

```
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
```

```
data_fake=pd.read_csv("/content/Fake.csv")
data_true=pd.read_csv("/content/True.csv")
```

```
data_fake.head()
```

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

```
data_true.head()
```

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

```
data_fake["class"] = 0
data_true["class"] = 1
```

```
data_fake.shape, data_true.shape
```

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

```
data_fake_manual_testing = data_fake.tail(10)
for i in range(23480, 23470,-1):
    data_fake.drop([i], axis = 0, inplace = True)

data_true_manual_testing = data_true.tail(10)
for i in range(21416, 21406,-1):
    data_true.drop([i], axis = 0, inplace = True)

data_fake.shape, data_true.shape

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

data_fake_manual_testing['class'] = 0
data_true_manual_testing['class'] = 1

<ipython-input-24-90008d39c97b>:1: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.
data_fake_manual_testing['class'] = 0
<ipython-input-24-90008d39c97b>:2: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.
data_true_manual_testing['class'] = 1
```

data_fake_manual_testing.head(10)

		title	text	subject	
23471	Seven Iranians freed in the prisoner swap have...	21st Century Wire says This week, the historic...	Middle-east	Jan	
23472	#Hashtag Hell & The Fake Left	By Dady Chery and Gilbert MercierAll writers ...	Middle-east	Jan	
23473	Astroturfing: Journalist Reveals Brainwashing ...	Vic Bishop Waking TimesOur reality is carefull...	Middle-east	Jan	
23474	The New American Century: An Era of Fraud	Paul Craig RobertsIn the last years of the 20t...	Middle-east	Jan	
23475	Hillary Clinton: 'Israel First' (and no peace ...	Robert Fantina CounterpunchAlthough the United...	Middle-east	Jan	
23476	McPain: John McCain Furious That Iran Treated ...	21st Century Wire says As 21WIRE reported earl...	Middle-east	Jan	
23477	JUSTICE? Yahoo Settles E-mail Privacy Class-ac...	21st Century Wire says It s a familiar theme. ...	Middle-east	Jan	
23478	Sunnistan: US and Allied 'Safe Zone' Plan to T...	Patrick Henningsen 21st Century WireRemember ...	Middle-east	Jan	
23479	How to Blow \$700 Million: Al Jazeera America F...	21st Century Wire says Al Jazeera America will...	Middle-east	Jan	
23480	10 U.S. Navy Sailors Held by Iranian Military ...	21st Century Wire says As 21WIRE predicted in ...	Middle-east	Jan	

```
data_true_manual_testing.head(10)
```

		title	text	subject	
21407	Mata Pires, owner of embattled Brazil builder ...	SAO PAULO (Reuters) - Cesar Mata Pires, the ow...	worldnews	Au	
21408	U.S., North Korea clash at U.N. forum over nuc...	GENEVA (Reuters) - North Korea and the United ...	worldnews	Au	
21409	U.S., North Korea clash at U.N. arms forum on ...	GENEVA (Reuters) - North Korea and the United ...	worldnews	Au	
21410	Headless torso could belong to submarine journ...	COPENHAGEN (Reuters) - Danish police said on T...	worldnews	Au	
21411	North Korea shipments to Syria chemical arms a...	UNITED NATIONS (Reuters) - Two North Korean sh...	worldnews	Au	
21412	'Fully committed' NATO backs new U.S. approach...	BRUSSELS (Reuters) - NATO allies on Tuesday we...	worldnews	Au	
21413	LexisNexis withdrew two products from Chinese ...	LONDON (Reuters) - LexisNexis, a provider of l...	worldnews	Au	
21414	Minsk cultural hub becomes haven from authorities	MINSK (Reuters) - In the shadow of disused Sov...	worldnews	Au	
21415	Vatican upbeat on possibility of Pope Francis ...	MOSCOW (Reuters) - Vatican Secretary of State ...	worldnews	Au	
21416	Indonesia to buy \$1.14 billion worth of Russia...	JAKARTA (Reuters) - Indonesia will buy 11 Sukh...	worldnews	Au	

```
data_merge = pd.concat([data_fake, data_true], axis = 0)
data_merge.head(10)
```

```



                                title                                text  subject
_
Donald Trump Sends Out Embarrassing New  Donald Trump just couldn t wish all Americans  ..  Decen
data_merge.columns

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

data = data_merge.drop(['title', 'subject', 'date'], axis =1)

data.isnull().sum()

text      0
class     0
dtype: int64
3  Racist Alabama Cops Brutalize Black Boy While...      ki  NEWS
data = data.sample(frac = 1)
_
A...      a...  NEWS
data.head()
```

	text	class	
19606		0	
6220	Is the primary process in the United States me...	0	
5194	OTTAWA (Reuters) - Canadian and U.S. officials...	1	
18715	Wow! Republicans had a chance to stand up to t...	0	
15411	What in the world is she thinking? She says th...	0	



```

data.reset_index(inplace = True)
data.drop(['index'], axis = 1, inplace = True)

data.columns

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

data.head()
```

	text	class	
0		0	
1	Is the primary process in the United States me...	0	
2	OTTAWA (Reuters) - Canadian and U.S. officials...	1	
3	Wow! Republicans had a chance to stand up to t...	0	
4	What in the world is she thinking? She says th...	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('\w*\d\w*', '', text)
return text

```

```
data['text'] = data['text'].apply(wordopt)
```

```

x = data['text']
y = data['class']

```

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

```
from sklearn.feature_extraction.text import TfidfVectorizer
```

```

vectorization = TfidfVectorizer()
xv_train = vectorization.fit_transform(x_train)
xv_test = vectorization.transform(x_test)

```

```
from sklearn.linear_model import LogisticRegression
```

```

LR = LogisticRegression()
LR.fit(xv_train, y_train)

```

```

▼ LogisticRegression
LogisticRegression()

```

```
pred_lr = LR.predict(xv_test)
```

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

```
0.9864527629233512
```

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

	precision	recall	f1-score	support
0	0.99	0.99	0.99	5883
1	0.98	0.99	0.99	5337
accuracy			0.99	11220
macro avg	0.99	0.99	0.99	11220
weighted avg	0.99	0.99	0.99	11220

```
from sklearn.tree import DecisionTreeClassifier
```

```

DT = DecisionTreeClassifier()
DT.fit(xv_train, y_train)

```

```

▼ DecisionTreeClassifier
DecisionTreeClassifier()

```

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

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

```
0.9963458110516934
```

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

	precision	recall	f1-score	support
0	1.00	1.00	1.00	5883
1	1.00	0.99	1.00	5337
accuracy			1.00	11220
macro avg	1.00	1.00	1.00	11220
weighted avg	1.00	1.00	1.00	11220

```
from sklearn.ensemble import GradientBoostingClassifier
```

```
GB = GradientBoostingClassifier(random_state = 0)
```

```
GB.fit(xv_train, y_train)
```

▼

GradientBoostingClassifier

GradientBoostingClassifier(random_state=0)

```
pred_gb = GB.predict(xv_test)
```

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

```
0.9957219251336898
```

```
print(classification_report(y_test, pred_gb))
```

	precision	recall	f1-score	support
0	1.00	0.99	1.00	5883
1	0.99	1.00	1.00	5337
accuracy			1.00	11220
macro avg	1.00	1.00	1.00	11220
weighted avg	1.00	1.00	1.00	11220

```
from sklearn.ensemble import RandomForestClassifier
```

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

```
RF.fit(xv_train, y_train)
```

▼

RandomForestClassifier

RandomForestClassifier(random_state=0)

```
pred_rf = RF.predict(xv_test)
```

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

0.9898395721925134

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

	precision	recall	f1-score	support
0	0.99	0.99	0.99	5883
1	0.99	0.99	0.99	5337
accuracy			0.99	11220
macro avg	0.99	0.99	0.99	11220
weighted avg	0.99	0.99	0.99	11220