Fake news prediction using Machine Learning

Pavan Gandavarapu
700741173
dept.Computer Science
University of Central Missouri
pxg11730@ucmo.edu

Sushmitha Virri
700742289
dept.Computer Science
University of Central Missouri
sxv22890@ucmo.edu

Divya Tejaswi Ejjurothu
700745894
dept.Computer Science
University of Central Missouri
dxe58940@ucmo.edu

Vishnu Ponugoti
700744508
dept.Computer Science
University of Central Missouri
vxp45080@ucmo.edu

Abstract—**The spread of fake news on social media and other platforms is a serious worry because it has the potential to have a negative influence on society and the country. There has already been a lot of research to identify it. The research includes creating a supervised machine learning algorithm, that can classify fake news as true or false for English language texts by using tools like Python, Scikit- Learn, and NLP for textual analysis. This paper analyses the research on fake news detection and explores the best traditional machine learning models. We suggest using the Python sci-kit-learn module to do tokenization and feature extraction of this procedure, which will result in feature extraction and vectorization. In this paper, we proposed two machine learning algorithms decision tree and Multinomial Naive Bayes algorithms to classify the news articles. To tokenize the text TFIDF vectorizer is used. We have conducted the experimental analysis of the models and found that the Naive Bayes classifier outperformed the Decision tree model in terms of accuracy. The accuracy of the Naive Bayes classifier is 92 percent whereas the decision tree accuracy is 88 percent. Furthermore, the performance of the model is verified using precision, recall and F1 score by publishing the classification report. To conduct experimental analysis, we have used the publicly available news dataset that is collected from the Kaggle repository.**

[1] *Index Terms*—**Multinomial Naive Bayes, Decision tree, TFIDF (Term Frequency - Inverse Document Frequency), tokenization, vectorization and classification report.**

I. INTRODUCTION

Nowadays information is everywhere but the authenticity of the data is lacking. According to the study, many platforms produce 60-80 percent of fake content. These metrics are concerning, as this data is produced due to a variety of reasons. This content in most cases is produced intentionally due to political pressure or some other reasons. Finding or filtering the content is a challenging task. Domains like machine learning and NLP provide solutions to these problems.

In the present time, the internet is the primary resource for information and access to the information is at our fingertips. However, the data present on the internet is not legitimate, most of the data is tabloid or fake content. Though humans understand the nature of the news after spending some

time on a particular news leads to wastage of time and misguides the users if they feel the news is real. So, instead of human efforts, NLP tools provide the solutions to this problem by understanding the nature of the text. These can classify the text as fake or real. Hybrid machine learning techniques like embedded methods perform well in this task. Fake news prediction applies to different industries. In this paper, fake news classification is applied to the Covid-19 dataset to know the percentage of false and real information. For the project, 4200 records related to Covid-19 are used to analyze the information category into fake and real. To classify the information Naive Bayes classifier, Support Vector Machine, and Passive-aggressive classifiers are implemented. One of the platforms that share information with a huge amount of people is Facebook. In this project, posts or news from Facebook are collected and the k-means clustering algorithm is applied to group the posts in the first instance where the news is clustered into group-I category and in the second instance after applying PCA (principal component analysis) we cluster the news into group-II. Together PCA and K-means cluster performed well and achieved good results. PCA gave a 95 percent of the confidence interval.

Online platforms and newsletters are responsible for the spread of fake information. In this paper, a simple logistic regression classifier is used to classify the fake and counterfeit information from the given input. With this simple model we achieved 80 percent of the accuracy. This model outperforms the hybrid models of machine learning in terms of overall accuracy.

In this project, we are implementing multiple machine learning algorithms to classify the news into real and fake. The project life cycle follows collecting the dataset, cleaning the dataset using regular expression modules, collecting vocabulary; this is achieved by count vectorizer and TFIDF vectorizer. After this process multiple machine learning models are applied to classify the news.

In this project, we can explore the techniques using the scikit-learn library. For the algorithm implementation we use scikit- learn library. For the task of preprocessing, we use

---

[1] https://github.com/PavanGandavarapu/ML_FINAL_PROJECT

the regex module in Python. For creating vocabulary, we used TF-IDF and count vectorizer from the scikit-learn library. For data visualizations seaborn and matplotlib libraries are used.

The problem type is a binary classification. The evaluation metrics in this project are accuracy, precision, F1 score, and recall. Accuracy is the default performance measure metric that gives the average of positive prediction and negative prediction scores. Whereas the precision gives the quality of the positive predictions, On the contrary recall gives the quality of negative predictions. On the other hand, F1 score gives the harmonic meaning of precision and recall. Accuracy is calculated on the test dataset using the accuracy function. The other parameters are published using classification reports.

II. MOTIVATION

In the past, if we look at the quality of the news there were only a handful of news creators globally. Though the news content is mundane at the same time they were useful to mankind to live a life. With the evolution of the internet, the perspective changed and there are plenty of sources even surpassing the required amount. In addition to this, it is a difficult question to determine which source is trustworthy. One of the challenges in content creation is spreading fake news through images, video, and audio formats on large platforms like Facebook and other social networking sites. The purpose of this task is to defame the person or a particular group. Sometimes this act is committed as political propaganda.

III. OBJECTIVES

The spread of fake news is executed through highly sophisticated bots which spread the news massively on social network platforms. Existing methods to tackle these problems are manual fact-checking, in this process we check the context of the news manually before publishing. The downside of manual fact checking is it is a slow process and cannot handle a huge amount of data at a time. Hence detecting the fake news is a challenging task since there is subtle difference between fake and real news in some cases which requires linguistic expertise. This is where the automation of this process comes into play. Artificial intelligence and NLP techniques solve these problems.

The main objectives of this project are:

- Applying TF IDF text vectorizer in the data preparation step
- Predicting the fake or legitimate news from the given input

The main features of this project are:

- Applying Term frequency vectorizer to overcome the challenges of a Countvectorizer.
- Count vectorizer gives the number of occurrences of the words in the document. If the sentences contain the stop words and other contexts of words automatically the weightage is given to the irrelevant words. Whereas the TFIDF vectorizer gives importance equally to the less frequent words and frequent words in the corpus.

IV. RELATED WORK

The main objective of the project is to classify the nature of the news as legitimate or fake. In recent years there has been a lot of fake news circulated on social media. The intention of this act could be nonchalant behaviors of the writers, targeting to defame the opposite person or the reasons can be any. There is a need to find the difference between the real and fake news as the reader needs to know the legitimacy of the fact that he reads. With this intention, we implement various machine learning classifiers to classify the news into legitimate or fake. This paper considered a dataset that contains various categories of news such as political and business news. There are 25000 articles in each category for a total of 50000 news articles. To classify the news, the Decision tree and Random Forest classifiers are deployed. After the experimental analysis, the accuracy score of the Random Forest model is 96 percent. In conclusion, we can say that Random Forest outperforms the Decision Tree model.[3].

With the expansion of social media, the authenticity of the news drastically reduced. Daily, there are many fake news circulating online. The dataset is collected from various sources. After scraping the web for news articles, the dataset is cleaned using various preprocessing techniques. In general, the articles contain various categories of news like political, business, and other categories. It is very difficult to analyze the multimodal categories of the news. But in this paper, a novel approach is proposed to categorize the news into fake news and real news. The framework used here was the History News Environment Perception framework.

Fake news Detection is the most challenging part of the digital era. The texts contain various jargons and dialects to predict the legitimacy of the news. The paper proposed probability methods to classify the categories of news [20]. Recent studies show that the accuracy of the model will be improved when we use non-linear dimensionality reduction methods on complex datasets. K-means clustering is used to categorize the legitimacy of the news. In this method, PCA (Principal Component Analysis) was used as the non-linear dimensionality reduction method. Generally, the dataset is complex. To reduce the complexity and dimensionality PCA uses the linear combination of the existing features without losing the information. The combination of PCA and K-means clustering proved to be the best to classify or predict the polarity of the news articles [5] [14].

Though we have multiple models for predicting the sentiments of the news articles. Still, there is room for improvement in terms of accuracy and other dimensions as well. To improve the accuracy, deep learning models are deployed. The deep learning models include GRU, CNN, and GCN models. These models focus on the semantics of the texts or articles instead of the accuracy parameter. In future work, focus is directed on the different approaches to compare the results of the exist-

ing methods like semi-supervised learning and unsupervised learning [9].

Existing system classification of fake news articles predicted the legitimacy of the news of different categories in general business or entertainment. In this study, we are proposing a state- of- the- art method to classify the legitimacy of the news during the outbreak of Covid-19. During the lockdown, there is an increased use of social media. With this world received a lot of fake news and this spread of tabloid news resulted in panic among the people. To curb these kinds of situations machine learning algorithms are proposed to predict the legitimacy of the news in the tweets collected from twitter [19].

In this digital era, we are consuming the news from various sources. The problem is there are no authorizations to these sources and a lot of tabloid news crops up during the election season. People are falling prey to this tabloid news. These kinds of tasks can damage the reputation of the institutions. To tackle this problem, hybrid machine learning methods are implemented to classify the legitimacy of the news. To build the hybrid classifier different machine learning classifiers are used and the results are obtained. In the results section an average voting of the results is taken [7].

With the emergence of fake news on the social media platform there is a lot of research going on in this section. Existing methods provide the classification ability but if we increase the complexity of the data they do not respond properly, and we get poor results out of them. To improve the results, we implement the Neural network architectures. In this study, LSTM i.e. Long Short Term Memory architecture is implemented to deal with the above problems. The performance of the model is compared with the machine learning models like SVM, and Naive Bayes classifiers. In the comparative analysis it is observed that the accuracy of the model LSTM is better than the other models in the existing system, and in the future, this model can be implemented to classify the realtime data samples [8].

In the past people read or gained information from the newspaper. Those days are gone with the improvement of the digital era. People are more dependent on electronic information than physical sources like books and newspapers. On these platforms, people openly share their views with no barriers or proof, and this shared information is majorly a personal opinion. Sometimes people share the information as a researched topic so we cannot judge the information's authenticity. The major challenge is finding a reliable source of information. In this paper, a logistic regression model is proposed to label the news articles as fake or real news. The model achieved 80 percent accuracy on the test dataset. There is a scope for improvement in the model tuning and future work proposes a tuned model to achieve better results [1] [2].

There is a plethora of publishers online to produce the news. But there is no scoring system to define the credibility of the news articles they produce. In this paper a model is proposed to score the publishers. It is a credibility rating system. This is achieved through two class boosted ensemble models and two class neural network models. For the credibility prediction an ensemble model is used. And classification of the news as legitimate or fake is done using a simple machine learning classifier. From the experimental analysis it is concluded that the ensemble method achieved satisfactory results in scoring the systems.[15] [18].

Fake news on social media spreads or flows like anything. There are many challenges in the digital world while consuming the news. For example, during the Covid-19 pandemic people are bombarded with fake information every day. In fact, the research showed that there is circulation of fake news every 30 mins on various platforms like WhatsApp, Facebook, and Twitter. The information is related to the deaths of the Covid-19 affected people, what to consume and what to avoid during the pandemic. In this study, a model is proposed to predict the nature of the data by evaluating the stiffness Index. The stiffness Index is calculated using the speed of the news spread in mins [11].

The study of this paper includes conducting an experimental analysis on the Covid-19 data. The data is collected from various websites using scraping methods. The collected data contains 4200 records of the data, and each record is labeled as fake or real. An exploratory data analysis is done to understand the semantics of the data. Since the dataset is small 3 compatible algorithms are implemented SVM (Support Vector Machine), Naive Bayes Classifier, and Passive aggressive classifier. By conducting the experimental analysis, the performance of each algorithm is compared. [16] [12].

The existing methods of classification are majorly implemented for global language English. The language semantics differ for one language to the other. So, there is no unified method to classify the news. Majorly people consume the news in local languages though English is a global language. In this paper a fake news prediction mechanism is proposed for classifying fake news in Sinhala language. To perform this task a hybrid classifier called Passive Aggressive classifier is used. To evaluate the performance of the model, accuracy parameters are used. To better understand the performance of the model precision, recall and F1 score are calculated. [10] [6].

The existing systems propose classification methods to differentiate the fake or real based on the patterns and vocabulary. This method is not effective. As per recent studies there are flexible methods to classify fake or real news. One of the effective methods to classify fake or real news is finding through the credibility of the publishers. This analysis mainly focuses on finding the credibility of the publishers [17].

In this study, comparative analysis of the various machine learning algorithms is performed. To do this task 3 machine learning algorithms are implemented which are consolidated from

the existing systems. The 3 algorithms are Multinomial Naive Bayes , Passive Aggressive Classifier and Decision Tree classifier. In the comparative study the algorithms are compared head to head and conclusion was that the Passive Aggressive Classifier performed best [13].

In this study, multiple models are proposed to predict the category of the news. In this paper the BERT models were implemented with the inspiration of Google's text-based models. In addition to the BERT models, SVM and Naive Bayes classifiers were also proposed. SVM classifier achieved 92 percent accuracy and the Naive Bayes classifier achieved 72 percent accuracy [4].

V. DATA DESCRIPTION

The dataset is collected from the open-source repository Kaggle. It contains two separate files one is fake news.csv and truenews.csv. The number of records in the fake news is 17k and the true news is 20k. Attributes of the dataset are the title of the news, body of the news, subject or the topic of the news, and time stamp of the news. The categories of the news are political and world news. The timestamp of the news ranges from 13 January 2016 to 31 December 2017.

	text	type
5449	Now that it s been established that Mike Pence...	0
15545	BAMAKO (Reuters) - A French military strike on...	1
4532	World-renowned writer/director Joss Whedon has...	0
4514	Were you worried that Donald Trump might fall ...	0
3978	(Reuters) - Ailing Puerto Rico would receive $...	1

Figure 1. Sample dataset with labels

	title	text	subject	date
0	As U.S. budget fight looms, Republicans flip t...	WASHINGTON (Reuters) - The head of a conservat...	politicsNews	December 31, 2017
1	U.S. military to accept transgender recruits o...	WASHINGTON (Reuters) - Transgender people will...	politicsNews	December 29, 2017
2	Senior U.S. Republican senator: 'Let Mr. Muell...	WASHINGTON (Reuters) - The special counsel inv...	politicsNews	December 31, 2017
3	FBI Russia probe helped by Australian diplomat...	WASHINGTON (Reuters) - Trump campaign adviser ...	politicsNews	December 30, 2017
4	Trump wants Postal Service to charge 'much mor...	SEATTLE/WASHINGTON (Reuters) - President Donal...	politicsNews	December 29, 2017

Figure 2. Sample dataset
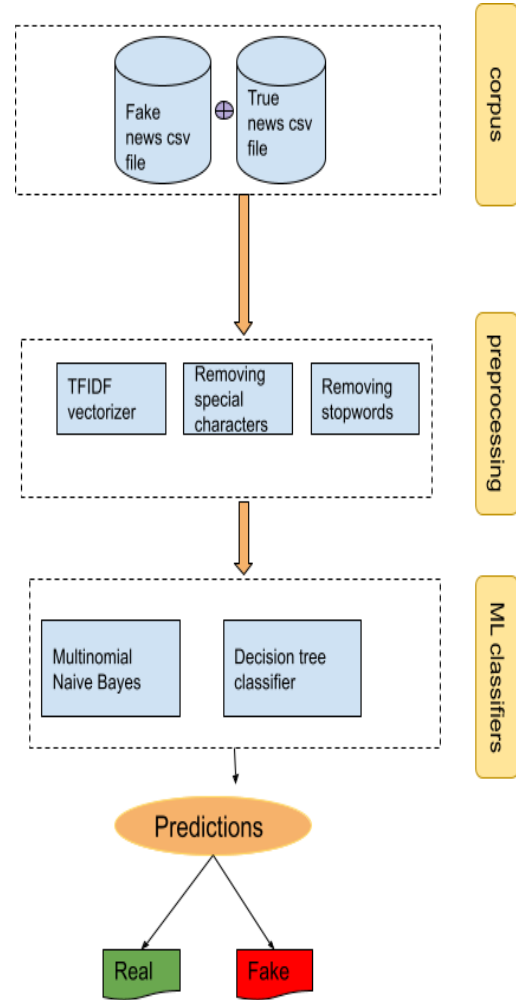
VI. PROPOSED FRAMEWORK

Figure 3. Workflow

The dataset contains true and fake news files separately. Using the Pandas library, we have labeled the true news as 1 and fake news as 0 and created a master dataset. In the data preprocessing step stop words are removed, and special characters and URL tags also were removed from the data. For this task, we have used the regular expression Python library. In the data preparation step, TFIDF vectorizer is applied to convert the text into vectors. TFIDF vectorizer is implemented using scikit-learn library. Data exploration is conducted to perform the text analysis. In the final step, we have used a machine learning classifier Multinomial Naive Bayes and decision tree classifier to classify the text into fake and real.

After the preprocessing data is split into training and testing. After the experimental analysis of test data, we can conclude that Multinomial Naive Bayes outperformed the decision tree.

Multinomial Naive Bayes scored 97 percent accuracy on the test data whereas the decision achieved 84 percent accuracy.
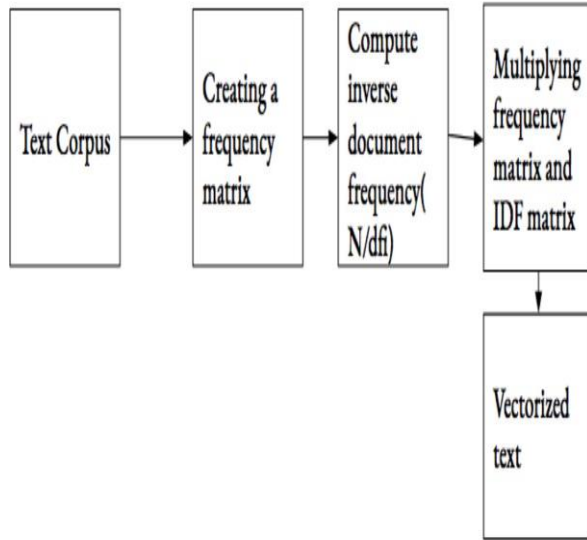
A. Methods

Figure 4. TFIDF vectorizer block diagram

Machine learning algorithms work well with numerical data. However, the NLP tasks such as language detection and fake news prediction contains huge amounts of text to be processed and needs to be converted into text vectors i.e., numerical values and this process is called text vectorization. TF-IDF vectorizer is the most popular method to convert the text into vectors.

TF-IDF stands for Term Frequency Inverse Document Frequency. This method works based on two factors which are Term Frequency and Document frequency.

Term frequency is the frequency of occurrence of the word in the document. The higher the frequency of the word higher is the weightage. The overall frequency of the words is calculated and stored as a matrix with rows representing the documents and columns representing unique words in the document.

Document frequency is the number of documents containing the word. If the document frequency of a word is high, then it means the specific word is the most common. Inverse document frequency is calculated to reduce the weight of the term if it is scattered all over the document.

Multinomial Naive Bayes algorithms follow a probabilistic approach to classify text documents. Multinomial Naive Bayes is preferred over the decision tree algorithms and distance based KNN algorithms. The main advantages of MNB are high training speed and the ability to produce accurate results with less amount of data [3]. Pseudo code of Multinomial Naive Bayes algorithm:

- Assuming S is the input sentence, D is the corpus of z-documents, C is set of m classes.
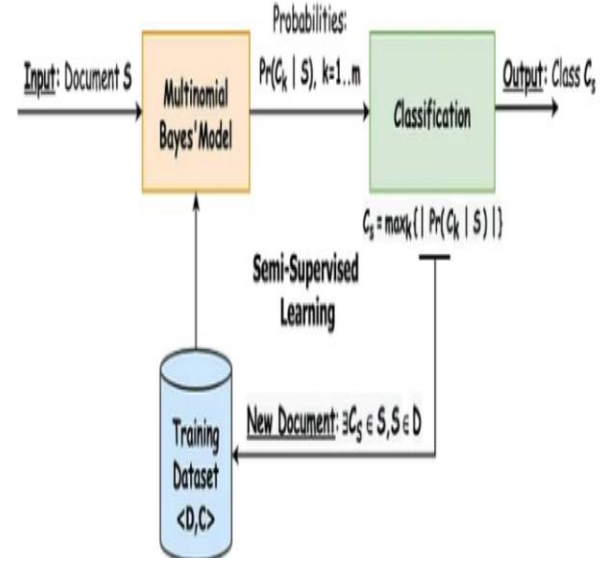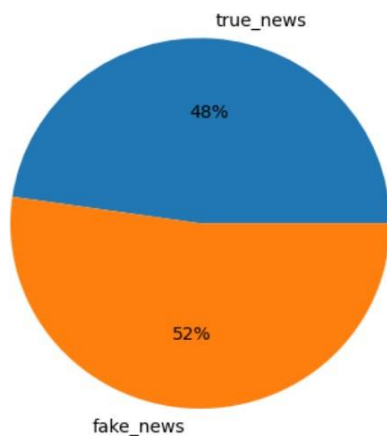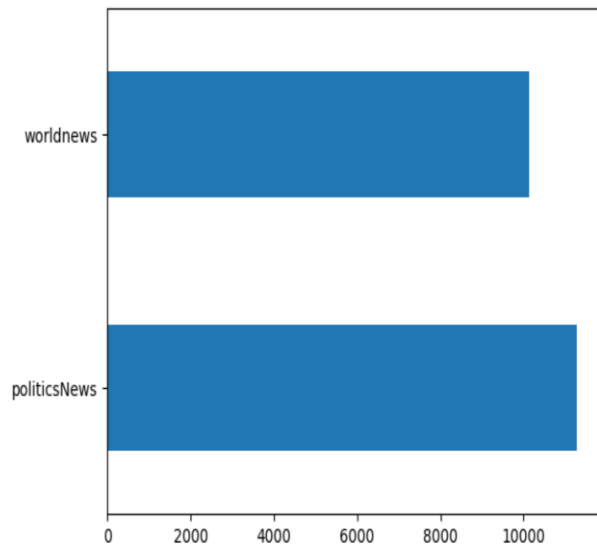- Split "S" samples into n-terms.

Figure 5. Multinomial Naive Bayes block diagram

- Compute the vector of features where is i is the frequency of the –th word.
- Evaluate the prior probability.
- Compute the posterior probability.
- Determine the class of the document.

B. Exploratory Data Analysis

In the exploratory data analysis, we have observed the under-lying structure of the data. The maximum distribution is between 75-85. To proceed with the models, we need to understand the distribution of each category. In the preliminary data analysis part, we explored the text as shown in the fig.6. The figure shows the count plotof news categories in which the highest number of news is political news and the lowest is world news. In the second step, we explored the target distribution as fake and real news. The dataset is balanced as shown in fig.7.There are 48 percent of real news and 52 percent of fake news in the corpus. For a better understanding of data at the word level we have plotted fig.8 and fig.9. which depict the word cloud of both categories. This is an important observation to note because applying the algorithmsdepends on the distribution of the data.

Figure 6. News category distribution

Figure 7. Targets distribution

Figure 8. Word cloud of fake news

Figure 9. Word cloud of real news

VII. RESULTS ANALYSIS

A. Decision tree classifier

In machine learning the classification task is divided into two simple tasks one is training the model and predicting the labels from test samples. There are many algorithms to do these tasks. But we need to choose the algorithms depending on the task at hand. The decision tree algorithm is one of the Machine learning supervised algorithms. This can be used for regression tasks and classification tasks. The mechanism of the decision tree algorithm is based on simple decision rules.

- Root Node: It represents the entire population or sample, and this further gets divided into two or more homogeneous sets.
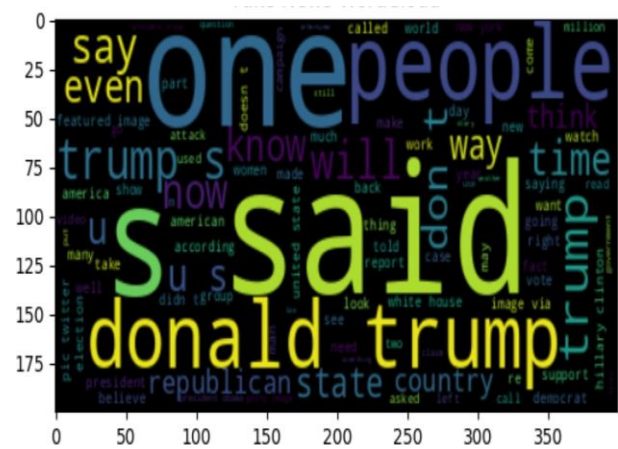- Splitting: It is a process of dividing a node into two or more sub-nodes.
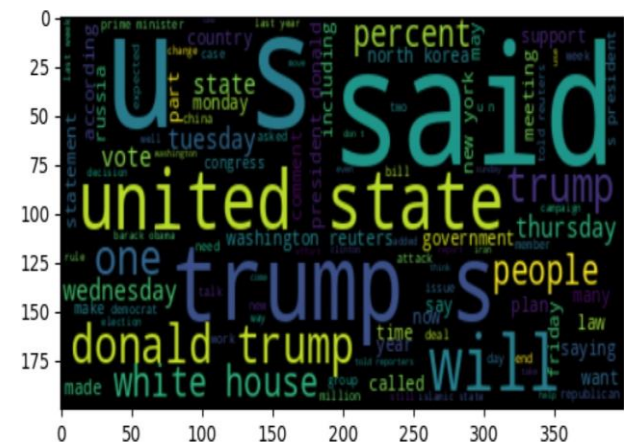- Decision Node: When a sub-node splits into further sub-nodes, then it is called the decision node.
- Parent and Child Node: A node, which is divided into sub-nodes is called a parent node of sub-nodes whereas sub-nodes are the child of a parent node.

B. Multinomial Naive Bayes

Multinomial Naive bayes classifier is a variant of the Naive Bayes classifier. All the variants of the Naive Bayes classifiers work based on the Bayes principle. This algorithm is vividly used in the NLP tasks like filtering spam or ham emails and predicting fake or real news. It calculates each tag's likelihood for a given sample and outputs the tag with the greatest chance.

	precision	recall	f1-score	support
0	0.85	0.84	0.84	4699
1	0.83	0.83	0.83	4281
accuracy			0.84	8980
macro avg	0.84	0.84	0.84	8980
weighted avg	0.84	0.84	0.84	8980

Figure 10. Decision tree classification report

Figure 13. Confusion matrix of Decision Tree

	precision	recall	f1-score	support
0	0.97	0.97	0.97	4689
1	0.96	0.97	0.96	4291
accuracy			0.97	8980
macro avg	0.97	0.97	0.97	8980
weighted avg	0.97	0.97	0.97	8980

These two algorithms are implemented using the sci-kit-learn library which is an opensource library for Python programming language. In the first stage, we have instantiated the objects of the algorithms and trained the models using a training set. After the training process test samples are used to predict the unseen samples. To measure the performance of the model we calculated the classification report and confusion matrix. The Classification report gives the parameters like precision, recall, and F1 score for a better understanding of true/false positives and true/false negatives. The Confusion matrix gives the quantity of true/false positives and true/false negatives. The Naive Bayes classifier achieved 97 percent on the test set and the Decision tree achieved 84 percent accuracy.

Figure 11. Multinomial Naive Bayes classification report

Figure 12. Confusion matrix of Multinomial Naive Bayes

VIII. Results summary

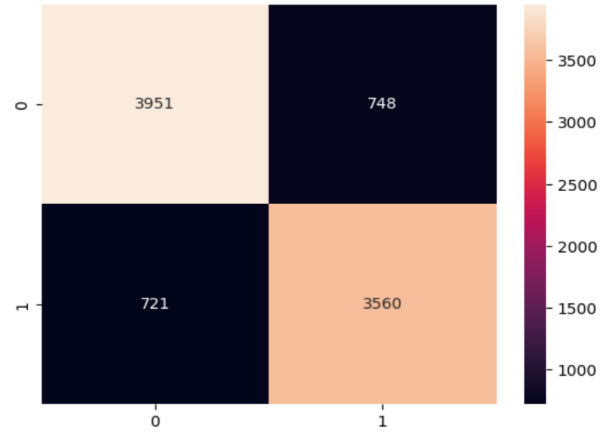In this paper, we have implemented 2 algorithms that are proven to be the best algorithms to deal with textualdata, the Decision tree algorithm and the Multinomial Naive Bayes Classifier algorithm.

REFERENCES

[1] Manisha Aluri, Divya Panchumarthi, Bhargav Boddupalli, Murali Krishna Enduri, Sumana G Sree, and Satish Anamalamudi. An empirical study on fake news prediction with machine learning methods. In *2022 14th International Conference on Computational Intelligence and Communication Networks (CICN)*, pages 419–423, 2022.

[2] Sawsan Alzubi and Feras M. Awaysheh. Edgefnf: Toward real-time fake news detection on mobile edge computing. In *2022 Seventh International Conference on Fog and Mobile Edge Computing (FMEC)*, pages 1–3, 2022.

[3] Bandi Ganesh and Dr. K. Anitha. Implementation of personality detection and accuracy prediction for identification of fake and truenews using decision tree and random forest algorithms. In *2022 International Conference on Business Analytics for Technology andSecurity (ICBATS)*, pages 1–5, 2022.

[4] Sai Rama Krishna Indarapu, Jahnavi Komalla, Dheeraj Reddy Inugala, Gowtham Reddy Kota, and Anjali Sanam. Comparative analysis of machine learning algorithms to detect fake news. In *2021 3rd International Conference on Signal Processing and Communication (ICPSC)*, pages 591–594, 2021.

[5] Rohit Kumar Kaliyar, Mohit Agrawal, and Anurag Goswami. Fndip: Fake news detection on social media using incompatible probabilistic method. In *2022 International Conference on Inventive Computation Technologies (ICICT)*, pages 417–422, 2022.

[6] N. Kousika, Deepa. S, Deephika. C, Dhatchaiyine. B M, and Amrutha. J. A system for fake news detection by using supervised learning model for social media contents. In *2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS)*, pages 1042–1047, 2021.

[7] Oussama Maakoul, Sabah Boucht, Karima EL HACHIMI, and Salma Azzouzi. Towards evaluating the covid'19 related fake news problem: Case of morocco. In *2020 IEEE 2nd International Conference on Electronics, Control, Optimization and Computer Science (ICECOCS)*, pages 1–6, 2020.

[8] Swornamalya. M. N, Kalishwari Meena. V, Sabu Verma, and S. Rajagopal. Nlp based fake news detection using hybrid machine learning techniques. In *2022 3rd International Conference on Electronics and Sustainable Communication Systems (ICESC)*, pages 818–822, 2022.

[9] R. Nomesh. and M.S. Saravanan. Classification of fake news on facebook a novel social network with k-means clustering approach for against principal component analysis method for better accuracy. In *2022 3rd International Conference on Smart Electronics and Communication (ICOSEC)*, pages 722–726, 2022.

[10] C. Rajalakshmi, T. Subika, K. Vaishali, and J. Shana. Fake news prediction on covid dataset using machine learning. In *2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, pages 1–7, 2021.

[11] Amit Neil Ramkissoon and Wayne Goodridge. Legitimacy: An ensemble learning model for credibility based fake news detection. In *2021 International Conference on Data Mining Workshops (ICDMW)*, pages 254–261, 2021.

[12] Amit Neil Ramkissoon and Wayne Goodridge. An energy-efficient ensemble-based computational social system for fake news detectionin manet messaging. In *2022 IEEE Eighth International Conference on Big Data Computing Service and Applications (BigDataService)*, pages 182–183, 2022.

[13] Amit Neil Ramkissoon and Shareeda Mohammed. An experimental evaluation of data classification models for credibility based fake news detection. In *2020 International Conference on Data Mining Workshops (ICDMW)*, pages 93–100, 2020.

[14] Saranya S S, Kanimozhi N, Kavitha M N, Atchayaprakassh K S, Bharani Kumar S, and Ragul K K. Authentic news prediction in machine learning using passive aggressivealgorithm. In *2022 Second International Conference on Artificial Intelligence and Smart Energy (ICAIS)*, pages 372–376, 2022.

[15] Apoorva Shete, Harshit Soni, Zen Sajnani, and Aishwarya Shete. Fake news detection using natural language processing and logistic regression.In *2021 2nd International Conference on Advances in Computing, Communication, Embedded and Secure Systems (ACCESS)*, pages 136–140, 2021.

[16] G Sreeraag and P G Shynu. Stiffness analysis for the prediction of fake news through online digital networks in india. In *2023 International Conference on Advances in Intelligent Computing and Applications (AICAPS)*, pages 1–6, 2023.

[17] Lalinda Udurawana, Ruvan Weerasinghe, and Randil Pushpananda. A hybrid approach for detection of fake news in sinhala text. In *2022 22nd International Conference on Advances in ICT for Emerging Regions (ICTer)*, pages 039–044, 2022.

[18] Asma Ul Hussna, Iffat Immami Trisha, Md. Sanaul Karim, and Md. Golam Rabiul Alam. Covid-19 fake news prediction on social media data. In *2021 IEEE Region 10 Symposium (TENSYMP)*, pages 1–5, 2021.

[19] Zhichao Wang. Deep learning methods for fake news detection. In *2022 IEEE 2nd International Conference on Data Science and Computer Application (ICDSCA)*, pages 472–475, 2022.

[20] Wencheng Yu, Jike Ge, Zhaoxu Yang, Yan Dong, Yujie Zheng, and Haojun Dai. Multi-domain fake news detection for history news environment perception. In *2022 IEEE 17th Conference on Industrial Electronics and Applications (ICIEA)*, pages 428–433, 2022.