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 to 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 λ = 20

Lasso λ = 0.001

Ridge:

- There's little decrease in the R^2 values (train: 0.00268 and test = 0.00763 after increasing)
- There's slight decrease in the MSE for train set (0.000393) and test set (0.0007802433334535425
- Top Predictors: '1stFlrSF', '2ndFlrSF', 'OverallQual', 'SaleCondition_Normal', 'Condition1_Norm' at rank 5.

Lasso:

- There's little decrease in the R^2 values (train: 0.00430 and test = 0.000943) after increasing
- There's slight decrease in the MSE for train set (0.00279) and test set (0.000446)
- TopPredictors'1stFlrSF','2ndFlrSF','OverallQual','SaleType_New', 'SaleCondition_Normal','Condition1_Norm','OverallCond' and 'Condition1_Norm' replaced by 'OverallCond' after λ update at 6th rank.

	Metric	Linear Regression	Ridge (λ = 20.0)	Lasso (λ = 0.001)
0	R2 Score (Train)	0.920159	0.917469	0.915108
1	R2 score (Test)	0.843716	0.851355	0.849445
2	RSS (Train)	13.546251	14.002608	14.403233
3	RSS (Test)	6.172395	5.870709	5.946124
4	MSE (Train)	0.108064	0.109869	0.111430
5	MSE (Test)	0.145891	0.142281	0.143192

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:

Ridge and Lasso models have very little difference in values of R-squared or MSE in both test and train datasets. As the R-Squared Value hasn't gone up after feature selection using the Lasso Model, it's evident that most predictors do impact the response variable. So, ridge is the better fit.

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:

Before drop:

```
'1stFlrSF','2ndFlrSF','OverallQual','SaleType_New','SaleCondition_Norma
l',
```

After drop:

```
'BsmtFinSF1', 'BsmtUnfSF', 'Neighborhood_Somerst', 'TotRmsAbvGrd', 'Ful
lBath'
```

From below table, we can see that RSS and MSE values are increased

	Lasso
BsmtFinSF1	0.071944
BsmtUnfSF	0.053934
Neighborhood_Somerst	0.053766
TotRmsAbvGrd	0.051647
FullBath	0.050405

	Metric	Lasso: λ = 0.002	Lasso: drop 5 var
0	R2 score (Train)	0.910800	0.879892
1	R2 score (Test)	0.848502	0.829830
2	RSS(Train)	15.134235	20.378201
3	RSS(Test)	5.983366	6.720798
4	MSE (Train)	0.114222	0.132542
5	MSE (Test)	0.143639	0.152234

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 simple model would usually have high bias and low variance whereas a complex model would have low bias and high variance. In either case, the total error would be high. We need is lowest total error i.e., low bias and high variance.

To achieve this, we use regularization where we add a penalty term to the model's cost function that shrinks the magnitude of the total coefficients towards 0, prevent ing the risk of overfitting.

If λ, the shrinkage penalty, is too small the model would remain overfit but if it's t oo high, it may end up with underfit model.

With regularization, may be the coefficients are more biased, but the variance of m odel may see a marked reduction. Training bias for significant reduction in varianc e.

Regularization does not improve the accuracy of on dataset, but it can improve the generalization performance, i.e., the performance on new, unseen data, which is ex actly we want.