

Using LASSO to select features

6 questions

1
point

1.

We learn weights on the entire house dataset, using an L1 penalty of $1e10$ (or $5e2$, if using scikit-learn). Some features are transformations of inputs; see the reading.

Which of the following features have been chosen by LASSO, i.e. which features were assigned nonzero weights? (Choose all that apply)

☐ yr_renovated

☐ waterfront

☐ sqft_living

☐ grade

☐ floors

1
point

2.

We split the house sales dataset into training set, test set, and validation set and choose the l_1 penalty that minimizes the error on the validation set.

In which of the following ranges does the best l_1 penalty fall?

- ☐ Between 0 and 100
 - ☐ Between 100 and 1000
 - ☐ Between 1000 and 10000
 - ☐ Between 10000 and 100000
 - ☐ Greater than 100000
-

1
point

3.

Using the best value of l_1 penalty as mentioned in the previous question, how many nonzero weights do you have?

18

1
point

4.

We explore a wide range of l1_penalty values to find a narrow region of l1_penalty values where models are likely to have the desired number of non-zero weights (max_nonzeros=7).

What value did you find for l1_penalty_min?

If you are using GraphLab Create, enter your answer in simple decimals without commas (e.g. 1131000000), rounded to nearest millions.

If you are using scikit-learn, enter your answer in simple decimals without commas (e.g. 4313), rounded to nearest integer.

3793000000

1
point

5.

We then explore the narrow range of l1_penalty values between l1_penalty_min and l1_penalty_max.

What value of l1_penalty in our narrow range has the lowest RSS on the VALIDATION set and has sparsity equal to max_nonzeros?

If you are using GraphLab Create, enter your answer in simple decimals without commas (e.g. 1131000000), rounded to nearest millions.

If you are using scikit-learn, enter your answer in simple decimals without commas (e.g. 4342), rounded to nearest integer.

3449000000

1
point

6.

Consider the model learned with the l_1 penalty found in the previous question. Which of the following features has non-zero coefficients? (Choose all that apply)

- ☐ sqft_living
 - ☐ bedrooms_square
 - ☐ sqft_lot_sqrt
 - ☐ bathrooms
 - ☐ floors
-

Submit Quiz

