**Mobile Price Classification Model**

*NOTE: The dataset which has been used in our project is taken from the “KAGGLE” named as “Mobile Price Classification”*

**TABLE OF CONTENTS**

Contents

Acknowledgement&Abstract&Aim\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_3

About the dataset&features(columns)\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_3-4

What is model and steps involved in model classification\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_4-5

Python libraries\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_5

Python Code and Outcomes\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_5-16

How to select the model?\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_16-19

How to evaluate the model?\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_19-20

Conclusion\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_21

**ACKNOWLEDGEMENT:** I would like to express my gratitude and deep regards to our teacher Mr.satish kumar for his guidance. Sir give guidance and I gain some knowledge in the field of Data Science with Python. And also, I express gratitude to the KAGGLE website, and YOUTUBE for providing mobile price classification dataset.

**ABSTRACT:**

* Now a days mobile phone is used largely by every person.There is no comparison between the elderly and younger people in the usage of mobile phones.
* The mobile phone is the most common selling and buying electronic device.
* The main goal of our model is”**Whether a cell phone with certain features would be expensive or inexpensive**”.
* Everyday,hundreds of thousands of mobiles are sold and brought.as a result, the mobile price prediction is a case study for the given problem form,namely,finding the best product.

**AIM OF OUR PROJECT:**

**“The aim is to classify the price range of obile phones and not to predict the actual price.”**

By giving the features as input it give how much price range is of our feature values.

**INPUT**: Collecting the data from the kaggle website.

**OUTPUT:** Whether a cell phone with certain features would be expensive or inexpensive.

**ABOUT DATASET:** Our dataset have two ***“csv”*** files

* "train.csv” (with 2000 rows and 21 columns)
* “test.csv” (with 1000 rows and 21 columns)

**FEATURES(ABOUT THE COLUMNS):**

1. battery\_power : Total energy a battery can store in one time measured in mAh

2. blue : Has bluetooth or not

3. clock\_speed : speed at which microprocessor executes instructions

4. dual \_sim : Has dual sim support or not

5. fc : Front Camera megapixels

6. four\_g : Has 4G or not

7. int\_memory : Internal Memory in Gigabytes

8. m\_dep : Mobile Depth in cm

9. mobile\_wt : Weight of mobile phone

10. n\_cores : Number of cores of processor

11. pc : Primary Camera megapixels

12. px\_height : Pixel Resolution Height

13. px\_width : Pixel Resolution Width

14. ram : Random Access Memory in Megabytes

15. sc\_h : Screen Height of mobile in cm

16. sc\_w : Screen Width of mobile in cm

17. talk\_time : longest time that a single battery charge will last when you are

18. three\_g : Has 3G or not

19. touch\_screen : Has touch screen or not

20. wifi : Has wifi or not

**What is a model**: A model is also known as a predictive modelling.it is commonly used statistical technique to predict future behavior.

Predictive modelling solutions are a form of data-mining technology that works by analyzing historical and current data and generating a model to help predict future outcomes.

**STEPS INVOLVED IN MODEL CLASSIFICATION:**

1. Collection of data and loading the data into dataframe

2. Exploring Data Analysis

3. Data Preprocessing

4. Data Visualization

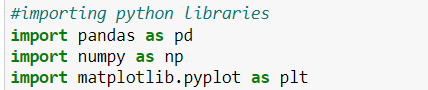
5. Separating the data into Featured and Target data

6. Model Implementation

7. Model Selection

8. Model Evaluation

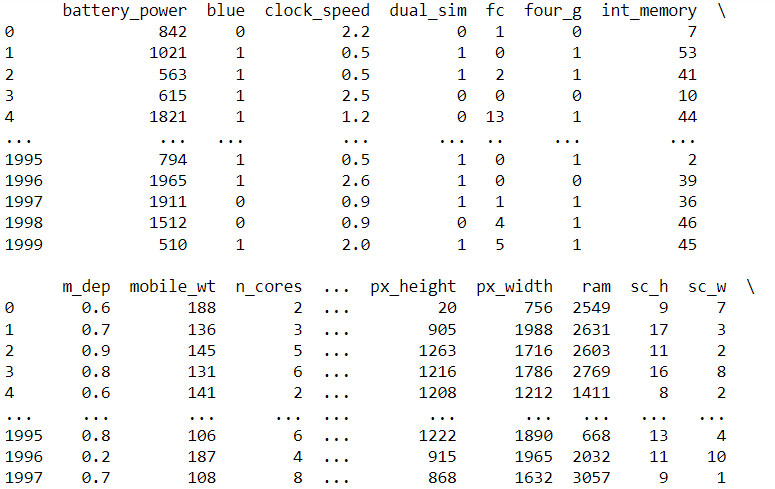
**IMPORTING THE PYTHON LIBRARIES IN OUR MODEL:**

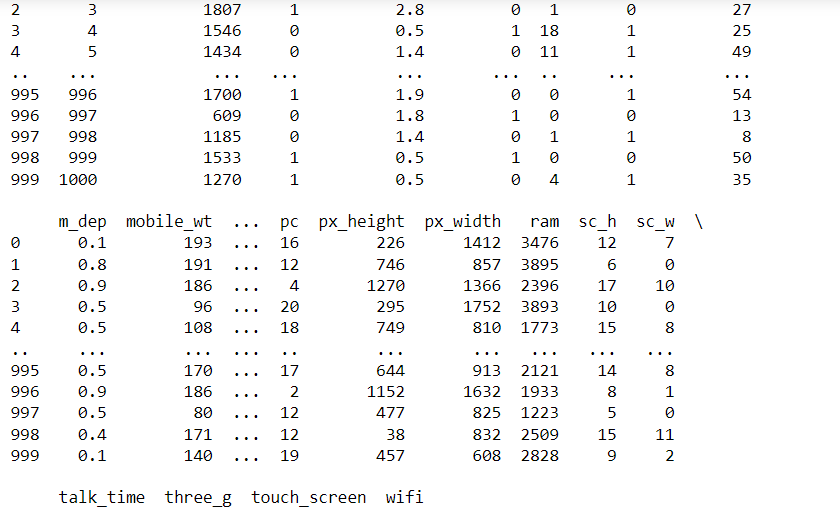
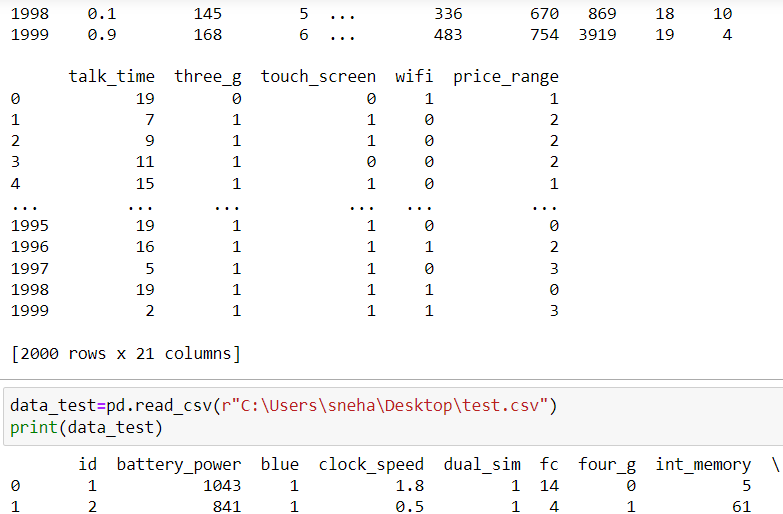


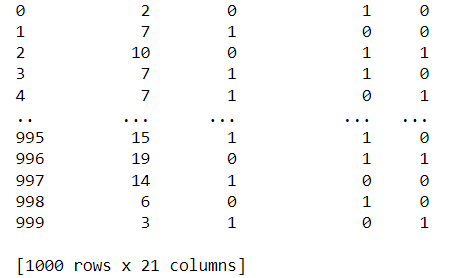
**1.READING THE DATA SET:**



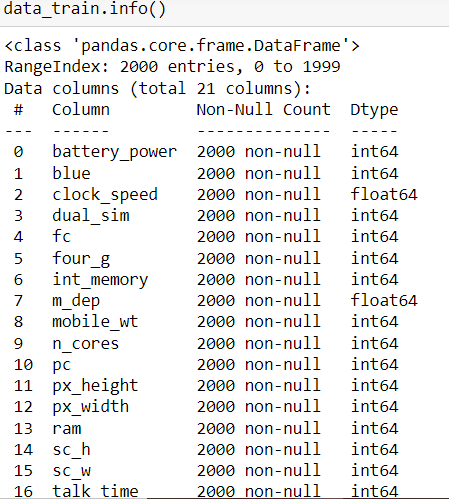
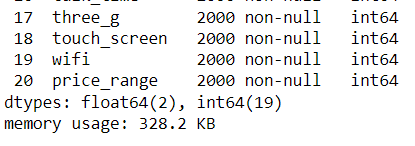
**OUTPUT:**

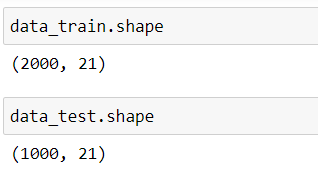
****

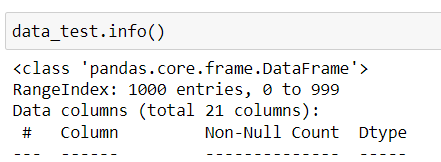
****

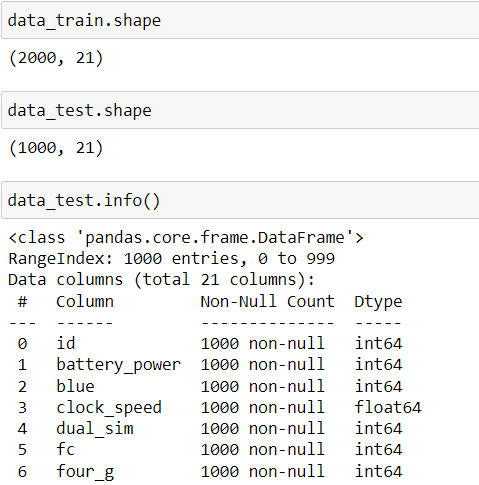
****

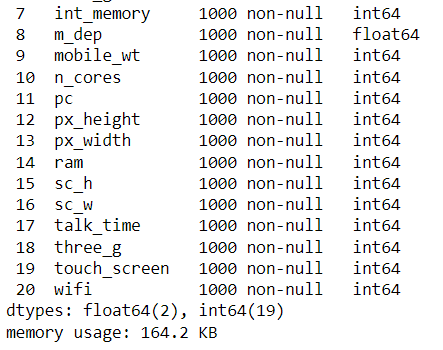
**2.EXPLORATORY DATA ANALYSIS:**

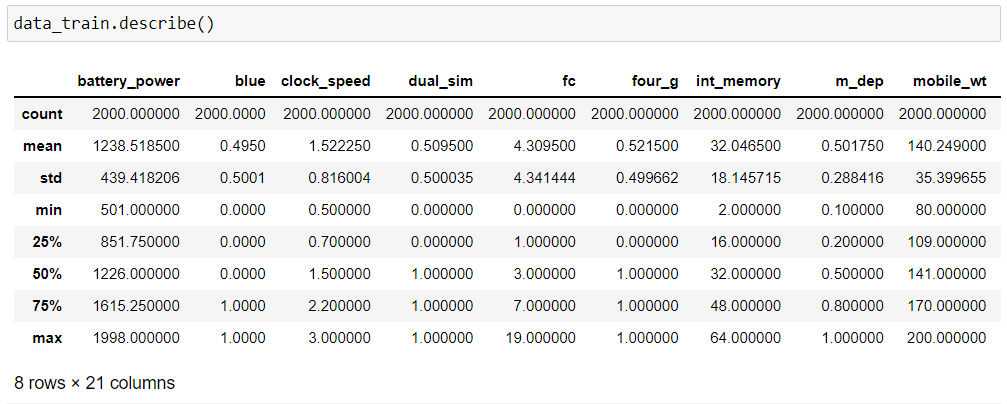




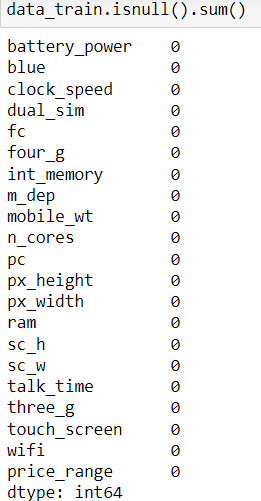




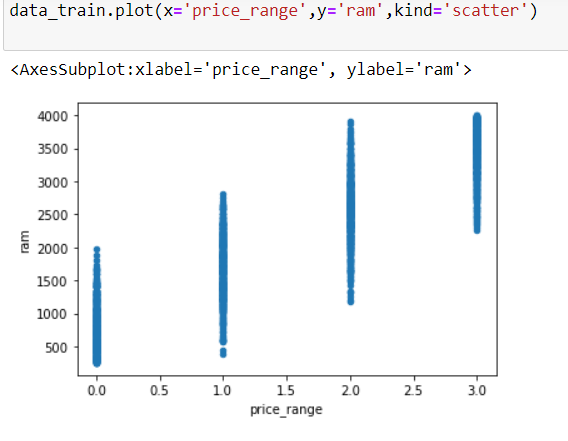
**DESCRIPTION OF DATA:**

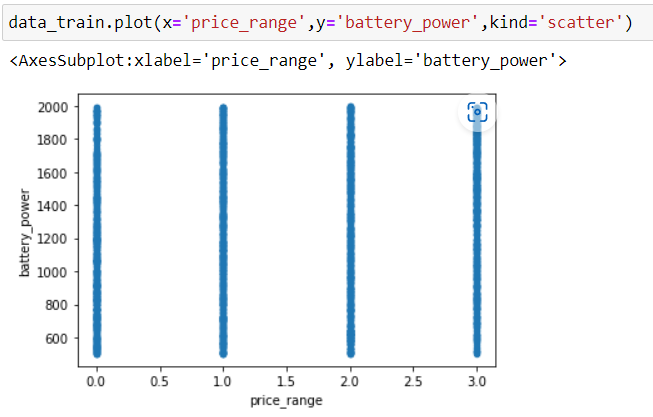


**3.DATA PROCESSING:**



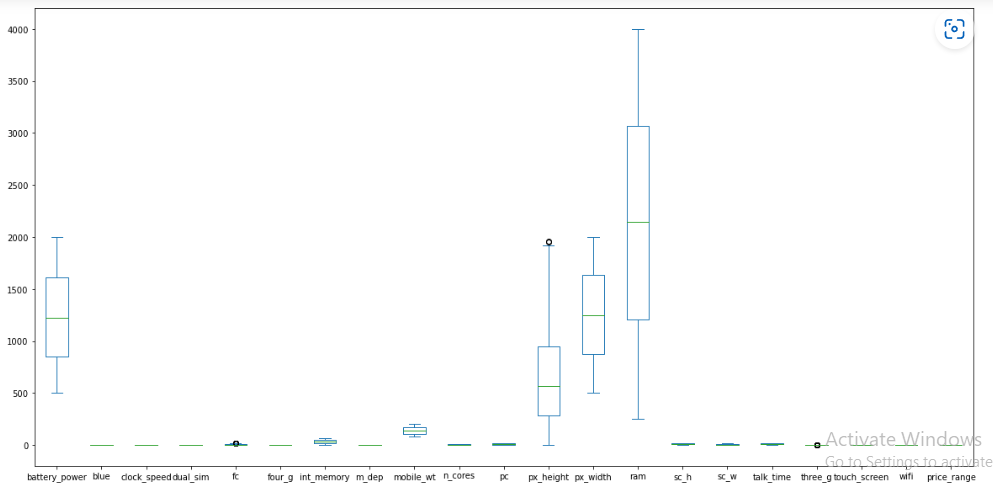
**4.DATA VISUALIZATION WITH SOME COLUMNS:**





**OUTLIERS:**

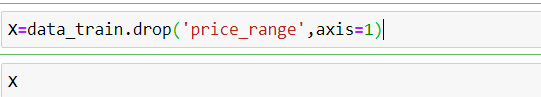


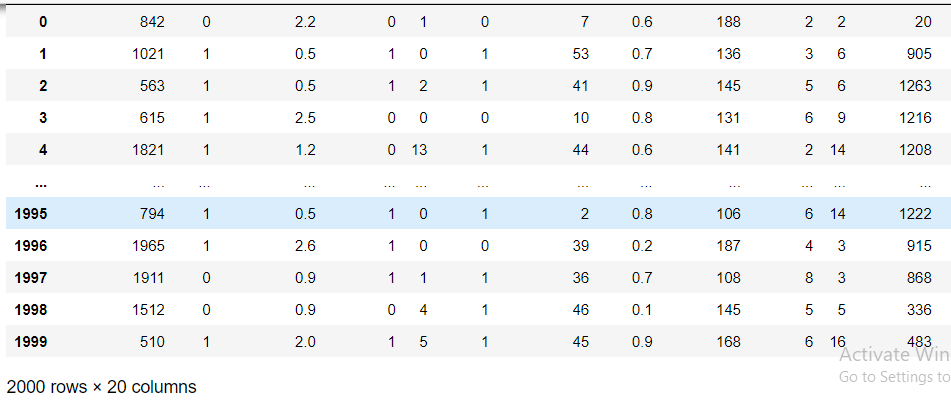


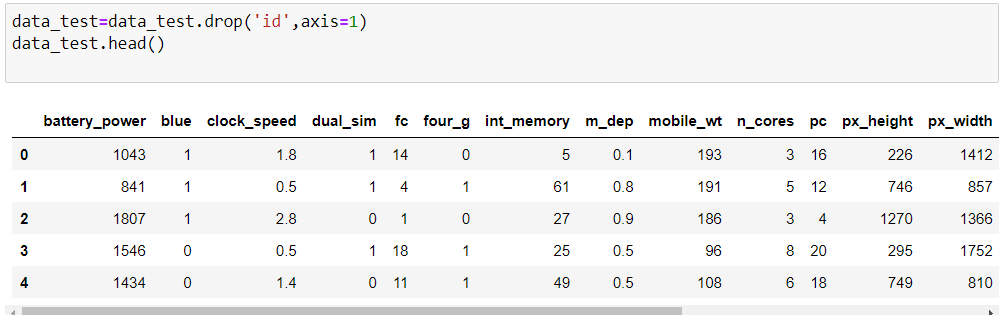
**5.MODEL IMPLEMENTATION:**

Separating the featured data and target data:

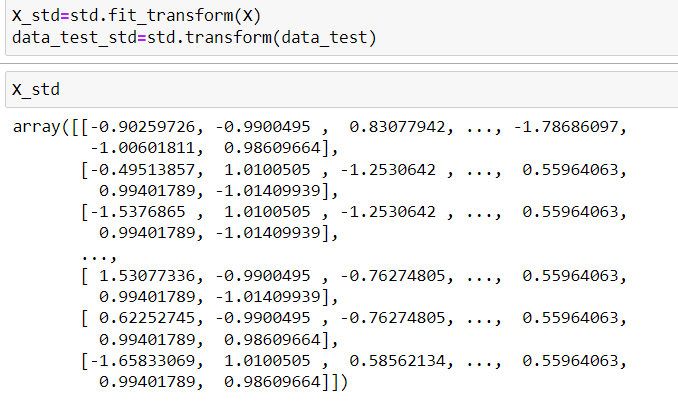
We split the whole dataset into featured or input features and target or output variable.





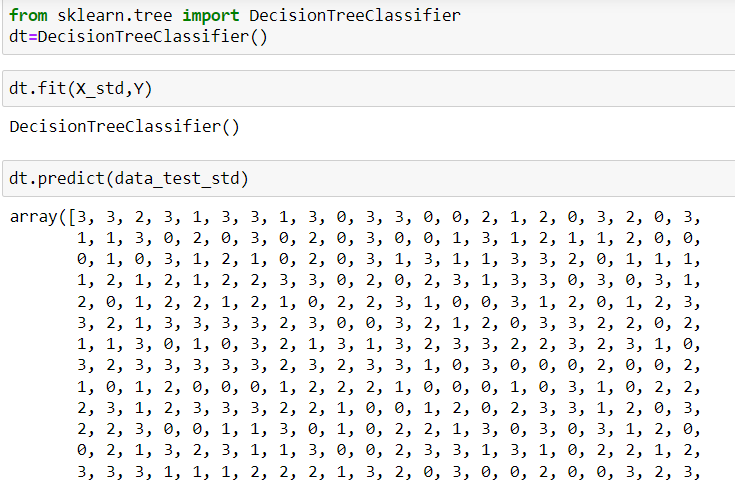


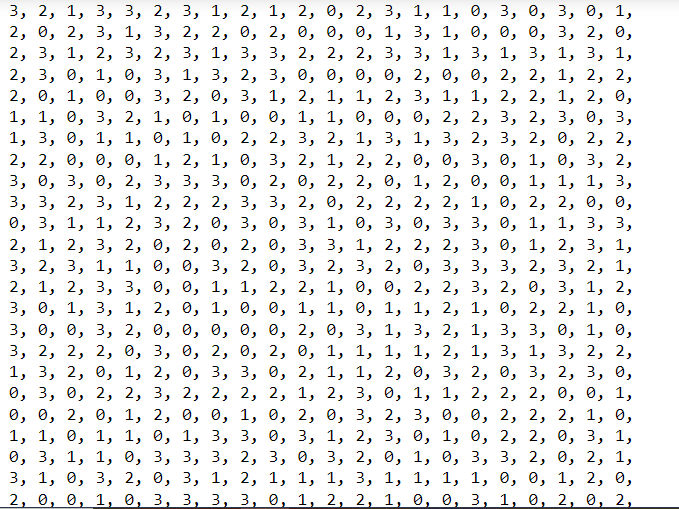
**FINDING THE STANDARD DEVIATION:**

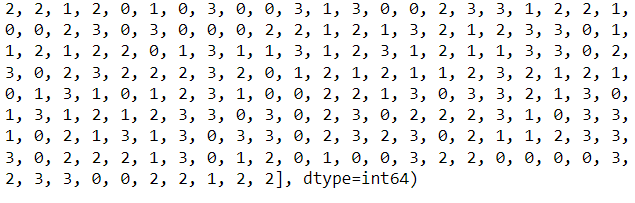


Splitiing the data into training and testing set:

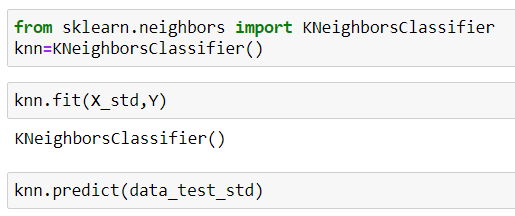
**TRAINING WITH DECISION TREE:**



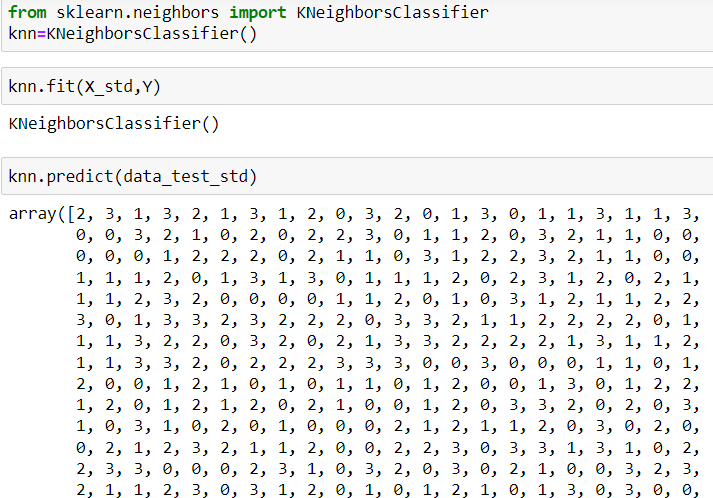


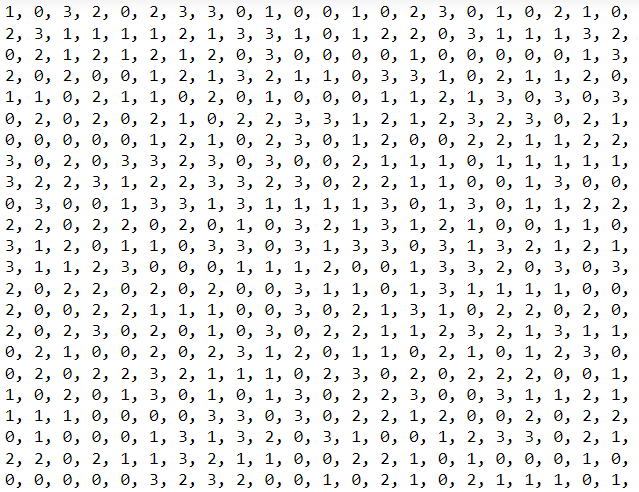


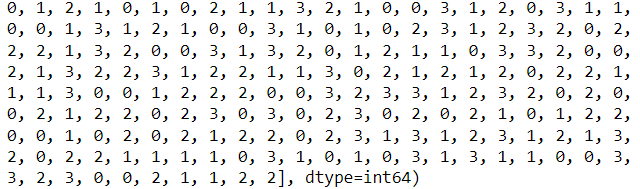
**TRAINING WITH KNN:**

****

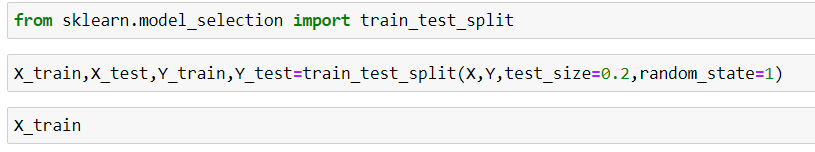
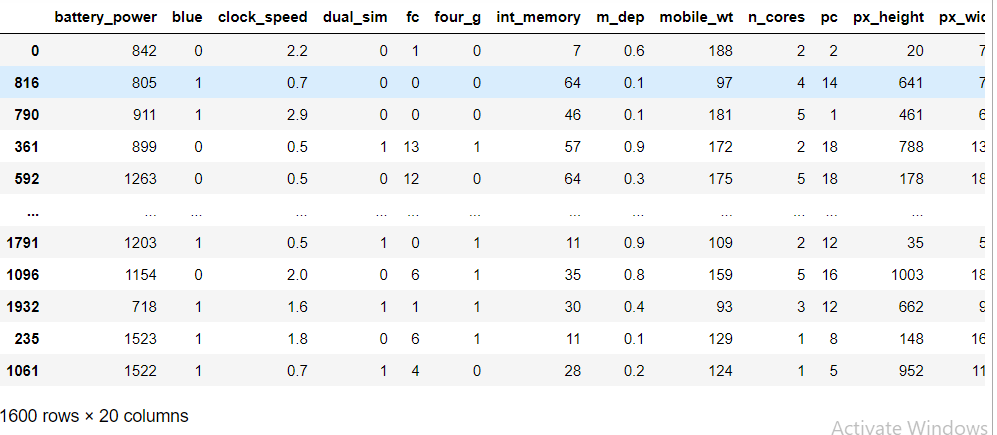
**OUTPUT:**

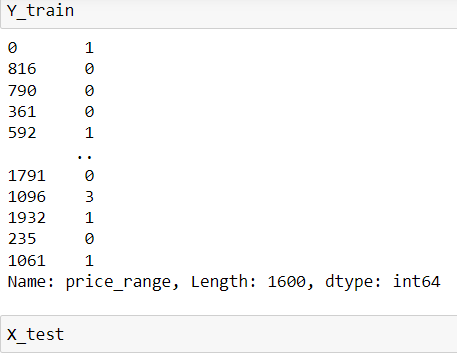


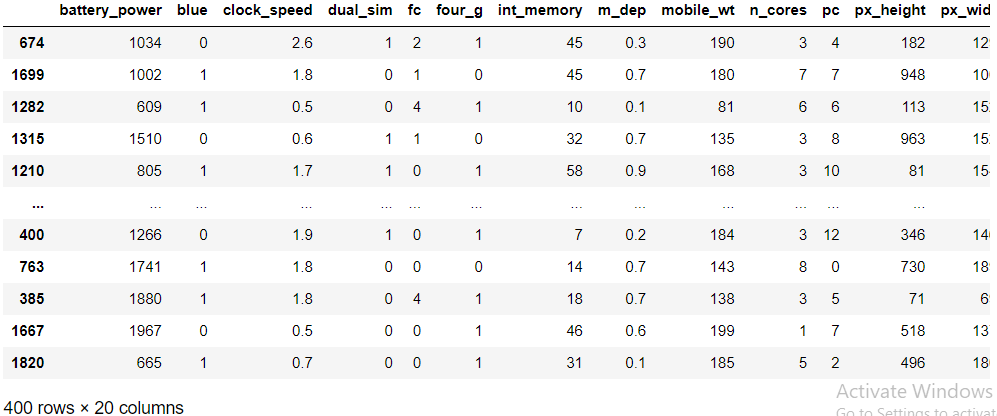


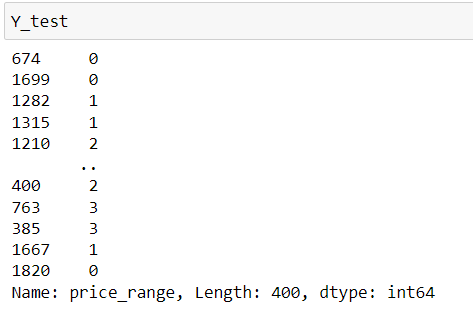


**TRAINING AND TESTING OF THE DATASET:**







**6.MODEL SELECTION:**

This is the most exciting phase in Applying Machine Learning to any Dataset. It is also known as Algorithm selection for Predicting the best results. Usually Data Scientists use different kinds of Machine Learning algorithms to the large data sets. But, at high level all those different algorithms can be **classified in two groups : supervised learning and unsupervised learning.**

**Supervised Learning:** Supervised learning is a type of system in which both input and desired output data are provided. Input and output data are labelled for classification to provide a learning basis for future data processing. Supervised learning problems can be further grouped into Regression and Classification problems.

* A regression problem is when the output variable is a real or continuous value, such as “salary” or “weight”.
* A classification problem is when the output variable is a category like filtering emails

“spam” or “not spam

**Unsupervised Learning:** Unsupervised learning is the algorithm using information that is neither classified nor labeled and allowing the algorithm to act on that information without guidance. Here our dataset we have the outcome variable or Dependent variable i.e Y having only two set of values, either 0 or 1 or 2 or 3. So we will use Classification algorithm of supervised learning

**We have different types of classification algorithms in Machine Learning:**

1) Logistic Regression

2) KNN Classification

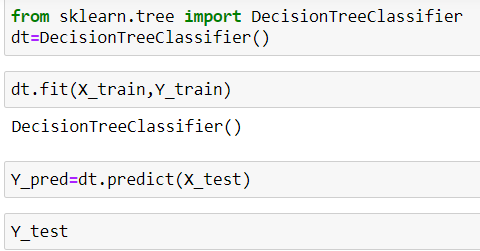
3) Support Vector Machine Algorithm

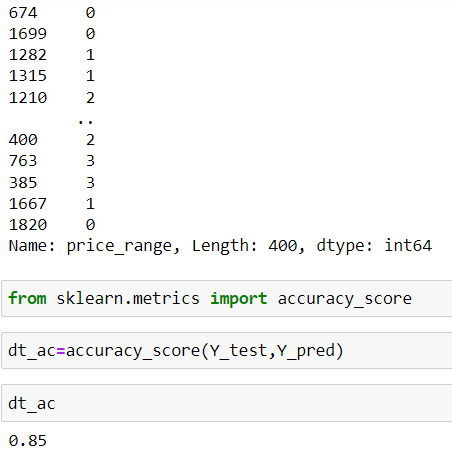
4) SVC Kernel Method

5) Random forest classifier

6) Decision Tree Classification Algorithm

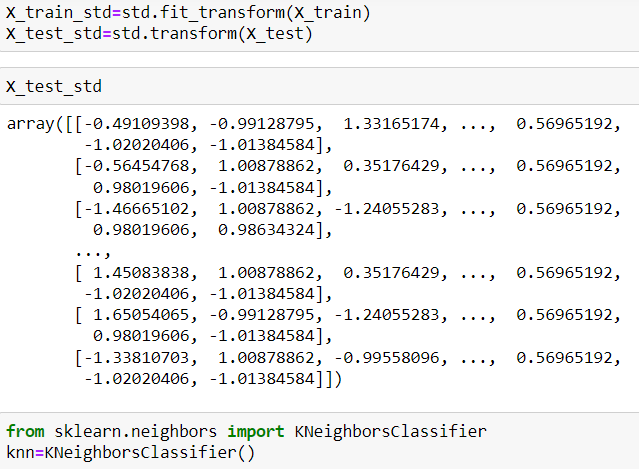
**Applying Decision Tree classifier method and it’s accuracy score:**

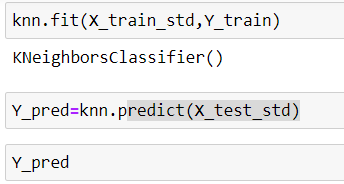


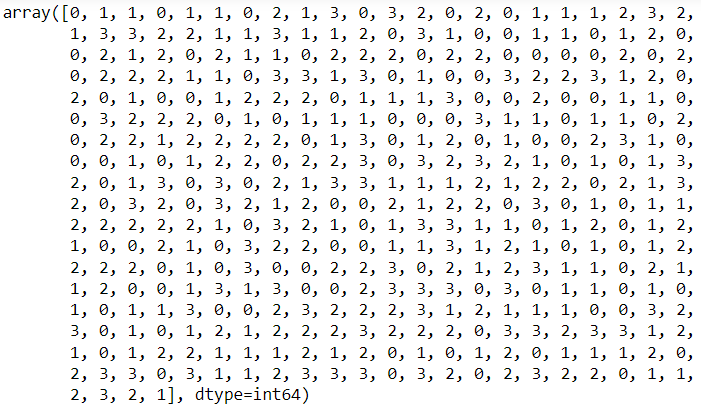


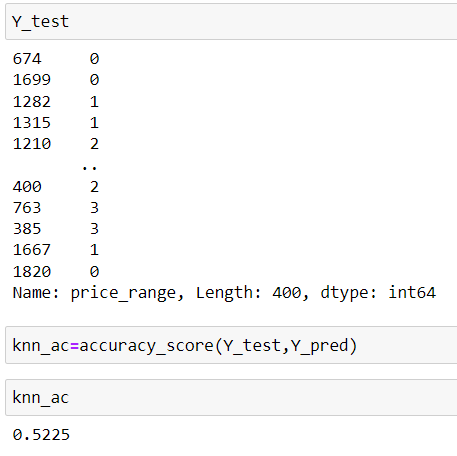
**Accuracy score by using decision tree classifier algorithm is 0.85**

**Applying knn method and it’s accuracy score:**









**Accuracy score by using knn classifier is: 0.5225**

By using different types of classification we have got different accuracy values.

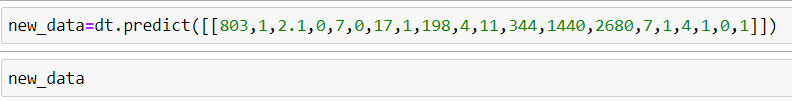
So finally we have built our classification model and we can see that

**“ Decision Tree classifier**” algorithm gives the best results for our dataset. Well its not always applicable to every dataset.

**7.MODEL EVALUATION:**

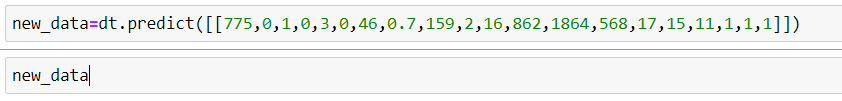
Model Evaluation is the process through which we quantify the quality of a system’s predictions.to do this, we measure the newly ttrained model performance on a new and independent data set.This model will compaer labeled data with it’s own prediction.

Finally,the output for our classification uses 20 features,and the output is:



**Output:**



2)

**Output:**



3)



**Output:**



**CONCLUSION:**

Finally, we can predict the mobile price range as an output 0 or 1 or 2 or3 (LOW,MEDIUM,HIGH,EXPENSIVE)ranges based on input features Of the mobile by using different classification algorithms of machine learning.

Among those classification,we can get the more accurate score by using

”**DECISION TREE** ” machine algorithm classification.

**NAME :** M.SNEHA

**ID :** S180005

**CLASS :** CSE-2A

**SUBJECT:** DATA SCIENCE WITH PYTHON

**(*MOBILE PRICE CLASSIFICATION MODEL*)**

g-mail:marpusneha678@gmail.com

https://github.com/marpusneha/Data-Science-Project.git