

FINANCIAL MARKET NEWS SENTIMENT ANALYSIS

Objective:

The objective of the project is to use Machine Learning model for analyzing the Sentiments in the Financial market news.*italicised text*

Data Source:

Github link-<https://github.com/YBIFoundation/Dataset/blob/main/Financial%20Market%20News.csv>

✚ Import Library:

```
import pandas as pd
```

```
import numpy as np
```

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

```
import seaborn as sns
```

✚ Import Data:

```
dataset=pd.read_csv("https://raw.githubusercontent.com/YBIFoundation/Dataset/main/Financial%20Market%20News.csv", encoding="ISO-8859-1")
```

✚ Describe Data:

```
dataset.head()
```

	Date	Label	News 1	News 2	News 3	News 4	News 5	News 6	News 7	News 8	...	News 16	News 17
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	Opera review: La Bohème	Pop review: Britney Spears	...	Finland 0 - 0 England	Healy a market mar
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	Review: The Northern Sinfonia	Review: Hysteria	...	Why Wenger will stick to his Gunners	Out of luck: England hit rock bottom
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	Poverty and resentment fuels Palestinian fury	Republican feud fear as dissident is killed	...	FTSE goes upwardly mobile	At this price? BF Amocc
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	Review: Classical	Review: New Contemporaries 2000	...	More cash on way for counties	Cairns carries Kiwis to victory
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	World music review: Marisa Monte	Art review: Hollingsworth/Heyer	...	Duisenberg in double trouble	Prudential charges

5 rows × 27 columns

```
dataset.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
```

```
dataset.describe()
```

	Label1
count	4101.000000
mean	0.528164
std	0.499267
min	0.000000
25%	0.000000
50%	1.000000
75%	1.000000
max	1.000000

▼ Data Visualisation:

dataset.shape

(4101, 27)

dataset.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')
```

▼ Data Preprocessing and Feature Extraction

```
' '.join(str(x) for x in dataset.iloc[1,2:27])
```

'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 Review: The Northern Sinfonia Review: Hysteria Review: The Guardsman Opera: The Marriage of Figaro Review: The Turk in Italy Deutsche spells out its plans for diversification Traders' panic sends oil prices skyward TV sport chief leaves home over romance Leader: Hi-tech twitch Why Wenger will stick to his Gunners Out of luck England hit rock bottom Wilkinson out of his depth Kinsella sparks Irish power play Brown banished as Scots rebound Battling Wales cling to lifeline Ethiopia close to sealing Boro move Man-to-man marking Match stats French referee at centre of storm is no stranger to controversy'

dataset.index

RangeIndex(start=0, stop=4101, step=1)

len(dataset.index)

4101

```
news = []
for row in range(0, len(dataset.index)):
    news.append(' '.join(str(x) for x in dataset.iloc[row, 2:27]))
```

type(news)

list

news[1]

'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 Review: The Northern Sinfonia Review: Hysteria Review: The Guardsman Opera: The Marriage of Figaro Review: The Turk in Italy Deutsche spells out its plans for diversification Traders' panic sends oil prices skyward TV sport chief leaves home over romance Leader: Hi-tech twitch Why Wenger will stick to his Gunners Out of luck England hit rock bottom Wilkinson out of his depth Kinsella sparks Irish power play Brown banished as Scots rebound Battling Wales cling to lifeline Ethiopia close to sealing Boro move Man-to-man marking Match stats French referee at centre of storm is no stranger to controversy'

X=news

type(X)

list

✓ Feature text conversion to Bag of Words

```
from sklearn.feature_extraction.text import CountVectorizer
```

```
cv=CountVectorizer(lowercase=True, ngram_range=(1,1))
```

✓ Define Target Variable (y) and Feature Variables (X)

```
X=cv.fit_transform(X)
```

```
X.shape
```

```
(4101, 48527)
```

```
y=dataset['Label']
```

```
y.shape
```

```
(4101,)
```

✓ Train and Split

```
from sklearn.model_selection import train_test_split
```

```
X_train,X_test,y_train,y_test=train_test_split(X,y,train_size=0.7,stratify=y,random_state=2529)
```

✓ Modelling

```
from sklearn.ensemble import RandomForestClassifier rf=RandomForestClassifier(n_estimators=200)
```

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

```
# Create an instance of RandomForestClassifier  
rf = RandomForestClassifier(n_estimators=200)
```

```
# Fit the model to the training data  
rf.fit(X_train, y_train)
```

```
▼      RandomForestClassifier  
RandomForestClassifier(n_estimators=200)
```

```
rf.fit(X_train,y_train)
```

```
▼      RandomForestClassifier  
RandomForestClassifier(n_estimators=200)
```

✓ Prediction

```
y_pred=rf.predict(X_test)
```

✓ Evaluation Matrix

```
from sklearn.metrics import confusion_matrix,accuracy_score,classification_report
```

```
confusion_matrix(y_test,y_pred)
```