# Machine Learning November Minor Project

## By-MANUPRIYA NAGAR

INTRODUCTION:

The price of a product is the most important attribute of marketing that product. One of those products where price matters a lot is a smartphone because it comes with a lot of features so that a company thinks a lot about how to price this mobile which can justify the features and also cover the marketing and manufacturing costs of the mobile.

Problem statement :

create a classification model to predict whether price range of mobile based on certain specification

Context:

An entrepreneur has started his own mobile company. He wants to give tough fight to big companies like Apple, Samsung etc. He does not know how to estimate price of mobiles his company creates. In this competitive mobile phone market, one cannot simply assume things. To solve this problem, he collects sales data of mobile phones of various companies. He wants to find out some relation between features of a mobile phone (e.g., RAM, Internal Memory etc) and its selling price. But he is not so good at Machine Learning. So, he needs your help to solve this problem. In this problem you do not have to predict actual price but a price range indicating how high the price is.

content

1. Data set
2. Remove handles
3. Apply the following models of dataset
4. Logistic Regression
5. KNN Classification
6. SVM Classifier with linear and rbf kernel
7. Decision Tree Classifier
8. Random Forest Classifier

Solution :-

**Used Libraries**

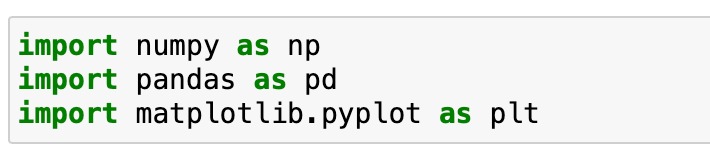
* **Panda**: In computer programming, pandas is a software library written for the Python programming language for data manipulation and analysis. In particular, it offers data structures and operations for manipulating numerical tables and time series. Pandas is mainly used for data analysis. Pandas allows importing data from various file formats such as comma-separated values, JSON, SQL, Microsoft Excel. Pandas allows various data manipulation operations such as merging, reshaping, selecting, as well as data cleaning, and data wrangling features.
* **Numpy**: Numpy is a library for the Python programming language, adding support for large, multi-dimensional arrays and matrices, along with a large collection of high-level mathematical functions to operate on these arrays.
* **Seaborn**: Seaborn is a Python data visualization library based on matplotlib. It provides a high-level interface for drawing attractive and informative statistical graphics.

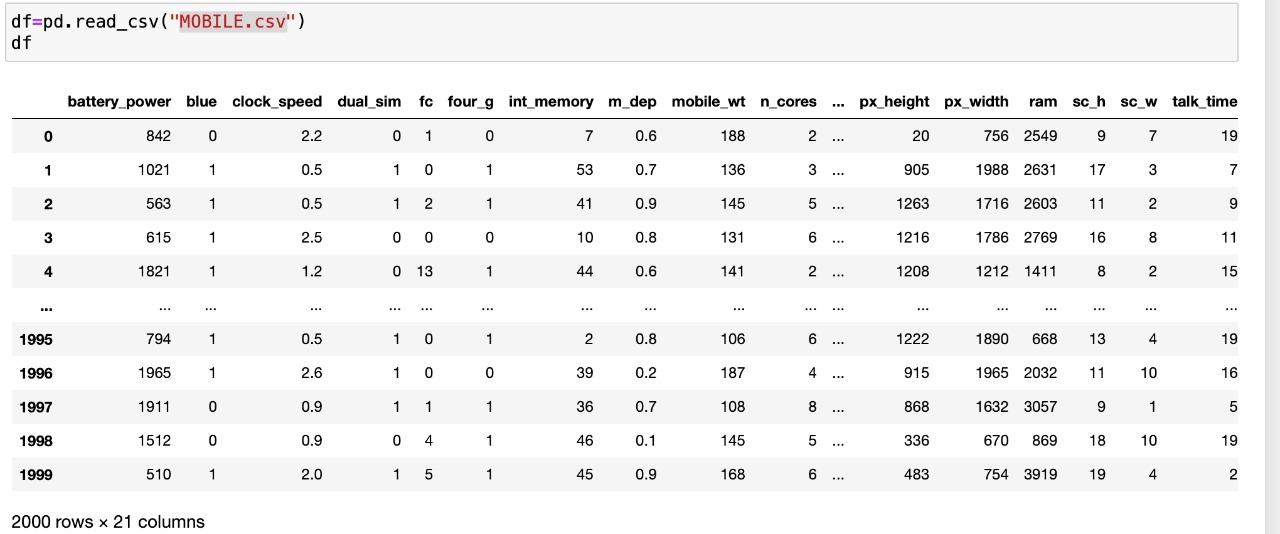
Data set:

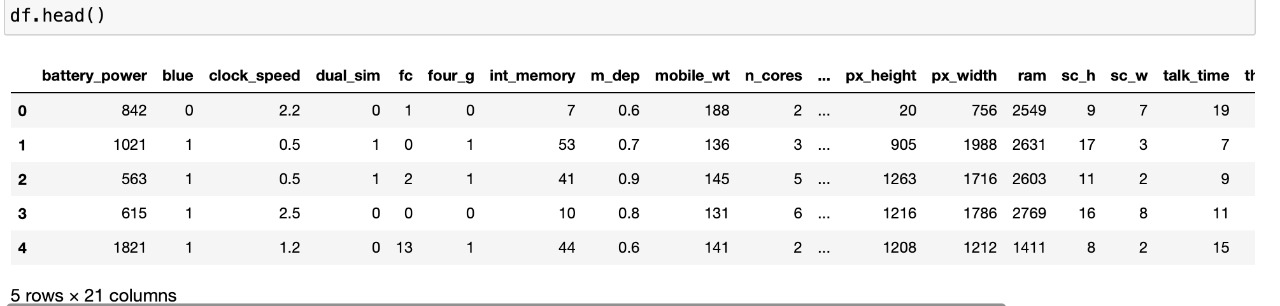
* **id**:ID
* **battery\_power**:Totalenergyabatterycanstoreinonetime   
  measured in MAh
* **blue**:Hasbluetoothornot
* **clock\_speed**:speedatwhichmicroprocessorexecutes instructions
* **dual\_sim**:Hasdualsimsupportornot
* **fc**:FrontCameramegapixels
* **four\_**g:Has4Gornot
* **int\_memory**:InternalMemoryinGigabytes
* **m\_dep**:MobileDepthincm
* **mobile\_wt**:Weightofmobilephone
* **n\_cores**:Numberofcoresofprocessor
* **pc**:PrimaryCameramegapixels
* **px\_height**:PixelResolutionHeight
* **px\_width**:PixelResolutionWidth
* **ram**:RandomAccessMemoryinMegabytes
* **sc\_h**:ScreenHeightofmobileincm
* **talk\_time**:longesttimethatasinglebatterychargewilllast when you are
* **three\_g**:Has3Gornot
* **touch\_screen**:Hastouchscreenornot **wif**:HasWi-Fiornot

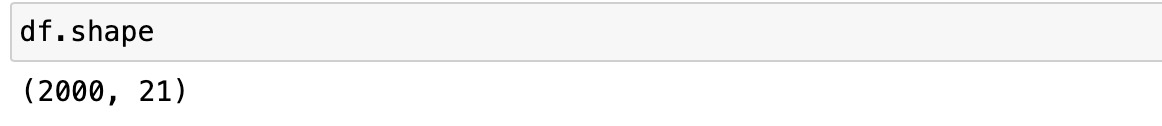
**Use: -**

* This kind of prediction will help companies estimate price of mobiles to give tough competition to other mobile manufacturer
* Also, it will be useful for Consumers to verify that they are paying best price for a mobile.

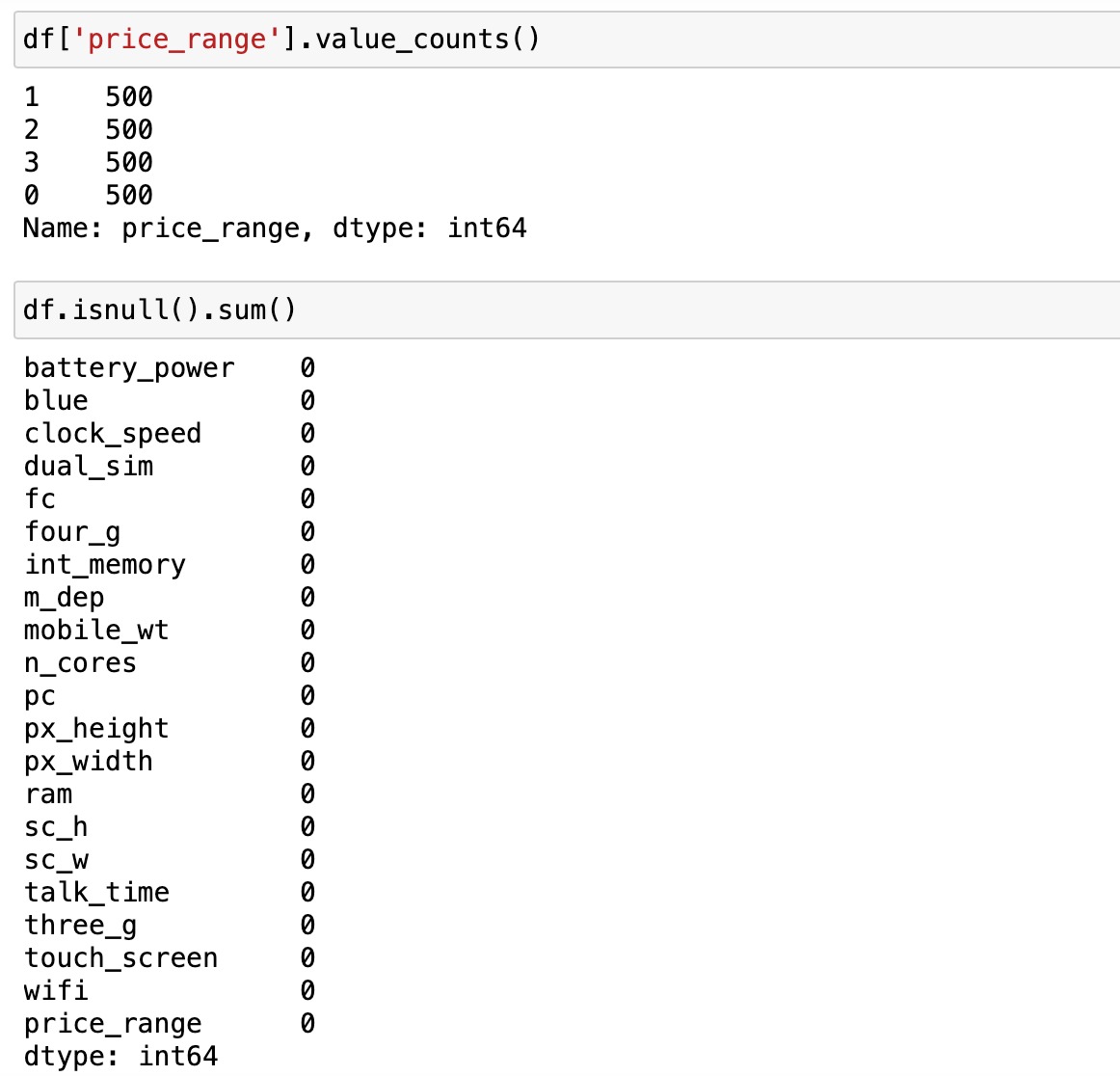




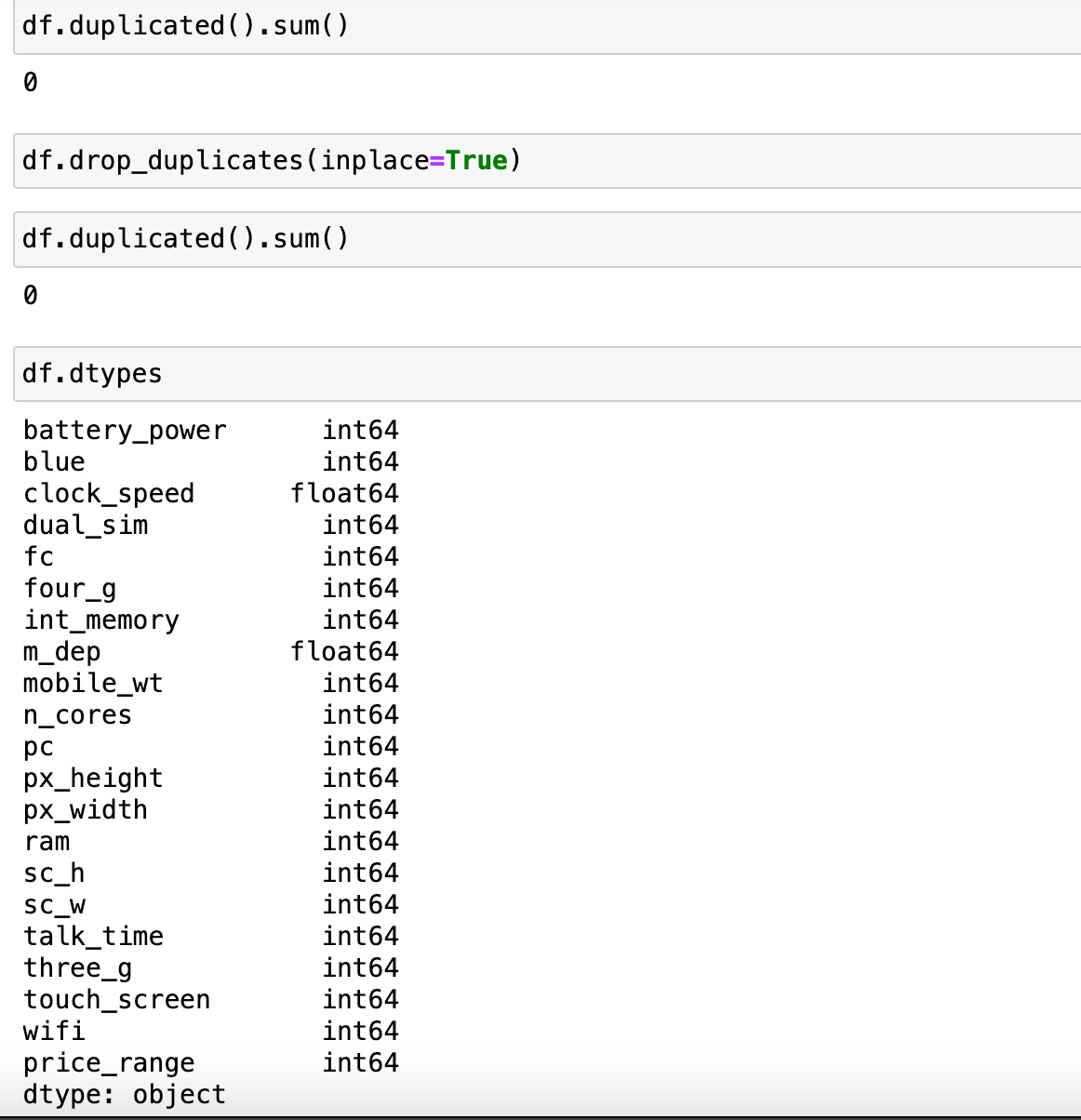




# TARGET VARIABLE & REMOVE HANDLE NULL VALUES(IF ANY)



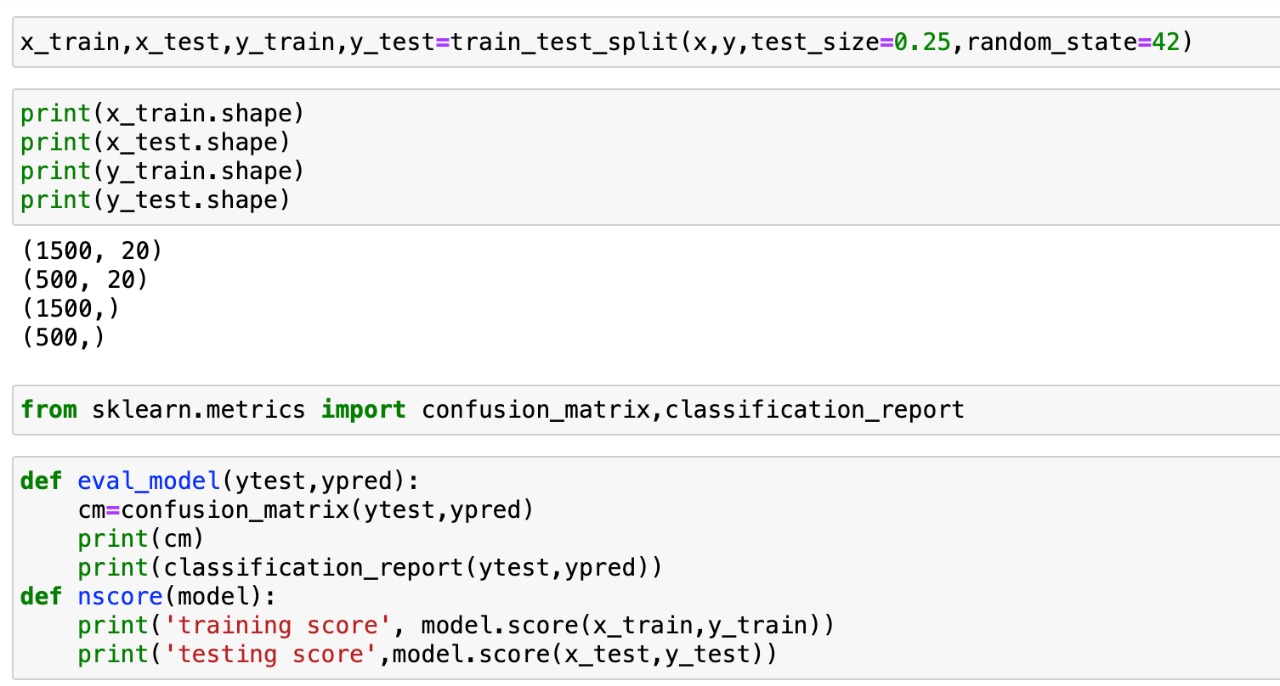
HANDLING DUPLICATES & CHECKING DATATYPES



## Selecting dependent(x) and independent(y) variables & spliting data into training and test data



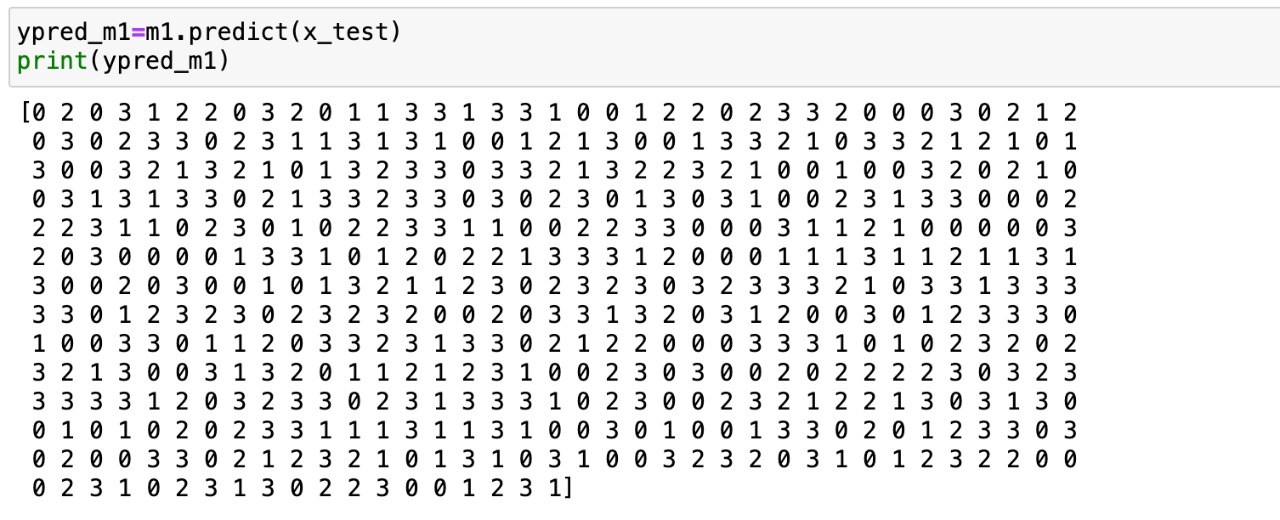
CONFUSION MATRIX

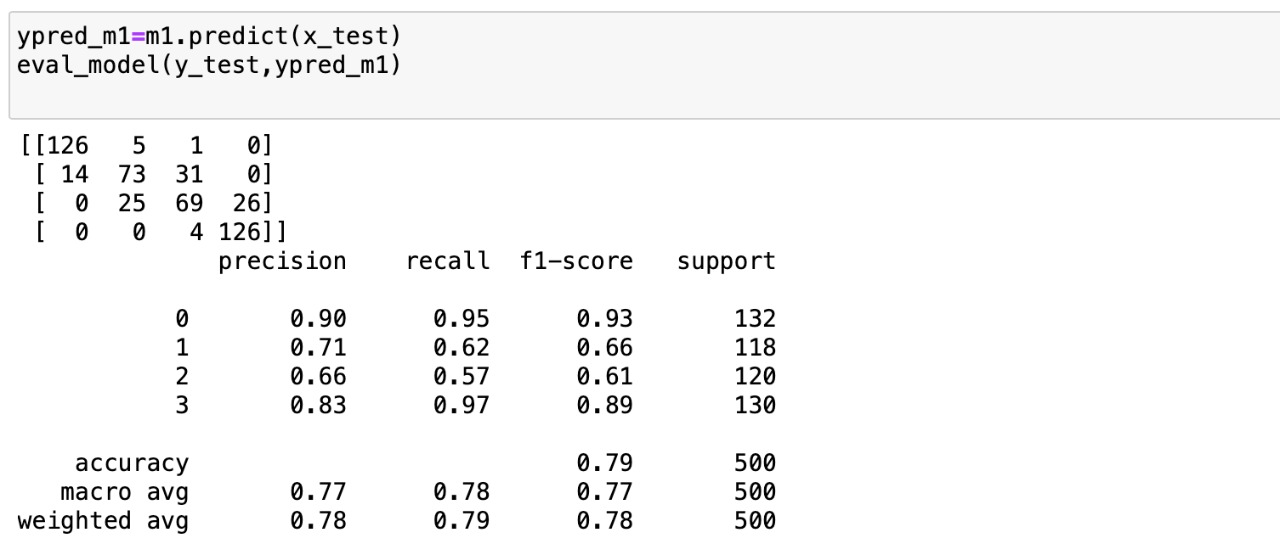


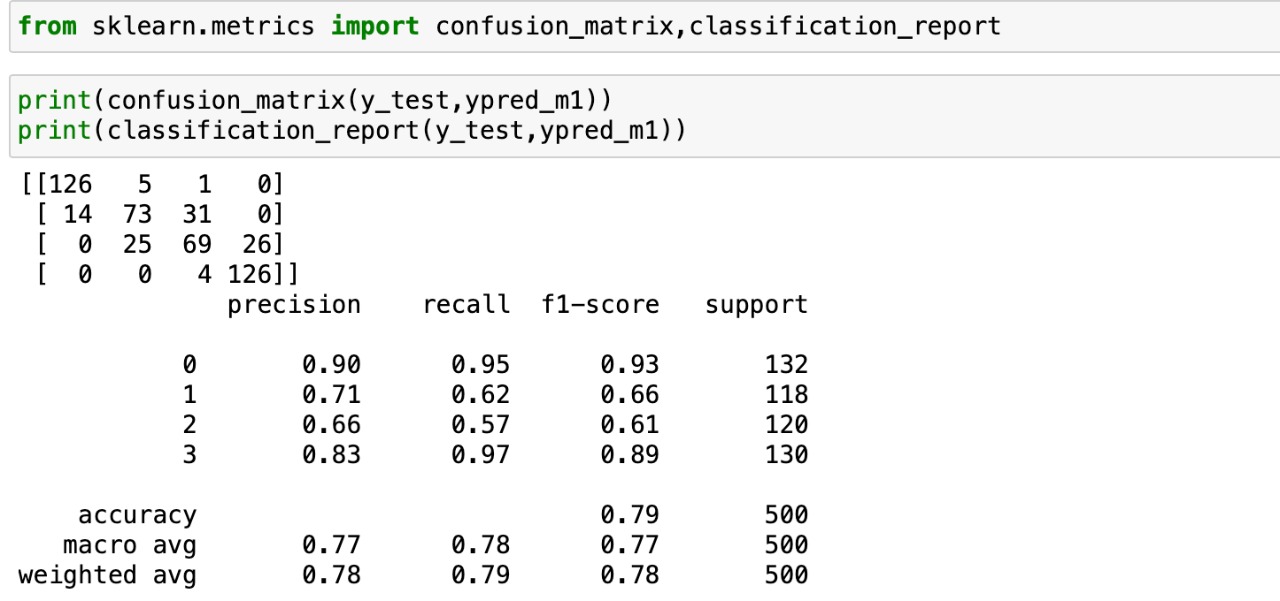
## Apply the following models on the training dataset and generate the predicted value for the test dataset

LOGISTIC REGRESSION



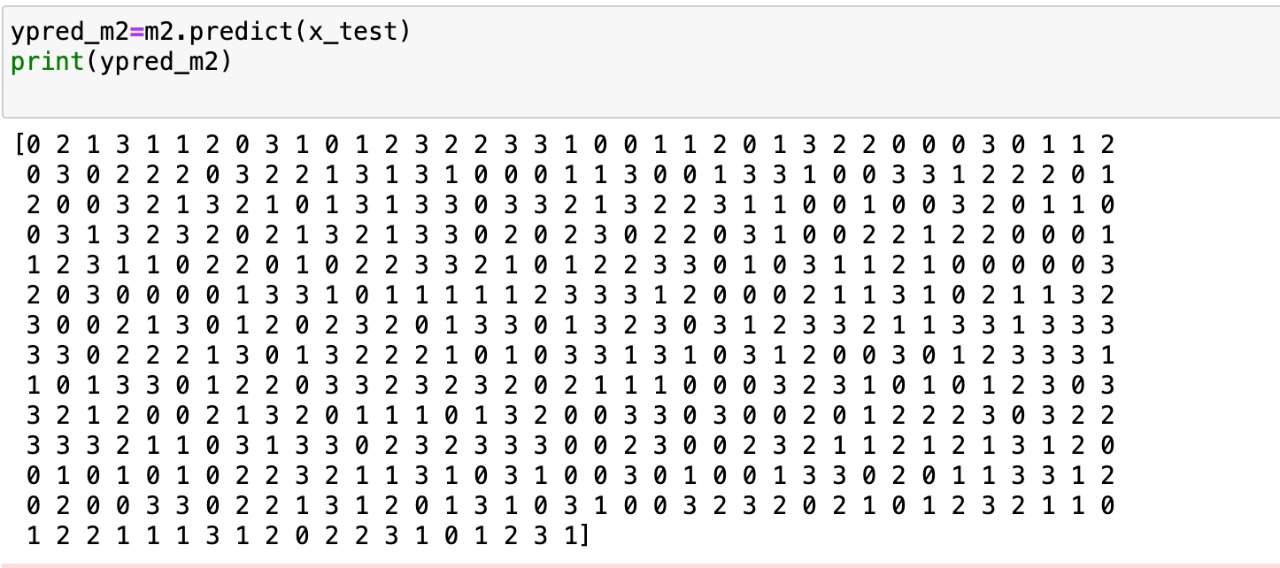




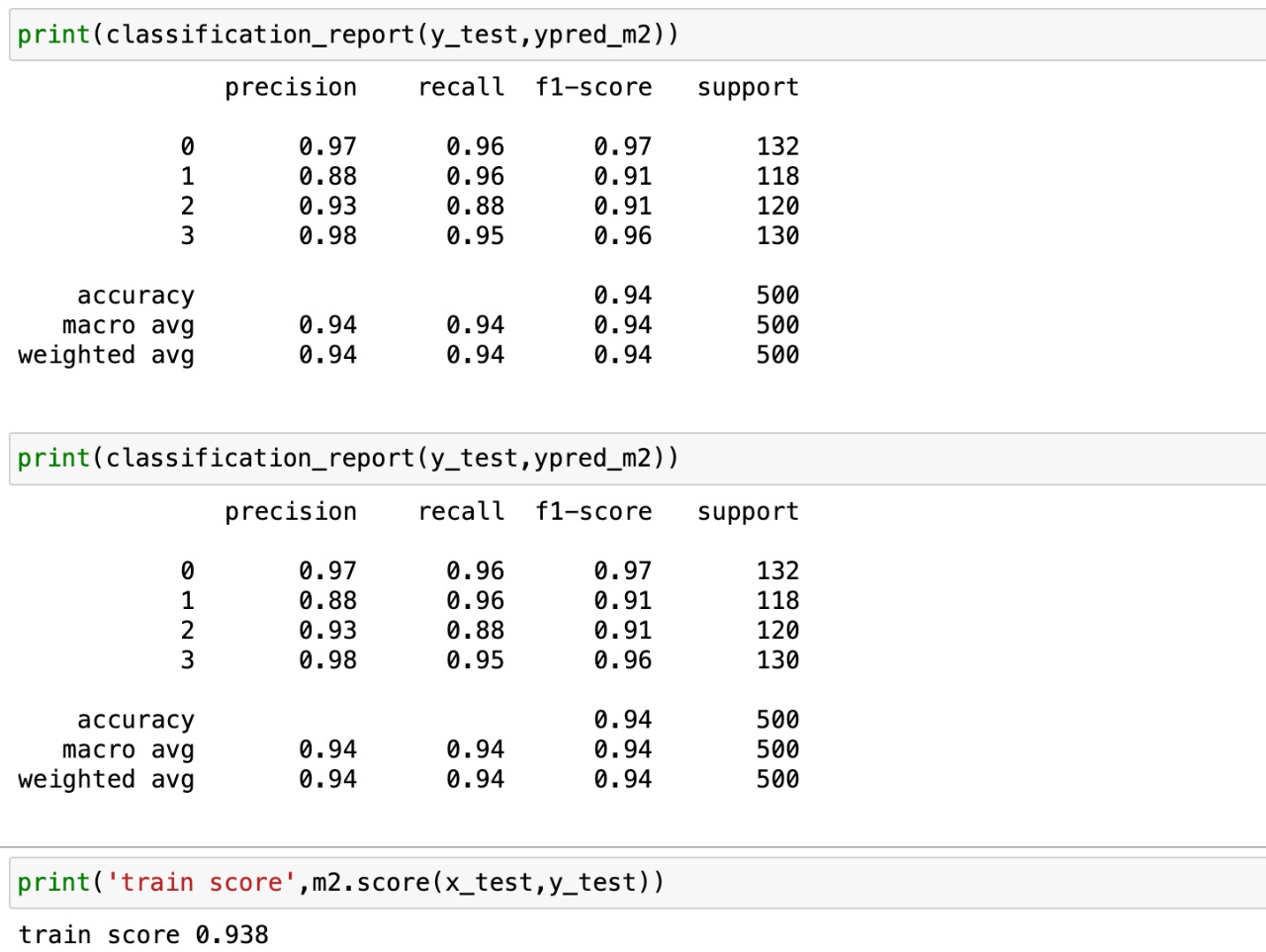


KNN classification

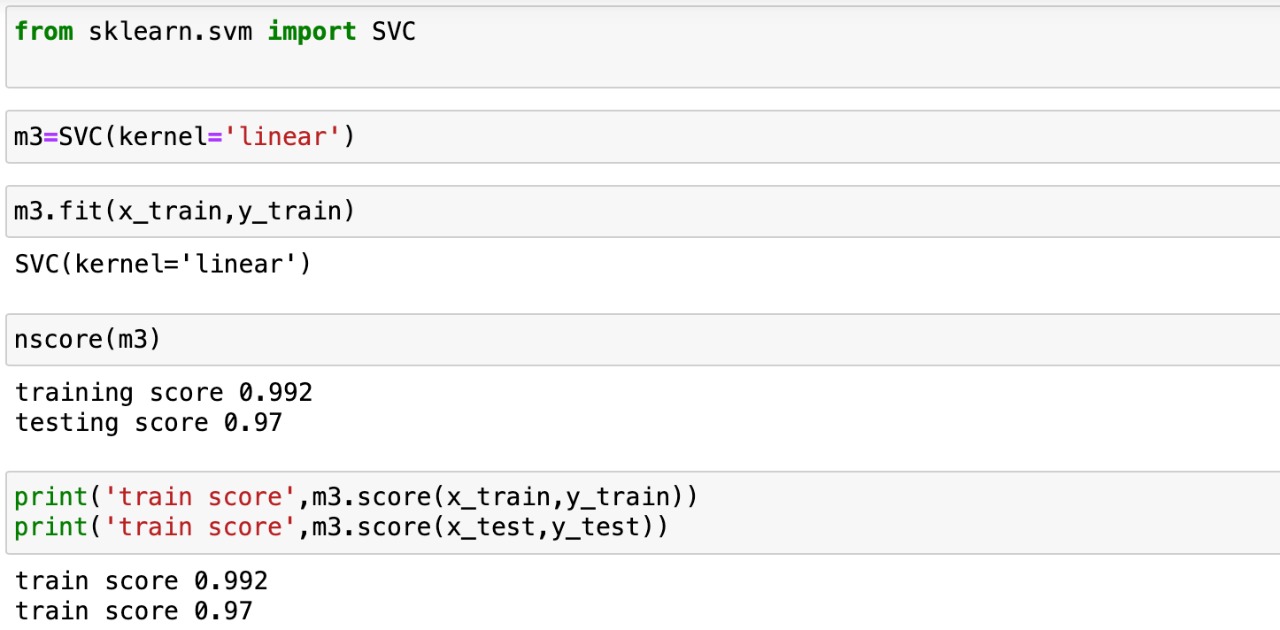


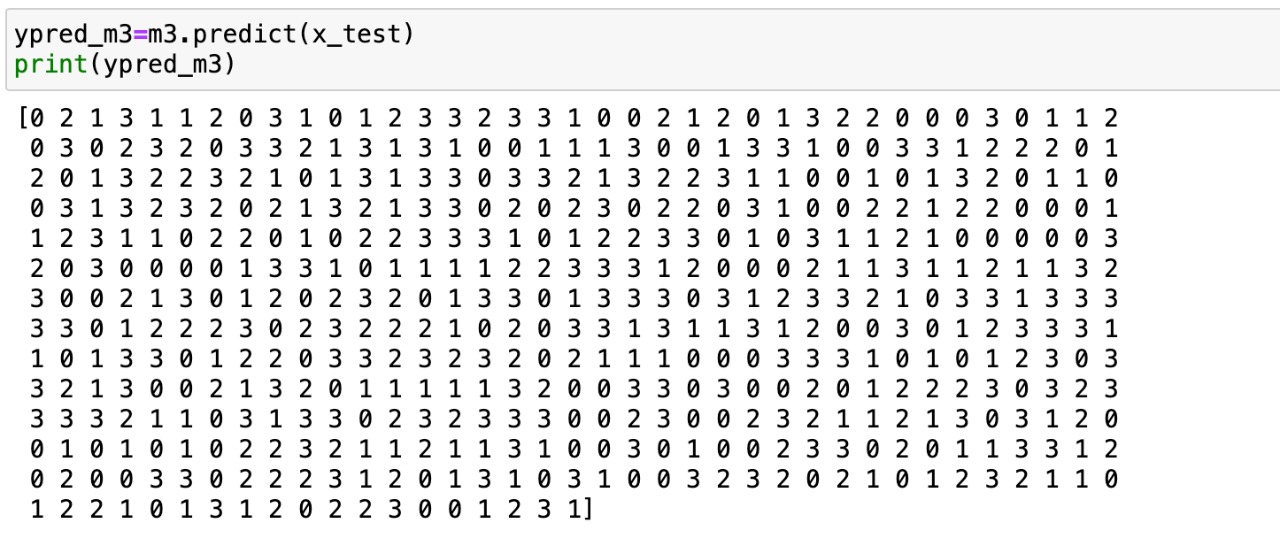


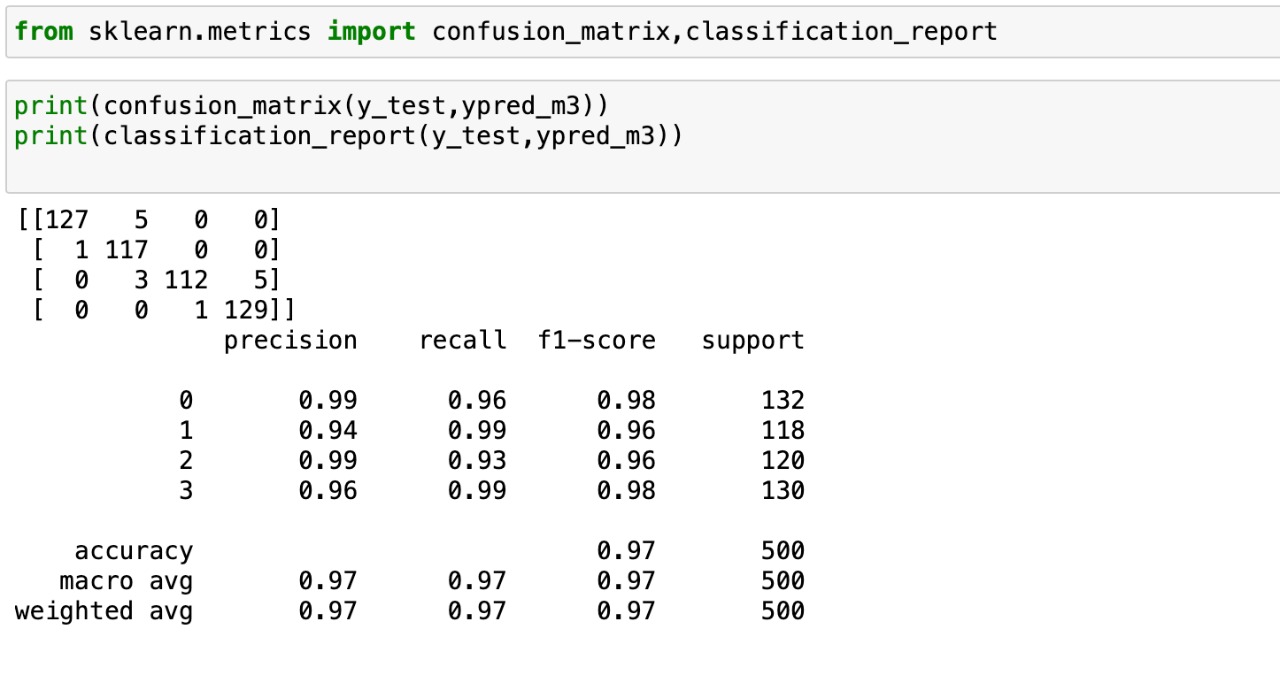


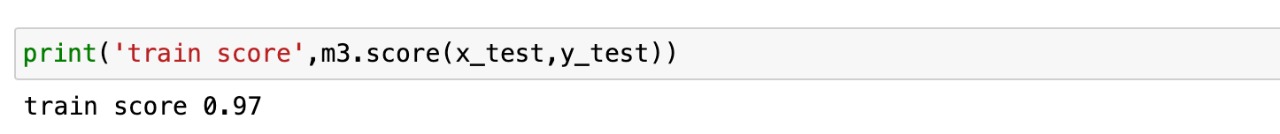


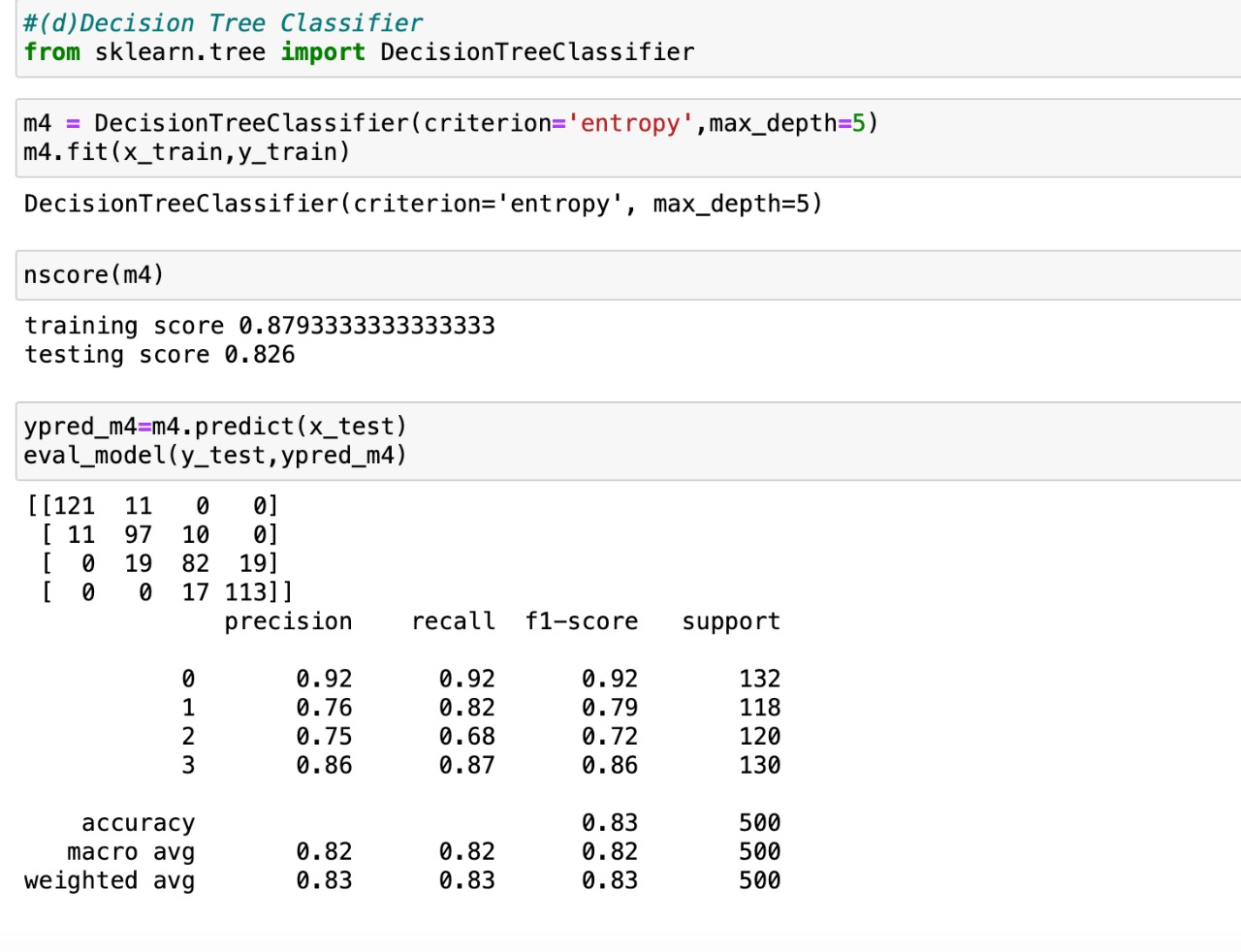
SVM classifier with linear and rbf kernal



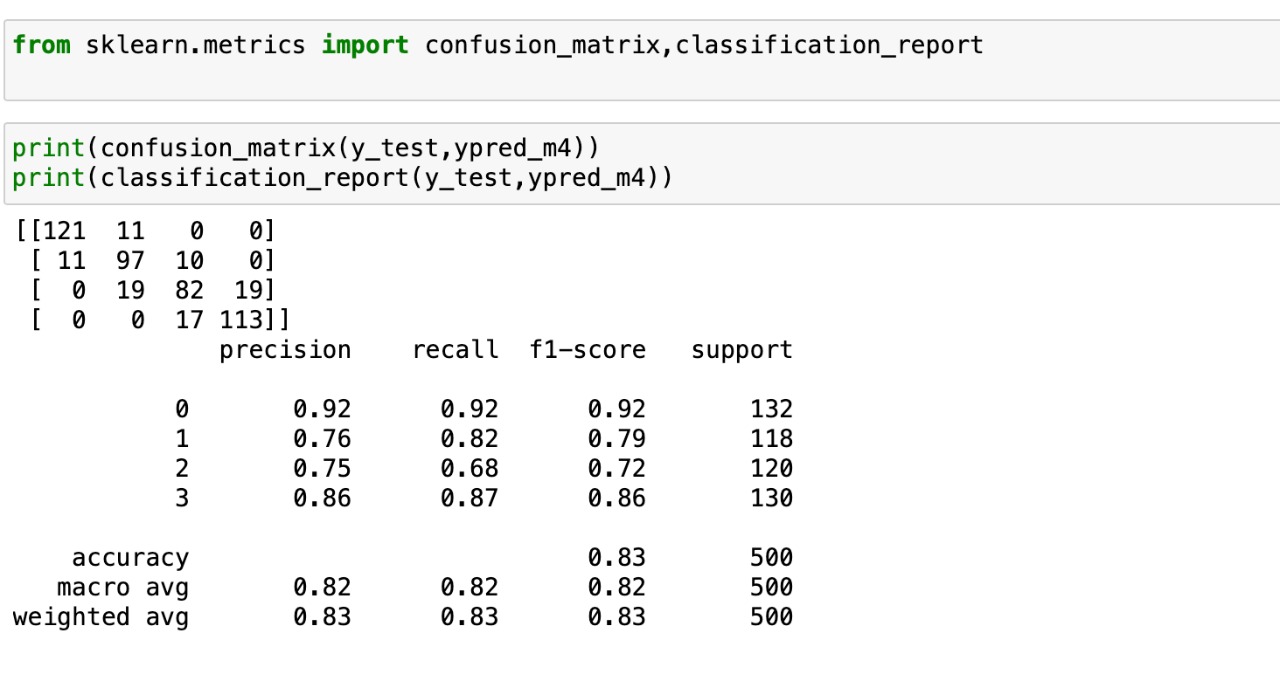


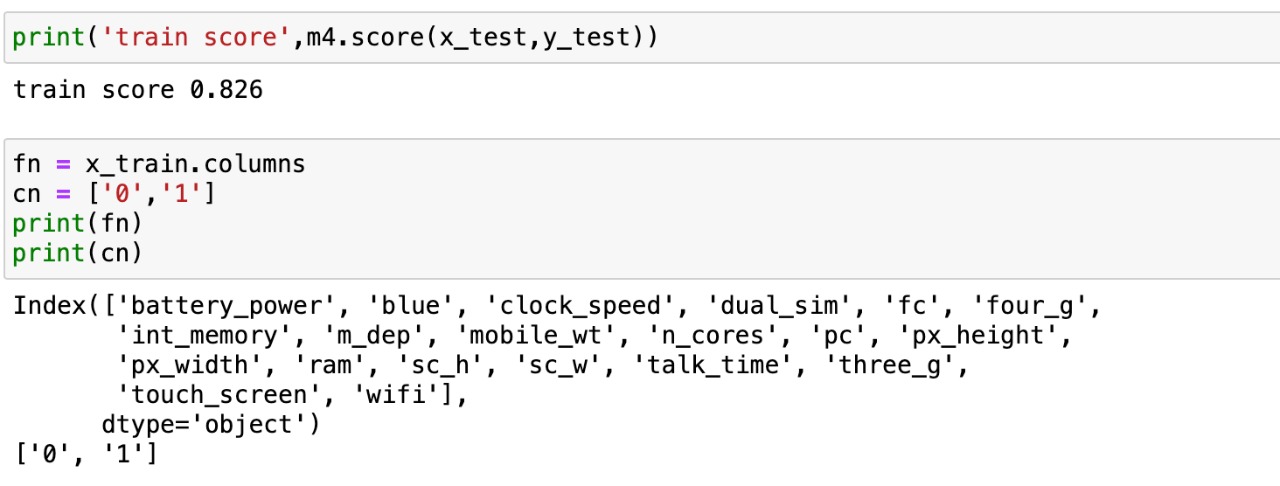




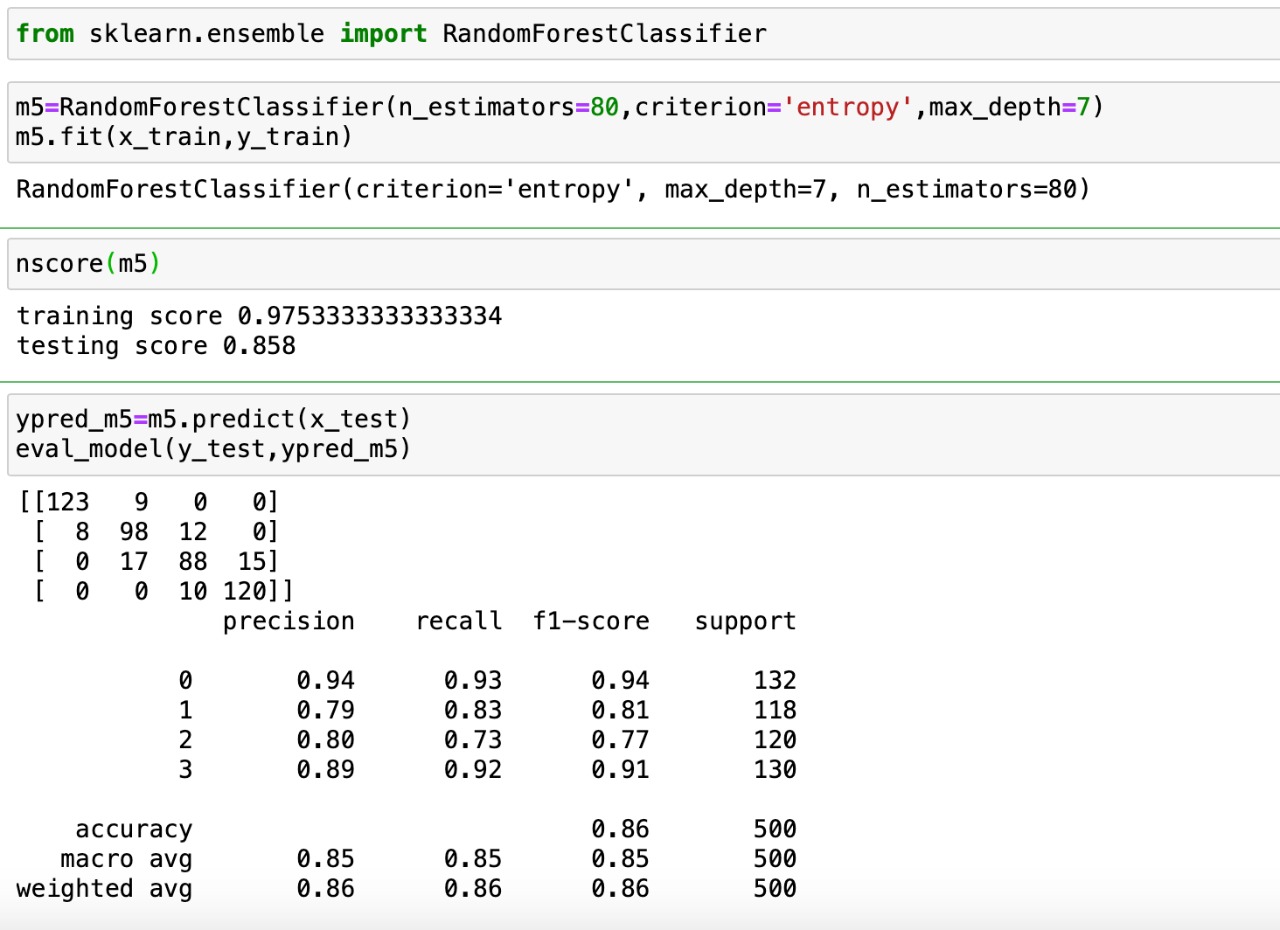




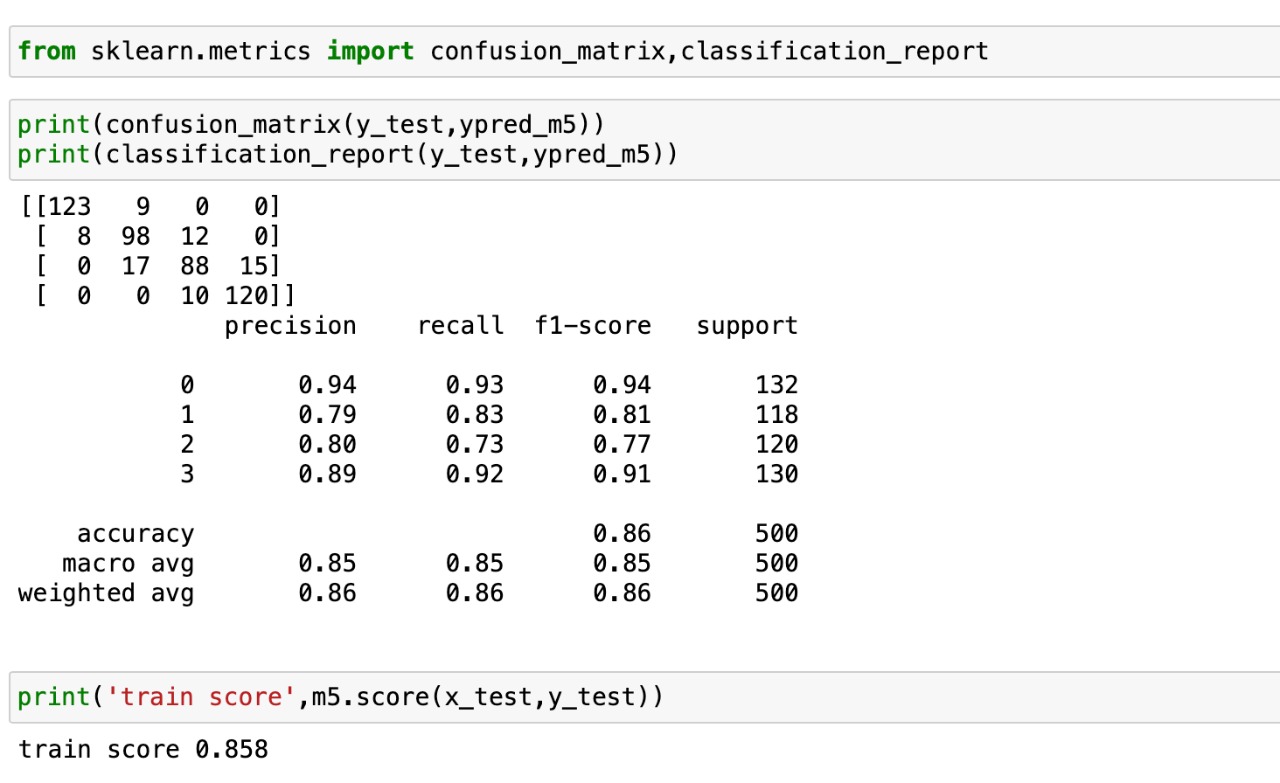


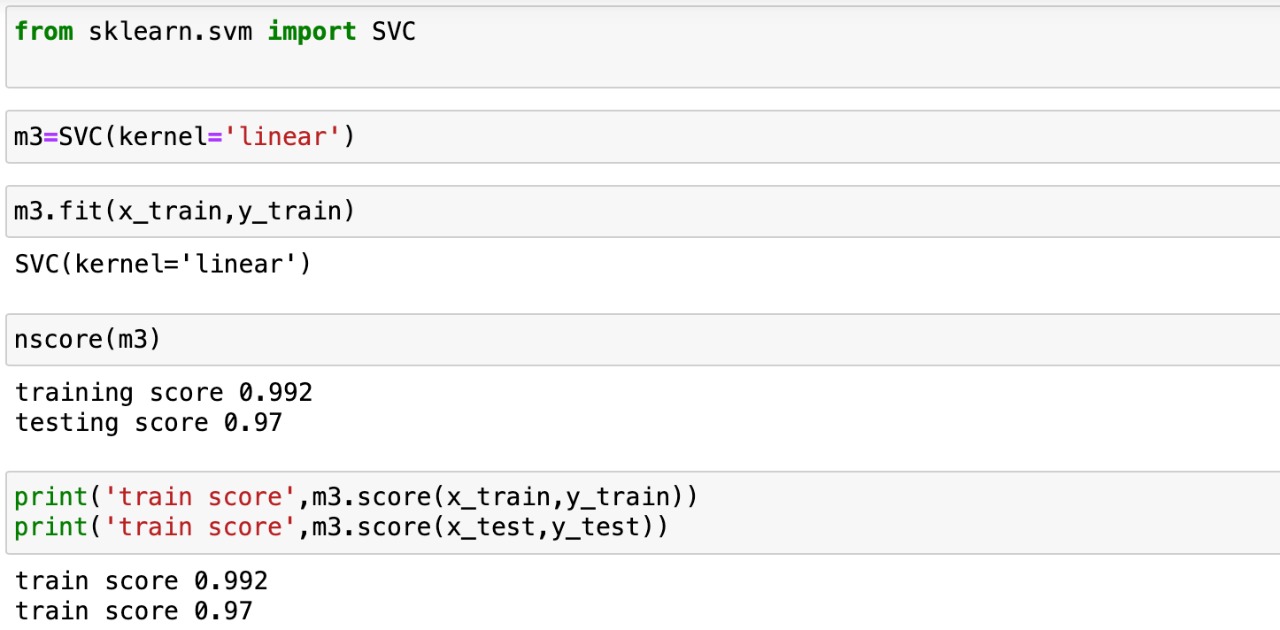


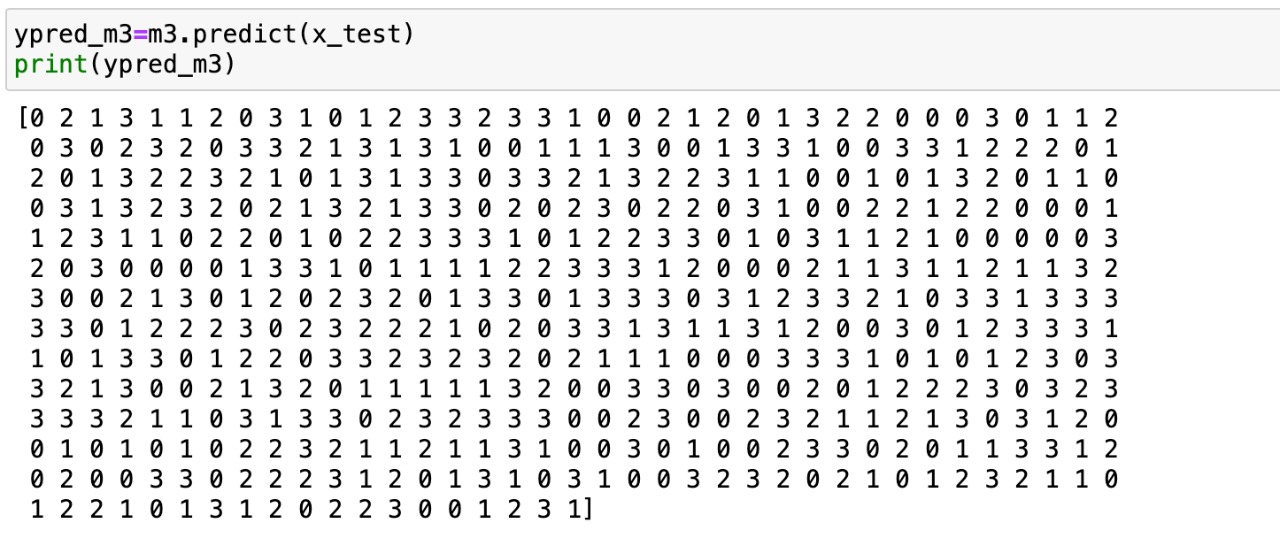
Random forest Classifier

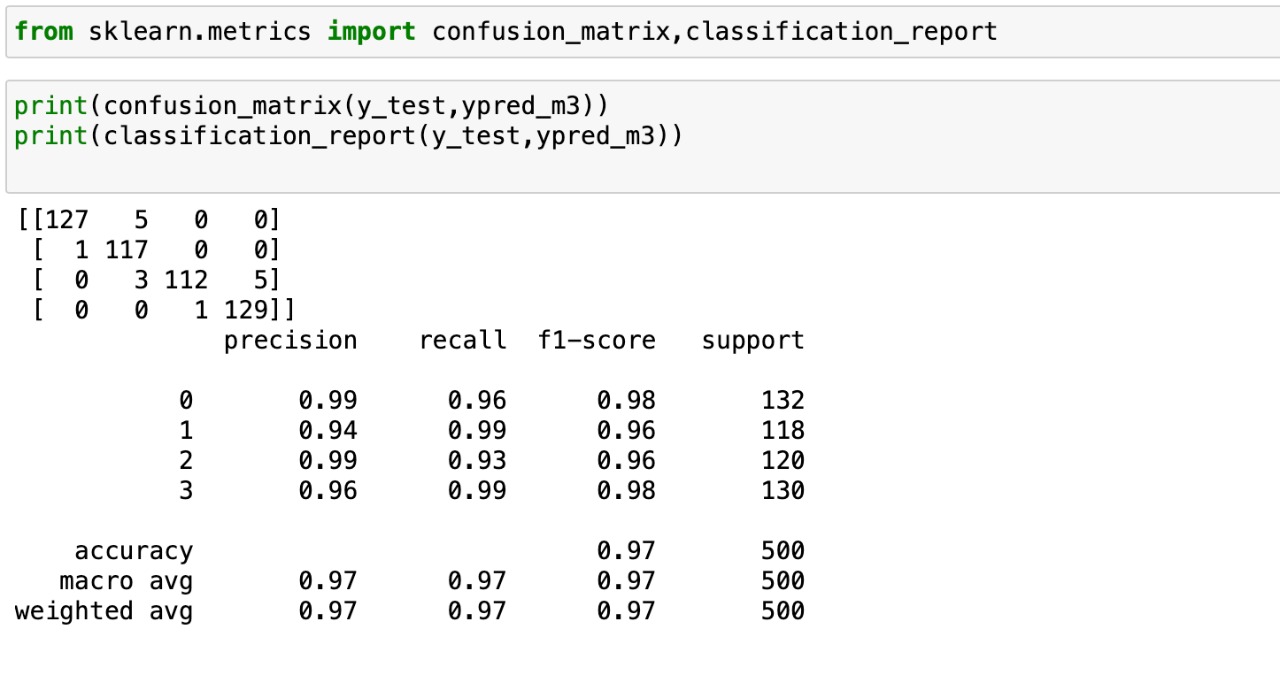


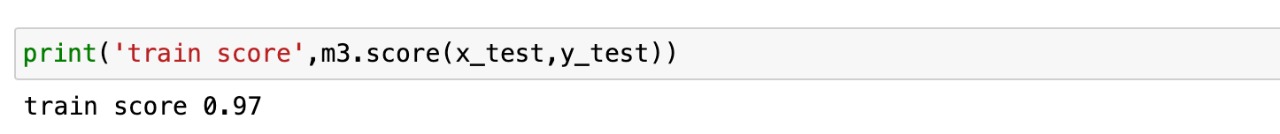












Source code link

https://drive.google.com/file/d/1-GTVExTFertxgH6nyg483HykjGsEWPUb/view?usp=sharing

Data set

<https://drive.google.com/file/d/1-Au_QjPYVvJv1hGFGqdleddGW8z78YYM/view?usp=share_link>

H