MACHINE LEARNING – WORKSHEET-1

Q1 to Q8

1. B) O(n)

2. B) Logistic Regression

3. B) Gradient Descent

4. C) Lasso

5. B) Mini-Batch Gradient Descent

6. A) True

7. C) it does not matter whether half is there or not.

8. C) Both of them

Q9 to Q11

9. A) We don't have to choose the learning rate.

   B) It becomes slow when number of features are very large.

   C) We need to iterate.

10.

11. C) It discovers causal relationship.

Q12 & Q13

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 needs to load the entire dataset into memory in order to take 1 step of gradient descent. Batch would be ok with the caveat that you have enough memory to load all the data. The normal equations method would not be a good choice because it is computationally inefficient. The main cause of the computational complexity comes from inverse operation on an (n x n) matrix.

13. 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.