In [1]:
```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
from sklearn.feature_extraction.text import TfidfVectorizer
from sklearn.linear_model import LogisticRegression
```

In [2]:
```python
dataframe_fake = pd.read_csv("Fake.csv")
dataframe_true = pd.read_csv("True.csv")
dataframe_fake.head()
```

Out[2]:

	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 [3]:
```python
dataframe_fake["class"] = 0
dataframe_true["class"] = 1
```

In [4]:
```python
dataframe_fake.shape, dataframe_true.shape
```

Out[4]: ((23481, 5), (21417, 5))

In [5]:
```python
# Correct index resetting after dropping rows
dataframe_fake_manual_testing = dataframe_fake.tail(10)
for i in range(23480, 23470, -1):
    dataframe_fake.drop([i], axis=0, inplace=True)

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

In [6]:
```python
# Concatenate and reset index
dataframe_merge = pd.concat([dataframe_fake, dataframe_true], axis=0)
dataframe_merge.reset_index(drop=True, inplace=True)
```

In [7]:
```python
dataframe = dataframe_merge.drop(["title", "subject", "date"], axis=1)
```

In [8]:
```python
# Check for null values
dataframe.isnull().sum()
```

Out[8]:
```
text     0
class    0
dtype: int64
```

In [9]:
```python
dataframe = dataframe.sample(frac=1)
dataframe.reset_index(inplace=True)
dataframe.drop(["index"], axis=1, inplace=True)
```

In [10]:
```python
def wordopt(t):
    tt = t.lower()
    t = re.sub('\[.*?\]', '', t)
    t = re.sub("\\W", " ", t)
    t = re.sub('https?://\S+|www\.\S+', '', t)
    t = re.sub('<.*?>+', '', t)
    t = re.sub('[%s]' % re.escape(string.punctuation), '', t)
    t = re.sub('\n', '', t)
    t = re.sub('\w*\d\w*', '', t)
    return t

dataframe["text"] = dataframe["text"].apply(wordopt)

x = dataframe["text"]
y = dataframe["class"]

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

In [11]:
```python
vectorization = TfidfVectorizer()
xv_train = vectorization.fit_transform(x_train)
xv_test = vectorization.transform(x_test)
```

In [12]:
```python
LR = LogisticRegression()
LR.fit(xv_train, y_train)
pred_lr = LR.predict(xv_test)
print("LR Classification Report:\n", classification_report(y_test, pred_lr))
```

```
LR Classification Report:
               precision    recall  f1-score   support

           0       0.99      0.99      0.99      5885
           1       0.99      0.99      0.99      5335

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

In [13]:
```python
from sklearn.tree import DecisionTreeClassifier


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

Out[13]:
```
▾ DecisionTreeClassifier
DecisionTreeClassifier()
```

In [14]:
```python
pred_dt = DT.predict(xv_test)
```

In [15]:
```python
print("Decision Tree Classification Report:\n", classification_report(y_test, pred_dt))
```

```
Decision Tree Classification Report:
               precision    recall  f1-score   support

           0       1.00      1.00      1.00      5885
           1       1.00      1.00      1.00      5335

    accuracy                           1.00     11220
   macro avg       1.00      1.00      1.00     11220
weighted avg       1.00      1.00      1.00     11220
```

```
In [16]: from sklearn.ensemble import GradientBoostingClassifier


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

Out[16]:
```
    ▼         GradientBoostingClassifier
    GradientBoostingClassifier(random_state=0)
```

```
In [17]: pred_gbc = GBC.predict(xv_test)
         GBC.score(xv_test, y_test)
```

Out[17]: 0.9967023172905526

```
In [18]: print(classification_report(y_test, pred_gbc))
```

```
                       precision    recall  f1-score   support

                   0       1.00      1.00      1.00      5885
                   1       1.00      1.00      1.00      5335

            accuracy                           1.00     11220
           macro avg       1.00      1.00      1.00     11220
        weighted avg       1.00      1.00      1.00     11220
```

```
In [19]: from sklearn.ensemble import RandomForestClassifier


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

Out[19]:
```
    ▼         RandomForestClassifier
    RandomForestClassifier(random_state=0)
```

```
In [20]: pred_rfc = RFC.predict(xv_test)
         print("Random Forest Classification Report:\n", classification_report(y_test, pred_rfc)
```

```
        Random Forest Classification Report:
                       precision    recall  f1-score   support

                   0       0.99      0.99      0.99      5885
                   1       0.99      0.99      0.99      5335

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

```
In [22]: CART = DecisionTreeClassifier(criterion='gini')
         CART.fit(xv_train, y_train)
```

Out[22]:
```
    ▼ DecisionTreeClassifier
    DecisionTreeClassifier()
```

In [23]:
```python
pred_cart = CART.predict(xv_test)
print("CART Classification Report:\n", classification_report(y_test, pred_cart))
```

```
CART Classification Report:
               precision    recall  f1-score   support

           0       1.00      1.00      1.00      5885
           1       1.00      1.00      1.00      5335

    accuracy                           1.00     11220
   macro avg       1.00      1.00      1.00     11220
weighted avg       1.00      1.00      1.00     11220
```

In [24]:
```python
ID3 = DecisionTreeClassifier(criterion='entropy')
ID3.fit(xv_train, y_train)
```

Out[24]:
```
▼          DecisionTreeClassifier
DecisionTreeClassifier(criterion='entropy')
```

In [25]:
```python
pred_id3 = ID3.predict(xv_test)
print("ID3 Classification Report:\n", classification_report(y_test, pred_id3))
```

```
ID3 Classification Report:
               precision    recall  f1-score   support

           0       1.00      1.00      1.00      5885
           1       1.00      1.00      1.00      5335

    accuracy                           1.00     11220
   macro avg       1.00      1.00      1.00     11220
weighted avg       1.00      1.00      1.00     11220
```

In [26]:
```python
def output_label(n):
    if n == 0:
        return "Fake News"
    elif n == 1:
        return "Not A Fake 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_DT = DT.predict(new_xv_test)
    pred_GBC = GBC.predict(new_xv_test)
    pred_RFC = RFC.predict(new_xv_test)

    print("\n\nLR Prediction: {} \nDT Prediction: {} \nGBC Prediction: {} \nRFC Predict
        output_label(pred_LR[0]),
        output_label(pred_DT[0]),
        output_label(pred_GBC[0]),
        output_label(pred_RFC[0])
    ))

news = str(input("Enter news for manual testing: "))
manual_testing(news)
```

Enter news for manual testing: WASHINGTON (Reuters) — The head of a conservative Republican faction in the U.S. Congress, who voted this month for a huge expansion of the national debt to pay for tax cuts, called himself a "fiscal conservative" on Sunday and urged budget restraint in 2018. In keeping with a sharp pivot under way among Republicans, U.S. Representative Mark Meadows, speaking on CBS' "Face the Nation," drew a hard line on federal spending, which lawmakers are bracing to do battle over in January. When they return from the holidays on Wednesday, lawmakers will begin trying to pass a federal budget in a fight likely to be linked to other issues, such as immigration policy, even as the November congressional election campaigns approach in which Republicans will seek to keep control of Congress. President Donald Trump and his Republicans want a big budget increase in military spending, while Democrats also want proportional increases for non-defense "discretionary" spending on programs that support education, scientific research, infrastructure, public health and environmental protection. "The (Trump) administration has already been willing to say: 'We're going to increase non-defense discretionary spending ... by about 7 percent,'" Meadows, chairman of the small but influential House Freedom Caucus, said on the program. "Now, Democrats are saying that's not enough, we need to give the government a pay raise of 10 to 11 percent. For a fiscal conservative, I don't see where the rationale is. ... Eventually you run out of other people's money," he said. Meadows was among Republicans who voted in late December for their party's debt-financed tax overhaul, which is expected to balloon the federal budget deficit and add about $1.5 trillion over 10 years to the $20 trillion national debt. "It's interesting to hear Mark talk about fiscal responsibility," Democratic U.S. Representative Joseph Crowley said on CBS. Crowley said the Republican tax bill would require the United States to borrow $1.5 trillion, to be paid off by future generations, to finance tax cuts for corporations and the rich. "This is one of the least ... fiscally responsible bills we've ever seen passed in the history of the House of Representatives. I think we're going to be paying for this for many, many years to come," Crowley said. Republicans insist the tax package, the biggest U.S. tax overhaul in more than 30 years, will boost the economy and job growth. House Speaker Paul Ryan, who also supported the tax bill, recently went further than Meadows, making clear in a radio interview that welfare or "entitlement reform," as the party often calls it, would be a top Republican priority in 2018. In Republican parlance, "entitlement" programs mean food stamps, housing assistance, Medicare and Medicaid health insurance for the elderly, poor and disabled, as well as other programs created by Washington to assist the needy. Democrats seized on Ryan's early December remarks, saying they showed Republicans would try to pay for their tax overhaul by seeking spending cuts for social programs. But the goals of House Republicans may have to take a back seat to the Senate, where the votes of some Democrats will be needed to approve a budget and prevent a government shutdown. Democrats will use their leverage in the Senate, which Republicans narrowly control, to defend both discretionary non-defense programs and social spending, while tackling the issue of the "Dreamers," people brought illegally to the country as children. Trump in September put a March 2018 expiration date on the Deferred Action for Childhood Arrivals, or DACA, program, which protects the young immigrants from deportation and provides them with work permits. The president has said in recent Twitter messages he wants funding for his proposed Mexican border wall and other immigration law changes in exchange for agreeing to help the Dreamers. Representative Debbie Dingell told CBS she did not favor linking that issue to other policy objectives, such as wall funding. "We need to do DACA clean," she said. On Wednesday, Trump aides will meet with congressional leaders to discuss those issues. That will be followed by a weekend of strategy sessions for Trump and Republican leaders on Jan. 6 and 7, the White House said. Trump was also scheduled to meet on Sunday with Florida Republican Governor Rick Scott, who wants more emergency aid. The House has passed an $81 billion aid package after hurricanes in Florida, Texas and Puerto Rico, and wildfires in California. The package far exceeded the $44 billion requested by the Trump administration. The Senate has not yet voted on the aid.

LR Prediction: Not A Fake News
DT Prediction: Not A Fake News
GBC Prediction: Not A Fake News
RFC Prediction: Not A Fake News

In [ ]: