A Review : Fake News Analysis and Detection

Mridula Arvind Halgekar*1, Prof. Vidya Kulkarni2

1MTech Scholar, Department of Computer Science, KLS Gogte Institute of Technology, Belagavi, Karnataka, India

2Professor Department of Computer Science, KLS Gogte Institute of Technology, Belagavi, Karnataka, India

ABSTRACT

With the growing world in terms of technology and population, the growth of technological use by the population has also increased. The technology has become a part of every human being's life. It is not just a part of his professional life but also a part of his personal life. There are so many things happening in the world that keeps the world changing. To grow along with this growing world, we need to keep ourselves updated. Media plays an important role in keeping the population updated. The world is kept updated irrespective of the location of the population reading the news and the location of the incident occurring. Fake news is the biggest drawback in this process. We believe what we see and what we read as it the only way to keep ourselves updated. So Fake news hampers the population and may result in unexpected incidents. So it is the need of the hour to understand the difference between real and fake news. This project is for fake news analysis and detection. A dataset of news is considered, pre processing is done and then the fake news and real news are predicted using random forest and xgboost algorithms.

Keywords : Fake News, Random Forest, Xgboost, Analysis, Detection.

I. INTRODUCTION

To keep ourselves updated with the happenings in the world, news plays an important role. There are various authors, publishers and sites that uodate news and information that help a lot of public to be updated with the existing conditions or theongoing conditions of the world. But if this news that is read by a mass of public in general is fake it posses a number of problems. Fake news has to be detected and avoided to avoid a lot of problems. The proposed Fake news Detector is a software that shows if a news that is presented is real or fake. As the society in general wants to be updated with the surrounding environments and the events happening at each part of the world, the media plays an important role in satisfying the society with the news. But unfortunately sometimes the media tends to spread news that is not real or fake. This misleads a mass of society and hence we face some very disastrous consequences.The proposed model helps to discriminate fake news from real news. To do this the basic rule of comparison is used. We will use the dataset of BBC news. This dataset includes genres like business, political, entertainment, technical and sports. Then the predictions are done for news to be real or fake.The main aim of the project is to determine if a project is real or fake based on some comparison and classification of the fake news. Some pre processing is done on the dataset so that the prediction becomes easy. Algorithms like random forest and xgboost are used. These algorithms helped predict the results with good accuracy.

Following is the survey done to understand some of the worked on projects of the related topic.

The paper continues with Literature survey, proposed model, conclusion and references.

II. LITERATURE

Fake news can be tracked by following some trails. These trails can be the tile or the text. The publishers that publish fake news can also be kept a check on. But this proposed system will use a dataset given, pre process it and then using the algorithms predict if the data or the news is fake or real. Some of the papers gone through and their findings are as follows. The paper published by Jaiwei Zhang, Bowen Dong, Philip S. Yu: FakeDetector: Effective Fake news Detection with deep diffusive neural networks discusses the problems of fake news and the way to tackle them. Here they have considered articles, creator and subjects to analyse the authentication of the news. The datasets that are created are done by taking data from tweets and fast check articles. Then the dataset is studied in detail. The article credibility and the author credibility is checked. Also the subject credibility and the publisher credibility is taken into consideration. Then using deep diffusive neural networks the model that they have proposed is followed. This study gives a wider perspective of how fake news can be handled.[1]In the paper done by SajjadAhmed, Knut Hinkelmann and Flavio Corradini: Combining machine learning in knowledge engineering to detect fake news in social network, they have studied the news that is put up in the social networking sites. They have given the basic idea of what exactly is fake news, how it affects certain areas and some ways to detect these fake news. Here they have proposed a number of models that can be used in detection of fake news. Each model has a unique way of gathering the data and then analysing it. The work of this paper helps in understanding that machine learning can be used in a number of ways to achieve the same goal.[2] The paper published by Xinyi Zhau and Reza Zafarani: Fake News: A survey of Research,

Detection Methods and opportunities, they have given a detailed view of what exactly is fake news. The life cycle of fake news. The creation of fake news and the propagation of fake news is also explained. They have put forth some examples of fake news in real life. The reasons behind generation of fake news and the effects of them are also discussed deeply. They have proposed some solution that include machine learning techniques to understand the fake news patterns and predict them. This paper helps in the deep analysis of the cause of the fake news and expands the criterions that should be considered to detect and predict fake news.[3]. In the paper published by Kai Shu, Amy Sliva, Suhang wong, Jiliang Tana and Huan Liu: Fake news detection on social media, A Data mining perspective, they have discussed some definitions of fake news and their propagation techniques. They have also discussed some currently used fake news detection systems that help in the detection of fake news. They have considered style, propagation, content , knowledge and stance based to predict the fake news. From this paper it was easier to understand the real problems faced by propagation of fake news and also some of the recent used methods to detect and analyse the fake news.[4]. The paper done by Nicole O'Brien: Machine Learning for detection of Fake news, discusses some of the basic causes and problems of fake news. The paper proposes a model for detecting the fake news using machine learning. She uses different algorithms and techniques to detect fake news. This paper relates to my proposed model in a lot of ways. This helps in widening the scope of the proposed model and also helps in understanding the machine learning technique in a wider perspective.[5]. The paper proposed by Nicole, Sophia, Georgia and Xavier: The language of fake news: opening the black box of deep learning based detector, discusses some not known techniques to understand the detection of fake news and the ways to perform analysis on them. They have put forth some previously carried out works on the same sector and they have given a detailed overview of their model that could help

detect fakenews. They have considered the sources, author and some other causes that generate the fake new and they have used these as a criterion to detect fake news.[6]. Julio C. S. Reis, Andr_e Correia, Fabr_ıcio Murai, Adriano Veloso, and Fabr_ıcio Benevenuto proposed a paper that discusses about fake news detection methods that can be used. They have discussed some features of a fake news detector that should be considered in forming a fake news detector. They have also discussed dome challenges that are caused by the fake news. This paper helps in understanding the fake news generation, prevention and model creation of a fake news detector.[7]. The paper proposed by Hadeer Ahmed, Issa Traore, and Sherif Saad: Detection of Online Fake News Using N-GramAnalysis and Machine Learning Techniques, explains how fake news can be detected using natural language processing. This paper gives a wider and a different view of how the same project can be handled using different approaches.[8]

These papers were studied to know different methods and techniques that were used in creating the same project.

III. PROPOSED MODEL

The model proposed here is a fake news analysing and a fake news detector model. For this, dataset from BBC is collected and then pre processing is done of the dataset. By this we merge two datasets and then create some independent columns for easier predictions. Random forest and Xgboost algorithms are used for the prediction. This model will be able to predict the fake news. The dataset contains news data belonging to different genres.

IV. CONCLUSION

The paper presented a survey on the different methods that are been used by different researchers to predict fake news. This has helped to understand the different methods and techniques that can be used to predict fake news. the work of so many researchers has helped in understanding the reasons of fake news, the creation, propagation and the effects of fake new. It has also helped in understanding the different criterions that have to be considered in creation a fake news detection model.

V. REFERENCES

[1]. Jaiwei Zhang, Bowen Dong, Philip S. Yu, "FakeDetector: Effective Fake news Detection with deep diffusive neural networks", 10 Aug 2019.

[2]. SajjadAhmed, Knut Hinkelmann and Flavio Corradini, "Combining machine learning in knowledge engineering to detect fake news in social network", Department of Computer Science, University of Camerino, Italy.

[3]. Xinyi Zhau and Reza Zafarani, "Fake News: A survey of Research, Detection Methods and opportunities", 2 Dec 2018.

[4]. Kai Shu, Amy Sliva, Suhang wong, Jiliang Tana and Huan Liu, "Fake news detection on social media, A Data mining perspective", Computer Science & Engineering, Arizona State University, Tempe, AZ, USA.

[5]. Nicole O'Brien, "Machine Learning for detection of Fake news." Massachusetts Institute of Technology 2018.

[6]. Nicole, Sophia, Georgia and Xavier, "The Language of Fake News: Opening the Black-Box of Deep Learning Based Detectors", 32nd Conference on Neural Information Processing Systems (NIPS 2018), Montréal, Canada.

[7]. Julio C. S. Reis, Andr_e Correia, Fabr_ıcio Murai, Adriano Veloso, and Fabr_ıcio Benevenuto, "Supervised learning of fake news".

[8]. Hadeer Ahmed, Issa Traore, and Sherif Saad, "Detection of Online Fake News Using N-GramAnalysis and Machine Learning Techniques", ECE Department, University of Victoria, Victoria, BC, Canada