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:

Optimal Value of alpha for ridge and lasso regression are:

Alpha for Ridge: 10

Alpha for Lasso: 0.001

When we double the value of alpha for ridge and lasso:

- The coefficients are reduced in case of Ridge and there are more coefficients becoming 0 in case of Lasso (eliminating more features)
- In case of Ridge: There is a slight increase in the mean squared error of train, for test it is almost the same and R2 value of train has a slight decrease whereas R2 value of test is almost same.
- In case of Lasso: There is a slight increase in the mean squared error and R2 value of train and test decreases slightly.
- Model is not changing drastically after alpha value is doubled.

After doubling the alpha value:

	Metric	Ridge Regression	Lasso Regression
0	R2 Score (Train)	0.941044	0.926097
1	R2 Score (Test)	0.867093	0.858539
2	RSS (Train)	8.431534	10.569148
3	RSS (Test)	11.135544	11.852265
4	MSE (Train)	0.091504	0.102448
5	MSE (Test)	0.160366	0.165446

Before doubling the alpha value:

	Metric	Ridge Regression	Lasso Regression
0	R2 Score (Train)	0.943965	0.934445
1	R2 Score (Test)	0.867650	0.864234
2	RSS (Train)	8.013780	9.375278
3	RSS (Test)	11.088907	11.375050
4	MSE (Train)	0.089208	0.096489
5	MSE (Test)	0.160030	0.162081

Below are the most important predictor after change is implemented:

For Ridge:

OverallQual	0.072855
GrLivArea	0.063309
Neighborhood_Crawfor	0.059235
YearBuilt	0.054681
OverallCond	0.054091
SaleCondition_Normal	0.047351
2ndFlrSF	0.047047
Exterior1st_BrkFace	0.043476
TotalBsmtSF	0.041362
Condition1_Norm	0.037899

For Lasso:

GrLivArea	0.126746
OverallQual	0.088333
YearBuilt	0.068988
OverallCond	0.054129
SaleType_New	0.044867
TotalBsmtSF	0.042504
Foundation_PConc	0.031494
Neighborhood_Crawfor	0.031107
FireplaceQu_Gd	0.029124
BsmtFinSF1	0.029028

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:

- Optimal Value of alpha for Ridge and Lasso regression are:
    - Lambda for Ridge: 10
    - Lambda for Lasso: 0.001
- R2 for Ridge and Lasso are:
    - Ridge: Train – 94.39, Test – 86.76
    - Lasso: Train – 93.44, Test – 86.42
- MSE for Ridge and Lasso are:
    - Ridge: Train – 0.00796, Test – 0.02561
    - Lasso: Train – 0.00931, Test – 0.02627

Looking at the above stats, it looks like both the models are doing a fairly good job, as the train R2 is in the range of 93-94.5 and test R2 is around 86.5 for the both the models. We can choose Lasso in this case, as it helps in feature elimination, which increases the model interpretation.

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:

Dropped the top 5 most important predictor variables (GrLivArea, OverallQual, SaleType_New, YearBuilt, Neighborhood_Crawfor) as per the Lasso model.

Got the below 5 most important variable on building the model again excluding the five most important predictor variables given above:

1. MSZoning_RL 0.326573
2. MSZoning_RH 0.322638
3. MSZoning_FV 0.316718
4. MSZoning_RM 0.260494
5. RoofStyle_Shed 0.201078

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:

- Simpler models are more generic and robust, although we have to compromise on accuracy.
- Complex models tend to memorise the data because of which there is a huge variance even with small change in training data and might not perform well on test data.
- There is a Bias-Variance trade-off, Simple models have low variance, high bias and complex models have low bias, high variance.
- It is important to find the balance in Bias and Variance, so that we make the model simple but not simpler. This can be achieved using Regularization.
- Regularization helps in managing the model complexity by shrinking the coefficients towards 0. It penalizes the model if it becomes more complex and helps in achieving optimum simpler model, by compromising a bit on Bias to reach a position where we have optimum Bias, Variance and minimum total error. This is known as Bias-Variance trade off.
- This point is known as Optimum Model Complexity, where Model is sufficiently simple to be generalisable and also sufficiently complex to be robust.

  As shown in the below graph, accuracy of the model is maintained by keeping the balance between Bias and Variance, with minimum total error.