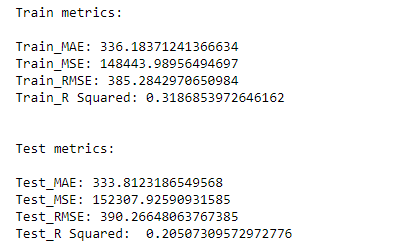
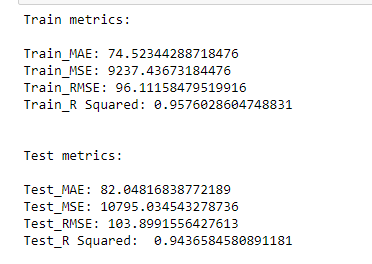
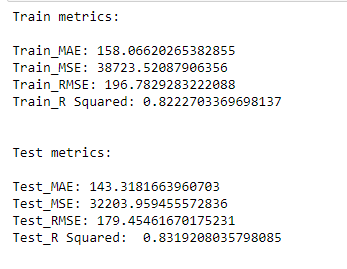
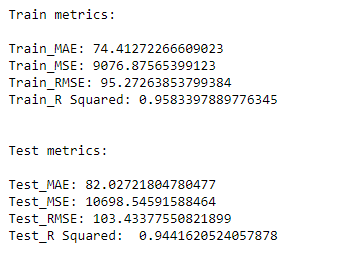
Data Preprocessing:

Encoding Categorical Variables.

How do we take care of ordinal Variables?

Do we need to do Feature Scaling



When I ran IV Scatterplots to check for Multicollinearity (by correlation), I found Limit and Rating being higly

correlated (r =1.0) and also, Income being correlated to Limit and Rating to the tune of r = 0.79

So I dropped Rating from the IV list, my train\_MSE increased to 96.11 and R Squared reduced to 95.76.

With Income and Limit being highly correlated as well, I should drop one more but should it be Income or Limit? It should

be Income because in DV Plots, DV - Balance is highly correlated with Limit.