

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:-(C)
2. True or False: Linear Regression is a supervised learning algorithm.
A) True
B) False
Ans:-(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:-(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:-(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:-(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:-(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:-(B)
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)

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

MACHINE LEARNING

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:-(A)

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

Ans:-(A,B)

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

14. Explain Linear Regression?

Ans:- Linear Regression is a statistical method used to find a relationship between a dependent variable (Y) and one or more independent variables (X). It is a supervised learning algorithm used to predict continuous numeric values.

The main goal of linear regression is to find the best fit line that can represent the relationship between the dependent and independent variables. This line is represented by the equation:

$$Y = \theta_0 + \theta_1 X_1 + \theta_2 X_2 + \dots + \theta_n X_n$$

where Y is the dependent variable, X_1, X_2, \dots, X_n are independent variables, θ_0 is the intercept, and $\theta_1, \theta_2, \dots, \theta_n$ are coefficients for independent variables.

The linear regression algorithm calculates the best fit line by minimizing the difference between the predicted values and the actual values of the dependent variable. The difference is measured using a cost function, such as the mean squared error (MSE). Linear regression can be used for both simple and multiple linear regression problems. Simple linear regression involves only one independent variable, while multiple linear regression involves two or more independent variables.

Linear regression also assumes that the relationship between the independent and dependent variables is linear, that the errors are normally distributed, and that the variance of the errors is constant (homoscedasticity). If these assumptions are not met, the results of the linear regression may not be accurate.

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

Ans:- Simple linear regression is a statistical method that is used to establish a linear relationship between two continuous variables, where one variable is considered as the independent variable (x) and the other variable is considered as the dependent variable (y). The goal of simple linear regression is to find the best-fitting line that describes the relationship between the two variables. This line is represented by the equation $y = \beta_0 + \beta_1 x$, where β_0 is the intercept, and β_1 is the slope.

Multiple linear regression, on the other hand, is used to model the relationship between a dependent variable and two or more independent variables. The goal of multiple linear

MACHINE LEARNING

regression is to find the best-fitting linear equation that describes the relationship between the dependent variable and the independent variables. The equation for multiple linear regression can be represented as $y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \dots + \beta_nx_n$, where x_1, x_2, \dots, x_n are the independent variables and $\beta_1, \beta_2, \dots, \beta_n$ are the coefficients of the independent variables.

In simple linear regression, there is only one independent variable, whereas in multiple linear regression, there are two or more independent variables. Additionally, the interpretation of the coefficients is different in both types of regression. In simple linear regression, the coefficient represents the change in the dependent variable for a unit change in the independent variable. In multiple linear regression, the coefficient represents the change in the dependent variable for a unit change in the corresponding independent variable, holding all other independent variables constant.

MACHINE LEARNING