**HRITIK YADAV**

**INTERNSHIP=34**

**WORKSHEET=3**

**MACHINE LEARNING**

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

1. In the linear regression equation *y = θ0 + θ1x*, *θ0* is the:

A) Slope of the line B) Independent variable

C) y intercept D) Coefficient of determination

ANSWER:-C) y intercept

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

A) True B) False

ANSWER:-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

ANSWER:-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

ANSWER:-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

ANSWER:-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

ANSWER:-B) y increases as x increases

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

ANSWER:-A) 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

ANSWER:-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

ANSWER:- B) RMSE

1. MAE

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.

Answer:-A) Linear regression is a supervised learning algorithm.

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

A) Ridge B) Lasso

C) Pruning D) Elastic Net

Answer:-A) Ridge , B) Lasso and D) Elastic Net

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

ANSWER:-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

C) Non-Independent D) Normality

ANSWER:-A) Linearity AND B) Homoscedasticity

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

1. Explain Linear Regression?

ANSWER:-Linear Regression is a supervised learning algorithm used to model the relationship between one or more independent variables (also known as features) and a continuous dependent variable (also known as target). The goal of linear regression is to find the best linear relationship between the independent variables and the dependent variable that can explain the variation in the target variable.

In its simplest form, linear regression assumes a linear relationship between the independent and dependent variables, represented by the equation:

y = β0 + β1\*x

where y is the dependent variable, x is the independent variable, β0 is the intercept, and β1 is the slope of the line.

The slope β1 represents the change in y per unit change in x, and the intercept β0 represents the value of y when x is equal to zero.

The goal of linear regression is to find the best values of β0 and β1 that minimize the sum of the squared differences between the predicted values of y and the actual values of y. This is typically done using a method called ordinary least squares (OLS) regression.

Linear regression can be extended to multiple linear regression, which involves modeling the relationship between multiple independent variables and the dependent variable. The equation for multiple linear regression is:

y = β0 + β1x1 + β2x2 + ... + βn\*xn

where x1, x2, ..., xn are the n independent variables, and β1, β2, ..., βn are the corresponding coefficients. The goal of multiple linear regression is to find the best values of the coefficients that minimize the sum of the squared differences between the predicted values of y and the actual values of y.

Linear regression is widely used in various fields such as finance, economics, engineering, and social sciences to analyze and model relationships between variables.

窗体底端

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

ANSWER:- Simple linear regression and multiple linear regression are both techniques used in linear regression analysis. The main difference between them is the number of independent variables used to predict the dependent variable.

Simple linear regression involves only one independent variable, whereas multiple linear regression involves two or more independent variables. In other words, simple linear regression models the relationship between a dependent variable and a single independent variable, while multiple linear regression models the relationship between a dependent variable and multiple independent variables.

In simple linear regression, the equation for the regression line is:

y = β0 + β1x

where y is the dependent variable, x is the independent variable, β0 is the intercept, and β1 is the slope of the line.

In multiple linear regression, the equation for the regression line is:

y = β0 + β1x1 + β2x2 + ... + βnxn

where x1, x2, ..., xn are the n independent variables, and β1, β2, ..., βn are the corresponding coefficients.

Simple linear regression is useful when there is only one independent variable that is thought to have a linear relationship with the dependent variable. It is commonly used in fields such as economics and finance to model the relationship between two variables, such as the relationship between housing prices and square footage.

Multiple linear regression is useful when there are multiple independent variables that are thought to have a linear relationship with the dependent variable. It is commonly used in fields such as social sciences and engineering to model more complex relationships between variables, such as the relationship between a person's age, income, and education level on their likelihood of developing a certain disease.

窗体底端