Fake News Detection

ABSTRACT

Misinformation presents a huge challenge in online society. As a result, there have been many attempts to identify and classify misinformation. Specifically, in social networking sites, blogs, as well as online newspapers.

With the widespread dissemination of information via digital media platforms, it is of utmost importance for individuals and societies to be able to judge the credibility of it. Fake news is not a recent concept, but it is a commonly occurring phenomenon in current times. The consequence of fake news can range from being merely annoying to influencing and misleading societies or even nations. A variety of approaches exist to identify fake news. By conducting a systematic literature review, we identify the main approaches currently available to identify fake news and how these approaches can be applied in different situations. Some approaches are illustrated with a relevant example as well as the challenges and the appropriate context in which the specific approach can be used.

INTRODUCTION

Paskin defines fake news as "particular news articles that originate either on mainstream media (online or offline) or social media and have no factual basis, but are presented as facts and not satire". The importance of combatting fake news is starkly illustrated during the COVID-19 pandemic. Social networks are stepping up in using digital fake news detection tools and educating the public towards spotting fake news.

At the time of writing, Facebook uses machine learning algorithms to identify false or sensational claims used in advertising for alternative cures, they place potential fake news articles lower in the news feed, and they provide users with tips on how to identify fake news themselves (Sparks and Frishberg 2020).

These measures are possible because different approaches exist that assist the detection of fake news. For example, platforms based on machine learning use fake news from the biggest media outlets, to refine algorithms for identifying fake news (Macaulay 2018).

What is fake news and how it works?

Fake news is not a new concept. Before the era of digital technology, it was spread through mainly yellow journalism with focus on sensational news such as crime, gossip, disasters and satirical news (Stein-Smith 2017). Since anyone can publish articles via digital media platforms, online news articles include well researched pieces but also opinion-based arguments or simply false information (Burkhardt 2017). There is no custodian of credibility standards for information on these platforms making the spread of fake news possible. To make things worse, it is by no means straightforward telling the difference between real news and semi-true or false news (Pérez-Rosas et al. 2018).

The nature of social media makes it easy to spread fake news, as a user potentially sends fake news articles to friends, who then send it again to their friends and so on. Comments

on fake news sometimes fuel its 'credibility' which can lead to rapid sharing resulting in further fake news (Albright 2017).

Social bots are also responsible for the spreading of fake news. Bots are sometimes used to target super-users by adding replies and mentions to posts. Humans are manipulated through these actions to share the fake news articles (Shao et al. 2018).

In today's world, it is normal to receive news from online sources like social media. News is often subjective to readers. We often choose to ingest content that appeals to the different emotions we have. So, considering this, the information that gets the most reach may not be real or accurate news. Additionally, real news may be twisted in transmission. A reader may end up with different versions of the same news. This may lead to information overload.

Why should one care?

A fake news article is designed to outrage and shock, causing some readers to share it on Facebook, Twitter, or another type of social media platform without questioning it. Sharing the article exposes it to more people who may be outraged by it, who also share it without question, and so on. This cycle continues until a sizeable number of people believe this fake story is the truth.

Exposure to misinformation can reduce trust in the media more broadly, making it tougher to know what fact or fiction in the future is. When we start to believe that there is the possibility that anything can be fake, it's easier to discount what is actually true. This presents a real concern about the impact of fake news on our children and young people.

According to the National Literacy Trust Fake News and Critical Literacy Report more than half of 12-15 year-olds go to social media as their regular source of news. And while only a third believe that social media stories are truthful, it is estimated that only 2% of school children have the basic critical literacy skills to tell the difference between real and fake news.

Half of the children asked, admitted being worried about fake news. Teachers surveyed on the matter noted a real increase in issues of anxiety, self-esteem, and a general skewing of world views. Generally, the trust children have in the news, social media interactions, and politicians being reliable sources is weakening.

Some fake stories can have a real impact on the lives of our children. The so-called "Anti-vaxxers" movement, the fake Momo scare, and the recent false news stories around the COVID-19 pandemic are all examples of different ways that fake news preys on our emotions and those of our children.

Children interviewed express a concern that when online they don't know who to trust, what is real, and which forms of knowledge are true. Nearly all children are now online, but many of them are not emotionally equipped to deal with the challenges of a fake news online culture. We cannot stop our kids using the internet nor should we, it is an incredible resource. It is important then that we teach them some basic rules so they can feel confident in the facts they find online.

What's being done about it?

Online giants like Google and Facebook are attempting to crack down on fake news by banning suspicious sites from advertising on their platforms and asking users to report dishonest articles. However, many critics feel that Google, Facebook, and other online services still aren't doing enough.

Unfortunately, fake news writers will likely continue to create new sites and methods to get around any digital roadblocks. This means the best way to prevent fake news from spreading is to teach users how to identify fake news themselves. That means you!

OBJECTIVE:

- The main objective is to detect the fake news, which is a classic text classification problem with a straight forward proposition. It is needed to build a model that can differentiate between "Real" news and "Fake" news. This is a binary outcome.

  Class 1, True News
  Class 0, Fake News

- Experiment with various Classification Models and see which yields greatest accuracy.

- Examine trends & correlations within our data

- Determine which features are important to identify fake news.

FEATURES AND PREDICTOR:

Our Predictor (Y, Fake News or not) is determined by 4 features (X):

1. Title: The title of the news

2. Text: The entire news

3. Subject: The type of news- politicsNews, worldnews, News, politics, left-news, Government News, US_News, Middle-east

4. Date: Date of the article

EXISTING SYSTEM

There exists a large body of research on the topic of machine learning methods for deception detection, most of it has been focusing on classifying online reviews and publicly available social media posts. Particularly since late 2016 during the American Presidential election, the question of determining 'fake news' has also been the subject of particular attention within the literature.

Conroy, Rubin, and Chen outlines several approaches that seem promising towards the aim of perfectly classify the misleading articles. They note that simple content-related n-grams and shallow parts-of-speech (POS) tagging have proven insufficient for the classification task, often failing to account for important context information. Rather, these methods have been shown useful only in tandem with more complex methods of analysis. Feng, Banerjee, and Choi [2] are able to achieve 85%-91% accuracy in deception related classification tasks using online review corpora.

Feng and Hirst implemented a semantic analysis looking at 'object:descriptor' pairs for contradictions with the text on top of Feng's initial deep syntax model for additional improvement. Rubin, Lukoianova and Tatiana analyze rhetorical structure using a vector space model with similar success. Ciampaglia et al. employ language pattern similarity networks requiring a pre-existing knowledge base.

PROPOSED SYSTEM

In this project several models are built based on the count vectorizer(i.e) word tallies relatives to how often they are used in other articles in your dataset) can help . Since this problem is a kind of text classification, implementing a Naive Bayes classifier will be best as this is standard for text-based processing. However I have tried to use other Classification models to determine if they detect fake news correctly or not. Since articles from different domains have a unique textual structure, it is difficult to train a generic algorithm that works best on all particular news domains. So in this project, I propose a solution to the fake news detection problem using the machine learning ensemble approach. Our study explores different textual properties that could be used to distinguish fake contents from real. By using those properties, I train a combination of different machine learning algorithms using various ensemble methods that are not thoroughly explored in the current literature. These techniques facilitate the training of different machine learning algorithms in an effective and efficient manner. The results validate the improved performance of our proposed technique using the 4 commonly used performance metrics (namely, accuracy, precision, recall, and F-1 score).

METHODOLOGY:

Data Cleaning and Pre-processing: There were 2 datasets present in Kaggle. One is for true news and the other is for fake news. I merged both the datasets to work on them. The website contains unfiltered data which must be filtered before the final data set can be used to train the model. In data cleaning step, first we check if there are any missing or junk values in the dataset for which we used the isnull() function.

Exploratory Data Analysis: Exploratory data analysis (EDA) is used to analyse and investigate the dataset and summarize its main characteristics if a news is Fake or not by employing data visualization methods. It helps determine the different subjects of the news based on the count and the comparison of true news and fake news.

Also used regular expressions (re) to treat text in a user defined function process.

MACHINE LEARNING ALGORITHMS:

Machine learning is an important component of the growing field of data science. Through the use of statistical methods, algorithms are trained to make classifications or predictions, and to uncover key insights in data mining projects. These insights subsequently drive decision making within applications and businesses, ideally impacting key growth metrics. Below are the various Machine Learning Models I implemented in my Project:

a) Logistic Regression:

Logistic regression is often used a lot of times in machine learning for predicting the likelihood of response attributes when a set of explanatory independent attributes are given. It is used when the target attribute is also known as a dependent variable having categorical values like yes/no or true/false, etc. It's widely used for solving classification problems. It falls under the category of supervised machine learning. It efficiently solves linear and binary classification problems. It is one of the most commonly used and easy to implement algorithms. It's a statistical technique to predict classes which are binary. When the target variable has two possible classes in that case it predicts the likelihood of occurrence of the event. In our dataset the target variable is categorical as it has only two classes-0 or 1.

b) Random Forest:

Random Forest is the most famous and it is considered as the best algorithm for machine learning. It is a supervised learning algorithm. To achieve more accurate and consistent prediction, random forest creates several decision trees and combines them together. The major benefit of using it is its ability to solve both regression and classification issues. When building each individual tree, it employs bagging and feature randomness in order to produce an uncorrelated tree forest whose collective forecast has much better accuracy than any individual tree's prediction. Bagging enhances accuracy of machine learning methods by grouping them together. In this algorithm, during the splitting of nodes it takes only random subset of nodes into an account. When splitting a node, it looks for the best feature from a random group of features rather than the most significant feature. This results into getting better accuracy. It efficiently deals with the huge datasets. It also solves the issue of overfitting in datasets. It works as follows: First, it'll select random samples from the provided dataset. Next, for every selected sample it'll create a decision tree and it'll receive a forecasted result from every created decision tree. Then for each result which was predicted, it'll perform voting and through voting it will select the best predicted result.

c) Decision Tree:

Decision Tree is a supervised learning technique that can be used for both classification and Regression problems, but mostly it is preferred for solving Classification problems. It is a tree-structured classifier, where internal nodes represent the features of a dataset, branches represent the decision rules and each leaf node represents the outcome.

In a Decision tree, there are two nodes, which are the Decision Node and Leaf Node. Decision nodes are used to make any decision and have multiple branches, whereas Leaf nodes are the output of those decisions and do not contain any further branches.
The decisions or the test are performed on the basis of features of the given dataset. It is a graphical representation for getting all the possible solutions to a problem/decision based on given conditions. It is called a decision tree because, similar to a tree, it starts with the root node, which expands on further branches and constructs a tree-like structure. In order to build a tree, we use the CART algorithm, which stands for Classification and Regression Tree algorithm. A decision tree simply asks a question, and based on the answer (Yes/No), it further split the tree into subtrees.

d) Naive Bayes:

It is a probabilistic machine learning algorithm which is mainly used in classification problems. It's based on Bayes theorem. It is simple and easy to build. It deals with huge datasets efficiently. It can solve complicated classification problems. The existence of a specific feature in a class is assumed to be independent of the presence of any other feature according to naïve Bayes theorem. It's formula is as follows : $P(S|T) = P(T|S) * P(S) / P(T)$ Here, T is the event to be predicted, S is the class value for an event. This equation will find out the class in which the expected feature for classification exists.

IMPLEMENTATION STEPS:

As we already discussed in the methodology section about some of the implementation methods. So, the language used in this project is Python programming.

We're running python code in Anaconda navigator's Jupyter notebook. Jupyter notebook is much faster than Python IDE tools like PyCharm or Visual studio for implementing ML algorithms. The advantage of Jupyter notebook is that it's really helpful for Data visualization and plotting some advance graphs like histogram and heatmap of correlated matrices. Let's revise implementation steps:

a) Dataset collection: The dataset is downloaded from Kaggle. https://www.kaggle.com/datasets/clmentbisaillon/fake-and-real-news-dataset

b) Importing Libraries: Numpy, Pandas, Scikit-learn, Matplotlib and Seaborn libraries are used.

c) Exploratory Data Analysis: To get more insights from data.

d) Data cleaning and pre-processing: Merged both datasets and checked for null and junk values using isnull ().

e) Feature Conversion: In this step, I converted my text data into Vectors using CountVectorizer

f) Model selection: First x is separated from y. X's are features or input variables of the dataset and y is dependent or target variable which is crucial to determine fake news. Then by importing model_selection function of the sklearn library, I split our x and y into train and test split using train_test_split () function of sklearn and kept 75% of the data for training and 25% for testing.

g) Applied ML models: In this step different models were applied on the same train and test data to determine their accuracy and created a confusion matrix of all models.

h) Model Testing: In this step I randomly input few news data to determine if the system is predicting it accurately or not

i) Deployment of all the models to determine if a news is fake or not

LIBRARIES

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
import warnings
warnings.filterwarnings('ignore')
from sklearn.model_selection import train_test_split
from sklearn.feature_extraction.text import CountVectorizer
from sklearn.linear_model import LogisticRegression
from sklearn.tree import DecisionTreeClassifier
from sklearn.ensemble import RandomForestClassifier
from sklearn.naive_bayes import MultinomialNB
from sklearn.metrics import classification_report,accuracy_score,plot_confusion_matrix
```

DATA WRANGLING

Reading datasets

```
df_true = pd.read_csv("True.csv")
df_fake = pd.read_csv("Fake.csv")
```

Inserting a column "class" as target feature

```
df_true["class"] = 1
df_fake["class"] = 0
```

Merging both dataframes

```
df=pd.concat([df_true ,df_fake ],axis=0 )
df
```

```
df.head()
```

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

```
df.shape
```

```
(44898, 5)
```

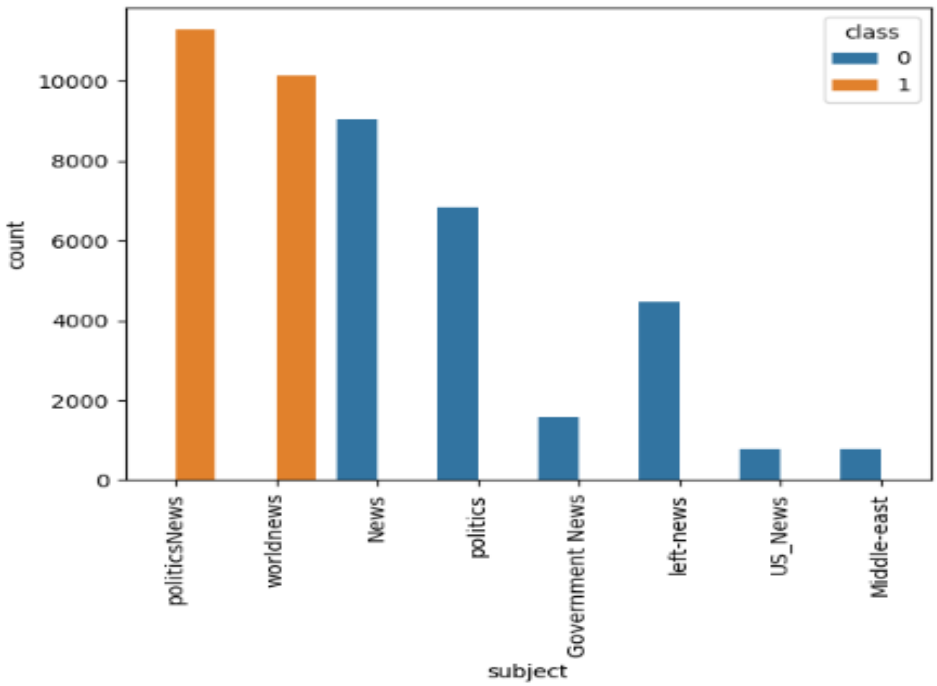
Checked Null Values

```
df.isnull().sum()
```

```
title      0
text       0
subject    0
date       0
class      0
dtype: int64
```

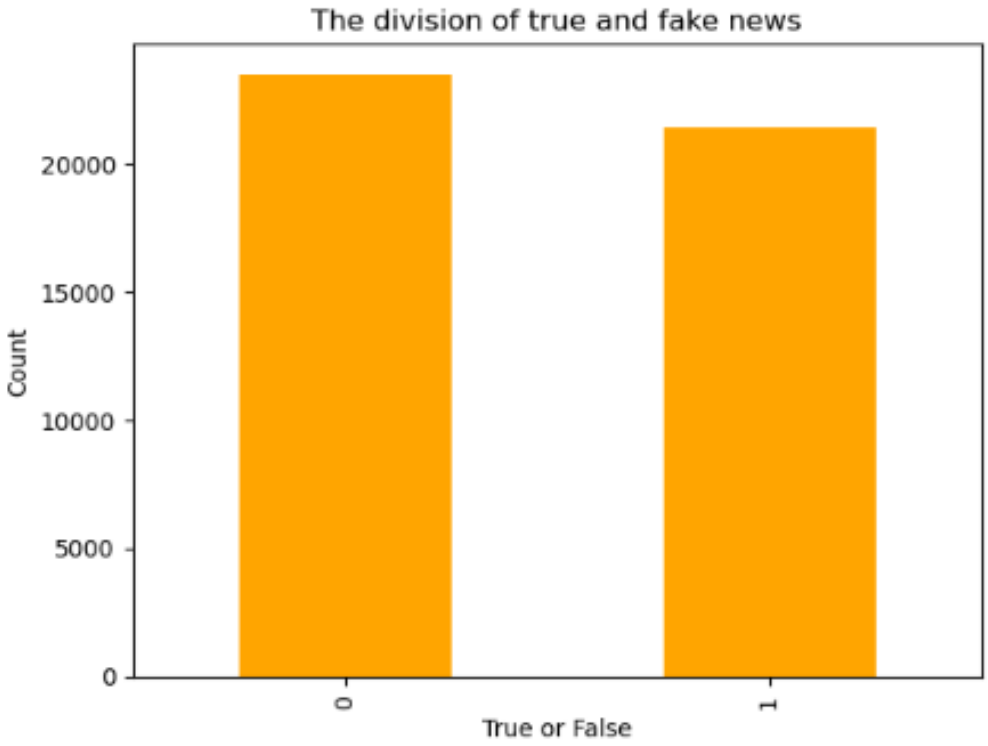
EXPLORATORY DATA ANALYSIS

```
sns.countplot(x='subject', hue='class',data=df)
plt.xticks(rotation=90)
plt.show()
```



Observation: We can see that the dataset consists mainly of True Political News. However the regular news and politics news have the highest number of fake news.

```
plt.title('The division of true and fake news')
plt.xlabel('True or False')
plt.ylabel('Count')
df['class'].value_counts().plot(kind='bar',color='orange')
plt.show
```



Observation: The count of fake news is higher than true news.

Removing unnecessary columns

```
df.drop(['subject','date'],axis=1,inplace=True)
df
```

Merging both columns title and text

```
df['text'] = df['title'] + " " + df['text']
df.drop(['title'],axis=1,inplace=True)
df
```

Creating a function to process the texts

```
import re
import string
def process(text):
    text = text.lower()
    text = re.sub('\[.*?\]', '', text)
    text = re.sub("\\W"," ",text)
    text = re.sub('https?://\S+|www\.\S+', '', text)
    text = re.sub('<.*?>+', '', text)
    text = re.sub('[%s]' % re.escape(string.punctuation), '', text)
    text = re.sub('\n', '', text)
    text = re.sub('\w*\d\w*', '', text)
    return text
```

MACHINE LEARNING AND PREDICTIVE ANALYTICS:

Prepare the data

To prepare data for modelling, just remember ASN (Assign,Split, Normalize).

- Assign text feature to x, & the class column to our classification predictor,y.

x=df['text']

y=df['class']

- Split: the data set into the Training set and Test set

from sklearn.model_selection import train_test_split
x_train,x_test,y_train,y_test=train_test_split(X,y,test_size=0.25)

Converting text to vectors

```
from sklearn.feature_extraction.text import CountVectorizer
```

```
cv = CountVectorizer()
```

```
x_train_cv = cv.fit_transform(x_train.values)
x_test_cv = cv.transform(x_test)
x_test_cv.toarray()
```

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

IMPLEMENTING MACHINE LEARNING MODELS

We will now train various Classification Models on the Training set & see which yields the highest accuracy.

We will compare the accuracy of the below algorithms:

Logistic Regression

Decision Trees

Random Forest

Naive Bayes Classifier.

Note: these are all supervised learning models.

Logistic Regression

```
from sklearn.linear_model import LogisticRegression
model_lr = LogisticRegression()
model_lr.fit(x_train_cv,y_train)
```

```
LogisticRegression()
```

```
pred_lr=model_lr.predict(x_test_cv)
```

```
print(model_lr.score(x_train_cv,y_train))
print(model_lr.score(x_test_cv, y_test))
```

```
0.9999703026163395
0.9968819599109131
```

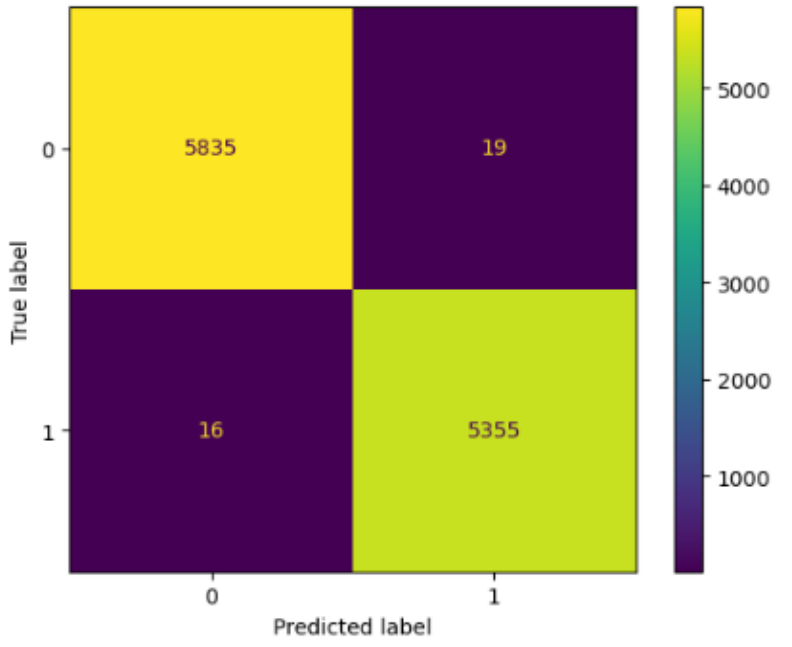
```
from sklearn.metrics import classification_report,accuracy_score,plot_confusion_matrix
print(classification_report(y_test, pred_lr))
```

```
              precision    recall  f1-score   support

           0       1.00      1.00      1.00      5854
           1       1.00      1.00      1.00      5371

    accuracy                           1.00     11225
   macro avg       1.00      1.00      1.00     11225
weighted avg       1.00      1.00      1.00     11225
```

```
plot_confusion_matrix(model_lr,x_test_cv,y_test)
plt.show()
```



Accuracy- 99.68%

Decision Tree Classifier

```
from sklearn.tree import DecisionTreeClassifier

model_dt = DecisionTreeClassifier()
model_dt.fit(x_train_cv, y_train)
```

```
DecisionTreeClassifier()
```

```
pred_dt=model_dt.predict(x_test_cv)
```

```
print(model_dt.score(x_train_cv,y_train))
print(model_dt.score(x_test_cv, y_test))
```

```
1.0
0.9971492204899778
```

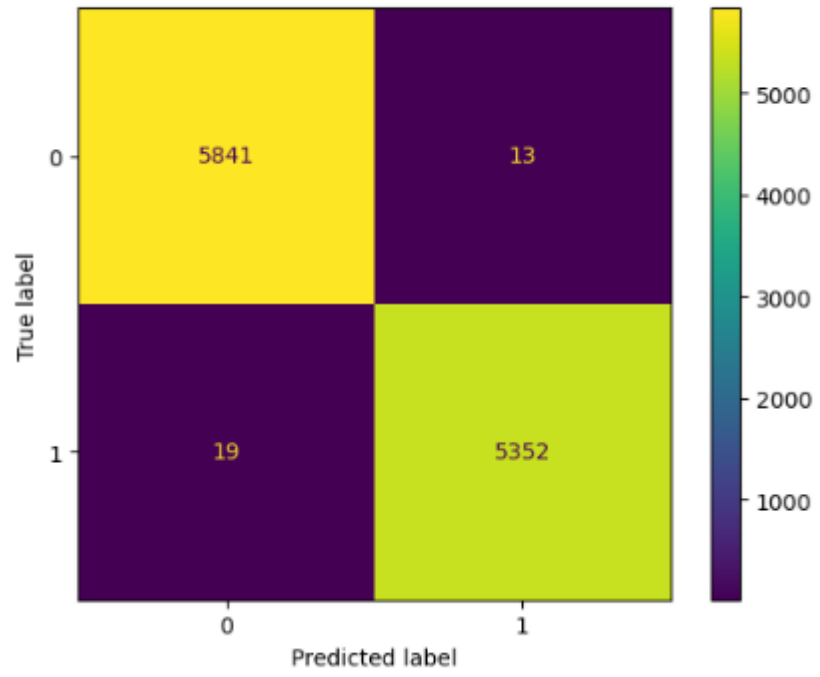
```
print(classification_report(y_test, pred_dt))
```

```
              precision    recall  f1-score   support

           0       1.00      1.00      1.00      5854
           1       1.00      1.00      1.00      5371

    accuracy                           1.00     11225
   macro avg       1.00      1.00      1.00     11225
weighted avg       1.00      1.00      1.00     11225
```

```
plot_confusion_matrix(model_dt,x_test_cv,y_test)
plt.show()
```



Accuracy- 99.7%

Random Forest Classifier

```
from sklearn.ensemble import RandomForestClassifier

model_rf = RandomForestClassifier(random_state=0)
model_rf.fit(x_train_cv, y_train)
```

```
RandomForestClassifier(random_state=0)
```

```
pred_rf=model_rf.predict(x_test_cv)
```

```
print(model_rf.score(x_train_cv,y_train))
print(model_rf.score(x_test_cv, y_test))
```

```
1.0
0.98913140311804
```

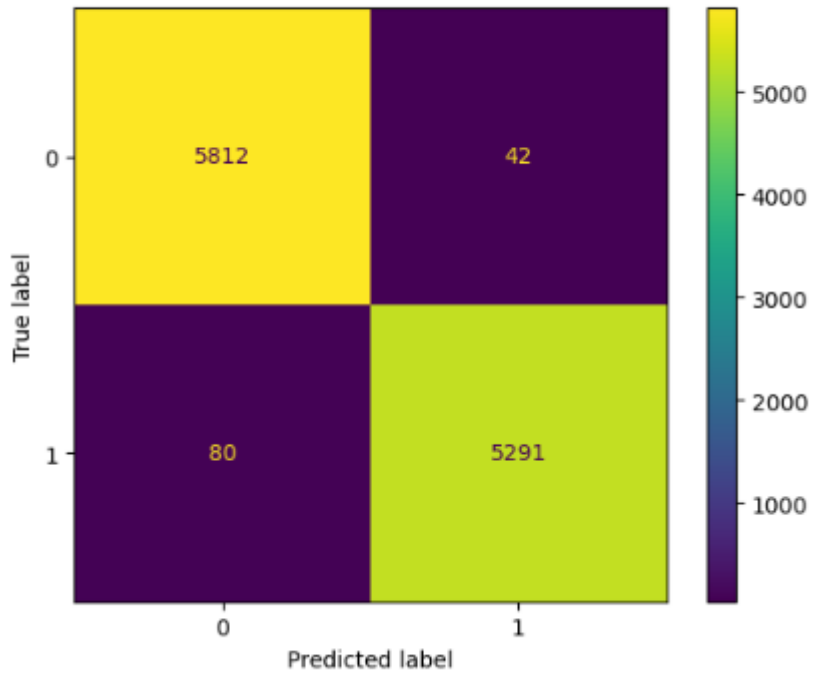
```
print(classification_report(y_test, pred_rf))
```

```
              precision    recall  f1-score   support

           0       0.99      0.99      0.99      5854
           1       0.99      0.99      0.99      5371

    accuracy                           0.99     11225
   macro avg       0.99      0.99      0.99     11225
weighted avg       0.99      0.99      0.99     11225
```

```
plot_confusion_matrix(model_rf,x_test_cv,y_test)
plt.show()
```



Accuracy-98.91%

Multinomial Naive Bayes

```
from sklearn.naive_bayes import MultinomialNB
```

```
model_nb = MultinomialNB()
model_nb.fit(x_train_cv,y_train)
```

```
MultinomialNB()
```

```
pred_nb=model_nb.predict(x_test_cv)
```

```
print(model_nb.score(x_train_cv,y_train))
print(model_nb.score(x_test_cv, y_test))
```

```
0.9633831259466041
0.951804008908686
```

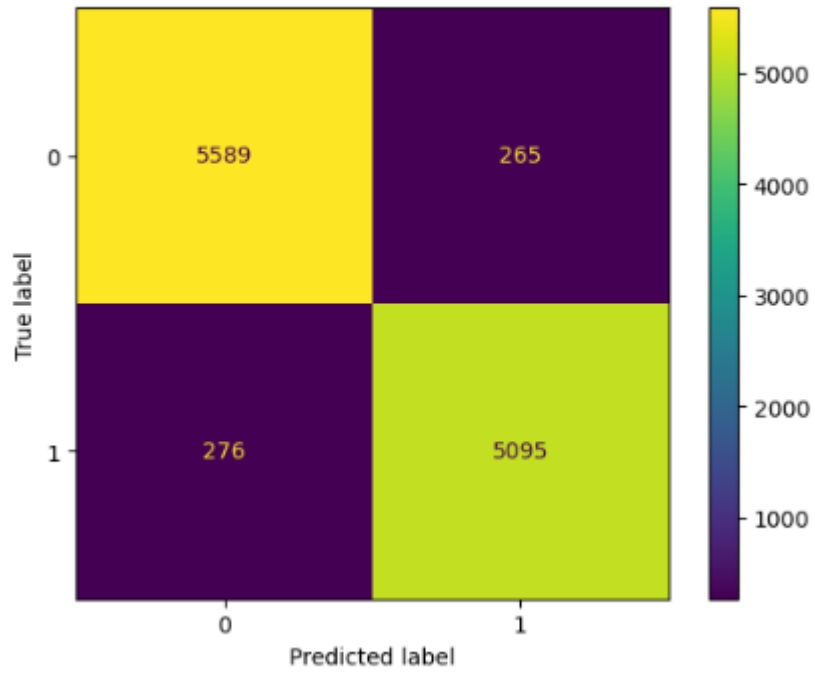
```
print(classification_report(y_test, pred_nb))
```

```
              precision    recall  f1-score   support

           0       0.95      0.95      0.95      5854
           1       0.95      0.95      0.95      5371

    accuracy                           0.95     11225
   macro avg       0.95      0.95      0.95     11225
weighted avg       0.95      0.95      0.95     11225
```

```
plot_confusion_matrix(model_nb,x_test_cv,y_test)
plt.show()
```



Accuracy-95.18

From comparing the 4 models, we can conclude that Model 2: Decision Tree yields the highest accuracy. With an accuracy of 99.7%.

We have precision, recall, f1-score and support for all the models used in the analysis.

Precision: "how many are correctly classified among that class"

Recall: "how many of this class you find over the whole number of element of this class"

F1-score: harmonic mean of precision and recall values. F1 score reaches its best value at 1 and worst value at 0. F1 Score = 2 x ((precision x recall) / (precision + recall))

Support: of samples of the true response that lie in that class.

MODEL TESTING

```
def output(n):
    if n == 0:
        return "Fake News"
    else:
        return "Not A Fake News"

def testing(news):
    testing_news = {"text":[news]}
    new_def_test = pd.DataFrame(testing_news)
    new_def_test["text"] = new_def_test["text"].apply(process)
    new_x_test = new_def_test["text"]
    new_cv_test = cv.transform(new_x_test)
    pred_lr = model_lr.predict(new_cv_test)
    pred_dt = model_dt.predict(new_cv_test)
    pred_rf = model_rf.predict(new_cv_test)
    pred_nb = model_nb.predict(new_cv_test)

    return print("\n\nLR Prediction: {} \nDT Prediction: {} \nRF Prediction: {} \nNB Prediction: {}".format(output(pred_lr[0]),
```

While there are other ways of measuring model performance (precision, recall, F1 Score, ROC Curve, etc), let's keep this simple and use accuracy as our metric.

To do this are going to see how the model performs on new data (test set) Accuracy is defined as: (fraction of correct predictions): correct predictions / total number of data points

ANALYSIS OF THE RESULT

We used precision, F1-score, recall and accuracy evaluation metrics for evaluating our models.

False Positive (FP) is when a model incorrectly predicts a positive outcome.

False Negative (FN) is when a model incorrectly predicts the negative outcome.

True Positive (TP) is when model correctly predicts a positive outcome.

True Negative (TN) is when a model correctly predicts a negative outcome.

$$Precision=TP/(TP+FP)$$

$$Recall = TP / (TP + FN)$$

Machine Learning Models	Accuracy
Logistic Regression	99.68%
Decision Tree	99.7%
Random Forest	98.91%
Naive Bayes	95.18%

PREDICTIONS:

Scenario: True and fake news articles were passed through this system to detect if it's predicting correctly or not using all machine learning models and they all determined correctly.

```
news = str(input())
testing(news)
```

```
BRUSSELS (Reuters) - NATO allies on Tuesday welcomed President Donald Trump s decision to commit more forces to Afghanistan, as
part of a new U.S. strategy he said would require more troops and funding from America s partners. Having run for the White Hou
se last year on a pledge to withdraw swiftly from Afghanistan, Trump reversed course on Monday and promised a stepped-up milita
ry campaign against Taliban insurgents, saying: Our troops will fight to win . U.S. officials said he had signed off on plan
s to send about 4,000 more U.S. troops to add to the roughly 8,400 now deployed in Afghanistan. But his speech did not define b
enchmarks for successfully ending the war that began with the U.S.-led invasion of Afghanistan in 2001, and which he acknowledg
ed had required an extraordinary sacrifice of blood and treasure . We will ask our NATO allies and global partners to suppor
t our new strategy, with additional troops and funding increases in line with our own. We are confident they will, Trump said.
That comment signaled he would further increase pressure on U.S. partners who have already been jolted by his repeated demands
to step up their contributions to NATO and his description of the alliance as obsolete - even though, since taking office, he
has said this is no longer the case. NATO Secretary General Jens Stoltenberg said in a statement: NATO remains fully committed
to Afghanistan and I am looking forward to discussing the way ahead with (Defense) Secretary (James) Mattis and our Allies and
international partners. NATO has 12,000 troops in Afghanistan, and 15 countries have pledged more, Stoltenberg said. Britain,
a leading NATO member, called the U.S. commitment very welcome . In my call with Secretary Mattis yesterday we agreed that de
spite the challenges, we have to stay the course in Afghanistan to help build up its fragile democracy and reduce the terrorist
threat to the West, Defence Secretary Michael Fallon said. Germany, which has borne the brunt of Trump s criticism over the s
cale of its defense spending, also welcomed the new U.S. plan. Our continued commitment is necessary on the path to stabilizin
g the country, a government spokeswoman said. In June, European allies had already pledged more troops but had not given detai
ls on numbers, waiting for the Trump administration to outline its strategy for the region.Nearly 16 years after the U.S.-led i
nvasion - a response to the Sept. 11 attacks which were planned by al Qaeda leader Osama bin Laden from Afghanistan - the count
ry is still struggling with weak central government and a Taliban insurgency. Trump said he shared the frustration of the Ameri
can people who were weary of war without victory , but a hasty withdrawal would create a vacuum for groups like Islamic State
and al Qaeda to fill.


LR Prediction: Not A Fake News
DT Prediction: Not A Fake News
RF Prediction: Not A Fake News
NB Prediction: Not A Fake News
```

```
news = str(input())
testing(news)
```

```
Vic Bishop Waking TimesOur reality is carefully constructed by powerful corporate, political and special interest sources in or
der to covertly sway public opinion. Blatant lies are often televised regarding terrorism, food, war, health, etc. They are fas
click to scroll output; double click to hide ion viewers to accept what have become destructive societal norms.The practice of manip
ulating and controlling public opinion with distorted media messages has become so common that there is a whole industry formed
around this. The entire role of this brainwashing industry is to figure out how to spin information to journalists, similar to
the lobbying of government. It is never really clear just how much truth the journalists receive because the news industry has
become complacent. The messages that it presents are shaped by corporate powers who often spend millions on advertising with th
e six conglomerates that own 90% of the media:General Electric (GE), News-Corp, Disney, Viacom, Time Warner, and CBS. Yet, thes
e corporations function under many different brands, such as FOX, ABC, CNN, Comcast, Wall Street Journal, etc, giving people th
e perception of choice As Tavistock s researchers showed, it was important that the victims of mass brainwashing not be aware
that their environment was being controlled; there should thus be a vast number of sources for information, whose messages coul
d be varied slightly, so as to mask the sense of external control. ~ Specialist of mass brainwashing, L. WolfeNew Brainwashing
Tactic Called AstroturfWith alternative media on the rise, the propaganda machine continues to expand. Below is a video of Shar
yl Attkisson, investigative reporter with CBS, during which she explains how astroturf, or fake grassroots movements, are use
d to spin information not only to influence journalists but to sway public opinion. Astroturf is a perversion of grassroots. As
troturf is when political, corporate or other special interests disguise themselves and publish blogs, start facebook and twitt
er accounts, publish ads, letters to the editor, or simply post comments online, to try to fool you into thinking an independen
t or grassroots movement is speaking. ~ Sharyl Attkisson, Investigative ReporterHow do you separate fact from fiction? Sharyl A
ttkisson finishes her talk with some insights on how to identify signs of propaganda and astroturfing These methods are used t
o give people the impression that there is widespread support for an agenda, when, in reality, one may not exist. Astroturf tac
tics are also used to discredit or criticize those that disagree with certain agendas, using stereotypical names such as conspi
racy theorist or quack. When in fact when someone dares to reveal the truth or questions the official story, it should spark
a deeper curiosity and encourage further scrutiny of the information.This article (Journalist Reveals Tactics Brainwashing Indu
stry Uses to Manipulate the Public) was originally created and published by Waking Times and is published here under a Creative
Commons license with attribution to Vic Bishop and WakingTimes.com. It may be re-posted freely with proper attribution, author
bio, and this copyright statement. READ MORE MSM PROPAGANDA NEWS AT: 21st Century Wire MSM Watch Files


LR Prediction: Fake News
DT Prediction: Fake News
RF Prediction: Fake News
NB Prediction: Fake News
```

Observation: Our model is working absolutely fine by using all algorithms.

CONCLUSIONS:

- We have trained & tested 4 models for this task. They all perform very well.
- Our Decision Tree algorithm yields the highest accuracy, 99.7%.
- This algorithm can now classify news whether it's fake or not with high accuracy.

REFERENCES:

https://www.kaggle.com/datasets/clmentbisaillon/fake-and-real-news-dataset

https://www.hindawi.com/journals/complexity/2020/8885861/

https://www.itm-conferences.org/articles/itmconf/pdf/2021/05/itmconf_icacc2021_03003.pdf

https://link.springer.com/chapter/10.1007/978-3-030-49264-9_2

https://guides.lib.umich.edu/fakenews

https://www.icfj.org/sites/default/files/2018-07/A%20Short%20Guide%20to%20History%20of%20Fake%20News%20and%20Disinformation_ICFJ%20Final.pdf

https://www.pantechsolutions.net/fake-news-detection-using-machine-learning