



PG DEPARTMENT OF COMPUTER APPLICATIONS

IMBD DATASET ANALYSIS

Submitted By,

Joshuval James

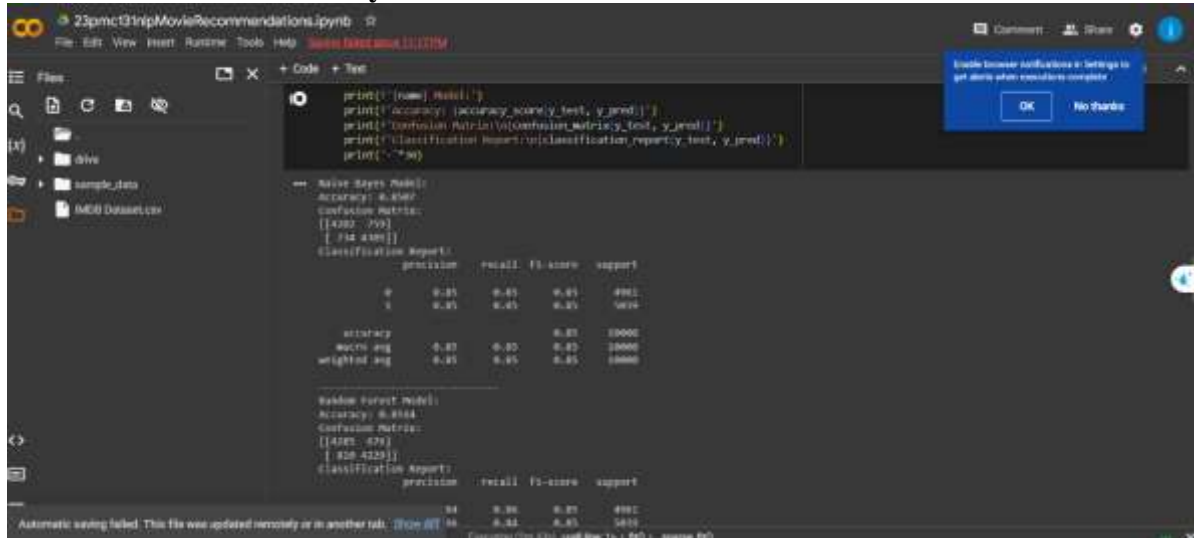
23PMC131

Submitted To,

Sr.Italia Joseph Maria

Assistant Professor

Screenshot of the accuracy:



```
print([name, model])
print('Accuracy: {accuracy_score(y_test, y_pred)}')
print('Confusion Matrix: {confusion_matrix(y_test, y_pred)}')
print('Classification Report: {classification_report(y_test, y_pred)}')
print('~*~')
```

```
Naive Bayes Model:
Accuracy: 0.875
Confusion Matrix:
[[4082  719]
 [ 234 4385]]
Classification Report:
      precision    recall  f1-score   support

0       0.85      0.85      0.85      4082
1       0.85      0.85      0.85      4604

 accuracy      0.87      0.87      0.87      8686
  weighted avg      0.85      0.85      0.85      8686

Random Forest Model:
Accuracy: 0.8164
Confusion Matrix:
[[4285  479]
 [ 828 4229]]
Classification Report:
      precision    recall  f1-score   support

0       0.86      0.86      0.86      4285
1       0.81      0.81      0.81      4314
```

According to the report I got the most accuracy in the ‘SVM’ algorithm (0.8755 ie,87.5%).

GitHub link:

<https://github.com/joshuval1017/movienlp>