IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 9, Issue 6, June 2021

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 7.542

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542 |

|| Volume 9, Issue 6, June 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0906027 |

Fake News Detection using Natural Language Processing and Deep Learning

Prof. B.B. Gite, Ajay Alure, Dnyaneshwar Karande

Department of Computer Engineering, Sinhgad Academy of Engineering, Kondhawa Bk, Pune, Maharashtra, India

ABSTRACT: The modern, ever-connected world by the aid of the Internet, has given stimulus to websites and platforms, publishing mass content and most of this, without any validation of sources. This gives growth to one of the topics of high research, Fake News Detection, which is studied to multiple aspects and approached with various solutions, which make an effort to solve this modern age problem to some extent. Though the existing methodologies have tackled the problem, there remains some scope for advancement. The following research is done to utilize that scope and provide a new, real-life, and more efficient system for Fake News Detection. The following paper proposes a combination of techniques from two disciplines of the Artificial Intelligence (AI) field viz. NLP and DL. The system can be said as based on exclusive algorithms such SBERT and GloVe, which are fine-tuned and used as encoders to convert text into vector for further processes. The system is stated to be having a combination strategy as it relies on outputs from two sub-systems for producing the final output. The final output expected from the following system is in the form of the predicted class of the input news title among True and Fake. Thus the system has a broader understanding of the presented problem, which in turn has provided more reliable performance than some of the existing systems for the same.

KEYWORDS: NLP, DL, AI, SBERT, GloVe, API

I. INTRODUCTION

Change being only constant thing in the world holds true in every field. This being quite evident in the shift that took place from people reading news from old and traditional news sources such as newspaper and printed Magazines to new platforms, prominently the Social media. With the increasing ease of getting connected to World Wide Web from nearly anywhere in the world, a greater number of people are consuming the news from web only. Though this advancement was beneficial in several ways, it also gave rise to many cons. Among many, Fake News topped the list. Fake News basically is the piece of information which is created by manipulating and changing the original news. This manipulation is done with the evil intention of giving a damage to any person, group or an institution. Fake News can has a great and adverse effect on the society as it can Damage the current government by destabilizing it, can spoil the Image of a person or could lead to the Crash of Wall Street in no time. Given the prevalence of this new phenomenon, “Fake news” was even named the word of the year by the Macquarie dictionary in 2016 [1]

Today, People are spending nearly 25 % of their time in their phone busy on social media. 350,000 tweets are sent per minute, indicating the involvement [2]. Hence if a News agency for instance alter any piece of news, same news will be read by large no of people at the same time, thus a single manipulation led to the transfer of a wrong information to large no people. Hence to keep check and control the spread of this manipulation, detection of such news became the need of time.

II. RELATED WORK

A piece of news article consists of 4 different things:

i. Source: Author or publisher of the news article.

ii. Headline: Title text describing the main topic of the article.

iii. Body Text: Main text elaborating the details of the news story.

iv. Media: Body content providing visual cues to frame the story.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542 |

|| Volume 9, Issue 6, June 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0906027 |

Note that the user is expected to enter the input as the News title only, in the text form. They can enter the input through the developed UI (user interface).

III. PROPOSED ALGORITHM

. The following research exploits advance technologies, for instance, Natural Language Processing (NLP) & Deep Learning (DL) which came into rescue as they can be used to build models which could be used to detect the Fake News.

A. NLP

Concepts focused entirely on learning the language presented and converting the text form of the input into machine usable format are included in NLP. NLP techniques help to find similarity between pair of words.

B. DL

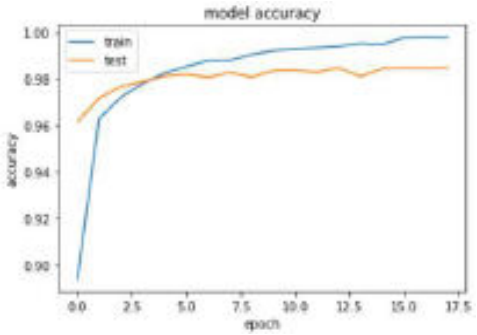
The non- linear relation between the independent variables and the dependent output is learned while training the model with Deep learning layers. Training a model with Deep leaning layer help to learn patter of occurrence of specific set of inputs and the resulting target class for that input.

- **Pseudo code**

Step 1: Generate all the possible routes.
Step 2: Calculate the TE_{node} for each node of each route using eq. (1).
Step 3: Check the below condition for each route till no route is available to transmit the packet.
if (RBE < = TE_{node})
Make the node into sleep mode.
else
Select all the routes which have active nodes
end
Step 4: Calculate the total transmission energy for all the selected routes using eq. (2).
Step 5: Select the energy efficient route on the basis of minimum total transmission energy of the route.
Step 6: Calculate the RBE for each node of the selected route using eq. (3).
Step 7: go to step 3.
Step 8: End.

IV. SIMULATION RESULTS

The DL model was trained on a dataset having data points covering 50,000 with distinct values of hyper-parameters. The dataset was split in 70:30 proportion as train dataset and test dataset. After training, the model yielded an accuracy of 98% on the test dataset as shown beneath. For more reliable evaluation we observed the Confusion Matrix of the predictions on the test dataset which symbolized an efficient performance by the DL model.



The NLP model results in the highest similarity value for the input news title and based on that it predicts the class of input news as either True or False. It's working shares similarity with unsupervised models. Still, its performance can be evaluated with peculiar matrices which will be the subsequent step to be done in this study.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542 |

|| Volume 9, Issue 6, June 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0906027 |

V. CONCLUSION AND FUTURE WORK

With the rise in popularity of social media, more people are turning to it for news instead of conventional news sources. However, social media has been used to spread false information, which has serious consequences on both individual users and culture as a whole. We concluded from our research that Natural Language Processing (NLP) & Deep Learning (DL) can be used to build models which could be used to detect the Fake News. Thus, we used these disciplines of artificial intelligence with some highly efficient techniques and tuned them to complete the various aspects of the system.

After covering various unsatisfactory points from the previous methodology there is a future scope in the proposed methodology too. By leading some deep research, the less efficient time consumption can be made optimal. Also, the obstacle in the prediction of input news title having specific proper nouns can be overcome by training with the records having the specific proper nouns. A dataset holding a random portion of the True news present in the True news dataset with each of its corresponding Fake news, is being arranged. Consequently, this appropriate dataset will be adopted in the overall evaluation of the proposed system.

REFERENCES

1. Kai Shu, Amy Sliva, Suhang Wang, Jiliang Tang, and Huan Liu. 2017. Fake News Detection on Social Media: A Data Mining Perspective. SIGKDD Explor. News l. 19, 1(June 2017), DOI:https://doi.org/10.1145/3137597.3137600
2. https://www.dsayce.com
3. Iftikhar Ahmad, Muhammad Yousaf, SuhailYousaf, Muhammad Ovais Ahmad, "Fake NewsDetection Using Machine Learning Ensemble Methods", Complexity, vol. https://doi.org/10.1155/2020/8885861
4. Thota, Aswini; Tilak, Priyanka; Ahluwalia, Simrat; and Lohia, Nibrat (2018) "Fake News Detection: A Deep Learning Approach," SMU Data Science Review: Vol. 1: No. 3, Article 10. Available at: https://scholar.smu.edu/datasciencereview/vol1/iss3/10
5. Gaurav Bhatt, Aman Sharma, Shivam Sharma, AnkushNagpal, Balasubramanian Raman, and Ankush Mittal. 2017. "On the Benefit of Combining Neural, Statistical and External Features for Fake News Identification," arXiv preprint arXiv:1712.03935 (2017)
6. Granik, Mykhailo, and VolodymyrMesyura. "Fake news detection using naive Bayes classifier." Electrical and Computer Engineering (UKRCON), 2017 IEEE First Ukraine Conference on. IEEE, 2017.
7. Nils Reimers, IrynaGurevych. 2019. "Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks," arXiv:1908.10084 (2019)
8. HrishikeshTelang, Shreya More, Yatri Modi, Lakshmi Kurup, “An empirical Analysis of Classification Models for Detection of Fake News Articles ”, © 2019 IEEE
9. Kyeong-hwan , Chang-sung Jeong, “Fake News Detection System using Article Abstraction” , ©2019 IEEE
10. Nils Reimers, IrynaGurevych (2019), “Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks”, arXiv:1908.10084v1 ©2019 IEEE
11. https://medium.com/analytics-vidhya/gloveembeddings-to-detect-fake-news-424accf3fa4c

INNO SPACE
SJIF Scientific Journal Impact Factor
Impact Factor: 7.542

doi cross ref

ISSN INTERNATIONAL STANDARD SERIAL NUMBER INDIA

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details