

Surprise Housing Assignment – Part II

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?

Ans: Optimal value for alpha is 0.001 for ridge and lasso regression. Upon doubling the value of alpha, the r^2 value decreases in ridge regression and becomes zero in the case of lasso regression. The most important predictor variable is **GrLivArea**

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?

I will choose ridge regression since it has better r^2 score

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?

TotalBsmtSF
1stFlrSF
2ndFlrSF
GarageArea
BsmtUnfSF

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?

Use RFE to choose the most important variable and use those variable in creating the models.