BIG DATA ANALYTICS AND APPLICATIONS

PROJECT

**CROP YIELD PREDICTION USING MACHINE LEARNING**

INTRODUCTION:

* Agriculture is that the backbone of each economy.
* Since the invention of recent innovative technologies and techniques the agriculture field is slowly degrading.
* Owing the cultivating techniques and thus the seasonal climatic conditions also are being modified against the element’s assets like soil, water, temperature etc.
* Crop yield prediction is very important agriculture problem.

ABSTRACT:

* Crop yield prediction involves predicting yield of the crop from obtainable historical information like weather parameter, soil parameter and historic crop yield.
* This system target predicting the yield of the crop supported the present information by mistreatment random forest algorithm.
* The prediction can help to the farmer to predict the yield of the crop before cultivating onto the agricultural field.

PROPOSED SYSTEM:

* This system mainly concentrates on weather forecasting, crop yield prediction and crop cost fore casting. These factors help the farmers to cultivate the best food crops.
* Farmers can adapt to climate changes to some degree by shifting planting dates. Choosing varieties with different growth duration, or changing crop rotations.
* The Statistical numeric data related to agriculture is undertaken.

DATASET DESCRIPTION:

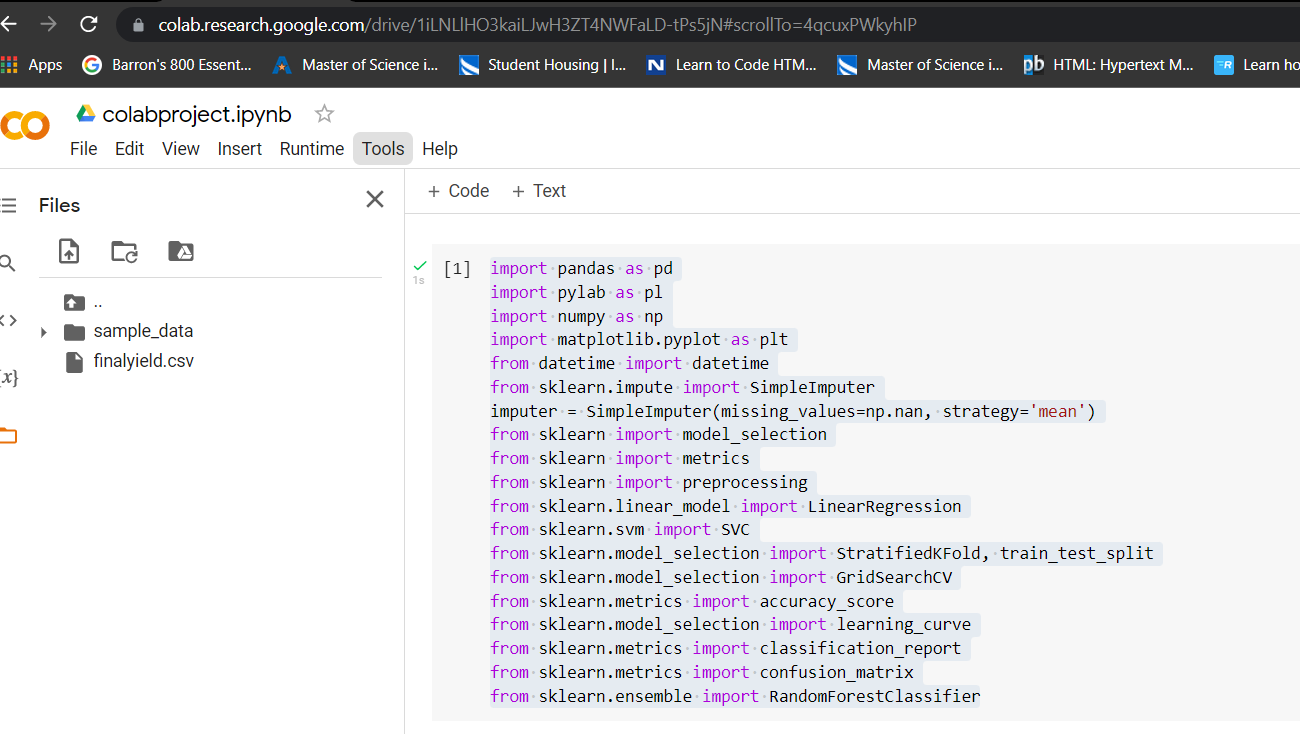
* The dataset contains over 3977 data input variables in which 3181 data inputs are for training and 796 for testing.
* The training data that is 3977 data input variables are classified into 11 classes.
* By using this dataset, we are going to build a model which will classify the crop yield.

METHODOLOGIES:

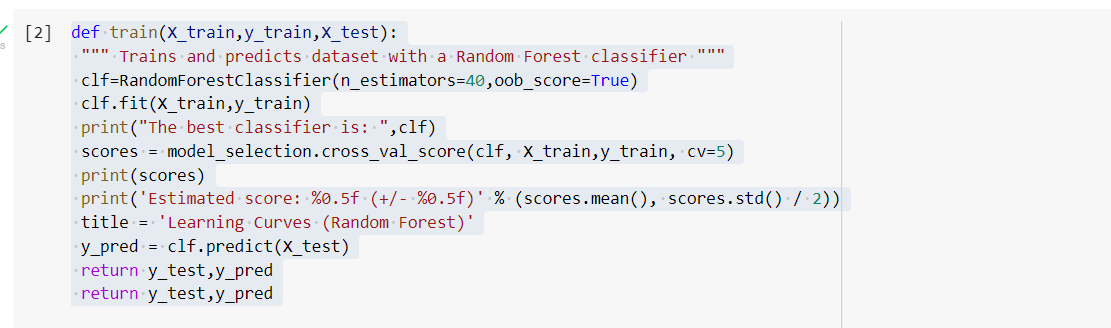
* Support Vector Machine
* Logistic Regression
* Random Forest Classification

CODING PART:

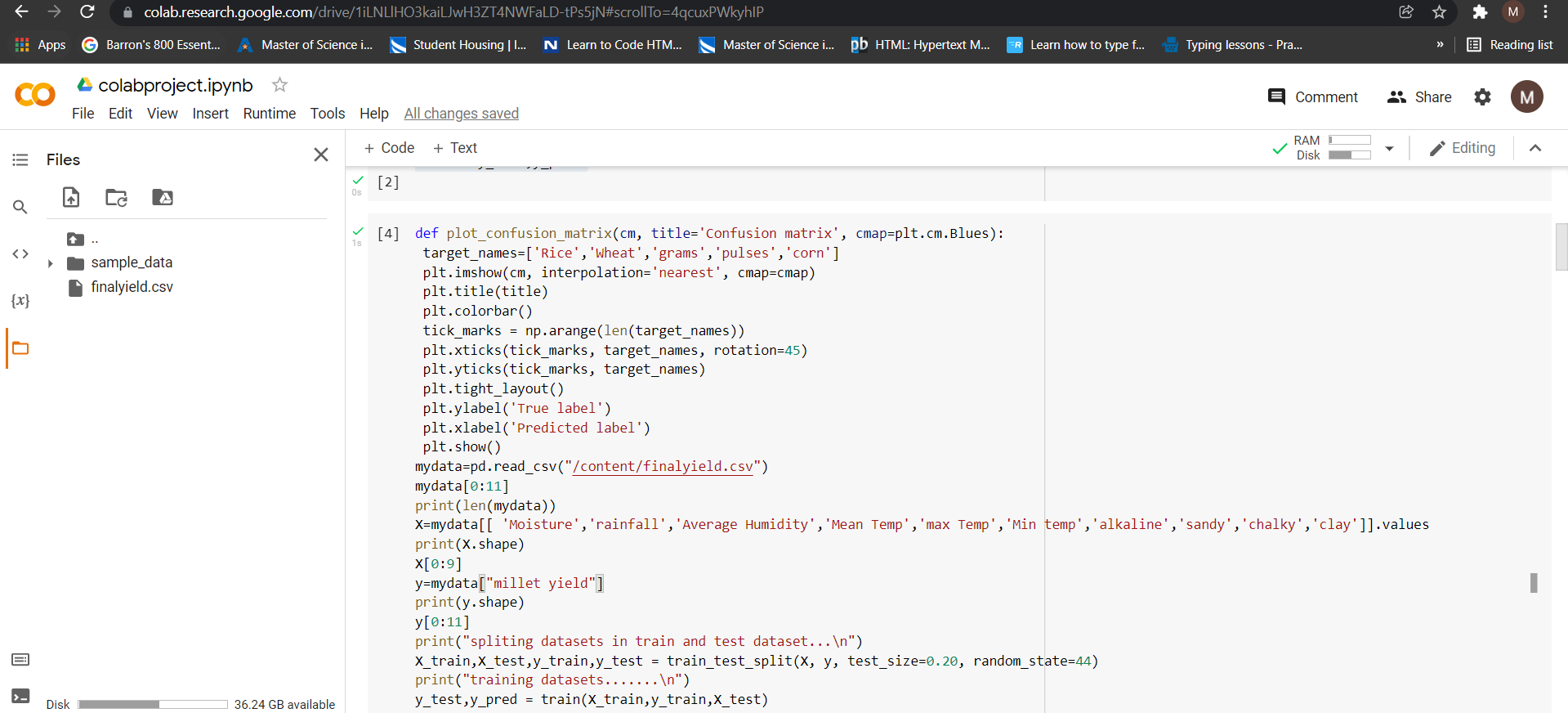
* First, I uploaded the dataset finalyield.csv file and imported required libraries and installed packages.

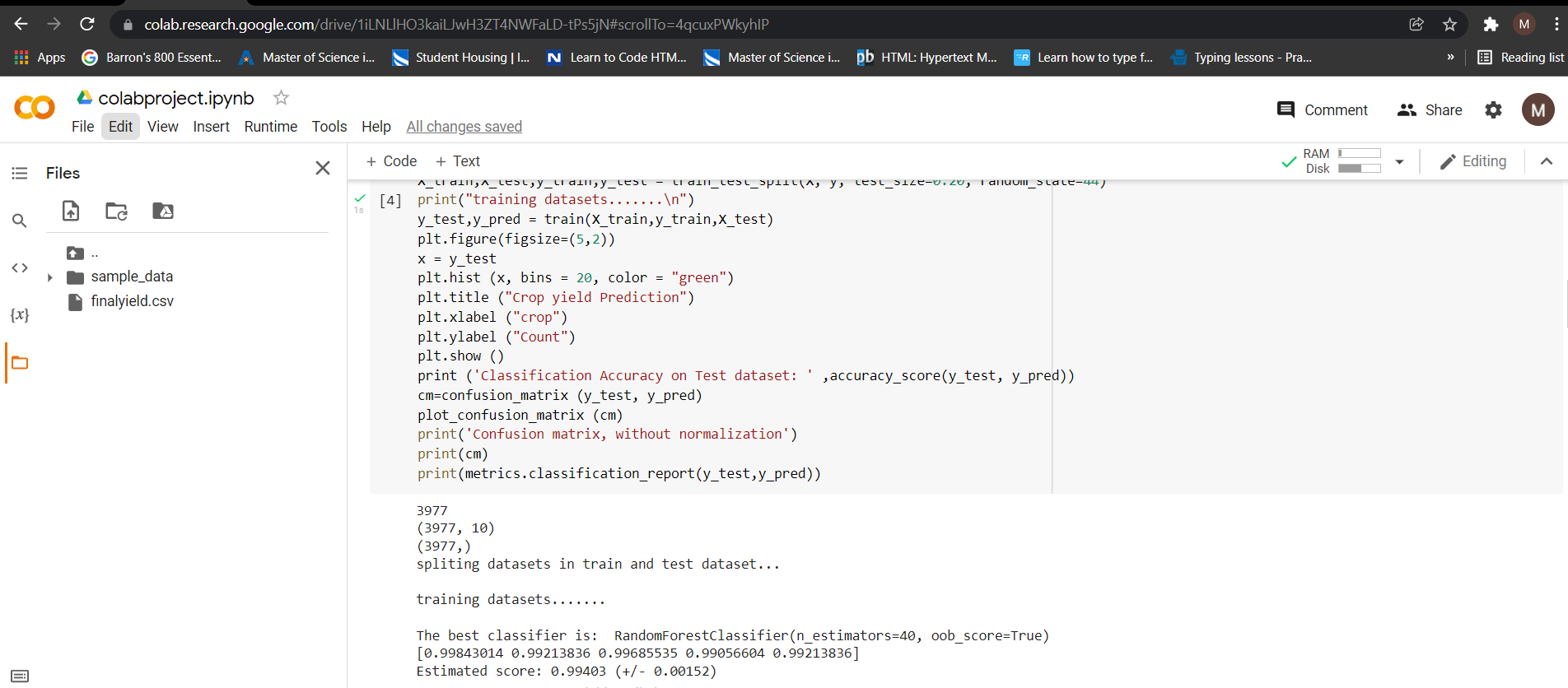


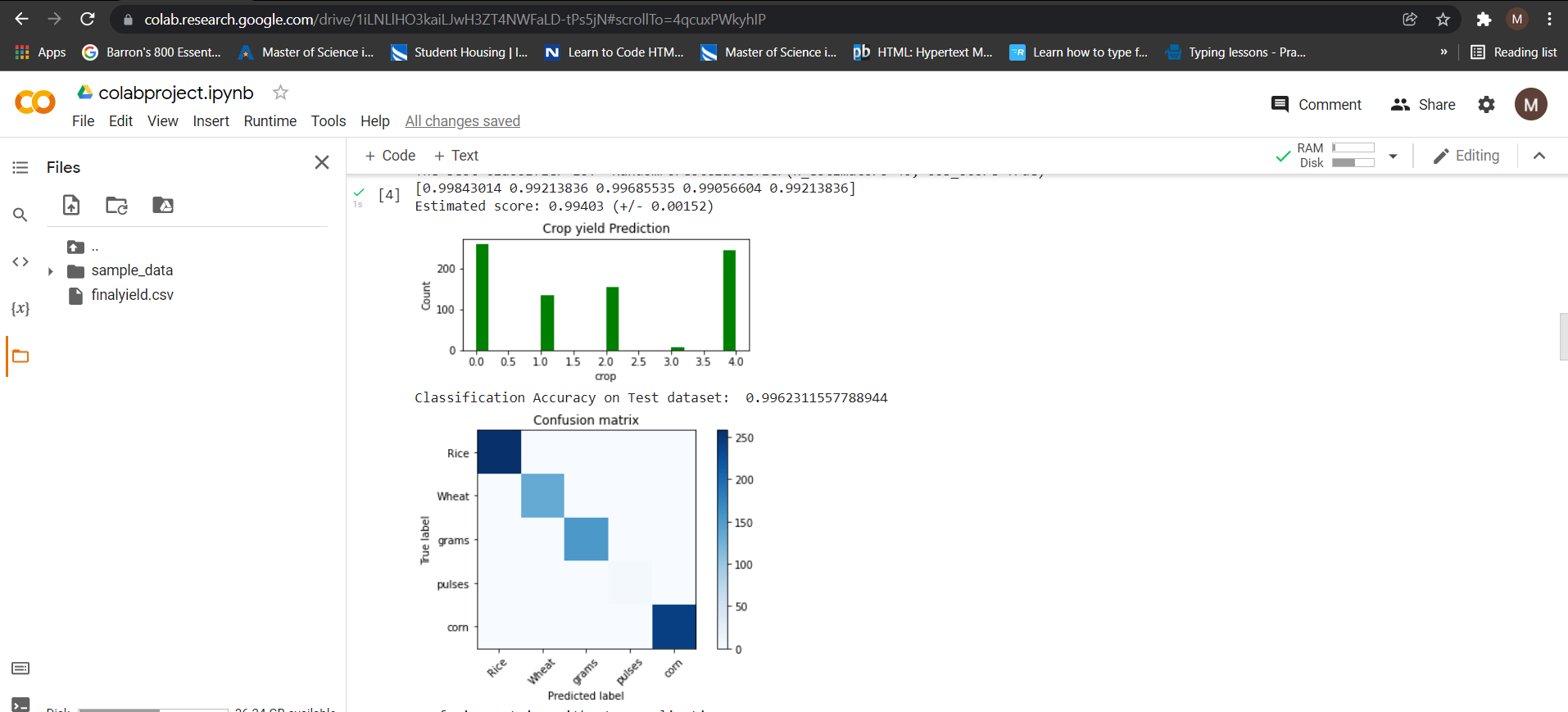
* Using Random Forest Classifier, we train the dataset and predict the dataset.

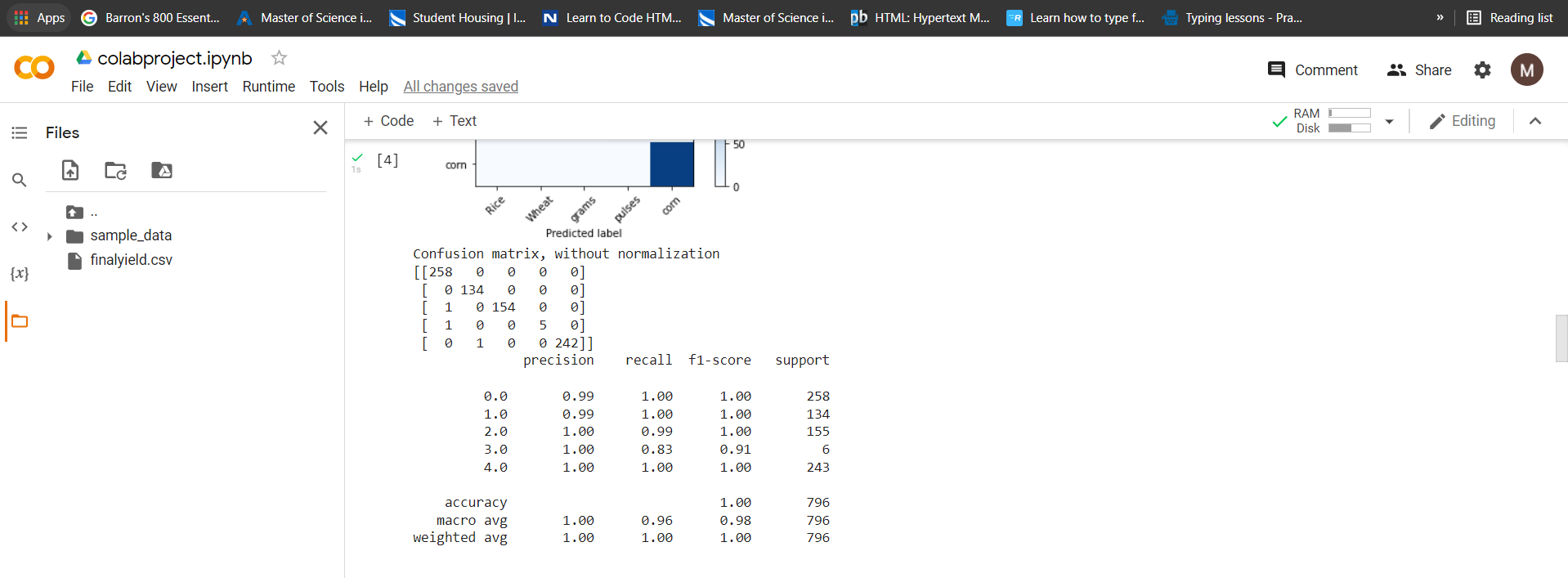


* Confusion Matrix for the given dataset with the target classes:



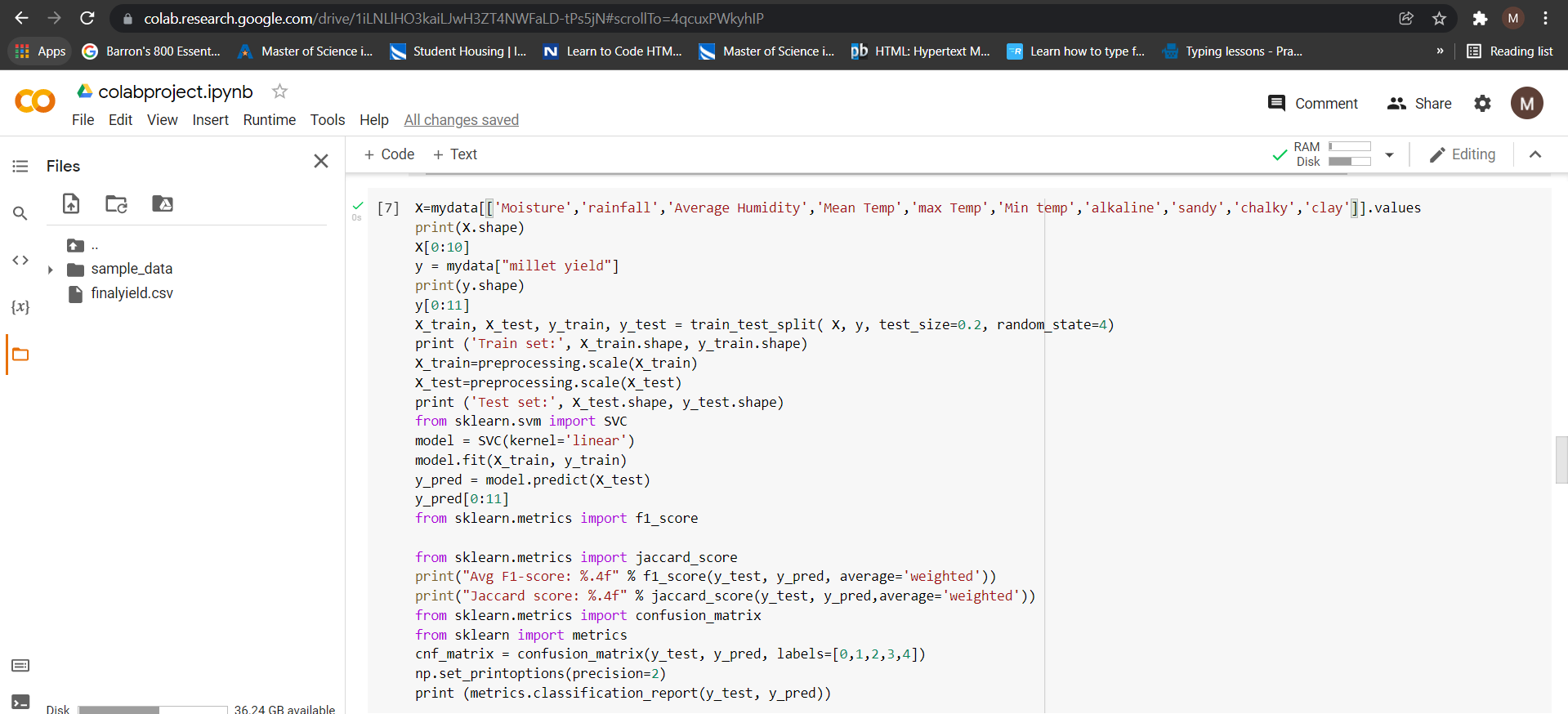


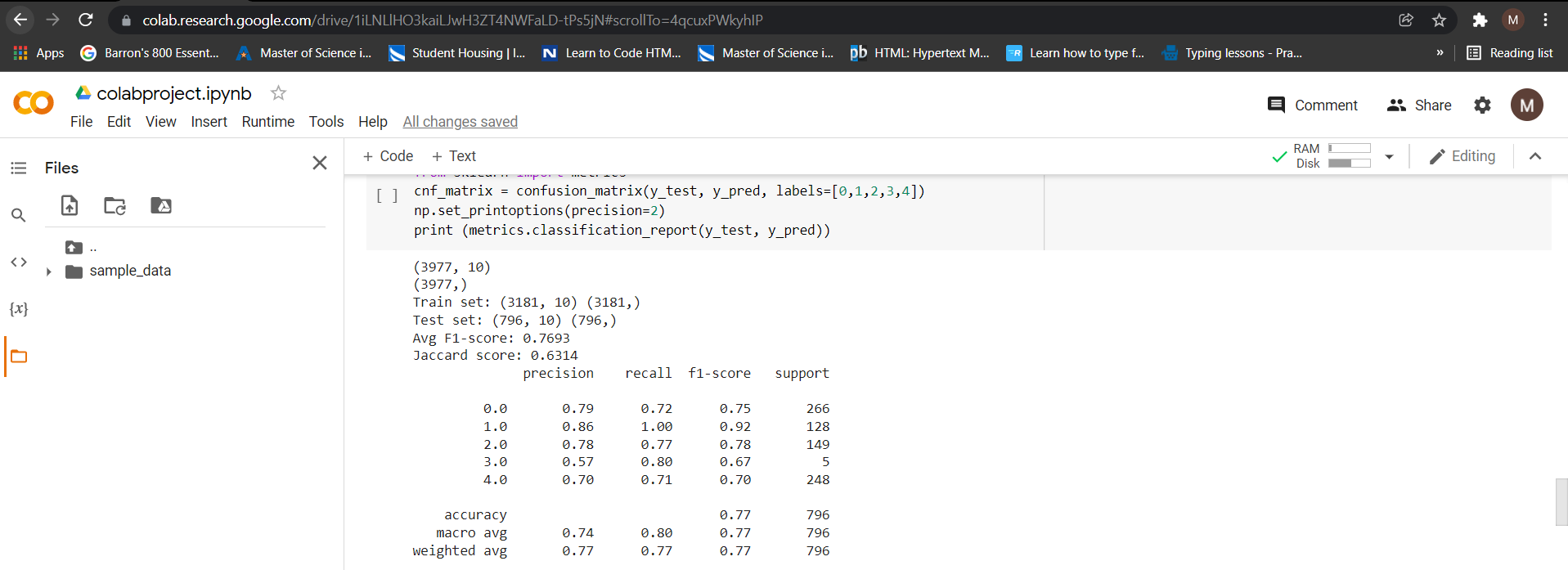




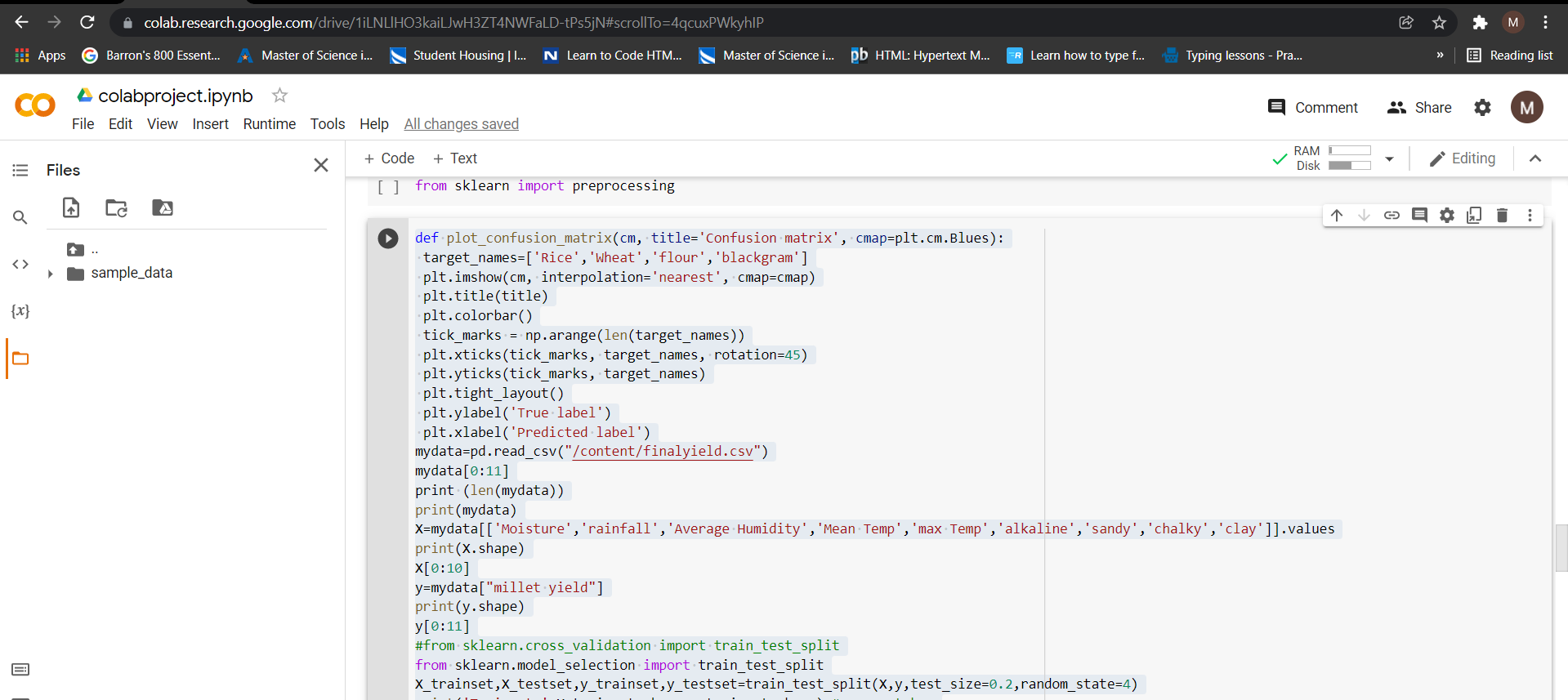
So, we got the accuracy 99% for the Random Forest classification for the given number of target classes.

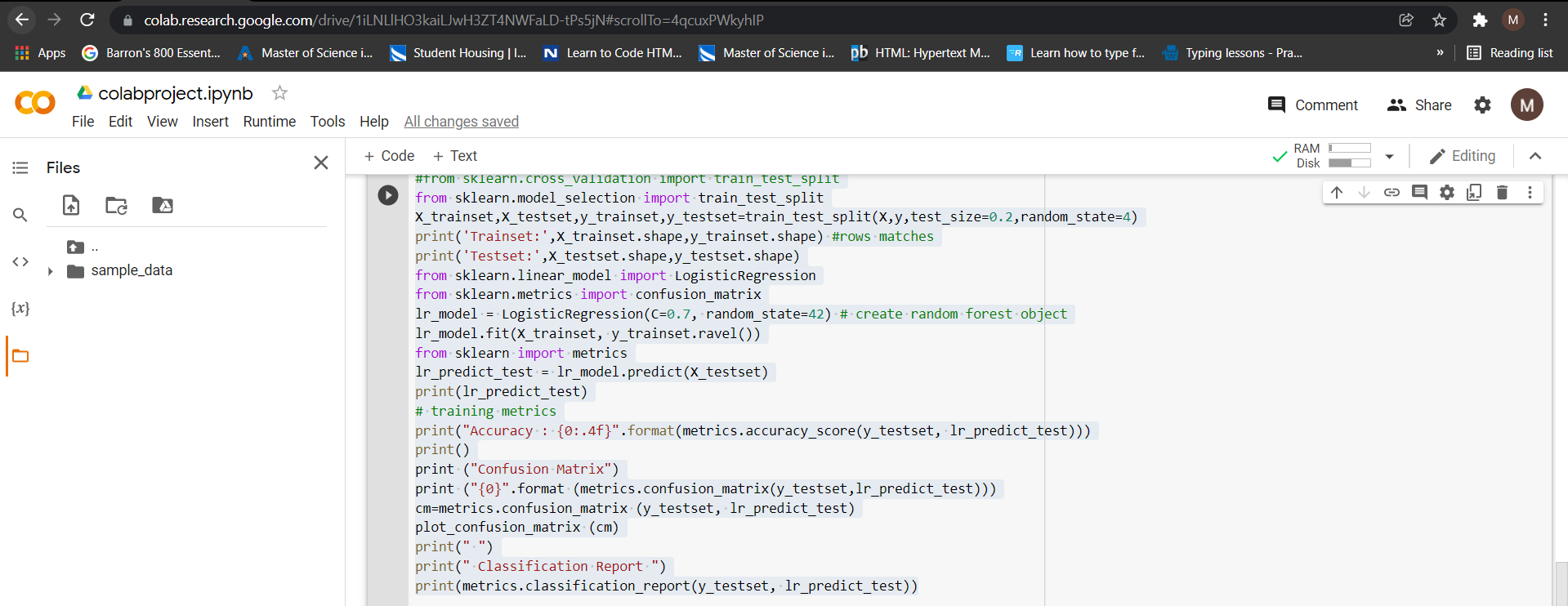
* Finding the accuracy of the given classification model using the support vector machine.

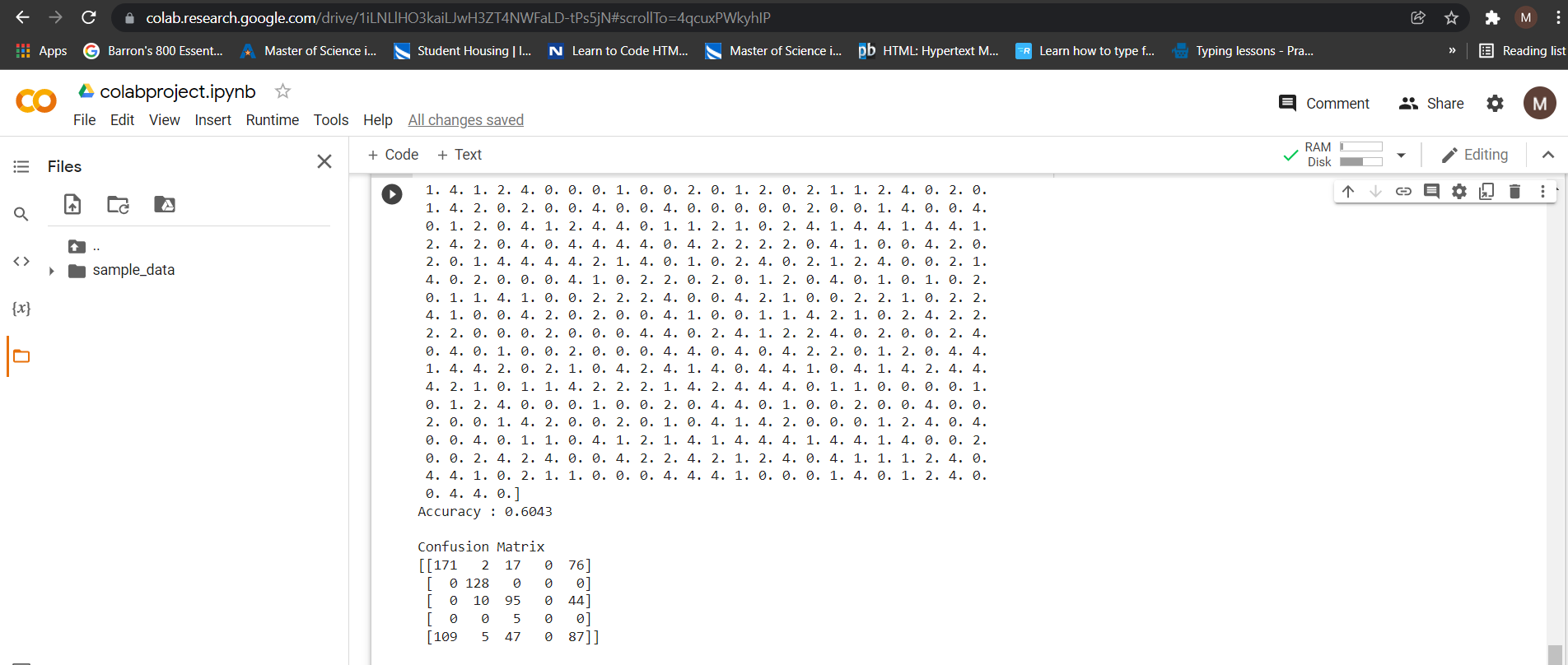


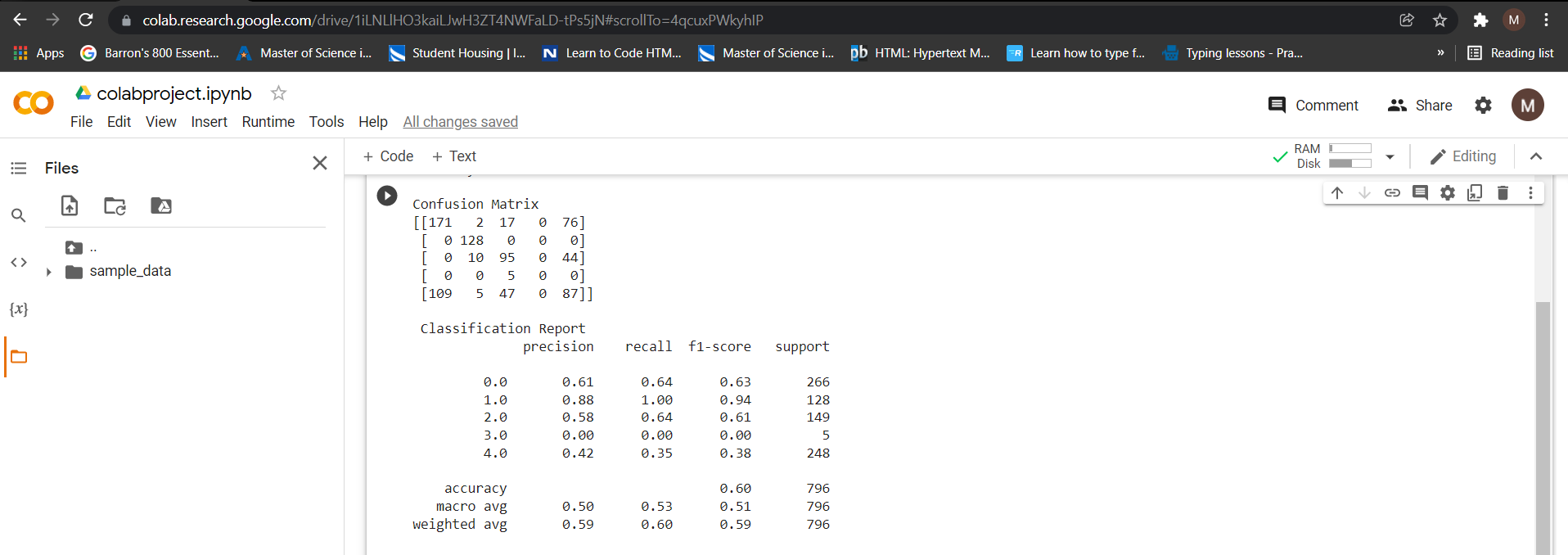


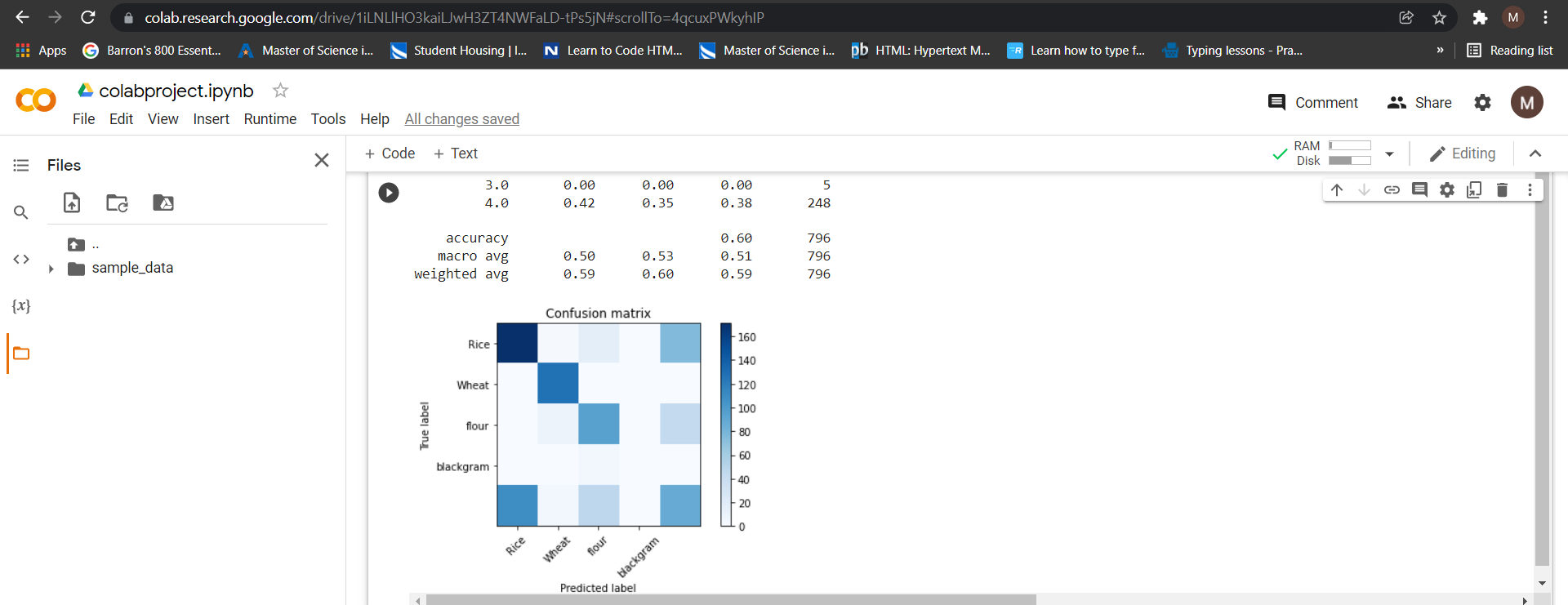
* Finding the accuracy of the given classification model using the logistic regression model.











RESULTS:

* Therefore, we can attain an accurate crop yield prediction using the Random Forest algorithm.
* Random Forest algorithm achieves a largest number of crop yield models with a lowest model.
* It is suitable for massive crop yield prediction in agricultural planning. This makes the farmers to take the right decision for right crop such that the agricultural sector will be developed by innovative ideas.