

FAKE NEWS DETECTION USING NLP

ABSTRACT:

The spreading of fake news has given rise to many problems in society. It is due to its ability to cause a lot of social and national damage with destructive impacts. Sometimes it gets very difficult to know if the news is genuine or fake. Therefore it is very important to detect if the news is fake or not.

"Fake News" is a term used to represent fabricated news or propaganda comprising misinformation communicated through traditional media channels like print, and television as well as non-traditional media channels like social media. Techniques of NLP and Machine learning can be used to create models which can help to detect fake news.

In this paper we have presented six LSTM models using the techniques of NLP and ML. The datasets in comma-separated values format, pertaining to political domain were used in the project. The different attributes like the title and text of the news headline/article were used to perform the fake news detection. The results showed that the proposed solution performs well in terms of providing an output with good accuracy, precision and recall. The performance analysis made between all the models showed that the models which have used GloVe and Word2vec method work better than the models using TF-IDF. Further, a larger dataset for better output and also other factors such as the author ,publisher of the news can be used to determine the

credibility of the news. Also, further research can also be done on images, videos, images containing text which can help in improving the models in future.

INTRODUCTION:

WHY FAKE NEWS DETECTOR:

Fake news is dangerous as it can deceive people easily and create a state of confusion among a community. This can further affect the society badly .

The spread of fake news creates rumors circulating around and the victims could be badly impacted. Recent reports showed that due to the rise of fake news that was being created online it had impacted the US Presidential Elections. Fake news might be created by people or groups who are acting in their own interests or those of third parties. The creation of misinformation is usually motivated by personal, political, or economic agendas. A false news about anyone can affect their day to day activities leaving them stressed and making them to feel unwanted. In order to overcome these type of scenarios and to provide the actual news, artificial intelligence and machine learning have played a crucial role. Here, detecting the fake news is actually a difficult tasks but after many failures, the developers have creating a structure to frame the correct one.

LITERATURE SURVEY:

[1] M. Graniket.al proposed a simple approach for the detection of fake news by using Naive Bayes Classifier. They tested it against a dataset of Facebook news posts. They also made use of the BuzzFeed news dataset. They achieved classification accuracy of approximately 74% on the test set.

[2] Niall J, Conroet.al designed a basic fake news detector that provides high accuracy for classification tasks. They used the linguistic cues approaches and network analysis approach in it. Both approaches adopt machine learning techniques for training classifiers to suit the analysis. They achieved an accuracy of 72% which could be improved. This could be done if the size of the input feature vector is reduced and also by performing cross-corpus analysis of the classification models.

[3] R. Barua et.al identified if a news article is real or misleading by using an ensemble technique using recurrent neural networks (LSTM and GRU). An android application was also developed for determining the sanctity of a news article. They tested this model on a large dataset which was prepared in their work. The limitation of this method was that it required the article to be of a particular size.

It would give wrong predictions if the article was not enough to generate a summary.

[4] B. Bhutani et.al used sentiment as an important feature to improve the accuracy of detecting fake news. They have used 3 different datasets. They used Count vectorizer, Tf-Idf vectorizer along with cosine similarity and Bi-grams , Tri-grams methods. The methods used to train the model are Naive Bayes and Random forest. They used different performance metrics to evaluate the model. They got an accuracy of 81.6%.

PROBLEM STATEMENT:

Since a lot of time is spent by users on social media and people prefer online means of information it has become difficult to know about the authenticity of the news. People acquire most of the

information by these means as it is free and can be accessed from anywhere irrespective of place and time. Since this data can be put out by anyone there is lack of accountability in it which makes it less trustable unlike the traditional methods of gaining information like newspapers or some trusted source. Fake news is dangerous as it can deceive people easily and create a state of confusion among a community. This can further affect the society badly .The spread of fake news creates rumors circulating around and the victims could be badly impacted.

PROPOSED SOLUTION:

As we have seen that the problem of spreading fake news is a serious issue therefore, there is a need to detect this fake news. The main aim of the project is to obtain a model which will help in detecting if a news article is fake or not. The problem of detecting fake news is a very difficult task and many researchers are trying to obtain a solution to it.

Since there are not many datasets which are available publicly to perform this task. We have considered three different datasets which will be merged together to obtain a master dataset which will help in training the models to find if a news is fake or not.

OVERVIEW:

First the dataset is taken from the respective website. From the dataset, the programmer reads the data in the database and starts to classify them through the NLP techniques. Once when the techniques are implemented, then the preprocessing process is done. This includes removing of unnecessary data, removing null values. After this, the data visualizing process is preceded where the developer

differentiates the real and fake news in a tabular format. To overcome the problem of detecting fake news this project proposes 6 similar LSTM models which are to be trained and each model will be fed with the different text vectors of news headline and news content. This will help in obtaining a good model which will tell if the news is true or it is fake. In this project we have used six similar LSTM models. Three text vectorization techniques are used which are GloVe, Word2vec and TF-IDF. The first LSTM model will be fed with the vectors of the title of the news using GloVe. The second model will be fed with the vectors of the content of the news using GloVe. Similarly, two models will be built using the Word2vec technique each for the title of the news and the content of the news respectively. Lastly, the LSTM model will be fed with the text vectors of the title of the news using TF-IDF and another model will be fed with the text vectors of the content of the news using TF-IDF. By doing so we can identify which technique gives better results and identify which model performs well. Lastly, the performance is measured using the performance metrics accuracy, precision and recall.

STEPS INVOLVED:

FETCHING DATASET:

This step involves getting the dataset from the respective website. First the dataset is downloaded from kaggle website <https://www.kaggle.com/datasets/clmentbisailon/fake-and-real-news-dataset> and then we open the dataset in jupyter notebook. After this, the dataset is read and displayed through the head() function. This head() function is responsible for displaying the dataset.

DATA PROCESSING:

This step involves the process where the collected data is taken and then the null values are checked. If there are no null values found, then we move on to the next step of deleting unwanted details. By doing this, the user can easily understand the entire function by looking at its structure itself.

This dataset consists of 4 columns which are the URLs of the news source, the Headline of the news, the Body of the news that is the content of the news, and the last column contains the Label of the news which tells whether the news is fake or not.

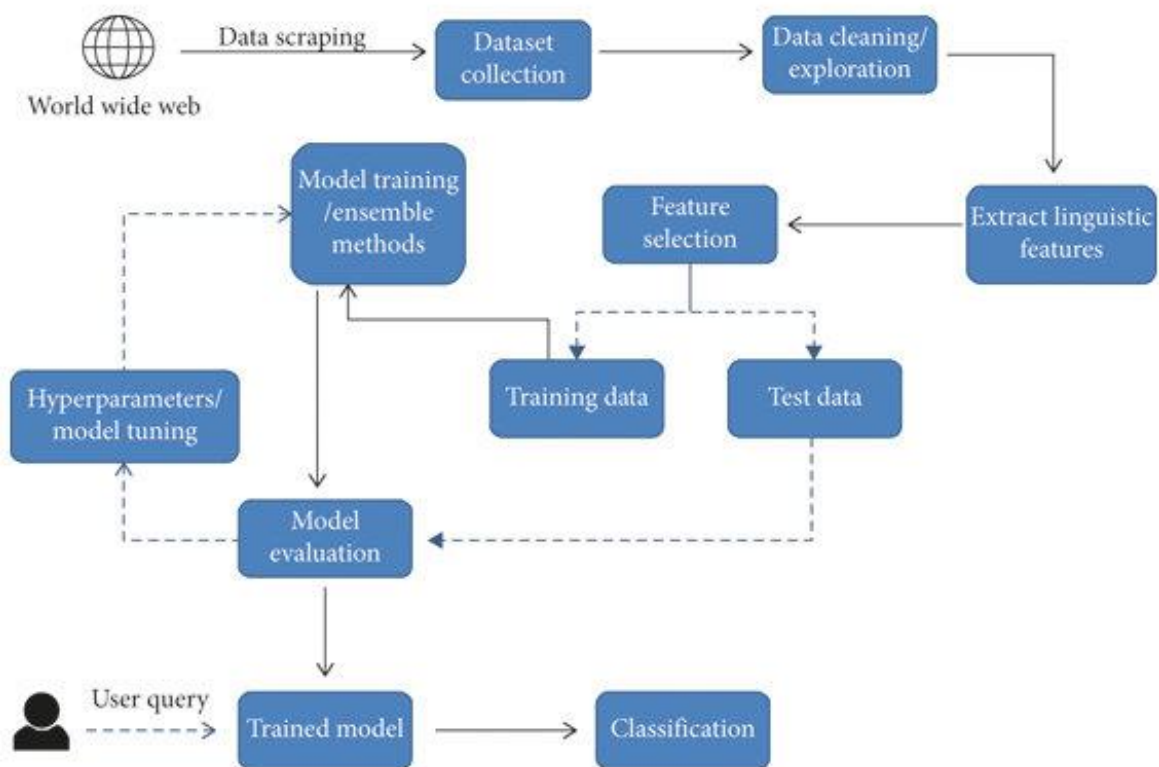
MERGING A DATASET:

Next, the two datasets are merged together to obtain a single dataset. After the merge, we obtained a dataset with 23471 records. Finally, we obtain a master dataset by merging the first dataset with the above merged dataset [dataset with 23471 records], hence the final obtained master dataset consists of 44898 records and five columns, Title, text, subject, date, and Class. In our program, we have removed the unwanted columns like title, subject, and date.

TRAIN-TEST-SPLIT:

This involves splitting up of training and testing dataset where the dataset is split in a way where 80% training data and 20% testing and splitting data.

FEATURE EXTRACTION:



The next step is feature extraction. Machine Learning algorithms learn from a predefined set of features from the training data to produce output for the test data. But the main problem in working with language processing is that machine learning algorithms cannot work on the raw text directly. So, we need some feature extraction techniques to convert text into a matrix(or vector) of features.

[Why fake news detector and how ML plays an important role in it?](#)

NEED A DETECTOR:

Nowadays due to the increase of various technological impacts a lot of people use it in a wrong way to misuse someone, to corrupt their privilege and to abuse them by spreading rumors through social media platforms. In order

to avoid this, we have developed the program in such a way that this finds out the unwanted or fake content about the person.

WHAT DOES ML DO:

Machine Learning Techniques have shown promising results in detecting fake news with the help of analysing vast amounts of data, in which it identifies patterns and it provides outcomes that are based on those patterns.

How the fake news is being detected through various methods in machine learning algorithm.

STEPS INVOLVED IN ML:

IMPORTING LIBRARIES AND DATASETS:

This step involves installation of the required libraries.

DATA PROCESSING:

In this step, we fetch the dataset and start to remove unwanted columns and null values. The first step is we read the dataset and next the unwanted values are removed which makes the processing simple and also neat.

PREPROCESSING AND ANALYSIS OF NEW COLUMN:

Once when the step 2 has ended the required information is

viewed and the datas are splitted into dependent and independent variables.

CONVERTING TEXT TO VECTORS:

This conversion step plays an important role for developing various preprocessing measures and for implementing various techniques It's necessary to proceed with this step so that both programmer and the user could understand the entire concept.

MODELING:

In this *program* we have concentrated mainly in the following five methods:

Logistic Regression:

This is a supervised machine learning algorithm which mainly focuses on predicting the probability for the targetted value. As this provides a result of 98.76%accuracy,we have given importance for this in our program.

Decision Tree Classifier:

Decision tree algorithm mainly focuses on grabbing the majority of possibility for any particular dataset.

We have used this as this classifier understands the pattern of the news pattern and provides the correct output whether the news is fake or a correct one.

Gradient Boost:

This provides a stronger result by combining and undersatnding the framwork of the dataset. As this provides an accurate value we have included this as a major part in our code.

Random Forest:

This is mainly used for predicting and classifying the data.As this is an ensemble learning technique it has the capability of finding the desired result by combing various possiblities from the decision tree classifier.

NATURAL PROCESSING LANGUAGE:

Natural language processing (NLP) is an interdisciplinary subfield of computer science and linguistics. It is primarily concerned with giving computers the ability to support and manipulate speech. It involves processing natural language datasets, such as text corpora or speech corpora, using either rule-based or probabilistic (i.e. statistical and, most recently, neural network-based) machine learning approaches. The goal is a computer capable of "understanding" the contents of documents, including the contextual nuances of the language within them. The technology can then accurately extract information and insights contained in the documents as well as categorize and organize the documents themselves.

Challenges in natural language processing frequently involve speech recognition, natural-language understanding, and natural-language generation.

HOW NATURAL LANGUAGE PROCESSING MAJOR ROLE IN DETECTING FAKE NEWS:

To identify bogus news, sentiment analysis using NLP can be effective strategy. NLP algorithms can be ascertain the intention and any biases of an author by analysing the emotions displayed in a news story or social media post. Fake news frequently preys on readers' emotions by using strong language or exaggeration.

IS NATURAL LANGUAGE PROCESSING ACTUALLY NEEDED?

NLP is important because it helps resolve ambiguity in language and adds useful numeric structure to the data for many downstream applications, such as speech recognition or text analytics.

TERMINOLOGIES:

FAKE NEWS:

we are defining “fake news” as those news stories that are false: the story itself is fabricated, with no verifiable facts, sources or quotes. Sometimes these stories may be propaganda that is intentionally designed to mislead the reader, or may be designed as “clickbait” written for economic incentives (the writer profits on the number of people who click on the story).

However, it’s important to acknowledge that “fake news” is a complex and nuanced problem, one that is far greater than the narrow definition above. The term itself has become politicized, and is widely used to discredit any opposing viewpoint. Some people use it to cast doubt on their opponents, controversial issues or the credibility of

some media organizations. In addition, technological advances such as the advent of social media enable fake news stories to proliferate quickly and easily as people share more and more information online. Increasingly, we rely on online information to understand what is happening in our world.

In order to handle these type of unwanted news,we have implemented specific NLP properties:

TF-IDF VECTORIZER:

TF(TERM FREQUENCY):

In the document, words are present so many times that is called term frequency. In this section, if you get the largest values it means that word is present so many times with respect to other words. when you get word is parts of speech word that means the document is a very nice match.

IDF(INVERSE DOCUMENT FREQUENCY):

In a single document, words are present so many times, but also available so many times in another document also which is not relevant. IDF is a proportion of how critical a term is in the whole corpus. collection of word Documents will convert into the matrix which contains TF-IDF features using Tf Idf Vectorizer.

In the previous phase, we have done the loading and preprocessing of data. Here we are dealing with the next step of how to develop a clean code for a program.

SOURCE CODE:

```
import pandas as pd
import numpy as np
```

```

import seaborn as sns
import matplotlib.pyplot as plt
from sklearn.model_selection import train_test_split
from sklearn.metrics import accuracy_score
from sklearn.metrics import classification_report
import re
import string

```

```

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

```

```
df_fake.head()
```

	title	text	subject	date
0	Donald Trump Sends Out Embarrassing New Year'...	Donald Trump just couldn t wish all Americans ...	News	December 31, 2017
1	Drunk Bragging Trump Staffer Started Russian ...	House Intelligence Committee Chairman Devin Nu...	News	December 31, 2017
2	Sheriff David Clarke Becomes An Internet Joke...	On Friday, it was revealed that former Milwauk...	News	December 30, 2017
3	Trump Is So Obsessed He Even Has Obama's Name...	On Christmas day, Donald Trump announced that ...	News	December 29, 2017
4	Pope Francis Just Called Out Donald Trump Dur...	Pope Francis used his annual Christmas Day mes...	News	December 25, 2017

```
df_true.head(5)
```

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

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

```
df_fake.shape, df_true.shape
```

((23481, 5), (21417, 5))

out:

```
# Removing last 10 rows for manual testing
df_fake_manual_testing = df_fake.tail(10)
for i in range(23480,23470,-1):
    df_fake.drop([i], axis = 0, inplace = True)
df_true_manual_testing = df_true.tail(10)
for i in range(21416,21406,-1):
    df_true.drop([i], axis = 0, inplace = True)
```

```
df_fake.shape, df_true.shape
```

Out:

((23471, 5), (21407, 5))

```
df_fake_manual_testing["class"] = 0
df_true_manual_testing["class"] = 1
```

```
df_fake_manual_testing.head(10)
```

	title	text	subject	date	Out: class
23471	Seven Iranians freed in the prisoner swap have...	21st Century Wire says This week, the historic...	Middle-east	January 20, 2016	0
23472	#Hashtag Hell & The Fake Left	By Dady Chery and Gilbert MercierAll writers ...	Middle-east	January 19, 2016	0

23473	Astroturfing: Journalist Reveals Brainwashing ...	Vic Bishop Waking TimesOur reality is carefull...	Middle-east	January 19, 2016	0
23474	The New American Century: An Era of Fraud	Paul Craig RobertsIn the last years of the 20t...	Middle-east	January 19, 2016	0
23475	Hillary Clinton: 'Israel First' (and no peace ...	Robert Fantina CounterpunchAlt hough the United...	Middle-east	January 18, 2016	0
23476	McPain: John McCain Furious That Iran Treated ...	21st Century Wire says As 21WIRE reported earl...	Middle-east	January 16, 2016	0
23477	JUSTICE? Yahoo Settles E-mail Privacy Class-ac...	21st Century Wire says It s a familiar theme. ...	Middle-east	January 16, 2016	0
23478	Sunnistan: US and Allied 'Safe Zone' Plan to T...	Patrick Henningsen 21st Century WireRemember ...	Middle-east	January 15, 2016	0
23479	How to Blow \$700 Million: Al Jazeera America F...	21st Century Wire says Al Jazeera America will...	Middle-east	January 14, 2016	0
23480	10 U.S. Navy Sailors Held by Iranian Military ...	21st Century Wire says As 21WIRE predicted in ...	Middle-east	January 12, 2016	0

```
df_true_manual_testing.head(10)
```

	title	text	subject	date	Out: class
21407	Mata Pires, owner of embattled Brazil builder ...	SAO PAULO (Reuters) - Cesar Mata Pires, the ow...	worldnews	August 22, 2017	1
21408	U.S., North Korea clash at U.N. forum over nuc...	GENEVA (Reuters) - North Korea and the United ...	worldnews	August 22, 2017	1
21409	U.S., North Korea clash at U.N. arms forum on ...	GENEVA (Reuters) - North Korea and the United ...	worldnews	August 22, 2017	1
21410	Headless torso could belong to submarine journ...	COPENHAGEN (Reuters) - Danish police said on T...	worldnews	August 22, 2017	1
21411	North Korea shipments to Syria chemical arms a...	UNITED NATIONS (Reuters) - Two North Korean sh...	worldnews	August 21, 2017	1
21412	'Fully committed' NATO backs new U.S. approach...	BRUSSELS (Reuters) - NATO allies on Tuesday we...	worldnews	August 22, 2017	1

21413	LexisNexis withdrew two products from Chinese ...	LONDON (Reuters) - LexisNexis, a provider of l...	worldnews	August 22, 2017	1
21414	Minsk cultural hub becomes haven from authorities	MINSK (Reuters) - In the shadow of disused Sov...	worldnews	August 22, 2017	1
21415	Vatican upbeat on possibility of Pope Francis ...	MOSCOW (Reuters) - Vatican Secretary of State ...	worldnews	August 22, 2017	1
21416	Indonesia to buy \$1.14 billion worth of Russia...	JAKARTA (Reuters) - Indonesia will buy 11 Sukh...	worldnews	August 22, 2017	1

```
df_manual_testing =
pd.concat([df_fake_manual_testing,df_true_manual_testing],
axis = 0)
df_manual_testing.to_csv("manual_testing.csv")
```

```
df_merge = pd.concat([df_fake, df_true], axis =0 )
df_merge.head(10)
```

	title	text	subject	date	Out: class
0	Donald Trump Sends Out Embarrassing New Year'...	Donald Trump just couldn t wish all Americans ...	News	December 31, 2017	0
1	Drunk Bragging Trump Staffer Started Russian ...	House Intelligence Committee Chairman Devin Nu...	News	December 31, 2017	0
2	Sheriff David Clarke Becomes An Internet Joke...	On Friday, it was revealed that former Milwauk...	News	December 30, 2017	0
3	Trump Is So Obsessed He Even Has Obama's Name...	On Christmas day, Donald Trump announced that ...	News	December 29, 2017	0
4	Pope Francis Just Called Out Donald Trump Dur...	Pope Francis used his annual Christmas Day mes...	News	December 25, 2017	0
5	Racist Alabama Cops Brutalize Black Boy While...	The number of cases of cops brutalizing and ki...	News	December 25, 2017	0

6	Fresh Off The Golf Course, Trump Lashes Out A...	Donald Trump spent a good portion of his day a...	News	December 23, 2017	0
7	Trump Said Some INSANELY Racist Stuff Inside ...	In the wake of yet another court decision that...	News	December 23, 2017	0
8	Former CIA Director Slams Trump Over UN Bully...	Many people have raised the alarm regarding th...	News	December 22, 2017	0
9	WATCH: Brand-New Pro-Trump Ad Features So Muc...	Just when you might have thought we d get a br...	News	December 21, 2017	0

```
df_merge.columns
```

Out:

```
Index(['title', 'text', 'subject', 'date', 'class'], dtype='object')
```

```
df = df_merge.drop(["title", "subject","date"], axis = 1)
df.isnull().sum()
```

Out:

```
text      0
class     0
dtype: int64
```

```
df = df.sample(frac = 1)
```

```
df.head()
```

Out:

		text	class
5391	Make no mistake about it: Ted Cruz is horrible...		0
2841	WASHINGTON (Reuters) - U.S. President Donald T...		1
8914	SAN JUAN (Reuters) - Puerto Rico's semi-public...		1
10062	WASHINGTON (Reuters) - A moderate Republican s...		1
19332	Streep s shameful attempt to paint Trump as a ...		0

```
df.reset_index(inplace = True)
```

```
df.drop(["index"], axis = 1, inplace = True)
```

```
df.columns
```

Out:

```
Index(['text', 'class'], dtype='object')
```

```
df.head()
```

Out:

	text	class
0	Make no mistake about it: Ted Cruz is horrible...	0
1	WASHINGTON (Reuters) - U.S. President Donald T...	1
2	SAN JUAN (Reuters) - Puerto Rico's semi-public...	1
3	WASHINGTON (Reuters) - A moderate Republican s...	1
4	Streep s shameful attempt to paint Trump as a ...	0

```
def wordopt(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
```

```
df["text"] = df["text"].apply(wordopt)
```

```
x = df["text"]
y = df["class"]
```

```
x_train, x_test, y_train, y_test = train_test_split(x, y,
test_size=0.25)
```

A. VECTORIZATION:

Vectorization is a methodology in NLP to map words or phrases from vocabulary to a corresponding vector of real numbers which is used to find word predictions, word similarities/semantics.

For curiosity, you surely want to check out this article on ‘ Why data are represented as vectors in Data Science Problems’.

To make documents’ corpora more relatable for computers, they must first be converted into some numerical structure. There are few techniques that are used to achieve this such as ‘Bag of Words’.

```
from sklearn.feature_extraction.text import TfidfVectorizer
vectorization = TfidfVectorizer()
xv_train = vectorization.fit_transform(x_train)
xv_test = vectorization.transform(x_test)
```

B. LOGISTIC REGRESSION:

Logistic regression is one of the most popular Machine Learning algorithms, which comes under the Supervised Learning technique. It is used for predicting the categorical dependent variable using a given set of independent variables. Logistic regression predicts the output of a categorical dependent variable. □ Logistic Regression is a significant machine learning algorithm because it has the ability to provide probabilities and classify new data using continuous and discrete datasets.

□ Logistic Regression can be used to classify the observations using different types of data and can easily determine the most effective variables used for the classification. □ Logistic regression uses the concept of predictive modeling as regression; therefore, it is called logistic regression, but is used to classify samples; Therefore, it falls under the classification algorithm.

```
from sklearn.linear_model import LogisticRegression
LR = LogisticRegression()
LR.fit(xv_train,y_train)
```

Out:

```
LogisticRegression()
```

```
pred_lr=LR.predict(xv_test)
LR.score(xv_test, y_test)
```

Out:

```
0.986541889483066
```

```
print(classification_report(y_test, pred_lr))
```

	precision	recall	f1-score	support
0	0.99	0.99	0.99	5849
1	0.98	0.99	0.99	5371
accuracy			0.99	11220
macro avg	0.99	0.99	0.99	11220
weighted avg	0.99	0.99	0.99	11220

C. DECISION TREE CLASSIFIER:

Decision Trees (DTs) are a non-parametric supervised learning method used for classification and regression. The goal is to create a model that predicts the value of a target variable by learning simple decision rules inferred from the data features. A tree can be seen as a piecewise constant approximation.

```
from sklearn.tree import DecisionTreeClassifier
DT = DecisionTreeClassifier()
DT.fit(xv_train, y_train)
```

Out:

```
DecisionTreeClassifier()
```

```
pred_dt = DT.predict(xv_test)
```

```
DT.score(xv_test, y_test)
```

Out:

```
0.9949197860962566
```

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

	precision	recall	f1-score	support
0	1.00	0.99	1.00	5849
1	0.99	1.00	0.99	5371
accuracy			0.99	11220
macro avg	0.99	0.99	0.99	11220
weighted avg	0.99	0.99	0.99	11220

D. GRADIENT BOOSTING CLASSIFIER:

Gradient Boosting is a popular boosting algorithm in machine learning used for classification and regression tasks. Boosting is one kind of ensemble Learning method which trains the model sequentially and each new model tries to correct the previous model. It combines several weak learners into strong learners.

```
from sklearn.ensemble import GradientBoostingClassifier
GBC = GradientBoostingClassifier(random_state=0)
GBC.fit(xv_train, y_train)
```

```
pred_gbc = GBC.predict(xv_test)
```

```
GBC.score(xv_test, y_test)
```

```
print(classification_report(y_test, pred_gbc))
```

E. RANDOM FOREST CLASSIFIER:

A random forest is a meta estimator that fits a number of decision tree classifiers on various sub-samples of the dataset and uses averaging to improve the predictive accuracy and control over-fitting. The sub-sample size is controlled with the `max_samples` parameter if `bootstrap=True` (default), otherwise the whole dataset is used to build each tree.

```
from sklearn.ensemble import RandomForestClassifier
RFC = RandomForestClassifier(random_state=0)
RFC.fit(xv_train, y_train)
```

```
pred_rfc = RFC.predict(xv_test)
```

```
RFC.score(xv_test, y_test)
```

```
print(classification_report(y_test, pred_rfc))
```

```
def output_lable(n):
    if n == 0:
        return "Fake News"
    elif n == 1:
        return "Not A Fake News"
def manual_testing(news):
    testing_news = {"text":[news]}
    new_def_test = pd.DataFrame(testing_news)
    new_def_test["text"] = new_def_test["text"].apply(wordopt)
    new_x_test = new_def_test["text"]
    new_xv_test = vectorization.transform(new_x_test)
    pred_LR = LR.predict(new_xv_test)
    pred_DT = DT.predict(new_xv_test)
    pred_GBC = GBC.predict(new_xv_test)
    pred_RFC = RFC.predict(new_xv_test)

    return print("\n\nLR Prediction: {} \nDT Prediction: {}
\nGBC Prediction: {} \nRFC Prediction: {}".format(pred_LR, pred_DT, pred_GBC, pred_RFC))
```

```
format(output_lable(pred_LR[0]),  
output_lable(pred_DT[0]),  
output_lable(pred_GBC[0]),  
Output_lable(pred_RFC[0])))
```

```
news = str(input())  
manual_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 House last year on a pledge to withdraw swiftly from Afghanistan, Trump reversed course on Monday and promised a stepped-up military campaign against Taliban insurgents, saying: Our troops will fight to win. U.S. officials said he had signed off on plans 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 benchmarks for successfully ending the war that began with the U.S.-led invasion of Afghanistan in 2001, and which he acknowledged had required an extraordinary sacrifice of blood and treasure. We will ask our NATO allies and global partners to support 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 despite 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 scale of its defense spending, also welcomed the new U.S. plan. Our continued commitment is necessary on the path to stabilizing the country, a government spokeswoman said. In June, European allies had already pledged more troops but had not given details on numbers, waiting for the Trump administration to outline its strategy.

for the region. Nearly 16 years after the U.S.-led invasion - a response to the Sept. 11 attacks which were planned by al Qaeda leader Osama bin Laden from Afghanistan - the country is still struggling with weak central government and a Taliban insurgency. Trump said he shared the frustration of the American 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

GBC Prediction: Not A Fake News

RFC Prediction: Not A Fake News

```
news = str(input())  
manual_testing(news)
```

Vic Bishop Waking Times Our reality is carefully constructed by powerful corporate, political and special interest sources in order to covertly sway public opinion. Blatant lies are often televised regarding terrorism, food, war, health, etc. They are fashioned to sway public opinion and condition viewers to accept what have become destructive societal norms. The practice of manipulating 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 the six conglomerates that own 90% of the media: General Electric (GE), News-Corp, Disney, Viacom, Time Warner, and CBS. Yet, these corporations function under many different brands, such as FOX, ABC, CNN, Comcast, Wall Street Journal, etc, giving people the 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 could be varied slightly, so as to mask the sense of external control. ~ Specialist of mass brainwashing, L. Wolfe

New Brainwashing Tactic Called Astroturf With alternative media on the rise, the propaganda machine continues to expand. Below is a video of Sharyl Attkisson, investigative reporter with CBS, during which she explains how astroturf, or fake grassroots movements, are used to spin information not only to influence journalists but to sway public opinion. Astroturf is a

perversion of grassroots. Astroturf is when political, corporate or other special interests disguise themselves and publish blogs, start facebook and twitter accounts, publish ads, letters to the editor, or simply post comments online, to try to fool you into thinking an independent or grassroots movement is speaking. ~ Sharyl Attkisson, Investigative Reporter

How do you separate fact from fiction? Sharyl Attkisson finishes her talk with some insights on how to identify signs of propaganda and astroturfing. These methods are used to give people the impression that there is widespread support for an agenda, when, in reality, one may not exist. Astroturf tactics are also used to discredit or criticize those that disagree with certain agendas, using stereotypical names such as conspiracy 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 Industry 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

GBC Prediction: Fake News

RFC Prediction: Fake News

```
news = str(input())
manual_testing(news)
```

SAO PAULO (Reuters) - Cesar Mata Pires, the owner and co-founder of Brazilian engineering conglomerate OAS SA, one of the largest companies involved in Brazil's corruption scandal, died on Tuesday. He was 68. Mata Pires died of a heart attack while taking a morning walk in an upscale district of S o Paulo, where OAS is based, a person with direct knowledge of the matter said. Efforts to contact his family were unsuccessful. OAS declined to comment. The son of a wealthy cattle rancher in the northeastern state of Bahia, Mata Pires' links to politicians were central to the expansion of OAS, which became Brazil's No. 4 builder earlier this decade, people familiar with his career told Reuters last year. His big break came when he befriended Antonio Carlos Magalh es, a popular

politician who was Bahia governor several times, and eventually married his daughter Tereza. Brazilians joked that OAS stood for Obras Arranjadas pelo Sogro - or Work Arranged by the Father-In-Law. After years of steady growth triggered by a flurry of massive government contracts, OAS was ensnared in Operation Car Wash which unearthed an illegal contracting ring between state firms and builders. The ensuing scandal helped topple former Brazilian President Dilma Rousseff last year. Trained as an engineer, Mata Pires founded OAS with two colleagues in 1976 to do sub-contracting work for larger rival Odebrecht SA - the biggest of the builders involved in the probe. Before the scandal, Forbes magazine estimated Mata Pires' fortune at \$1.6 billion. He dropped off the magazine's billionaire list in 2015, months after OAS sought bankruptcy protection after the Car Wash scandal. While Mata Pires was never accused of wrongdoing in the investigations, creditors demanded he and his family stay away from the builder's day-to-day operations, people directly involved in the negotiations told Reuters at the time. He is survived by his wife and his two sons.

LR Prediction: Not A Fake News

DT Prediction: Not A Fake News

GBC Prediction: Not A Fake News

RFC Prediction: Not A Fake News

OBSERVED RESULT:

The model we have made is producing accurate results, considering the accuracy of all the models, which was almost 99%, so we can say machine learning can be used as a tool for detecting fake news.

CONCLUSION:

By detecting fake news articles before they are widely disseminated, machine learning algorithms can prevent the harm caused by fake news. However, it is important to use diverse datasets and other techniques, such as fact-checking,

to verify the authenticity of news articles.

Fake news have increased in recent years and it has caused a lot of harm to the society. This research project aimed to develop a model using the techniques of NLP and ml to detect if a news article/headline is fake or not and identify which methods give better output. In this paper, we have presented six models and three different methods were used for feature extraction. We have used different attributes like the title and text of the news to perform fake news detection. For future work we can work on larger dataset and also future research can be done on images , videos which can help in improving the models. The version of this template is v2. Most of the formatting instructions in this document have been compiled by causal productions from the iee latex style files. Causal productions offers both a4 templates and us letter templates for latex and microsoft word. The latex templates depend on the official ieeetran.cls and ieeetran.bst files, whereas the microsoft word templates are self-contained. Causal productions has used its best efforts to ensure that the templates have the same appearance