

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: Optimal alpha values for ridge and lasso regression are 200.0 and 200.0

If the value is doubled for ridge then it increases the regularization and reduces the coefficients. This affects the bias of model and makes it less complex.

On the other hand, Lasso regression can also reduce some coefficients to even 0 which makes the model even simpler.

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: Lasso regression would be the better choice as it automatically selects the most important predictors.

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: The new top 5 most important predictors are OverallCond, TotRmsAbvGrd, MasVnrArea, Functional, and Fireplaces.

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: To ensure the model is robust we can follow few steps such as by handling outliers, by removing irrelevant or highly correlated features. It needed to ensure that it does not overfit.

Implications of Model accuracy are a well regularized model will be more stable, it avoids overfitting. If the model is more generalized it will be more biased and less variance.