MACHINE LEARNING

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

1. In the linear regression equation $y = \theta_0 + \theta_1 X$, θ_0 is the:
   A) Slope of the line
   B) Independent variable
   C) y intercept
   D) Coefficient of determination

Ans:- **B) Independent variable**

2. True or False: Linear Regression is a supervised learning algorithm.
   A) **True**
   B) False

Ans:- (A) **True**

3. In regression analysis, the variable that is being predicted is:
   A) the independent variable
   B) the dependent variable
   C) usually denoted by x
   D) usually denoted by r

Ans:- **B) the dependent variable**

4. Generally, which of the following method(s) is used for predicting continuous dependent variables?
   A) Logistic Regression
   B) Linear Regression
   C) Both
   D) None of the above

Ans:- **B) Linear Regression**

5. The coefficient of determination is:
   A) the square root of the correlation coefficient
   B) usually less than zero
   C) the correlation coefficient squared
   D) equal to zero

Ans- C) the correlation coefficient squared

6. If the slope of the regression equation is positive, then:
   A) y decreases as x increases
   B) y increases as x increases
   C) y decreases as x decreases
   D) None of these

Ans:- **C) the correlation coefficient squared**

7. Linear Regression works best for:
   A) linear data
   B) non-linear data
   C) both linear and non-linear data
   D) None of the above

Ans:- C) both linear and non-linear data

8. The coefficient of determination can be in the range of:
   (A) **0 to 1**
   B) -1 to 1
   C) -1 to 0
   D) 0 to infinity

Ans:- (A) **0 to 1**

In Q9 to Q13, more than one options are correct, Choose all the correct options:

9. Which of the following evaluation metrics can be used for linear regression?
   A) Classification Report
   B) RMSE
   C) ROC curve
   D) MAE

Ans- **B) RMSE**

10. Which of the following is true for linear regression?
    A) **Linear regression is a supervised learning algorithm.**
    B) Linear regression supports multi-collinearity.
    C) Shape of linear regression's cost function is convex.
    D) Linear regression is used to predict discrete dependent variable.

Ans:- **Linear regression is a supervised learning algorithm.**

MACHINE LEARNING

11. Which of the following regularizations can be applied to linear regression?
    A) **Ridge** B) **Lasso**
    C) Pruning D) Elastic Net

Ans- **A)Ridge**
B) Lasso

12. Linear regression performs better for:
    A) Large amount of training samples with small number of features.
    B) Same number of features and training samples
    C) Large number of features
    D) The variables which are drawn independently, identically distributed

13. Which of the following assumptions are true for linear regression?
    (A) Linearity **B) Homoscedasticity**
    C) Non-Independent D) Normality

Ans:- **B) Homoscedasticity**

MACHINE LEARNING

Q14 and Q15 are subjective answer type questions, Answer them briefly.

14. Explain Linear Regression?

Ans:- Linear regression analysis is used to predict the value of a variable based on the value of another variable. The variable you want to predict is called the dependent variable. The variable you are using to predict the other variable's value is called the independent variable. It is a data analysis technique that predicts the value of unknown data by using another related and known data value. It mathematically models the unknown or dependent variable and the known or independent variable as a linear equation. In statistics, a regression equation (or function) is linear when it is linear in the parameters. While the equation must be linear in the parameters, you can transform the predictor variables in ways that produce curvature. This form of analysis estimates the coefficients of the linear equation, involving one or more independent variables that best predict the value of the dependent variable. Linear regression fits a straight line or surface that minimizes the discrepancies between predicted and actual output values. There are simple linear regression calculators that use a “least squares” method to discover the best-fit line for a set of paired data. You then estimate the value of X (dependent variable) from Y (independent variable).

15. What is difference between simple linear and multiple linear regression?

Ans:- Simple linear regression has only one x and one y variable. Multiple linear regression has one y and two or more x variables. For instance, when we predict rent based on square feet alone that is simple linear regression. Regression as a tool helps pool data together to help people and companies make informed decisions. There are different variables at play in regression, including a dependent variable—the main variable that you're trying to understand—and an independent variable—factors that may have an impact on the dependent variable.

In order to make regression analysis work, you must collect all the relevant data. It can be presented on a graph, with an x-axis and a y-axis.

There are several main reasons people use regression analysis:

To predict future economic conditions, trends, or values

To determine the relationship between two or more variables

To understand how one variable change when another changes

There are many different kinds of regression analysis. For the purpose of this article, we will look at two: linear regression and multiple regression.