Case Study: Advance Regression

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?

Answer 1:

For Lasso Regression: Optimal Value of Alpha: 0.01. For Ridge Regression: Optimal Value of Alpha: 2

If we double the value of alpha in Ridge Regression, then the model will apply the penalty to the curve and try to make the model more generalize and simple which in turn generates more error for the test and train dataset.

If we double the value of alpha in Lasso Regression, then the model get penalized which causes more coefficients to become 0, which causes the r-squared error to go up.

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? **Answer-2:**

Both these techniques are the regularization technique which helps in determining the prediction power for the target variables as to how much accurate the predicted value would be.

In Ridge Regression technique – it uses the hyperparameter called lambda as a penalty multifplied by the square of the magnitude of the coefficients which is identified as the cross validation.
The penalty is lambda times sum of squares of the coefficients, hence the coefficients that have greater value gets penalized.

As we increase the value of lambda the variance of model increases and bias remains constant.
Ridge Regression includes all the feature in the final model

In Lasso Regression technique – Penalty is absolute value of the magnitude of coefficents which is identified by cross validation.

As the lambda value increases Lasso shrinks the co-efficients towards 0. Lasso also does the feature selections.

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?
Answer-3

They are: GrLiveArea, OverallQual, OverallCond, TotalBsmtSF, GarageArea

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?
Answer-4
Based on the Occum's principle the model should be simpler.

For simple models the accuracy decreases but it will be more robust and generalizable.
It can also be understood by Bias-Variance Tradeoff.

Bias and Variance
We considered the example of a model memorising the entire training data set. If you change the data set slightly, this model will also need to change drastically. The model is, therefore, unstable and sensitive to changes in training data, and this is called high variance.

The 'variance' of a model is the variance in its output on some test data with respect to the changes in the training data. In other words, variance here refers to
the degree of changes in the model itself with respect to changes in the training data.

Bias quantifies how accurate the model is likely to be on future (test) data. Extremely simple models are likely to fail in predicting complex real-world phenomena. Simplicity has its own disadvantages.

Imagine solving digital image processing problems using simple linear regression when much more complex models such as neural networks are typically successful for such problems. We say that a linear model has a high bias because it is quite simple to be able to learn the complexity involved in the task.

Ideally, we want to reduce both bias and variance because the expected total error of a model is the sum of the errors in bias and variance,

Bias Variance Tradeoff
In practice, however, we often cannot have a model with a low bias and a low variance. As the model complexity increases, the bias reduces, whereas the variance increases and, hence, the trade-off