

### 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 for ridge regression is 20 and for Lasso regression is 0.001
- After doubling the alpha values
  - After ridge regression, from R2\_square values for train and test data - 0.8776 626982958321 and 0.8676541782728279, model performs well on both train and test data
  - After Lasso regression, from R2\_square values for train and test data - 0.9071 and 0.85648, model performs well on train but not that much on test data when compared to ridge
- Most important predictor variables are 'LotFrontage', 'LotArea', 'MasVnrArea', 'BsmtFinSF1', 'TotalBsmtSF', '1stFlrSF', 'LowQualFinSF', 'GrLivArea', 'GarageArea', 'WoodDeckSF', 'OpenPorchSF', 'EnclosedPorch', 'OverallQual\_8', 'OverallQual\_9', 'OverallQual\_10'

### 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

From ridge regression lambda value is 20 and for Lasso regression it is 0.001. We can choose 0.001 as optimal lambda value since the model performs well on both train and test data and Lasso moved a few feature variables to zero.

### 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

A generalisable and robust model shouldn't be sensitive to changes in feature variables. A model should perform well on test data. Have to handle outliers during training the model since these outliers affect the accuracy of model predictions.