Name: - Nidhi Charde

Batch No:-Internship 19

MACHINE LEARNING

1. The computational complexity of linear regression is:

A) (n 2.4)

B) (n)

C) (n 2)

D) (n 3)

Answer- B) (n)

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 – C) Polynomial Regression

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 – B) Gradient Descent.

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– 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 – C) Batch Gradient Descent

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

A) True

B) False

Answer – B) False

7. The cost function of linear regression can be given as ($w0$, $w1$) = 1 $2m$ ∑ ($w0$ + $w1x$ (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– 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?

A) Regression

B) Correlation

C) Both of them

D) None of these

Answer– B) Correlation

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

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

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

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

Answer – C) It discovers causal relationship.

D) No inference can be made from regression line.

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

Answer For Millions of Features, first we have to understand the behaviour and correlation between them. After that we can perform the VIF (variance Inflation Factor) or Dimensionally Reduction process PCA (Principle Component Analytics).

After that divide feature into the subset and using ensemble method

Random Forest and SVM should be use

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

Answer – The normal equations method does not require normalizing the feature, 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.