

FAKE-NEWS DETECTION USING PYTHON

Phase 4: Development Part 2

Name:A.Amishek Daniel

Reg.No:961221205004

1.Fake news Detection:



1.0.1 Importing required library

Here I am going to importing some of the required library, if extra library is required to install It will be install later on.

```
[70]: 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
from sklearn.feature_extraction.text import TfidfVectorizer
import re
import string
```

1.0.2 Inserting fake and real dataset

```
[43]: fake    pd.read_csv( D:\IBM\Fake.csv')
      true    pd.read_csv( D:\IBM\True.csv')
```

```
C:\Users\Admin\AppData\Local\Temp\ipykernel_11724\129791993.py:1: DtypeWarning:
Columns
(4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27,28
,29
,30,31,32,33,34,35,36,37,38,39,40,41,42,43,44,45,46,47,48,49,50,51,52,
53,54,55,5
6,57,58,59,60,61,62,63,64,65,66,67,68,69,70,71,72,73,74,75,76,77,78,79
,80,81,82,
83,84,85,86,87,88,89,90,91,92,93,94,95,96,97,98,99,100,101,102,103,104
,105,106,1
07,108,109,110,111,112,113,114,115,116,117,118,119,120,121,122,123,124
,125,126,1
27,128,129,130,131,132,133,134,135,136,137,138,139,140,141,142,143,144
,145,146,1
47,148,149,150,151,152,153,154,155,156,157,158,159,160,161,162,163,16
4 ,165,166,1 67,168,169,170,171) have mixed types. Specify dtype
option on import or set low_memory=False. fake =
pd.read_csv('D:\IBM\Fake.csv')
```

```
[6]: fake.head(5)
```

```
[6]:                                     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 News all
Americans ...

1 House Intelligence Committee News
Chairman Devin Nu...

2 On Friday, it was revealed that News former
Milwauk...

3 On Christmas day, Donald Trump News announced
that ...

4 Pope Francis used his annual News
Christmas Day mes... date Unnamed: 4 Unnamed: 5
Unnamed: 6 Unnamed: 7
Unnamed: 8 \

0 December 31, 2017 NaN NaN NaNNaN
NaN

1 December 31, 2017 NaN NaN NaNNaN
NaN

2 December 30, 2017 NaN NaN NaNNaN
NaN

3 December 29, 2017 NaN NaN NaNNaN
NaN

4 December 25, 2017 NaN NaN NaNNaN
NaN

Unnamed: 9 ... Unnamed: 162 Unnamed: 163
Unnamed: 164 Unnamed: 165 \

0 NaN ... NaN NaN NaN NaN

1 NaN ... NaN NaN NaN NaN

2 NaN ... NaN NaN NaN NaN

3 NaN ... NaN NaN NaN NaN

4 NaN ... NaN NaN NaN NaN

Unnamed: 166 Unnamed: 167 Unnamed: 168
Unnamed: 169 Unnamed: 170 \

0 NaN NaN NaN NaN NaN

1 NaN NaN NaN NaN NaN

2 NaN NaN NaN NaN NaN

3 NaN NaN NaN NaN NaN

4 NaN NaN NaN NaN NaN
Unnamed: 171

0 NaN

1 NaN

2 NaN

```

3         NaN
4         NaN
[5 rows x 172 columns]
[8]: true.head(5)
[8]:                                     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

```

Inserting a column called “class” for fake and real news dataset to categories fake and true news.

```
[11]: fake["class"]    0
      true["class"]   1
```

Removing last 10 rows from both the dataset, for manual testing

```
[12]: fake.shape, true.shape
```

```
[12]: ((23502, 173), (21417, 5))
```

```
[13]: fake_manual_testing  fake.tail(10)
      for i in range(23480,23470,-1):
          fake.drop([i], axis = 0, inplace = True)
      true_manual_testing  true.tail(10)
      for i in range(21416,21406,-1):
          true.drop([i], axis = 0, inplace = True)
```

```
[14]: fake.shape, true.shape
```

```
[14]: ((23492, 173), (21407, 5))
```

Merging the manual testing dataframe in single dataset and save it in a csv file

```
[44]: fake_manual_testing["class"]    0
      true_manual_testing["class"]   1
```

```
[16]: fake_manual_testing.head(10)
```

```
[16]:
```

	title \
23492	Seven Iranians freed in the prisoner swap have...
23493	#Hashtag Hell & The Fake Left
23494	Astroturfing: Journalist Reveals Brainwashing ...
23495	The New American Century: An Era of Fraud
23496	Hillary Clinton: 'Israel First' (and no peace ...
23497	McPain: John McCain Furious That Iran Treated ...
23498	JUSTICE? Yahoo Settles E-mail Privacy Class-ac...
23499	Sunnistan: US and Allied 'Safe Zone' Plan to T...
23500	How to Blow \$700 Million: Al Jazeera America F...
23501	10 U.S. Navy Sailors Held by Iranian Military ...

```

text      subject \
23492 21st Century Wire says This week, the historic... Middle-east 23493
By Dady Chery and Gilbert MercierAll writers ... Middle-east
23494 Vic Bishop Waking TimesOur reality is carefull... Middle-east
23495 Paul Craig RobertsIn the last years of the 20t... Middle-east
23496 Robert Fantina CounterpunchAlthough the United... Middle-east
23497 21st Century Wire says As 21WIRE reported earl... Middle-east
23498 21st Century Wire says It s a familiar theme. ... Middle-east
23499 Patrick Henningsen 21st Century WireRemember ... Middle-east
23500 21st Century Wire says Al Jazeera America will... Middle-east
23501 21st Century Wire says As 21WIRE predicted in ... Middle-east date
      Unnamed: 4 Unnamed: 5 Unnamed: 6 Unnamed: 7 \
23492 January 20, 2016      NaN      NaN      NaN      NaN
23493 January 19, 2016      NaN      NaN      NaN      NaN
23494 January 19, 2016      NaN      NaN      NaN      NaN
23495 January 19, 2016      NaN      NaN      NaN      NaN
23496 January 18, 2016      NaN      NaN      NaN      NaN
23497 January 16, 2016      NaN      NaN      NaN      NaN
23498 January 16, 2016      NaN      NaN      NaN      NaN
23499 January 15, 2016      NaN      NaN      NaN      NaN
23500 January 14, 2016      NaN      NaN      NaN      NaN
23501 January 12, 2016      NaN      NaN      NaN      NaN
Unnamed: 8 Unnamed: 9 ... Unnamed: 163 Unnamed: 164 Unnamed: 165 \
23492      NaN      NaN ...      NaN      NaN      NaN
23493      NaN      NaN ...      NaN      NaN      NaN
23494      NaN      NaN ...      NaN      NaN      NaN
23495      NaN      NaN ...      NaN      NaN      NaN
23496      NaN      NaN ...      NaN      NaN      NaN
23497      NaN      NaN ...      NaN      NaN      NaN
23498      NaN      NaN ...      NaN      NaN      NaN
23499      NaN      NaN ...      NaN      NaN      NaN
23500      NaN      NaN ...      NaN      NaN      NaN
23501      NaN      NaN ...      NaN      NaN      NaN
Unnamed: 166 Unnamed: 167 Unnamed: 168 Unnamed: 169 Unnamed: 170 \
23492      NaN      NaN      NaN      NaN      NaN
23493      NaN      NaN      NaN      NaN      NaN
23494      NaN      NaN      NaN      NaN      NaN
23495      NaN      NaN      NaN      NaN      NaN
23496      NaN      NaN      NaN      NaN      NaN
23497      NaN      NaN      NaN      NaN      NaN
23498      NaN      NaN      NaN      NaN      NaN
23499      NaN      NaN      NaN      NaN      NaN

```

23500	NaN	NaN	NaN	NaN	NaN
23501	NaN	NaN	NaN	NaN	NaN

```

      Unnamed: 171 class
23492      NaN 0
23493      NaN 0
23494      NaN 0
23495      NaN 0
23496      NaN 0
23497      NaN 0
23498      NaN 0
23499      NaN 0
23500      NaN 0
23501      NaN 0
[10 rows x 173 columns]

```

```
[11]: df_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

```

```

pd.concat([df_fake_manual_testing,df_true_manual_testing],
axis = 0)

```

```

[12]:
df_manual_testing =
df_manual_testing.to_csv("manual_testing.csv")

```

Merging the main fake and true dataframe

```

[17]: marge pd.concat([fake, true], axis =0 )
marge.head(10)

```

```

[17]:
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 News all
Americans ...
1 House Intelligence Committee News
Chairman Devin Nu...
2 On Friday, it was revealed that News former
Milwauk...
3 On Christmas day, Donald Trump News
announced that ...
4 Pope Francis used his annual News
Christmas Day mes...

```


5 The number of cases of cops News
brutalizing and ki...

6 Donald Trump spent a good News
portion of his day a...

7 In the wake of yet another court News
decision that...

8 Many people have raised the News
alarm regarding th...

9 Just when you might have thought News we d
get a br...

date Unnamed: 4 Unnamed: 5 Unnamed: 6 Unnamed: 7
Unnamed: 8 \

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

Unnamed: 9 ... Unnamed: 163 Unnamed: 164
Unnamed: 165 Unnamed: 166 \

0	NaN ... NaN NaN NaN NaN
1	NaN ... NaN NaN NaN NaN
2	NaN ... NaN NaN NaN NaN
3	NaN ... NaN NaN NaN NaN
4	NaN ... NaN NaN NaN NaN
5	NaN ... NaN NaN NaN NaN
6	NaN ... NaN NaN NaN NaN
7	NaN ... NaN NaN NaN NaN
8	NaN ... NaN NaN NaN NaN
9	NaN ... NaN NaN NaN NaN

```

    Unnamed: 167 Unnamed: 168 Unnamed: 169 Unnamed: 170 Unnamed: 171
    class
0      NaN      NaN      NaN      NaN      NaN 0
1      NaN      NaN      NaN      NaN      NaN 0
2      NaN      NaN      NaN      NaN      NaN 0
3      NaN      NaN      NaN      NaN      NaN 0
4      NaN      NaN      NaN      NaN      NaN 0
5      NaN      NaN      NaN      NaN      NaN 0
6      NaN      NaN      NaN      NaN      NaN 0
7      NaN      NaN      NaN      NaN      NaN 0
8      NaN      NaN      NaN      NaN      NaN 0
9      NaN      NaN      NaN      NaN      NaN 0
    10 rows x 173 columns]

```

```
[18]: marge.columns
```

```

[18]: Index(['title', 'text', 'subject', 'date', 'Unnamed: 4', 'Unnamed:
5', 'Unnamed: 6', 'Unnamed: 7', 'Unnamed: 8', 'Unnamed: 9',
...
'Unnamed: 163', 'Unnamed: 164', 'Unnamed: 165', 'Unnamed: 166',
'Unnamed: 167', 'Unnamed: 168', 'Unnamed: 169', 'Unnamed: 170',
'Unnamed: 171', 'class'], dtype='object',
length=173)

```

“title”, “subject” and “date” columns is not required for detecting the fake news, so I am going to drop the columns. [19]:

```
[20]: df.isnull().sum()
```

```

                                df = marge.drop(["title", "subject", "date"], axis = 1)
[20]: text      0
    Unnamed: 4    44878
    Unnamed: 5    44879
    Unnamed: 6    44879
    Unnamed: 7    44879
    ...
    Unnamed: 168  44897
    Unnamed: 169  44897
    Unnamed: 170  44897
    Unnamed: 171  44897 class
    0
    Length: 170, dtype: int64

```

Randomly shuffling the dataframe

```
[21]:
```

```
[22]: df.head()
df = df.sample(frac = 1)
```

```
[22]: text Unnamed: 4 \
16097 BUENOS AIRES (Reuters) - Argentina s President... NaN
2842I woke up this morning completely unfazed over... NaN
5912BERLIN (Reuters) - German Chancellor Angela Me... NaN
21886If you have a system that rewards you for bei... NaN
5629WASHINGTON (Reuters) - U.S. Republicans on Fri... NaN
    Unnamed: 5 Unnamed: 6 Unnamed: 7 Unnamed: 8 Unnamed: 9 Unnamed:
    10 \
16097 NaN NaN NaN NaN NaN NaN 2842 NaN NaN NaN NaN NaN NaN
5912 NaN NaN NaN NaN NaN NaN 21886 NaN NaN NaN NaN NaN NaN
5629      NaN      NaN      NaN      NaN      NaN      NaN
    Unnamed: 11 Unnamed: 12 ... Unnamed: 163 Unnamed: 164 Unnamed: 165
    \
16097      NaN      NaN ...      NaN      NaN      NaN
2842      NaN      NaN ...      NaN      NaN      NaN
5912      NaN      NaN ...      NaN      NaN      NaN
21886      NaN      NaN ...      NaN      NaN      NaN
5629      NaN      NaN ...      NaN      NaN      NaN
    Unnamed: 166 Unnamed: 167 Unnamed: 168 Unnamed: 169 Unnamed: 170
    \
16097      NaN      NaN      NaN      NaN      NaN      NaN
2842      NaN      NaN      NaN      NaN      NaN      NaN
5912      NaN      NaN      NaN      NaN      NaN      NaN
21886      NaN      NaN      NaN      NaN      NaN      NaN
5629      NaN      NaN      NaN      NaN      NaN      NaN
    Unnamed: 171 class
16097      NaN      1
2842      NaN      0
5912      NaN      1
21886      NaN      0
```

```
5629          NaN      1
```

```
[5 rows x 170 columns]
```

```
[23]: df.reset_index(inplace=True)
df.drop(["index"], axis = 1, inplace = True)
```

```
[24]: df.columns
```

```
[24]: Index(['text', 'Unnamed: 4', 'Unnamed: 5', 'Unnamed: 6', 'Unnamed:
7',
          'Unnamed: 8', 'Unnamed: 9', 'Unnamed: 10', 'Unnamed: 11',
          'Unnamed: 12', ...,
          'Unnamed: 163', 'Unnamed: 164', 'Unnamed: 165', 'Unnamed: 166',
          'Unnamed: 167', 'Unnamed: 168', 'Unnamed: 169', 'Unnamed: 170',
          'Unnamed: 171', 'class'], dtype='object',
          length=170)
```

```
[25]: df.head()
```

```
[25]:          text Unnamed: 4 Unnamed: 5 \
0 BUENOS AIRES (Reuters) - Argentina s President... NaN      NaN
1 I woke up this morning completely unfazed over... NaN      NaN
2 BERLIN (Reuters) - German Chancellor Angela Me... NaN      NaN
3 If you have a system that rewards you for bei... NaN      NaN
4 WASHINGTON (Reuters) - U.S. Republicans on Fri... NaN      NaN
Unnamed: 6 Unnamed: 7 Unnamed: 8 Unnamed: 9 Unnamed: 10 Unnamed: 11 \
0      NaN NaN NaN NaN NaN NaN
1      NaN NaN NaN NaN NaN NaN
2      NaN NaN NaN NaN NaN NaN
3      NaN NaN NaN NaN NaN NaN
4      NaN NaN NaN NaN NaN NaN
Unnamed: 12 ... Unnamed: 163 Unnamed: 164 Unnamed: 165 Unnamed: 166 \
0      NaN ...      NaN      NaN      NaN      NaN
1      NaN ...      NaN      NaN      NaN      NaN
2      NaN ...      NaN      NaN      NaN      NaN
3      NaN ...      NaN      NaN      NaN      NaN
4      NaN ...      NaN      NaN      NaN      NaN
Unnamed: 167 Unnamed: 168 Unnamed: 169 Unnamed: 170 Unnamed: 171 class 0
NaN NaN NaN NaN NaN 1
1      NaN NaN NaN NaN NaN 0
2      NaN NaN NaN NaN NaN 1
3      NaN NaN NaN NaN NaN 0
```

```
4          NaN NaN NaN NaN NaN 1
[5 rows x 170 columns]
```

Creating a function to convert the text in lowercase, remove the extra space, special chr., ulr and links.

```
[46]: def word_drop(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
```

```
[47]: df["text"] = df["text"].apply(word_drop)
```

Defining dependent and independent variable as x and y

```
[48]: x = df[ text ]
      y = df[ class ]
```

Splitting the dataset into training set and testing set.

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

Convert text to vectors

```
[71]: from sklearn.feature_extraction.text import TfidfVectorizer
```

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

1.0.3 1. Logistic Regression

```
[73]: from sklearn.linear_model import LogisticRegression
```

```
[78]: LR = LogisticRegression()
      LR.fit(xv_train, y_train)
```

```
C:\Users\Admin\anaconda3\anaconda3\Lib\site-
packages\sklearn\utils\validation.py:605: FutureWarning: is_sparse is
deprecated
and will be removed in a future version. Check `isinstance(dtype,
pd.SparseDtype)` instead.
    if is_sparse(pd_dtype):
C:\Users\Admin\anaconda3\anaconda3\Lib\site-
packages\sklearn\utils\validation.py:614: FutureWarning: is_sparse is
deprecated
and will be removed in a future version. Check
`isinstance(dtype, pd.SparseDtype)` instead. if
is_sparse(pd_dtype) or not is_extension_array_dtype(pd_dtype):
```

```
[78]: LogisticRegression()
```

```
[75]: pred_lr LR predict(xv_test)
```

```
[76]: LR score(xv_test, y_test)
```

```
C:\Users\Admin\anaconda3\anaconda3\Lib\site-
packages\sklearn\utils\validation.py:605: FutureWarning: is_sparse is
deprecated
and will be removed in a future version. Check
`isinstance(dtype, pd.SparseDtype)` instead. if
is_sparse(pd_dtype):
C:\Users\Admin\anaconda3\anaconda3\Lib\site-
packages\sklearn\utils\validation.py:614: FutureWarning: is_sparse is
deprecated
and will be removed in a future version. Check `isinstance(dtype,
pd.SparseDtype)` instead.
    if is_sparse(pd_dtype) or not is_extension_array_dtype(pd_dtype):
C:\Users\Admin\anaconda3\anaconda3\Lib\site-
packages\sklearn\utils\validation.py:605: FutureWarning: is_sparse is
deprecated
and will be removed in a future version. Check `isinstance(dtype,
pd.SparseDtype)` instead.
    if is_sparse(pd_dtype):
C:\Users\Admin\anaconda3\anaconda3\Lib\site-
packages\sklearn\utils\validation.py:614: FutureWarning: is_sparse is
deprecated
and will be removed in a future version. Check `isinstance(dtype,
pd.SparseDtype)` instead. if is_sparse(pd_dtype) or not
is_extension_array_dtype(pd_dtype):
C:\Users\Admin\anaconda3\anaconda3\Lib\site-
packages\sklearn\utils\validation.py:605: FutureWarning: is_sparse
is
deprecated
```

```

and will be removed in a future version. Check `isinstance(dtype,
pd.SparseDtype)` instead.
    if is_sparse(pd_dtype):
C:\Users\Admin\anaconda3\anaconda3\Lib\site-
packages\sklearn\utils\validation.py:614: FutureWarning: is_sparse is
deprecated
and will be removed in a future version. Check
`isinstance(dtype, pd.SparseDtype)` instead. if
is_sparse(pd_dtype) or not is_extension_array_dtype(pd_dtype):

```

```
[76]: 0.985924276169265
```

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

```

precision recall f1-score support
      0      0.99      0.99      0.99      5821
      1      0.98      0.99      0.99      5404

```

```

      accuracy              0.99      11225
    macro avg      0.99      0.99      0.99      11225
weighted avg      0.99      0.99      0.99      11225
avg

```

```

C:\Users\Admin\anaconda3\anaconda3\Lib\site-
packages\sklearn\utils\validation.py:605: FutureWarning: is_sparse is
deprecated
and will be removed in a future version. Check `isinstance(dtype,
pd.SparseDtype)` instead.
    if is_sparse(pd_dtype):
C:\Users\Admin\anaconda3\anaconda3\Lib\site-

```

```

and will be removed in a future version. Check `isinstance(dtype,
pd.SparseDtype)` instead. if is_sparse(pd_dtype) or not
is_extension_array_dtype(pd_dtype):
C:\Users\Admin\anaconda3\anaconda3\Lib\site-
packages\sklearn\utils\validation.py:605: FutureWarning: is_sparse is
deprecated
and will be removed in a future version. Check
`isinstance(dtype, pd.SparseDtype)` instead. if
is_sparse(pd_dtype):
C:\Users\Admin\anaconda3\anaconda3\Lib\site-
packages\sklearn\utils\validation.py:614: FutureWarning: is_sparse is
deprecated
and will be removed in a future version. Check `isinstance(dtype,
pd.SparseDtype)` instead. if is_sparse(pd_dtype):
C:\Users\Admin\anaconda3\anaconda3\Lib\site-
packages\sklearn\utils\validation.py:605: FutureWarning: is_sparse is
deprecated
and will be removed in a future version. Check `isinstance(dtype,
pd.SparseDtype)` instead.
    if is_sparse(pd_dtype):
C:\Users\Admin\anaconda3\anaconda3\Lib\site-
packages\sklearn\utils\validation.py:614: FutureWarning: is_sparse is
deprecated
and will be removed in a future version. Check
`isinstance(dtype, pd.SparseDtype)` instead. if
is_sparse(pd_dtype) or not is_extension_array_dtype(pd_dtype):

```

1.0.4 2. Decision Tree Classification

[79]:

[80]:

```

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

```

from sklearn.tree import DecisionTreeClassifier

```

C:\Users\Admin\anaconda3\anaconda3\Lib\site-
packages\sklearn\utils\validation.py:605: FutureWarning: is_sparse is
deprecated
and will be removed in a future version. Check `isinstance(dtype,
pd.SparseDtype)` instead.
    if is_sparse(pd_dtype):
C:\Users\Admin\anaconda3\anaconda3\Lib\site-
packages\sklearn\utils\validation.py:614: FutureWarning: is_sparse is
deprecated

```



```
packages\sklearn\utils\validation.py:614: FutureWarning: is_sparse is deprecated
```

```
and will be removed in a future version. Check `isinstance(dtype, pd.SparseDtype)` instead.
```

```
if is_sparse(pd_dtype) or not is_extension_array_dtype(pd_dtype):
```

```
[80]: DecisionTreeClassifier()
```

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

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

```
[36]: 0.9945632798573975
```

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

```
precision recall f1-score support
```

0	1.00	1.00	1.00	5821
1	1.00	1.00	1.00	5404

accuracy			1.00	11225
macro avg	1.00	1.00	1.00	11225
weighted avg	1.00	1.00	1.00	11225

```
avg
```

```
C:\Users\Admin\anaconda3\anaconda3\Lib\site-
```

```
packages\sklearn\utils\validation.py:605: FutureWarning: is_sparse is deprecated
```

```
and will be removed in a future version. Check `isinstance(dtype, pd.SparseDtype)` instead.
```

```
if is_sparse(pd_dtype):
```

```
C:\Users\Admin\anaconda3\anaconda3\Lib\site-
```

```
packages\sklearn\utils\validation.py:614: FutureWarning: is_sparse is deprecated
```

```
and will be removed in a future version. Check
```

```
`isinstance(dtype, pd.SparseDtype)` instead. if
```

```
is_sparse(pd_dtype) or not is_extension_array_dtype(pd_dtype):
```

```
C:\Users\Admin\anaconda3\anaconda3\Lib\site-
```

```
packages\sklearn\utils\validation.py:605: FutureWarning: is_sparse is deprecated
```

```
and will be removed in a future version. Check `isinstance(dtype, pd.SparseDtype)` instead.
```

```
if is_sparse(pd_dtype):
```

```
C:\Users\Admin\anaconda3\anaconda3\Lib\site-
```

```
and will be removed in a future version. Check `isinstance(dtype,
```

```
pd.SparseDtype)` instead. if is_sparse(pd_dtype) or not
```

```
is_extension_array_dtype(pd_dtype):
```

```
C:\Users\Admin\anaconda3\anaconda3\Lib\site-
```

```
packages\sklearn\utils\validation.py:605: FutureWarning: is_sparse is
```

```

deprecated
and will be removed in a future version. Check `isinstance(dtype,
pd.SparseDtype)` instead.
    if is_sparse(pd_dtype):
C:\Users\Admin\anaconda3\anaconda3\Lib\site-
packages\sklearn\utils\validation.py:614: FutureWarning: is_sparse is
deprecated
and will be removed in a future version. Check
`isinstance(dtype, pd.SparseDtype)` instead. if
is_sparse(pd_dtype) or not is_extension_array_dtype(pd_dtype):

```

1.0.5 3. Gradient Boosting Classifier

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

```
[39]: GBC = GradientBoostingClassifier(random_state=0)
      GBC.fit(xv_train, y_train)
```

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

```
[40]: pred_gbc = GBC.predict(xv_test)
```

```
[41]: GBC.score(xv_test, y_test)
```

```
[41]: 0.9955436720142602
```

```
[42]: print(classification_report(y_test, pred_gbc))
```

```

precision recall f1-score support

      0      1.00      0.99      1.00      5883
      1      0.99      1.00      1.00      5337

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

```

1.0.6 4. Random Forest Classifier

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

```
[84]: RFC = RandomForestClassifier(random_state=0)
      RFC.fit(xv_train, y_train)
```

```

C:\Users\Admin\anaconda3\anaconda3\Lib\site-
packages\sklearn\utils\validation.py:605: FutureWarning: is_sparse is
deprecated

```

```
packages\sklearn\utils\validation.py:614: FutureWarning: is_sparse is deprecated
```

```
and will be removed in a future version. Check `isinstance(dtype, pd.SparseDtype)` instead.
```

```
    if is_sparse(pd_dtype):
```

```
C:\Users\Admin\anaconda3\anaconda3\Lib\site-
```

```
packages\sklearn\utils\validation.py:614: FutureWarning: is_sparse is deprecated
```

```
and will be removed in a future version. Check
```

```
`isinstance(dtype, pd.SparseDtype)` instead. if
```

```
is_sparse(pd_dtype) or not is_extension_array_dtype(pd_dtype):
```

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

```
[85]: pred_rfc    RFC predict(xv_test)
```

```
[88]: RFC score(xv_test, y_test)
```

```
C:\Users\Admin\anaconda3\anaconda3\Lib\site-
```

```
packages\sklearn\utils\validation.py:605: FutureWarning: is_sparse is deprecated
```

```
and will be removed in a future version. Check
```

```
`isinstance(dtype, pd.SparseDtype)` instead. if
```

```
is_sparse(pd_dtype):
```

```
C:\Users\Admin\anaconda3\anaconda3\Lib\site-
```

```
packages\sklearn\utils\validation.py:614: FutureWarning: is_sparse is deprecated
```

```
and will be removed in a future version. Check `isinstance(dtype, pd.SparseDtype)` instead.
```

```
    if is_sparse(pd_dtype) or not is_extension_array_dtype(pd_dtype):
```

```
C:\Users\Admin\anaconda3\anaconda3\Lib\site-
```

```
packages\sklearn\utils\validation.py:605: FutureWarning: is_sparse is deprecated
```

```
and will be removed in a future version. Check
```

```
`isinstance(dtype, pd.SparseDtype)` instead. if
```

```
is_sparse(pd_dtype):
```

```
C:\Users\Admin\anaconda3\anaconda3\Lib\site-
```

```
and will be removed in a future version. Check `isinstance(dtype, pd.SparseDtype)` instead. if is_sparse(pd_dtype) or not
```

```
is_extension_array_dtype(pd_dtype):
```

```
C:\Users\Admin\anaconda3\anaconda3\Lib\site-
```

```
packages\sklearn\utils\validation.py:605: FutureWarning: is_sparse is deprecated
```

```
and will be removed in a future version. Check `isinstance(dtype, pd.SparseDtype)` instead.
```

```
    if is_sparse(pd_dtype):
```

```
C:\Users\Admin\anaconda3\anaconda3\Lib\site-
```

```
packages\sklearn\utils\validation.py:614: FutureWarning: is_sparse is
```

```

deprecated
and will be removed in a future version. Check
`isinstance(dtype, pd.SparseDtype)` instead. if
is_sparse(pd_dtype) or not is_extension_array_dtype(pd_dtype):

```

```
[88]: 0.9882405345211581
```

```
[89]: print(classification_report(y_test, pred_rfc))
```

```

precision recall f1-score support
      0      0.99      0.99      0.99      5821
      1      0.99      0.98      0.99      5404

```

```

      accuracy              0.99      11225
      macro avg      0.99      0.99      0.99      11225
      weighted      0.99      0.99      0.99      11225
      avg

```

```

C:\Users\Admin\anaconda3\anaconda3\Lib\site-
packages\sklearn\utils\validation.py:605: FutureWarning: is_sparse is
deprecated

```

```

and will be removed in a future version. Check
`isinstance(dtype, pd.SparseDtype)` instead. if
is_sparse(pd_dtype):

```

```

C:\Users\Admin\anaconda3\anaconda3\Lib\site-
packages\sklearn\utils\validation.py:614: FutureWarning: is_sparse is
deprecated

```

```

and will be removed in a future version. Check `isinstance(dtype,
pd.SparseDtype)` instead. if is_sparse(pd_dtype) or not
is_extension_array_dtype(pd_dtype):

```

```

C:\Users\Admin\anaconda3\anaconda3\Lib\site-
packages\sklearn\utils\validation.py:605: FutureWarning: is_sparse is
deprecated

```

```

and will be removed in a future version. Check `isinstance(dtype,
pd.SparseDtype)` instead.
    if is_sparse(pd_dtype):

```

```

C:\Users\Admin\anaconda3\anaconda3\Lib\site-
packages\sklearn\utils\validation.py:614: FutureWarning: is_sparse is
deprecated

```

```

and will be removed in a future version. Check `isinstance(dtype,
pd.SparseDtype)` instead. if is_sparse(pd_dtype) or not
is_extension_array_dtype(pd_dtype):

```

```

C:\Users\Admin\anaconda3\anaconda3\Lib\site-
packages\sklearn\utils\validation.py:605: FutureWarning: is_sparse is
deprecated

```

```

and will be removed in a future version. Check `isinstance(dtype,
pd.SparseDtype)` instead.

```

```
packages\sklearn\utils\validation.py:614: FutureWarning: is_sparse is
deprecated
```

```
    if is_sparse(pd_dtype):
C:\Users\Admin\anaconda3\anaconda3\Lib\site-
packages\sklearn\utils\validation.py:614: FutureWarning: is_sparse is
deprecated
and will be removed in a future version. Check
`isinstance(dtype, pd.SparseDtype)` instead. if
is_sparse(pd_dtype) or not is_extension_array_dtype(pd_dtype):
```

2 Model Testing With Manual Entry

2.0.1 News

```
[8]: import pandas as pd

def output_lable(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(word_drop)
    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)

    return print("\n\nLR Prediction: {} \nDT Prediction: {} \nGBC Prediction: {} \nRFC Prediction: {}".format(output_lable(pred_LR[0]),
    ↪                                     output_lable(pred_DT[0])
    ↪                                     output_lable(pred_GBC[0])
    ↪                                     output_lable(pred_RFC[0])))

[ ]: news = str(input())
manual_testing(news)
model = pickle.load(open("model.pkl", "rb"))
```

```
[ ]:
```