

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 Value of Ridge = 10

Optimal Value of Lasso = 0.001

After doubling the alpha value the training and test score for ridge had slight decrease whereas for Lasso we could see considerable decrease.

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:

We can choose Lasso because it has feature selection. We can remove more features and even then the score does not have huge difference.

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:

Top 5 features are MasVnrType\_Stone, HouseStyle\_2.5Unf, MasVnrType\_BrkFace, MasVnrType\_None, GrLivArea

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:

R<sup>2</sup> score must be  $> 0.70$

p-value must  $< 0.05$

VIF must  $< 5$