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

    o **A) Least Square Error**
    (Least Squares Error is the common method used to find the best fit line.)

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

    o **A) Linear regression is sensitive to outliers**
    (Outliers can significantly affect the performance of linear regression models.)

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

    o **B) Negative**
    (A negative slope indicates that the line falls as you move from left to right.)

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

    o **B) Correlation**
    (Correlation measures the strength and direction of a linear relationship and is symmetric.)

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

    o **C) Low bias and high variance**
    (Overfitting typically occurs when a model has low bias but high variance.)

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

    o **B) Predictive model**
    (Predictive models are used when the output is a label or prediction.)

7. **Lasso and Ridge regression techniques belong to _____?**

    o **D) Regularization**
    (Both Lasso and Ridge regression are techniques used for regularization.)

8. **To overcome with imbalance dataset which technique can be used?**

    o **D) SMOTE**
    (SMOTE is a technique used to handle imbalanced datasets by generating synthetic samples.)

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

    o **A) TPR and FPR**
    (The ROC curve plots the True Positive Rate (TPR) against the False Positive Rate (FPR).)

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

- o **B) False**
  (A higher AUC indicates a better model. The area under the curve should be larger for a better model.)

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

    - o **B) Apply PCA to project high-dimensional data**
      (Principal Component Analysis (PCA) is a method for feature extraction.)

Multiple Choice Questions:

12. **Which of the following is true about the Normal Equation used to compute the coefficient of the Linear Regression?**

    - o **A) We don't have to choose the learning rate.**

    - o **B) It becomes slow when the number of features is very large.**

    - o (The Normal Equation does not require choosing a learning rate and can become computationally expensive with many features.)

    - o **C) We need to iterate.**
      (This is not true for the Normal Equation, as it provides a direct solution without iteration.)

Subjective Answer Type Questions:

13. **Explain the term regularization?**

    - o Regularization is a technique used in machine learning to prevent overfitting by adding a penalty to the loss function. This penalty discourages the model from fitting the noise in the training data, thereby improving its generalization to unseen data. Common regularization techniques include Lasso (L1 regularization) and Ridge (L2 regularization).

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

    - o Algorithms that use regularization include:

        - ▪ **Ridge Regression (L2 Regularization):** Adds a penalty proportional to the square of the magnitude of coefficients.

        - ▪ **Lasso Regression (L1 Regularization):** Adds a penalty proportional to the absolute value of the magnitude of coefficients.

        - ▪ **Elastic Net:** Combines both L1 and L2 regularization penalties.

15. **Explain the term error present in linear regression equation?**

    - o In linear regression, the term "error" refers to the difference between the predicted values and the actual observed values of the dependent variable. It can be represented as the residuals, which are the vertical distances between the data points and the regression line. The goal of linear regression is to minimize these errors, usually by minimizing the sum of the squared residuals (Least Squares Error).