Fake News Detection

July 9, 2025

1 Fake News Detection Analysis

(By: Gautam Yadav)

1.1 Datasets Description

- **Fake.csv:** A full training dataset containing fake news.
- **True.csv:** A full training dataset containing true news.

Attributes: (Both of the datasets have same attributes).

- **Title:** The title of a news article.
- **Text:** Full news in text.
- **Subject:** Subject or category of a news.
- **Date:** Date of news published.

Class:

- **Fake news:** 0
- **True news:** 1

1.2 Import Libraries

```python
[1]: 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
```

1.3 Load Datasets

```python
[2]: data_fake = pd.read_csv("C:\\Users\\DELL\\Downloads\\Fake.csv")
     data_true = pd.read_csv("C:\\Users\\DELL\\Downloads\\True.csv")
```

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

```
[3]:                                                        title  \
     0   Donald Trump Sends Out Embarrassing New Year'…
     1   Drunk Bragging Trump Staffer Started Russian …
     2   Sheriff David Clarke Becomes An Internet Joke…
     3   Trump Is So Obsessed He Even Has Obama's Name…
     4   Pope Francis Just Called Out Donald Trump Dur…

                                                  text subject  \
     0   Donald Trump just couldn t wish all Americans …    News
     1   House Intelligence Committee Chairman Devin Nu…    News
     2   On Friday, it was revealed that former Milwauk…    News
     3   On Christmas day, Donald Trump announced that …    News
     4   Pope Francis used his annual Christmas Day mes…    News

                 date
     0   December 31, 2017
     1   December 31, 2017
     2   December 30, 2017
     3   December 29, 2017
     4   December 25, 2017
```

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

```
[4]:                                                        title  \
     0   As U.S. budget fight looms, Republicans flip t…
     1   U.S. military to accept transgender recruits o…
     2   Senior U.S. Republican senator: 'Let Mr. Muell…
     3   FBI Russia probe helped by Australian diplomat…
     4   Trump wants Postal Service to charge 'much mor…

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

                 date
     0   December 31, 2017
     1   December 29, 2017
     2   December 31, 2017
     3   December 30, 2017
     4   December 29, 2017
```

1.4 Data Preprocessing

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

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

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

```
[7]: data_fake_manual_testing = data_fake.tail(10)
     data_fake = data_fake.iloc[:-10]

     data_true_manual_testing = data_true.tail(10)
     data_true = data_true.iloc[:-10]
```

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

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

```
[9]: data_fake_manual_testing["class"] = 0
     data_true_manual_testing["class"] = 1
```

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

```
[10]:                                                    title  \
      23471  Seven Iranians freed in the prisoner swap have…
      23472                       #Hashtag Hell & The Fake Left
      23473  Astroturfing: Journalist Reveals Brainwashing …
      23474          The New American Century: An Era of Fraud
      23475  Hillary Clinton: 'Israel First' (and no peace …
      23476  McPain: John McCain Furious That Iran Treated …
      23477  JUSTICE? Yahoo Settles E-mail Privacy Class-ac…
      23478  Sunnistan: US and Allied 'Safe Zone' Plan to T…
      23479  How to Blow $700 Million: Al Jazeera America F…
      23480  10 U.S. Navy Sailors Held by Iranian Military …


                                                      text      subject  \
      23471  21st Century Wire says This week, the historic…  Middle-east
      23472   By Dady Chery and Gilbert MercierAll writers …  Middle-east
      23473  Vic Bishop Waking TimesOur reality is carefull…  Middle-east
      23474  Paul Craig RobertsIn the last years of the 20t…  Middle-east
      23475  Robert Fantina CounterpunchAlthough the United…  Middle-east
      23476  21st Century Wire says As 21WIRE reported earl…  Middle-east
      23477  21st Century Wire says It s a familiar theme. …  Middle-east
      23478  Patrick Henningsen  21st Century WireRemember …  Middle-east
      23479  21st Century Wire says Al Jazeera America will…  Middle-east
      23480  21st Century Wire says As 21WIRE predicted in …  Middle-east
```

```
              date  class
23471  January 20, 2016      0
23472  January 19, 2016      0
23473  January 19, 2016      0
23474  January 19, 2016      0
23475  January 18, 2016      0
23476  January 16, 2016      0
23477  January 16, 2016      0
23478  January 15, 2016      0
23479  January 14, 2016      0
23480  January 12, 2016      0
```

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

[11]:
```
                                                 title  \
21407  Mata Pires, owner of embattled Brazil builder …
21408  U.S., North Korea clash at U.N. forum over nuc…
21409  U.S., North Korea clash at U.N. arms forum on …
21410  Headless torso could belong to submarine journ…
21411  North Korea shipments to Syria chemical arms a…
21412  'Fully committed' NATO backs new U.S. approach…
21413  LexisNexis withdrew two products from Chinese …
21414  Minsk cultural hub becomes haven from authorities
21415  Vatican upbeat on possibility of Pope Francis …
21416  Indonesia to buy $1.14 billion worth of Russia…


                                                  text     subject  \
21407  SAO PAULO (Reuters) - Cesar Mata Pires, the ow…  worldnews
21408  GENEVA (Reuters) - North Korea and the United …  worldnews
21409  GENEVA (Reuters) - North Korea and the United …  worldnews
21410  COPENHAGEN (Reuters) - Danish police said on T…  worldnews
21411  UNITED NATIONS (Reuters) - Two North Korean sh…  worldnews
21412  BRUSSELS (Reuters) - NATO allies on Tuesday we…  worldnews
21413  LONDON (Reuters) - LexisNexis, a provider of l…  worldnews
21414  MINSK (Reuters) - In the shadow of disused Sov…  worldnews
21415  MOSCOW (Reuters) - Vatican Secretary of State …  worldnews
21416  JAKARTA (Reuters) - Indonesia will buy 11 Sukh…  worldnews


                  date  class
21407  August 22, 2017      1
21408  August 22, 2017      1
21409  August 22, 2017      1
21410  August 22, 2017      1
21411  August 21, 2017      1
21412  August 22, 2017      1
21413  August 22, 2017      1
21414  August 22, 2017      1
```

```
21415   August 22, 2017        1
21416   August 22, 2017        1
```

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

```
[12]:                                                     title  \
      0    Donald Trump Sends Out Embarrassing New Year'…
      1    Drunk Bragging Trump Staffer Started Russian …
      2    Sheriff David Clarke Becomes An Internet Joke…
      3    Trump Is So Obsessed He Even Has Obama's Name…
      4    Pope Francis Just Called Out Donald Trump Dur…
      5    Racist Alabama Cops Brutalize Black Boy While…
      6    Fresh Off The Golf Course, Trump Lashes Out A…
      7    Trump Said Some INSANELY Racist Stuff Inside …
      8    Former CIA Director Slams Trump Over UN Bully…
      9    WATCH: Brand-New Pro-Trump Ad Features So Muc…

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

                       date  class
      0    December 31, 2017      0
      1    December 31, 2017      0
      2    December 30, 2017      0
      3    December 29, 2017      0
      4    December 25, 2017      0
      5    December 25, 2017      0
      6    December 23, 2017      0
      7    December 23, 2017      0
      8    December 22, 2017      0
      9    December 21, 2017      0
```

1.5 Remove Unused Columns and Merge Columns

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

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

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

1.6 Check NaN Values

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

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

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

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

```
[17]:                                                     text  class
      3274    The Republican National Committee released a b…      0
      19433   KIEV/NEW YORK (Reuters) - Ukraine s Petro Poro…      1
      3945    On Friday, a jury did something that will cert…      0
      11979   ANKARA (Reuters) - Turkish President Tayyip Er…      1
      7300    CHICAGO (Reuters) - U.S. President-elect Donal…      1
```

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

```
[19]: data.columns
```

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

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

```
[20]:                                                 text  class
      0  The Republican National Committee released a b…      0
      1  KIEV/NEW YORK (Reuters) - Ukraine s Petro Poro…      1
      2  On Friday, a jury did something that will cert…      0
      3  ANKARA (Reuters) - Turkish President Tayyip Er…      1
      4  CHICAGO (Reuters) - U.S. President-elect Donal…      1
```

```
[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
```

```
<>:3: SyntaxWarning: invalid escape sequence '\['
<>:5: SyntaxWarning: invalid escape sequence '\S'
<>:9: SyntaxWarning: invalid escape sequence '\w'
<>:3: SyntaxWarning: invalid escape sequence '\['
<>:5: SyntaxWarning: invalid escape sequence '\S'
<>:9: SyntaxWarning: invalid escape sequence '\w'
C:\Users\DELL\AppData\Local\Temp\ipykernel_9044\1031775046.py:3: SyntaxWarning:
invalid escape sequence '\['
  text = re.sub('\[.*?\]', '', text)
C:\Users\DELL\AppData\Local\Temp\ipykernel_9044\1031775046.py:5: SyntaxWarning:
invalid escape sequence '\S'
  text = re.sub('https?://\S+|www\.\S+', '', text)
C:\Users\DELL\AppData\Local\Temp\ipykernel_9044\1031775046.py:9: SyntaxWarning:
invalid escape sequence '\w'
  text = re.sub('\w*\d\w*', '', text)
```

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

[23]:
```python
x = data['text']
y = data['class']
```

1.7 Feature Extraction

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

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

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

1.8 Models

1.9 Logistic Regression

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

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

[26]: LogisticRegression()

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

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

[28]: 0.9841354723707665

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

```
              precision    recall  f1-score   support

           0       0.99      0.98      0.98      5924
           1       0.98      0.99      0.98      5296

    accuracy                           0.98     11220
   macro avg       0.98      0.98      0.98     11220
weighted avg       0.98      0.98      0.98     11220
```

1.10 Decision Tree Classifier

```
[30]: from sklearn.tree import DecisionTreeClassifier

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

```
[30]: DecisionTreeClassifier()
```

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

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

```
[32]: 0.9949197860962566
```

```
[33]: print(classification_report(y_test, pred_dt))
```

```
              precision    recall  f1-score   support

           0       0.99      1.00      1.00      5924
           1       1.00      0.99      0.99      5296

    accuracy                           0.99     11220
   macro avg       1.00      0.99      0.99     11220
weighted avg       0.99      0.99      0.99     11220
```

1.11 Gradient Boosting Classifier

```
[34]: from sklearn.ensemble import GradientBoostingClassifier

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

```
[34]: GradientBoostingClassifier(random_state=0)
```

```
[35]: pred_gb = GB.predict(xv_test)
```

```
[36]: GB.score(xv_test, y_test)
```

```
[36]: 0.9959001782531194
```

```
[37]: print(classification_report(y_test, pred_gb))
```

```
              precision    recall  f1-score   support

           0       1.00      0.99      1.00      5924
           1       0.99      1.00      1.00      5296

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

1.12 Random Forest Classifier

```
[38]: from sklearn.ensemble import RandomForestClassifier

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

```
[38]: RandomForestClassifier(random_state=0)
```

```
[39]: pred_rf = RF.predict(xv_test)
```

```
[40]: RF.score(xv_test, y_test)
```

```
[40]: 0.9899286987522281
```

```
[41]: print(classification_report(y_test, pred_rf))
```

```
              precision    recall  f1-score   support

           0       0.99      0.99      0.99      5924
           1       0.99      0.99      0.99      5296

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

1.13 Models Testing

```
[89]:  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_GB = GB.predict(new_xv_test)
           pred_RF = RF.predict(new_xv_test)

           return print("\n\nLR Prediction: {} \nDT Prediction: {} \nGB Prediction: {}␣
       ↪\nRF Prediction: {}".format(output_label(pred_LR[0]),

       ␣
       ↪                            output_label(pred_GB[0]),

       ␣
       ↪                            output_label(pred_DT[0]),

       ␣
       ↪                            output_label(pred_RF[0])))
```

```
[97]:  news = str(input())
       manual_testing(news)
```

 The Pentagon is considering a Boeing proposal to supply Ukraine with cheap,
small precision bombs fitted on to abundantly available rockets, allowing Kyiv
to strike far behind Russian lines, according to a Reuters report. US and
allied military inventories are shrinking, and Ukraine faces an increasing need
for more sophisticated weapons as the war drags on. Boeing's proposed system,
dubbed Ground-Launched Small Diameter Bomb (GLSDB), is one of about a half-dozen
plans for getting new munitions into production for Ukraine and America's
eastern European allies, industry sources told the news agency. GLSDB could be
delivered as early as spring 2023, according to a document reviewed by Reuters
and three people familiar with the plan. It combines the GBU-39 Small Diameter
Bomb (SDB) with the M26 rocket motor, both of which are common in US
inventories. Although a handful of GLSDB units have already been made, there
are many logistical obstacles to formal procurement. The Boeing plan requires a
price discovery waiver, exempting the contractor from an in-depth review that
ensures the Pentagon is getting the best deal possible. Any arrangement would
also require at least six suppliers to expedite shipments of their parts and
services to produce the weapon quickly. Although the US has rebuffed requests

for the 185-mile (297km) range Atacms missile, the GLSDB's 94-mile (150km) range would allow Ukraine to hit valuable military targets that have been out of reach and help it continue pressing its counterattacks by disrupting Russian rear areas. GLSDB is made jointly by Saab AB and Boeing Co and has been in development since 2019, well before the invasion, which Russia calls a "special operation". In October, SAAB chief executive Micael Johansson said of the GLSDB: "We are imminently shortly expecting contracts on that." According to the document - a Boeing proposal to US European - has small, folding wings that allow it to glide more than 100km if dropped from an aircraft and targets as small as 3ft in diameter.

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

```
[95]: news = str(input())
      manual_testing(news)
```

 Many have downplayed allegations of Russian war crimes or even claimed that the war is a supposed "hoax". In one widely shared video, a news reporter could be seen standing in front of lines of body bags, one of which was moving. However, the footage did not show invented war casualties in Ukraine, but a "Fridays for Future" climate change protest in Vienna in February, three weeks before the invasion began. Days later, another viral video of a mannequin claimed to be proof that Ukrainian authorities had "staged" the mass killing of civilians in the town of Bucha. The misleading clip showed a prosthetic doll being dressed and prepared by two men. Nadezhda, an assistant director for a Russian television programme, confirmed to Euronews that the video showed their film set near St. Petersburg and not Ukrainian military personnel. "The information being given [to Russian citizens] is one-sided, it has nothing to do with reality, it is as badly done as any fake," she told Euronews. Euronews The mannequin was being prepared for a television scene in Vsevolozhsk in Russia's Leningrad region.Euronews Other examples of Ukraine war misinformation have centred on "crisis actors" - people who are supposedly hired to act out the role of terrified or deceased war victims. One false claim suggested that a well-known beauty blogger had "acted" as the pregnant victim of a deadly attack on a maternity hospital in the city of Mariupol on 9 March. Russia has shifted its stance on the attack, accusing Ukrainian Azov nationalists of staging a "hoax" bombing at a "non-operational" hospital. Its unfounded claims were later removed by both Facebook and Twitter.

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

RFC Prediction: Fake News