MACHINE LEARNING

In Q1 to Q8, only one option is correct, Choose the correct option:

1. D
2. C
3. B
4. C
5. D
6. A
7. C
8. C

In Q9 to Q11, more than one options are correct, Choose all the correct options:

9. A, B
10. A, C
11. C, D

Q12 and Q13 are subjective answer type questions, Answer them briefly.

Q12. 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, it will be very, very computationally expensive Instead we can use Gradient Descent.

Q13. Which algorithms will not suffer or might suffer, if the features in training set have very different scales?

Answer: The Gradient Descent suffers from features of different scales, because the model will take a longer time to reach the global maximum. We can always scale the features to eliminate this problem.The normal equations method does not require normalizing the features, so it remains unaffected by features in the training set having very diffent scales. Features scaling is required for the various gradient descent algorithms. feature scaling will help gradient descent coverge quicker. The cost function will have the shape of an elongated bowl, so the Gradient Descent Algorithms will take a long to converge. To solve this you should scale the data before training the model.