

Machine learning worksheet

1> B

2> C

3> B

4>C

5>D

6> B

7>D

8> B

9> A, B, C

10 > B, D

11>C

12> we can use batch gradient descent, stochastic gradient descent, or mini-batch gradient descent. SGD and MBGD would work the best because neither of them need to load the entire dataset into memory

13>

Feature scaling is required for the various gradient descent algorithms. Feature scaling will help gradient descent converge quicker. The normal equations method does not require normalizing the features, so it remains unaffected by features in the training set having very different scales.