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?

Sol: **For lasso regression,** the first step to build a lasso model is to find the optimal lambda. For lasso regression, the alpha value is 1. The output is the best cross-validated lambda, which comes out to be 0.001. It determines the weighting to be used. In case of ridge regression, the value of alpha is zero.

When we double the value of alpha for ridge regression, we will take the value of 10^{th} model which will apply more penalty on the curve and try to make the model simplier and fit every data on the dataset. Similarly when we increase the value of alpha for lasso, we try to penalize more of our model coefficients of the variable to zero because when we increase the value, it decreases.

==========

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?

Sol: The model we will choose to apply will depend on the use case.If we have too many variables and one of our primary goal is feature selection, then we will use **Lasso**.If we don't want to get too large coefficients and reduction of coefficient magnitude is one of our prime goals, then we will use **Ridge Regression**.

==========

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?

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

Variance is error model. When model tries to over learn from data. High variance means model performs exceptionally well on training data as it has very well trained on this data but performs poor on testing data.

==========

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?

Sol:

- A model is **robust** when any variation in the data does not affect its performance much. **A generalizable** model is able to adapt properly to new, previously unseen data, drawn from the same distribution as the one used to create the model.
- In other words, the model should not be too complex in order to be robust and generalizable. To make sure a model is robust and generalizable, we have to **take care it doesn't overfit**.
- This is because an overfitting model has very high variance and a smallest change in data affects the model prediction heavily. Such a model will identify all the patterns of a training data, but fail to pick up the patterns in unseen test data.
- If we look at it from the prespective of **Accuracy**, a too complex model will have a very high accuracy. So, to make our model more robust and generalizable, we will have to decrease variance which will lead to some bias.
- Addition of bias means that accuracy will decrease. In general, we have to find strike some balance between model accuracy and complexity. This can be achieved by Regularization techniques like Ridge Regression and Lasso

==========