

Assignment-3

Q1. Sentiment Analysis

Feature Vector:-

[illegible]

Bing Lio Feature:-

	Positive-WordCount	Negative-WordCount	Pos-Neg-Diff-WordCount
0	0	0	0
1	0	0	0
2	1	1	0
3	0	2	2
4	0	0	0

MPQA:-

	Count_of_negative(mpqa)	Count_of_positive(mpqa)
0	0	0
1	0	1
2	1	4
3	5	2
4	0	1

Sentiment140:-

	Score(NRC-sentiment140)
0	2.801
1	0.342
2	-3.334
3	3.262
4	-4.011

NRC-Hashtag:-

```
Score(NRC-hashtag)
0      4.100
1     -1.337
2     -2.701
3      2.573
4     -5.368
```

N-gram:-

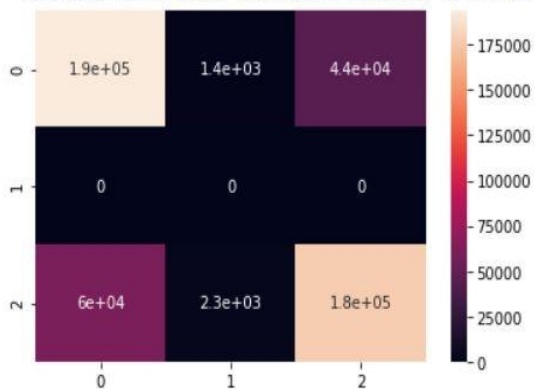
```
[[0 0 0 ... 0 0 0]
 [0 0 0 ... 0 0 0]
 [0 0 0 ... 0 0 1]
 ...
 [0 0 0 ... 0 0 0]
 [0 1 0 ... 0 0 0]
 [0 0 0 ... 0 0 0]]
```

Naive-Bayes:-

```
Confusion_matrix is :-
[[193711.  1390.  44103.]
 [    0.    0.    0.]
 [ 60433.  2334. 178029.]]
```

```
Classes Precision Recall
0  0.762210 0.809815
4  0.801456 0.739335
```

```
Macro average F1 score is:- 0.7781870612270346
Accuracy is :- 0.7744583333333334
Precision is :- 0.5212218348464259
Recall is :- 0.5163834747598104
/usr/local/lib/python3.6/dist-packages/sklearn/me
_warn_prf(average, modifier, msg_start, len(res
```



SVM:-

Confusion_matrix is :-

```
[[179082.    0.  60122.]
 [    0.    0.    0.]
 [ 47723.    0. 193073.]]
```

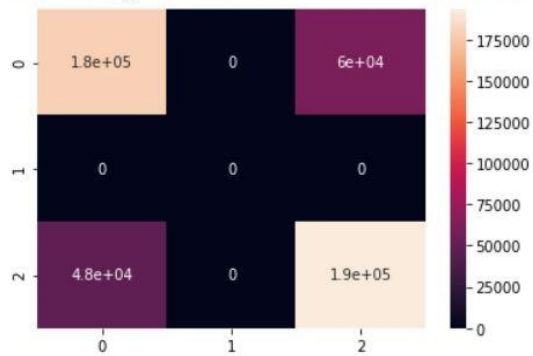
Classes	Precision	Recall
0	0.789586	0.748658
4	0.762547	0.801811

Accuracy is :- 0.7753229166666666

Precision is :- 0.7760662106372096

Recall is :- 0.7752347705408744

Macro average F1 score is:- 0.7756502677784326



Decision Tree:-

Confusion_matrix is :-

```
[[158535.    0.  80669.]
 [    0.    0.    0.]
 [ 86702.    0. 154094.]]
```

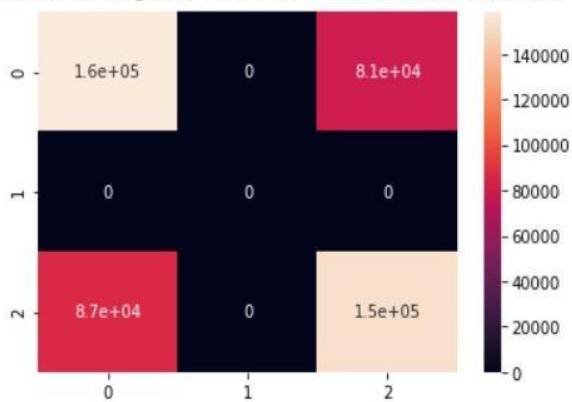
Classes	Precision	Recall
0	0.646456	0.662761
4	0.656381	0.639936

Accuracy is :- 0.6513104166666667

Precision is :- 0.6514187007096428

Recall is :- 0.6513482677549309

Macro average F1 score is:- 0.6513834823283389

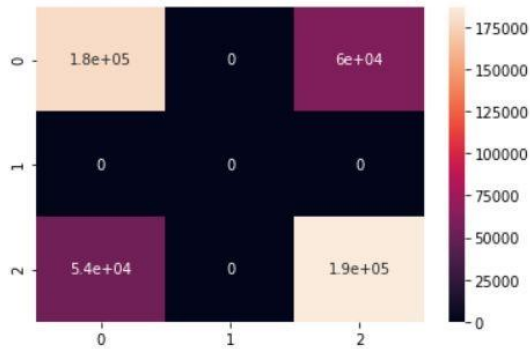


MLP:-

```
Confusion_matrix is :-  
[[179237.    0.  59967.]  
 [    0.    0.    0.]  
 [ 54003.    0. 186793.]]
```

Classes	Precision	Recall
0	0.768466	0.749306
4	0.756982	0.775731

```
Accuracy is :- 0.7625625  
Precision is :- 0.7627242254612054  
Recall is :- 0.7625186780561583  
Macro average F1 score is:- 0.7626214379085156
```

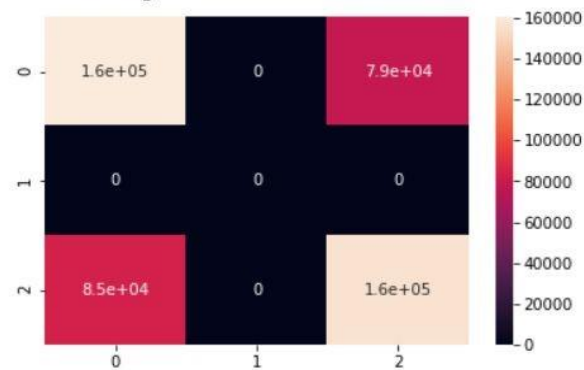


TF-IDF -decision tree

```
Confusion_matrix is :-  
[[160523.    0.  78681.]  
 [    0.    0.    0.]  
 [ 84522.    0. 156274.]]
```

Classes	Precision	Recall
0	0.655076	0.671072
4	0.665123	0.648989

```
Accuracy is :- 0.65999375  
Macro average F1 score is:- 0.6600648599211598
```



Q2. Emotion Analysis

ANGER

PEARSON COEFFICIENT:

SVM :

(0.3822372423369907,

Decision tree :

0.4144469609823051

MLP:

0.1073758509834736

SPEARMANS COEFFICIENT

SVM :

-0.4098899760865696,

Decision tree:

0.40733088582220556

MLP:

0.22308006471787512

JOY

PEARSON COEFFICIENT

SVM

0.19485844976798644

DCT

0.08699489709962026

MLP

0.1534077819441137

SPEARMANS COEFFICIENT

SVM

0.23741492342689663

DCT

0.0939532512457811

MLP

0.34760872703943824