# CSE 4020 - MACHINE LEARNING

# Lab 29+30

# Digital Assignment-5

# Submitted by: Alokam Nikhitha(19BCE2555)

# Random Forest

**Question:**

Use random forest regression as part of ensemble learning to predict the amount of petrol consumption by studying different traits of a particular place.

**Dataset Used:**

petrol\_consumption.csv

**Procedure:**

-Using pandas, we first import the dataset into our workspace.

-Next we define the set of dependent and independent attributes.

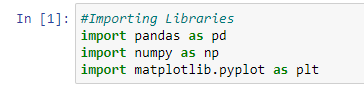
- We then import the random forest regressor from sklean rn.ensemble and train our model using the independent and dependent attributes.

- Next, we have printed the results of independent set as predicted by our regressor.

- Lastly, To check for the performance of our dataset, we have printed all the evaluation metrics

Since it has less Number of Rows we haven’t split the dataset

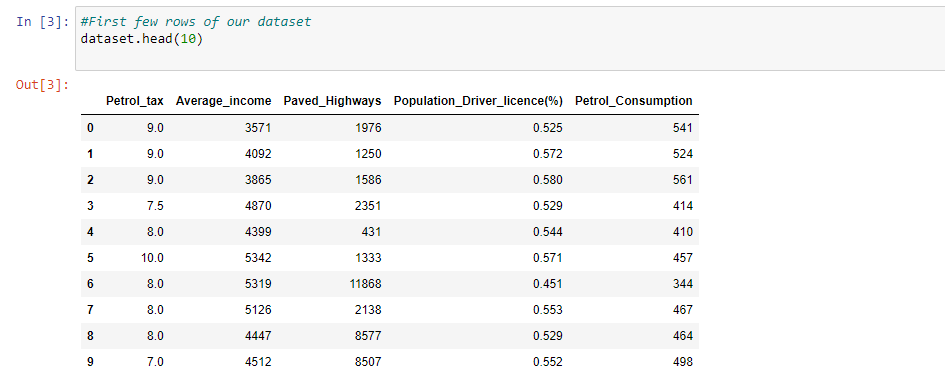
**Code Snippets and Explanation:**

****

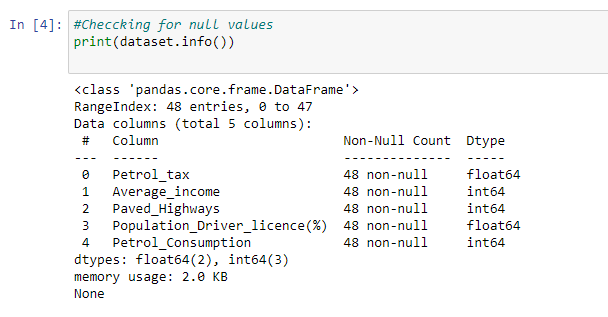
**Here we are importing the required Libraries**

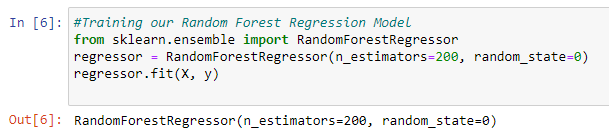
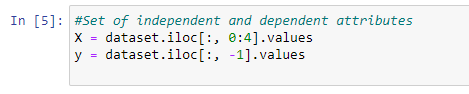
****

**Using Pandas we are importing the data**

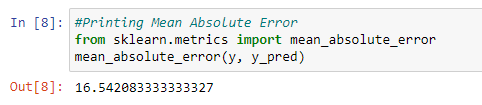
****

**Printing the first few rows.**

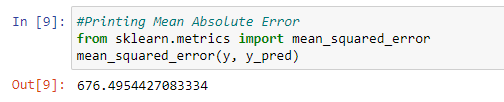
****

****

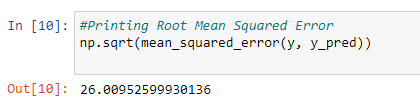
We have Defined set of Dependent and Independent attributes.The n\_estimators here indicate the number of decision trees that we are using to train our random forest regressor. Hence we are using 200 decision trees for prediction. For final value we have used the average value of each decision tree to find the final consumption of petrol of a particular region.

****

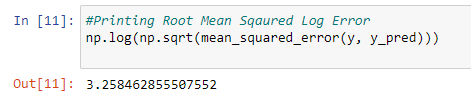
**Printing the Mean Absolute Error**

****

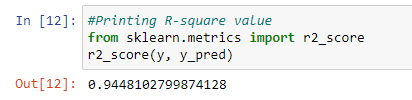
**Printing the Mean Squared Error**

****

**Printing the Root Mean Squared Error**

****

**Printing the Root Mean Sqaured Log Error**

****

**Printing the R-square value**

**Results and Conclusions:**

**Mean Absolute Error from cell8 is** 16.542083333333327

**Mean absolute error from cell 9 is** 676.4954427083334

**Root Mean Squared Error from cell10 is** 26.00952599930136

**Root Mean Squared Log Error from cell11 is** 3.258462855507552

**R-square value from cell12 is** 0.9448102799874128