

### 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

I have scaled the data using RobustScaler and then applied Ridge and Lasso regression on the data.

The optimal values(alphas) for Lasso and Ridge regressions are 0.1 and 0.004 for Ridge and Lasso respectively.

With these alpha parameters, I was able to achieve a R2(R-Squared) of 0.91 and 0.87 for train and test data respectively.

### 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

I will choose Lasso regression because it is able to do feature selection and reduce the number of features.

Also, in this case, the difference between the R2 values for train and test data is less for Lasso regression as compared to 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 3

These are the top 5 most important predictor variables for Lasso regression:

LotArea, YearBuilt, OverallQual, GrLivArea, OverallCond, Type 1 finished square feet are the top 5 most important predictor variables for Lasso regression.

('LotArea', 0.2981),

('YearBuilt', 0.2889),

('OverallQual', 0.2874),

('LogGrLivArea', 0.2727),

('GrLivArea', 0.1475),

('OverallCond', 0.1177),

These are top 5 important predictor variables for Lasso regression after dropping the above features.

('2ndFlrSF', 0.4201),

('Log1stFlrSF', 0.2789),

('YearRemodAdd', 0.2279),  
( 'Functional\_Typ', 0.2124),  
( 'CentralAir\_Y', 0.1958),  
( 'TotalBsmtSF', 0.182),

#### 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

A model is robust and generalisable if it is able to perform well on unseen data. The variance may reduce and the bias may increase if the model is not robust and generalisable but its accuracy may be affected.