

## **MACHINE LEARNING WORKSHEET**

Q1) B

Q2) B

Q3) B

Q4) C

Q5) D

Q6) A

Q7) C

Q8) B

Q9) A and B

Q10) A and C

Q11) C and D

Q12) We can use stochastic gradient descent or mini-batch gradient descent as they do not need to load the whole dataset into memory for taking one step of gradient descent. We can also use batch gradient descent if we have enough memory to load the entire data.

Q13) If the features in your training set have very different scales, the gradient descent will take a long time to reach the global maximum as the cost function will have an elongated bowl shape. This can be solved by scaling the data.