

Assignment-based Subjective Questions

1. What is the optimal value of alpha for ridge and lasso regression? What will be the changes in the model if you choose to 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 for ridge and lasso regression models are 0.9 and 0.001 respectively. If we choose to double the value of alpha for both, there will be more regularization and hence more penalty applied. The coefficients will be more reduced to zero. There might be the case of overfitting when you increase alpha from optimal value and thus it may be the case here.

The most important predictor variables are: MiscVal, BsmtHdHd Bath, LowQualFinSF, BsmtFullBath and HalfBath.

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?

Ans. I would use Lasso regression model as it will help eliminate some of the features that are less important and helps understand the model better and make predictions.

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?

Ans. GrLivArea, OverallQual, OverallCond, TotalBsmtSF, GarageArea.

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?

Ans. Simple technique to make model more robust is to make it simple and understandable. To make it more robust check the Bias Variance trade off graph. Too much importance should not be given to outliers to increase the accuracy instead eliminate such outliers to make the model fit better on the test data. Hence conducting a thorough EDA before building helps get rid of such outliers and makes our model more robust.