Assignment Part-II Questions

Que_1: What is the optimal value of alpha for ridge and lasso regression? What will be the changes in the model if you choose to 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 values of alpha for *Ridge Regression: 10, Lasso Regression: 0.001*

After doubled the values of alpha (Ridge), we can see that R2_Score of Train set decreases from 0.94 to 0.93. and followed by Test set 0.91 to 0.90.

After doubled the values of alpha (Lasso), we can see that R2_Score of Train set decreases from 0.94 to 0.92. and followed by Test set 0.91 to 0.90.

Que_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: Well, it depends on the situation that we are in.

- If our primary goal is feature selection, then we will use **Lasso**.
- If we don't want to get too large coefficients and reduction of coefficient magnitude is one of our prime goals, then we will use **Ridge Regression**.

Que_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: The top five features when we dropped existing top 5 features, “GrLivArea”, “SaleType_New”, “Neighborhood_Crawfor”, “OverallQual”, “Functional_Typ”.

Que_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: Basically, a model is called **Robust** when any deflection or changes in data does not affect the performance of model and in terms of its Accuracy. Another property generalizable, for that we need to make sure that our model does not **overfit.** In order to achieve robustness and generalizable model we need balance between Bias and variance tradeoff.