MACHINE LEARNING WORKSHEET-1

1. B

2. C

3. B

4. C

5. D

6. A

7. C

8. C

9. A and B

10. A and D

11. C and D

12. Which Linear Regression training algorithm can we use if we have a training set with millions of features?

ANSWER: If you have a training set with millions of features you can use Stochastic Gradient Descent or Mini-batch Gradient Descent, and perhaps Batch Gradient Descent if the training set fits in memory. But you cannot use the Normal Equation because the computational complexity grows quickly (more than quadratically) with the number of features.

13. Which algorithms will not suffer or might suffer, if the features in training set have very different scales?

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