Batch: A3   **Experiment Number:8**

Roll Number: 16010421073   **Name:Keyur Patel**
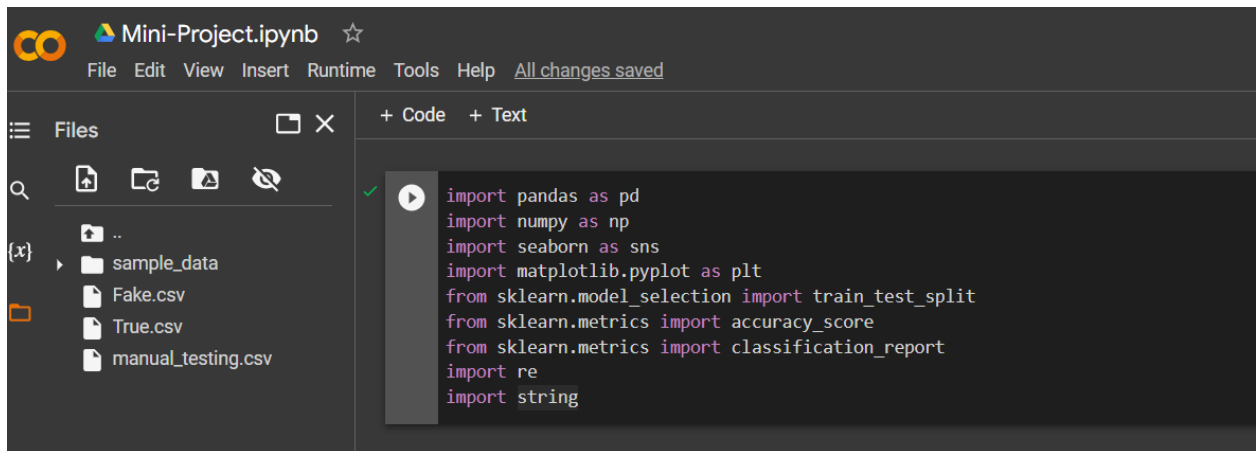
Aim of the Experiment: Mini Project

Output/Result:

Problem Statement:

Fake news and misinformation in today's digital age has become a significant concern. False information, deliberately or inadvertently spread through various online platforms, can have far-reaching consequences, including political, social, and economic impact. Therefore, the development of effective tools to identify and combat fake news is essential. Machine learning models can play a crucial role in automating the process of fake news detection.

Develop a machine learning model that can accurately and efficiently identify fake news articles from genuine ones, based on the content and context of the text.

Dataset Details(US News):

1) True News dataset
2) Fake News dataset
3) Manual Testing dataset(consisting of true and fake news)

CODE :

❖ **Importing Libraries**

❖ **Importing Fake news and True news Dataset**

```
[12] df_fake = pd.read_csv("Fake.csv")
     df_true = pd.read_csv("True.csv")
```

❖

```
df_fake.head()
```

	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

```
[14] df_fake.head(5)
```

	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

❖ **Inserting a column "class" as target feature (0 & 1 representing fake and true news)**

```
[15] df_fake["class"] = 0
     df_true["class"] = 1
```

```
    df_fake.shape, df_true.shape
```

```
    ((23481, 5), (21417, 5))
```

❖ **Removing last 10 rows from fake and true dataset for manual testing**

```
[17] #removing last 10 rows for manual testing
     df_fake_manual_testing = df_fake.tail(10)
     for i in range(23480,23470,-1):
         df_fake.drop([i], axis = 0, inplace = True)


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

```
[18] df_fake.shape, df_true.shape

     ((23471, 5), (21407, 5))
```

❖

```
[19] df_fake_manual_testing["class"] = 0
     df_true_manual_testing["class"] = 1

     <ipython-input-19-3aaf8ec2aad1>:1: SettingWithCopyWarning:
     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
       df_fake_manual_testing["class"] = 0
     <ipython-input-19-3aaf8ec2aad1>:2: SettingWithCopyWarning:
     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
       df_true_manual_testing["class"] = 1
```

Manual testing dataset for fake news..

❖

```
df_fake_manual_testing.head(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

+ Code   + Text

(A Constituent College of Somaiya Vidyavihar University)

❖

Manual testing dataset for True news..

```
df_true_manual_testing.head(10)
```

	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	1
21408	U.S., North Korea clash at U.N. forum over nuc...	GENEVA (Reuters) - North Korea and the United ...	worldnews	August 22, 2017	1
21409	U.S., North Korea clash at U.N. arms forum on ...	GENEVA (Reuters) - North Korea and the United ...	worldnews	August 22, 2017	1
21410	Headless torso could belong to submarine journ...	COPENHAGEN (Reuters) - Danish police said on T...	worldnews	August 22, 2017	1
21411	North Korea shipments to Syria chemical arms a...	UNITED NATIONS (Reuters) - Two North Korean sh...	worldnews	August 21, 2017	1
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

Manual testing dataset for true news

❖ **Concatenating two dataframes manual testing fake and true dataframes and merging true and fake dataframes creating a single manual_testing.csv.**

```
[22] df_manual_testing = pd.concat([df_fake_manual_testing,df_true_manual_testing], axis = 0)
     df_manual_testing.to_csv("manual_testing.csv")
```

Concatenating two dataframes manual testing fake and true to one dataframe which is manual testing and converting it to csv format.

```
[23] df_merge = pd.concat([df_fake, df_true], axis =0 )
     df_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

merging true and fake dataframes.

❖

❖ **Removing columns which are not required. Then shuffling the dataframe manual_testing.csv**

merging true and fake dataframes.

```
[24] df_merge.columns

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

Removing columns which are not required

```
[25] df = df_merge.drop(["title", "subject","date"], axis = 1)
```

```
[26] df.isnull().sum()

        text     0
        class    0
        dtype: int64
```

Random shuffling the dataframe

```
[27] df = df.sample(frac = 1)
```

❖

```
df.head()
```

	text	class
17971	BEIRUT (Reuters) - Islamic State took control ...	1
19309	SAN FRANCISCO (Reuters) - Facebook Inc (FB.O) ...	1
11130	CHICAGO (Reuters) - Former U.S. House Speaker ...	1
10740	Rush Limbaugh nails it in the transcript below...	0
8195	You may or may not know this, but the Democrat...	0

to reset the index of a DataFrame to the default integer index and remove any additional index column that was created during the reset process.

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

```
[30] df.columns

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

❖
❖
❖
❖

❖ **Creating a function to process the texts.(regular expressions)**

```
df.head()
```

	text	class
0	BEIRUT (Reuters) - Islamic State took control ...	1
1	SAN FRANCISCO (Reuters) - Facebook Inc (FB.O) ...	1
2	CHICAGO (Reuters) - Former U.S. House Speaker ...	1
3	Rush Limbaugh nails it in the transcript below...	0
4	You may or may not know this, but the Democrat...	0

Creating a function to process the texts

```
[32] 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
```

```
[33] df["text"] = df["text"].apply(wordopt)
```

❖ **Defining dependent and independent variables .Then splitting training and testing of datasets 25%(testing dataset)**

Defining dependent and independent variables.

```
x = df["text"]
y = df["class"]
```

+ Code   + Text

Splitting Training and Testing of datasets

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

Convert text to vectors

```
[36] from sklearn.feature_extraction.text import TfidfVectorizer

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

❖

Implementing machine learning algorithms

❖

❖ **Logistic regression**

```
Logistic Regression

[37]  from sklearn.linear_model import LogisticRegression

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

      ▾ LogisticRegression
      LogisticRegression()


[38]  pred_lr=LR.predict(xv_test)


[39]  LR.score(xv_test, y_test)

      0.9877896613190731


[40]  print(classification_report(y_test, pred_lr))

                    precision    recall  f1-score   support

                0       0.99      0.99      0.99      5871
                1       0.99      0.99      0.99      5349

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

❖ **Decision Tree classification**

```
Decision tree classification
                                              + Code    + Text

[41]  from sklearn.tree import DecisionTreeClassifier

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

      ▾ DecisionTreeClassifier
      DecisionTreeClassifier()


[42]  pred_dt = DT.predict(xv_test)


[43]  DT.score(xv_test, y_test)

      0.995632798573975


[44]  print(classification_report(y_test, pred_dt))

                    precision    recall  f1-score   support

                0       1.00      1.00      1.00      5871
                1       1.00      0.99      1.00      5349

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

❖ **Gradient booster classfier**

```
Gradient Boosting Classifier

[▶] from sklearn.ensemble import GradientBoostingClassifier
5m
    GBC = GradientBoostingClassifier(random_state=0)
    GBC.fit(xv_train, y_train)

[➡]  ▼        GradientBoostingClassifier
    GradientBoostingClassifier(random_state=0)


[46] pred_gbc = GBC.predict(xv_test)
0s

[47] GBC.score(xv_test, y_test)
0s
    0.9958110516934047

[48] print(classification_report(y_test, pred_gbc))
0s
                  precision    recall  f1-score   support

              0       1.00      0.99      1.00      5871
              1       0.99      1.00      1.00      5349

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

❖ **Random forest Classifier**

```
Random Forest Classifier

[49] from sklearn.ensemble import RandomForestClassifier
1m
    RFC = RandomForestClassifier(random_state=0)
    RFC.fit(xv_train, y_train)

[➡]  ▼        RandomForestClassifier
    RandomForestClassifier(random_state=0)


[50] pred_rfc = RFC.predict(xv_test)
2s

[51] RFC.score(xv_test, y_test)
0s
    0.9906417112299465

[52] print(classification_report(y_test, pred_rfc))
0s
                  precision    recall  f1-score   support

              0       0.99      0.99      0.99      5871
              1       0.99      0.99      0.99      5349

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

Model Testing

```
Model Testing

[54] 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(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)

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

<u>Results</u>

```
[55] news = str(input())
     manual_testing(news)

     BRUSSELS (Reuters) - NATO allies on Tuesday welcomed President Donald Trump s decision to commit more forces to Afghanistan, as part of a new U.S. strate

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

```
news = str(input())
manual_testing(news)

SAO PAULO (Reuters) - Cesar Mata Pires, the owner and co-founder of Brazilian engineering conglomerate OAS SA, one of the largest companies involved in B

LR Prediction: Not A Fake News
DT Prediction: Not A Fake News
GBC Prediction: Not A Fake News
RFC Prediction: Not A Fake News
```
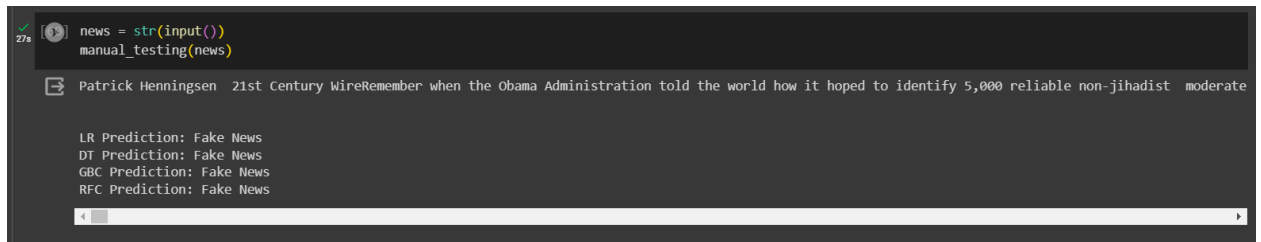
```
news = str(input())
manual_testing(news)

Vic Bishop Waking TimesOur reality is carefully constructed by powerful corporate, political and special interest sources in order to covertly sway publi

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

(A Constituent College of Somaiya Vidyavihar University)

```
news = str(input())
manual_testing(news)
```

```
Patrick Henningsen  21st Century WireRemember when the Obama Administration told the world how it hoped to identify 5,000 reliable non-jihadist  moderate

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

Outcomes: CO3: Comprehend radial-basis-function (RBF) networks and Kernel learning method

---

Conclusion (based on the Results and outcomes achieved):

Thus we successfully did a mini project for fake news detection consisting of
True and fake datasets and merged it into manual_testing dataframe.

References:

Books,Journals etc.