

PS9

Anais Ouedraogo

April 11, 2023

]

The training data has 75 variables and 404 observations. There are 61 more X variables than in the original housing data.

Using LASSO model the optimal value of lambda is 0.00569. In sample RMSE=0.175 and out of sample RMSE=0.180.

Using ridge regression, the optimal value of lambda is 0.00356. In sample RMSE=0.172 and out of sample RMSE=0.180.

It will not be possible to estimate a simple linear regression model on a data set that has more columns than rows. The model underfit the data