



Rudraksh : 23225034

A14 Machine Learning ESE-1

First using pandas, we read the CSV and create a dataframe.

Then we use the `Head()` ~~com~~ method to see a few rows of data.

The given question asks us to predict the Fuel consumption & so our variable to be predicted is fuel consumption. The fuel consumption is a range of numbers so it's a Regression problem.

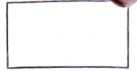
Our dependant variable is Fuel consumption. Then we need to find out the list of independent variables that can help us predict the value of Fuel consumption.

We start by seeing the correlation b/w the fuel consumption and other variables.

We have considered the columns

Engine Size, Cylinders, Fuel consumption, CO emissions & transmission labels.

Our transmission values are in string, aimed to encode them into Int for corr to work. we will Encode using label encode.



Now using Linear Regression, we will find the line of best fit for these ~~var~~ points.

for training the model we used sklearn library for train test split, linear regression as ~~as a~~ ~~as a~~ ~~as a~~ column don't contain Null values (checked using value counts) we don't need to clean the data as such.

After training, we got an accuracy of 1 for all the variables.