

## Objective Questions

**1. Which of the following methods do we use to find the best fit line for data in Linear Regression?**

- A) Least Square Error
  - B) Maximum Likelihood
  - C) Logarithmic Loss
  - D) Both A and B
- Answer:** A) Least Square Error

**2. Which of the following statement is true about outliers in linear regression?**

- A) Linear regression is sensitive to outliers
  - B) Linear regression is not sensitive to outliers
  - C) Can't say
  - D) None of these
- Answer:** A) Linear regression is sensitive to outliers

**3. A line falls from left to right if a slope is \_\_\_\_\_?**

- A) Positive
  - B) Negative
  - C) Zero
  - D) Undefined
- Answer:** B) Negative

**4. Which of the following will have symmetric relation between dependent variable and independent variable?**

- A) Regression
  - B) Correlation
  - C) Both of them
  - D) None of these
- Answer:** B) Correlation

**5. Which of the following is the reason for overfitting condition?**

- A) High bias and high variance
  - B) Low bias and low variance
  - C) Low bias and high variance
  - D) None of these
- Answer:** C) Low bias and high variance

**6. If output involves labels, then that model is called as:**

- A) Descriptive model
  - B) Predictive model
  - C) Reinforcement learning
  - D) All of the above
- Answer:** B) Predictive model

**7. Lasso and Ridge regression techniques belong to \_\_\_\_\_?**

- A) Cross-validation
- B) Removing outliers
- D) Regularization
- C) SMOTE

**Answer:** D) Regularization

**8. To overcome an imbalanced dataset which technique can be used?**

- A) Cross-validation
- B) Regularization
- C) Kernel
- D) SMOTE

**Answer:** D) SMOTE

**9. The AUC Receiver Operator Characteristic (AUCROC) curve is an evaluation metric for binary classification problems. It uses \_\_\_\_\_ to make a graph?**

- A) TPR and FPR
- B) Sensitivity and precision
- C) Sensitivity and Specificity
- D) Recall and precision

**Answer:** A) TPR and FPR

**10. In the AUC Receiver Operator Characteristic (AUCROC) curve, for the better model area under the curve should be less.**

- A) True
- B) False

**Answer:** B) False

**11. Pick the feature extraction method from below:**

- A) Constructing a bag of words from an email
- B) Apply PCA to project high-dimensional data
- C) Removing stop words
- D) Forward selection

**Answer:** B) Apply PCA to project high-dimensional data

---

## **Multi-Select Question**

**12. Which of the following is true about the Normal Equation used to compute the coefficient of Linear Regression? (Select all that apply)**

- A) We don't have to choose the learning rate.
- B) It becomes slow when the number of features is very large.
- C) We need to iterate.
- D) It does not make use of the dependent variable.

**Answer:** A) We don't have to choose the learning rate, B) It becomes slow when the number of features is very large.

---

## Subjective Questions

### 13. Explain the term regularization.

Regularization is a technique used in machine learning to prevent overfitting by adding a penalty term to the cost function. This penalty discourages overly complex models and helps to ensure that the model generalizes well to new data. Common types of regularization include L1 (Lasso) and L2 (Ridge).

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

The main algorithms used for regularization are Ridge Regression, which uses L2 regularization, and Lasso Regression, which uses L1 regularization. Elastic Net is another algorithm that combines both L1 and L2 regularization.

### 15. Explain the term error present in the linear regression equation.

In linear regression, the error refers to the difference between the actual values and the predicted values generated by the model. This error, also known as the residual, indicates how well the model's predictions align with the actual data. Reducing error is crucial for improving the accuracy of the linear regression model.