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?

Answer:

The optimal value of Alpha for Ridge and Lasso Regression is the point where the Variance is considerably decreased and Bias is increased within the permissible limit

If the Alpha is more doubled regression is more regularized which makes more co-efficient to near Zero or Zero. Bias will be increased and variance will be decreased

The more important predictors will be non-zero co-efficient parameters

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?

Answer:

For Ridge regression the chosen lambda value is 10, because the Negative mean Absolute error of test and train data is minimum at 10

For Lasso regression the chosen lambda value is 0.5, because the R2 value error of test and train data is high at 0.5

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?

Answer:

- 1.BsmtFinSF1
- 2. BsmtFinSF2
- 3. BsmtUnfSF
- 4. WoodDeckSF
- 5. ScreenPorch

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?

Answer:

Cross-Validation, Data Splitting, Feature Engineering, Regularization, Hyperparameter Tuning, Ensemble Methods, Data Preprocessing, Feature Scaling, Regular Monitoring can make the model robust and generalised

Accuracy of the model might decrease slightly but the variance will be decreased and model performing well on unknown data can be increased