ML) techniques. The common challenge in most of these works is related to feature extraction from a considered dataset. The task of manual extraction of features requires a lot of time and effort, with a limited efficiency in detecting rumors for most of these works. Recently, deep neural networks have been proposed to simplify the extraction of features and to provide a strong ability for abstract representation learning.[2] The identification of rumors obtains a high-level of consensus among annotators. Although the research on rumor detection is in full swing, there is quite a few research on the impact of rumors. Some social media providers even hire senior journalists working 24h and 7d every week to maintain an official account , which exposes new rumors regularly, in order to minimize the negative impact of rumors on their platform. To the best of our knowledge, most existing works regarding rumor impact are solely based on prior knowledge or various other assumptions or even human power. This paper proposes a RISM model that can predict the impact of a new rumor, which thus provides a basis for rumors filtering. ;An Yunjie, describes The SMB have very important status and plays an important role in our country's economic development. But their existence and development are determined by technological innovation. It is urgent for that SMB to need technological innovation, but their economic and technical strength is weak, the imitating innovation model is the first selection, the cooperation innovation model comes next. Independent technological innovation model doesn't suit to them at present time.[4] Penetration testing helps to secure networks, and highlights the security issues. In this paper investigate different aspects of penetration testing including tools, attack methodologies, and defense strategies. More specifically, we performed different penetration tests using a private networks, devices, and virtualized systems and tools. We predominately used tools within the Kali Linux suite. The attacks we performed included: smartphone penetration testing, hacking phones Bluetooth, traffic sniffing, hacking WPA Protected Wifi, Man-in-the-Middle attack, spying (accessing a PC microphone), hacking phones Bluetooth, and hacking remote PC via IP and open ports using advanced port scanner. The results are then summarized and discussed. The paper also outlined the detailed steps and methods while conducting these attacks [5] In connection with mass introduction of robots in various spheres of activity, and also absence of due attention to such factor as safety, the probability of unauthorized access to their blocks of management raises. Typical decisions on safety maintenance by introduction of information security systems are not suitable for the robotized platforms in view of their capacities limitation. Thus, there is a problem of providing protection of the control unit of robotic platforms from various threats. The results of the analysis of possible approaches to the formation of a list of threats and vulnerabilities of the robotic platform are shown. The use of mathematical models to describe the spreading of ideas and information has now been established for half a century. There are many rumor spreading models were coming up with researchers. Daley and Kendal came up with a rumor model named DK model, which divided population of interest into three groups: people who know and spread the rumor, people who do not know the rumor and people who know but do not

transmit the rumor. After that, there had sprung up several rumor transmissions models, such as SIHR rumor spreading model, SIR rumor spreading model, Ising model, model with various contact interactions and model with variable forgetting rate. Flavio Chierichetti designed a classic preferential attachment model for rumor transmitting in social networks. Rumors and epidemics are popularly treated as similar. A kind of researchers had done their researches on rumors and epidemics. [4] There has been an increasing research interest on methods for combating propaganda, and this interest persists despite the reality that these two areas are not yet well established and there are no systematic reviews available for further research. Systematic Literature Reviews (SLRs), are employed under circumstances when there is insufficient empirical evidence available, or when the topic area is too vast for an SLR to be conducted. Thus, systematic mapping was conducted because systematic mapping would give an extensive, wide and detailed overview of the research area where it would map out and categorize existing literature; which would fit the objectives of this study. For conducting the review, the guidelines followed were described by Petersen in his paper on conducting systematic mapping. This paper aims at presenting not only an overview but also an understanding of how mitigation effort contributes to different areas. [5] In general, to detect false news can use a detection approach (SD) in the approach that automatically defines whether the authors of a text supporting the target provided or against it can be implemented by using articles, headlines from the news with a variety of methods. In addition to the SD approach, marrow approach and comparisons can be used as an alternative, the way this approach works is to summarize the news and compare it with actual news to classify the news as false news to determine the model. In this paper, deep learning is the solution to the problem especially the programming because if the green side of the accuracy is very good. In addition, deep learning can also be implemented for large amounts of data, Recurrent Neural Network (RNN).

III. METHODOLOGY

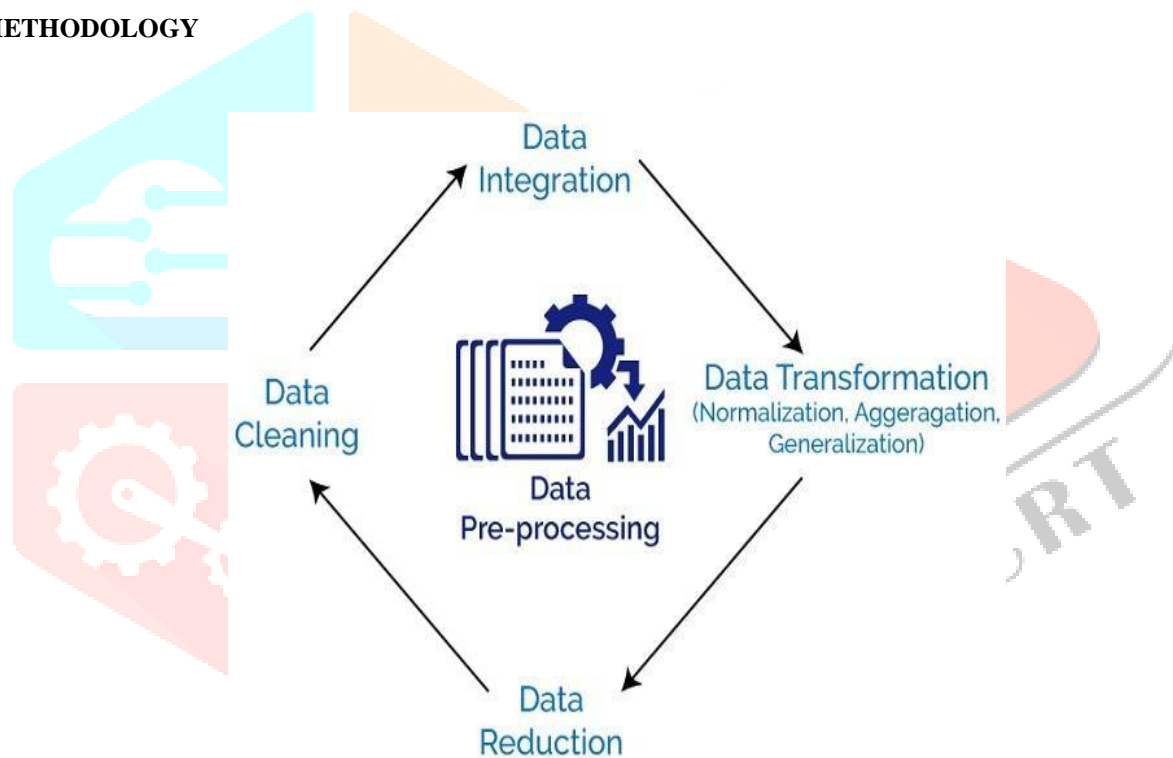


Fig 1: Architecture of Rumors prediction

The system will take post as an input from the user, tokenization and steaming will be done on the post using NLP. After that Keywords will be extracted from the post and it will be compared with a model and model will predict if the given post is fake or true and it will classify the post.

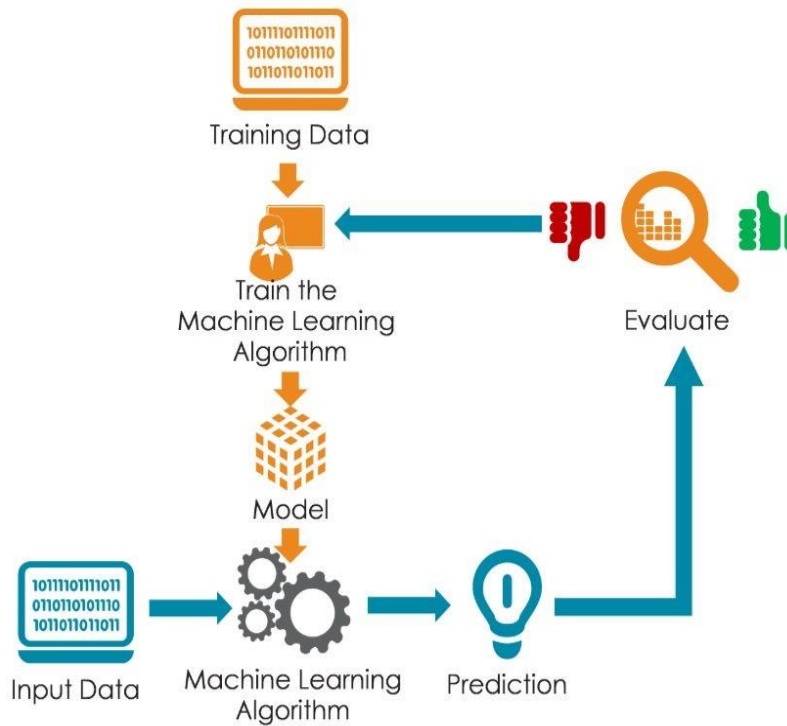


Fig 3: Training a model

Models used:

1. Naïve Bayes
2. Support vector machine
3. Neural Network
4. Long Short-term memory (LSTM)

Naïve Bayes: Classification technique based on Bayes theorem with an assumption of independence among predict. Convert datasets into a frequency table, create likelihood table by finding probabilities. Use naïve Bayesian equation to calculate posterior probability for each class.

$$P(c | x) = \frac{P(x | c)P(c)}{P(x)}$$

Likelihood
Class Prior Probability

Posterior Probability
Predictor Prior Probability

$$P(c | X) = P(x_1 | c) \times P(x_2 | c) \times \dots \times P(x_n | c) \times P(c)$$

Support Vector Machine (SVM)

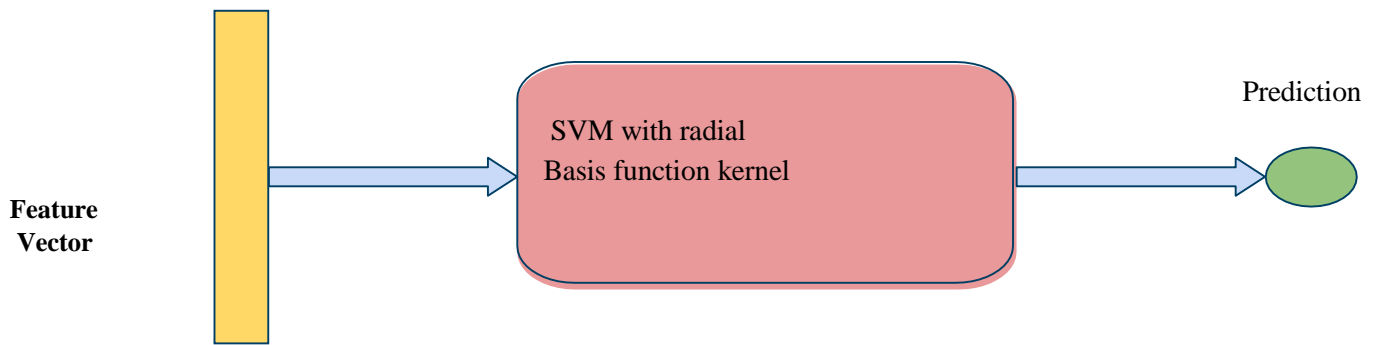


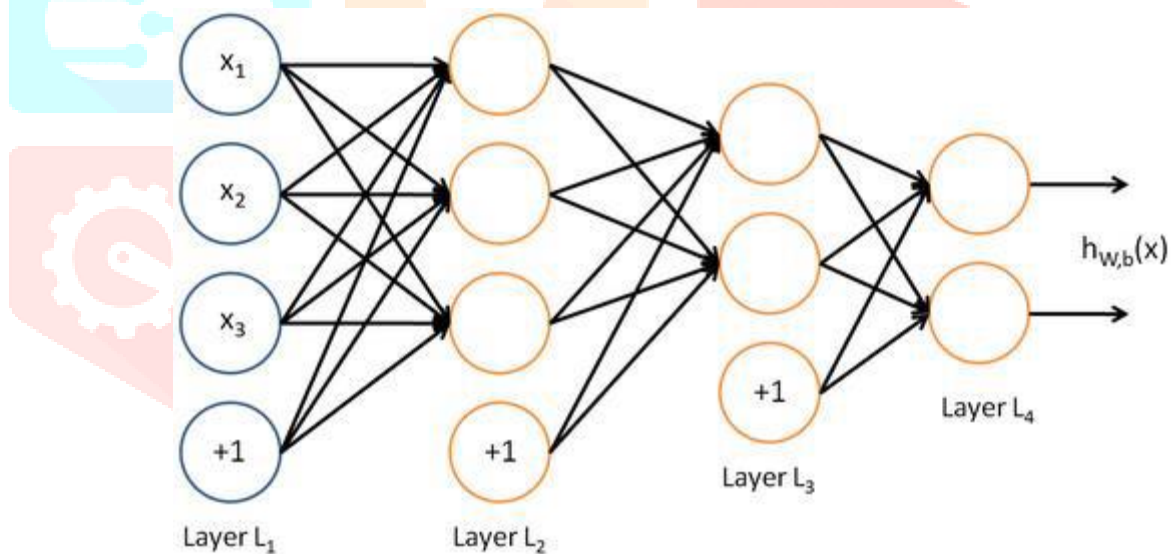
Fig 4: Support Vector Machine

Neural Network:

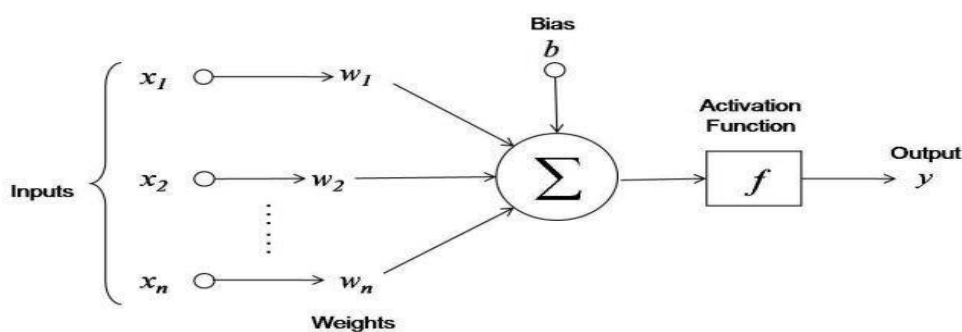
Artificial Neural Network (ANN) is a machine learning algorithm used for classification, regression and clustering problems. It is the building block of deep neural networks.

ANN uses different layers of mathematical processing. It has several units organized in several layers. A single unit is called a neuron. The input units in the input layer receives various information from the outside world as inputs. From here the data goes to the hidden unit to transform the data into a form in which output units can use it.

The below figure has 1 input layer with 4 input units, 2 hidden layers with the first hidden layer having 4 neurons and the second having 3 neurons and finally 1 output layer with 2 output units.



Initially, a neuron adds the values of every neuron from the previous layer it is connected to. In the below figure the neuron has got 3 inputs. These 3 values will be multiplied by the weights and added together (w_1, w_2, w_3). Weights are the values that determine the strength between two neurons. Those are the values that will be altered in the learning process. Then a bias value will be added with the previous summation. After all these summations the neuron applies an activation function to the value.



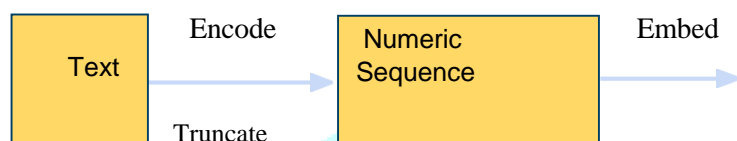
That's basically what a neuron does. Receives values from connected neurons multiplies it by respective weights, adds them and applies an activation function. Then it sends it to the other neurons. After every neuron from a layer has done it passes to the next layer. Finally, the last value obtained should be able to predict the desired output. At first, the neural network's predictions will be random. But as each epoch passes and trains the model on that the output should be, the prediction will become closer and closer to the correct value.

This above working of single-neuron should be carried out throughout the network and is carried out by a process called forward pass/propagation. Using that we start with the activation of input units and then we forward propagate that to the hidden layer and compute the activation of the hidden layer and sort of forward propagate so we that can compute the activation of the output layer.

Then to determine the optimum weights/parameters a process called back-propagation is done. In that, the weights are optimized so that the neural network can learn how to correctly map random inputs to outputs.

Fully connected layer

Long Short- Term Memory (LSTM)



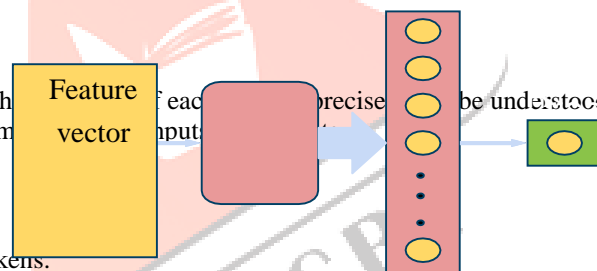
Prediction

LSTM

Fig 5: Long Short-Term Memory

Data flow diagram

The DFD is about the flow of task and execution. DFD is written so that data flow diagram (DFD) illustrates how data is processed by a system.



- User will provide news as input to the system.
- NLP will process the post and generate keywords and tokens.
- Keywords will be extracted and compared with a model and the model will predict the output.

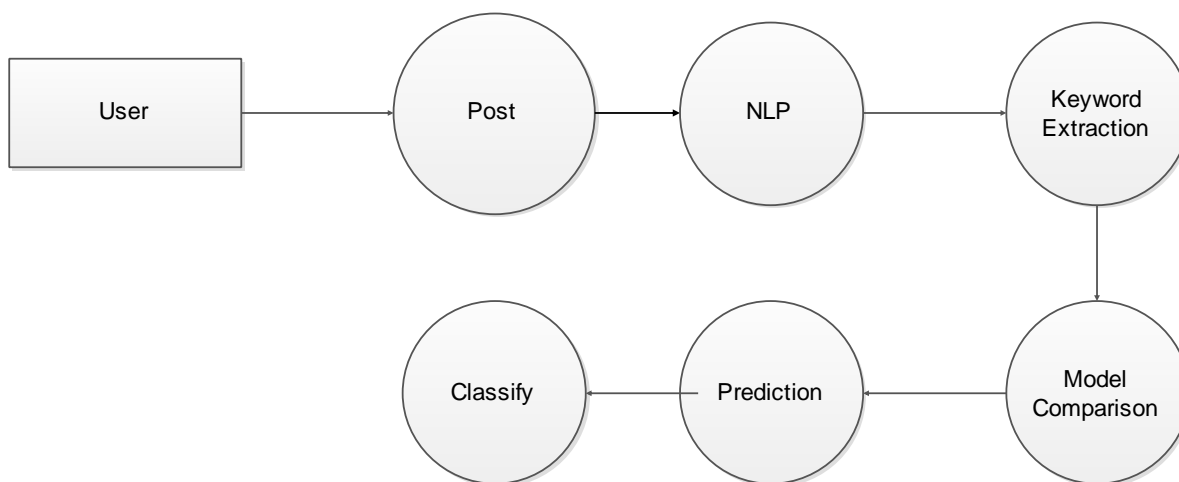


Fig 6: Data flow diagram of rumors prediction

Algorithm NLP:

Step1: The term morphology is concerned with the interplay between words and their relationship with other words

Step2: Syntax defines how words and sentences are put together

Step3: Semantics is the study of the meaning of words and groups of words

Step4: Pragmatics is used to explain the content of spoken expressions

Step5: And lastly, phonology covers the acoustic structure of spoken language and is essential for language recognition

CONCLUSION

It is significant to find the accuracy of news which is available on internet. In the project, the components for recognizing Fake news are discussed. A mindfulness that not all, the fake news will propagate via web-based networking media. Currently, to test out the proposed method of Naïve Bayes classifier, SVM, and NLP are used. In future, ensuing algorithm may provide better results with hybrid approaches for the same purpose fulfilment. The mentioned system detects the fake news on the based on the models applied. Also it had provided some suggested news on that topic which is very useful for any user. The System is helpful to public and most importantly social media in order to avoid rumors. System classifies the Rumors into FAKE or REAL NEWS.

Future work

In the future, the efficiency and accuracy of the prototype can be enhanced to a certain level, and also enhance the user interface of the proposed model.

REFERENCES

- [1] MOHAMMED AL-SAREM^{1,2}, WADII BOULILA (Senior Member, IEEE), MUNA AL-HARBY, JUNAID QADIR, (Senior Member, IEEE), AND ABDULLAH ALSAEEDI, “Deep Learning-Based Rumor Detection on Microblogging Platforms: A Systematic Review”, 2020.
- [2] Jun Yin, Shaowu Liu, Qian Li, Guandong Xu, Advanced Analytics Institute, School of Computer Science University of Technology Sydney, “Prediction and Analysis of Rumour’s Impact on Social Media”, 2019.
- [3] Hui Zhang, Ji Li, Yueliang Xiao, “Hadoop Cellular Automata for Identifying Rumor in Social Networks”, 2018.
- [4] Aimi Nadrah Maseria, Azah Anir Normana, Christopher Ifeanyi Ekeab, Atif Ahmadc, Nurul Nuha Abdul Molokd Department of Information System, Faculty of Computer Science and Information Technology, University of Malaya, Kuala Lumpur, “Socio-Technical Mitigation Effort to Combat Cyber Propaganda: A Systematic Literature Mapping”, 2018.
- [5] Ghinadya, Suyanto Suyanto, School of Computing Telkom University Bandung, Indonesia, “Synonyms-Based Augmentation to Improve Fake News Detection using Bidirectional LSTM”, 2017
- [6] X. Liu, A. Nourbakhsh, Q. Li, R. Fang, et al., “Real-time rumor debunking on twitter”, Proceedings of the 24th ACM International Conference on Information and Knowledge Management, ACM, Melbourne, Australia, pp.1867–1870, 2015.
- [7] Fan Yang, Xiaohui Yu, Yang Liu, et al., “Automatic detection of rumor on Sina Weibo”, Proceedings of the ACM SIGKDD Workshop on Mining Data Semantics, Beijing, China, Page 13, 2015.
- [8] X. Zhou, J. Cao, Z. Jin, et al., “Realtime news certification system on sina weibo”, Proceedings of the 24th International Conference on World Wide Web, Florence, Italy, pp.983–988, 2014.
- [9] S. Muthukrishnan, “Data streams: Algorithms and applications”. Foundations and Trends in Theoretical Computer Science, Vol.1, No.2, pp.117–236, 2014.

- [10] K. Wu, S. Yang and K. Zhu, “False rumours detection on Sina Weibo by propagation structures”, 2015 IEEE 31st International Conference on Data Engineering, Seoul, South Korea, pp.651–662, 2013.
- [11] Shihan Wang and Takao Terano, “Detecting rumour patterns in streaming social media”, IEEE International Conference on Big Data, Guimi, Santa Clara, CA, USA, pp.2709–2715, 2013.
- [12] C. Castillo, M. Mendoza and B. Poblete, “Information credibility on Twitter”, The 20th International Conference on World Wide Web, Hyderabad, India, pp.675–684, 2012.
- [13] S. Petrovic, M. Osborne, R. McCreddie, et al., “Can Twitter replace Newswire for breaking news?”, Proceedings of the Seventh International AAAI Conference on Weblogs and social media., Massachusetts, USA, pp.713–716, 2012.
- [14] 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, Florence, Italy, pp.1395–1405, 2012.
- [15] G. Cai, H. Wu and R. Lv, “Rumours detection in Chinese via crowd responses”, IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), Beijing, China, pp.912–917, 2011.
- [16] S. Kwon, M. Cha, K. Jung, et al., “Prominent features of rumor propagation in online social media”, 13th International Conference on Data Mining (ICDM), IEEE, Dallas, TX, USA, pp.1103–1108, 2011.
- [17] V. Qazvinian, E. Rosengren, D.R. Radev, et al., “Rumour has it: Identifying misinformation in Microblogs”, Proceedings of the Conference on Empirical Methods in Natural Language Processing, Edinburgh, UK, pp.1589–1599, 2010.

