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?

Ridge and Lasso Regression works on same principle and they try to penalize beta coefficients to zero.

$$\sum_{i=1}^{M}(y_i - \hat{y_i})^2 = \sum_{i=1}^{M}\left(y_i - \sum_{j=0}^{p} w_j \times x_{ij}\right)^2 + \lambda \sum_{j=0}^{p} w_j^2$$

Cost function for ridge regression

Ridge regression puts constraint on the coefficients (w). The penalty term (lambda) regularizes the coefficients such that if the coefficients take large values the optimization function is penalized. Because of shrinking coefficients to zero and helps to reduce model complexity and multi collinearity. Lasso Regression helps to bring some coefficients to zero and thereby reduces over fitting.

For Comparing Alpha = 0,0.01 and 2 in Ridge Regression the value at 2 showed higher accuracy of 0.819 compared to 0.818 at alpha = 0.01 and 0.816 at alpha = 2

For Comparing Alpha = 0.01 and 100 in Lasso Regression the value at 0.01 showed higher accuracy of 0.819 compared to 0.818 at alpha = 100

Whenever double the value of alpha for Ridge over certain optimum value of alpha model starts under fitting the data thereby the value of r2 decreases and getting more error on training and test data

Whenever double the value of alpha for Lasso over certain optimum value of alpha model starts under fitting the data thereby the value of r2 decreases or even to zero and most of the coefficients value will decrease and getting more error on training and test data

MSZoning_FV , MSZoning_RL , Neighborhood_Crawfor , MSZoning_RH , MSZoning_RM are the most important variable after the changes has been implemented for ridge regression

GrLivArea, OverallQual, OverallCond, TotalBsmtSF, BsmtFinSF1 are the most important variable after the changes has been implemented for Lasso regression

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?

Lasso Regression would be chosen as optimal ones since it would help in feature elimination and the model will be more robust

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?

Those 5 most important predictor variables that will be excluded are :-
1. GrLivArea
2. OverallQual
3. OverallCond
4. TotalBsmtSF
5. GarageArea

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?

The model should be as simple as possible and it should also understand Bias- Variance Trade off. Whenever the model is simple in nature the bias will be more and variance will be very less. Accuracy of the model implications are the simpler and generalised model performs equally on training and test data.

Bias: Bias is error in model, when the model is weak to learn from the data. High bias means model is unable to learn details in the data. Model performs poor on training and testing data.

Variance: Variance is error in model, when model tries to over learn from the data. High variance means model performs exceptionally well on training data as it has very well trained on this of data but performs very poor on testing data as it was unseen data for the model.
It is important to have balance in Bias and Variance to avoid overfitting and under-fitting of data.