MACHINE LEARNING WORKSHEET – 3

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**

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

A) True B) False

Answer : **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

Answer : **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

Answer : **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

Answer : **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

Answer : **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

Answer : **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

Answer : **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

Answer : **B,C,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.

Answer : **A,C**

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

A) Ridge B) Lasso

C) Pruning D) Elastic Net

Answer : **A,B,D**

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**

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

A) Linearity B) Homoscedasticity

C) Non-Independent D) Normality

Answer : **A,B,D**

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

14. Explain Linear Regression?

Answer :

Linear regression is a type of supervised learning algorithm used for predicting a continuous dependent variable based on one or more independent variables. It assumes a linear relationship between the dependent variable and the independent variables, and aims to find the best fit line that can describe the relationship between the two.

The simplest form of linear regression, known as simple linear regression, involves only one independent variable. The equation for simple linear regression can be expressed as:

$y = b0 + b1 * x$

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

Answer :

The key 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 (X) that is used to predict the dependent variable (Y). The relationship between X and Y is assumed to be linear and is represented by a straight line. The equation for simple linear regression is $Y = \beta0 + \beta1*X$, where $\beta0$ is the intercept, $\beta1$ is the slope, X is the independent variable, and Y is the dependent variable.

In multiple linear regression, there are multiple independent variables (X1, X2, X3, etc.) that are used to predict the dependent variable (Y). The relationship between the independent variables and the dependent variable is still assumed to be linear, but the equation is more complex than simple linear regression