

Lasso and Model Order Selection. In-class Exercise 1 Solution

EL-GY 6143 Intro Machine Learning. Prof. Sundeep Rangan

Question

Complete the inclass exercise in demo2_housing.ipynb:

In-class exercise: Re-train the model with the variables with the largest 50 values in `coeff_ls`. What is the test R^2 ?

Solution

One possible solution is as follows

```
1 # TODO
2
3 # Sort the coefficients in descending magnitude
4 I = np.argsort(np.abs(coeff_ls))
5 I = np.flipud(I)
6 I = I[:50]
7
8 # Get the data matrix elements for the top 50 coefficients
9 Xtr_top = Xtr1[:,I]
10 Xts_top = Xts1[:,I]
11
12 # Fit on the training data
13 reg_top = LinearRegression()
14 reg_top.fit(Xtr_top, ytr1)
15
16 # Score on the test data
17 yhat1 = reg_top.predict(Xts_top)
18 rsq = r2_score(yts1, yhat1)
19
20 print(rsq)
```

0.8923472995633146