

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?

- Lasso alpha value : {'alpha': 0.001}
- Ridge alpha value : {'alpha': 0.2}

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?

- Lasso alpha value : {'alpha': 0.001}
- Ridge alpha value : {'alpha': 0.2}

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?

After comparing both the model we can see that the below Features are best explaining the Dataset

- MSSubClass
- RoofMatl_Membran
- MSZoning_RL
- MSZoning_FV
- MSZoning_RH
- MSZoning_RM
- Condition2_PosA
- RoofMatl_WdShngl

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?

Implement the following ways to make your predictive model more resilient to outliers:
Should use a model that is immune to outsiders.

Tree-based models are usually not affected by outliers, whereas regression-based models are. If you are doing a predictive evaluation, try a non-parametric test instead of a parametric test.