MACHINE LEARNING WORKSHEET - 1 (Answers)

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

1. The computational complexity of linear regression is:

Answer: B) (n)

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

Answer: B) Logistic Regression

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

Answer: B) Gradient Descent

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

Answer: C) Lasso

5. Which gradient descent algorithm always gives optimal solution?

A) Stochastic Gradient Descent
B) Mini-Batch Gradient Descent
C) Batch Gradient Descent
D) All of the above

Answer: D) All of the above

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

Answer: A) True

7. The cost function of linear regression can be given as $(w_{0,1}) = \frac{1}{2m}\sum(w_0 + w_1 x^{(i)} - y^{(i)})^2_{mi=1}$. The half term at start is due to:

Answer: A) scaling cost function by half makes gradient descent converge faster.

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

Answer: 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?

B) It becomes slow when number of features are very large. ,
D) It does not make use of dependent variable.

Answer: B , D

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

Answer: A, C

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

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

Answer: C, D

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?

Answer:

By using the "Batch gradient descent, stochastic gradient descent, or Mini-batch gradient descent" we can fix the millions of training set features.

Stochastic slope plunge & Mini-batch slope plunge would work the leading because neither of them should stack the whole dataset into memory to require 1 step of slope plunge. The group would be alright with the caveat of simply have sufficient memory to stack all the information.

The typical conditions strategy would not be a great choice since it is computationally wasteful. The most cause of the computational complexity comes from the reverse operation on an (n x n) matrix.

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

Answer:

The ordinary conditions strategy does not require normalizing the highlights, 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.