

```
[1]: import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
import seaborn as sns
import io
%matplotlib inline
import warnings
warnings.filterwarnings("ignore")

In [2]: filepath2 = r"C:\Users\91623\OneDrive\Desktop\projects\finacial news\news.csv"
df= pd.read_csv(filepath2)
print(df)

0      01-01-2010      0      McIlroy's men catch cold from Gudjonsson      News 1
1      02-01-2010      0      Warning from history points to crash
2      03-01-2010      0      Comment: Why Israel's peaceniks feel betrayed
3      04-01-2010      1      750,000-a-goal Weah aims parting shot
4      05-01-2010      1      Leeds arrive in Turkey to the silence of the fans
...
4096  20-03-2021      0      Barclays and RBS shares suspended from trading...
4097  21-03-2021      1      2,500 Scientists To Australia: If You Want To ...
4098  22-03-2021      1      Explosion At Airport In Istanbul
4099  23-03-2021      1      Jamaica proposes marijuana dispensers for tour...
4100  24-03-2021      1      A 171-year-old woman in Mexico City finally re...

      News 2
0      Obituary: Brian Walsh
1      Investors flee to dollar haven
2      Court deals blow to seizure of drug assets
3      Newcastle pay for Fletcher years
4      One woman's vision offers loan lifeline
...
4096  Pope says Church should ask forgiveness from g...
4097  The personal details of 112,000 French police ...
4098  Yemeni former president: Terrorism is the offs...
4099  Stephen Hawking says public life has 'stupidity'...
4100  IMF chief backs Athens as permanent Olympic host

      News 3
0      Workplace blues leave employers in the red
1      Banks and tobacco in favour
2      An ideal target for spooks
3      Brown sent to the stands for Scotland qualifier
4      Working Lives: How world leaders worked
...
4096  Poland 'shocked' by xenophobic abuse of Poles ...
4097  S&P cuts United Kingdom sovereign credit r...
4098  UK must accept freedom of movement to access E...
4099  Boris Johnson says he will not run for Tory pa...
4100  The president of France says if Brexit won, so...

      News 4
0      Classical review: Rattle
1      Review: Llama Farmers
2      World steps between two sides intent on war
3      Tourists wary of breaking new ground
4      Working Lives: Tricks of the trade
...
4096  There will be no second referendum, cabinet ag...
4097  Huge helium deposit found in Africa
4098  Devastated: scientists too late to captive bre...
4099  Six gay men in Ivory Coast were abused and for...
4100  British Man Who Must Give Police 24 Hours' Not...

      News 5
0      Dance review: Merce Cunningham
1      War jitters lead to sell-off
2      What the region's papers say
3      Canary Wharf climbs into the FTSE 100
4      Working Lives: six-hour days, long lunches and...
...
4096  Scotland welcome to join EU, Merkel ally says
4097  CEO of the South African state broadcaster qui...
4098  British Labor Party leader Jeremy Corbyn loses...
4099  Switzerland denies citizenship to Muslim immig...
4100  100+ Nobel laureates urge Greenpeace to stop o...

      News 6
0      Genetic tests to be used in setting premiums
1      Your not-so-secret history
2      Comment: Fear and rage in Palestine
3      Review: Bill Bailey
4      Pop review: We Love UK
...
4096  Sterling dips below Friday's 31-year low amid ...
4097  Brexit: Cost investors $2 trillion, the worst o...
4098  A Muslim shop in the UK Was Just Firebombed B...
4099  Palestinian terrorist stabs israeli teen girl ...
4100  Brazil: Huge spike in number of police killing...

      News 7
0      Opera review: La Boh me
1      Review: The Northern Sinfonia
2      Poverty and resentment fuels Palestinian fury
3      Review: Classical
4      World music review: Marisa Monte
...
4096  No negative news about South African presiden...
4097  Hong Kong democracy activists call for return ...
4098  Mexican Authorities Sexually Torture Women in ...
4099  Puerto Rico will default on $1 billion of debt...
4100  Austria's highest court annuls presidential el...

      News 8
0      Pop review: Britney Spears
1      Review: Hysteria
2      Review: New Contemporaries 2000
3      Art review: Hollingsworth/Heyer
...
4096  Surge in Hate Crimes in the U.K. Following U.K...
4097  Brexit: Iceland president says UK can join 'tr...
4098  UK shares and pound continue to recover
4099  Republic of Ireland fans to be awarded medal f...
4100  Facebook wins privacy case, can track any Belg...

      News 16
0      Finland 0 - 0 England
1      Why Wenger will stick to his Gunners
2      FTSE goes upwardly mobile
3      More cash on way for counties
4      Duisenberg in double trouble
...
4096  German lawyers to probe Erdogan over alleged w...
4097  US, Canada and Mexico pledge 50% of power from...
4098  Escape Tunnel, Dug by Hand, Is Found at Holoca...
4099  Googles free wifi at Indian railway stations i...
4100  The United States has placed Myanmar, Uzbekist...

      News 17
0      Healy a marked man
1      Out of luck England hit rock bottom
2      At this price? BP Amoco
3      Cairns carries Kiwis to victory
4      Pru to cut pension charges
...
4096  Boris Johnson says the UK will continue to "in...
4097  There is increasing evidence that Australia is...
4098  The land under Beijing is sinking by as much a...
4099  Mounting evidence suggests 'hobbits' were wi...
4100  S&P revises European Union credit rating t...

      News 18
0      Happy birthday Harpers & Queen
1      Wilkinson out of his depth
2      Go fish
3      Year of Blanchflower's flourish when Spurs sto...
4      Art review: Paul Graham
...
4096  Richard Branson is calling on the UK governmen...
4097  Richard Branson, the founder of Virgin Group, ...
4098  Car bomb and Anti-Islamic attack on Mosque in ...
4099  The men who carried out Tuesday's terror attac...
4100  India gets $1 billion loan from World Bank for...

      News 19
0      Win unlimited access to the Raindance film fes...
1      Kinsella sparks Irish power play
2      Bosnian Serb blows himself up to evade law
3      New direct approach brings only pay-pep-blues
4      Shearer shot sparks Boro humiliation
...
4096  Turkey 'sorry for downing Russian jet'
4097  37,000-yr-old skull from Borneo reveals surpri...
4098  Emaciated lions in Talz Zoo are trapped in blo...
4099  Calls to suspend Saudi Arabia from UN Human Ri...
4100  U.S. sailors detained by Iran spoke too much u...

      News 20
0      Labour pledges 800m to bridge north-south divide
1      Brown banished as Scots rebound
2      Orange float delayed to 2001
3      Third Division round-up
4      Ridsdale's lingering fears as Leeds revisit Tu...
...
4096  Edward Snowden lawyer vows new push for pardon...
4097  Palestinians stone Weingart Wall worshippers; do...
4098  Rupert Murdoch describes Brexit as 'wonderful'...
4099  More Than 100 Nobel Laureates Call Out Greenpe...
4100  Mass fish kill in Vietnam solved as Taiwan ste...

      News 21
0      Wales: Lib-Lab pact firm despite resignation
1      Battling Wales cling to lifeline
2      Angry factory workers root out fear, favours a...
3      Second Division round-up
4      Champions League: Rangers v Galatasaray
...
4096  Brexit opinion poll reveals majority don't wan...
4097  Jean-Claude Juncker asks Farage: Why are you h...
4098  More than 40 killed in Yemen suicide attacks
4099  British pedophile sentenced to 85 years in US ...
4100  Philippines president Rodrigo Duterte urges pe...

      News 22
0      Donald Dewar
1      Ehiogu close to sealing Boro move
2      Smith defied advice on dome payout
3      First Division round-up
4      Champions League: Lazio v Arsenal
...
4096  Conservative MP Leave Campaigner: "The leave c...
4097  "Romanians for Remainians" offering a new home...
4098  Google Found Disastrous Symantec and Norton Vu...
4099  US permitted 1,200 offshore fracks in Gulf of ...
4100  Spain arrests three Pakistanis accused of prom...

      News 23
0      Regenerating homes regenerates well-being in ...
1      Man-to-man marking
2      Xerox takes the axe to jobs
3      McLean ends his career with a punch
4      Lazio 1 - 1 Arsenal
...
4096  Economists predict UK recession, Scotland weake...
4097  Brexit: Gibraltar in talks with, further to st...
4098  Extremist violence on the rise in Germany: Dom...
4099  We will be swimming in ridicule - French beach...
4100  Venezuela, where anger over food shortages is ...

      News 24
0      Win 100 worth of underwear
1      Match stats
2      Comment: Refugees in Britain
3      Heskey grabs triple crown
4      England in Pakistan
...
4096  New EU 'superstate plan by France, Germany: G...
4097  8 Suicide Bombers Strike Lebanon
4098  BBC News: Labour MPs pass Corbyn no-confidence...
4099  UEFA says no minutes of silence for Istanbul v...
4100  A Hindu temple worker has been killed by three...

      News 25
0      TV guide: Random views
1      French referee at centre of storm is no strang...
2      Maverick who sparked the new intifada
3      Weah on his way as City march on
4      England given olive-branch reception
...
4096  Pakistani clerics declare transgender marriage...
4097  Mexico's security forces routinely use 'sexual...
4098  Tiny New Zealand town with 'too many jobs' lau...
4099  Law Enforcement Sources: Gun Used in Paris Ter...
4100  Ozone layer hole seems to be healing - US amp...

[4101 rows x 27 columns]
```

```
In [3]: df.head()

Out[3]:
```

	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 Boh me	Pop review: Britney Spears	...	Finland 0 - 0 England	Healy a marked man
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? BP Amoco
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	Pru to cut pension charges

5 rows x 27 columns

```
In [23]: df.tail()
```

```
Out[23]:
```

	Date	Label	News 1	News 2	News 3	News 4	News 5	News 6	News 7	News 8	...	News 16	News 17
4096	20-03-2021	0	Barclays and RBS shares suspended from trading...	Pope says Church should ask forgiveness from g...	Poland 'shocked' by xenophobic abuse of Poles ...	There will be no second referendum, cabinet ag...	Scotland welcome to join EU, Merkel ally says	Sterling dips below Friday's 31-year low amid ...	No negative news about South African President...	Surge in Hate Crimes in the U.K. Following U.K...	...	German lawyers to probe Erdogan over alleged w...	Boris Johnson says the UK will continue to "in...
4097	21-03-2021	1	2,500 Scientists To Australia: If You Want To ...	The personal details of 112,000 French police ...	S&P cuts United Kingdom sovereign credit r...	Huge helium deposit found in Africa	CEO of the South African state broadcaster qui...	Brexit cost investors \$2 trillion, the worst o...	Hong Kong democracy activists call for return ...	Brexit: Iceland president says UK can join 'tr...	...	US, Canada and Mexico pledge 50% of power from...	There is increasing evidence that Australia is...
4098	22-03-2021	1	Explosion At Airport In Istanbul	Yemeni former president: Terrorism is the offs...	UK must accept freedom of movement to access E...	Devastated: scientists too late to captive bre...	British Labor Party leader Jeremy Corbyn loses...	A Muslim Shop in the UK Was Just Firebombed Wh...	Mexican Authorities Sexually Torture Women in ...	UK shares and pound continue to recover	...	Escape Tunnel, Dug by Hand, Is Found at Holoca...	The land under Beijing is sinking by as much a...
4099	23-03-2021	1	Jamaica proposes marijuana dispensers for tour...	Stephen Hawking says pollution and 'stupidity'...	Boris Johnson says he will not run for Tory pa...	Six gay men in Ivory Coast were abused and for...	Switzerland denies citizenship to Muslim immig...	Palestinian terrorist stabs israeli teen girl ...	Puerto Rico will default on \$1 billion of debt...	Republic of Ireland fans to be awarded medal f...	...	Googles free wifi at Indian railway stations i...	Mounting evidence suggests 'hobbits' were wi...
4100	24-03-2021	1	A 117-year-old woman in Mexico City finally re...	IMF chief offers Athens as permanent Olympic host	The president of France says if Brexit won, so...	British Man Who Must Give Police 24 Hours' Not...	100+ Nobel laureates urge Greenpeace to stop o...	Brazil: Huge spike in number of police killing...	Austria's highest court annuls presidential el...	Facebook wins privacy case, can track any Belg...	...	The United States has placed Myanmar, Uzbekist...	S&P revises European Union credit rating t...

5 rows x 27 columns

```
In [4]: 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
```

```
In [5]: df.shape
```

```
Out[5]: (4101, 27)
```

```
In [6]: ' '.join(str(x) for x in df.iloc[1,2:27])
```

```
Out[6]: "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: New Contemporaries 2000 Art review: Hollingsworth/Heyer Why Wenger will stick to his Gunners FTSE goes upwardly mobile At this price? BP Amoco"

In [7]: df.index
```

```
Out[7]: RangeIndex(start=0, stop=4101, step=1)
```

```
In [8]: news = []
for row in range(0,len(df.index)):
    news.append(' '.join(str(x) for x in df.iloc[row,2:27]))
```

```
In [9]: type(news)
```

```
Out[9]: list
```

```
In [10]: news[0]
```

```
Out[10]: "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 Boh me Pop review: Britney Spears Review: Hysteria Review: The Northern Sinfonia Review: Hysteria Review: New Contemporaries 2000 Art review: Hollingsworth/Heyer Why Wenger will stick to his Gunners FTSE goes upwardly mobile At this price? BP Amoco"

In [11]: x=news
```

```
In [12]: type(x)
```

```
Out[12]: list
```

```
In [13]: from sklearn.feature_extraction.text import CountVectorizer
```

```
In [14]: cv = CountVectorizer(lowercase = True ,ngram_range =(1,1))
```

```
In [15]: x = cv.fit_transform(x)
```

```
In [16]: x.shape
```

```
Out[16]: (4101, 48463)
```

```
In [17]: y = df['Label']
```

```
In [18]: y.shape
```

```
Out[18]: (4101,)
```

```
In [19]: from sklearn.model_selection import train_test_split
x_train,x_test,y_train,y_test =train_test_split(x,y, test_size=0.3,random_state=2529)
x_train.shape,x_test.shape,y_train.shape,y_test.shape
```

```
Out[19]: ((2870, 48463), (1231, 48463), (2870,), (1231,))
```

```
In [20]: from sklearn.ensemble import RandomForestClassifier
rfc=RandomForestClassifier()
rfc.fit(x_train,y_train)
rfc_pred=rfc.predict(x_test)
print(confusion_matrix(y_test,y_pred))
print(classification_report(y_test,y_pred))
```

```
In [21]: from sklearn.metrics import confusion_matrix ,classification_report
```

```
In [22]: print(confusion_matrix(y_test,rfc_pred))
print(classification_report(y_test,rfc_pred))
```

		precision	recall	f1-score	support
	0	0.51	0.24	0.33	594
	1	0.53	0.79	0.63	637
accuracy				0.52	1231
macro avg		0.52	0.51	0.48	1231
weighted avg		0.52	0.52	0.49	1231

```
In [25]: from sklearn.linear model import LogisticRegression
lr=LogisticRegression()
lr.fit(x_train,y_train)
y_pred = lr.predict(x_test)
print(confusion_matrix(y_test,y_pred))
print(classification_report(y_test,y_pred))
```

		precision	recall	f1-score	support
	0	0.50	0.47	0.47	594
	1	0.53	0.54	0.53	637
accuracy				0.50	1231
macro avg		0.51	0.50	0.39	1231
weighted avg		0.51	0.52	0.40	1231

```
In [27]: from sklearn import tree
clf = tree.DecisionTreeClassifier()
clf.fit(x_train,y_train)
clf_pred=clf.predict(x_test)
print(confusion_matrix(y_test,clf_pred))
print(classification_report(y_test,clf_pred))
```

		precision	recall	f1-score	support
	0	0.48	0.46	0.47	594
	1	0.52	0.54	0.53	637
accuracy				0.50	1231
macro avg		0.50	0.50	0.39	1231
weighted avg		0.50	0.50	0.50	1231

```
In [28]: from sklearn.svm import SVC
svc= SVC()
svc.fit(x_train,y_train)
svc_pred=svc.predict(x_test)
from sklearn.metrics import confusion_matrix ,classification_report
print(confusion_matrix(y_test,svc_pred))
print(classification_report(y_test,svc_pred))
```

		precision	recall	f1-score	support
	0	0.51	0.07	0.12	594
	1	0.52	0.94	0.67	637
accuracy				0.52	1231
macro avg		0.51	0.50	0.39	1231
weighted avg		0.51	0.52	0.40	1231

```
In [ ] :
```

```
In [ ] :
```