**SUBJECTIVE QUESTIONS ADVANCED REGRESSION ASSIGNMENT**

**Question 1**

*What is the optimal value of alpha for ridge and lasso regression? What will be the changes in the model if you choose double the value of alpha for both ridge and lasso? What will be the most important predictor variables after the change is implemented?*

**Ans**. Optimal value of alpha in Ridge regression will lie between 15 to 30 whereas in Lasso Regression, it will lie till 50. If we double values of alpha, the model will underfit and be a simple model with increased bias. So, we should strike a balance between low values of alpha giving us high variance and overfitting and higher values of alpha making high bias and underfitted models. Most important variables will be numeric variables like Lot Frontage, TotalBsmtSF, FullBath, GarageCars, GarageArea, Fireplace, OpenPorch etc.

**Question 2**

*You have determined the optimal value of lambda for ridge and lasso regression during the assignment. Now, which one will you choose to apply and why?*

**Ans**. I shall apply Lasso Regression as it carries out feature elimination as well. As there are a large number of predictor variables in the model, I shall use Lasso Regression to reduce the number of predictor variables. An optimal solution in this regard will be to use Elastic Net which uses values of alpha as well as L1/L2 ratio that is, ratio of Lasso and Ridge Regression for hyper parameter tuning.

**Question 3**

*After building the model, you realised that the five most important predictor variables in the lasso model are not available in the incoming data. You will now have to create another model excluding the five most important predictor variables. Which are the five most important predictor variables now?*

**Ans.** Five most important variables are Lot Frontage, TotalBsmtSF, FullBath, GarageCars, GarageArea, Fireplace, OpenPorch, TotRmsAbvGrd.

**Question 4**

*How can you make sure that a model is robust and generalisable? What are the implications of the same for the accuracy of the model and why?*

**Ans.** Hyper parameter tuning is essential for making this model robust. Accuracy of the model will go up with low values of lambda whereas it will also tend to overfit with high variance. On the other hand with high values of lambda, the model will be simpler but at the cost of accuracy and will have a high bias.