FAKE NEWS DETECTION USING PYTHON

Team Members: Aditya Chettipalli, Ahteshamuddin Mohammed, Ganesh Gude, Nasiruddin Khazi.

Abstract:

Fake news has become a significant problem today, with the rise of social media and online platforms. The spread of false information can have profound consequences, including shaping public opinion, influencing elections, and even inciting violence.

Detecting and combating fake news is crucial, and various techniques have been developed to identify and verify the authenticity of news articles. Some common approaches involve fact-checking by trained journalists or utilizing automated fact-checking tools.

Machine learning algorithms have also been developed to automatically detect fake news articles. These algorithms typically involve analyzing the content of the article and identifying patterns that are characteristic of fake news, such as sensationalist headlines, biased language, or unsupported claims.

To achieve this, machine learning models are trained on a labeled dataset of examples that distinguish between real and fake news articles. These models are then used to classify new articles as either real or fake on unlabeled data.

Table of Contents

INTRODUCTION

Fake news is a growing problem in today's digital age, where information spreads rapidly through social media and online platforms. The extensive circulation of false information can have severe consequences, including shaping public opinion, influencing political outcomes, and even inciting violence. The rapid pace at which news spreads in the digital age makes it difficult to verify the authenticity of news articles, making it easier for false information to spread. This problem has been further worsened by the fact that people are more likely to share sensational news, irrespective of whether it is true or false.

Several approaches have been developed to combat fake news, including fact-checking by trained journalists and automated fact-checking tools. However, these approaches are time-consuming and often rely on human intervention, making them impractical for handling the large volume of news articles published daily. Machine learning algorithms offer an efficient solution to detect and combat fake news automatically.

Machine learning algorithms involve analyzing the content of news articles and identifying patterns that are characteristic of fake news. The algorithms are trained on labeled datasets that distinguish between real and fake news articles. The labeled dataset is used to train the machine learning model to identify the patterns that differentiate between real and fake news articles. The trained model is then used to classify new articles as either real or fake, with a high degree of accuracy.

The potential impact of machine learning algorithms in reducing the spread of fake news is immense. By detecting and flagging fake news articles, these algorithms can prevent the spread of false information, limit its impact on public opinion, and reduce the potential harm caused by the spread of fake news. However, the effectiveness of these algorithms is highly dependent on the quality of the labeled dataset used for training. A well-labeled dataset is essential to train accurate machine learning models, and therefore, it is critical to invest in developing high-quality labeled datasets for machine learning algorithms.

The previously published solutions used different datasets and used bag of words (Baarir, 2021) or TF-IDF vectorizer (Sharma, 2020 & Ahmad, 2020 & Jain,2019) to extract text features and trained and did performance analysis on various machine learning algorithms. In this paper, we are going to perform feature extraction using TF-IDF vectorizer on a new dataset and we will discuss various machine learning algorithms that are used to detect fake news articles, highlighting their strengths and limitations. Finally, we discuss the potential impact of machine learning algorithms in combating fake news and suggest future research directions to further improve their effectiveness.

BACKGROUND

The wide spread of fake news had caused several issues in the past as well, In December 2016, a shooting occurred at Comet Ping Pong, a pizza shop in Washington D.C., due to false tweets that claimed the shop was involved in a pedophile sex ring involving Hillary Clinton and her campaign. This conspiracy theory gained traction on anonymous bulletin board sites and social media after the announcement of the FBI's investigation into Clinton's use of private email during her tenure as Secretary of State. Despite social media banning related posts, threats against the shop increased, and a man from North Carolina showed up with a rifle intending to rescue children he believed were trapped in the shop. This incident highlights the dangerous consequences of the spread of fake news and misinformation on social media.

The above-mentioned incident is just one example of how a fake tweet can lead to a chain of events, some fake news articles have created a rift between different communities in India, which lead to few deaths and several injuries. These types of incidents will continue to occur if we do not contain the spread of fake news circulation on the various websites or social media handles.

As part of our work, we wanted to test and find a best machine learning model that classifies news articles as fake or true and then extend it on to larger scale to implement it on one of the major social media platforms like Twitter as part of future work.

LITERATURE REVIEW

The widespread of fake news has become a critical issue, few works have been published to address this issue in the past.

Jain et al. (2019) proposed a smart system for fake news detection using machine learning, where they employed a hybrid approach that combined content-based and user-based features. They trained their model on a dataset of real and fake news articles and achieved an accuracy of 92.4%. The authors also developed a Chrome extension for their system, which can be used to flag potentially fake news articles while browsing the web.

Also, in 2019 in the research conducted by Manzoor et al. (2019), Fake news detection using machine learning approaches: A systematic review. Manzoor and Shingla did not train their own model. Instead, they reviewed 32 research papers published between 2016 and 2019 that used machine learning techniques for fake news detection. Manzoor and Singla analyzed the reviewed papers and reported on the performance of the different machine learning algorithms for fake news detection. They also identified the strengths and weaknesses of the existing approaches and highlighted future research directions in the field of fake news detection.

Sharma et al. (2020) proposed a fake news detection system that utilized the Naïve Bayes and Support Vector Machine (SVM) algorithms. They extracted features such as the length of the article, the number of named entities, and the presence of specific words to train their model. The authors achieved a classification accuracy of 89.16% on their dataset of real and fake news articles. Ahmad et al. (2020) developed an ensemble-based approach for fake news detection, where they combined different machine learning algorithms such as SVM, Random Forest, and Multilayer Perceptron (MLP). They evaluated their model on a dataset of 12,000 articles and achieved an accuracy of 94.5%.

In another study, Sharma et al. (2020) proposed a model that used several machine learning algorithms, including logistic regression, random forests, and k-nearest neighbors, to classify news articles as fake or real. The model achieved an accuracy of 87.32% on a dataset of news articles. Ahmad et al. (2019) used an ensemble of machine learning algorithms, including decision trees, k-nearest neighbors, and random forests, to classify news articles. The model achieved an accuracy of 91.47% on a dataset of news articles.

Baarir and Djeffal (2021) presented a machine learning-based approach to detect fake news, where they utilized a combination of content-based and metadata-based features. They trained their model on a dataset of real and fake news articles and achieved a high accuracy rate of 96% on the testing set. Manzoor and Singla (2019) conducted a systematic review of various machine learning-based approaches for fake news detection. They analyzed the strengths and limitations of different models, including traditional machine learning algorithms and deep learning methods, and provided insights into the future research directions in this field.

METHODOLOGY

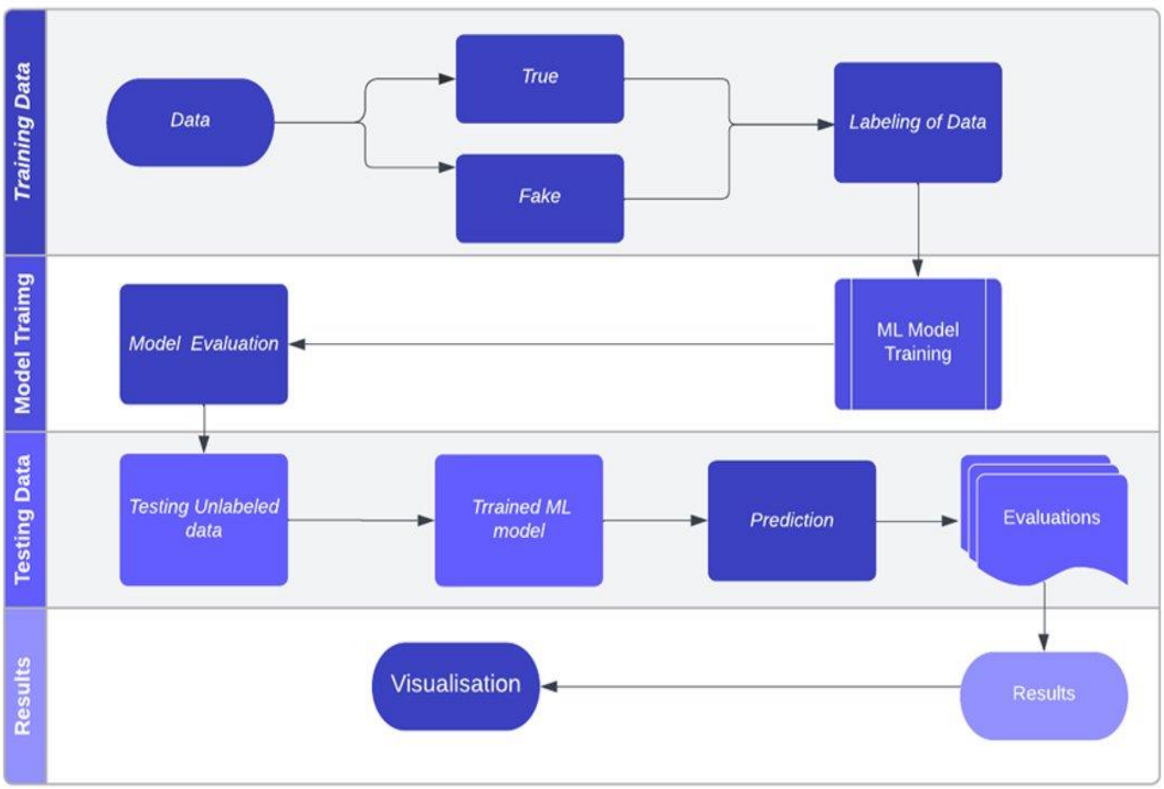


Fig-1: **Flowchart of the project**

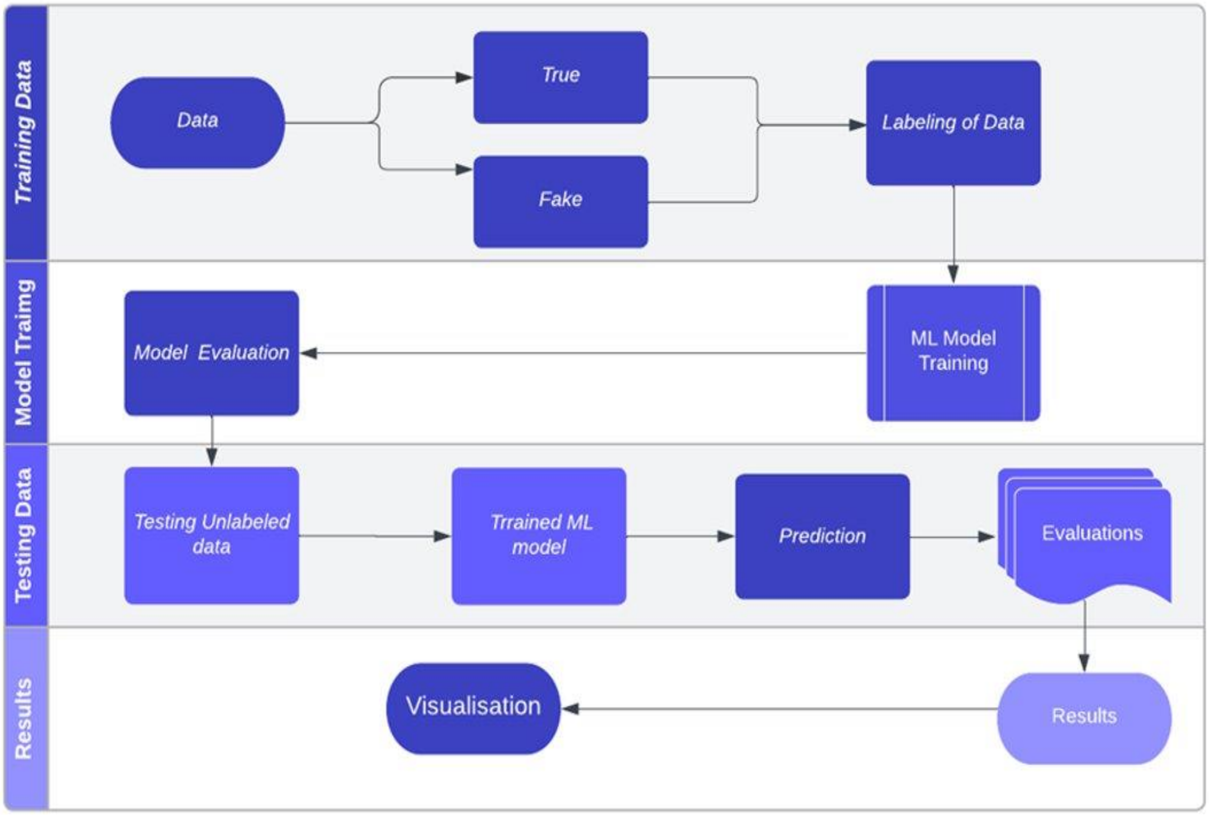


Fig-1: **Flowchart of the project**

1.1. WORKING PROCESS:

- Training data is labeled
- The Machine Learning models are trained with the labeled dataset.
- Model is evaluated.
- Testing dataset (unlabeled) is given to the Machine learning models to predict.
- Based on the predictions, results are evaluated.
- Visualization is done to know the insights of data.
- Model comparison based on visualization and results.

WORKING PROCESS

- Training data is labeled
- The Machine Learning models are trained with the labeled dataset.
- Model is evaluated.
- Testing dataset (unlabeled) is given to the Machine learning models to predict.
- Based on the predictions, results are evaluated.
- Visualization is done to know the insights of data.
- Model comparison based on visualization and results.

DATA COLLECTION

After finalizing a topic for our project study and conducted a thorough search for relevant data from various sources such as Kaggle and academic journals. We obtained a dataset from the Kaggle website which will be used for our analysis. The datasets are as follows:

Dataset on World news:

Fake News Detection Datasets | Kaggle

DATA PROCESSING

The dataset includes both actual and fraudulent news stories. The genuine articles were retrieved via crawling articles from the news website "Reuters.com" for this dataset, which was compiled from sources in the real world. The phony news pieces were gathered from a variety of sources. The false news stories were gathered from shady websites that Wikipedia and the American fact-checking group PolitiFact had identified as being unreliable. The dataset includes a variety of articles on various subjects, including tweets, however the majority of the articles cover political and international news themes.

Details of Dataset

1. **Title**: Title of the concerned news.
2. **Text**: The news content.
3. **Subject**: Category of news.
4. **Date:** Date on which news was published.

Data Processing:

- We have segregated the Kaggle dataset into two parts for both the Fake dataset and True dataset. One part of them is labeled and further used for training the ML models. Another part of both the datasets is concatenated and shuffled and used for the testing of the trained ML models.

```python
# Load the true news dataset
true_data = pd.read_csv(r'C:\Users\khazi\Desktop\archive\News _dataset\Truedata.csv')
true_data['label'] = 1

# Load the fake news dataset
fake_data = pd.read_csv(r'C:\Users\khazi\Desktop\archive\News _dataset\Fakedata.csv')
fake_data['label'] = 0

# Concatenate the datasets
data = pd.concat([true_data, fake_data], ignore_index=True)
data = data.sample(frac=1, random_state=42)
data
```

Out[1]:

	title	text	subject	date	label
13819	OBAMA'S WAR ON COPS Takes Toll On Black Commun...	Communities like Ferguson, Baltimore and now M...	left-news	Aug 15, 2016	0
2325	China's Chongqing vows purge of former leader'...	BEIJING (Reuters) - The southwestern Chinese m...	worldnews	September 30, 2017	1
18167	JAPANESE SCHOOLS DON'T EMPLOY JANITORS...Why Ame...	Watch NPR employee and Afghanistan refugee (wh...	left-news	Feb 15, 2017	0
2698	Frankfurt starts evacuation before attempt to ...	FRANKFURT (Reuters) - Frankfurt emergency serv...	worldnews	September 2, 2017	1
7883	Virginia Republicans seek to block governor's ...	(Reuters) - Virginia Republican leaders said t...	politicsNews	August 31, 2016	1
...	...	...	...	...	...
11964	Here's How The Bundy Terrorists Left The Malh...	On March 23, KOIN in Burns, Oregon, got a beh...	News	March 24, 2016	0
21575	THIS VIDEO WILL DESTROY "Black Lives Matter"...S...		left-news	Sep 22, 2016	0
5390	Catalan secessionists mull snap election as Ma...	MADRID (Reuters) - Secessionist leaders in Cat...	worldnews	October 24, 2017	1
860	Former Trump security aide was Russia blackmai...	WASHINGTON (Reuters) - Former Acting U.S. Atto...	politicsNews	May 8, 2017	1
15795	PAUL RYAN: American Companies Will Close Witho...	Paul Ryan is just like all the other establish...	politics	Dec 25, 2015	0

Fig-2: Dataset after cleaning and labeling

Feature Extraction:

The main concept behind every Machine learning model that we have used is extracting of features such as headlines, informal words etc. from training data and based on the extracted features making predictions on testing data. The concept which used in our models for feature extraction is TF-IDF Vectorizer

TF-IDF Vectorizer:

It accounts for the frequency of each word across all documents. Tf-Idf Vectorizer also computes a weight for each word based on how frequently it appears across all documents. This helps to give less weight to words that are common across all documents, such as "the", "and", "of", etc.

TF-IDF assigns a numerical value to each term in a document or corpus based on its frequency and importance in the text. The approach consists of two main components:

Term Frequency (TF): This measures the frequency of a term (word or phrase) in a document. It is calculated as the number of times a term appears in a document divided by the total number of terms in the document.

Inverse Document Frequency (IDF): This measures the importance of a term in the entire corpus of documents. It is calculated as the logarithm of the total number of documents divided by the number of documents that contain the term.

The final TF-IDF score for a term is obtained by multiplying its TF and IDF values. Terms with higher TF-IDF scores are considered more important and informative than terms with lower scores.

While there are other feature extraction methods available for text data, Tf-Idf Vectorizer is often considered to be one of the best methods due to its effectiveness and efficiency for handling high-dimensional, sparse data in text-based tasks.

MACHINE LEARNING MODELS AND TRAINING

Machine learning models for fake news detection typically involve analyzing the content of news articles and identifying patterns that are characteristic of fake news. These patterns can include various linguistic and contextual features that distinguish real news from fake news.

One common approach for fake news detection involves feature extraction, where various linguistic and contextual features are extracted from news articles. These features can include the frequency of certain words or phrases, the presence of named entities or dates, and the sentiment expressed in the article.

Once these features have been extracted, they are used to train a machine learning model. When the latest news data is inputted into the trained machine learning model, it analyzes the extracted features and predicts whether the article is real, or fake based on the patterns learned during training.

We have used some well-known Machine learning models for our fake news prediction which are widely accepted for classification and assumption tasks.

Random Forest Model: It is a popular machine learning algorithm used for classification, regression, and other tasks. It is an ensemble learning method that involves building multiple decision trees and combining their outputs to make a final prediction.

In the random forest algorithm, each decision tree is built using a random subset of the training data and a random subset of the features. This helps to reduce overfitting and improve the model's generalization ability.

To make a prediction for a new data point, the random forest algorithm feeds the data point through each decision tree and tallies up the number of times the data point is classified as each class. The final prediction is then based on the majority vote of the decision trees.

Support Vector Machine: The SVM algorithm involves finding a hyperplane that best separates the data into different classes. In a binary classification problem, this hyperplane separates the data into two classes by maximizing the margin, which is the distance between the hyperplane and the closest data points of each class. The data points closest to the hyperplane are called support vectors.

SVM can handle both linearly separable and non-linearly separable data by using a technique called kernel trick. In the kernel trick, the data is transformed into a higher-dimensional space, where it may become linearly separable. The SVM algorithm then finds the hyperplane in this higher-dimensional space.

Naïve Bayes: Naive Bayes assumes that the features in the data are independent of each other, given the class label. This is known as the "naive" assumption since it is often not true in practice. Despite this simplifying assumption, Naive Bayes can still perform well on a wide range of classification tasks.

To train a Naive Bayes classifier, the algorithm first calculates the probability of each class label given the training data. It then calculates the conditional probability of each feature given each class label. This is done using the training data, by counting the occurrences of each feature in each class.

When a new data point is inputted into the trained Naive Bayes classifier, the algorithm calculates the probability of each class label given the features in the data point. The final prediction is then based on the class label with the highest probability.

TESTING

TESTING DATA

For testing data we have taken half data from Fake news dataset and half data from the True dataset. Then these datasets have been concatenated to a single testing dataset. And the rows of this dataset have been shuffled within the same CSV.

```
# Predict the labels of the testing data using the trained SVM model
test_y_pred = model.predict(test_X)

test_y_pred
```

Out[54]: `array([0, 1, 0, ..., 1, 1, 0], dtype=int64)`

Labelling test data with predicted lable's

In [55]:

```
test_data['label'] = test_y_pred
test_data
```

Out[55]:

	title	text	subject	date	label
13821	One Crucial Detail On Budweiser's New 'Patrio...	The country woke up to confusion today as Budw...	News	May 10, 2016	0
2325	Immigration overhaul backers sing, dance, chan...	WASHINGTON (Reuters) - Dancing, singing, chant...	politicsNews	April 18, 2016	1
14115	BREAKING: 28 YR OLD PALESTINIAN MUSLIM Feras M...	28 year old Abdullah Faris Freatekh, aka Feras...	politics	Oct 12, 2016	0
2698	Senate blocks bill that would override state G...	(Reuters) - The U.S. Senate on Wednesday block...	politicsNews	March 16, 2016	1

VISUALIZATION

CONFUSION MATRIX

For visualization we have considered the number of instances where the model has failed to

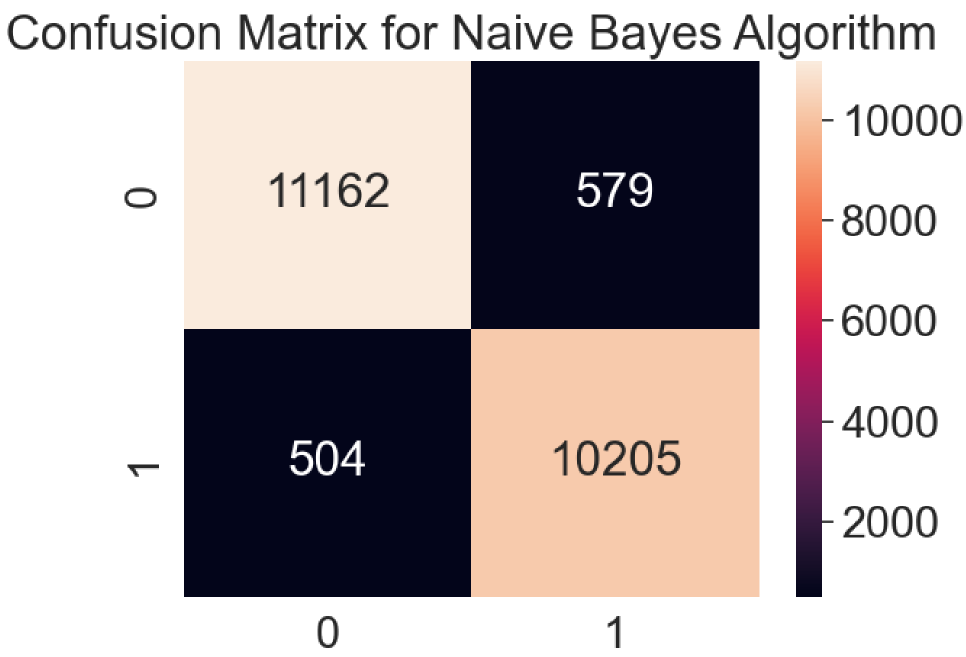


Fig: Confusion Matrix for Naïve bayes algorithm

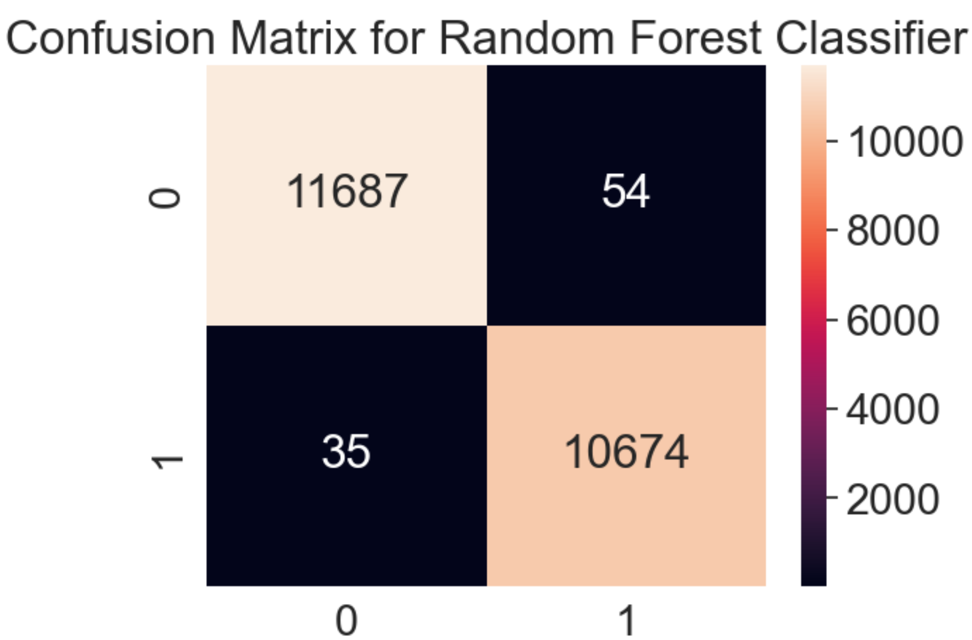


Fig: Confusion Matrix for Random Forest Classifier algorithm

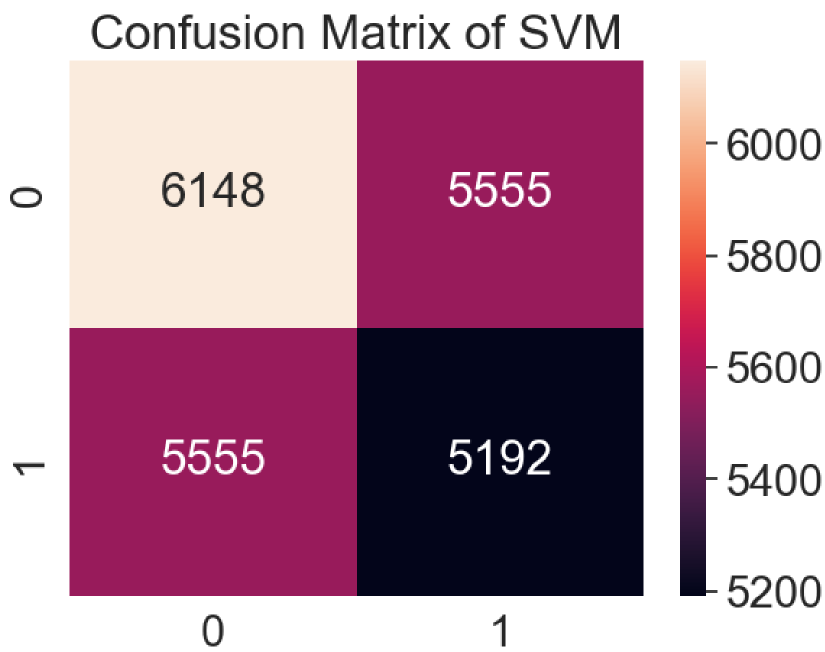


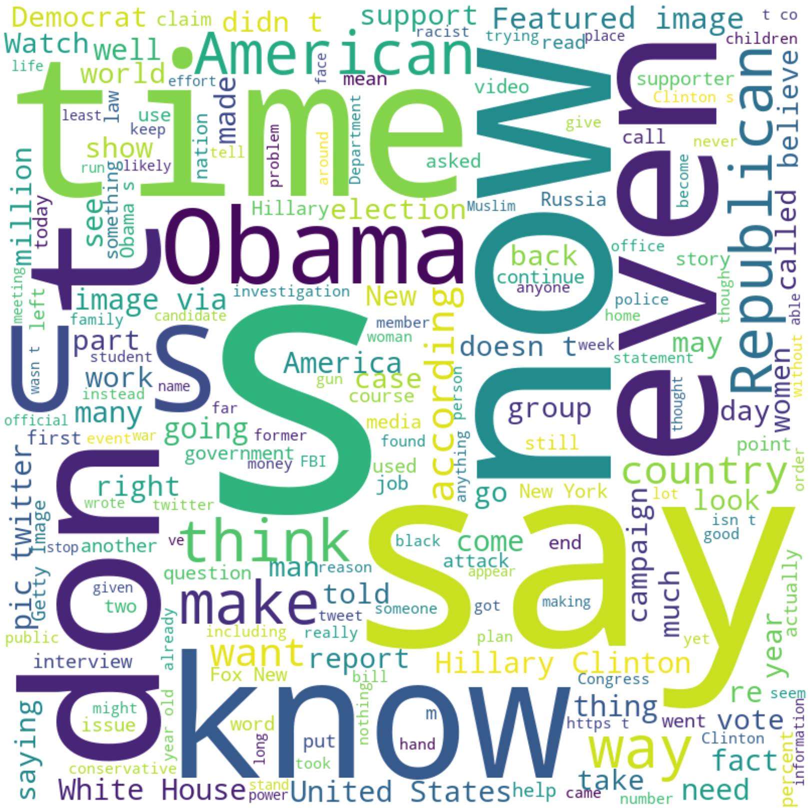
Fig: Confusion Matrix for Support Vector Machine (SVM) algorithm

predict the correct output. Based on the number of instances of the model, we created a heatmap and some other graphs to point out potential differences among the models and note the significant reasons for the difference in their performance.

TOP 10 WORDS

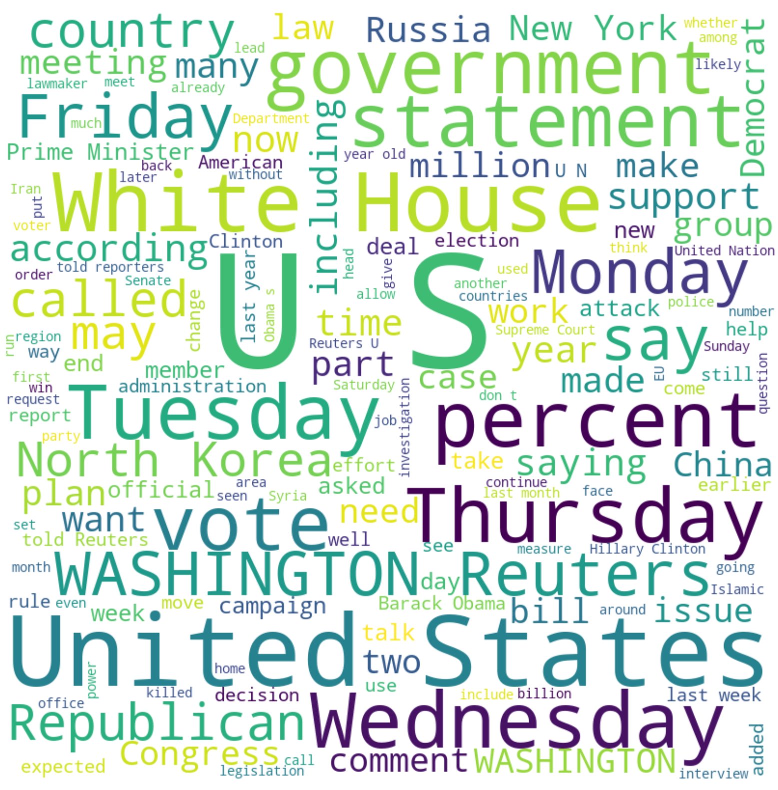
Common words after removing top 10 occuring words for the given data.

Fake news:



Democrat
claim
didn t
support
Featured
image
t co
Watch
well
American
racist
trying
read
place
children
life
time
now
even
world
face
mean
Republican
believe
effort
video
supporter
Clinton s
least
law
use
made
keep
give
show
nation
problem
Department
around
call
never
run
likely
tell
asked
become
million
something
Obama s
Hillary
election
Muslim
Russia
today
see
don t
Obama
back
office
story
called
continue
anyone
meeting
left
image via
investigation
New
police
though
family
candidate
member
home
able
U
S
part
woman
doesn t
week
may
wasn t
student
America
statement
without
work
case
gun
instead
name
course
person
women
official
many
far
group
day
first
event
war
going
media
found
thought
point
government
former
still
order
twitter
money
FBI
used
country
anything
job
according
right
go
New York
lot
look
wrote
twitter
image
campaign
isn t
stop
another
ve
think
black
end
good
attack
question
man
reason
appear
say
actually
pic
Getty
given
make
told
making
two
tweet
someone
got
much
public
already
including
really
plan
yet
year
interview
want
report
Hillary Clinton
know
Fox New
bill
Congress
re
seem
year old
might
thing
m
https t
went
vote
saying
issue
word
nothing
long
put
hand
way
Clinton
information
conservative
took
fact
percent
White House
stand
power
United States
help
came
take
number
need

True News:



country
law
Russia
New York
whether
lead
among
meeting
many
government
Democrat
lawmaker
meet
already
likely
Friday
much
Department
statement
now
Prime Minister
back
year old
later
American
without
million
U N
make
Iran
put
White House
support
voter
including
new
group
according
Clinton
deal
election
think
United Nation
U
S
order
told reporters
head
used
Monday
Senate
last year
give
another
called
change
allow
countries
police
number
Obama s
attack
run
region
may
time
work
help
say
way
Supreme Court
Reuters U
year
Sunday
end
member
part
question
first
administration
EU
case
made
still
win
Saturday
come
request
don t
report
job
Tuesday
percent
party
investigation
North Korea
effort
take
saying
China
plan
area
official
Syria
continue
seen
asked
last month
earlier
face
need
set
want
vote
Thursday
told Reuters
well
see
measure
Hillary Clinton
month
WASHINGTON
day
Reuters
going
rule
even
move
Barack Obama
Islamic
week
campaign
around
bill
issue
talk
United
States
two
home
power
use
killed
decision
include
billion
office
last week
Republican
Wednesday
added
Congress
call
expected
legislation
comment
WASHINGTON
interview

RESULTS

MODEL COMPARISON

Confusion Matrix: - Confusion matrix is a table comparing its predicted output values with the actual output values and it is used to evaluate the performance of a classification model. It is a common way of evaluating the accuracy of machine learning algorithms.

A confusion matrix consists of four outcomes, they are True Positive (TP), False Positive (FP), True Negative (TN), and False Negative (FN). Each of these results reflects a different case in which the classification of a specific occurrence was either properly or wrongly predicted by the model.

Evaluate the results of a classification model using a confusion matrix, we can calculate various performance metrics. The choice of performance metrics depends on the specific problem and the goals of the model.

Performance Metrics:

1. <u>Accuracy</u>: This metric calculates the percentage of correct predictions made by the model. Accuracy is calculated as

Accuracy = (TP + TN) / (FP + FN+ TP + TN).

2. <u>Precision</u>: This metric calculates the proportion of correctly predicted positive instances out of all instances predicted as positive. It is calculated as

Precision = TP / (TP + FP).

3. <u>Recall</u>: This metric calculates the proportion of correctly predicted positive instances out of all actual positive instances. It is calculated as

Recall = TP / (TP + FN).

4. <u>F1 score</u>: This metric is the harmonic mean of precision and recall. It provides a balanced measure of both metrics. It is calculated as

F1 score = 2 * (precision * recall) / (precision + recall).

These metrics provide different perspectives on the model's performance and can be used to compare different models or to choose the best model for a specific problem.

Misclassified Instances: For the machine learning model and its confusion matrix, Misclassified instances are the data points which were classified incorrectly by the model.

If the model predicts the news to be true even though its false and vice versa is a misclassified instance. Identifying misclassified instances is important as it gives us insights into the model's weaknesses and where to make improvements.

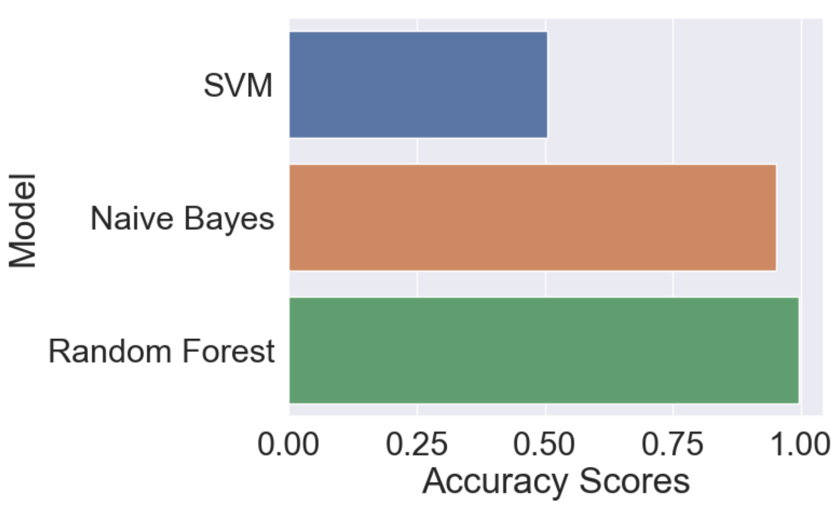
After identifying misclassified instances, the model can be further retrained, tuned and new features can be added to improve its performance and reduce the misclassified instances to happen.

Model Name	Accuracy	Precision	Recall	F1-Score	Misclassified instances
Support Vector Machine (SVM)	0.505	0.483	0.483	0.483	11110
Naïve Baye's	0.951	0.946	0.952	0.949	1083
Random Forest	0.995	0.994	0.997	0.995	89

Table: Performance comparison between 3 Machine Learning Models

FINDINGS

From the Metrics calculated above for all the models,



Support Vector Model: SVM classifier is at 52% Accuracy in predicting the type of news with more misclassified instances up to 11000+. By this we can say that SVM classifier has performed poorly in predicting Fake news and true news from the testing dataset. From 11703 fake news instances it has predicted 6148 news instances as fake and from 10747 true news instances only 5192 instances were predicted to be true. 48%-50% are falsely predicted by this model.

Naïve Bayes Model: The Naïve Bayes Model performed good in predicting the nature of the news with high accuracy and less misclassified instances. From 11703 fake news It has predicted 579 fake news instances as true, and from10747 true news instances 504 instances predicted as fake. It has an accuracy of 95% and 1000+ misclassified instances.

Random forest Model: From the metrics calculated, with high accuracy and precision Random Forest model has topped the chart when compared with Naïve Bayes model and SVM model. It has an accuracy of 99.6% which is higher in all three cases. Misclassified instances are 80+. It has predicted 35 true news instances as fake and 54 instances of fake news as true ones.

DISCUSSIONS

The elements for identifying fake news are explained in this project. Currently, we have tested three algorithms which are Naive Bayes, Support Vector Machine (SVM), and Random Forest Classifier. The results of this project usually show how accurate the machine learning models were at identifying whether news articles are real or fake. A high accuracy score shows that the models are working effectively and can be used to correctly identify whether articles are true or fake.

This information can be used to draw conclusions about the effectiveness of different machine learning algorithms for identifying fake news, the significance of feature engineering in enhancing classification performance, and the potential drawbacks of using a particular dataset or approach for fake news detection.

With more time given, this project might be enhanced with more time to incorporate a larger and more varied dataset as well as sophisticated natural language processing (NLP) techniques for feature engineering. The research might also be expanded to include real-time monitoring of news sources for the detection of false news like data from Twitter or Facebook data.

The study's shortcomings could relate to bias in the training data, restrictions on the dataset's size or representativeness, and potential overfitting of the machine learning models. When analyzing the findings of any initiative to identify fake news, it is crucial to carefully evaluate these constraints and potential causes of mistake.

Additionally, we considered using Twitter tweets as a dataset using Twitter API (Application Programming Interface) access, but sadly, only one member of our team was granted access. And only 1500 tweets could be extracted each month, which wasn't enough to train and test the models, so we used the Kaggle dataset instead.

REFRENCES

1. Baarir, N. F., & Djeffal, A. (2021, February). Fake news detection using machine learning. In 2020 2nd International Workshop on Human-Centric Smart Environments for Health and Well-being (IHSH) (pp. 125-130). IEEE.
2. Manzoor, S. I., & Singla, J. (2019, April). Fake news detection using machine learning approaches: A systematic review. In 2019 3rd international conference on trends in electronics and informatics (ICOEI) (pp. 230-234). IEEE.
3. Sharma, U., Saran, S., & Patil, S. M. (2020). Fake news detection using machine learning algorithms. International Journal of Creative Research Thoughts (IJCRT), 8(6), 509-518.
4. Ahmad, I., Yousaf, M., Yousaf, S., & Ahmad, M. O. (2020). Fake news detection using machine learning ensemble methods. Complexity, 2020, 1-11.
5. Jain, A., Shakya, A., Khatter, H., & Gupta, A. K. (2019, September). A smart system for fake news detection using machine learning. In 2019 International conference on issues and challenges in intelligent computing techniques (ICICT) (Vol. 1, pp. 1-4). IEEE.

INDIVIDUAL CONTRIBUTION TO THE PROJECT:

Aditya Chettipalli

- Looked through several websites in search of the project idea and discovered the Kaggle dataset.
- Worked on labeling, categorizing, and structuring the dataset's data.
- Contributed to the creation of the presentations and of the final report's results and discussions section.

Ahteshamuddin Mohammed

- Worked with Aditya on data processing and aided with data cleaning.
- Worked in model testing and evaluation of the outcomes using various metrics.
- Assisting with various areas of the final report and preparing the presentations.

Ganesh Gude

- Collaborated with Ahteshamuddin in testing the model and analyzing the results.
- Worked together with Nasiruddin to visualize the results.
- Contributed to the creation of presentations and the Introduction, background, and literature review sections in the final report.

Nasiruddin Khazi

- Worked on Model training and evaluation for all three machine learning models.
- Made use of various Python packages to visualize the outcomes and interpretation of results.