COVID-19 Fake News Detection Using Naïve Bayes Classifier

O. Akande[1], G. Egwuonwu[2], and W. Ajayi[3]
[1,3]Member, Computer Professional Association of Nigeria, Babcock University, Nigeria
[2]Graduate, Babcock University, Nigeria

***Abstract*—** This study aimed at detecting fake information relating to COVID-19 using Naïve Bayes. The advent of social media which is made available through the internet provides platforms on which news can be disseminated and reach a large number of audiences in seconds. This opportunity comes with its challenges, of which one major one is the possibility of spreading fake news quickly. Detection of fake news is a binary classification problem that is handled with machine learning techniques that learn on their own. Naive Bayes is one of the well-known machine learning classifiers that is used in resolving text classification problems. This algorithm is applicable regardless of the number of inputs. It was used in this work to build a model which can distinguish fake news from real ones. For the moderately-sized COVID-19 dataset, an accuracy of 96.7 percent was achieved. With a very large dataset Multimodal, Naïve Bayes will perform better.

***Keywords*—** Algorithm, COVID-19, Fake News, Machine Learning, Naïve Bayes.

I. INTRODUCTION

The authors in [1] 2006, stated that in any human information there will be untrue elements. The advent of social media made available through the internet provides platforms on which news can be disseminated and reach a large number of audiences in seconds. This opportunity comes with its challenges, of which one major one is the possibility of spreading fake news quickly. This is more dangerous when it comes to health and security issues. This paper focuses on health issues, specifically COVID-19 which held the whole world at ransomed from 2019 till date.

Each time someone reads a news report, he or she is susceptible to coming across fake news. Artificial intelligence (AI), machine learning (ML), rule-based techniques, and statistical techniques are being used by researchers to detect fake news so that recipients will not continue to spread it. These text classification algorithms are used in sentiment analysis, topic classification, fake news detection etc. There are many AI, ML, statistical and rule-based techniques that have been developed to do this.

The rule-based technique involved a lot of domain knowledge from which pattern can be deduced so that rules or inference engines may infer whether a word is in one category or the other. Rule-based use precedent and antecedent to form rules [2]. For beginners who lacked adequate knowledge about a domain, this technique may be difficult to use. ML techniques for natural language processing are naïve Bayes and support vector machine (SVM). SVM is an algorithm that can classify data, even complex data into two dimensional (x-axis, y-axis) or multidimensional (x-axis, y-axis, z-axis) planes with the use of hyperplane. It can be trained on small datasets but produces a better result on complex datasets i.e. when datasets have both homogeneous and heterogeneous attributes.

Statistical techniques are not algorithm-based techniques but are mathematically based. The information must be preprocessed to be concise thus feature extraction techniques such as Average Neighbourhood Margin Margimization, and Principal component analysis (PCA) to achieve good classification. All these techniques and others like Biased Discriminant Analysis cannot handle non-linear or unstructured datasets and cannot work or scale well on large datasets [3].

In this current work ML technique called Naïve Bayes is used. Naïve Bayes variant that is used is Multimodal because it works well with text datasets unlike the traditional naïve Bayes algorithm [3]. It is used here because it is a common algorithm used for text classification and is easy to use [4].

II. RELATED WORKS

The author in [5] found out that deep learning performs better on datasets, in their case COVID-19 dataset because of dataset inconsistencies which baseline models cannot handle. Such inconsistencies are because some news information may not be totally true or false. They proposed that to overcome this hybrid model should be used where the initial model will only learn important parameters from subclasses and the final model will learn from a combination of subclasses parameters, this will improve the accuracy of the model. Koirala work used logistic regression, embedding with dense layer followed by embedding layer with LSTM

layer and then bidirectional LSTM. The author concluded that the deep learning technique performs better than baseline techniques.

In [6], the TUDublin team set up a challenge to develop a model that will detect fake and real news about COVID-19. They used the ensembles method using models namely; Support Vector Machine (SVM), a combination of Logistic Regression and Naïve Bayes, Logistic Regression, Bidirectional Long Short-Term Memory, and Naïve Bayes.

The work is divided into two stages: preprocessing and model building. For the preprocessing, term-frequency inverse-document frequency (TFDIF) for machine learning models, PorterStemmer for neural network models and finally for each model the words are converted into lower cases. In the first stage of modelling, each model was implemented separately and then two ensembles were created one for classic machine learning and the other for neural network models. Finally, the two ensembles were combined as a single ensemble. It was found that ensemble models especially the ones involving machine learning and neural network performs better than single models, while ensemble with purely classical machine learning model or neural network performs lesser than the ensemble of machine learning and neural network.

The authors in [7] built a model for detecting fake news on social media. The machine learning algorithms used were support vector machine (SVM) and Naïve Bayes. To build the model news collected were preprocessed and feature extraction was done by pretrained algorithms for text extraction such as CounterVectorizer(CV). The dataset was split into train and test. SVM and Naïve Bayes worked on the training dataset to build a model. SVM uses hyperplane for binary division of datapoints while multimodal naïve Bayes was used to check if the news is fake or real.

The authors in [8] represented the problem of detecting fake news as a binary classification problem. The paper used the relationship between words (i.e., word meaning and the context where it appears in a text) to detect whether a word is fake or not. When the relationship is positive it means fake but if negative it implies real. Different pretrained models which can convert text to numbers were used.

The pre-trained algorithms used were TFDIF (term frequency-inverse document frequency), CountVectorizer (CV) and Word2Vec (W2V). CV was found better. The output of each pre-trained algorithm was passed into five neural network algorithms for actual classification into fake and real. Out of the five algorithms used artificial neural network (ANN) and long-short term memory (LSTM) performed better. It was also found that LSTM with CV performed much better.

The authors in [9] used the transformer ensemble method to build a model for covid-19 fake news detection. Ensemble models were BERT, ALBERT and XLNET. Firstly, Natural language processing (NLP) was used to preprocess the text to become tokenized, then fake news detection models were built using traditional NLP models like linear regression, support vector machines, passive-aggressive classifiers. After this, the models were built using deep learning techniques such as LSTM, BiLSTM with attention, Convolutional Neural Network (CNN) and lastly CNN with BILSTM. Lastly, transformer models ensembled were; Bidirectional Encoder Representations from Transformers (BERT), enhanced version of BERT (XLNet), and A lite BERT (ALBERT). The transformer ensembled performs better than all others.

III. METHODOLOGY

COVID19 Fake News Dataset NLP Kaggle here which includes three columns and 2140 rows. The first thing is to deal with missing data. Here columns with missing data are eliminated. Then large sentences are broken down into smaller chunks through the process of Tokenization. Tokenization in Python is defined to be the splitting up of a larger body of text into smaller lines, words or even creating words for another language. From the various tokenization functions derived from the natural language toolkit (nltk) module, word tokenization is adopted.

Stemming was done to reduce the words to their root form. Then the dataset was split into the train (80% of the dataset) and test dataset (20%of the dataset). Stop words, punctuations and special characters are removed from the two datasets because they do not make much meaning. Feature extraction was done to remove the dimensionality in terms of attributes in the column of the dataset to a manageable one.

The next step is the vectorization of the text into vectors of numbers. This is done using CountVectorizer. This step is necessary because the dataset must be formatted into a form that a machine learning algorithm or deep learning algorithm can use.

The format is mostly in number form. Multimodal Naïve Bayes algorithm was used to build the model for detecting fake news about Covid-19 using the training dataset. Figure 1 following is the methodology.

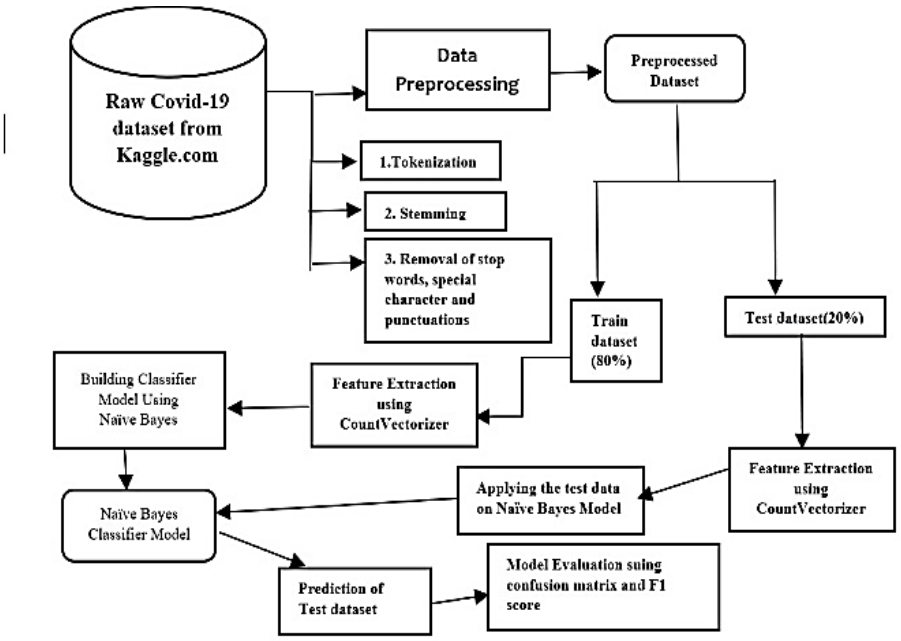

Figure 1: *Methodology*

IV. IMPLEMENTATION

The model is then tested on the test dataset and it was found to have 96.7% accuracy and a bit of overfitting. A piece of news was inputted into the model and it was found to be fake. F1 score was used for the accuracy. F1 score on 1 is 0.98 and on 0 is 0.5.

V. FUTURE WORKS

Further study will be done to accommodate news or more specifically COVID-19 news that were in picture, video and audio forms. In addition, other forms of Machine Learning classifiers were not applied in the building of the model, in the future we intend to use other machine learning algorithms and measure their respective level of accuracy. Also, we intend to build this model to a production level such that it can be available as a mobile or web application, and can also work offline.

VI. CONCLUSION

The research methodology is clearly shown in Fig.1 this will help anyone who wants to repeat this work. The naïve Bayes model built can detect if the news is fake or real. Detection of fake news is highly needed in this age of social media.

REFERENCES

[1] S. Prachusilpa, A. Oumtanee, A. Satiman (2006). A study of the dissemination of health information via the internet. (PMID, Ed.) *Stud Health Technol Inform*.

[2] MonkeyLearn. (2021, October 18). *Text classification*. Retrieved from MonkeyLearn: https://monkeylearn.com/text-classification/

[3] M. Thangaraj and M. Sivakami. (2018). TEXT CLASSIFICATION TECHNIQUES: A LITERATURE REVIEW. *International Journal Information, Knowledge and Management*, 117-135. Retrieved from https://www.ijikm.org/Volume13/IJIKMv13p117-135Thangaraj3803.pdf

[4] K. Vasa (2016). Text Classification through Statistical and Machine Learning Methods: A Survey. *International Journal of Engineering Development and Research*, 655-658.

[5] A. Koirala (2020). COVID-19 Fake News Classification with Deep Learning. 1-6.

[6] S. A. Elena (2021). TUDublin team at Constraint@AAAI2021 -COVID19 Fake News Detection. *viewed at arxiv.org*, 1-8. Retrieved 2021, from https://arxiv.org/pdf/2101.05701.pdf

[7] A. Jain, A. Shakya, H. Khatter and A. K. Gupta. (2019). A smart System for Fake News Detection Using Machine Learning. *2019 International Conference on Issues and Challenges in Intelligent Computing Techniques (ICICT)* (pp. 1-4). Ghaziabad, India: IEEE. doi:10.1109/ICICT46931.2019.8977659

[8] S. Vijayaraghavan, Y. Wang, Z. Guo, J. Voong, W. Xu, A. Nasseri, J. Cai, L. Li, K. Vuong, E. Wadhwa. (2020). Fake News Detection with Different Models. *CoRR, abs-2003-04978*. Retrieved from https://arxiv.org/abs/2003.04978

[9] Sunil Gundapu and Radhika Mamidi. (2021). Transformer based Automatic COVID-19 Fake News Detection System. *viewed at arxiv.org*, 1-12. Retrieved from https://arxiv.org/pdf/2101.00180.pdf