

Subjective Questions

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?

Ans:

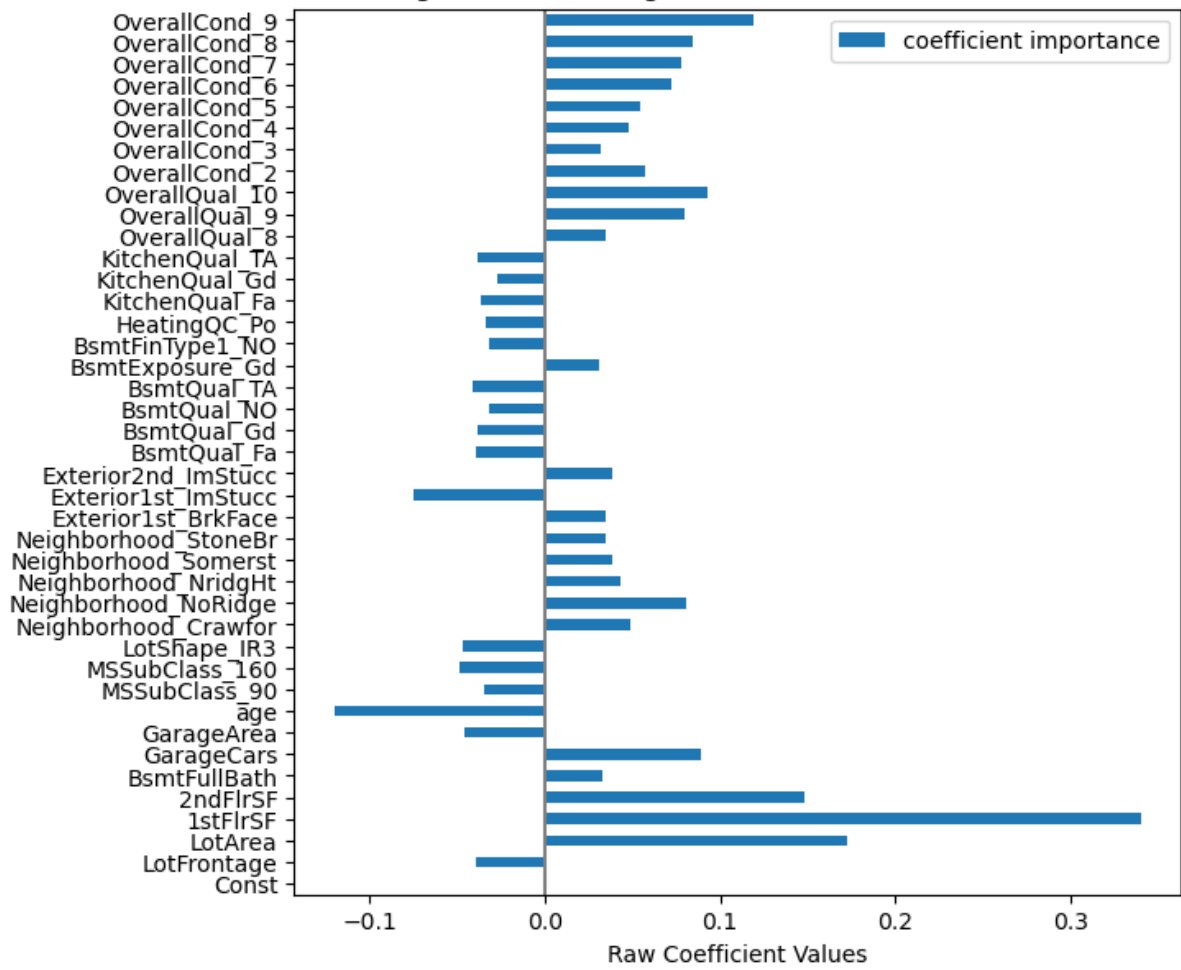
Alpha for ridge and lasso regression is 0.001

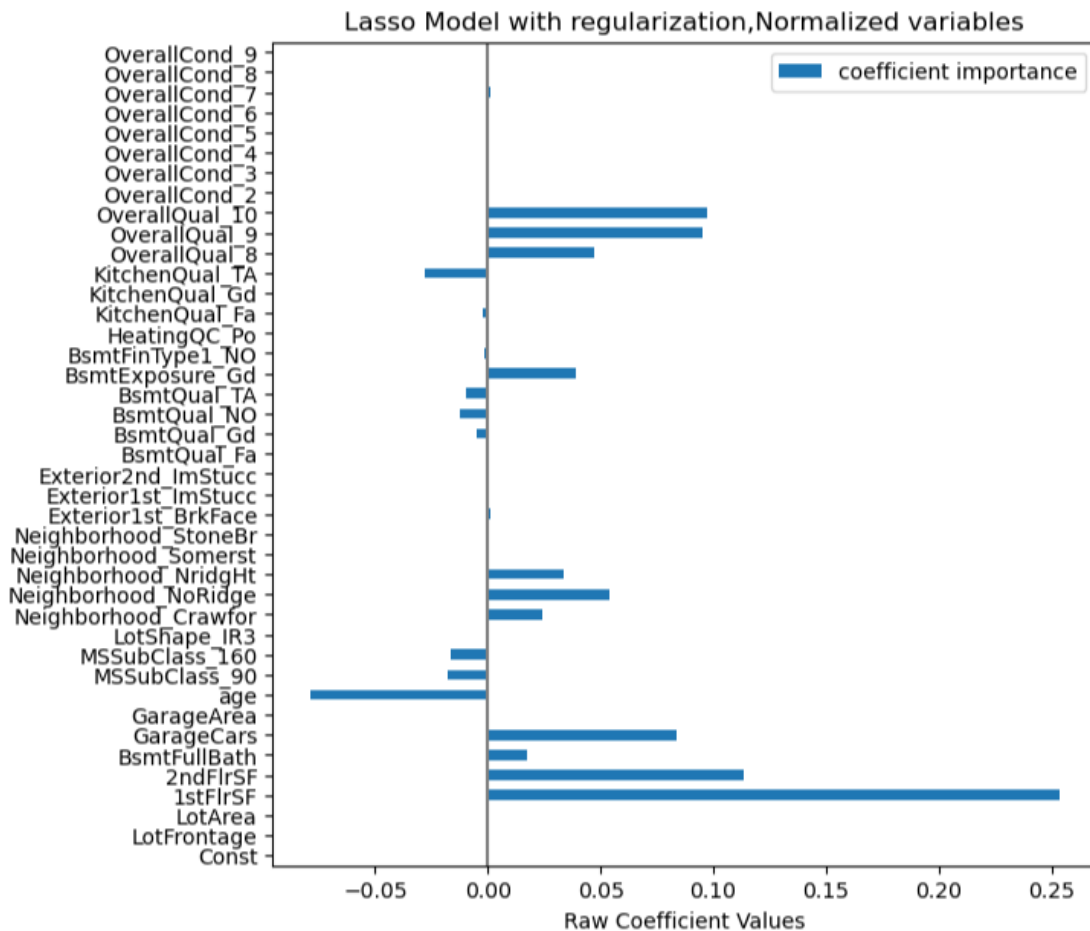
When alpha/lambda is doubled to 0.002, there is not much change in ridge coefficients but for lasso, most of feature coefficients are 0s when compared with previous values (which were near to +/-0.01..).

The most important predictor variables after the change is implemented:

"1stFlrSF"

Ridge Model with regularization, Normalized variables





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?

Ans:

Lambda/alpha chosen for both ridge and lasso regression are 0.001

Because r^2 scores and RMSE for both training and test sets are with less difference and coef. are minimal values.

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?

Ans:

Top 5 predictor variables before:

```
1stFlrSF      :0.253522
2ndFlrSF      :0.113366
OverallQual_10:0.097391
OverallQual_9 :0.094984
age           :-0.078841
```

After removing top 5 features:

LASSO

Lamda:0.001

Training Data:

r2score: 0.7047142382911329

RMSE:0.05965539055961897

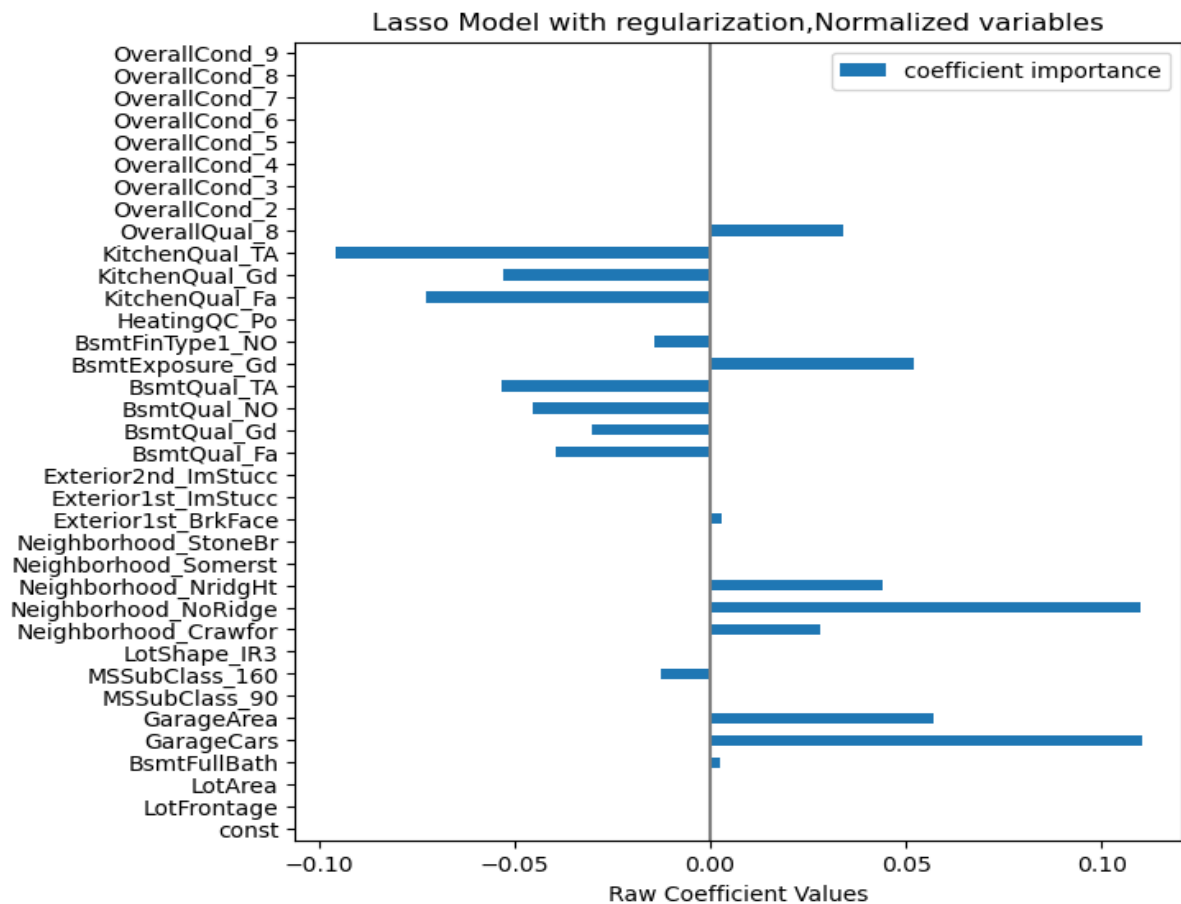
Testing Data:

r2score: 0.668442642034726

RMSE:0.06414678017773935

Top5 features after removal:

```
GarageCars      :0.11020291
Neighborhood_NoRidge :0.10989839
KitchenQual_TA   :-0.09570708
KitchenQual_Fa   :-0.072652
GarageArea       :0.05692849
```



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?

Ans:

In context of Lasso and Ridge regression models, we can make the model robust and generalisable by adjusting the hyperparameter (lambda or alpha).

Model will minimise the coefficient values of the features and sometimes coefficients will be zero or near to zero.

Thus, model will become robust and generalised.

The accuracy of the model will reduce (a little) when we are generalising by removing or minimising the impact of the features and need to avoid both overfitting and underfitting.