Introduction to NLP

Assignment 4 Report

:- Hitesh Goel
2020115003

Sentiment Analysis

Confusion Matrices and Classification Reports

Train Data

```
Classification Report:
              precision    recall  f1-score   support

           0       0.62      0.70      0.66      4037
           1       0.70      0.61      0.65      4507

    accuracy                           0.65      8544
   macro avg       0.66      0.66      0.65      8544
weighted avg       0.66      0.65      0.65      8544


Confusion Matrix:
[[2843 1194]
 [1776 2731]]
```

Test Data:

```
Classification Report:
              precision    recall  f1-score   support

           0       0.53      0.57      0.55      1099
           1       0.54      0.50      0.52      1111

    accuracy                           0.54      2210
   macro avg       0.54      0.54      0.54      2210
weighted avg       0.54      0.54      0.54      2210


Confusion Matrix:
[[628 471]
 [554 557]]
```
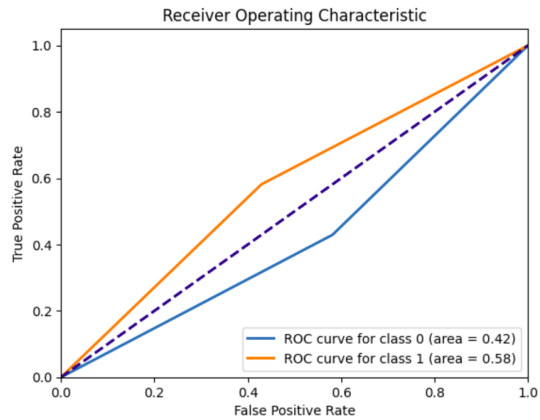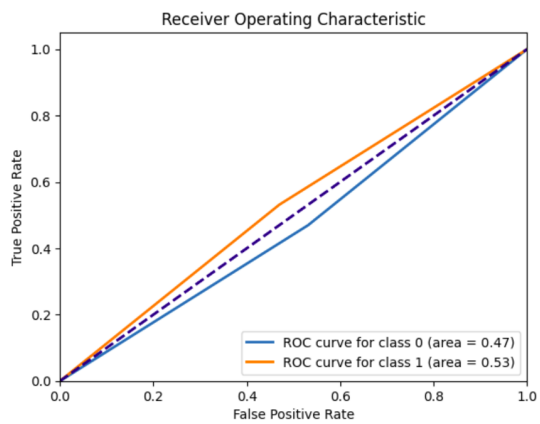
ROC Curves:

Train Data:

Receiver Operating Characteristic

Test Data:

Receiver Operating Characteristic

NLI

Confusion Matrices and Classification Report:

Train Data:

```
Classification Report:
              precision    recall  f1-score   support

           0       0.39      0.39      0.39     13525
           1       0.35      0.37      0.36     12122
           2       0.41      0.40      0.40     14353

    accuracy                           0.39     40000
   macro avg       0.38      0.38      0.38     40000
weighted avg       0.39      0.39      0.39     40000


Confusion Matrix:
[[5235 3968 4322]
 [3756 4483 3883]
 [4475 4180 5698]]
```
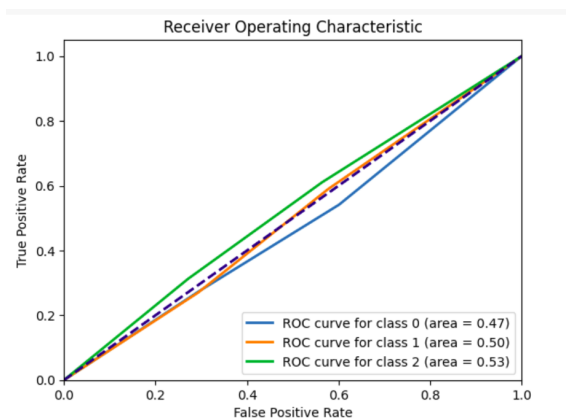
Test Data:

	precision	recall	f1-score	support
class 1	0.37	0.46	0.41	14208
class 2	0.33	0.32	0.32	12764
class 3	0.39	0.31	0.35	15097
accuracy			0.36	42069
macro avg	0.36	0.36	0.36	42069
weighted avg	0.37	0.36	0.36	42069

ROC Curves:

Train Data:

Test Data: