Module 5 – Model Evaluation:

Question 1: In the following plot, the vertical access shows the mean square error andthe horizontal axis represents the order of the polynomial. The red line represents the training error the blue line is the test error. What is the best order of the polynomial given the possible choices in the horizontal axis?

- 2
- **8**
- 16

Question 2: What is the 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_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)

Question 3: What is the output of cross_val_score(lre, x_data, y_data, cv=2)?

- 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

Question 4: What is the code to create a ridge regression object "RR" with an alpha term equal 10

- RR=LinearRegression(alpha=10)
- **RR=Ridge(alpha=10)**
- RR=Ridge(alpha=1)

Question 5: 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

- alpha=[1,10,100]
- **[{'alpha': [1,10,100]}]**
- [{'alpha': [0.001,0.1,1, 10, 100, 1000,10000,100000,100000],'normalize':[True,False]}]