DELHI TECHNOLOGICAL UNIVERSITY
(FORMERLY DELHI COLLEGE OF ENGINEERING)

SHAHBAD DAULATPUR, BAWANA ROAD
DELHI-110042

CO (PR: PATTERN RECOGNITION)

Submitted to:

Professor Anurag Goel

Submitted by:

Aadhar Dutta || Dhruv Yadav

(2K18/IT/001) || (2K18/SE/056)

Report

ON

COMPARE AND CONTRAST BETWEEN THE CLASSIFIERS EXPLAINED

SUBMITTED BY:

AADHAR DUTTA || DHRUV YADAV

2K18/IT/001 || 2K18/SE/056

In Adequate Fulfillment for the Award of the Degree of

Bachelors of Technology

From

DELHI TECHNOLOGICAL UNIVERSITY

Bawana Rd, Shahbad Daulatpur Village, Rohini, Delhi, 110042

This report has been written to compare and contrast the result of the two classifiers that have been built:

1. Namingly classifiers produced when priors are 0.5 & 0.5, respectively
2. When the priors are 0.1 & 0.9, respectively

To Compare and Contrast the two, we will discuss the possible differences between the Evaluation metrics of both, which will factually explain the differences.

Evaluation metrics of both are explained below:

1.) Prior 0.5 & 0.5

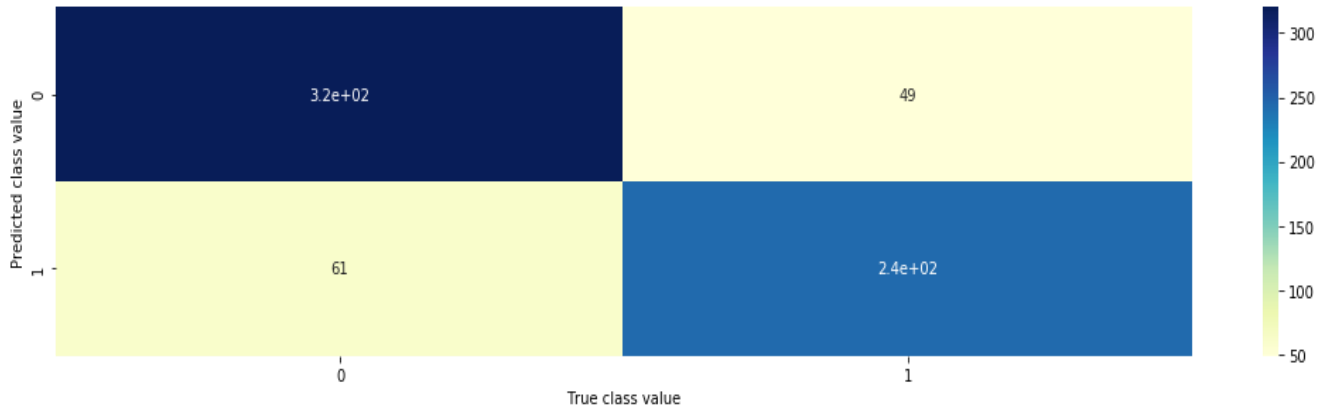
A.) Confusion matrix:

Firstly we will talk about the confusion matrix,

A Confusion matrix is an N x N matrix used for evaluating the performance of a classification model, where N is the number of target classes. The matrix compares the actual target values with those predicted by the machine learning model. This gives us a holistic view of how well our classification model performs and what kinds of errors it is making.

For the classifier built between the priors 0.5 & 0.5, we get:

```
320     49
61     244
```



B.) Classification Accuracy

Classification accuracy is a metric that summarizes the performance of a classification model as the number of correct predictions divided by the total number of predictions. It is easy to calculate and intuitive to understand, making it the most common metric used for evaluating classifier models.

```
[ ]  print("Accuracy of the model is : ", score(matrix), " %")

     Accuracy of the model is :  83.67952522255193  %
```

C.) Classification Report

A Classification report is used to measure the quality of predictions from a classification algorithm. ... The report shows the paramount classification metrics precision, recall and f1-score on a per-class basis. The metrics are calculated by using true and false positives, true and false negatives.

```
[ ] report = classification_report(Y_test, y_pred, target_names=["Class 0","Class 1"])
    print(report)
```

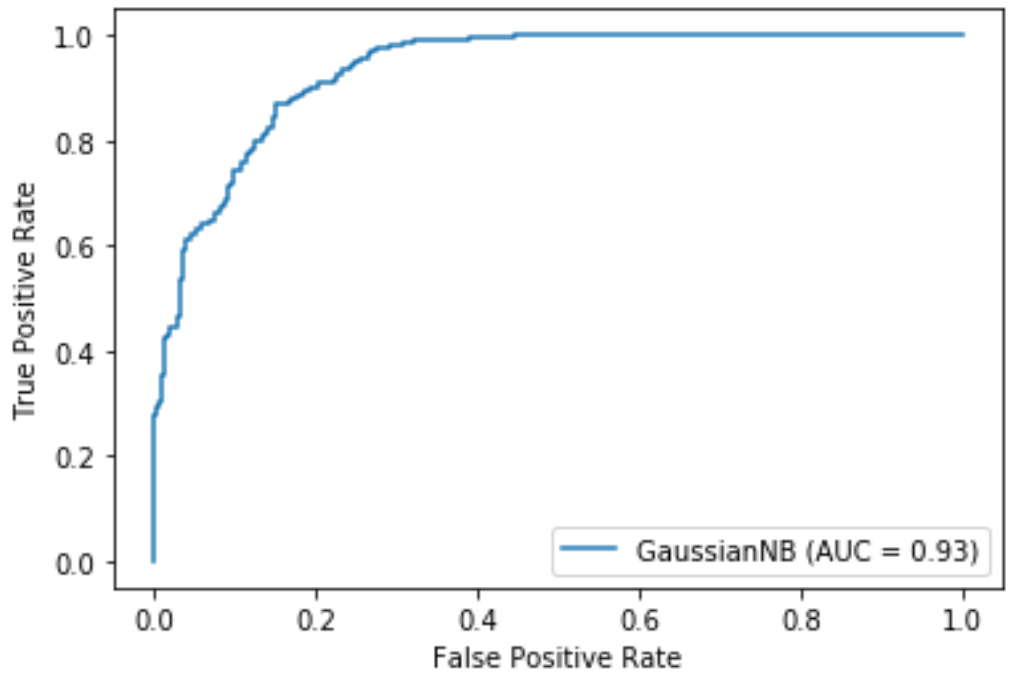
```
                precision    recall  f1-score   support

       Class 0       0.84      0.87      0.85       369
       Class 1       0.83      0.80      0.82       305

      accuracy                           0.84       674
     macro avg       0.84      0.83      0.83       674
  weighted avg       0.84      0.84      0.84       674
```

D.) ROC Curve Between FAR Vs GAR

A ROC curve (receiver operating characteristic curve) shows a classification model's performance at all classification thresholds. This curve plots two parameters: True Positive Rate. False Positive Rate.

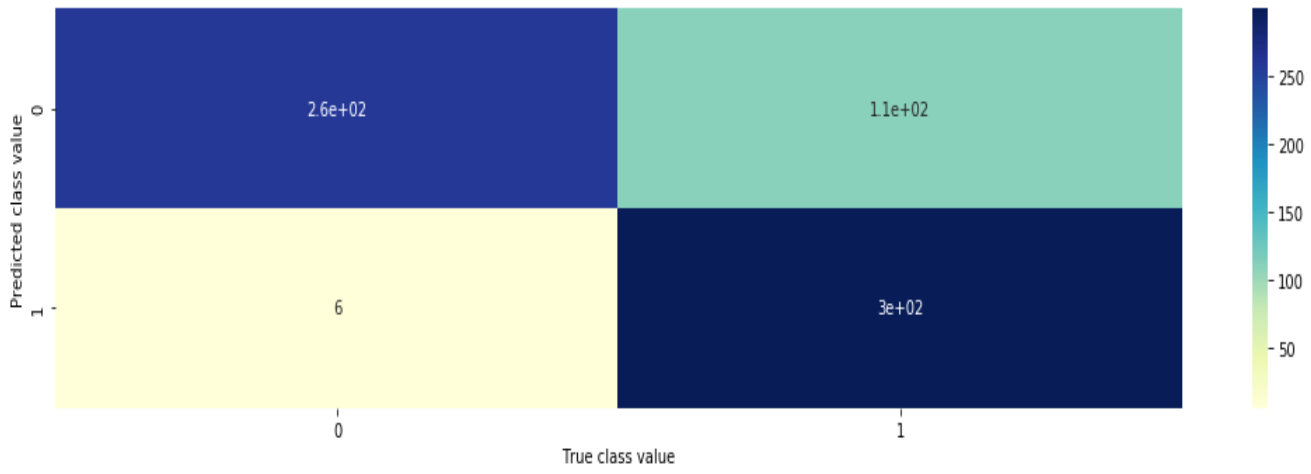


2.) Priors 0.1 & 0.9

A.) Confusion matrix:

For the classifier built for priors 0.1 & 0.9, we get:

```
258     111
6       299
```



B.) Classification Accuracy

```
[ ]  print("Accuracy of the model is : ", score(matrix), " %")

     Accuracy of the model is :  82.64094955489614  %
```

C.) Classification Report

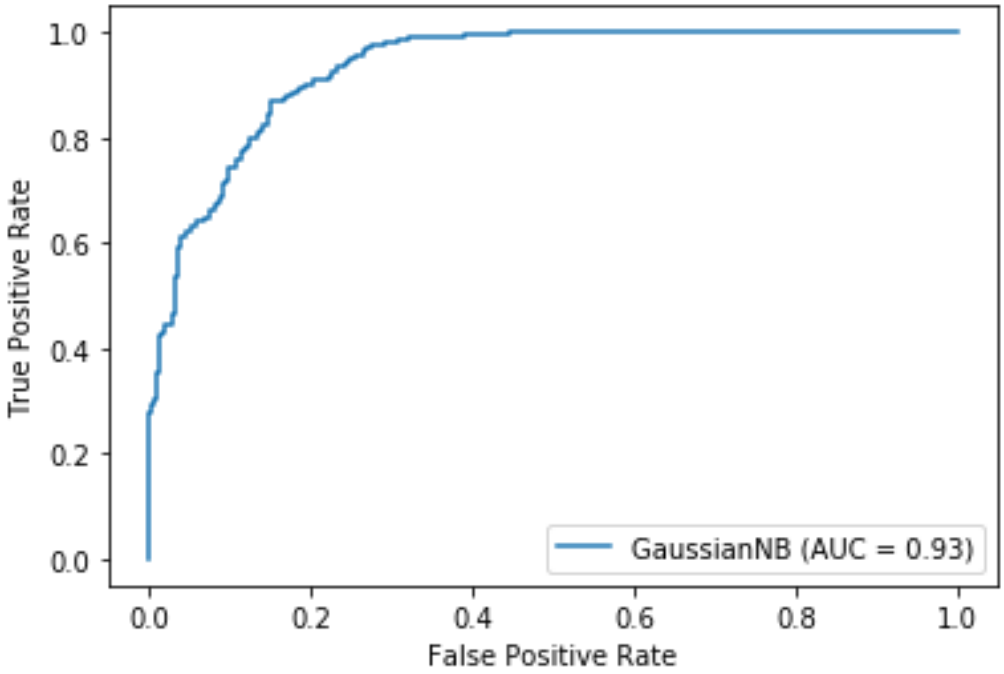
```
[ ]  report = classification_report(Y_test, y_pred, target_names=["Class 0","Class 1"])
     print(report)
```

```
                precision    recall  f1-score   support

       Class 0       0.98      0.70      0.82       369
       Class 1       0.73      0.98      0.84       305

      accuracy                           0.83       674
     macro avg       0.85      0.84      0.83       674
  weighted avg       0.87      0.83      0.82       674
```

D.) ROC Curve Between FAR Vs GAR



The contrast between the two classifiers is as below:

1. Confusion Matrix: The first classifier recognized class 0 more accurately than class 1, and the second classifier(priors[0.1,0.9]) identified class 1 more accurately than class 0.

2. Classification Accuracy : The first classifier(priors[0.5,0.5]) has more accuracy (83.67 %) than the second classifier (82.64 %). The difference is very less.
3. Classification Report: Giving very high prior(0.9) to class 1 leads to a very accurate classification of class 1. The recall of class 1 is much higher in the second classifier(prior [0.1,0.9]) than in the first classifier.
4. ROC Curve : The area under the ROC curve for both the classifiers were the same (0.93).