2.

Cross-validation aims to split the data into 5 groups: 4 training and 1 testing sets. We will iterate through the dataset and use each index as the testing set while training the rest of the index values. The training will produce a more accurate training and testing error parameter. Cross-validation allows us to use multiple predictions for a more accurate data prediction compared to traditional prediction methods that uses singular training values. Although cross-validation may result in extra computational steps than the traditional train-test split.

|  |  |
| --- | --- |
| Degree | Test Error |
| 1 |  |
| 2 |  |
| 3 |  |
| 4 |  |
| 5 |  |
| 6 |  |
| 7 |  |
| 8 |  |
| 9 |  |
| 10 |  |
| 11 |  |
| 12 |  |
| 13 |  |
| 14 |  |
| 15 |  |
| 16 |  |
| 17 |  |
| 18 |  |
| 19 |  |
| 20 |  |

Degree, Test Error

0, 0.10256410256410253

1, 0.1282051282051282

2, 0.08333333333333337

3, 0.10897435897435892

4, 0.08012820512820518

5, 0.09294871794871795

6, 0.09294871794871795

7, 0.11538461538461542

8, 0.13782051282051277

9, 0.07371794871794868

10, 0.08653846153846156

11, 0.09935897435897434

12, 0.08653846153846156

13, 0.13782051282051277

14, 0.08974358974358976

15, 0.10576923076923073

16, 0.1217948717948718

17, 0.1217948717948718

18, 0.08653846153846156

19, 0.08653846153846156

Runs, Degree

0,2.0

1,3.0

2,4.0

3,3.0

4,3.0

5,3.0

6,4.0

7,4.0

8,2.0

9,4.0

10,5.0

11,3.0

12,5.0

13,3.0

14,3.0

15,4.0

16,4.0

17,3.0

18,3.0

19,4.0

|  |  |  |
| --- | --- | --- |
|  | Test Errors | Best Degrees |
| Mean | 0.102 | 3.75 |
| Standard deviation | 0.0207 | 1.34 |