

**COMSATS UNIVERSITY ISLAMABAD,**

**WAH CAMPUS**

**SUBMITTED TO:**

**Dr. M. Khalil Afzal**

**SUBMITTED BY:**

**Assignment:** LR, SVM and Decision Tree.

Muhammad Uzair (FA17-BSE-042-6A).

**Date:**02-07-2020.

# Introduction:

According to instruction. I pick the datasets of various makes of cars from various years that gathered by  [UC Irvine Machine Learning Data Repository](http://www.ics.uci.edu/~mlearn/MLRepository.html). And data set contain total eight features of car. It’s helpful for making prediction of cars according to price or as per your wish or also use to test the accuracy of dataset. I use this training and testing dataset on python.

Here is my GitHub repository link that contain my implemented algorithm and I also implement one algorithm on Weka tool for my personal experience.

<https://github.com/Muhammad-uzair029/ML_Assignment_02>

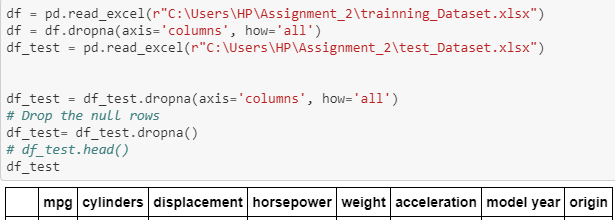
Methodology and implementation:

I use to Linear regression, Naïve base and Decision tree on this algorithm. I will show you one by one how above this algorithm react to this dataset.

Linear Regression

First thing; I implement the multivariable linear regression for more accuracy and prediction.

* After importing pandas, matplolib, and linear regression libraries I import the train and test datasets.



Then I create a dataFrame (that make the subset from original dataset) using panda library and use training dataset for splitting. And I add the three columns (cylinder, acceleration and model year).

Then import the linear model for linear regression and matplotlib also.

Then, I import the library for test\_train\_split for splitting the train and test size and random state.

I use random state variable for splitting the data and I create split the data into three subsets.

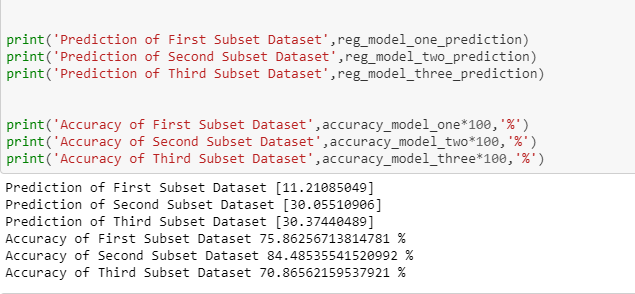
* In the first subset I add the first 80 rows for training the model.
* In the second subset I add the 160 rows for training the model.
* In the third subset I add the next 80 rows up to 140 for training the model.

And fit the model one by one and predict the value according to our test dataset values X and Y. and find the prediction on “mpg” of dataset.

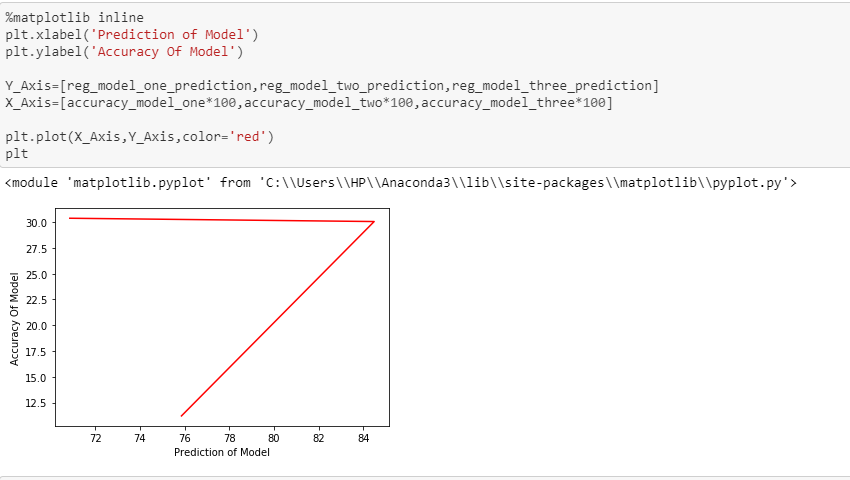
Then also check the accuracy of each splitting dataset



* Here is my print of accuracy and prediction of my above three models.



* Here is plot graph and indicate the three corner or points because I have created three subsets.



Naive Bayes



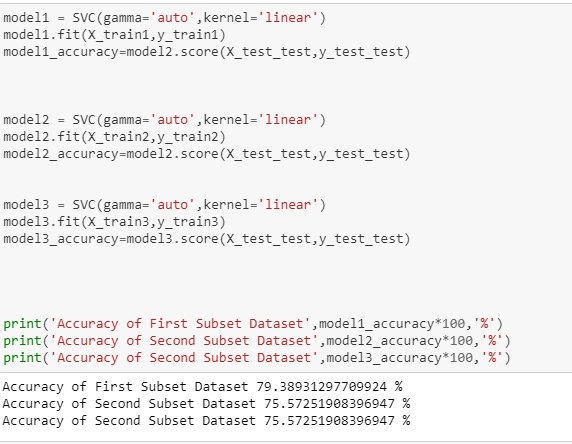
After importing library; I create the subsets object for test and split methods of training and testing data.

On line 43. I implement test and train split on test dataset for prediction of origin value.

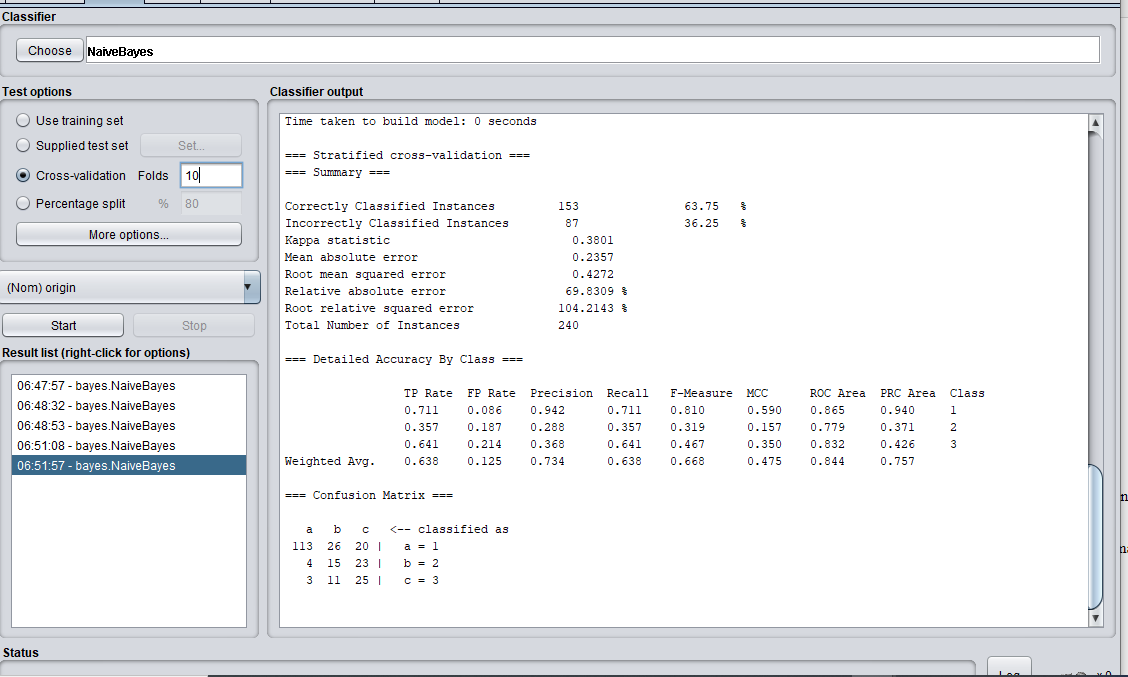
On line 44. I implement test and train split on train dataset for creating model. And also split the datasets into subset.



* I set the parameter kernel to linear for smoothness and gamma is auto value set. Then, I create the three models for accuracy checking and pass the testing dataset for checking accuracy.

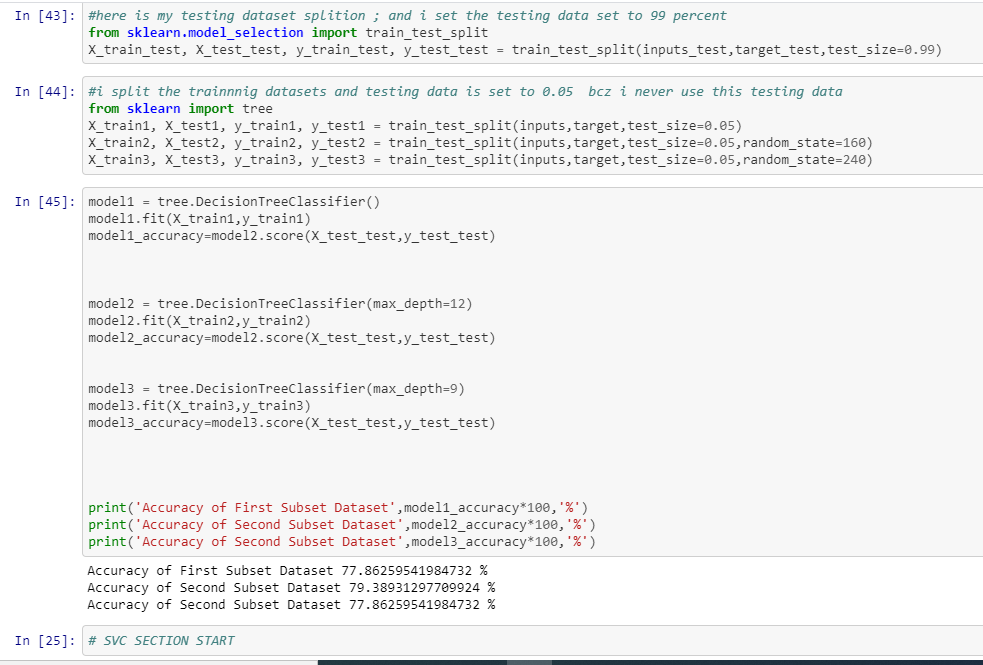


### And I Also implement the “Naïve bayes” on Weka for my personal experience.

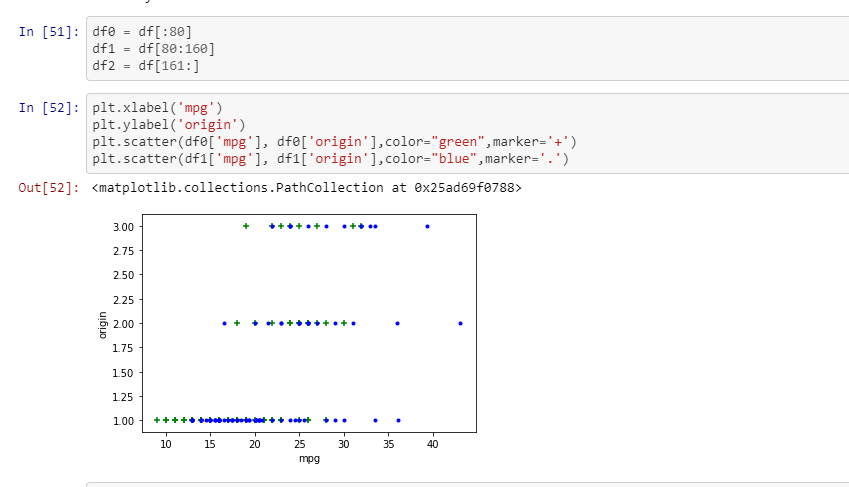


Decision Tree

* Here are my decision tree implementations of decision tree that very similar implementation as above naive bayes. Same as naïve bayes I create the test and split methods for train and test datasets and then create the three model and then find the accuracy of each model



* Here is a plot graph of Naïve bayes and decision tree that shows the three subsets between origin and mpg values. The separate the mpg value according to origin. Because we have three (1,2,3) of origin on this data set.

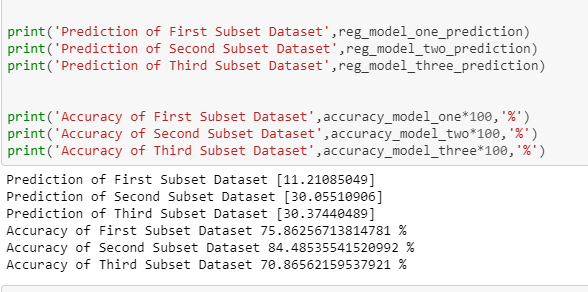


# Findings and results:

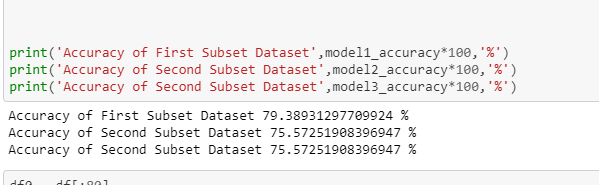
The prediction of three models in linear regression is very accurate , u can test my model from getting code of my repository and in case of naïve bayes and decision tree the accuracy of decision three is little bit more than naïve base.

Naïve bayes accuracy is less than 80 percent on the other hand on decision tree the accuracy touch the 80 percent.

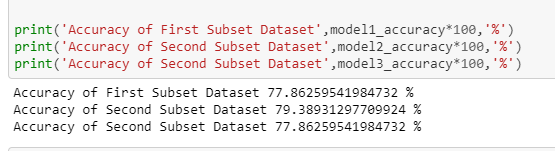
### Linear Regression prediction and Accuracy:



### Naïve Bayes Accuracy:



### Decision Tree Accuracy:



# The End