

## MACHINE LEARNING

**In Q1 to Q11, only one option is correct, choose the correct option:**

1. Which of the following methods do we use to find the best fit line for data in Linear Regression?  
A) Least Square Error
2. Which of the following statement is true about outliers in linear regression?  
A) Linear regression is sensitive to outliers
3. A line falls from left to right if a slope is \_\_\_\_\_?  
B) Negative
4. Which of the following will have symmetric relation between dependent variable and independent variable?  
B) Correlation
5. Which of the following is the reason for over fitting condition?  
B) Low bias and low variance
6. If output involves label then that model is called as:  
B) Predictive modal
7. Lasso and Ridge regression techniques belong to \_\_\_\_\_?  
C) Regularization
8. To overcome with imbalance dataset which technique can be used?  
D) SMOTE
9. The AUC Receiver Operator Characteristic (AUCROC) curve is an evaluation metric for binary classification problems. It uses \_\_\_\_\_ to make graph?  
A) TPR and FPR
10. In AUC Receiver Operator Characteristic (AUCROC) curve for the better model area under the curve should be less.  
A) True
11. Pick the feature extraction from below:  
B) Apply PCA to project high dimensional data

**In Q12, more than one options are correct, choose all the correct options:**

12. Which of the following is true about Normal Equation used to compute the coefficient of the Linear Regression?  
A) We don't have to choose the learning rate.  
B) It becomes slow when number of features is very large.  
C) We need to iterate.
-

**MACHINE LEARNING**

**Q13 and Q15 are subjective answer type questions, Answer them briefly.**

13. Explain the term regularization?
  14. Which particular algorithms are used for regularization?
  15. Explain the term error present in linear regression equation?
-

## **MACHINE LEARNING**

### 1. Explain the term regularization?

Regularization is the technique which can be used to adjust the Machine learning model in way that it will reduce errors and the chances of overfitted and underfitted model. There are two techniques in Regularization, Ridge and Lasso Regularization.

1. Ridge Regularization – It adds the squared magnitude of the coefficient as the penalty term to the loss function.
2. Lasso Regularization – It adds the absolute value of magnitude of the coefficient as a penalty term to the loss function.

### 2. Which particular algorithms are used for regularization?

There are two algorithms which can be used for regularization, Ridge and Lasso Regularization.

1. Ridge Regularization – It is also called L2 norm or Ridge regression. When using this technique, we add the sum of weight's square to a loss function.

$$\text{Formula} = w_1[f_1]^2 + w_2[f_2]^2 + \dots + w_n[F_n]^2$$

2. Lasso Regularization – It is also called L1 norm or Lasso regression. It uses absolute weight values for normalization.

$$\text{Formula} = w_1[f_1] + w_2[f_2] + \dots + w_n[F_n]$$

### 3. Explain the term error present in linear regression equation?

The term error present in linear regression represents the difference between the actual output/label and input/features.

E.g: There is a model to predict the sales for specific area and there is difference between the expected sales at a particular area and the sales that was actually occurred.

---