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 Values of Alpha for Ridge regression is **2.1** and Lasso Regression is **120**. We do not see huge variation when after doubling the alpha values for both Ridge and Lasso regression. However, there are very slight changes in **Mean Squared Error** and **R2Score**.

	r2_score_train	r2_score_test	RSS_train	RSS_test	MSE_train	MSE_test
Linear_regression	0.767897	0.734106	1.427648e+12	8.126794e+11	37393.630329	43074.730975
Ridge_alpha1	0.764667	0.742112	1.447518e+12	7.882100e+11	37652.956425	42421.295103
Ridge_alpha2	0.759081	0.743120	1.481875e+12	7.851287e+11	38097.185064	42338.294349
Lasso_alpha1	0.761744	0.747248	1.465495e+12	7.725136e+11	37886.053629	41996.781256
Lasso_alpha2	0.752531	0.752366	1.522164e+12	7.568705e+11	38611.600133	41569.397349

Top 5 features before and after changing the alpha

Before change	After Change
KitchenQual_Fa	KitchenQual_Fa
Neighborhood_NoRidge	Neighborhood_NoRidge
BsmtQual_Fa	KitchenQual_TA
KitchenQual_TA	BsmtQual_Fa
Neighborhood_StoneBr	Neighborhood_StoneBr

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 Values of Alpha for Ridge regression is **2.1** and Lasso Regression is **120**.
With this Optimal Values of Alpha obtained, MSE and R2Score are almost similar (varied in 0.001 range – actual values in screenshot above).

As we have a very minute better test_R2Score for Lasso Regression and also Lasso shrinked coefficients of few variables to zero, which reduces number of features which in turn reduces model complexity.

With the above-mentioned advantages of Lasso over Ridge, we'll choose **Lasso Regression** for final modelling with Alpha value as 120.

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:

Below are the five top most features with alpha = 120 for Lasso Regression before deleting top 5 and after deleting top 5 features

Before Deleting top 5	After Deleting top 5, New top 5 features
KitchenQual_Fa	BsmtQual_None
Neighborhood_NoRidge	BsmtQual_TA
BsmtQual_Fa	KitchenQual_Gd
KitchenQual_TA	BsmtQual_Gd
Neighborhood_StoneBr	BldgType_Twnhs

Question-4

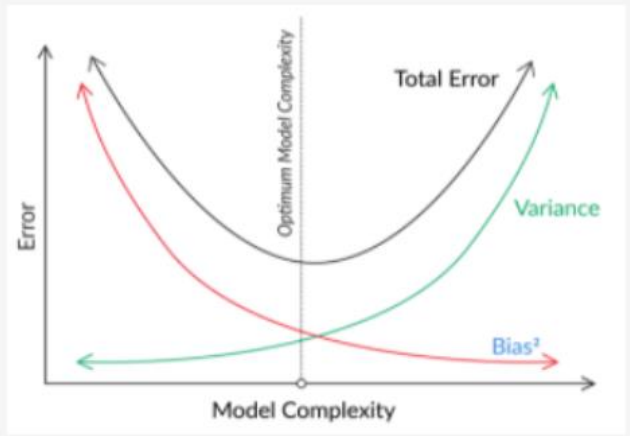
How can you make sure that a model is robust and generalisable? What are the implications of the same for the accuracy of the model and why?

Answer:

Per, Occam's Razor— given two models that show similar 'performance' in the finite training or test data, we should pick the one that makes fewer on the test data due to following reasons: -

- Simpler models are usually more 'generic' and are more widely applicable
- Simpler models require fewer training samples for effective training than the more complex ones and hence are easier to train.
- Simpler models are more robust.
- Complex models tend to change wildly with changes in the training data

Models will be more robust when it is Simple. As we see the image on the right side.



- If our model is too simple, it won't be able to explain more variance (i.e., model will fail to understand the pattern behind the data).
- If our model is too complex, instead of detecting the generalized pattern, it'll memorize each and every data point which we call "Overfitting".

To keep the balance, we can do Regularization to reduce the overfitting problem. A model which keeps the balance will be more robust and generalizable.