

**MACHINE LEARNING**

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

1. In the linear regression equation  $y = \theta_0 + \theta_1 x$ ,  $\theta_0$  is the:

Ans:

B) Independent variable

2. True or False: Linear Regression is a supervised learning algorithm.

Ans:

A) True

3. In regression analysis, the variable that is being predicted is:

Ans:

B) the dependent variable

4. Generally, which of the following method(s) is used for predicting continuous dependent variables?

Ans:

B) Linear Regression

5. The coefficient of determination is:

Ans:

D) equal to zero

6. If the slope of the regression equation is positive, then:

Ans:

B) y increases as x increases

7. Linear Regression works best for:

Ans:

D) None of the above

8. The coefficient of determination can be in the range of:

Ans:

B) -1 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?

Ans:

B) RMSE

D) MAE

---

## MACHINE LEARNING

10. Which of the following is true for linear regression?

Ans:

- A) Linear regression is a supervised learning algorithm.
- D) Linear regression is used to predict discrete dependent variable.

11. Which of the following regularizations can be applied to linear regression?

Ans:

- A) Ridge
- B) Lasso

12. Linear regression performs better for:

Ans:

- A) Large amount of training samples with small number of features.
- B) Same number of features and training samples

13. Which of the following assumptions are true for linear regression?

Ans:

- A) Linearity
- B) Homoscedasticity

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

14. Explain Linear Regression?

Ans:

- Linear regression is a statistical technique that is used to model the relationship between a dependent variable and one or more independent variables. The goal of linear regression is to find the line of best fit that best describes the relationship between the dependent variable and the independent variables.
- In a simple linear regression model, there is only one independent variable. The line of best fit is represented by the equation:

$$y = \beta_0 + \beta_1 x$$

where  $y$  is the dependent variable,  $x$  is the independent variable,  $\beta_0$  is the intercept, and  $\beta_1$  is the regression coefficient that represents the slope of the line.

- For example, let's say you want to model the relationship between the number of hours studied ( $x$ ) and the test scores ( $y$ ) of a group of students. In this case, the dependent variable is the test score ( $y$ ), and the independent variable is the number of hours studied ( $x$ ). You could use linear regression to find the line of best fit that best describes the relationship between test scores and hours studied.

The results of the regression analysis could provide a prediction equation of the form:

$$y = \beta_0 + \beta_1 x$$

## **MACHINE LEARNING**

where  $\beta_0$  represents the average test score for students who have studied 0 hours, and  $\beta_1$  represents the expected change in test scores for each additional hour of study.

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

Ans:

- Simple linear regression is regression where there is only one linear independent variable, and the highest exponent applied to the single independent variable is 1 (which is often omitted since anything to the power of 1 is just itself). For example, when only considering temperature when analyzing rate of melting of ice, the only independent variable is the temperature, and it would be linear (a straight line).
  - Multiple linear regression is where there is more than one independent variable. It could be 2, 3, 4, 5, 6, 10, 100, 10000, a trillion... hypothetically, there is no limit to the number of independent variables, so long as the number is at least two. All of the exponents are 1 (which is often omitted since anything to the power of 1 is just itself). For example, when considering both temperature and pressure when analyzing rate of melting of ice, the only independent variable is the temperature, and it would be linear (a straight line).
-