RESEARCH ARTICLE OPEN ACCESS

A STUDY ON FAKE NEWS ANALYSIS

Monashri D
M.Tech, Data Science and Engineering, BITS PILANI
Monashri.deena@gmail.com

ABSTRACT

Everyone knows about fake news, so no introduction is needed. Due to the rapid development of social media platforms like Facebook, Twitter, WhatsApp, we have seen and seen a significant increase in internet usage in recent years. YouTube should not be overlooked as one of the main channels for disseminating misinformation. We cannot deny that we all fall for such information. These websites and apps offer a variety of benefits, including the ability to share content that improves and benefits the general public. The main drawback of fake news is that it often spreads like wildfire. The primary motive for the rapid spread of misinformation is to advance your company or your interests politically or financially. Sentiment analysis, a subset of information extraction and retrieval, is used in fake news. Many computer scientists have researched this problem over the years since it occurred in people's daily life. They developed many computational methods and algorithms that are used to solve many of our everyday problems simultaneously. There are many workable approaches thanks to the efforts of researchers in deep learning, neural networks, and other related fields. As a first step, you should identify and understand the source of the news. This is due to the fact that erroneous information is frequently broadcast by the media. When they admitted to spreading false information, they publicly apologized. Spreading false information for entertainment is wrong. The recent coronavirus announcement is a prime example. As the deadly virus spread across the globe, people began spreading the myth that experts predicted the disease would be eradicated by summer. However, it eventually turns out to be the most dangerous of all winters. People don't expect fake news to spread like this because it has a negative psychological effect on them. People are more likely to experience mental illness or discomfort if exposed to false information.

KEYWORDS:*Face book, fake news, media, Twitter,Internet*

I INTRODUCTION

Many scholars claim that once one is caught in this vulnerability, it is difficult to escape from it. Thanks to the Internet, people have the confidence to access information, create their own ideas, and engage with issues of cultural interest. 53% of US people say they often or sometimes obtain their news from online sources, while 59% of Twitter users and 54% of Facebook users say the same. This is according to data from the Pew Research Centre Journalism Project. Over time, the prevalence of distributing false information has steadily grown. Large digital firms like Facebook, Twitter, and YouTube have grown rapidly over the last 10 years, demonstrating this quick development. During the 2016 US presidential election, the issue of false news was made very clear. The widespread distribution of erroneous information has a detrimental effect, not only on the reputation of politicians and their political parties, but also on the reputation of other areas, such as athletics, medicine, and the scientific community [8]. The financial market is another key industry that is touched by the situation. Here, we get that even the smallest rumour has the potential to drastically alter the market and eventually lose the owners money. One of the key reasons why false information travels so fast throughout the globe is because we are entirely reliant on the news we get from social media or any other news source. Numerous examples demonstrate how inaccurate the news is that receives the greatest attention. One such indicator is the widespread dissemination of false and erroneous information about the coronavirus. In recent years, machine learning algorithms have excelled in a variety of fields. The logistic regression method, the TF-IDF vector method, and the random forest classification strategy are three machine learning techniques that are especially good at identifying and categorizing news as true or untrue. According to our study, we frequently assess the credibility of a news report using a particular set of data. The property of news amount is defined, and the answer variable may be fake or real. The main findings of our research are as follows:

The four machine learning models developed in this study are more successful than earlier studies, according to the research, which employs pre-processed data and substantial data mining to discriminate between false and true news.

The suggested method may be used to determine if news is false or true for a number of additional datasets.
• Extensive data mining and pre-processing processes are applied in our work to distinguish between fake and real news.
• According to the research, the four machine learning models proposed for this project are more efficient than previous studies.
• The proposed approach can help identify fake or genuine news for other types of datasets.

II LITERATURE REVIEW

It has become difficult to correctly verify the validity of data, content, and distribution as false news and phony data have become so pervasive. Numerous scholarly projects, some of which have been significant, have been centred on this issue. While some people have studied machine learning, others have studied deep learning. However, there haven't been many studies done on mood analysis or sentiment data.
4-g model using word frequency and TF-IDF to weed out pseudoscientific material. Ahmed H., Traray, and Saad S. Contrarily, neither for fake news nor for real news, nonlinear machine learning models outperformed linear models. Research constraints are less of an issue when you use the largest n-gram.
Two significant groups of false and bogus news identification techniques are reviewed by the writers, Conroy NK, Rubin VL, and Chen. The first session covered language methods, including the extraction of false communication material and the dissection of language patterns in connection to binary communication. The second sort of misdirection we've observed has a lot to do with network technologies because it makes use of network data, such as message metadata or structured information analysis searches, which can be combined to produce comprehensive measures of misdirection. We take notice of the potential of a novel mixed strategy that fuses artificial intelligence, semantic tags, and network-based social data.
Hussain's DME-DM study on Sentiment Analysis (SA) using Natural Language Processing yielded 41 NLP papers. Investigations are ongoing to identify false sites and false reviews as well as handle false, misleading or fake news. In addition, the average accuracy rate decreases as sensation difficulty is explored. Future work that can be done is discussed in this article. According to the paper, the focus should be on expanding the test circuit so that in the future you can probe the inputs more consistently and successfully.
According to Bondielli A. &Marcelloni F., which examined several methods used to complete these tasks and highlighted how difficult it is to obtain relevant information, items believed to help identify false, fraudulent or rumour methods were thoroughly tested. Research is limited in that it can only explore and report many explanations for misinformation and poorly written rumours. Second, the study's selection of important data used to identify fake news was imprecise and impaired the performance of machine learning models.
In their article, "Fake News Detection," Bali A.P.S. and co-author's M. Fernandes, S. Chaubey, and M. Goel discussed the results of a study on the topic. Different features were extracted from the headlines and contents of the three data set representatives and then evaluated using NLP and ML. Gradient boosting was found to be the most effective classification method. Accuracy and F1 scores for seven potential learning algorithm pairs were all below 90%.
According to the study of Faustini P. and Covoes T., one-class classifier (OCC) should be used to define received news by creating a training dataset with only false samples. This case study analyses data from Twitter and WhatsApp and focuses specifically on the Brazilian political scene in the run-up to the 2018 general election. The fact-checking part of this research requires a lot of human effort and is very expensive and time-consuming.

III STATEMENT OF THE PROBLEM

We are addressing the key problems faced in today's scenario on reading and receiving news. We are generally not capable of analysing the news to determine its optimality. In this project we will be implementing machine leaning models such as logistic regression, inverse-frequency vector terminology (TF-IDF), and random forest classification, which are used for both visualisation and exploration. We can conclude our results by analysing the accuracy of the different models that we have used.

IV Dataset and Methodology

This section outlines the tools and strategies we applied in this study to separate false news from the selected data set. The datasets and their accompanying metadata are also discussed in Section 4.1. Section 4.2 covers data pre-

treatment, Section 4.3 addresses data exploration, and Section 4.4 examines the primary methods and methodologies for overcoming this problem.

4.1. Dataset Description and Architecture

The data collection utilised in this investigation comprises both true news and fraudulent news. More than 20,000 instances of fraudulent and authentic news may be discovered in each file in the collection. The dataset takes into consideration the article's title, subject, topic, and publication date. It also contains data from the authentic and fraudulent news data sets utilised for Ahmed, Traray and Saad. Figure: 1 provides a bar chart comparing the count of news with respect to published month

After identifying which year had more false news than genuine news using the date section, we went on to create two new features: "month," which is shown in Figure 2, and "year," which is shown in Figure 3. The 2015 dataset material is completely wrong. By the eighth month, there was more fake news than real news. After that, the real news stories increased. In summary, month = 8 means that the report is more likely to be incorrect.

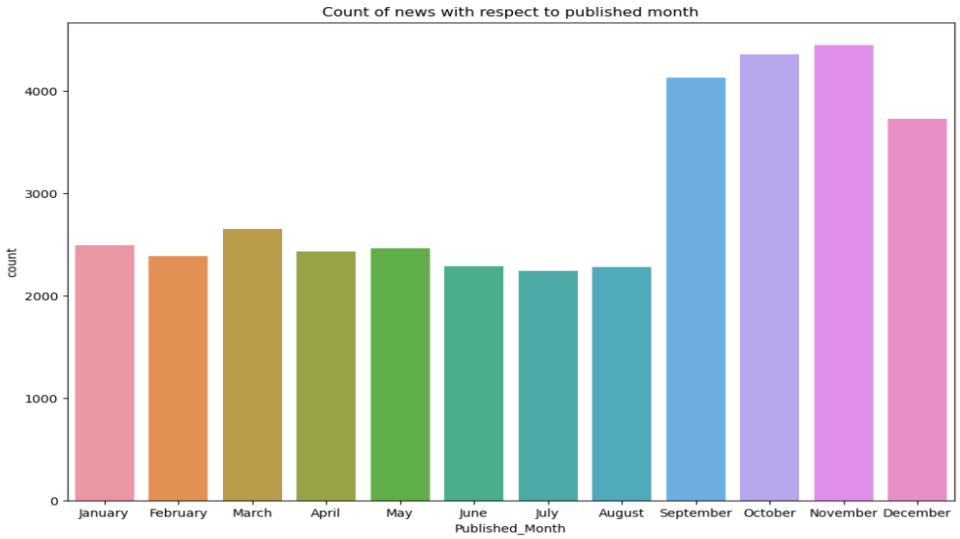

Figure: 1

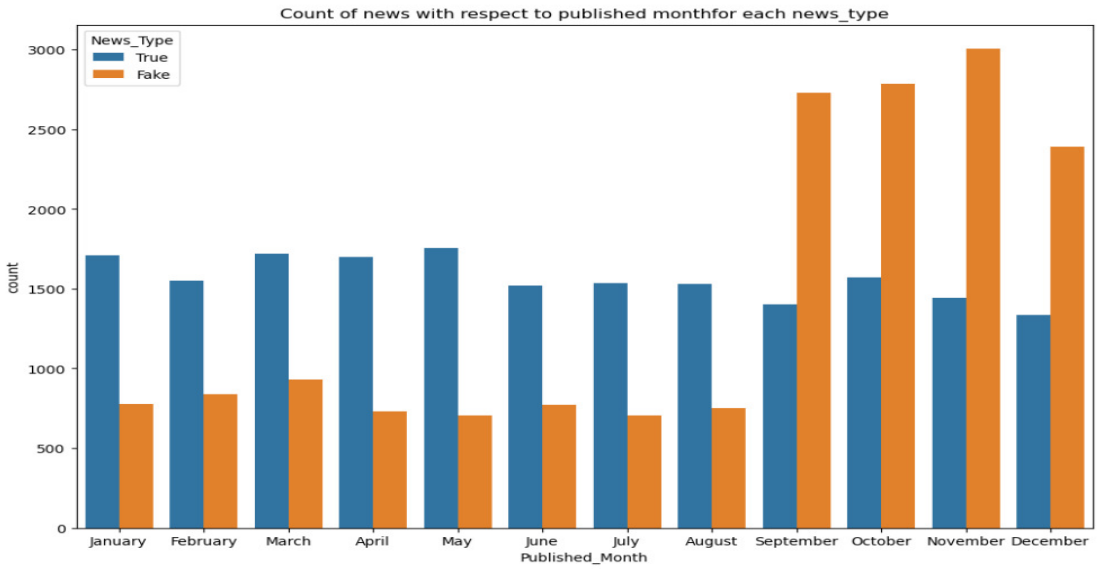

Figure: 2 (shows us the comparison between the comparison of the Ture and fake news with respect to published month)

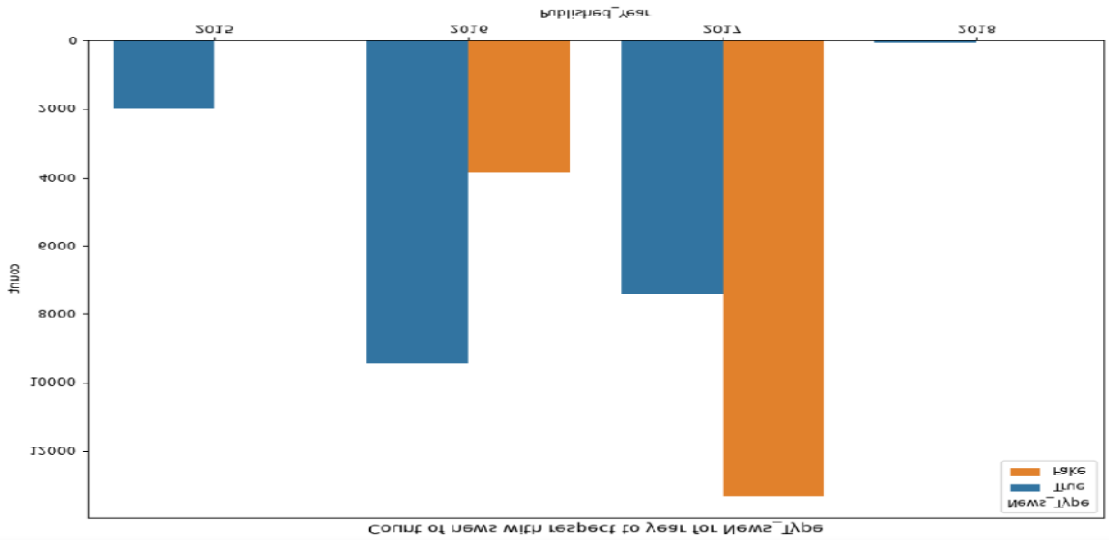

Figure: 3 (depicts us the count of each news with respect to the year of publish.)

The system architecture for the stages in our method is shown in Figure: 4 below. The data set is analysed, pre-processed, segmented, tested, and then put through a total of four different machine learning classification models before being used in the experimentation with the test set.

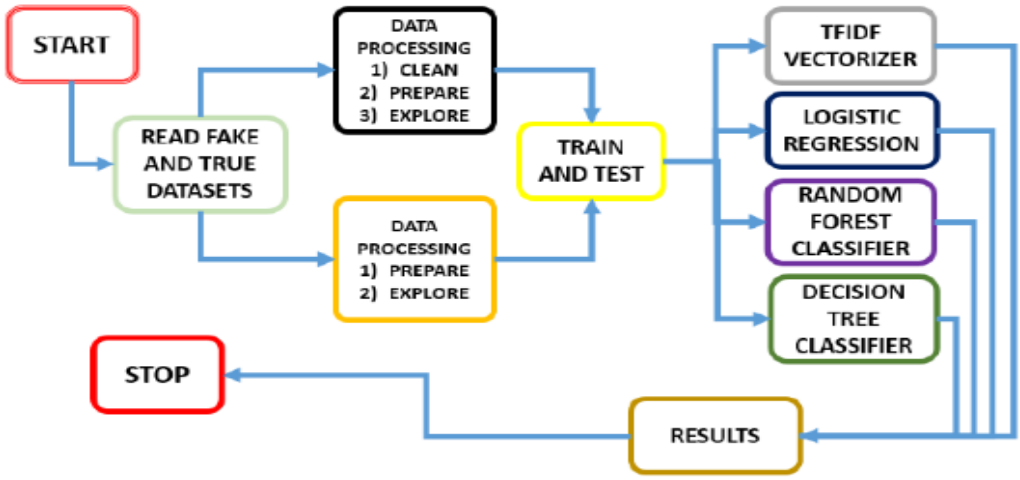

Figure: 4

4.2 Data Pre-Processing

The pre-processing of the data is essential before the training, testing, and modelling stages. Before initiating these operations, authentic and false news are merged. We eliminated the columns from the datasets that did not need to be processed during the dataset cleansing procedure. Stop words and punctuation have been removed. Stop words are frequently occurring words. Examples include "I," "yes," "like," "should," "that," etc. Lowercase letters are created from uppercase letters. The data set is reasonably good after cleaning and the exploration phase begins. However, both clean and unclean datasets are thoroughly explored for more in-depth research. Both real and fake datasets are combined into a data frame for the exploration process to facilitate processing. The table below shows the total number of samples of fake and legitimate news (Table: 1)

S.No	Article Title	Frequency
1	Fake news articles	23,481
2	Real news articles	21,417

Table: 1

4.3. Data Exploration

In the data exploration phase, patterns and insights from both false and genuine news are discovered by examining and visualising the data. We created various graphs using the Python packages matplotlib and seaborn. We first created word clouds with instances of true and misleading news. The word clouds included dataset-related keywords. Figure: 5(a) shows the word cloud for most frequently occurring words with remarks like "Trump", "President", "Obama", "People", "American", "One", "New", "Hillary", "White", "Donald" and so on, of the true dataset. Figure: 5(b) contains most frequently occurring words of the fake news dataset with the item titles including "House", "Washington", "Said", "Government", "State", "Reuters", "Election", "Party" and so on. Figure: 6(a) and 6(b) depicts us the less frequently occurring words in Ture and Fake dataset respectively.

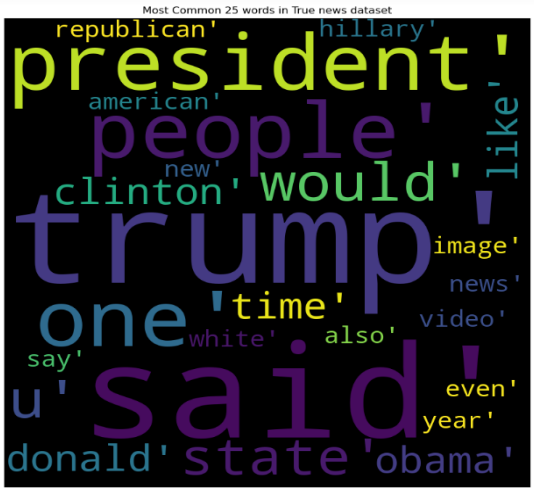

Figure: 5(a)

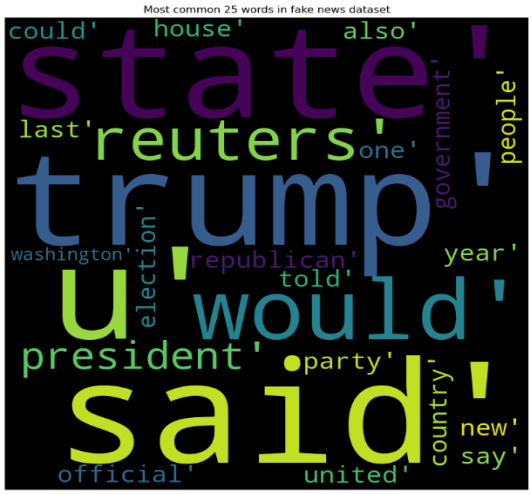

Figure: 5(b)

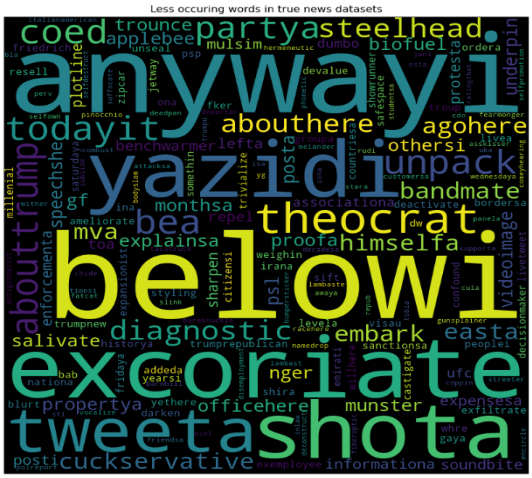

Figure: 6(a)

Figure: 6(b)

Additionally, a bar graph showing the number of different news topics has been drawn. After cleansing the information, we can see that the Middle East news has the greatest numbers. Figure: 7 displays a word count graph and a graph of the article's topic. (Article content).

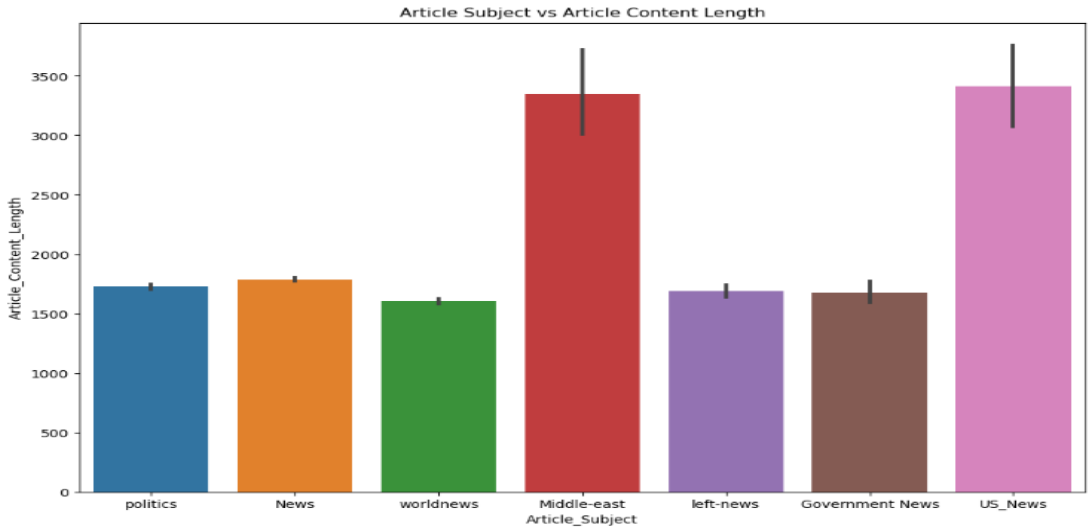

Figure: 7

Figure: 8(a) and 8(b) showcases us the News Count-Published month with respect to news subject. From the graph we can say that the politics news is high in frequency than others, expect for months like Sep, Oct, Nov, Dec where world news is higher in count and with respect to year, politics news are higher in frequency with respect to all years

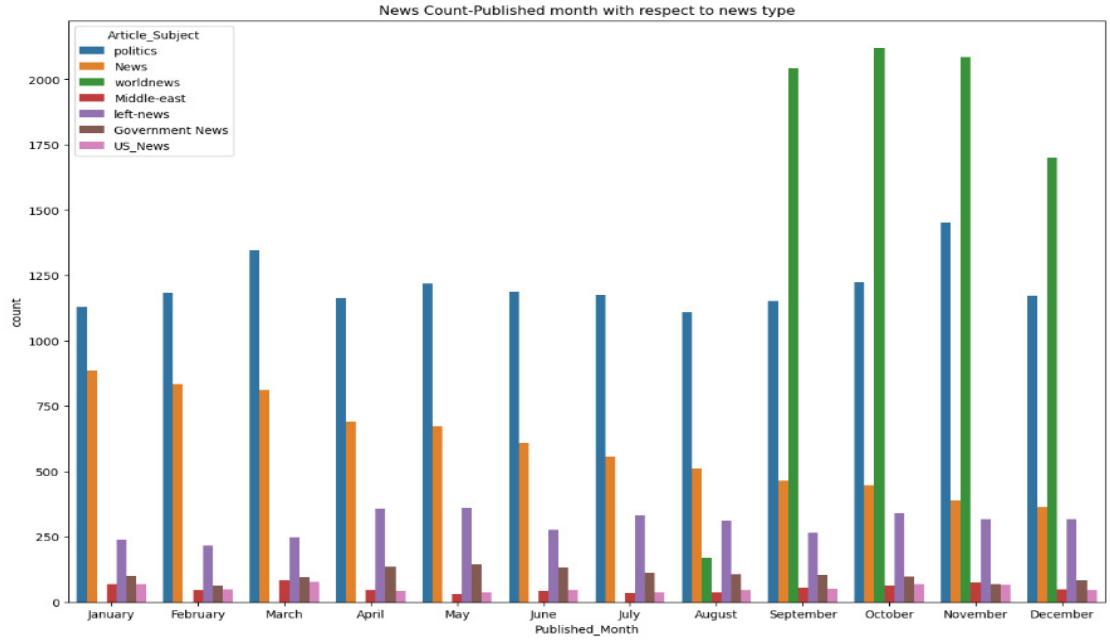

Figure: 8 (a)

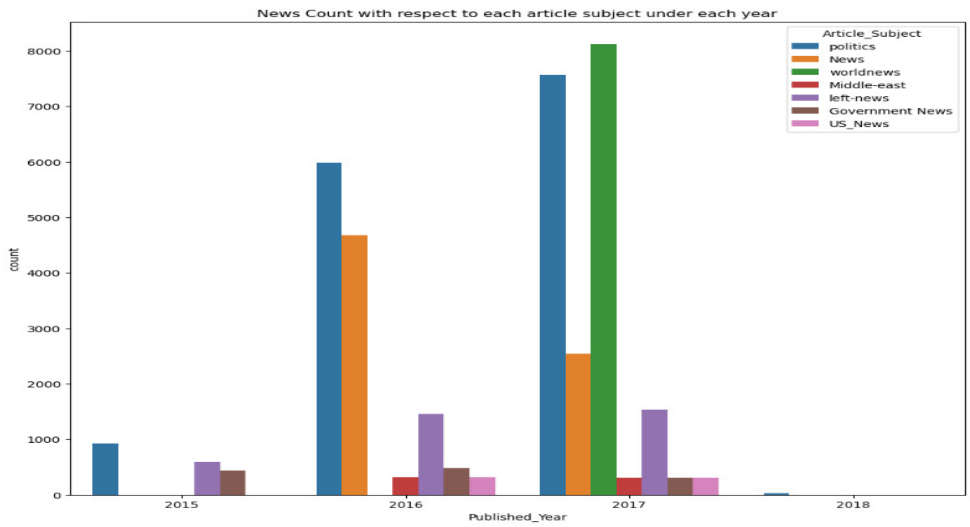

Figure: 8 (b)

V OUR APPROACH

The methods we used to assess fake and real news are discussed in this section.

5.1 TF-IDF Vectorizer

Scikit Rail, Python tool used. The ideal way to perform any work with the TF-IDF vector model is to use this tool. Terms of particular interest in a document or material are represented by TF-IDF vectors, which are part of this technique. The next aspect of this approach is how important it is to repeat words. (TF). Shows how often the phrase appears in the dataset (we found this information while doing a data search). The formula for finding the TF is shown in Equation (1):

$$\mathrm{TF}(t,d) = \frac{Number\ of\ times\ t\ occurs\ in\ a\ document\ 'd'}{Total\ word\ count\ of\ document\ 'd'}$$

The IDF, or inverted document frequency, is the next parameter to find to determine if the model is working properly. It is used to measure how important a word is throughout the information. The formula for IDF is shown in Equation (2):

$$\mathrm{IDF}(t) = \log_e\left(\frac{Total\ number\ of\ documents}{Number\ of\ documents\ with\ term\ t\ in\ it}\right)$$

The strength of the Israel Defense Forces should be ascertained as the next step. The inverse document frequency built in word frequency represents the TF-IDF.the formula of which is shown in Equation (3):

$$\text{TF-IDF}(t,d) = \text{TF}(t,d) * \text{IDF}(t)$$

The TF-IDF algorithm took the feature shape and rated the most pertinent terms from the real and false news in our dataset. This led to increased effectiveness. Second, we employ a technique called the vector TF-IDF approach.

The TF-IDF vectorizer maintains a (trickle) word event recurrent network and maps consecutive words to stress files using an in-memory lexicon (Python dictionary). The TF-IDF vector is composed of the compressed data and the saved repetition weights.

5.2. Logistic Regression

The logistic regression method is the third strategy we use to make this model work properly. Machine learning requires logistic regression, which can determine the relationship between probability (probability) and probability (outcome) for a given outcome. When the predictor value is binary, the logistic regression algorithm is used. For example, he responds correctly or incorrectly when he guesses a number. The probability (outcome) and significance (probability) of a given outcome are related using logistic regression. sklearnlinear_model can be used to load a logistic regression model.

We can determine the accuracy of the model so that we can use it for our future evaluation.

```
fitted Logistic Regression model
The training dataset accuracy is: 0.9645443404634582


classification report
              precision    recall  f1-score   support

           0       0.96      0.96      0.96      4282
           1       0.97      0.96      0.97      4694

    accuracy                           0.96      8976
   macro avg       0.96      0.96      0.96      8976
weighted avg       0.96      0.96      0.96      8976


Accuracy of model:  0.9635695187165776
```

Currently, the obtained accuracy of Logistic Regression model is **0.9645443404634582**

5.3 Random Forest Classifier

The hyperparameters of a decision tree or class classifier are almost identical to those of a random forest. When creating branches, this method increases the arbitrariness of the model. A random forest predictor is, first and foremost, a way to generate multiple alternative trees and combine them to produce more accurate and reliable predictions. The decision tree or hyperparameters of a crushing classifier are almost identical to a random forest. When creating branches, this method increases the arbitrariness of the model. There are many random trees that can yield a value, and the true result of this classification is a value with more points. Additionally, how do we load a linear model from Sklearn..


```
fitted Random forest classifier
The training dataset accuracy is: 1.0


classification report
              precision    recall  f1-score   support

           0       0.98      0.97      0.98      4282
           1       0.97      0.98      0.98      4694

    accuracy                           0.98      8976
   macro avg       0.98      0.98      0.98      8976
weighted avg       0.98      0.98      0.98      8976


Accuracy of model:  0.9772727272727273
```

We can get the accuracy of the model used, which will be used later for our observation.
Currently, the obtained accuracy of Logistic Regression model is **0.9772727272727273**

5.4 Decision Tree Classifier

This prediction is among the finest in machine learning, as far as we know. It is possible to use unsupervised learning methods, which are common with decision trees, to solve categorization and regression issues. Functions as anticipated. Tree structures called order trees have a goal variable with a specific collection of characteristics. Decision trees that are based on the Gini indicator perform well and are produced rapidly. Our concluding machine learning technique is categorization using decision trees. Popular autonomous learning techniques for categorization and regression issues include decision trees. A decision tree is effective at spotting false news as well. The Sklearn tree model must be loaded in order to use the decision tree classification.

```
fitted Decision Tree Classifier model
The training dataset accuracy is: 1.0


classification report
              precision    recall  f1-score   support

           0       0.94      0.91      0.92      4282
           1       0.92      0.95      0.93      4694

    accuracy                           0.93      8976
   macro avg       0.93      0.93      0.93      8976
weighted avg       0.93      0.93      0.93      8976


Accuracy of model:  0.9283645276292335
```

We can get the accuracy of the model used, which will be used later for our observation.
currently, the obtained accuracy of Logistic Regression model is **0.9283645276292335**

5.5 Multinomial Naive Bayes Algorithm

One of the variations of the Naive Bayes algorithm used in machine learning is called polynomial Naive Bayes and is excellent for use on polynomial dispersive datasets. This method can be used to predict the label of a text when there are many groups to classify the text. It does this by calculating the probability of each input text label and producing the label with the highest probability as output.
We can determine the accuracy of the model so that we can use it for our future evaluation.
Currently, the obtained accuracy of Logistic Regression model is **0.9166666666666666**


```
fitted MNB model
The training dataset accuracy is: 0.9133522727272727


classification report
              precision    recall  f1-score   support

           0       0.91      0.92      0.91      4282
           1       0.93      0.91      0.92      4694

    accuracy                           0.92      8976
   macro avg       0.92      0.92      0.92      8976
weighted avg       0.92      0.92      0.92      8976


Accuracy of model:  0.9166666666666666
```

VI EXPERIMENTAL RESULTS

This part is divided into two distinct sections, Sections 6.1 and 6.2, which discuss the experimental setup and the results of our study.

6.1. Experimental Setup

In Jupiter Notebook, four principles are put into practice. We used Python 3.5 and later for this. We used numpy, pandas, scikit learning, natural language tool kit (NLTK), matplotlib and seabarn as training and testing tools. We split the data set, with an 80:20 split between training and test sets.

We also created a different dataset by combining the combined data from the real news dataset and the false news dataset. To determine if the news is fake or true, we decided to create a feature called Human Test. A random news text is used as input and our goal is to determine if the news is real or fake by comparing it with the entire group. (both real and real dataset)

6.2. RESULTS

F1 score was used to evaluate the summary and results of Scikit library classification of confusion matrix, precision and recall.

In the test sample, logistic regression was first evaluated. The logistic regression model has an accuracy of 97%, which is almost perfect. The total number of cases of fake news identified by the algorithm was 4,709, and 4,222 cases of real news. However, it produced 25 fake news stories and 24 fake news stories, indicating that these news examples contained both true and false information.

Second, the test dataset was used to evaluate the random forest classification model. The accuracy of the model is 98%, which is the best among the models evaluated. A total of 4,644 cases of fake news and 4,248 cases of legitimate news were identified by the algorithm.

Third, 92% success was achieved with decision tree classification. A total of 4,688 cases of fake news and 4,210 cases of legitimate news were identified by the algorithm.

Finally, we used a multinomial naive base model, which had a success rate of 91%. In total, the algorithm identified 4,716 fake news and 4,235 legitimate news. 15 cases of true false news and 14 cases of true false news indicate that these news samples contain both true and false information.

Figure 9 (a–d) presents the confusion matrix of false and true news datasets for logistic regression, random forest, decision tree and naive polynomial cells algorithms.

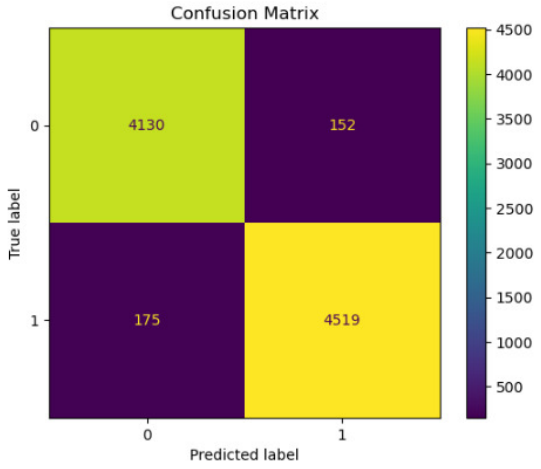

Figure: 9 (a)

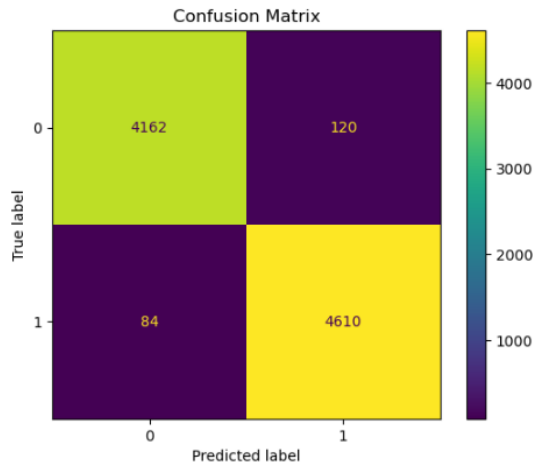

Figure: 9 (b)

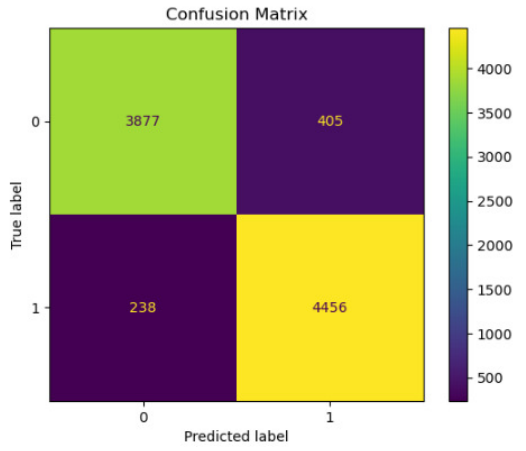

Figure: 9 (c)

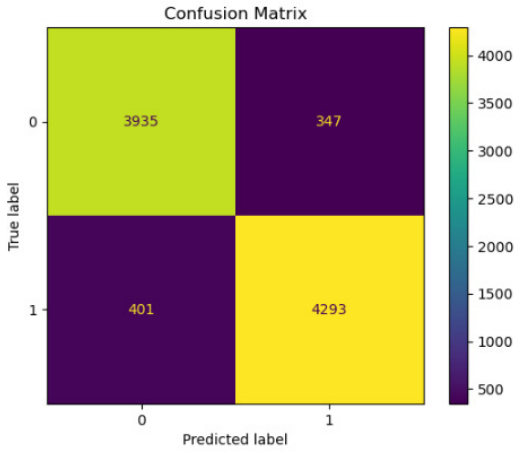

Figure: 9 (d)

Figure 9 Confusion matrix for real, fake, actual and estimated data sets: (a) Confusion matrix for logistic regression; (b) mixing matrix of random forest classification model; (c) confusion matrix of the decision tree classification model; and (d) the confusion matrix in the polynomial-naïve Bayes model.

VIISUMMARY OF THE RFESULTS

To assess the accuracy of all applied models based on F-score, we summarized our research. The accuracy of a model on a data set is measured by the F-score, also known as the F1-score. It is used to evaluate binary classification algorithms that classify cases as "positive" or "negative".

The harmonic mean of model accuracy and recovery, called the F-score, is a way to combine model accuracy and recovery.

F-score Formula is shown in Equation: 4

$$F_1 = \frac{2}{\frac{1}{\text{recall}} \times \frac{1}{\text{precision}}} = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}}$$
$$= \frac{\text{tp}}{\text{tp} + \frac{1}{2}(\text{fp} + \text{fn})}$$

F-score Formula Symbols Explained

Presision	The proportion of true positive cases among those classified as positive by the algorithm is called precision positive. In other words, the ratio of true positives to false positives + true positives.
Recall	Recall The percentage of cases from all positive samples classified as positive is called recall, also known as susceptibility. Rather, it is the ratio of true positives to true positives + false negatives.
tp	Number of true positives classified by the model.
fn	Number of false negatives classified by the model.
fp	Number of false positives classified by the model.

The F1 score obtained by the above results are shown in the following Table (Table: 2)

And we have used a graph to show us the result of the accuracy of the model for data before and after cleaning process.

The line plot graph is depicted in the below figure; Figure 10

S.No	Classifier Model	Accuracy (text data)	Accuracy (text+non-text) Before cleaning
1	Multinomial NB	0.916667	0.834893
2	Gaussian NB	0.901404	0.916667
3	Logistic Regression	0.96357	0.968471
4	Random Forest	0.977273	0.985851

Table: 2

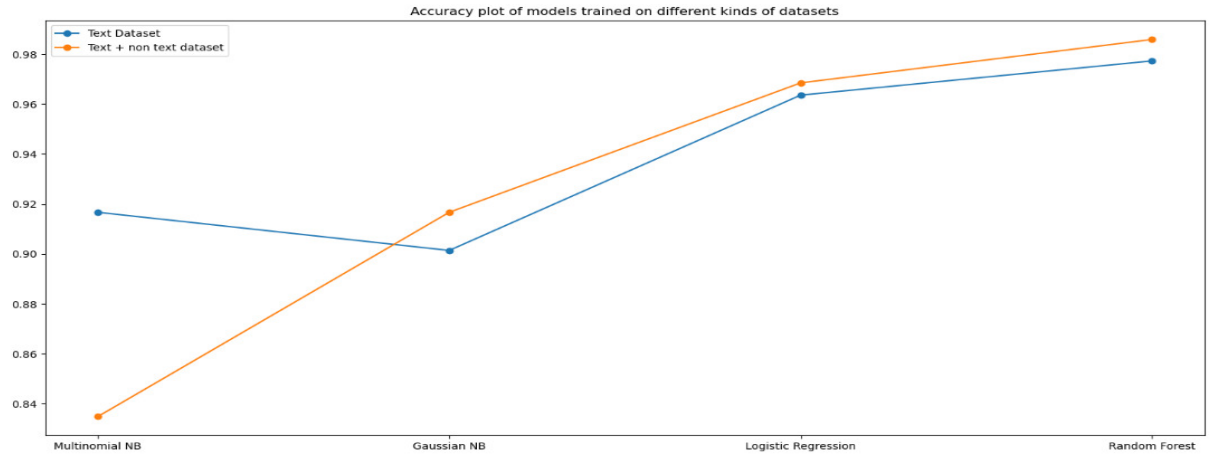

Figure: 10

VIII DISCUSSION

The results of the present experiment show that all algorithms produced great results, indicating the effectiveness of the research study method in distinguishing between true and false news.

As this is the authors' first attempt at such a project, the present study's success rate of over 90% is a good starting point. This study proves that fake news can be detected quickly and effectively. The current study's ability to deliver the best results is highly successful as the majority of research articles consider the results to be more than 80% effective. A recent study reveals that the spread of fake news is not a big problem in society. This research also found that cleaning the data set is the most important step to take in studies of a similar nature. Fake news spreads for a number of reasons. Our article explains how to combat fake news.

We believe the next step in this search is to create a virtual user interface. A good GUI is crucial when creating an app because it needs to look attractive. People can copy and paste any text into the GUI to see its classification results using the GUI.

Shows how technology has made life easier and more difficult. If we look at consumer needs, we can see that one of the main consumer needs is the difference between fake and real news. Other user needs include technical options and decision support. Users can distinguish between real news and fake news. The tools used in this research project are machine learning techniques.

These methods that can be used after loading the necessary tools include vector TF-IDF, random forest classifier, logistic regression, decision tree classifier and polynomial NB methods.

This approach was chosen because the algorithms used can produce very accurate results. The figure below shows the difference between the different programs tried over the last three years. (11)

Sr. No.	Machine Learning Models	Accuracy
1	TF-IDF-Vectorizer, Logistic Regression, Random Forest Classifier, Decision Tree Classifier	99.45%
3	Random forest algorithm, Perez-LSVM, Linear SVM, multilayer perceptron, bagging classifiers, boosting classifiers, KNN	99%, 99%,98%, 98%, 98%, 88%
4	LSTM and BI-LSTM Classifier	91.51%
5	Term Frequency-Inverted Document Frequency (TF-IDF) and Support Vector Machine (SVM)	95.05%

Figure: 11

The precision of other publications is clearly lower than our study, as shown in the figure above. Shows how flawless our results are. One of the disadvantages of the study was that the databases were not very large. Only four machine learning algorithms were used in the study.

IX.CONCLUSIONS

Every kind of news is generated through our social media and most of it is false. Usually, when two competing worlds collide on an issue, we wonder which is right. In trying to decide which source to place our trust in, we put ourselves in a difficult position. Cleaning the information is very important and we have covered that in the discussion section as well. This is critical because it affects the results of the study. As we can see from counting the number of times each statement appears in the dataset, the words trump and article appear more frequently after cleaning the data. Words such as the, are, and are frequently found in uncleaned datasets.When used without other categories, these phrases lack personality and are considered useless. To generate reliable results, databases need to be cleaned. The last point the authors would like to make is that while spreading misinformation sometimes makes people happy, it often makes people sad. Don't spread false news. We used some amazing machine learning techniques in our study and the results were amazing. The software has demonstrated nearly 99% flawless accuracy.As a result of this study, Internet addicts no longer need to worry about fake news. The epilogue to the given work has some shortcomings and limitations. This happens if the information is not balanced or cleaned up because the results are not accurate and may not be useful. In terms of computing speed, Aggregate Data Structure Spark Machine Learning can deliver superior results. Moreover, the newly leveraged LSTM false news databases can be used with deep learning-enabled big data models.

X.FUTURE SCOPE

The meaning of fake news is still debated and there is still some ambiguity, so marketing experts are also trying to explain it from a marketing perspective. Additionally, the line between error and fraud is sometimes blurred. Therefore, fake news in marketing should be a subject of future study as a distinct marketing idea. Next, to understand the ethical responsibilities of marketing practitioners, ethics—both individual and organizational—must be examined.

Media plays a big role in spreading and spreading fake news. Therefore, future studies may focus on social media as a vehicle and platform for disinformation. While the role of social media in disseminating false news has been examined in psychology, communication, and journalism, marketing and customer studies have not. Future study could fill this gap in several ways. For example, they can see why people spread misinformation when it comes to marketing. Although there are different reasons in other lines of study, the motivation of people in business still needs more attention. Additionally, they can learn whether fake news is credible and why people believe it.

Without a question, false news has a negative impact on society, and marketing should take this into consideration. Misinformation is common in the branding and commercial industries. This document classifies some marketing materials as false news, including commercials, public relations, concealed advertising, and promotions. It may be suitable to conduct further research on the businesses and brands connected to false news as well as the impacts of fake news on brands and brands.Further research into how fake news impacts consumers may be helpful. At the social level, the majority of study on the impacts of false news has been conducted. Fake news, on the other hand, typically happens at the level of businesses and people when it comes to marketing and consumer behavior. Future research should concentrate on how businesses create their marketing campaigns and develop strategies to counter false news.In other words, marketing and consumer behavior are neglected in favor of research in psychology, politics, or communication. You can learn how people view or respond to false news by doing research on it. It may be possible to gain insight into how individuals as customers react to false information by considering other individual variables, such as how bogus news affects consumers' feelings.

XI. REFERENCES

[1] Ahmed, W., Vidal-Alaball, J., Downing, J. and Seguí, F.L. (2020), “COVID-19 and the 5G conspiracy theory: social network analysis of twitter data”, Journal of Medical Internet Research, Vol. 22 No. 5, p. e19458.

[2] Aleinikov, A.V., Miletskiy, V.P., Pimenov, N.P. and Strebkov, A.I. (2019), “The fake-news phenomenon and transformation of information strategies in the digital society”, Scientific and Technical Information Processing, Vol. 46 No. 2, pp. 117-122.

[3] Allcott, H. and Gentzkow, M. (2017), “Social media and fake news in the 2016 election”, Journal of Economic Perspectives, Vol. 31 No. 2, pp. 211-236.

[4] Asudeh, A., Jagadish, H.V., Wu, Y. and Yu, C. (2020), "On detecting cherry-picked trendlines", Proceedings of the VLDB Endowment, Vol. 13 No. 6, pp. 939-952.

[5] BBC (2020), "Coronavirus: outcry after trump suggests injecting disinfectant as treatment", available at: www.bbc.com/news/world-us-canada-52407177

[6] Belanche, D., Casaló, L.V., Flavián, C. and Schepers, J. (2020), "Service robot implementation: a theoretical framework and research agenda", The Service Industries Journal, Vol. 40 Nos 3/4, pp. 203-225.

[7] Berthon, P., Pehlivan, E., Yalcin, T. and Rabinovich, T. (2019), "True, fake and alternative: a topology of news and its implications for brands", Journal of Product and Brand Management, Vol. 29 No. 2, pp. 144-149.

[8] Blake-Turner, C. (2019), "Fake news, relevant alternatives, and the degradation of our epistemic environment", Inquiry, pp. 1-21.

