Advanced Regression Assignment – Subjective Questions

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: The optimal value of alpha for ridge regression was 1.0 whereas for lasso regression, the optimal value of alpha came out to be 0.001.
On doubling the alpha values, the coefficient values would change slightly.

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 choose to apply the Lasso regression because the r2_scores for both the models are similar but Lasso would have an upper edge as RFE can be done easily in it, if we want to further optimize the model.

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: For the current model, **MSSubClass, BsmtCond, LotArea, LotShape, BsmtExposure***
Are the most important predictor variables.

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: We should pick the model which is relatively simple as it requires fewer datapoints in the training set and has less chances of overfitting. It should be noted that the model shouldn't be too simple. A model can be ensured to be robust and generalizable with regularization techniques.