

```
In [1]: import pandas as pd
import numpy as np
from sklearn.model_selection import train_test_split
from sklearn.metrics import classification_report
import re
import string

In [2]: df_fake = pd.read_csv("Fake.csv")
df_true = pd.read_csv("True.csv")

In [3]: df_fake.head(10)

Out[3]:
   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
5 Racist Alabama Cops Brutalize Black Boy While... The number of cases of cops brutalizing and ki... News December 25, 2017
6 Fresh Off The Golf Course, Trump Lashes Out A... Donald Trump spent a good portion of his day a... News December 23, 2017
7 Trump Said Some INSANELY Racist Stuff Inside ... In the wake of yet another court decision that... News December 23, 2017
8 Former CIA Director Slams Trump Over UN Bully... Many people have raised the alarm regarding th... News December 22, 2017
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

In [4]: df_true.head(10)

Out[4]:
   title text subject 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. Muel... WASHINGTON (Reuters) - The special counsel inv... politicsNews December 31, 2017
3 FBI Russia probe helped by Australian diplom... WASHINGTON (Reuters) - Trump campaign adviser ... politicsNews December 30, 2017
4 Trump wants Postal Service to charge 'much mor... SEATTLEWASHINGTON (Reuters) - President Donal... politicsNews December 29, 2017
5 White House, Congress prepare for talks on spe... WEST PALM BEACH, Fla./WASHINGTON (Reuters) - T... politicsNews December 29, 2017
6 Trump says Russia probe will be fair, but time... WEST PALM BEACH, Fla (Reuters) - President Don... politicsNews December 29, 2017
7 Factbox: Trump on Twitter (Dec 29) - Approval ... The following statements were posted to the ve... politicsNews December 29, 2017
8 Trump on Twitter (Dec 28) - Global Warming The following statements were posted to the ve... politicsNews December 29, 2017
9 Alabama official to certify Senator-elect Jone... WASHINGTON (Reuters) - Alabama Secretary of St... politicsNews December 28, 2017

In [5]: df_fake["class"]=0
df_true["class"]=1

In [6]: df_fake.shape,df_true.shape

Out[6]: ((23481, 5), (21417, 5))

In [7]: 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)

In [8]: df_manual_testing = pd.concat([df_fake_manual_testing,df_true_manual_testing],axis=0)
df_manual_testing.to_csv("manual_testing.csv")

In [9]: df_marge=pd.concat([df_fake,df_true],axis=0)
df_marge.head(10)

Out[9]:
   title text subject date 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

In [10]: df=df_marge.drop(["title","subject","date"],axis=1)
df.head(10)

Out[10]:
   text class
0 Donald Trump just couldn t wish all Americans ... 0
1 House Intelligence Committee Chairman Devin Nu... 0
2 On Friday, it was revealed that former Milwauk... 0
3 On Christmas day, Donald Trump announced that ... 0
4 Pope Francis used his annual Christmas Day mes... 0
5 The number of cases of cops brutalizing and ki... 0
6 Donald Trump spent a good portion of his day a... 0
7 In the wake of yet another court decision that... 0
8 Many people have raised the alarm regarding th... 0
9 Just when you might have thought we d get a br... 0

In [11]: df=df.sample(frac=1)

In [12]: df.head(10)

Out[12]:
   text class
18414 MANCHESTER, England (Reuters) - British Prime ... 1
12080 VATICAN CITY (Reuters) - Tens of thousands of ... 1
16705 Wow! These people are clueless! Scary! 0
15925 BEIRUT/AMMAN/ANKARA (Reuters) - Syria s opposi... 1
19280 MEXICO CITY (Reuters) - For a country still in... 1
23390 21st Century Wire says NATO s ugly stepchild, ... 0
21807 I watched this exchange last night and couldn ... 0
18378 BERLIN (Reuters) - More than half of Germans, ... 1
8931 WASHINGTON (Reuters) - The Commodity Futures T... 1
3454 (Reuters) - World leaders, Democratic members ... 1

In [13]: df.isnull().sum()

Out[13]:
text 0
class 0
dtype: int64

In [14]: def word_drop(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\dw+', '',text)
return text

In [15]: df["text"]=df["text"].apply(word_drop)

In [16]: df.head(10)

Out[16]:
   text class
18414 manchester reuters british prime minist... 1
12080 vatican city reuters tens of of people m... 1
16705 wow these people are clueless scary 0
15925 beirut amman ankara reuters syria s opposi... 1
19280 mexico city reuters for a country still in... 1
23390 21st century wire says nato s ugly turkey h... 0
21807 i this exchange last night t believe the in... 0
18378 berlin reuters more than half of germans ... 1
8931 washington reuters the futures commissio... 1
3454 reuters members of congress others u... 1

In [17]: x=df["text"]
y=df["class"]

In [18]: x_train,x_test,y_train,y_test=train_test_split(x,y,test_size=.25)

In [19]: from sklearn.feature_extraction.text import CountVectorizer

In [20]: vectorization=CountVectorizer()
xv_train=vectorization.fit_transform(x_train)
xv_test=vectorization.transform(x_test)

In [21]: from sklearn.linear_model import LogisticRegression

In [22]: LR=LogisticRegression()
LR.fit(xv_train,y_train)

C:\Users\irishi\anaconda3\lib\site-packages\sklearn\linear_model\_logistic.py:763: ConvergenceWarning: lbfgs failed to converge (status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
https://scikit-learn.org/stable/modules/linear_model.html#logistic-regression
n_iter_1 = _check_optimize_result(

Out[22]: LogisticRegression()

In [23]: pred_LR=LR.predict(xv_test)

In [24]: print(classification_report(y_test,pred_LR))

precision recall f1-score support
0 1.00 1.00 1.00 5856
1 1.00 1.00 1.00 5364
accuracy 1.00 1.00 1.00 11220
macro avg 1.00 1.00 1.00 11220
weighted avg 1.00 1.00 1.00 11220

In [25]: from sklearn.tree import DecisionTreeClassifier

In [26]: DT=DecisionTreeClassifier()
DT.fit(xv_train,y_train)

Out[26]: DecisionTreeClassifier()

In [27]: DT.score(xv_test,y_test)

Out[27]: 0.9986131907398378

In [28]: pred_DT=DT.predict(xv_test)

In [29]: print(classification_report(y_test,pred_DT))

precision recall f1-score support
0 1.00 1.00 1.00 5856
1 1.00 1.00 1.00 5364
accuracy 1.00 1.00 1.00 11220
macro avg 1.00 1.00 1.00 11220
weighted avg 1.00 1.00 1.00 11220

In [31]: from sklearn.ensemble import GradientBoostingClassifier

In [32]: GBC=GradientBoostingClassifier()
GBC.fit(xv_train,y_train)

Out[32]: GradientBoostingClassifier()

In [33]: GBC.score(xv_test,y_test)

Out[33]: 0.9953654188948307

In [34]: pred_GBC=GBC.predict(xv_test)

In [35]: print(classification_report(y_test,pred_GBC))

precision recall f1-score support
0 1.00 0.99 1.00 5856
1 0.99 1.00 1.00 5364
accuracy 1.00 1.00 1.00 11220
macro avg 1.00 1.00 1.00 11220
weighted avg 1.00 1.00 1.00 11220

In [36]: from sklearn.ensemble import RandomForestClassifier

In [37]: RFC=RandomForestClassifier()
RFC.fit(xv_train,y_train)

Out[37]: RandomForestClassifier()

In [38]: RFC.score(xv_test,y_test)

Out[38]: 0.9893048128342246

In [39]: pred_RFC=RFC.predict(xv_test)

In [40]: print(classification_report(y_test,pred_RFC))

precision recall f1-score support
0 0.99 0.99 0.99 5856
1 0.99 0.99 0.99 5364
accuracy 0.99 0.99 0.99 11220
macro avg 0.99 0.99 0.99 11220
weighted avg 0.99 0.99 0.99 11220

In [44]: def manual_testing(news):
testing_news=({"text": [news]})
new_def_test=pd.DataFrame(testing_news)
new_def_test["text"]=new_def_test["text"].apply(word_drop)
new_x_test=new_def_test["text"]
new_xv_test=vectorization.transform(new_x_test)
pred_LR=LR.predict(xv_test)
pred_DT=DT.predict(xv_test)
pred_GBC=GBC.predict(xv_test)
pred_RFC=RFC.predict(xv_test)
return print("\n\nLR Prediction: {} \nDT Prediction: {} \nGBC Prediction: {} \nRFC Prediction: {}".format(pred_LR,
pred_DT,
pred_GBC,
pred_RFC))

In [46]: 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 ensna red 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 bu ilder'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: [0 0 0 ... 0 0 1]
DT Prediction: [0 0 0 ... 0 0 1]
GBC Prediction: [0 0 0 ... 0 0 1]
RFC Prediction: [0 0 0 ... 0 0 1]

In [ ]:

In [ ]:
```