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

```
true = pd.read_csv('True.csv')
fake = pd.read_csv('Fake.csv')
```

```
true.head()
```

	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. 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
4	Trump wants Postal Service to charge 'much mor...	SEATTLE/WASHINGTON (Reuters) - President Donal...	politicsNews	December 29, 2017

```
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

```
true['label']=1
fake['label']=0
```

```
news = pd.concat([true,fake])
```

```
news.isnull().sum()
```

	0
title	0
text	0
subject	0
date	0
label	0

```
news = news.drop(['title','subject','date'],axis=1)
```

```
news.head()
```

	text	label
0	WASHINGTON (Reuters) - The head of a conservat...	1
1	WASHINGTON (Reuters) - Transgender people will...	1
2	WASHINGTON (Reuters) - The special counsel inv...	1
3	WASHINGTON (Reuters) - Trump campaign adviser ...	1
4	SEATTLE/WASHINGTON (Reuters) - President Donal...	1

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

```
1 news.head()
```

	text	label
6018	Beebe, Arkansas mayor, Mike Robertson, loves J...	0
9141	Many are suggesting that the former comedian, ...	0
5690	WASHINGTON (Reuters) - The U.S. Treasury said ...	1
17219	PARIS (Reuters) - President Donald Trump s haw...	1
4053	Colin Powell is a Republican, but he voted for	0

```
1 news.reset_index(inplace=True)
```

```
1 news.head()
```

	index	text	label
0	6018	Beebe, Arkansas mayor, Mike Robertson, loves J...	0
1	9141	Many are suggesting that the former comedian, ...	0
2	5690	WASHINGTON (Reuters) - The U.S. Treasury said ...	1
3	17219	PARIS (Reuters) - President Donald Trump s haw...	1
4	4053	Colin Powell is a Republican, but he voted for	0

```
1 news.drop(['index'],axis=1,inplace=True)
```

```
1 news.head()
```

	text	label
0	WASHINGTON (Reuters) - U.S. President Donald T...	1
1	At least 26 people have been killed inside a T...	0
2	It turns out that William F. Buckley s own son...	0
3	Commies carrying flags fought with the Austin ...	0
4	WASHINGTON (Reuters) - Congress on Wednesday o	1

Next steps: Generate code with news | View recommended plots | New interactive sheet

```
1 import re
```

```
1 def wordopt(text):
2     #convert into lowercase
3     text = text.lower()
4
5     #Remove URLs
6     text = re.sub(r'https?://\S+|www\.\S+', '', text)
7
8     #Remove HTML Tags
9     text = re.sub(r'<.*?>+', '', text)
10
11    #Remove punctuations
12    text = re.sub(r'[^\w\s]','',text)
13
14    #Remove digits
15    text = re.sub(r'\d','',text)
16
17    #Remove newline characters
18    text = re.sub(r'\n','',text)
19
20    return text
```

```
1 news['text'] = news['text'].apply(wordopt)
```

```
1 news['text']
```

	text
0	washington reuters us president donald trump ...
1	at least people have been killed inside a tex...
2	it turns out that william f buckley s own son ...
3	commies carrying flags fought with the austin ...
4	washington reuters congress on wednesday over...
...	...
44893	just another tolerant violent liberal shooting...
44894	please rt until we identify this disgusting n...
44895	prague reuters the antiestablishment ano part...
44896	reuters the united states and china are targe...
44897	some people like to create solutions to proble...

44898 rows × 1 columns

```
1 x = news['text']
2 y = news['label']
```

```
1 x
```

	text
0	washington reuters us president donald trump ...
1	at least people have been killed inside a tex...
2	it turns out that william f buckley s own son ...
3	commies carrying flags fought with the austin ...
4	washington reuters congress on wednesday over...
...	...
44893	just another tolerant violent liberal shooting...
44894	please rt until we identify this disgusting n...
44895	prague reuters the antiestablishment ano part...
44896	reuters the united states and china are targe...
44897	some people like to create solutions to proble...

44898 rows × 1 columns

dtype: object

```
1 y
```

	label
0	1
1	0
2	0
3	0
4	1
...	...
44893	0
44894	0
44895	1
44896	1
44897	0

44898 rows × 1 columns

dtype: int64

```
from sklearn.model_selection import train_test_split
```

```
x_train,x_test,y_train,y_test = train_test_split(x,y,test_size=0.3)
```

```
x_train.shape
```

```
(31428,)
```

```
x_test.shape
```

```
(13470,)
```

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

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

```
xv_train
```

```
<31428x174834 sparse matrix of type '<class 'numpy.float64'>'
        with 6451458 stored elements in Compressed Sparse Row format>
```

```
xv_test
```

```
<13470x174834 sparse matrix of type '<class 'numpy.float64'>'
        with 2728959 stored elements in Compressed Sparse Row format>
```

```
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)
```

```
0.9875278396436525
```

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

```
              precision    recall  f1-score   support

           0       0.99      0.99      0.99      7079
           1       0.98      0.99      0.99      6391

    accuracy                           0.99     13470
   macro avg       0.99      0.99      0.99     13470
weighted avg       0.99      0.99      0.99     13470
```

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

```
DTC = DecisionTreeClassifier()
DTC = DTC.fit(xv_train,y_train)
```

```
pred_dtc=DTC.predict(xv_test)
```

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

```
0.9956941351150705
```

```
print(classification_report(y_test,pred_dtc))
```

```
              precision    recall  f1-score   support

           0       1.00      1.00      1.00      7079
           1       1.00      1.00      1.00      6391

    accuracy                           1.00     13470
   macro avg       1.00      1.00      1.00     13470
weighted avg       1.00      1.00      1.00     13470
```

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

```
rfc = RandomForestClassifier()
rfc.fit(xv_train,y_train)
```

```
▾ RandomForestClassifier
RandomForestClassifier()
```

```
predict_rfc=rfc.predict(xv_test)
```

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

```
0.9870081662954714
```

```
print(classification_report(y_test,predict_rfc))
```

```
              precision    recall  f1-score   support

           0       0.99      0.99      0.99      7079
           1       0.98      0.99      0.99      6391

    accuracy                           0.99     13470
   macro avg       0.99      0.99      0.99     13470
weighted avg       0.99      0.99      0.99     13470
```

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

```
gbc=GradientBoostingClassifier()
gbc.fit(xv_train,y_train)
```

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

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

```
0.9958426132145508
```

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

```
              precision    recall  f1-score   support

           0       1.00      0.99      1.00      7079
           1       0.99      1.00      1.00      6391

    accuracy                           1.00     13470
   macro avg       1.00      1.00      1.00     13470
weighted avg       1.00      1.00      1.00     13470
```

```
def output_label(n):
  if n==0:
    return "It's a Fake News"
  elif n == 1:
    return "It's a Genuine 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_dtc = DTC.predict(new_xv_test)
  pred_rfc = rfc.predict(new_xv_test)
  #pred_gbc = gbc.predict(new_xv_test)
  return print("\n\nLR Prediction:{} \nDTC Prediction:{} \nRFC Prediction:{} \nGBC Prediction:{}".format(output_label(pred_lr[0]),output
```

```
 news_article = str(input())
```

```
U.S. launches effort to reduce reliance on imports or critical minerals
```

```
manual_testing(news_article)
```

```


LR Prediction:It's a Fake News
DTC Prediction:It's a Fake News
RFC Prediction:It's a Fake News
GBC Prediction:It's a Fake News
```

```
Start coding or generate with AI.
```