

## Internal

### 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:**

**We have applied MinMaxScaler before proceeding with Model Built. So after applying MinMaxScaler.**

- Optimal Value for Alpha with **Lasso as 0.0001**
- Optimal Value for Alpha with **Ridge as 5**
- If we **double the value of Alpha** for Ridge and Lasso, **test score is getting reduced.**
- After change is implemented, we found that **important predictor variables** for **Lasso** as ['GrLivArea', 'OverallQual', 'GarageCars', 'Neighborhood\_NoRidge', 'TotRmsAbvGrd', 'MSSubClass', 'Neighborhood\_StoneBr', 'Fireplaces', 'Neighborhood\_NridgHt', 'BsmtQual\_Gd', 'BsmtQual\_TA', 'BsmtExposure\_Gd', 'KitchenQual\_TA', 'FullBath', 'Exterior1st\_BrkFace'] and **Ridge** as ['OverallQual', '2ndFlrSF', 'GrLivArea', 'Neighborhood\_NoRidge', 'TotRmsAbvGrd', 'GarageCars', 'Neighborhood\_StoneBr', 'FullBath', 'BsmtQual\_TA', 'BsmtQual\_Gd', 'KitchenQual\_TA', 'Fireplaces', '1stFlrSF', 'Neighborhood\_NridgHt', 'KitchenQual\_Gd']

### 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:**

I will choose Lasso Regression with Alpha (Lambda) as 0.0001 and Feature count as 15 which has better test score 0.83 and having Low Variance where I got test score as 0.81 for Ridge regression.

### 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:**

As first five predictors in Lasso Regression 'GrLivArea', 'OverallQual', 'GarageCars', 'Neighborhood\_NoRidge', 'TotRmsAbvGrd', I will choose **next five predictor** variables as **'MSSubClass', 'Neighborhood\_StoneBr', 'Fireplaces', 'Neighborhood\_NridgHt'**, in Lasso Regression Model.

## Internal

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:**

As we observed in the assignment, we got High Training data Score for OLS model but score is decreased drastically on Test data which indicates that our OLS model was overfitting on training data. So we added penalty terms on the model by applying Lasso and Ridge Regularization techniques where model score is not changed much for Training data and Test Data. So our model became robust and generalizable with Regularization technique.