machine learning

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

ANS: Since there are lots of features, we cannot use Normal Equations
(it will be very, very computationally expensive).
 Instead we can use Gradient Descent.

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

ANS: 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.