

# Faultless Decision Making for False Information in Online: A Systematic Approach

Yasir Babiker Hamdan,  
International University of Africa (IUA),  
Khartoum, Sudan.  
[yasir20ap@iua.edu.sd](mailto:yasir20ap@iua.edu.sd)

Prof. Sathish,  
Department of EEE,  
Eritrea Institute of Technology,  
Eritrea.  
[sathesh4you@gmail.com](mailto:sathesh4you@gmail.com)

**Abstract-** An identifying the news are real or fake instantly with high accuracy is a challenging work. The deep learning algorithm is implementing here to acquire very accurate separation of real and fake news rather than other methods. This research work constructs naïve bayes and CNN classifiers with Q-learning decision making. The two different approaches detect fake news in online and it gives to decision making section which is designed at tail in our research. The deep decision making section compares the input and make the decision wisely and it provides the more accurate output rather than single classifiers in deep learning. This research work comprises compare between our proposed works with single classifiers.

**Keywords:** *Naïve Bayes, Convolutional Neural Network, Decision making*

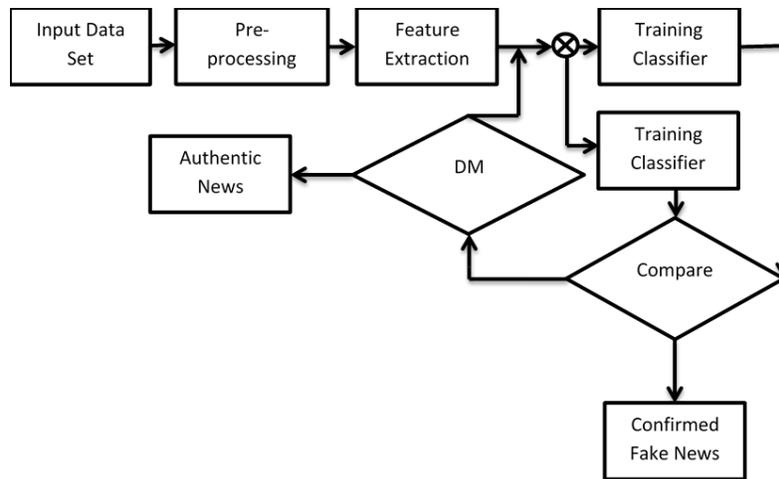
## 1. INTRODUCTION

Information is priceless on social media; recently, accessing internet and social media is increasing rapidly in a worldwide. Misinformation or fake news is spreading abundant these days in internet to making public nuisance, popularity between peoples, and extra [1]. This can have harmful effects on our society and it should be controlled. This false information can create erroneous society and panic situation in any nation easily. The user of accessing the internet increases day by day. So this social media news can reach to those people in their daily life [2]. The figure 1 shows one of the examples of spreading fake news in the society. Russia allow to run riot more than 500 lions on the street to control the people to stay in their home itself for COVID'19. This will create panic situation in the society and needs some medication for old age people during this time. So there is research gap of fast detection of fake news in the social media [3, 4]. Many research uses naïve bayes classification for fake news detection to give better accurate results. The research proposes classifying Facebook posts fake or authentic by naïve bayes classification. Here the model was trained with 11K articles of four columns name index, title, content and label were labeled as authentic or fake [5].



**Figure 1** Example of Fake News

The figure 2 shows process flow of proposed model. Here the feature extraction and training classification will be done in same section name deep learning. Comparing two training classifier (TC) and provides “confirmed fake news” while both are same negative output. Comparing section will direct to deep decision making section for remaining all other possibilities. The Even though both are positive output from training classifiers, deep decision section will check less number of iteration to get greater accuracy for faultiness decision making of false information in an online.



**Figure 2** Block diagram for overview of proposed model

The joining of decision making (DM) section provides better accuracy for the proposed model.

**Motivation**

In more populated country, the growth of fake news contributes disaster to democracy and integrity. Recently, intruders are targeting the innocent people to familiarize the fake news in the name of humor memes and changing their state of mind further. This one is motivated us to do detect fake news in social media with higher accuracy in quicker. The greater accuracy of finding fake news in social media algorithm is a mandate research currently.

## 2. RELATED WORKS

In the year 2016, the word “Fake News” was popularized during the time of US presidential election campaigns. There are more research articles about fake news detection after this campaign [6]. The naïve bayes, decision tree, state vector machines were implemented for fake news detection. Also random forest ensemble methods are used. Here 4Lakh news articles are using from trust worthy news sources [9, 10]. The deep convolutional neural network algorithm is implementing for detecting the fake news. There is absence of adding any auxiliary features in further and lack of contain news video images shows in figure 1 [11, 12]. The deep neural network is used and LSTM and Feed-forward Neural Network and word embedding methods are used to increase the accuracy with the help of word2vec and GloVe models. Here they added auxiliary features to the data. But the domain name was limited with 10 due to increase the speed of the process [13, 14, 15]. They were unsuccessful in the accuracy of the classification. Peng Qi et al introduce multi domain visual neural network model to detect the fake news with the use of optical data [16]. The frequency domain sub network contains two fully connected layers with 16 and 64 neurons. The over fitting problems are solving by dropout layer with the rate 0.4 is added. Each fully connected layer is connected by a dropout layer with the rate of 0.5. They classified with weibo dataset contains 4749 fake news article and 4779 real news articles. The 84.5% of accuracy was yielded by proposed method. They add more data set from the Twitter Multimedia data set as future work [17, 9, 2]. Yin-Fu et al proposes ensemble learning model based on self-adaptive harmony search (SAHS) algorithm to detect fake news with highest accuracy. The proposed algorithm is working well to optimize the weight of combined model. The accuracy is less due to it contain double edge sword model as classifier. They concluded that cross domain intractability issues will be addressed in future work [18, 19, 14, 26]. Kyeong-hwan Kim et al compares many machine learning classifiers techniques and mainly focuses on the short sentences news for produce a credibility score for the classification. Also here the dataset contains 3 columns named author, content and output label [20, 21, 8, 22].

Emerson Yoshiaki Okano et al evaluated the Hierarchical Attention Network (HAN) model for Brazilian Portuguese fake news from reference dataset. It comprises of 3600 fake news articles and 3600 authentic news articles for detecting the fake news. They achieved 97% accuracy to detect fake news automatically [22, 7, 19, 23]. The three level HAN was proposed by sneha singhania et al. The trained model comprises word encoder and attention, sentence encoder and attention, headline encoder and attention which are used to create news vector with effective representation of an article. It provides good accuracy than other state of art models. They are improving detection technique for fake news in a real time online social media automatically as a further development [24].

The research paper proposes fake news net data set includes multidimensional information. News specific and tweet specific attributes used for training the mode with the help of decision tree and random forest [25]. This CNN research has technique that news in the form of an image cannot be verified in the model [12]. Data mining stage is integrated in this system. This is used to add supplementary data to the considered news (LSTM FF). MVNN is used to fuse the visual information of frequency and pixel domains. Visual features are extracted from different semantic levels [18]. The model must be manually initialized to keep the sources up to date (HAN). Performance suffers when long sentences or words outside vocabulary is encountered [25].

## RESEARCH GAP

All above discussed researches highlighting mainly data mining or machine learning sections or data mining and machine learning section or both as well as. There is research gap for decision making to improve the accuracy of the system whereas finalized from the machine learning study.

## PROPOSED SOLUTION

In this research, we established decision making section for further accuracy improvement in the system after comparison of ML algorithm. If same like pole output of ML algorithm, the decision will be easy and compares it with data mining section. If the different result from ML algorithm, the decision will conclude with aid of data mining section after one more time of classification from both ML algorithms. It has tested and found good accuracy improvement for the model.

### 3. ORGANIZATION OF THE RESEARCH

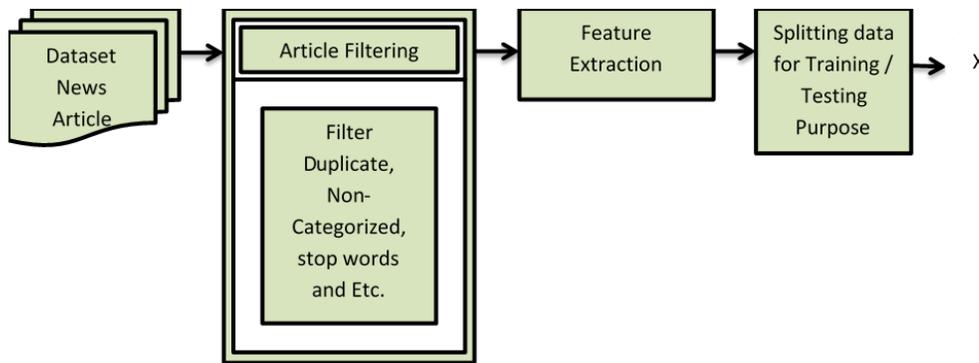
This research article covers introduction, related work, and motivation, proposed methodology, results and discussion, conclusion in a row for the systematic approach to faultiness decision making for false information in online.

### 4. PROPOSED METHOD

Our proposed algorithm have three sections includes data mining, deep learning, decision making sections as follows.

#### a) Data Mining Section

This section consist of loading the input dataset, removal of duplicate data, non-categorized data, stop words, NAN value from the text column of the news article [19,27]. There are processes of splitting the data for training and testing after feature extraction.



**Figure 4** Process of Data Mining Section

#### Step 1:

*Prepare dataset using web scraping from reliable and neutral news sources and preprocessing*

#### Step 2:

Feature extraction for weighting term (TF-Inverse Document Frequency)

$$\text{Term Frequency} = \frac{\text{No. of times 't' appears in a document}}{\text{total number of term in the document}}$$

$$IDF(t) = \log_e \left( \frac{\text{Total No of Documents}}{\text{No of Documens with 't' in it}} \right)$$

The data set contains ID, heading, author, content or text and fake or real classification column.

#### b) Machine Learning Section

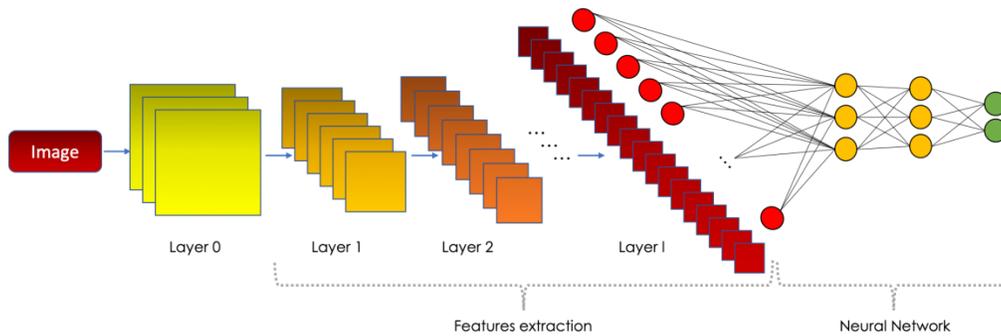
In the training process flow, the collection of preprocessed labeled news articles will be trained. Currently, this detection model is built and completing the learning section. This one will be classified by CNN and Naïve bayes method separately.

**Step 3:**

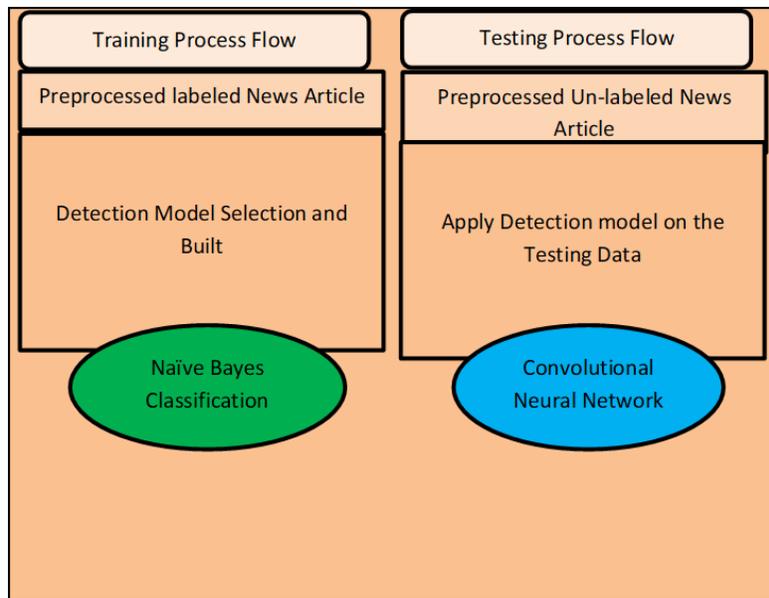
**Part 3(a)** Separately train a CNN and Naïve bayes method as they have been observed to give highest accuracy has a memory that gives the errors to the previous layers.

**Part 3(b)** For one lined texts, web mining will be done to substantiate the information and for image  
 The simple equation of naïve bayes theorem is defined as,

$$P\left(\frac{c}{x}\right) = \frac{P\left(\frac{x}{c}\right)p(c)}{p(x)}$$



**Figure 5** Convolutional Neural Network classification

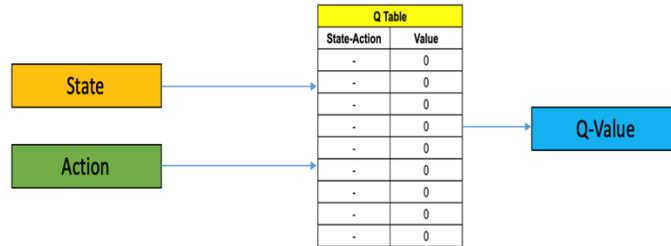


**Figure 6** Process of Machine learning Section

**c) Decision making Section**

The decision making section accomplish deep learning from positive and negative reinforcement after combining the two effective algorithm to detect the fake news articles in an online. Generally speaking, the reinforcement techniques are teaching the machine learning algorithm that how to act in a certain environment. There

is construction of Q table for state and action to deliver the Q value to agent. More number of iteration provides a good Q table for good accurate results. The figure 7 & 8 shows the creation of Q table and deep Q learning construction details.



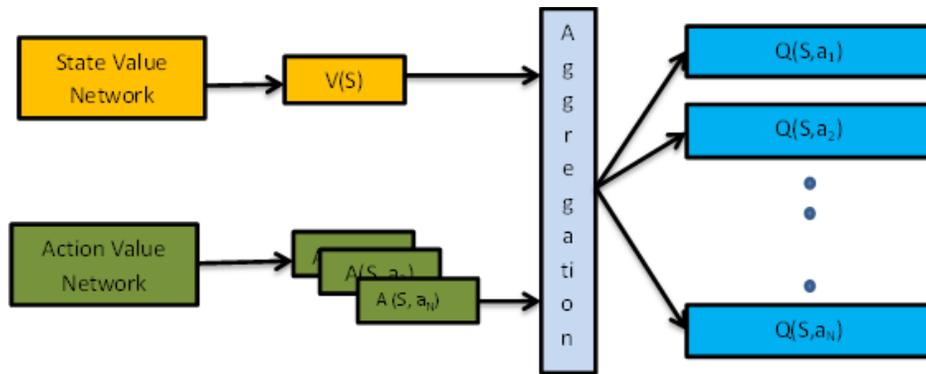
**Figure 7** Creation of Q table for Decision making

**Step 4:**

Part 4(a) Initialize Q table

Part 4(b) Choose and perform an Action

Part 4(c) Measure the reward and update the Q table to get precise results further



**Figure 8** Construction of Deep Q- Learning Decision making

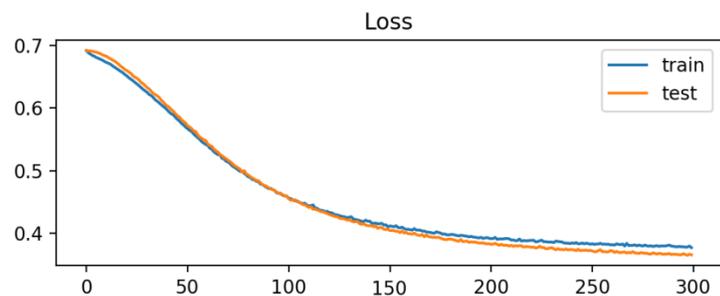
**5. RESULTS AND DISCUSSIONS**

The figure 9&10 shows loss and accuracy dataset during training and testing respectively. These responses are almost similar with each other. The following formulas are used to measure the metric measurement of the proposed model.

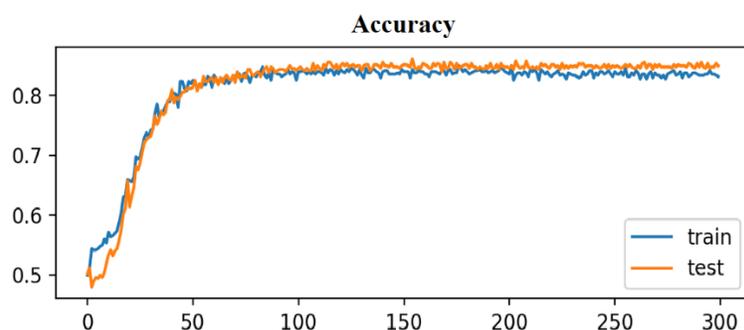
$$Precision = \frac{TP}{TP + FP}$$

$$Sensitivity = \frac{TP}{TP + FN}$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$



**Figure 9** Loss during training and testing the data



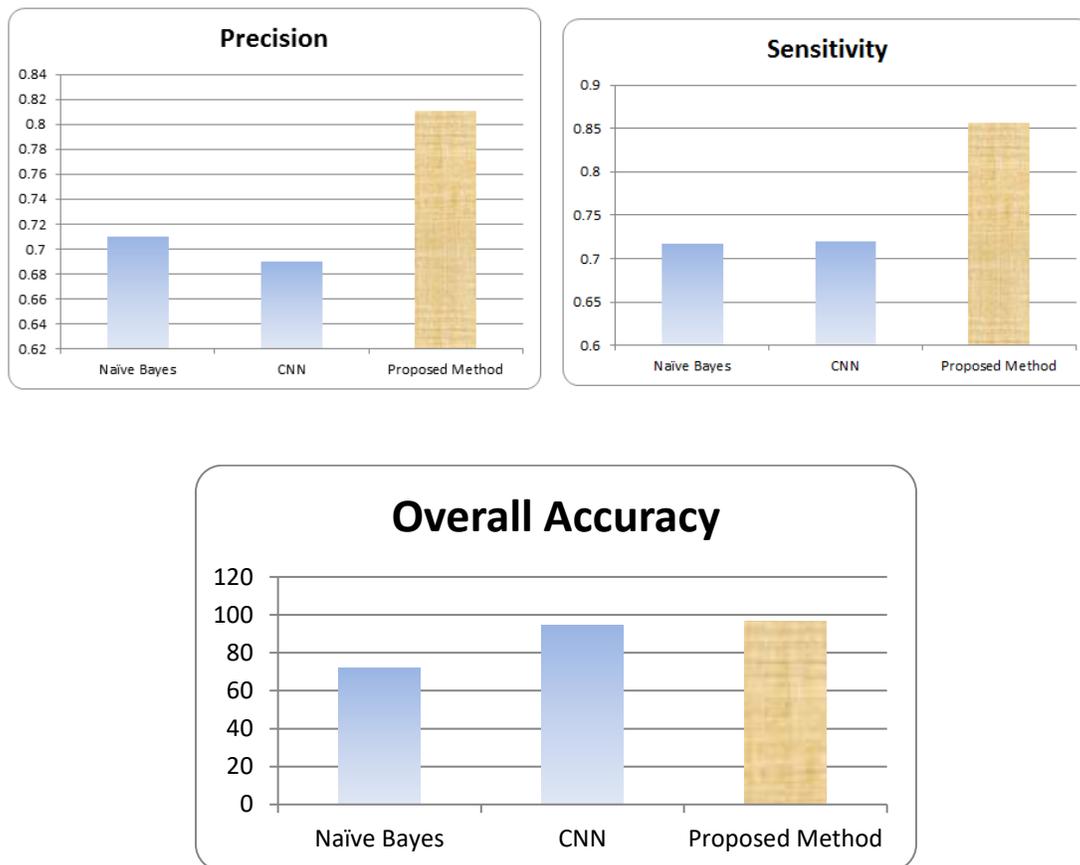
**Figure 10** Accuracy during training and testing the data

We have tabulated metric measurements for different machine learning algorithm and shown in table 1. The sensitivity measures the actual positive samples from the decision. The precision parameter provides accurate the positive prediction in the predicted dataset. The overall performance of the model can be identified by accuracy metric. We have done lot of iteration to construct an accurate Q-table. The learning from positive and negative reinforcement, machine learns how to act in a certain environment. In order to maximize the rewards, it will learn from mistakes.

**Table 1** Comparison of Machine learning algorithms with measuring metrics

S. No	Methods / Metric	Precision	Sensitivity	Accuracy
1	Naïve Bayes	0.71	0.717	71.84%
2	CNN	0.69	0.72	94.27%
3	Proposed Method	0.811	0.857	96.8%

The following figure 11 shows the overall performance of the proposed algorithm. From these graphs there is a clarity of proposed model overall performance as well as faultiness decision making.



**Figure 11** Overall accuracy of proposed Method

## 6. CONCLUSION

Our proposed model involves of comparing accuracy with single classifier. The accuracy level has improved and executing time is increased due to the more processes in the deep Q-learning. The executing time is an unsatisfactory result in our research work. In future, we will be evaluating the proposed algorithm with different reinforcement learning methods such as R-learning, TD learning for decision making. While increasing of raw data for classification, the executing time is more compared with single classifier in the model. So the proposed method will be analyzed with fast decision making algorithm in future. The revised algorithm has to implement for large records for training and testing data. Our proposed algorithm needs optimization techniques to tune decision making parameters in the model for better framework for improved accuracy.

## ACKNOWLEDGEMENT

We are most grateful to Mr. Efreem Yamane, Department of Electrical & Electronics Engineering, Mainefhi College of Engineering, for his assistance to collect the data and complete this research work.

## REFERENCES

- [1] Roy, A., Basak, K., Ekbal, A., & Bhattacharyya, P. (2018). "A Deep Ensemble Framework for Fake News Detection and Classification" ArXiv, abs/1811.04670.
- [2] N. X. Nyow and H. N. Chua, "Detecting Fake News with Tweets' Properties," *2019 IEEE Conference on Application, Information and Network Security (AINS)*, Pulau Pinang, Malaysia, 2019, pp. 24-29, doi: 10.1109/AINS47559.2019.8968706.
- [3] Pathak, Ajeet & Mahajan, Aditee & Singh, Keshav & Patil, Aishwarya & Nair, Anusha. (2020). "Analysis of Techniques for Rumor Detection in Social Media" *Procedia Computer Science*. 167. 2286-2296. 10.1016/j.procs.2020.03.281.
- [4] Tijare, Poonam. (2019). "A Study on Fake News Detection Using Naïve Bayes, SVM, Neural Networks and LSTM"
- [5] A. Jain and A. Kasbe, "Fake News Detection," *2018 IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS)*, Bhopal, 2018, pp. 1-5, doi: 10.1109/SCEECS.2018.8546944.
- [6] Ajao, Oluwaseun & Bhowmik, Deepayan & Zargari, Shahrzad. (2018). Fake News Identification on Twitter with Hybrid CNN and RNN Models. 10.1145/3217804.3217917.
- [7] Habib, Ammara & Asghar, Dr. Muhammad & Khan, Adil & Habib, Anam & Khan, Aurangzeb. (2019). "False information detection in online content and its role in decision making: a systematic literature review" *Social Network Analysis and Mining*. 9. 10.1007/s13278-019-0595-5.
- [8] Granik, Mykhailo & Mesyura, Volodymyr. (2017). "Fake news detection using naive Bayes classifier" 900-903. 10.1109/UKRCON.2017.8100379.
- [9] S. Helmstetter and H. Paulheim, "Weakly Supervised Learning for Fake News Detection on Twitter," *2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, Barcelona, 2018, pp. 274-277, doi: 10.1109/ASONAM.2018.8508520.
- [10] Drif, Ahlem & Ferhat Hamida, Zineb & Giordano, Silvia. (2019). "Fake News Detection Method Based on Text-Features"
- [11] T S, Steni & P S, SREEJA. (2020). "Fake News Detection on Social Media-A Review" *Test Engineering and Management*. 83. 12997-13003.
- [12] Kaliyar, Rohit & Goswami, Anurag & Narang, Pratik & Sinha, Soumendu. (2020). FNDNet- A Deep Convolutional Neural Network for Fake News Detection. *Cognitive Systems Research*. 61. 10.1016/j.cogsys.2019.12.005.
- [13] Pierri, Francesco & Ceri, Stefano. (2019). "False News On Social Media: A Data-Driven Survey"
- [14] Randika, Banura. (2020). "The Misinformation Era: Review on Deep Learning Approach to Fake News Detection" 10.6084/m9.figshare.13299440.v1.
- [15] S, Deepak & Chitturi, Bhadrachalam. (2020). Deep neural approach to Fake-News identification. *Procedia Computer Science*. 167. 2236-2243. 10.1016/j.procs.2020.03.276.
- [16] Alam, Shahid & Ravshanbekov, Abdulaziz. (2019). "Sieving Fake News From Genuine: A Synopsis"
- [17] P. Qi, J. Cao, T. Yang, J. Guo and J. Li, "Exploiting Multi-domain Visual Information for Fake News Detection," *2019 IEEE International Conference on Data Mining (ICDM), Beijing, China, 2019, pp. 518-527, doi: 10.1109/ICDM.2019.00062.*
- [18] Huang, Yin-Fu & Chen, Po-Hong. (2020). Fake News Detection Using an Ensemble Learning Model Based on Self-adaptive Harmony Search Algorithms. *Expert Systems with Applications*. 159. 113584. 10.1016/j.eswa.2020.113584.
- [19] Nguyen, John, "USING DEEP LEARNING AND LINGUISTIC ANALYSIS TO PREDICT FAKE NEWS WITHIN TEXT" (2020). *Master's Projects*. 931. [https://scholarworks.sjsu.edu/etd\\_projects/931](https://scholarworks.sjsu.edu/etd_projects/931).
- [20] K. Kim and C. Jeong, "Fake News Detection System using Article Abstraction," *2019 16th International Joint Conference on Computer Science and Software Engineering (JCSSE)*, Chonburi, Thailand, 2019, pp. 209-212, doi: 10.1109/JCSSE.2019.8864154.
- [21] Aphiwongsophon, Supanya and P. Chongstitvatana. "Detecting Fake News with Machine Learning Method." *2018 15th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON)* (2018): 528-531.

[22] Yang, Kai-Chou et al. "Fake News Detection as Natural Language Inference." ArXiv abs/1907.07347 (2019): n. pag.

[23] Okano, Emerson & Liu, Zebin & Ji, Donghong & Ruiz, Evandro. (2020). "Fake News Detection on Fake.Br Using Hierarchical Attention Networks" 10.1007/978-3-030-41505-1\_14.

[24] Singhanian, Sneha & Fernandez, Nigel & Rao, Shrisha. (2017). 3HAN: A Deep Neural Network for Fake News Detection. 10.1007/978-3-319-70096-0\_59.

[25] 'Bilateral Multi-Perspective Matching for Natural Language Sentences' - Zhiguo Wang, Wael Hamza, Radu Florian IBM T.J. Watson Research Center

[26] N. X. Nyow and H. N. Chua, "Detecting Fake News with Tweets' Properties," 2019 IEEE Conference on Application, Information and Network Security (AINS), Pulau Pinang, Malaysia, 2019, pp. 24-29, doi: 10.1109/AINS47559.2019.8968706.

[27] Xinyi Zhou and Reza Zafarani. 2020. "A Survey of Fake News: Fundamental Theories, Detection Methods, and Opportunities" ACM Comput. Surv, 53, 5, Article 109 (October 2020), 40 pages. DOI:<https://doi.org/10.1145/3395046>