INTERNATIONAL JOURNAL OF ADVANCE RESEARCH, IDEAS AND INNOVATIONS IN TECHNOLOGY

ISSN: 2454-132X
Impact Factor: 6.078
(Volume 7, Issue 3 - V7I3-2041)
Available online at: https://www.ijariit.com

A study on analyzing fake news through Neural Network Models

Ashwini Durgadas Kalpande
ashwinikalpande97@gmail.com
Sanmati Engineering College, Washim, Maharashtra

Sachin A. Vyawhare
s.vyawhare@sanmati.in
Sanmati Engineering College, Washim, Maharashtra

ABSTRACT

The social world is very dependable on real and fake news and to grasps publicity many unreal news was circulated. And to identify the fake news, it becomes necessary to find out the solution to this serious problem. Many people focus on to explore the false information rather than the true information. The overall increase in this unreal circulation makes the potential increase in the market. But it is also difficult to identify the difference between real news and fake news. So it become necessary to create some model o understands the nuisance in natural language by defining some network in the system. Furthermore the concept of NLP can be used to proceed further for extension. Previous methods have the intension for detection but it limits with the comparison art in between real and fake news because the source they referred was on digital media. To rectify this issue we try to implement the part using some model creation using neural network and once the neural network was designed successfully we will try to focus on NLP in future work to predict the fake news.

Keywords: *Deep Learning, Natural Language, Detection Models, Predict, Binary*

1. INTRODUCTION

Decorating the news so as to further more get the publicity and comments is called as FAKE NEWS. And circulation of that news was a part of misinterpretation and cause many blenders in the market flow and the traditional media used are like print, and television as well as non-traditional media channels like social media and various related applications. For damaging of a reputation of any firm, brand and any other recognize personality the fake news was generated. Many research's was conducted to find an importance of false news over true news and it was found that false news circulation was much more faster than true news. Nobody expecting to search for the source of such fake news and the misleading of such information was continued further. So detecting the fake news from an authenticate source was a challenge t build the system by using Artificial Intelligence or Neural network or Natural Language processing. Summarizing of the news was also the challenging way to detect the news as fake and real after comparing it with actual news. A unique way of identifying the news is by extracting the feature of sentence and comparing it with the dataset model become an important expect to predict the news as fake and real.

So we called it bas stances in the sentence and while evaluating it we build various dataset model to compare these instances and named it as fake and real. Also the classification of those features in terms of words phrases and punctuation required to identify the sentence.

Let take some example for classification
1. Misusing the Data *"Have a Beer, It's Good for Your Brain,"* reported Inc. But you should wait a minute before you grab a pint (or two). The study was done on mice — not people. And the amount of beer was the equivalent of 28 kegs in humans.
2. Imprecise and Sloppy *"1 in 5 CEOs are Psychopaths, Study Finds."* But the headline is wrong. The research was based on a survey of professionals in the supply chain industry, not CEOs.
3. "The corona vaccine is curing the patient" The news content is perfectly fine but the understanding part to the user is quite different. It publicly committed that the vaccine cure corona but if we see the actual fact in it says that vaccine only boost the immunity power, the curing vaccination are working different in nature. Many examples can be taken to rectify the exact understanding of the news but the main problem is how to find out the accuracy for it from an authentic source. The media of news circulation was not limited to traditional sources but also from non-traditional sources also.

Table 1. Stance labels in training dataset

Stance category	Percentage	Description
Agree	7.36%	Headline agrees with the claim made in the news article
Disagree	1.68%	Headline disagrees with the claim made in the news article
Discuss	17.82%	Headline discusses same topic as news article
Unrelated	73.13%	Headline does not discuss same topic as news article

© 2021, www.IJARIIT.com All Rights Reserved

2. INTRODUCTION TO NEURAL NETWORK ARCHITECTURES

Various networks can be implemented to design a model in order to predict the system for fake news and real news. Following are some network description that is available in reference for model designing:

2.1 Dense Neural Network (DNN)

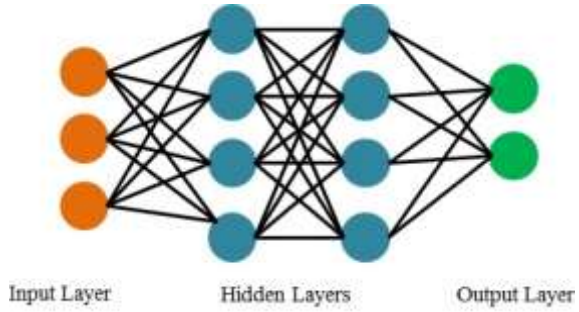

Fig. 1 Dense Neural Network

It allows us to pass the input as a sequence of words and comprises of input layer, an output layer and hidden layer.

2.1.1 Convolution Neural Network (CNN): It is similar to simple ordinary neural network the only difference is it was build with a set of artificial neurons called as nodes which have some weights and some bias pattern. It consists of an input and an output layer and multiple hidden layers. And this hidden layers includes the convolution, pooling and various normalization layers.

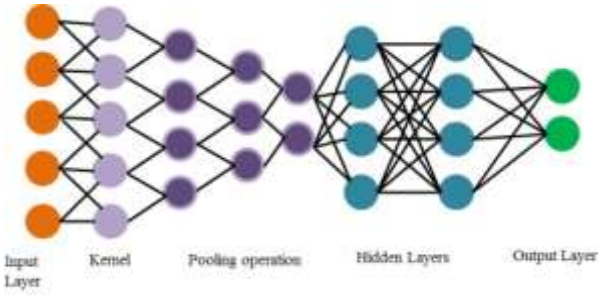

Fig. 2 Convolution Neural Network

2.1.2 Recurrent Neural Network (RNN): This layer plays an important role in understanding language processing part and consist of an input layer, output layer and various recurrent hidden layer with some memory gates.

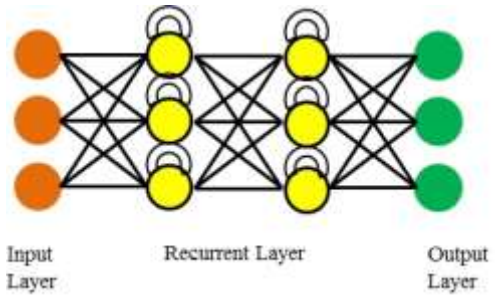

Fig.3 Recurrent Neural Network

2.1.3 Data Preprocessing: When we are extracting the feature of any sentence the processing of the data is very important since it includes various features which helps to extract the feature from a particular query sentence. The deep learning algorithm was implemented to perform the extraction operation from the sentence. Various method of extraction was presented further so as to understand the concept of feature classification.

2.1.4 Stop Word Removal: This process involves removing the most common words from the sentence sometime we called it as articles. Removal of stop word from the sentence does not affect the outcome of the result. It indirectly helps to reduce the time consuming part which was occurred during comparison procedure. Some common stop words are "and", "but", "an", "a", "from", "in", "of"", but", "an", "the". In coding part we use NLTK tool kit for processing.

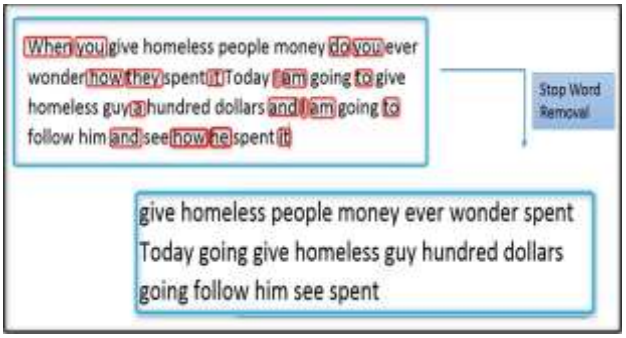

Fig. 4 Example for Stop Word removal

2.1.5 Punctuation Removal: Grammatical context in the sentence includes comma, fullstop exclamatory amrk, question and may does not required to predict the sentence as fake or real. So removing that part become easy to predict the sentence and it reduces the time processing part also.

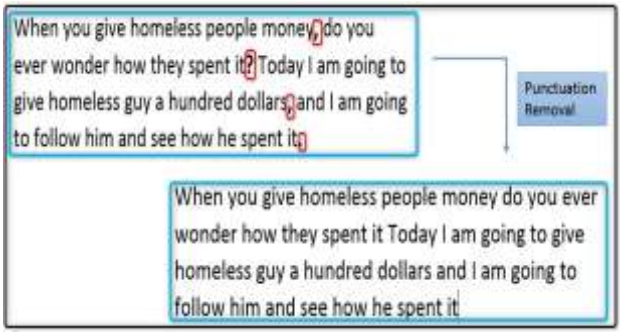

Fig. 5 Example for punctuation removal

3. PROPOSED METHODS

The sources also play a vital role of generation of news and those sources prediction is also an important parameter in fake news detection. Dataset on the other is easily available from various internet sources and due to which it become a challenging part to identify the valid fake source.

3.1 Static Search Implementation

In this search we try to implement some algorithm to finalized the model which will suit proper for detection of fake news depend on various dataset..

Step 1: Features in the sentence was extracted.
Step 2: Classifier was design from the extracted feature for model comparison.
Step 3: This classifier was compared with the dataset and them finalized to compare with input sentence.
Step 4: In overall this classification and algorithm the logistic regression was very confined and created the model which was providing the good result for identification
Step 6: And finally the model design was used to compare the probability or we can say the prediction of truth news.

© 2021, www.IJARIIT.com All Rights Reserved

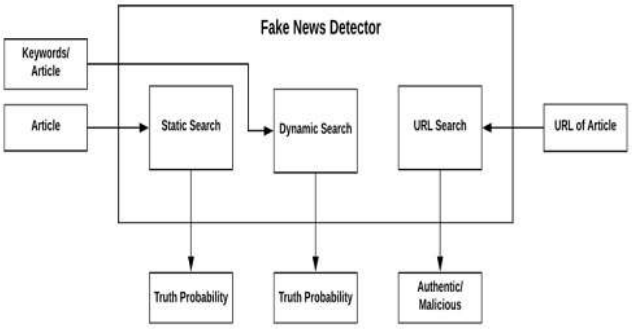

Fig. 6: System Design

4. CONCLUSION

So by creating such system the dependability on fake news was reduced to some extent and the news was published will be the real news. The identification procedure implemented in this section was providing the best solution to this serious issue which will help to minimize the circulation of fake news. Whereas the difficulty which was identifying in accordance with fake and real news was decreases and one can identify the source of the news also. The best method implemented for designing the model by virtue of which it become easy to identify and predict the new properly and will not effect for any branding, status issue or any privacy issue.

5. REFERENCES

[1] Kai Shu, Amy Sliva, Suhang Wang, Jiliang Tang, and Huan Liu, "Fake News Detection on Social Media: A Data Mining Perspective" arXiv:1708.01967v3 [cs.SI], 3 Sep 2017

[2] Kai Shu, Amy Sliva, Suhang Wang, Jiliang Tang, and Huan Liu, "Fake News Detection on Social Media: A Data Mining Perspective" arXiv:1708.01967v3 [cs.SI], 3 Sep 2017

[3] M. Granik and V. Mesyura, "Fake news detection using naive Bayes classifier," 2017 IEEE First Ukraine Conference on Electrical and Computer Engineering (UKRCON), Kiev 2017, pp. 900-903.

[4] Fake news websites. (n.d.) Wikipedia. [Online]. Available: https://en.wikipedia.org/wiki/Fake_news_website. Accessed Feb. 6, 2017 [5] Cade Metz. (2016, Dec. 16). The bittersweet sweepstakes to build an AI that destroys fake news.

[5] Conroy, N., Rubin, V. and Chen, Y. (2015). "Automatic deception detection: Methods for finding fake news" at Proceedings of the Association for Information Science and Technology, 52(1), pp.1-4.

[6] Markines, B., Cattuto, C., & Menczer, F. (2009, April). "Social spam detection". In Proceedings of the 5th International Workshop on Adversarial Information Retrieval on the Web (pp. 41-48)

[7] Rada Mihalcea , Carlo Strapparava, The lie detector: explorations in the automatic recognition of deceptive language, Proceedings of the ACL-IJCNLP

[8] Kushal Agarwalla, Shubham Nandan, Varun Anil Nair, D. Deva Hema, "Fake News Detection using Machine Learning and Natural Language Processing," International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-7, Issue-6, March 2019

[9] H. Gupta, M. S. Jamal, S. Madisetty and M. S. Desarkar, "A framework for real-time spam detection in Twitter," 2018 10th International Conference on Communication Systems & Networks (COMSNETS), Bengaluru, 2018, pp. 380-383 [11] M. L. Della Vedova, E. Tacchini, S. Moret, G. Ballarin, M. DiPierro and L. de Alfaro, "Automatic Online Fake News Detection Combining Content and Social Signals," 2018 22nd Conference of Open Innovations Association (FRUCT), Jyvaskyla, 2018, pp. 272- 279.

[10] C. Buntain and J. Golbeck, "Automatically Identifying Fake News in Popular Twitter Threads," 2017 IEEE International Conference on Smart Cloud (SmartCloud), New York, NY, 2017, pp. 208-215.

[11] S. B. Parikh and P. K. Atrey, "Media-Rich Fake News Detection: A Survey," 2018 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR), Miami, FL, 2018, pp. 436-441

[12] Scikit-Learn- Machine Learning In Python [15] Dataset- Fake News detection William Yang Wang. " liar, liar pants on _re": A new benchmark dataset for fake news detection. arXiv preprint arXiv:1705.00648, 2017.

[13] Titcomb, J., Carson, J.: www.telegraph.co.uk. Fake news: What exactly is it – and how can you spot it?

[14] Allcott, H., Gentzkow, M.: Social media and fake news in the 2016 election Technical report, National Bureau of Economic Research (2017)

[15] Langin, K.: http://www.sciencemag.org. Fake news spreads faster than true news on Twitter—thanks to people, not bots (2018)

[16] Wardle, C.: Fake News. It's complicated. First Draft New (2017). https://firstdraftnews.com/fake-news-complicated/

[17] Throne, J., Chen, M., Myrianthous, G., Pu, J., Wang, X., Vlachos, A.: 2017. Fake News Detection using Stacked Ensemble of Classifiers. In ACL.

[18] Davis, R., Proctor, C.: 2017. Fake News, Real Consequences: Recruiting Neural Networks for the Fight against Fake News. https://web.stanford.edu/class/cs224n/reports/2761239.pdf

[19] SemEval-2016: Semantic Evaluation Exercises International Workshop on Semantic Evaluation (SemEval-2016).

[20] Sobhani, P., Inkpen, D., Zhu, X.: A Dataset for Multi-Target Stance Detection. http://www.aclweb.org/anthology/E17-2088.

[21] Pang, B., Lee, L.: 2008. Opinion mining and sentiment analysis. http://www.cs.cornell.edu/home/llee/omsa/omsa.pdf

[22] Collobert, R., Weston, J., Bottou, L., Karlen, M., Kavukcuoglu, M., Kuksa, P.: 2011. Natural Language Processing (Almost) from Scratch. Journal of Machine Learning Research 12 (2011) 2493-2537.

[23] Johnson, J.: 2016. The Five Types of Fake News. https://www.huffingtonpost.com/dr-john- johnson/the-five-types-of-fake-ne_b_13609562.html

[24] As fake news spreads lies, more readers shrug at the truth. In https://www.nytimes.com/ 2016/12/06/us/ fake-news-partisan-republican-democrat.html.

© 2021, www.IJARIIT.com All Rights Reserved