MID-TERM\_ASSIGNMENT:

So the data set that was given to me was a car classification dataset which is basically a categorical data set.

Now let us go through the steps/stages:

1. EDA: here I loaded the csv. File and did a bit of data exploration using various codes(.head(), .info(), .sample(), etc)
2. Next I went on to create a heat map of the columns on dataset with each other. It shows Pearson's correlation coefficient of column w.r.t other columns
3. Next I went to do the hot-encoding by firstly installing sklearn and importing the preprocessing function.
4. ,Next I did the logistic regression through which I came to see that the accuracy rate is 0.6647398843930635.
5. Next I went for the multi-layer neural network and confusion matrices.  
   here from sklearn.neural\_network I imported MLPClassifier and trained the test data-sets.
6. Here we changed the hidden-layer sizes till we got the required accuracy.
7. We finally did the confusion matrix for logistic regression and neural networks.
8. We found out the False positive rates and False negative rates for both the models.
9. I would like to conclude that since it was a categorical data neural network performed better.