

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: y intercept(option C)
2. True or False: Linear Regression is a supervised learning algorithm.

A) True
B) False

Ans: True (option A)
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: the dependent variable(option B)
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: Linear Regression(option-B)
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: the correlation coefficient squared D) equal to zero(option c)
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: y increases as x increases(option-B)
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: Linear Data (option A)
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:- 0 to 1 (option A)

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: RMSE and MAE (option B and D)
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: A)Linear regression is a supervised learning algorithm B) Linear regression supports multi-collinearity. C) Shape of linear regression's cost function is convex.(option A,B,C are correct)

MACHINE LEARNING

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

- A) Ridge
- B) Lasso
- C) Pruning
- D) Elastic Net

Ans: ridge and lasso(option A and B)

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

Ans: Large amount of training samples with small number of features.(option A) Large number of features (option c)

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

- A) Linearity
- B) Homoscedasticity
- C) Non-Independent
- D) Normality

Ans: Linearity and Non-independent (option A and C)

MACHINE LEARNING

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

14. Explain Linear Regression?

Ans: Linear regression is a statistical technique used to establish a relationship between a dependent variable and one or more independent variables. In its simplest form, linear regression involves finding the best-fit line through a set of data points. This line can then be used to make predictions about the values of the dependent variable based on the values of the independent variables.

In a linear regression model, the dependent variable is typically denoted as "y", while the independent variable(s) are denoted as "x". The equation for a simple linear regression model can be written as:

$$y = a + bx$$

Where "a" represents the intercept, or the point at which the line intersects the y-axis, and "b" represents the slope of the line. The slope indicates how much y changes for every unit change in x.

In a multiple linear regression model, there are multiple independent variables, and the equation is expanded to include a coefficient for each variable:

$$y = a + b_1x_1 + b_2x_2 + \dots + b_nx_n$$

The goal of linear regression is to find the values of "a" and the "b"s that minimize the difference between the predicted values of y and the actual values of y for a given set of data. This is typically done using a method called "least squares," which involves minimizing the sum of the squared differences between the predicted and actual values of y.

Once the line of best fit has been established, it can be used to make predictions about the values of y for new values of x. This makes linear regression a useful tool for forecasting and trend analysis in many different fields, including economics, finance, and science

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

Ans: The main difference between simple linear regression and multiple linear regression is the

number of independent variables used in the model.

In simple linear regression, there is only one independent variable, denoted as "x", and a single

dependent variable, denoted as "y". The goal of the model is to establish a linear relationship

between x and y. The equation for a simple linear regression model is:

$$y = a + bx$$

Where "a" is the intercept, and "b" is the slope of the line.

In multiple linear regression, there are two or more independent variables, denoted as x_1 , x_2 , x_3 ,

etc., and a single dependent variable, denoted as "y". The goal of the model is to establish a linear

relationship between the independent variables and the dependent variable. The equation for a

multiple linear regression model is:

$$y = a + b_1x_1 + b_2x_2 + b_3x_3 + \dots$$

MACHINE LEARNING

Where "a" is the intercept, and "b1", "b2", "b3", etc. are the coefficients that represent the effect of
each independent variable on the dependent variable
