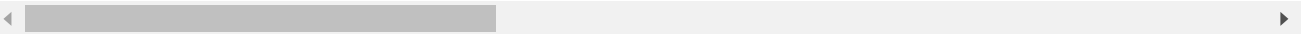


```
import pandas as pd
import numpy as np
df = pd.read_csv('https://raw.githubusercontent.com/YBI-Foundation/Dataset/main/Financial%
df.head()
```

	Date	Label	News 1	News 2	News 3	News 4	News 5	News 6	
0	01-01-2010	0	McIlroy's men catch cold from Gudjonsson	Obituary: Brian Walsh	Workplace blues leave employers in the red	Classical review: Rattle	Dance review: Merce Cunningham	Genetic tests to be used in setting premiums	revisi
1	02-01-2010	0	Warning from history points to crash	Investors flee to dollar haven	Banks and tobacco in favour	Review: Llama Farmers	War jitters lead to sell-off	Your not-so-secret history	Revisi
2	03-01-2010	0	Comment: Why Israel's peaceniks feel betrayed	Court deals blow to seizure of drug assets	An ideal target for spooks	World steps between two sides intent on war	What the region's papers say	Comment: Fear and rage in Palestine	Revisi
3	04-01-2010	1	£750,000-a-goal Weah aims parting shot	Newcastle pay for Fletcher years	Brown sent to the stands for Scotland qualifier	Tourists wary of breaking new ground	Canary Wharf climbs into the FTSE 100	Review: Bill Bailey	Revisi
4	05-01-2010	1	Leeds arrive in Turkey to the silence of the fans	One woman's vision offers loan lifeline	Working Lives: How world leaders worked	Working Lives: Tricks of the trade	Working Lives: six-hour days, long lunches and...	Pop review: We Love UK	Revisi

5 rows × 27 columns



```
df.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 4101 entries, 0 to 4100
Data columns (total 27 columns):
#   Column      Non-Null Count  Dtype
---  ---
0    Date        4101 non-null   object
```

```

1  Label      4101 non-null  int64
2  News 1     4101 non-null  object
3  News 2     4101 non-null  object
4  News 3     4101 non-null  object
5  News 4     4101 non-null  object
6  News 5     4101 non-null  object
7  News 6     4101 non-null  object
8  News 7     4101 non-null  object
9  News 8     4101 non-null  object
10 News 9     4101 non-null  object
11 News 10    4101 non-null  object
12 News 11    4101 non-null  object
13 News 12    4101 non-null  object
14 News 13    4101 non-null  object
15 News 14    4101 non-null  object
16 News 15    4101 non-null  object
17 News 16    4101 non-null  object
18 News 17    4101 non-null  object
19 News 18    4101 non-null  object
20 News 19    4101 non-null  object
21 News 20    4101 non-null  object
22 News 21    4101 non-null  object
23 News 22    4101 non-null  object
24 News 23    4100 non-null  object
25 News 24    4098 non-null  object
26 News 25    4098 non-null  object
dtypes: int64(1), object(26)
memory usage: 865.2+ KB

```

```
df.shape
```

```
(4101, 27)
```

```
df.columns
```

```

Index(['Date', 'Label', 'News 1', 'News 2', 'News 3', 'News 4', 'News 5',
      'News 6', 'News 7', 'News 8', 'News 9', 'News 10', 'News 11', 'News 12',
      'News 13', 'News 14', 'News 15', 'News 16', 'News 17', 'News 18',
      'News 19', 'News 20', 'News 21', 'News 22', 'News 23', 'News 24',
      'News 25'],
      dtype='object')

```

```
' '.join(str(x) for x in df.iloc[0,2:27])
```

'McIlroy's men catch cold from Gudjonsson Obituary: Brian Walsh Workplace blues leave employers in the red Classical review: Rattle Dance review: Merce Cunningham Genetic tests to be used in setting premiums Opera review: La Bohème Pop review: Britney Spears Theatre review: The Circle Wales face a fraught night Under-21 round-up Smith off to blot his copybook Finns taking the mickey Praise wasted as Brown studies in jury options Ireland wary of minnows Finland 0 - 0 England Healy a marked man Hannu

```
n = len(df)
```

```
n
```

```
4101
```

```

news = []
for i in range(0,n):
    news.append(''.join(str(x) for x in df.iloc[i,2:27]))

news[2]

'Comment: Why Israel's peaceniks feel betrayedCourt deals blow to seizure of drug as
setsAn ideal target for spooksWorld steps between two sides intent on warWhat the re
gion's papers sayComment: Fear and rage in PalestinePoverty and resentment fuels Pal
estinian furyRepublican feud fear as dissident is killedThe flood: Low-lying homes m
ay become uninsurableSynagogues and schools firebombedFlood: Alerts remain as clean-
up bill nears £2bnBitter divisions that halted talks within sight of a lasting peace
No headline'T have iust killed your husband' wife was toldTorrid time assuredETSE on

```

```
x = news
```

```

from numpy.ma.core import count
from sklearn.feature_extraction.text import CountVectorizer
cv=CountVectorizer()
x=cv.fit_transform(x)

```

```
x.shape
```

```
(4101, 108682)
```

```
y = df['Label']
```

```

from sklearn.model_selection import train_test_split
xtrain,xtest,ytrain,ytest = train_test_split(x,y,test_size =0.2,random_state=42,stratify =

```

```
from sklearn.ensemble import RandomForestClassifier
```

```

from scipy.sparse.construct import random
rf = RandomForestClassifier(n_estimators=300,random_state=2529,max_depth=32,n_jobs=-1)
rf.fit(xtrain,ytrain)
rf.score(xtest,ytest)

```

```
0.535931790499391
```

```

from sklearn.metrics import confusion_matrix,classification_report
ypred = rf.predict(xtest)
print('classification report: \n',classification_report(ytest,ypred))
print('confusion matrix: \n',confusion_matrix(ytest,ypred))

```

```

↳ classification report:

```

	precision	recall	f1-score	support
0	0.54	0.10	0.16	387
1	0.54	0.93	0.68	434
accuracy			0.54	821

macro avg	0.54	0.51	0.42	821
weighted avg	0.54	0.54	0.44	821

confusion matrix:

```
[[ 37 350]
```

```
[ 31 403]]
```

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

✓ 0s completed at 1:23 PM

