MACHINE LEARNING

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

1. The computational complexity of linear regression is:

B) $O(n)$

2. Which of the following can be used to fit non-linear data?

B) Logistic Regression

3. Which of the following can be used to optimize the cost function of Linear Regression?

B) Gradient Descent

4. Which of the following method does not have closed form solution for its coefficients?

C) Lasso

5. Which gradient descent algorithm always gives optimal solution?

A) Stochastic Gradient Descent

6. Generalization error measures how well a model performs on training data.

A) True

7. The cost function of linear regression can be given as $J(w0, w1) = 1/2m \sum (w0 + w1x(i) - y(i))^2$ $i=1$.

The half term at start is due to:

C) it does not matter whether half is there or not.

8. Which of the following will have symmetric relation between dependent variable and independent variable?

B) Correlation

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

9. Which of the following is true about Normal Equation used to compute the coefficient of the Linear Regression?

A) We don't have to choose the learning rate.

B) It becomes slow when number of features are very large.

C) We need to iterate.

D) It does not make use of dependent variable.

10. Which of the following statement/s are true if we generated data with the help of polynomial features with 5 degrees of freedom which perfectly fits the data?

A) Linear Regression will have high bias and low variance.

B) Linear Regression will have low bias and high variance.

C) Polynomial with degree 5 will have low bias and high variance.

D) Polynomial with degree 5 will have high bias and low variance.

11. Which of the following sentence is false regarding regression?

A) It relates inputs to outputs.

B) It is used for prediction.

C) It discovers causal relationship.

D) No inference can be made from regression line.

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

12. Which Linear Regression training algorithm can we use if we have a training set with millions of features?

Ans) We can use a Stochastic Gradient Descent, or a Mini-batch Gradient Descent and may be Batch Gradient Descent if the training set fits in memory.
Do not use the Normal Equation, because the computational complexity grows quickly.

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

Ans) 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, we should scale the data before training the model.

Note that the Normal Equation will work just fine without scaling