MACHINE LEARNING ASSIGNMENT - 10

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

Ans: C) y intercept

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

Ans: 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

Ans: 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

Ans: 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

Ans: A) the square root of the correlation coefficient

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) 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

Ans: 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

Ans: 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

Ans: B) RMSE -Root Mean Square Error (RMSE)

D) MAE -Mean Absolute Error (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.

Ans: A) Linear regression is a supervised learning algorithm.

C) Shape of linear regression's cost function is convex.

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

Ans: A) Ridge B) Lasso D) Elastic Net

Ridge - L2 Regularization

Lasso - L1 Regularization

Elastic Net – combines both L1 and L2.

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) Large amount of training samples with small number of features

D) The variables which are drawn independently, identically distributed

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

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

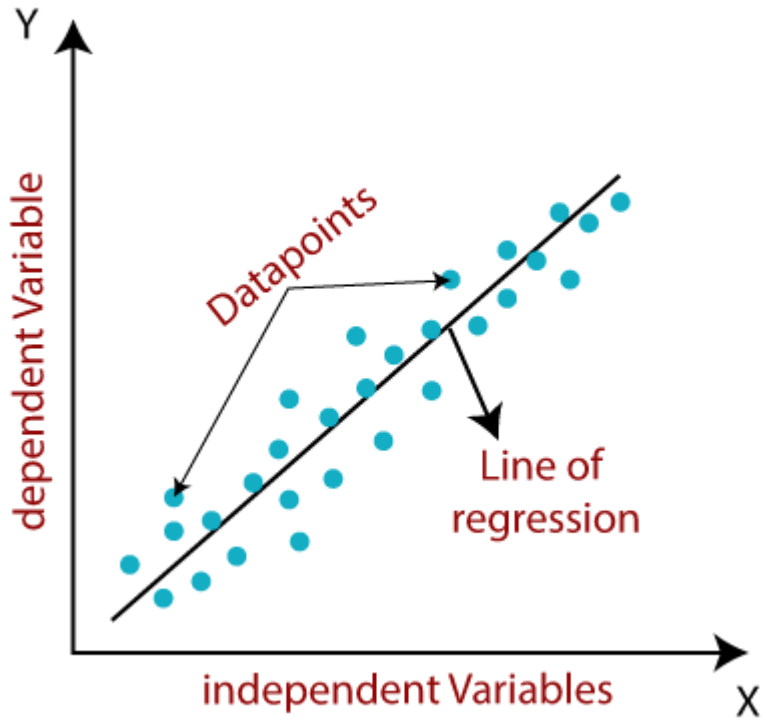
Ans: A) Linearity B) Homoscedasticity D) Normality

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

14. Explain Linear Regression?

Ans: Linear regression algorithm **shows a linear relationship between a dependent (y) and one or more independent (y) variables**, hence called as linear regression.

Linear regression is one of the easiest and most p image:



Mathematically, we can represent a linear regression as: $y= a_0+a_1x+ \varepsilon$

pular Machine Learning algorithms. It is a statistical method that is used for predictive analysis. Linear regression makes predictions for continuous/real or numeric variables such as **sales, salary, age, product price,** etc.

Linear regression algorithm shows a linear relationship between a dependent (y) and one or more independent (y) variables, hence called as linear regression. Since linear regression shows the linear relationship, which means it finds how

the value of the dependent variable is changing according to the value of the independent variable.

The linear regression model provides a sloped straight line representing the relationship between the variables. Consider the below image:

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

Ans: Simple linear regression has only one x and one y variable.

Multiple linear regression has one y and two or more x variables.

1. Simple linear regression comprises one independent and one dependent variable, whereas, multiple linear regression comprises two or more independent variables and one dependent variable.

2. It is assumed that for any specific set of values of the independent variable, the regression equation is associated with a random error, e. The random errors are normally distributed with mean 0 and standard deviation, σ. Finally, the random errors are independent of each other.

3. Find the general formula of the multiple regression as shown below:

$\hat{y}=b0+b1x1+b2x2+...+bkxk$

where k denotes the number of predictor variables, b0,b1,...bk denote the coefficients of the independent variables x1,x2,..,xk

4. The multiple coefficient of determination, R2 determines the total variation explained jointly by the independent variables. The Multiple R determines the strength and magnitude of the linear relationship of the dependent and the independent variables in the study.