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:
   C) y intercept
2. True or False: Linear Regression is a supervised learning algorithm.
   A) True
3. In regression analysis, the variable that is being predicted is:
   A) the independent variable
4. Generally, which of the following method(s) is used for predicting continuous dependent variables?
   B) Linear Regression
5. The coefficient of determination is:
   C) the correlation coefficient squared
6. If the slope of the regression equation is positive, then:
   B) y increases as x increases
7. Linear Regression works best for:
   A) linear data
8. The coefficient of determination can be in the range of
   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?
   B) **RMSE** **D) MAE**
10. Which of the following is true for linear regression?
    A) Linear regression is a supervised learning algorithm.
    C) Shape of linear regression's cost function is convex.
11. Which of the following regularizations can be applied to linear regression?
    A) Ridge **B) Lasso**
    D) Elastic Net
12. Linear regression performs better for:
    A) Large amount of training samples with small number of features.
13. Which of the following assumptions are true for linear regression?
    A)Linearity
    B) Homoscedasticity
    D) Normality

MACHINE LEARNING

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

14. Explain Linear Regression?

Ans: Linear regression is a supervised machine learning algorithm that is used to model the relationship between a dependent variable (also called the response variable) and one or more independent variables (also called the explanatory variables or predictors). The goal of linear regression is to find the line of best fit that describes the relationship between the variables.

The line of best fit is determined by finding the values of the coefficients (also called weights or parameters) that minimize the sum of the squared errors between the predicted values of the dependent variable and the actual values. The coefficients represent the slope and intercept of the line.

The equation for linear regression can be written as:

$y = \beta0 + \beta1x1 + \beta2x2 + \ldots + \beta n*xn$

where y is the dependent variable, x1, x2, …, xn are the independent variables, β0 is the intercept, and β1, β2, …, βn are the coefficients for the independent variables.

Linear regression can be used for both simple linear regression, which involves only one independent variable, and multiple linear regression, which involves two or more independent variables.

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

Ans: Simple linear regression and multiple linear regression are two types of linear regression models used to model the relationship between a dependent variable and one or more independent variables.

Simple linear regression involves only one independent variable, and the equation for the model is:

$y = \beta0 + \beta1*x + \varepsilon$

where y is the dependent variable, x is the independent variable, β0 is the intercept, β1 is the coefficient for the independent variable, and ε is the error term.

Multiple linear regression, on the other hand, involves two or more independent variables, and the equation for the model is:

$y = \beta0 + \beta1x1 + \beta2x2 + \ldots + \beta n*xn + \varepsilon$

MACHINE LEARNING

where y is the dependent variable, x1, x2, …, xn are the independent variables, β0 is the intercept, and β1, β2, …, βn are the coefficients for the independent variables.

The main difference between the two is the number of independent variables used in the model. Simple linear regression is used when there is only one independent variable and a linear relationship is assumed between the dependent variable and the independent variable. Multiple linear regression is used when there are two or more independent variables and a linear relationship is assumed between the dependent variable and the independent variables.

In terms of interpretation, simple linear regression produces a straight line on a two-dimensional plot, while multiple linear regression produces a hyperplane in a higher-dimensional space. Additionally, multiple linear regression is typically more complex than simple linear regression and requires more data to train the model accurately