

▼ Fake News Detection

Importing necessary library

```
import numpy as np
import pandas as pd
from sklearn.model_selection import train_test_split as tts
from sklearn.feature_extraction.text import TfidfVectorizer
from sklearn.linear_model import PassiveAggressiveClassifier
from sklearn.metrics import accuracy_score, confusion_matrix
```

Reading data from data file

```
data=pd.read_csv('/content/drive/MyDrive/news/news.csv')
print(data.shape)
data.head()
```

(6335, 4)

	Unnamed: 0		title	text	label
0	8476	You Can Smell Hillary's Fear	Daniel Greenfield, a Shillman Journalism Fello...		FAKE
1	10294	Watch The Exact Moment Paul Ryan Committed Pol...	Google Pinterest Digg LinkedIn Reddit Stumbleu...		FAKE
2	3608	Kerry to go to Paris in gesture of sympathy	U.S. Secretary of State John F. Kerry said Mon...		REAL

```
data['text'][2]
```

'U.S. Secretary of State John F. Kerry said Monday that he will stop in Paris later this week, amid criticism that no top American officials attended Sunday's unity march against terrorism.\n\nKerry said he expects to arrive in Paris Thursday evening, as he heads home after a week abroad. He said he will fly to France at the conclusion of a series of meetings scheduled for Thursday in Sofia, Bulgaria. He plans to meet the next day with Foreign Minister Laurent Fabius and President Francois Hollande, then return to Washington.\n\nThe visit by Kerry, who has family and childhood ties to the country and speaks fluent French, could address some of the criticism

Splitting data into input and output

```
def mapLabel(x):
    if x=='REAL':
        return 1
    return 0
```

```
X=data['text']
```

```

y=data['label']

#mapping real to 1 and fake to 0
y=y.map(mapLabel)
print((X.shape,y.shape))
y

((6335,), (6335,))
0      0
1      0
2      1
3      0
4      1
..
6330    1
6331    0
6332    0
6333    1
6334    1
Name: label, Length: 6335, dtype: int64

```

Splitting data to train and test dataset

```

X_train,X_test,y_train,y_test=tts(X,y,test_size=0.2,random_state=7)

X_train.shape,X_test.shape,y_train.shape,y_test.shape

((5068,), (1267,), (5068,), (1267,))

```

Vectorizing Text to Numericals

```

class TextVector:
    def __init__(self):
        self.vect=TfidfVectorizer(stop_words='english',max_df=0.7)
    def findTerms(self,x):
        self.vect.fit(x)
    def toVector(self,x):
        return self.vect.transform(x)

#vectorizer --> that converts text to numericals
vectorizer=TextVector()
vectorizer.findTerms(X_train)
X_train_tfidf=vectorizer.toVector(X_train)

```

Building a model

```

pac=PassiveAggressiveClassifier(max_iter=50)
pac.fit(X_train_tfidf,y_train)

```

Prediction on test set

Analyzing the output

Accuracy score is 92.5808997632202

- ▼ Interface for users

Enter news textU.S. Secretary of State John F. Kerry said Monday that he will stop i
prediction is >>>>>>>>>>>>
Real

✓ 5s completed at 4:38 PM ● ✕