```python
In [3]:  import pandas as pd
         import numpy as np
```

```python
In [4]:  df = pd.read_csv('Data.csv', encoding = 'ISO-8859-1')
         df.head()
```

Out[4]:

Top1	Top2	Top3	Top4	Top5	Top6	Top7	Top8	...	Top
A rance to tions': tracts the...	Scorecard	Hughes' instant hit buoys Blues	Jack gets his skates on at ice-cold Alex	Chaos as Maracana builds up for United	Depleted Leicester prevail as Elliott spoils E...	Hungry Spurs sense rich pickings	Gunners so wide of an easy target	...	Flint injury pi on woe Engla
ecard	The best lake scene	Leader: German sleaze inquiry	Cheerio, boyo	The main recommendations	Has Cubie killed fees?	Has Cubie killed fees?	Has Cubie killed fees?	...	On t critical l
ventry ght on er by Flo	United's rivals on the road to Rio	Thatcher issues defence before trial by video	Police help Smith lay down the law at Everton	Tale of Trautmann bears two more retellings	England on the rack	Pakistan retaliate with call for video of Walsh	Cullinan continues his Cape monopoly	...	Sou Melbour (Austral
ilgrim s how gress	Thatcher facing ban	McIlroy calls for Irish fighting spirit	Leicester bin stadium blueprint	United braced for Mexican wave	Auntie back in fashion, even if the dress look...	Shoaib appeal goes to the top	Hussain hurt by 'shambles' but lays blame on e...	...	Pu adm Yeltsin q to give h a head
itches and rlocks	Beckham off but United survive	Breast cancer screening	Alan Parker	Guardian readers: are you all whingers?	Hollywood Beyond	Ashes and diamonds	Whingers - a formidable minority	...	Mc everywhe UD

```python
In [5]:  train = df[df['Date'] < '20150101']
         test = df[df['Date'] > '20141231']


         data = train.iloc[:, 2:27]
         data.replace("[^a-zA-z]", ' ', regex = True , inplace = True)
```

```
#renaming column.
list1 = [i for i in range(25)]
new_Index = [str(i) for i in list1]
data.columns= new_Index
data.head()
```

Out[5]:

	0	1	2	3	4	5	6	7
0	A hindrance to operations extracts from the…	Scorecard	Hughes instant hit buoys Blues	Jack gets his skates on at ice cold Alex	Chaos as Maracana builds up for United	Depleted Leicester prevail as Elliott spoils E…	Hungry Spurs sense rich pickings	Gunners so wide of an easy target	g St((
1	Scorecard	The best lake scene	Leader German sleaze inquiry	Cheerio boyo	The main recommendations	Has Cubie killed fees	Has Cubie killed fees	Has Cubie killed fees	H fur F Han
2	Coventry caught on counter by Flo	United s rivals on the road to Rio	Thatcher issues defence before trial by video	Police help Smith lay down the law at Everton	Tale of Trautmann bears two more retellings	England on the rack	Pakistan retaliate with call for video of Walsh	Cullinan continues his Cape monopoly	M put
3	Pilgrim knows how to progress	Thatcher facing ban	McIlroy calls for Irish fighting spirit	Leicester bin stadium blueprint	United braced for Mexican wave	Auntie back in fashion even if the dress look…	Shoaib appeal goes to the top	Hussain hurt by shambles but lays blame on e…	Eng dec di
4	Hitches and Horlocks	Beckham off but United survive	Breast cancer screening	Alan Parker	Guardian readers are you all whingers	Hollywood Beyond	Ashes and diamonds	Whingers a formidable minority	p(

5 rows × 25 columns

In [6]:
```
#converting the news to lower case.
for index in new_Index:
    data[index] = data[index].str.lower()
data.head()
```

Out[6]:

	0	1	2	3	4	5	6	7
0	a hindrance	scorecard	hughes instant	jack gets his	chaos as maracana builds	depleted leicester	hungry spurs	gunners so wide of	r

	0	1	2	3	4	5	6	7	
	a hindrance to operations extracts from the…		hit buoys blues	skates on at ice cold alex	up for united	prevail as elliott spoils e…	sense rich pickings	an easy target	gl str d
1	scorecard	the best lake scene	leader german sleaze inquiry	cheerio boyo	the main recommendations	has cubie killed fees	has cubie killed fees	has cubie killed fees	ho furic fo l hann
2	coventry caught on counter by flo	united s rivals on the road to rio	thatcher issues defence before trial by video	police help smith lay down the law at everton	tale of trautmann bears two more retellings	england on the rack	pakistan retaliate with call for video of walsh	cullinan continues his cape monopoly	mc puts c r
3	pilgrim knows how to progress	thatcher facing ban	mcilroy calls for irish fighting spirit	leicester bin stadium blueprint	united braced for mexican wave	auntie back in fashion even if the dress look…	shoaib appeal goes to the top	hussain hurt by shambles but lays blame on e…	eng deca dis
4	hitches and horlocks	beckham off but united survive	breast cancer screening	alan parker	guardian readers are you all whingers	hollywood beyond	ashes and diamonds	whingers a formidable minority	p pa

5 rows × 25 columns

In [11]:
```python
headlines = []
for row in range(len(data.index)):
    headlines.append(" ".join(str(x) for x in data.iloc[row,0:25]))
```

In [12]:
```python
headlines[0]
```

Out[12]:
```
'a  hindrance to operations   extracts from the leaked reports scorecard hughes  instant hit buoys blues jack gets his skates on at ice cold alex chaos as maracana builds up for united depleted leicester prevail as elliott spoils everton s party hungry spurs sense rich pickings gunners so wide of an easy target derby raise a glass to strupar s debut double southgate strikes  leeds pay the penalty hammers hand robson a youthful lesson saints party like it s        wear wolves have turned into lambs stump mike catches testy gough s taunt langer escapes to hit       flintoff injury piles on woe for england hunters threaten jospin with new battle of the somme kohl s successor drawn into scandal the difference between men and women sara denver  nurse turned solicitor diana s landmine crusade put tories in a panic yeltsin s resignation caught opposition flat footed russian roulette sold out recovering a title'
```

In [20]:
```python
headlines[1]
```

Out[20]: 'scorecard the best lake scene leader german sleaze inquiry cheerio boyo the main recommendations has cubie killed fees has cubie killed fees has cubie killed fees hopkins furious at foster s lack of hannibal appetite has cubie killed fees a tale of two tails i say what i like and i like what i say elbows eyes and nipples task force to assess risk of asteroid collision how i found myself at last on the critical list the timing of their lives dear doctor irish court halts ira man s extradition to northern ireland burundi peace initiative fades after rebels reject mandela as mediator pe points the way forward to the ecb campaigners keep up pressure on nazi war crimes suspect jane ratcliffe yet more things you wouldn t know without the movies millennium bug fails to bite'

In [26]:
```python
from sklearn.feature_extraction.text import CountVectorizer
from sklearn.ensemble import RandomForestClassifier
```

In [27]:
```python
Countvectors = CountVectorizer(ngram_range = (2,2))
traindataset = Countvectors.fit_transform(headlines)
```

In [31]:
```python
randomForest = RandomForestClassifier(n_estimators = 500, criterion = 'entropy')
randomForest.fit(traindataset, train['Label'])
```

Out[31]: RandomForestClassifier(criterion='entropy', n_estimators=500)

In [34]:
```python
test_transform = []
for row in range(len(test.index)):
    test_transform.append(" ".join(str(x) for x in test.iloc[row, 2:27]))
test_dataset = Countvectors.transform(test_transform)
predictons = randomForest.predict(test_dataset)
```

In [36]:
```python
from sklearn.metrics import classification_report, confusion_matrix, accuracy_score
matrix = confusion_matrix(test['Label'], predictons)
print(matrix)
score = accuracy_score(test['Label'], predictons)
print(score)
report = classification_report(test['Label'], predictons)
print(report)
```

```
[[137  49]
 [  5 187]]
0.8571428571428571
              precision    recall  f1-score   support
```

```
           0       0.96      0.74      0.84       186
           1       0.79      0.97      0.87       192

    accuracy                           0.86       378
   macro avg       0.88      0.86      0.85       378
weighted avg       0.88      0.86      0.85       378
```

In [ ]:

```
           0       0.96      0.74      0.84       186
           1       0.79      0.97      0.87       192

    accuracy                           0.86       378
   macro avg       0.88      0.86      0.85       378
weighted avg       0.88      0.86      0.85       378
```