```
In [1]:  import pandas as pd
         import numpy as np
```

```
In [3]:  true = pd.read_csv('/content/sample_data/True.csv')
```

```
In [4]:  fake = pd.read_csv('/content/sample_data/Fake.csv')
```

```
In [5]:  true.head()
```

Out[5]:

	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

```
In [6]:  fake.head()
```

Out[6]:

	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

```
In [7]:  true['label']=1
```

```
In [8]:  fake['label']=0
```

```
In [9]:  true.head()
```

Out[9]:

	title	text	subject	date	label
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

In [10]:
```python
news = pd.concat([fake,true],axis=0)
```

In [11]:
```python
news.head()
```

Out[11]:

	title	text	subject	date	label
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

In [12]:
```python
news.tail()
```

	title	text	subject	date	label
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

In [13]: `news.isnull().sum()`

Out[13]:

	0
title	0
text	0
subject	0
date	0
label	0

dtype: int64

In [14]: `news = news.drop(['title','subject','date'],axis= 1)`

In [15]: `news.head()`

Out[15]:

	text	label
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

In [16]: `news = news.sample(frac=1)`

```
In [17]:  news.head()
```

Out[17]:

	text	label
13506	TEGUCIGALPA (Reuters) - Honduras faced a growi...	1
2082	Congress is chest-deep in their leak investiga...	0
2171	The following statements were posted to the ve...	1
17073	Will the new 3-story Islamic Museum include pi...	0
16479	Listen up people! This is great! Lou Dobbs get...	0

```
In [18]:  news.reset_index(inplace=True)
```

```
In [19]:  news.head()
```

Out[19]:

	index	text	label
0	13506	TEGUCIGALPA (Reuters) - Honduras faced a growi...	1
1	2082	Congress is chest-deep in their leak investiga...	0
2	2171	The following statements were posted to the ve...	1
3	17073	Will the new 3-story Islamic Museum include pi...	0
4	16479	Listen up people! This is great! Lou Dobbs get...	0

```
In [20]:  news.drop(['index'],axis=1,inplace=True)
```

```
In [21]:  news.head()
```

Out[21]:

	text	label
0	TEGUCIGALPA (Reuters) - Honduras faced a growi...	1
1	Congress is chest-deep in their leak investiga...	0
2	The following statements were posted to the ve...	1
3	Will the new 3-story Islamic Museum include pi...	0
4	Listen up people! This is great! Lou Dobbs get...	0

```
In [22]:  import re
```

```
In [23]:  def wordopt(text):

              #convert into lowercase
              text = text.lower()
```

```python
    #remove URLs
    text = re.sub(r'https?://\s+|www\.\s+','',text)

    #remove html tags
    text = re.sub(r'[^\w\s]','',text)

    #remove digits
    text = re.sub(r'\d','',text)

    #remove newline characters
    text = re.sub(r'\n', '', text)

    return text
```

In [24]: `news['text'] = news['text'].apply(wordopt)`

In [25]: `news['text']`

Out[25]:

	text
0	tegucigalpa reuters honduras faced a growing ...
1	congress is chestdeep in their leak investigat...
2	the following statements were posted to the ve...
3	will the new story islamic museum include pict...
4	listen up people this is great lou dobbs gets ...
...	...
44893	war veteran johnny joey jones praised presid...
44894	united nations reuters russia vetoed on frida...
44895	when loretta lynch was confirmed as the new di...
44896	thanks to george soros at least couches in th...
44897	washington reuters any tax reform plan that i...

44898 rows × 1 columns

dtype: object

In [26]:
```python
x = news['text']
y = news['label']
```

In [27]: `x`

	text
0	tegucigalpa reuters honduras faced a growing ...
1	congress is chestdeep in their leak investigat...
2	the following statements were posted to the ve...
3	will the new story islamic museum include pict...
4	listen up people this is great lou dobbs gets ...
...	...
44893	war veteran johnny joey jones praised presid...
44894	united nations reuters russia vetoed on frida...
44895	when loretta lynch was confirmed as the new di...
44896	thanks to george soros at least couches in th...
44897	washington reuters any tax reform plan that i...

44898 rows × 1 columns

dtype: object

y

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

44898 rows × 1 columns

dtype: int64

```python
In [29]: from sklearn.model_selection import train_test_split
```

```python
In [30]: x_train,x_test,y_train,y_test = train_test_split(x,y,test_size=0.3)
```

```python
In [31]: x_train.shape
```

```
Out[31]: (31428,)
```

```python
In [32]: x_test.shape
```

```
Out[32]: (13470,)
```

```python
In [34]: from sklearn.feature_extraction.text import TfidfVectorizer
```

```python
In [35]: vectorization = TfidfVectorizer()
```

```python
In [36]: xv_train = vectorization.fit_transform(x_train)
```

```python
In [37]: xv_test = vectorization.transform(x_test)
```

```python
In [38]: xv_train
```

```
Out[38]: <Compressed Sparse Row sparse matrix of dtype 'float64'
                 with 6468162 stored elements and shape (31428, 177696)>
```

```python
In [39]: xv_test
```

```
Out[39]: <Compressed Sparse Row sparse matrix of dtype 'float64'
                 with 2716332 stored elements and shape (13470, 177696)>
```

```python
In [40]: from sklearn.linear_model import LogisticRegression

         LR = LogisticRegression()    # define model

         LR.fit(xv_train, y_train)    # train model
```

```
Out[40]:  ▼ LogisticRegression ⓘ ⓘ

         LogisticRegression()
```

```python
In [41]: pred_lr = LR.predict(xv_test)
```

```python
In [42]: LR.score(xv_test,y_test)
```

```
Out[42]: 0.9861915367483296
```

```python
In [43]: 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      7002
           1       0.98      0.99      0.99      6468

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

In [44]:
```python
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      7002
           1       0.98      0.99      0.99      6468

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

In [47]:
```python
from sklearn.tree import DecisionTreeClassifier
```

In [48]:
```python
DTC = DecisionTreeClassifier()
```

In [49]:
```python
DTC.fit(xv_train,y_train)
```

Out[49]:
▼ DecisionTreeClassifier ① ⑦

DecisionTreeClassifier()

In [50]:
```python
pred_dtc = DTC.predict(xv_test)
```

In [51]:
```python
DTC.score(xv_test,y_test)
```

Out[51]: 0.9956198960653304

In [52]:
```python
print(classification_report(y_test,pred_dtc))
```

```
              precision    recall  f1-score   support

           0       1.00      1.00      1.00      7002
           1       0.99      1.00      1.00      6468

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

```python
In [55]:   from sklearn.ensemble import RandomForestClassifier
```

```python
In [56]:   rfc = RandomForestClassifier()
```

```python
In [58]:   rfc.fit(xv_train,y_train)
```

Out[58]: ▾ RandomForestClassifier ① ⑦

RandomForestClassifier()

```python
In [59]:   predict_rfc = rfc.predict(xv_test)
```

```python
In [60]:   rfc.score(xv_test,y_test)
```

Out[60]: 0.9867854491462509

```python
In [61]:   print(classification_report(y_test,predict_rfc))
```

```
               precision    recall  f1-score   support

           0       0.99      0.99      0.99      7002
           1       0.99      0.99      0.99      6468

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

```python
In [62]:   from sklearn.ensemble import GradientBoostingClassifier
```

```python
In [63]:   gbc = GradientBoostingClassifier()
```

```python
In [64]:   gbc.fit(xv_train, y_train)
```

Out[64]: ▾ GradientBoostingClassifier ① ⑦

GradientBoostingClassifier()

```python
In [67]:   pred_gbc = gbc.predict(xv_test)
```

```python
In [68]:   gbc.score(xv_test, y_test)
```

Out[68]: 0.9959910913140312

```python
In [69]:   print(classification_report(y_test, pred_gbc))
```

```
              precision    recall  f1-score   support

           0       1.00      0.99      1.00      7002
           1       0.99      1.00      1.00      6468

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

In [ ]: