

Question 1

What is the result of the following code?

```
cross_val_predict(lr2e, x_data, y_data, cv=3)
```

1 / 1 point

Performs multiple out-of-sample evaluations

The predicted values of the test data using cross-validation

Calculates the free parameter alpha

The average R2 on the test data for each of the two folds

Correct

Correct! The method `cross_val_predict()` predicts values using cross-validation.

Question 2

How would you organize the values 1, 10, and 100 as possible values of alpha for Grid Search?

1 / 1 point

```
parameter = alpha(1,10,100)
```

```
parameter=[1,10,100]
```

```
parameter = [{'alpha': [1,10,100]}]
```

```
parameter = Ridge(alpha=[1,10,100])
```

Correct

Correct! This is the correct syntax to create the variable 'parameter' for Grid Search.

Question 3

You do the following steps with a data set:

Divide a data set into testing and training sets.

Create a linear model with the training set.

Find the average R2 value on your training data. It is found to be 0.5.

Perform a 100th-order polynomial transform on your data.

Use these transformed values to train another model.

Find the new value for R2. It is found to be 0.99.

Which of the following statements is correct?

1 / 1 point

You should use the simpler model

100-th order polynomial will work better on the rest of your data

You should use your test data to test the model further

Create another linear model with all of the data and compare results

Correct

Correct! The results of your training data are not the best indicator of how your model performs.

4.

Question 4

What is the purpose of “folding” your data sets? 1 / 1 point

Folds are used for cross-validation

To find R2 values on a training set and a test set of data

To find the actual predicted values of the model before calculating R2

Folding is used primarily for polynomial transformations

Correct

Correct! By creating folds, you iterate on your training and testing data using different combinations of the data set and compare results.

Question 5

In the following image, the blue curve represents a model, the blue dots represent the data, and the orange curve represents the true function. Which of the following is true about the model?

0 / 1 point

No conclusions can be drawn about the model

The model is a good fit

It displays underfitting

It displays overfitting