```python
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
```

In [1]:

In [2]:
```python
data_fake = pd.read_csv('Fake.csv.zip')
data_true = pd.read_csv('True.csv.zip')
```

In [3]:
```python
data_fake.head()
```

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 |

In [4]:
```python
data_true.head()
```

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. 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 [5]:
```python
data_fake["class"] = 0
data_true["class"] = 1
```

In [6]:
```python
data_fake.shape, data_true.shape
```

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

In [7]:
```python
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)
```

Loading [MathJax]/extensions/Safe.js

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

In [8]: `data_fake.shape, data_true.shape`

Out[8]: `((23471, 5), (21407, 5))`

In [9]:
```
data_fake_manual_testing['class'] = 0
data_true_manual_testing['class'] = 0
```

```
C:\Users\Jananisha\AppData\Local\Temp\ipykernel_5944\408239254.py:1: SettingWithCopyWarn
ing:
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.html#returning-a-view-versus-a-copy
  data_fake_manual_testing['class'] = 0
C:\Users\Jananisha\AppData\Local\Temp\ipykernel_5944\408239254.py:2: SettingWithCopyWarn
ing:
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.html#returning-a-view-versus-a-copy
  data_true_manual_testing['class'] = 0
```

In [10]: `data_fake_manual_testing.head(10)`

Out[10]:

| | title | text | subject | date | 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 CounterpunchAlthough 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 |

In [11]: `data_true_manual_testing.head(10)`

Loading [MathJax]/extensions/Safe.js

| | title | text | subject | date | class |
|---|---|---|---|---|---|
| **21407** | Mata Pires, owner of embattled Brazil builder ... | SAO PAULO (Reuters) - Cesar Mata Pires, the ow... | worldnews | August 22, 2017 | 0 |
| **21408** | U.S., North Korea clash at U.N. forum over nuc... | GENEVA (Reuters) - North Korea and the United ... | worldnews | August 22, 2017 | 0 |
| **21409** | U.S., North Korea clash at U.N. arms forum on ... | GENEVA (Reuters) - North Korea and the United ... | worldnews | August 22, 2017 | 0 |
| **21410** | Headless torso could belong to submarine journ... | COPENHAGEN (Reuters) - Danish police said on T... | worldnews | August 22, 2017 | 0 |
| **21411** | North Korea shipments to Syria chemical arms a... | UNITED NATIONS (Reuters) - Two North Korean sh... | worldnews | August 21, 2017 | 0 |
| **21412** | 'Fully committed' NATO backs new U.S. approach... | BRUSSELS (Reuters) - NATO allies on Tuesday we... | worldnews | August 22, 2017 | 0 |
| **21413** | LexisNexis withdrew two products from Chinese ... | LONDON (Reuters) - LexisNexis, a provider of l... | worldnews | August 22, 2017 | 0 |
| **21414** | Minsk cultural hub becomes haven from authorities | MINSK (Reuters) - In the shadow of disused Sov... | worldnews | August 22, 2017 | 0 |
| **21415** | Vatican upbeat on possibility of Pope Francis ... | MOSCOW (Reuters) - Vatican Secretary of State ... | worldnews | August 22, 2017 | 0 |
| **21416** | Indonesia to buy $1.14 billion worth of Russia... | JAKARTA (Reuters) - Indonesia will buy 11 Sukh... | worldnews | August 22, 2017 | 0 |

In [12]:
```python
data_merge = pd.concat([data_fake, data_true], axis = 0)
data_merge.head(10)
```

| | 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 [13]:
```python
data_merge.columns
```

Out[13]:
```
Index(['title', 'text', 'subject', 'date', 'class'], dtype='object')
```

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

Loading [MathJax]/extensions/Safe.js

```
In [15]:  data.isnull().sum()

Out[15]:  text    0
          class   0
          dtype: int64

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

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

Out[17]:

|        | text | class |
|--------|------|-------|
| 3557   | WASHINGTON (Reuters) - U.S. lawmakers introduc... | 1 |
| 1197   | White House deputy press secretary Sarah Hucka... | 0 |
| 20107  | ANKARA (Reuters) - Turkey remains a dependable... | 1 |
| 847    | Now that Donald Trump s Russia scandal has bee... | 0 |
| 21345  | The USPS is drowning in debt, has lost every o... | 0 |

```
In [18]:  data.reset_index(inplace = True)
          data.drop(['index'], axis = 1, inplace = True)

In [19]:  data.columns

Out[19]:  Index(['text', 'class'], dtype='object')

In [20]:  data.head()
```

Out[20]:

|   | text | class |
|---|------|-------|
| 0 | WASHINGTON (Reuters) - U.S. lawmakers introduc... | 1 |
| 1 | White House deputy press secretary Sarah Hucka... | 0 |
| 2 | ANKARA (Reuters) - Turkey remains a dependable... | 1 |
| 3 | Now that Donald Trump s Russia scandal has bee... | 0 |
| 4 | The USPS is drowning in debt, has lost every o... | 0 |

```
In [21]:  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

In [22]:  data['text'] = data['text'].apply(wordopt)

In [23]:  x = data['text']
          y = data['class']

In [24]:  x_train, x_test, y_train, y_test = train_test_split(x,y, test_size=0.25)
```

Loading [MathJax]/extensions/Safe.js

```
In [25]:  from sklearn.feature_extraction.text import TfidfVectorizer

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

```
In [26]:  from sklearn.linear_model import LogisticRegression

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

Out[26]:  ▼ LogisticRegression

          LogisticRegression()

```
In [27]:  pred_lr = LR.predict(xv_test)
```

```
In [28]:  LR.score(xv_test, y_test)
```

Out[28]:  0.9871657754010695

```
In [29]:  print(classification_report(y_test, pred_lr))
```

```
                        precision    recall  f1-score   support

                   0        0.99      0.99      0.99      5817
                   1        0.99      0.99      0.99      5403

            accuracy                            0.99     11220
           macro avg        0.99      0.99      0.99     11220
        weighted avg        0.99      0.99      0.99     11220
```

```
In [31]:  from sklearn.tree import DecisionTreeClassifier

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

Out[31]:  ▼ DecisionTreeClassifier

          DecisionTreeClassifier()

```
In [32]:  pred_dt = DT.predict(xv_test)
```

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

Out[33]:  0.996078431372549

```
In [34]:  print(classification_report(y_test, pred_lr))
```

```
                        precision    recall  f1-score   support

                   0        0.99      0.99      0.99      5817
                   1        0.99      0.99      0.99      5403

            accuracy                            0.99     11220
           macro avg        0.99      0.99      0.99     11220
        weighted avg        0.99      0.99      0.99     11220
```

```python
In [35]:    from sklearn.ensemble import GradientBoostingClassifier

            GB = GradientBoostingClassifier(random_state = 0)
            GB.fit(xv_train, y_train)
```

Out[35]:    ▼          GradientBoostingClassifier

            GradientBoostingClassifier(random_state=0)

```python
In [36]:    pred_gb = GB.predict(xv_test)
```

```python
In [37]:    GB.score(xv_test, y_test)
```

Out[37]:    0.9948306595365419

```python
In [39]:    print(classification_report(y_test, pred_gb))
```

```
                  precision    recall  f1-score   support

               0       1.00      0.99      1.00      5817
               1       0.99      1.00      0.99      5403

        accuracy                           0.99     11220
       macro avg       0.99      0.99      0.99     11220
    weighted avg       0.99      0.99      0.99     11220
```

```python
In [40]:    from sklearn.ensemble import RandomForestClassifier

            RF = RandomForestClassifier(random_state = 0)
            RF.fit(xv_train, y_train)
```

Out[40]:    ▼          RandomForestClassifier

            RandomForestClassifier(random_state=0)

```python
In [45]:    pred_rf = RF.predict(xv_test)
```

```python
In [46]:    RF.score(xv_test, y_test)
```

Out[46]:    0.9881461675579323

```python
In [47]:    print(classification_report(y_test, pred_rf))
```

```
                  precision    recall  f1-score   support

               0       0.99      0.99      0.99      5817
               1       0.99      0.98      0.99      5403

        accuracy                           0.99     11220
       macro avg       0.99      0.99      0.99     11220
    weighted avg       0.99      0.99      0.99     11220
```

```python
In [76]:    def output_label(n):
                if n == 0:
                    return "Fake News"
                elif n == 1:
                    return "Not A Fake News"
```

Loading [MathJax]/extensions/Safe.js `sting(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_GB = GB.predict(new_xv_test)
        pred_RF = RF.predict(new_xv_test)

        return print("\n\nLR Prediction: {} \nDT Prediction: {} \nGBC Prediction: {} \nRFC P
```

In [77]:
```
news = str(input())
manual_testing(news)
```

The news on Republic TV about Jama Masjid being in dark due to non-payment of electricity bills will make it to any top ten list of fake news. This was fake news that originated from various Hindutva handles and Postcard News, a renowned fake news website. Republic TV sent a fact finding team to the site, which snooped around outside Imam Bukhari's residence counting cars and recording their make but didn't ring the bell to verify if the story was true. It also did not occur to the team to ask those around whether the mosque is usually lit at night and what time the lights are normally switched off at night. Funnily the reporter pointed out to a lit board at the gate but it didn't occur to him how the board could be lit when the electricity was cut off. On the basis of this so-called investigation, ignoring the clarifications tweeted by BSES, Republic TV broke the news about BSES giving a jolt to Jama Masjid due to non-payment of over four crores. The fake news was exposed by Alt News in its article after which the channel quietly deleted its tweet and video without any apology or explanation.


LR Prediction: Fake News
DT Prediction: Fake News
GBC Prediction: Fake News
RFC Prediction: Fake News