

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: The optimal value of alpha for Ridge is 2 and for Lasso it is 0.001. With these alphas the R2 of the model was approximately 0.83. After doubling the alpha values in the Ridge and Lasso, the prediction accuracy remains around 0.78

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?

▪ The optimum lambda value in case of Ridge and Lasso is as follows:-
• Ridge - 2
• Lasso - 0.0001
▪ The Mean Squared Error in case of Ridge and Lasso are:
• Ridge - 0.0018396090787924262
• Lasso - 0.0018634152629407766
▪ The Mean Squared Error of both the models are almost same

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?

The five most important predictor variables in the current lasso model is:-
1. Total_sqr_footage
2. GarageArea
3. TotRmsAbvGrd
4. OverallCond
5. LotArea
We build a Lasso model in the Jupiter notebook after removing these attributes from the dataset. The R2 of the new model without the top 5 predictors drops to .73
The Mean Squared Error increases to 0.0028575670906482538

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?

As Per, Occam's Razor— given two models that show similar 'performance' in the finite training or test data, we should pick the one that makes fewer on the test data due to following reasons:-
▪ Simpler models are usually more 'generic' and are more widely applicable
▪ Simpler models require fewer training

samples for effective training than the more complex ones and hence are easier to train. ▪ Simpler models are more robust. o Complex models tend to change wildly with changes in the training data set o Simple models have low variance, high bias and complex models have low bias, high variance o Simpler models make more errors in the training set. Complex models lead to overfitting — they work very well for the training samples, fail miserably when applied to other test samples Therefore, to make the model more robust and generalizable, make the model simple but not simpler which will not be of any use.