

```
import numpy as np
import pandas as pd
import re
from nltk.corpus import stopwords
from nltk.stem.porter import PorterStemmer
from sklearn.feature_extraction.text import TfidfVectorizer
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LogisticRegression
from sklearn.metrics import accuracy_score
```

```
import nltk
nltk.download('stopwords')
```

```
[nltk_data] Downloading package stopwords to /root/nltk_data...
[nltk_data]   Unzipping corpora/stopwords.zip.
True
```

```
from google.colab import files
```

```
uploaded=files.upload()
```

 train.csv.zip

- **train.csv.zip**(application/x-zip-compressed) - 38841253 bytes, last modified: 7/15/2021 - 100% done
Saving train.csv.zip to train.csv.zip

```
for fn in uploaded.keys():
    print('User uploaded file"{name}" with length {length} bytes'.format (name=fn,length=len(up
    User uploaded file"train.csv.zip" with length 38841253 bytes
```

```
uploaded
```

```
import pandas as pd
```

```
news_dataset = pd.read_csv('train.csv.zip')
```

Double-click (or enter) to edit

```
news_dataset.isnull().sum()
```

```
id          0
title      558
author     1957
text       39
```

```
label      0
dtype: int64
```

Double-click (or enter) to edit

```
news_dataset=news_dataset.fillna('')
news_dataset.head(20)
```

	id	title	author	
0	0	House Dem Aide: We Didn't Even See Comey's Let...	Darrell Lucas	House Dem Aide: We Didn't Even See Comey's Let...
1	1	FLYNN: Hillary Clinton, Big Woman on Campus - ...	Daniel J. Flynn	Ever get the feeling your li
2	2	Why the Truth Might Get You Fired	Consortiumnews.com	Why the Truth Might Get You Fired Octo
3	3	15 Civilians Killed In Single US Airstrike Hav...	Jessica Purkiss	Videos 15 Civilians Killed In S
4	4	Iranian woman jailed for fictional unpublished...	Howard Portnoy	Print \nAn Iranian woman I sent
5	5	Jackie Mason: Hollywood Would Love Trump if He...	Daniel Nussbaum	In these trying times, Jackie
6	6	Life: Life Of Luxury: Elton John's 6 Favorite ...		Ever wonder how Britain's me
7	7	Benoît Hamon Wins French Socialist Party's Pre...	Alissa J. Rubin	PARIS — France chose an i
8	8	Excerpts From a Draft Script for Donald Trump'...		Donald J. Trump is schu make a
9	9	A Back-Channel Plan for Ukraine and Russia, Co...	Megan Twohey and Scott Shane	A week before Michael resigned
10	10	Obama's Organizing for Action Partners with So...	Aaron Klein	Organizing for Action, th gro
11	11	BBC Comedy Sketch "Real Housewives of ISIS" Ca...	Chris Tomlinson	The BBC produced spoof on t Hous
12	12	Russian Researchers Discover Secret Nazi Milit...	Amando Flavio	The mystery surrounding 1 Reich :

Double-click (or enter) to edit

```
news_dataset['content']=news_dataset['author']+' '+news_dataset['title']
news_dataset.head(20)
```

	id	title	author	text	label	
0	0	House Dem Aide: We Didn't Even See Comey's Let...	Darrell Lucas	House Dem Aide: We Didn't Even See Comey's Let...	1	Da Der
1	1	FLYNN: Hillary Clinton, Big Woman on Campus - ...	Daniel J. Flynn	Ever get the feeling your life circles the rou...	0	FLYN
2	2	Why the Truth Might Get You Fired	Consortiumnews.com	Why the Truth Might Get You Fired October 29, ...	1	Con Wh
3	3	15 Civilians Killed In Single US Airstrike Hav...	Jessica Purkiss	Videos 15 Civilians Killed In Single US Aistr...	1	J
4	4	Iranian woman jailed for fictional unpublished...	Howard Portnoy	Print \nAn Iranian woman has been sentenced to...	1	Irar
5	5	Jackie Mason: Hollywood Would Love Trump if He...	Daniel Nussbaum	In these trying times, Jackie Mason is the Voi...	0	I Hc
6	6	Life: Life Of Luxury: Elton John's 6 Favorite ...		Ever wonder how Britain's most iconic pop pian...	1	Lif
7	7	Benoît Hamon Wins French Socialist Party's Pre...	Alissa J. Rubin	PARIS — France chose an idealistic, traditi...	0	Aliss Ha
8	8	Excerpts From a Draft Script for Donald Trump'...		Donald J. Trump is scheduled to make a highly ...	0	Exce
						M.

seperating the data and label

```
x=news_dataset.drop(columns='label',axis=1)
y=news_dataset['label']
print(x)
print(y)
```

```

id    ...    content
0      0    ...    Darrell Lucas House Dem Aide: We Didn't Even S...
1      1    ...    Daniel J. Flynn FLYNN: Hillary Clinton, Big Wo...
2      2    ...    Consortiumnews.com Why the Truth Might Get You...
3      3    ...    Jessica Purkiss 15 Civilians Killed In Single ...
4      4    ...    Howard Portnoy Iranian woman jailed for fictio...
...    ...    ...
20795  20795 ...    Jerome Hudson Rapper T.I.: Trump a 'Poster Chi...
20796  20796 ...    Benjamin Hoffman N.F.L. Playoffs: Schedule, Ma...
20797  20797 ...    Michael J. de la Merced and Rachel Abrams Macy...
```

```

20798 20798 ... Alex Ansary NATO, Russia To Hold Parallel Exer...
20799 20799 ... David Swanson What Keeps the F-35 Alive

```

```
[20800 rows x 5 columns]
```

```
0      1
```

```
1      0
```

```
2      1
```

```
3      1
```

```
4      1
```

```
..
```

```
20795  0
```

```
20796  0
```

```
20797  0
```

```
20798  1
```

```
20799  1
```

```
Name: label, Length: 20800, dtype: int64
```

stemming

```
port_stem=PorterStemmer()
```

```
def stemming(content):
```

```
    stemmed_content = re.sub('[^a-zA-Z]', ' ', content)
```

```
    stemmed_content = stemmed_content.lower()
```

```
    stemmed_content = stemmed_content.split()
```

```
    stemmed_content = [port_stem.stem(word) for word in stemmed_content if not word in stopwords]
```

```
    stemmed_content = ' '.join(stemmed_content)
```

```
    return stemmed_content
```

```
news_dataset['content'] = news_dataset['content'].apply(stemming)
```

```
print(news_dataset['content'])
```

```
0      darrel lucu hous dem aid even see come letter...
```

```
1      daniel j flynn flynn hillari clinton big woman...
```

```
2      consortiumnew com truth might get fire
```

```
3      jessica purkiss civilian kill singl us airstri...
```

```
4      howard portnoy iranian woman jail fiction unpu...
```

```
...
```

```
20795  jerom hudson rapper trump poster child white s...
```

```
20796  benjamin hoffman n f l playoff schedul matchup...
```

```
20797  michael j de la merc rachel abram maci said re...
```

```
20798  alex ansari nato russia hold parallel exercis ...
```

```
20799  david swanson keep f aliv
```

```
Name: content, Length: 20800, dtype: object
```

```
X=news_dataset['content'].values
```

```
Y=news_dataset['label'].values
```

```
Vectorizer = TfidfVectorizer()
```

```
vectorizer = TfidfVectorizer()  
Vectorizer.fit(X)  
X=Vectorizer.transform(X)
```

Double-click (or enter) to edit

```
import matplotlib.pyplot as plt
```

```
X_train,X_test,Y_train,Y_test =train_test_split(X,Y,test_size=0.2,stratify=Y,random_state=2)
```

```
model=LogisticRegression()  
model.fit(X_train, Y_train)
```

```
LogisticRegression(C=1.0, class_weight=None, dual=False, fit_intercept=True,  
                    intercept_scaling=1, l1_ratio=None, max_iter=100,  
                    multi_class='auto', n_jobs=None, penalty='l2',  
                    random_state=None, solver='lbfgs', tol=0.0001, verbose=0,  
                    warm_start=False)
```

```
X_train_prediction=model.predict(X_train)  
train_data_accuracy=accuracy_score(X_train_prediction,Y_train)
```

```
X_test_prediction = model.predict(X_test)  
test_data_accuracy = accuracy_score(X_test_prediction,Y_test)
```

```
print(test_data_accuracy)
```

```
0.9790865384615385
```

```
X_new=X_test[0]  
prediction = model.predict(X_new)  
print(prediction)  
if(prediction == 0):  
    print("the news is real")  
else:  
    print("the news is Fake")
```

```
[1]  
the news is Fake
```

[+ Code](#)[+ Text](#)

✓ 0s completed at 3:53 PM

● ✕