**REPORT**

Hemanth Sai Muthyala

1319766

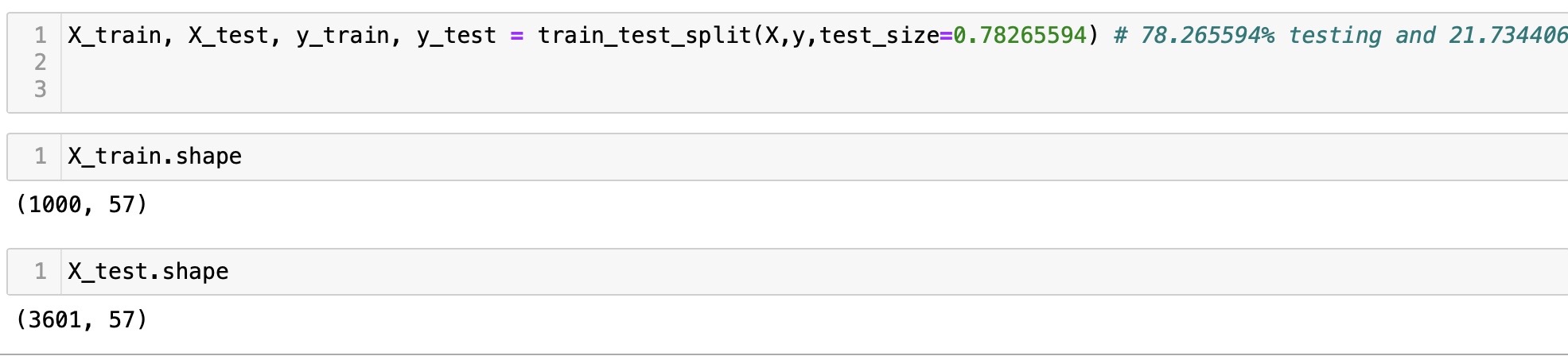
1. Compare the accuracies of the fused model with AdaBoost Ensemble with Decision Tree as the base learner.

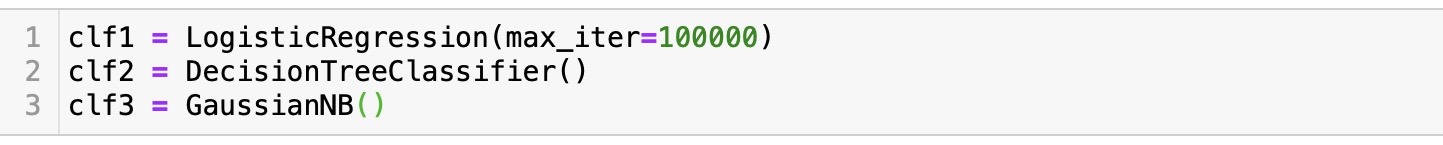
**Accuracies of the Fused model and AdaBoost Ensemble with Decision Tree.**

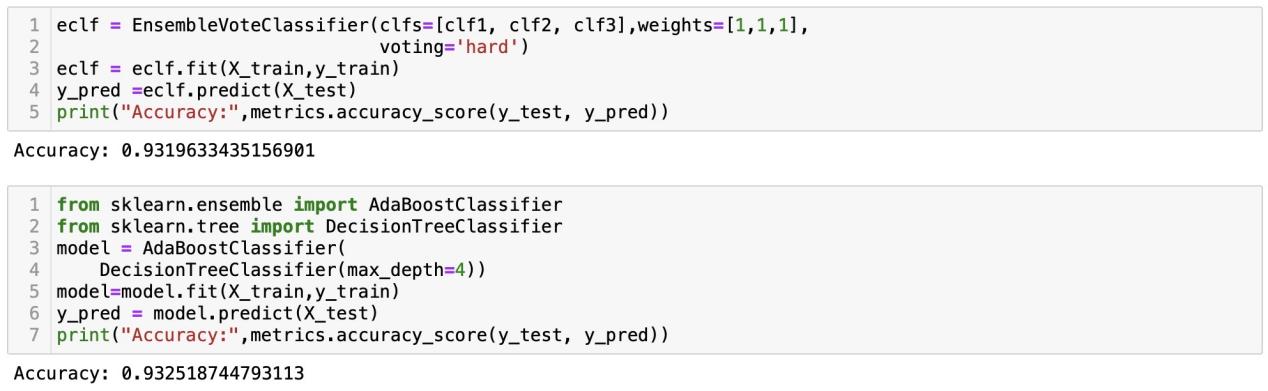
We got an accuracy score of 93.19 percent for fused model and 93.25 for AdBoost Ensemble.

1000 instances for training.

3601 instances for testing.







**Observations :**

As we can clearly see the accuracy of the AdaBoostClassifier is greater than Fused model by 0.0555

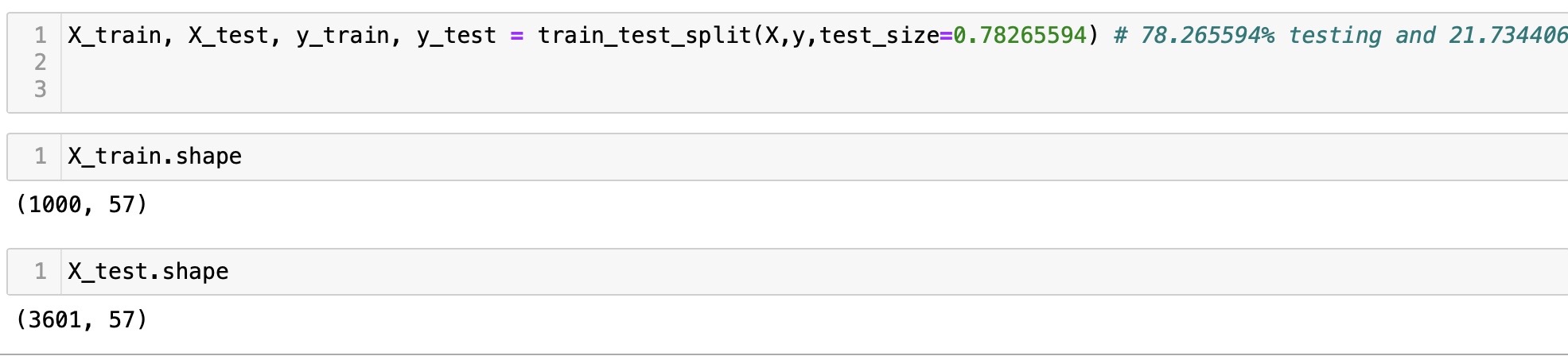
2.Compare the accuracies of the fused model with Random Forest (with 1000 base learners).

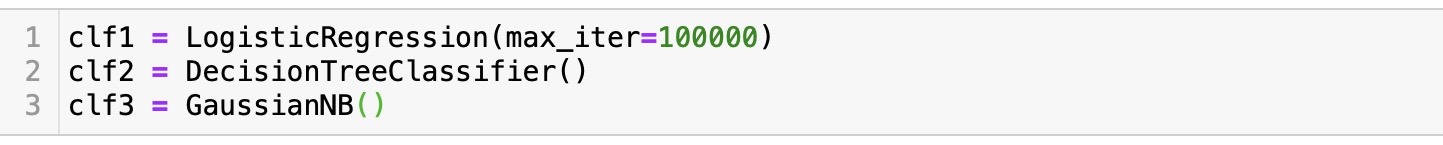
**Accuracies of the Fused model and Random Forest .**

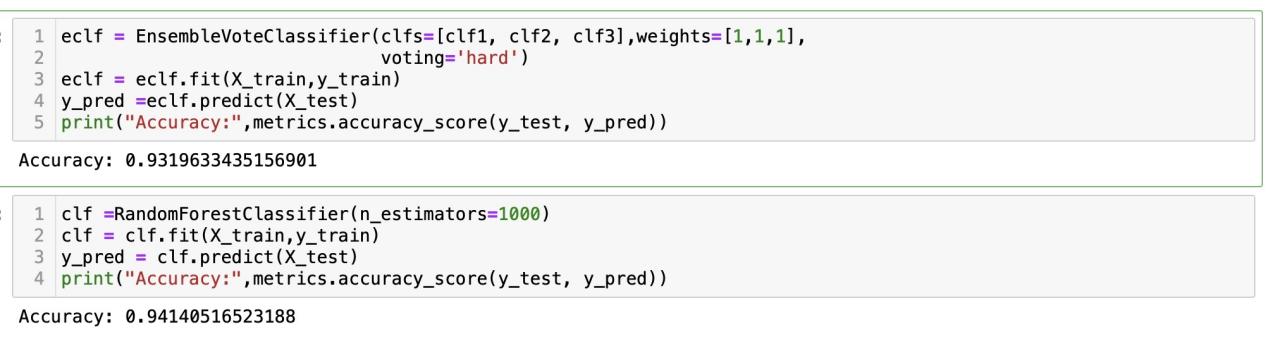
We got an accuracy score of 93.19 percent for Fused model and 94.14 for Random Forest Classifier.

1000 instances for training.

3601 instances for testing.



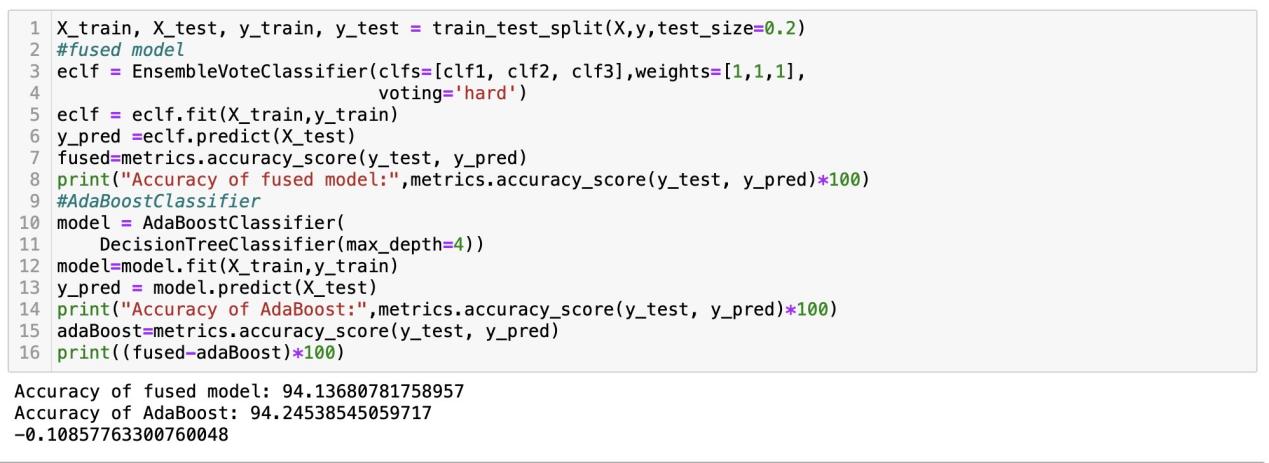
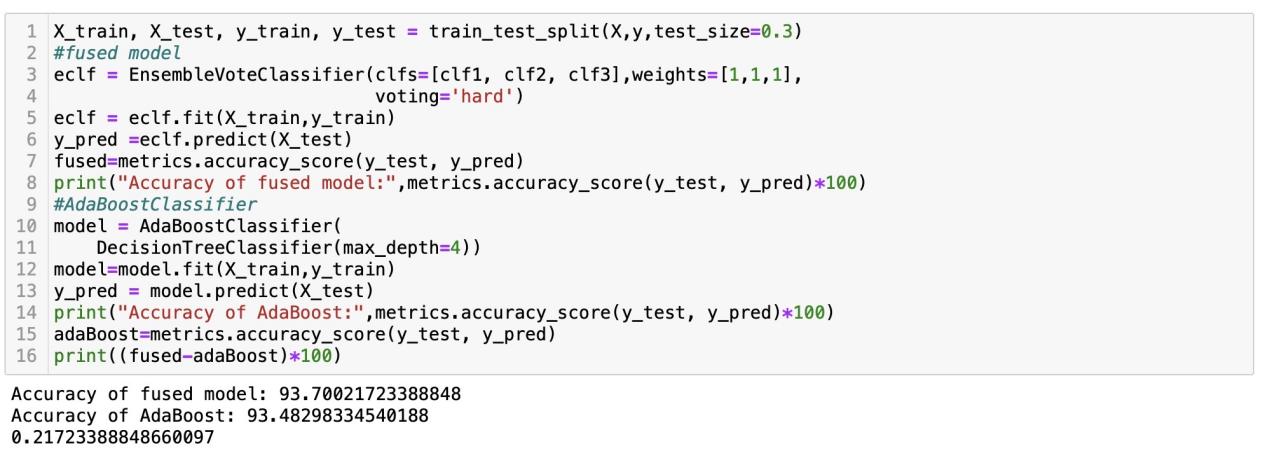
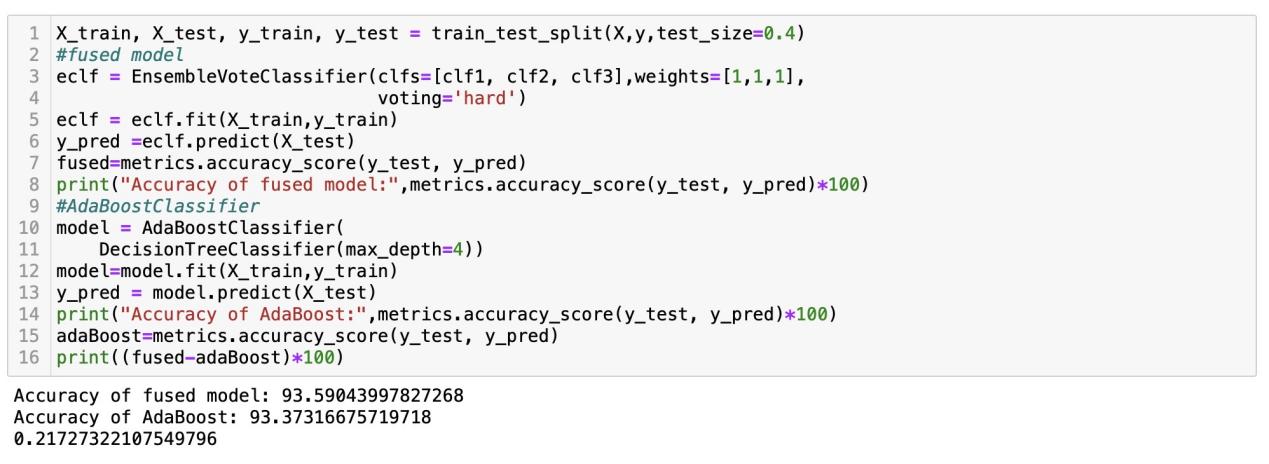
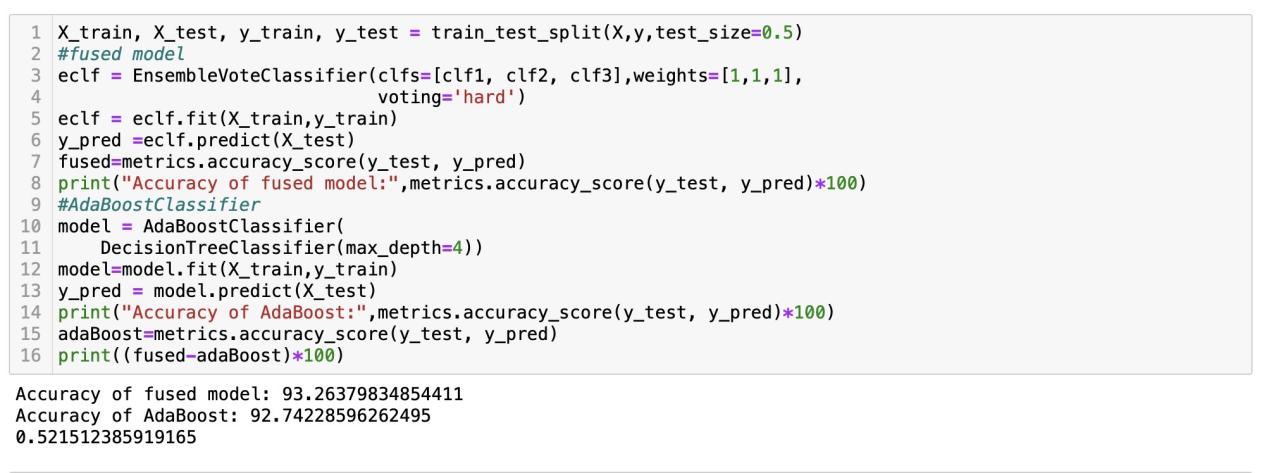


****

**Observations :**

As we can clearly see the accuracy of the RandomForestClassifier is greater than Fused model by 0.9441

1. Compare their accuracies of fused classifier and the AdaBoost Ensemble with Decision Tree as the base learner with the following training-test splits: 50%-50%, 60%-40%, 70%-30%, and 80%-20%.



**Observations :**

we can clearly see the difference between accuracies of the fused model and AdaBosstClassifier is being decreased with increase in training data.

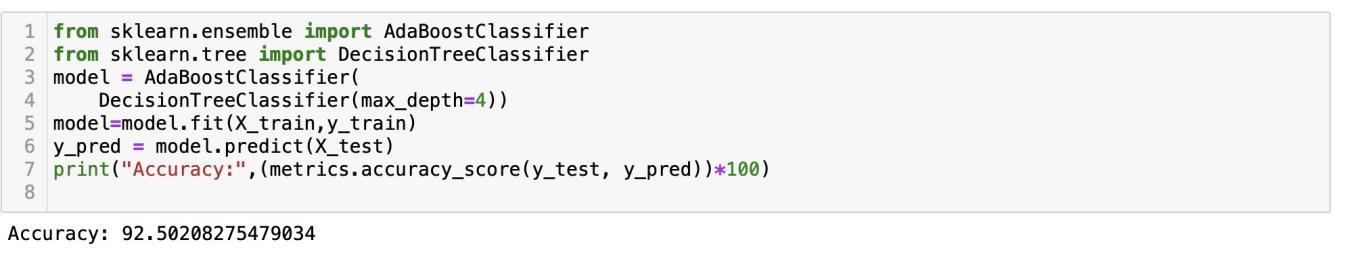
At first we considered the training split of 50% , difference of accuracies of the models are 0.52151.

But later on increasing the training split ,difference between accuracies of the models are decreased.At 80% of training split, AdaBoostClassifier has greater accuracy than Fused model.

**Classification accuracy:**

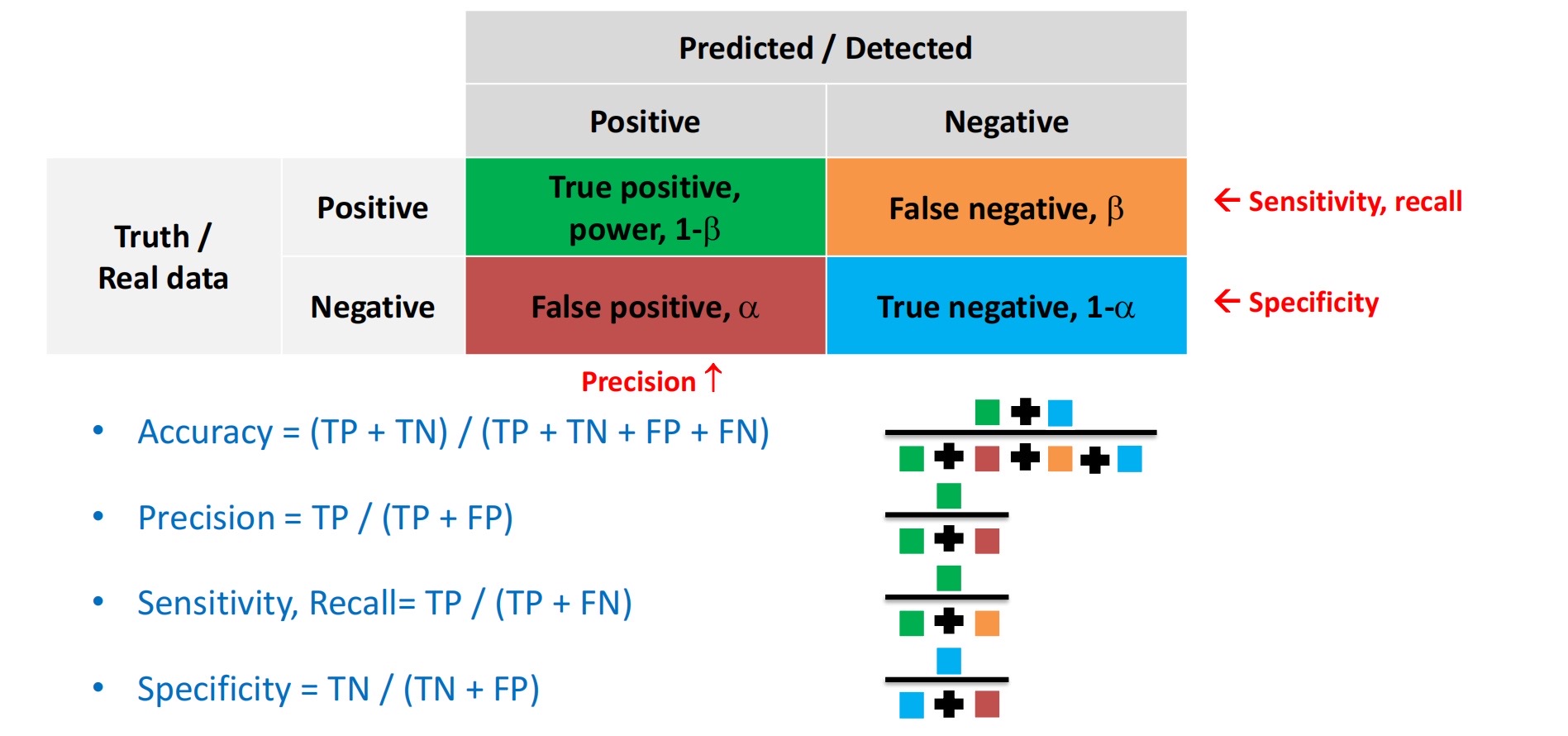
Accuracy is one metric for evaluating classification models. Informally, accuracy is the fraction of predictions our model got right. Formally, accuracy has the following definition: Accuracy = **Number of correct predictions / Total number of predictions** made,multiplied by 100 to turn it into a percentage.

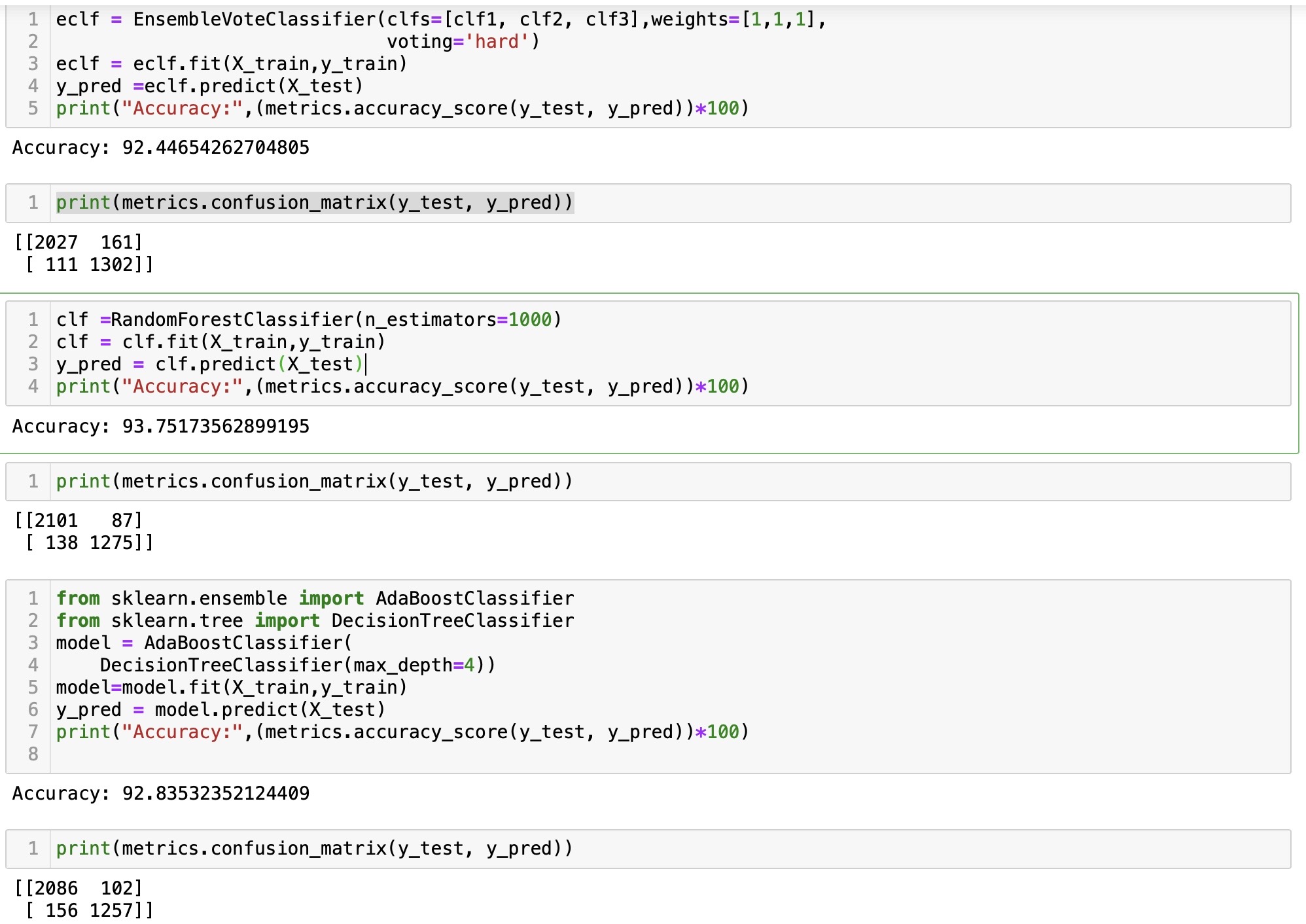


****

**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.





**Comparison**:

By comparing with all the models, Random Forest Classifier has highest accuracy with 94%.