

MACHINE LEARNING

LAB EXPERIMENT-3

TOPIC CHOSEN: Sentimental Analysis – FAANG Stock news

<https://www.kaggle.com/code/bayunova/faang-stock-news>

DATASET NAME: FAANG_STOCK_NEWS.csv

DATASET DETAILS:

The dataset contains the news of FAANG such as (Facebook, Amazon, Apple, Netflix, Google). It is collected from finviz website did web scrapping.

COLUMNS:

Ticker, data, time, headlines, neg, neu , pos, compound

TASKS:

Create a machine learning model predicts if a stock is buying or selling or hold based on news headlines and sentimental analysis.

1. IMPORTING LIBRARIES:

STEP 1:

```
import numpy as np

import pandas as pd import seaborn as sns
import nltk nltk.download('stopwords')
from textblob import TextBlob from
wordcloud import WordCloud import re
import string string.punctuation import
matplotlib.pyplot as plt from
nltk.stem.porter import PorterStemmer
from nltk.stem import WordNetLemmatizer
```

2. DATA EXTRACTION:

STEP 2:

```
stock = pd.read_csv('FAANG_STOCK_NEWS.csv')
```

STEP 3:

```
stock.shape
```

OUTPUT:

```
(500, 9)
```

STEP 4:

```
stock.head()
```

OUTPUT:

	Unnamed: 0	ticker	date	time	headline	neg	neu	pos	compound
0	0	FB	2021-09-22	06:30AM	2 Growth Stocks to Buy Hand Over Fist If the M...	0.000	0.608	0.392	0.7003
1	1	FB	2021-09-22	02:23AM	UPDATE 1- Facebook wraps up deals with Australi...	0.167	0.833	0.000	-0.3400
2	2	FB	2021-09-22	12:36AM	Facebook wraps up deals with Australian media ...	0.179	0.821	0.000	-0.3400
3	3	FB	2021-09-22	12:27AM	Facebook wraps up deals with Australia media f...	0.000	1.000	0.000	0.0000
4	4	FB	2021-09-21	10:34PM	Facebook overpaid FTC fine as quid pro quo to ...	0.099	0.659	0.242	0.3818

STEP 5:

```
stock = stock.drop('Unnamed: 0', 1)
stock.head()
```

OUTPUT:

	ticker	date	time	headline	neg	neu	pos	compound
0	FB	2021-09-22	06:30AM	2 Growth Stocks to Buy Hand Over Fist If the M...	0.000	0.608	0.392	0.7003
1	FB	2021-09-22	02:23AM	UPDATE 1-Facebook wraps up deals with Australi...	0.167	0.833	0.000	-0.3400
2	FB	2021-09-22	12:36AM	Facebook wraps up deals with Australian media ...	0.179	0.821	0.000	-0.3400
3	FB	2021-09-22	12:27AM	Facebook wraps up deals with Australia media f...	0.000	1.000	0.000	0.0000
4	FB	2021-09-21	10:34PM	Facebook overpaid FTC fine as quid pro quo to ...	0.099	0.659	0.242	0.3818

STEP 6:

```
stock.describe()
```

OUTPUT:

	neg	neu	pos	compound
count	500.000000	500.000000	500.000000	500.000000
mean	0.056344	0.822262	0.121406	0.100331
std	0.111799	0.171060	0.157284	0.346107
min	0.000000	0.306000	0.000000	-0.743000
25%	0.000000	0.706750	0.000000	0.000000
50%	0.000000	0.821000	0.000000	0.000000
75%	0.000000	1.000000	0.231000	0.307000
max	0.519000	1.000000	0.694000	0.941300

STEP 7:

```
stock.info()
```

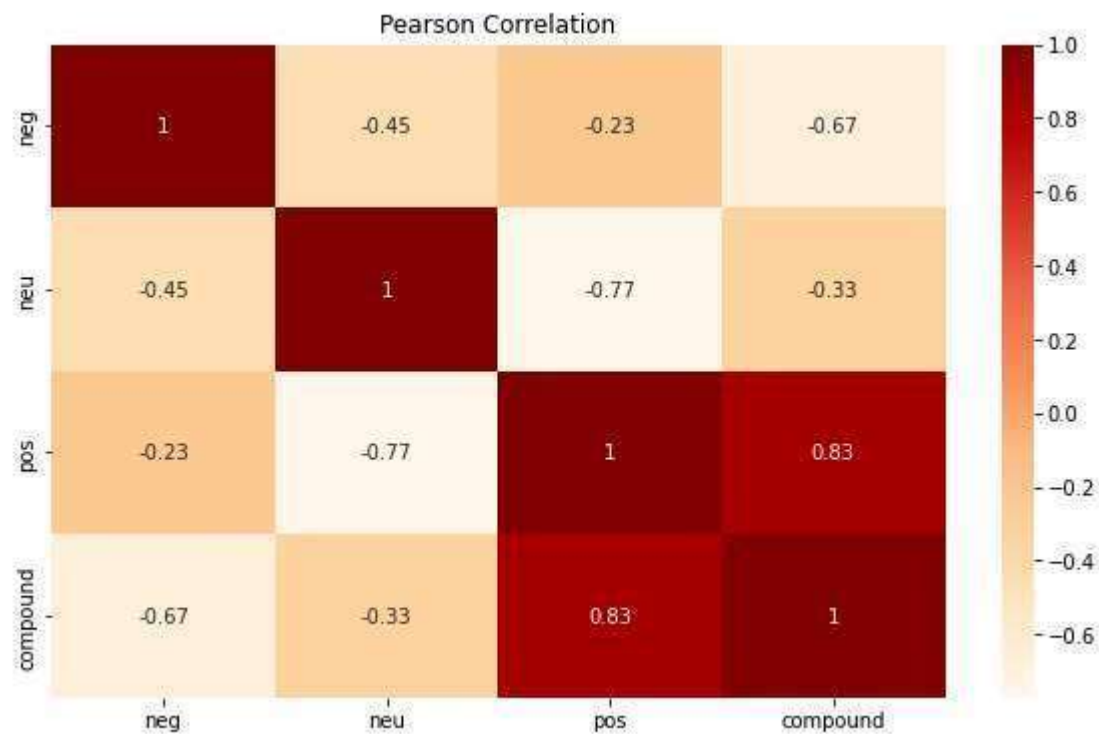
OUTPUT:

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 500 entries, 0 to 499
Data columns (total 8 columns):
 #   Column      Non-Null Count  Dtype
---  -
 0   ticker      500 non-null    object
 1   date        500 non-null    object
 2   time        500 non-null    object
 3   headline    500 non-null    object
 4   neg         500 non-null    float64
 5   neu         500 non-null    float64
 6   pos         500 non-null    float64
 7   compound    500 non-null    float64
dtypes: float64(4), object(4)
memory usage: 31.4+ KB
```

STEP 8:

```
plt.figure(figsize = (10,6))
sns.heatmap(stock.corr(), annot = True, cmap = 'OrRd')
plt.title("Pearson Correlation")
plt.show()
```

OUTPUT:



3. TEXT PROCESSING:

STEP 9:

```
headline = stock[['ticker', 'headline']] headline.head()
```

OUTPUT:

	ticker	headline
0	FB	2 Growth Stocks to Buy Hand Over Fist If the M...
1	FB	UPDATE 1-Facebook wraps up deals with Australi...
2	FB	Facebook wraps up deals with Australian media ...
3	FB	Facebook wraps up deals with Australia media f...
4	FB	Facebook overpaid FTC fine as quid pro quo to ...

STEP 10 (REMOVING PUNCTUATION):

OUTPUT:

```
def remove_punctuation(text):
```

```
    punctuationfree = "".join([i for i in text if i not in string.punctuation])    return
```

```
punctuationfree #storing the punctuation free text headline['clean_headline'] =
```

```
headline['headline'].apply(lambda x:remove_punctuation(x)) headline.head()
```

	ticker	headline	clean_headline
0	FB	2 Growth Stocks to Buy Hand Over Fist If the M...	2 Growth Stocks to Buy Hand Over Fist If the M...
1	FB	UPDATE 1-Facebook wraps up deals with Australi...	UPDATE 1Facebook wraps up deals with Australia...
2	FB	Facebook wraps up deals with Australian media ...	Facebook wraps up deals with Australian media ...
3	FB	Facebook wraps up deals with Australia media f...	Facebook wraps up deals with Australia media f...
4	FB	Facebook overpaid FTC fine as quid pro quo to ...	Facebook overpaid FTC fine as quid pro quo to ...

STEP 11 (SETTING LOWER CASE):

```
headline['headline_lower'] = headline['clean_headline'].apply(lambda x: x.lower())
```

```
headline.head()
```

OUTPUT:

	ticker	headline	clean_headline	headline_lower
0	FB	2 Growth Stocks to Buy Hand Over Fist If the M...	2 Growth Stocks to Buy Hand Over Fist If the M...	2 growth stocks to buy hand over fist if the m...
1	FB	UPDATE 1-Facebook wraps up deals with Australi...	UPDATE 1Facebook wraps up deals with Australia...	update 1facebook wraps up deals with australia...
2	FB	Facebook wraps up deals with Australian media ...	Facebook wraps up deals with Australian media ...	facebook wraps up deals with australian media ...
3	FB	Facebook wraps up deals with Australia media f...	Facebook wraps up deals with Australia media f...	facebook wraps up deals with australia media f...
4	FB	Facebook overpaid FTC fine as quid pro quo to ...	Facebook overpaid FTC fine as quid pro quo to ...	facebook overpaid ftc fine as quid pro quo to ...

STEP 12 (TOKENIZATION):

```
def tokenization(text):
```

```
    tokens = re.split('W+', text)
```

```
    return tokens
```

```
#applying function to the column headline['headline_tokenied'] =
```

```
headline['headline_lower'].apply(lambda x: tokenization(x))
```

```
headline.head()
```

OUTPUT:

	ticker	headline	clean_headline	headline_lower	headline_tokenied
0	FB	2 Growth Stocks to Buy Hand Over Fist If the M...	2 Growth Stocks to Buy Hand Over Fist If the M...	2 growth stocks to buy hand over fist if the m...	[2 growth stocks to buy hand over fist if the ...
1	FB	UPDATE 1-Facebook wraps up deals with Australi...	UPDATE 1Facebook wraps up deals with Australia...	update 1facebook wraps up deals with australia...	[update 1facebook wraps up deals with australi...
2	FB	Facebook wraps up deals with Australian media ...	Facebook wraps up deals with Australian media ...	facebook wraps up deals with australian media ...	[facebook wraps up deals with australian media...
3	FB	Facebook wraps up deals with Australia media f...	Facebook wraps up deals with Australia media f...	facebook wraps up deals with australia media f...	[facebook wraps up deals with australia media ...
4	FB	Facebook overpaid FTC fine as quid pro quo to ...	Facebook overpaid FTC fine as quid pro quo to ...	facebook overpaid ftc fine as quid pro quo to ...	[facebook overpaid ftc fine as quid pro quo to...

STEP 13 (STEMMING):

```
stopwords = nltk.corpus.stopwords.words('english') stopwords[0:10]
```

```
['i', 'me', 'my', 'myself', 'we', 'our', 'ours', 'ourselves', 'you', "you're"]
```

```
def remove_stopwords(text):
```



```

output= [i for i in text if i not in stopwords]
return output

```

```

#applying the function headline['no_stopwords']= headline['headline_tokenied'].apply(lambda
x:remove_stopwords(x)) headline.head()

```

	ticker	headline	clean_headline	headline_lower	headline_tokenied	no_stopwords
0	FB	2 Growth Stocks to Buy Hand Over Fist If the M...	2 Growth Stocks to Buy Hand Over Fist If the M...	2 growth stocks to buy hand over fist if the m...	[2 growth stocks to buy hand over fist if the ...	[2 growth stocks to buy hand over fist if the ...
1	FB	UPDATE 1-Facebook wraps up deals with Australi...	UPDATE 1Facebook wraps up deals with Australia...	update 1facebook wraps up deals with australia...	[update 1facebook wraps up deals with australi...	[update 1facebook wraps up deals with australi...
2	FB	Facebook wraps up deals with Australian media ...	Facebook wraps up deals with Australian media ...	facebook wraps up deals with australian media ...	[facebook wraps up deals with australian media...	[facebook wraps up deals with australian media...
3	FB	Facebook wraps up deals with Australia media f...	Facebook wraps up deals with Australia media f...	facebook wraps up deals with australia media f...	[facebook wraps up deals with australia media ...	[facebook wraps up deals with australia media ...
4	FB	Facebook overpaid FTC fine as quid pro quo to ...	Facebook overpaid FTC fine as quid pro quo to ...	facebook overpaid ftc fine as quid pro quo to ...	[facebook overpaid ftc fine as quid pro quo to...	[facebook overpaid ftc fine as quid pro quo to...

```

#defining the object for stemming porter_stemmer = PorterStemmer()

```

```

#defining a function for stemming def stemming(text):
stem_text = [porter_stemmer.stem(word) for word in text]
return stem_text

```

```

headline['headline_stemmed'] = headline['no_stopwords'].apply(lambda x: stemming(x)) headline.head()

```

OUTPUT:

	ticker	headline	clean_headline	headline_lower	headline_tokenied	no_stopwords	headl
0	FB	2 Growth Stocks to Buy Hand Over Fist If the M...	2 Growth Stocks to Buy Hand Over Fist If the M...	2 growth stocks to buy hand over fist if the m...	[2 growth stocks to buy hand over fist if the ...	[2 growth stocks to buy hand over fist if the ...	[2 gro to buy fist if
1	FB	UPDATE 1-Facebook wraps up deals with Australia...	UPDATE 1Facebook wraps up deals with Australia...	update 1facebook wraps up deals with australia...	[update 1facebook wraps up deals with australia...	[update 1facebook wraps up deals with australia...	[upda 1face up de austr
2	FB	Facebook wraps up deals with Australian media ...	Facebook wraps up deals with Australian media ...	facebook wraps up deals with australian media ...	[facebook wraps up deals with australian media...	[facebook wraps up deals with australian media...	[facel up de austr
3	FB	Facebook wraps up deals with Australia media f...	Facebook wraps up deals with Australia media f...	facebook wraps up deals with australia media f...	[facebook wraps up deals with australia media ...	[facebook wraps up deals with australia media ...	[facel up de austr

STEP 14:

```
#selection data df = headline[['ticker',
'clean_headline']]
df.head()
```

	ticker	clean_headline
0	FB	2 Growth Stocks to Buy Hand Over Fist If the M...
1	FB	UPDATE 1Facebook wraps up deals with Australia...
2	FB	Facebook wraps up deals with Australian media ...
3	FB	Facebook wraps up deals with Australia media f...
4	FB	Facebook overpaid FTC fine as quid pro quo to ...

```
#create function to get subjectivity def
getSubjectivity(text):
    return TextBlob(text).sentiment.subjectivity

#create function to get polarity def
getPolarity(text):
    return TextBlob(text).sentiment.polarity
```

```
#apply function to data
```

```
df['subjectivity'] = df['clean_headline'].apply(getSubjectivity)  
df['polarity'] = df['clean_headline'].apply(getPolarity) df.head()
```

0	FB	2 Growth Stocks to Buy Hand Over Fist If the M...	0.0	0.000000
1	FB	UPDATE 1Facebook wraps up deals with Australia...	0.0	0.000000
2	FB	Facebook wraps up deals with Australian media ...	0.0	0.000000
3	FB	Facebook wraps up deals with Australia media f...	0.0	0.000000
4	FB	Facebook overpaid FTC fine as quid pro quo to ...	0.5	0.416667

```
#create function to get sentiment data
```

```
def getSentiment(score):    if score <  
0:
```

```
    return 'Negative'
```

```
elif score == 0:
```

```
    return 'Neutral'
```

```
else:
```

```
    return 'Positive'
```

```
#apply function to data df['sentiment'] = df['polarity'].apply(getSentiment) df.head()
```

OUTPUT:

```
df['sentiment'].value_counts()
```

```
Neutral    272  
Positive   173  
Negative    55  
Name: sentiment, dtype: int64
```

```
#get headline of sentiment neutral nl_headline
```

```
= df[df.sentiment == 'Neutral'] nl_headline =
```

```
nl_headline['clean_headline']
```

```
nl_headline.head()
```

```
0    2 Growth Stocks to Buy Hand Over Fist If the M...
1    UPDATE 1Facebook wraps up deals with Australia...
2    Facebook wraps up deals with Australian media ...
3    Facebook wraps up deals with Australia media f...
7    Facebook says Wall Street Journal report alleg...
Name: clean_headline, dtype: object
```

```
#get headline of sentiment positive
```

```
p_headline = df[df.sentiment == 'Positive']
```

```
p_headline = p_headline['clean_headline']
```

```
p_headline.head()
```

```
4    Facebook overpaid FTC fine as quid pro quo to ...
5    Is GBTC Stock A Good Buy As Bitcoin Slumps Aft...
6    Facebook Oversight Board to Review Special Rul...
8    These Are The 5 Best Stocks To Buy And Watch Now
13   Top Research Reports for Facebook Walmart Nov...
Name: clean_headline, dtype: object
```

```
#get headline of sentiment negative
```

```
n_headline = df[df.sentiment == 'Negative']
```

```
n_headline = n_headline['clean_headline']
```

```
n_headline.head()
```

```
22   Facebook spent over 13 billion on safety secur...
23   Facebook spent over 13 bln on safety security ...
28   Facebook Says It Has Spent 13 Billion on Safet...
32   Why Apple Facebook Amazon and Other Tech Stock...
39   TikTok Parent Forced To Restrict Screen Time F...
Name: clean_headline, dtype: object
```

```
def score(x):
```

```
    return TextBlob(x).sentiment
```

```
df['score'] = df['clean_headline'].apply(score)
df.score.head()
```

```
<ipython-input-37-f233604538ba>:4: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user\_guide/indexing.html#returning-a-view-versus-a-copy
df['score'] = df['clean_headline'].apply(score)
0      (0.0, 0.0)
1      (0.0, 0.0)
2      (0.0, 0.0)
3      (0.0, 0.0)
4  (0.4166666666666667, 0.5)
Name: score, dtype: object
```

STEP 15 (NEUTRAL):

```
neutral = " ".join(df[df.sentiment == 'Neutral']['clean_headline'].values) w =
WordCloud(width = 700, height = 400, random_state = 10, max_font_size = 100,
background_color = 'white').generate(neutral)
```

```
plt.figure(figsize = (10,6)) plt.imshow(w,
interpolation = "bilinear")
plt.title("Wordcloud of Neutral Headline")
plt.axis('off')
plt.show()
```

OUTPUT:



STEP 17 (POSITIVE):

```
positive = " ".join(df[df.sentiment == 'Positive']['clean_headline'].values) w = WordCloud(width =
700, height = 400, random_state = 10, max_font_size = 100, background_color =
'white').generate(positive)
```

```
plt.figure(figsize = (10,6))
plt.imshow(w, interpolation = "bilinear")
plt.title("Wordcloud of Positive Headline")
plt.axis('off')
plt.show()
```

OUTPUT:

