**REPORT**

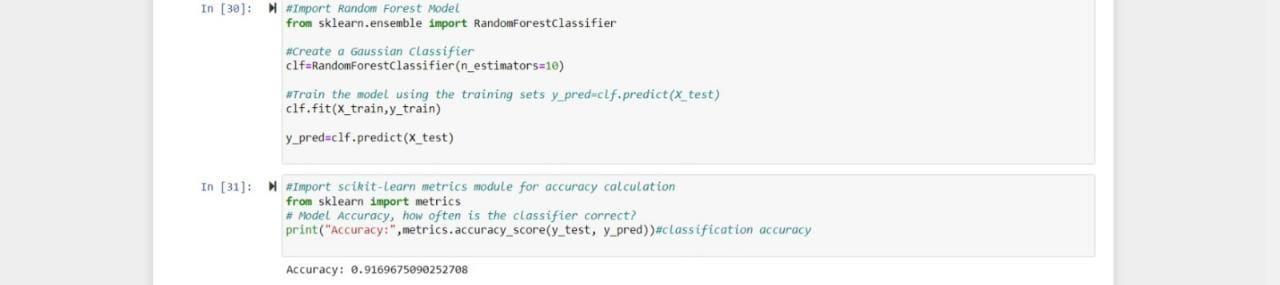
Hemanth Sai Muthyala

1319766

**Accuracies of the Random Forest classifier as a function of the number of base learners.**

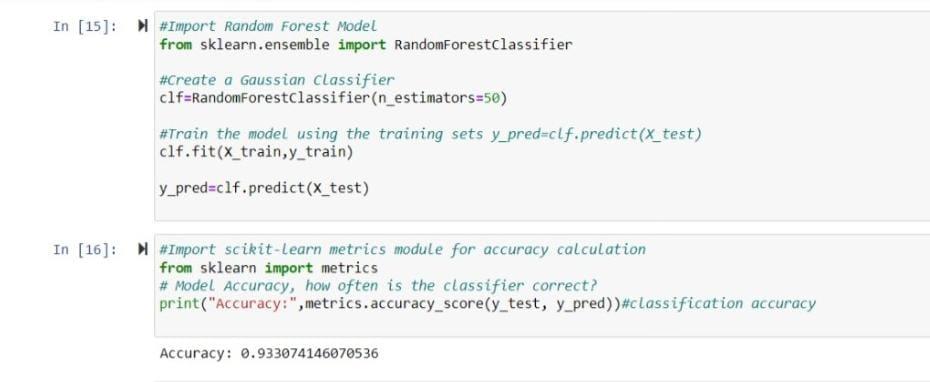
For base learners of 10

We got an accuracy score of 91 percent



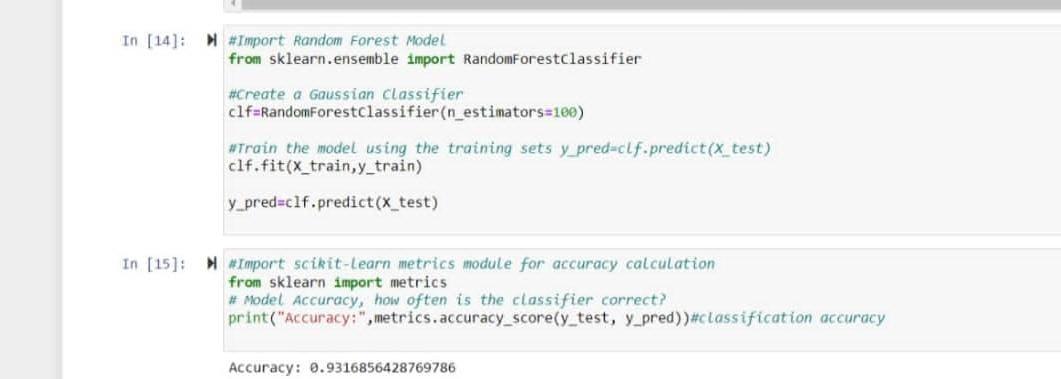
For base learners of 50

We got accuracy score of 93 percent



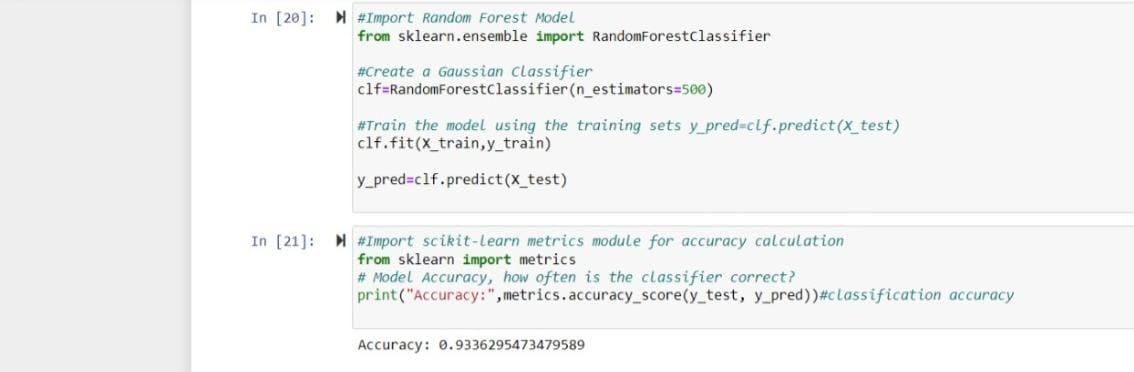
For base learners of 100

We got accuracy score of 93 percent



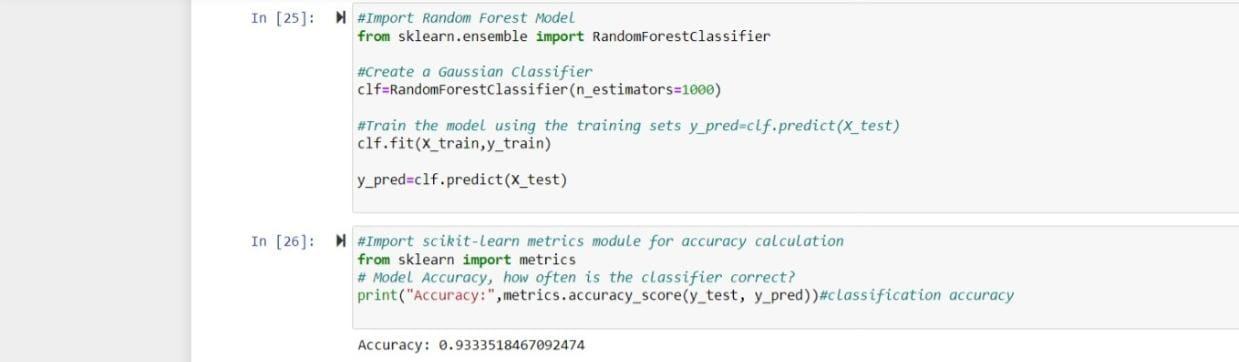
For base learners of 500

We got the accuracy of 93 percent.



For base learner of 1000

We got the same accuracy of 93 percent



**Observations :**

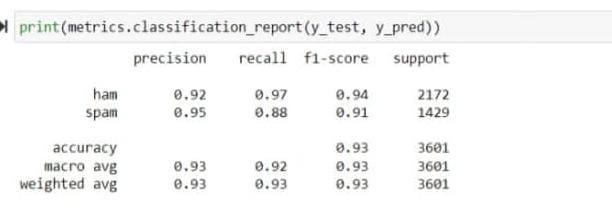
As we can clearly see the accuracy of the model is being increased with increase in the base learners. At first we considered the base learners as fifty the accuracy of

the model of at 91 percent. But later on increasing the base learners the accuracy of the model is increased to 93 percent.

At some point of time the accuracy of model remain same at 93 percent. From 100 observations the model remains 93 percent accuracy. More the accuracy Better the performance of model.

**Classification Report** :

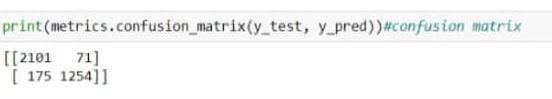
A Classification report is used to measure the quality of predictions from a classification algorithm. How many predictions are True and how many are False. More specifically, True Positives, False Positives, True negatives and False Negatives are used to predict the metrics of a classification



In classification report, the f1-score accuracy is 93 percentage.

**Confusion Matrix** :

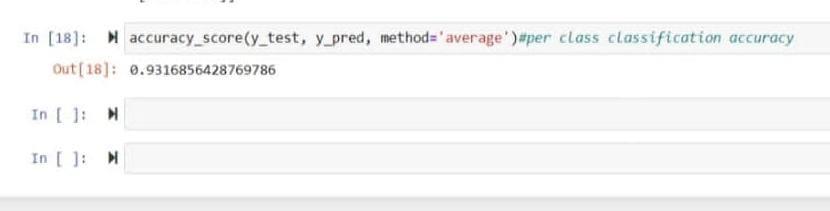
A confusion matrix is a way to express how many of a classifier's predictions were correct, and when incorrect, where the classifier got confused (hence the name!). In the confusion matrices below, the rows represent the true labels and the columns represent predicted labels. Values on the diagonal represent the number (or percent, in a normalized confusion matrix) of times where the predicted label matches the true label. Values in the other cells represent instances where the classifier mislabeled an observation; the column tells us what the classifier predicted, and the row tells us what the right label was. This is a convenient way to spot areas where the model may need a little extra training.



In confusion matrix the accuracy is 93 percentage.

**Accuracy Score** :

Accuracy is one metric for evaluating classification models. Informally, accuracy is the fraction of predictions our model got right. Formally, accuracy has the following definition: Accuracy = **Number of correct predictions Total number of predictions.**



By using y\_test,y\_pred,method=”average” the accuracy is 93 percentage.

Without using average method the accuracy is 93 percentage.

**Comparison**:

By using any method the accuracy is same as 93 percentage.