

Quiz: Model Refinement

1. What is the correct use of the "train_test_split" function such that 40% of the data samples will be utilized for testing, the parameter "random_state" is set to zero, and the input variables for the features and targets are x_data, y_data respectively?

1 point

- ☐ train_test_split(x_data, y_data, test_size=0, random_state=0.4)
- ☒ train_test_split(x_data, y_data, test_size=0.4, random_state=0)
- ☐ train_test_split(x_data, y_data)

2. What is the output of the following code?

1 point

```
1 cross_val_score(lre, x_data, y_data, cv=2)
```

- ☐ This function finds the free parameter alpha
- ☒ The average R^2 on the test data for each of the two folds
- ☐ The predicted values of the test data using cross-validation

3. What is the code to create a ridge regression object **RR** with an alpha term equal to 10?

1 point

- ☐

```
1 RR=LinearRegression(alpha=10)
```
- ☒

```
1 RR=Ridge(alpha=10)
```
- ☐

```
1 RR=Ridge(alpha=1)
```

4. What dictionary value would we use to perform a grid search for the following values of alpha? 1, 10, 100. No other parameter values should be tested

1 point

☐ 1 `alpha=[1,10,100]`

☒ 1 `[{'alpha': [1,10,100]}]`

☐ 1 `[{'alpha': [0.001,0.1,1, 10, 100, 1000,10000,100000,100000], 'normalize':[True, False]}]`

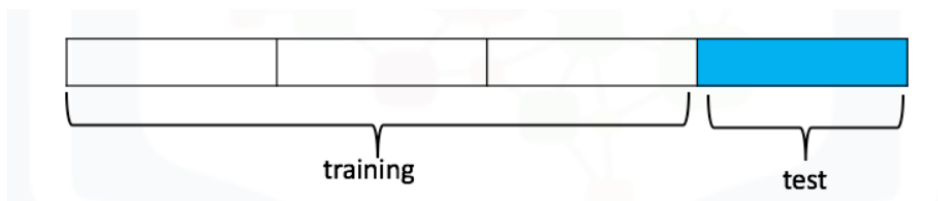
5. You train a ridge regression model, you get a R^2 of 1 on your training data and you get a R^2 of 0 on your validation data; what should you do?

1 point

- ☐ Nothing, your model performs flawlessly on your validation data
- ☐ Your model is under fitting; so perform a polynomial transform
- ☒ Your model is overfitting, so increase the parameter alpha

6. Consider the following diagram of 4 fold cross-validation. From the diagram how many folds are used for training?

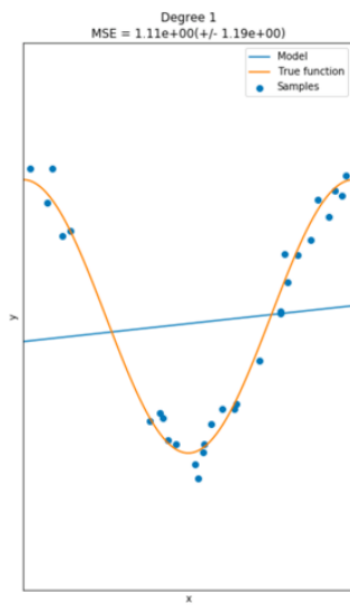
1 point



- ☒ 3
- ☐ 4
- ☐ 1

7. The following is an example of what?

1 point



- ☐ Overfitting
- ☐ Perfect fit
- ☒ Underfitting