

Question 1

What is the optimal value of alpha for ridge and lasso regression? How does adjusting the value of alpha impact the models, specifically if you choose double the alpha for both ridge and lasso? Additionally, what predictor variables become most crucial after this adjustment?

Answer

For ridge regression, the optimal alpha value is determined to be 10, while for lasso regression, it stands at 1.

When doubling the alpha for both regressions, there is a noticeable impact on the rsquare values. For ridge regression, the rsquare decreases for the training set and remains constant for the test set. On the other hand, in lasso regression, both the train and test set rsquare values decrease.

Post-adjustment, the most influential predictor variables are as follows:

- OverallQual
- TotalBsmtSF
- Neighborhood
- Exterior1st
- LotArea
- CentralAir

Question 2

After determining the optimal lambda for ridge and lasso regression, which one will you choose to apply, and why?

Answer

I opt for Lasso regression over Ridge because of the numerous features involved. Removing features with lower coefficients seems fitting in managing the complexity of the model and preventing overfitting.

Question 3

Upon realizing that the five most important predictor variables in the lasso model are unavailable in the incoming data, you must create another model excluding these variables. What are the five most crucial predictor variables in this new scenario?

Answer

Excluding the five most important predictor variables from the lasso model results in a focus on the following:

OverallQual
LotArea
Neighborhood
Exterior1st
BsmtQual

Question 4

How can you ensure that a model is robust and generalizable? What implications does this have for the model's accuracy, and why?

Answer

Ensuring a model's robustness and generalizability involves achieving a high r-square value, indicating a good fit. While the model may not precisely produce a linear line, it closely aligns with actual values, avoiding overfitting and ensuring accuracy in predictions.