MACHINE LEARNING

Answers of Q1 to Q8

1. **(B) $O(n)$**
2. **(C) Polynomial Regression**
3. **(B) Gradient Descent**
4. **(C) Lasso**
5. **(C) Batch Gradient Descent**
6. **(A) True**
7. **(A) Scaling cost function by half makes gradient descent converge faster**
8. **(B) Correlation**

Answers of 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.

10. **(A) Linear Regression will have high bias and low variance.**
    (C) Polynomial with degree 5 will have low bias and high variance.

11. **(C) It discovers causal relationship.**
    (D) No inference can be made from regression line.

12. **Linear Regression training algorithm we can use if we have a training set with millions of features are-**
    (a) Batch gradient descent
    (b) Stochastic gradient descent
    (c) Mini-batch gradient descent

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.**

Submitted By
Raushan Kumar
(Intern Data Science)