Machine Learning

1. D
2. C
3. B
4. C
5. D
6. A
7. A
8. B
9. B,C
10. D
11. A,B
12. Which Linear Regression training algorithm can we use if we have a training set with millions of features?

Answer: If we have training set with millions of features we can use**Stochastic Gradient Descent or Mini-batch Gradient Descent**, and perhaps Batch Gradient Descent if the training set fits in memory.

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

Answer: If the features in training set have very different scales, the cost function will have the shape of an elongated bowl, so the Gradient Descent algorithms will take a long time to converge. To solve this you should scale the data before training the model. Note that the Normal Equation will work just fine without scaling.