Assignment –II

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 – Optimal value of alpha in lasso regression is 0.33 and in ridge is 0.7 .Below are the list of important predictor variables after the change is implemented

```
17  YearBuilt  3.065154e-03
18  YearRemodAdd  2.668123e-03
60    GarageArea  3.203223e-04
44     GrLivArea  2.853395e-04
68   ScreenPorch  2.699255e-04
```

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?

Ans- We will be choosing lasso regression as it involves feature selection and it will create a simpler model. Higher the alpha more the regularization more the underfitting and Lower the alpha lesser the regularization more the overfitting

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 dropping 5 most important predictors and rebuilding model , the next 5 important predictors are

```
54    GarageYrBlt  0.003541
40       2ndFlrSF  0.000382
39       1stFlrSF  0.000334
60     WoodDeckSF  0.000244
34    TotalBsmtSF  0.000234
```

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?

We can make model robust and generalised by making it a simpler model .If the bias is high varience will be low and vice versa. We need to regularise the model with different alphas and choose the best way with the help of different metrics available like R2 Score .