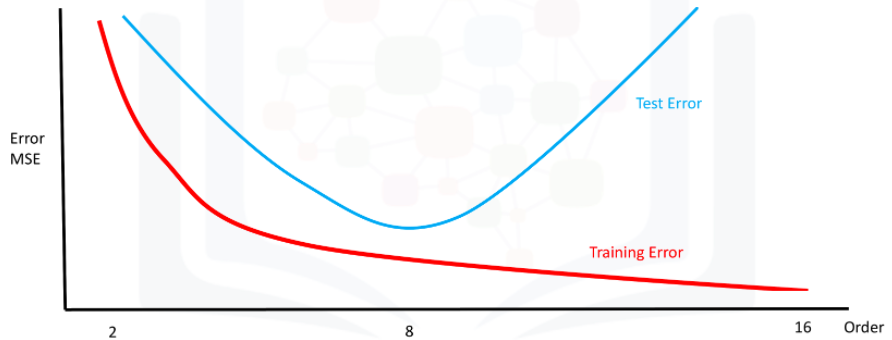
**Quiz: Model Refinement**

**TOTAL POINTS 7**

1.Question 1

In the following plot, the vertical axis shows the mean square error and the horizontal axis represents the order of the polynomial. The red line represents the training error and the blue line is the test error. What is the best order of the polynomial given the possible choices in the horizontal axis?





8



2



16

1 point

2.Question 2 x

What is the output of the following code?



1

cross\_val\_score(lre, x\_data, y\_data, cv=2)



The predicted values of the test data using cross-validation



This function finds the free parameter alpha



The average R^2 on the test data for each of the two folds

1 point

3.Question 3

What is the output of the following code?



1

cross\_val\_predict (lr2e, x\_data, y\_data, cv=3)



The predicted values of the test data using cross-validation



The average R^2 on the test data for each of the two folds



This function finds the free parameter alpha

1 point

4.Question 4 x

What dictionary value would we use to perform a grid search to determine if normalization should be used and for testing the following values of alpha? 1,10, 100





1

2

alpha=[1,10,100]

normalize=[True,False]





1

[{'alpha':[1,10,100],'normalize':[True,False]} ]





1

{'alpha': [1,10,100]}]

1 point

5.Question 5

You have a linear model; the average R^2 value on your training data is 0.5, you perform a 100th order polynomial transform on your data then use these values to train another model. After this step, your average R^2 is 0.99; which of the following comments is correct?



You should always use the simplest model



The results on your training data is not the best indicator of how your model performs; you should use your test data to get a better idea

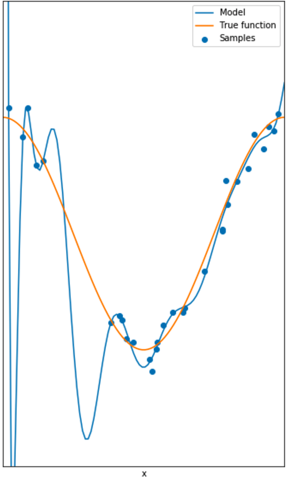


100-th order polynomial will work better on unseen data

1 point

6.Question 6

The following is an example of what?





Underfitting



Overfitting

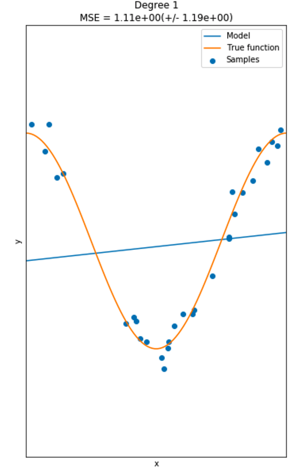


Perfect fit

1 point

7.Question 7

The following is an example of what?





Overfitting



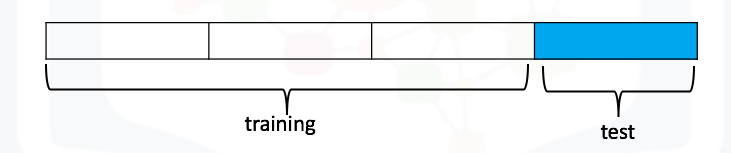
Perfect fit



Underfitting

1 point

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





1



4



3

1 point

9.

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?



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)

1 point

10.

What is the correct use of the "train\_test\_split" function such that 40 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?



train\_test\_split(x\_data, y\_data)



train\_test\_split(x\_data, y\_data, test\_size=0.40, random\_state=0)



train\_test\_split(x\_data, y\_data, test\_size=40, random\_state=0)

1 point

11. 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

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]} ]

1 point

12.

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?



Your model is under fitting; so perform a polynomial transform



Your model is overfitting, so increase the parameter alpha



Nothing, your model performs flawlessly on your validation data

1 point