Assignment 39

Machine Learning

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

1. The computational complexity of linear regression is:
A) $O(n\ 2.4\ )$ B) $O(n)$ C) $O(n\ 2\ )$ D) $O(n\ 3\ )$

Answer is D

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

A) Lasso Regression B) Logistic Regression

C) Polynomial Regression D) Ridge Regression

Answer is C

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

A) Entropy B) Gradient Descent

C) Pasting D) None of the above.

Answer is B

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

A) extrapolation B) Ridge

C) Lasso D) Elastic Nets

Answer is C

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 is D

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

A) True B) False

Answer is D

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

The half term at start is due to:

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

B) presence of half makes it easy to do grid search.

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

D) None of the above.

Answer is A

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

A) Regression B) Correlation

C) Both of them D) None of these

Answer is C

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.

Answer is A,B

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.

Answer is A,D

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.

Answer is 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:

There are 3 algorithms – 1) Batch gradient descent 2) Stochastic gradient descent 3) Mini-batch gradient descent.

The best algorithm which can be used with the mentioned huge features are Stochastic gradient descent & Mini-batch gradient descent. As both the algorithms does not need to load the entire data.

Batch gradient descent is not recommended as it requires loads of memory to load all the data.

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

Scales

Answer:

The normal equation methods do not affect the training set which already have different scales, although Feature scale might suffer also the Gradient Descent algorithms will take a long time to converge.

Feature scales have different scales. So the data should be scaled accordingly before training the model.