Q1. Which of the following methods do we use to find the best fit line for data in Linear Regression?

- **Correct Answer: A) Least Square Error**

Q2. Which of the following statement is true about outliers in linear regression?

- **Correct Answer: A) Linear regression is sensitive to outliers**

Q3. A line falls from left to right if a slope is _____?

- **Correct Answer: B) Negative**

Q4. Which of the following will have a symmetric relation between dependent variable and independent variable?

- **Correct Answer: B) Correlation**

Q5. Which of the following is the reason for overfitting condition?

- **Correct Answer: C) Low bias and high variance**

Q6. If output involves a label then that model is called as:

- **Correct Answer: B) Predictive model**

Q7. Lasso and Ridge regression techniques belong to _____?

- **Correct Answer: D) Regularization**

Q8. To overcome an imbalanced dataset which technique can be used?

- **Correct Answer: D) SMOTE**

Q9. The AUC Receiver Operator Characteristic (AUCROC) curve is an evaluation metric for binary classification problems. It uses _____ to make a graph?

- **Correct Answer: A) TPR and FPR**

Q10. In AUC Receiver Operator Characteristic (AUCROC) curve for the better model area under the curve should be less.

- **Correct Answer: B) False**

Q11. Pick the feature extraction from below:

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

Q12. Which of the following is true about the Normal Equation used to compute the coefficient of the Linear Regression?

- **Correct Answers:**
  - A) We don’t have to choose the learning rate.
  - B) It becomes slow when the number of features is very large.

Subjective Questions

Q13. Explain the term regularization?

- **Answer:**
  - Regularization is a technique used in machine learning to prevent overfitting by adding a penalty to the model's loss function for having large coefficients. This ensures the model generalizes well to new data by discouraging overly complex models that fit noise in the training data. Regularization helps maintain balance by keeping the model simpler and more robust. Two common types of regularization are L1 (Lasso) and L2 (Ridge).

Q14. Which particular algorithms are used for regularization?

- **Answer:**
  - **Lasso Regression:** Uses L1 regularization, which adds an absolute value penalty to the loss function, potentially setting some coefficients to zero, thus performing feature selection.
  - **Ridge Regression:** Uses L2 regularization, adding a squared value penalty to the loss function, which reduces the magnitude of coefficients but doesn't eliminate them completely.
  - **Elastic Net:** Combines L1 and L2 penalties to benefit from both regularization techniques.

Q15. Explain the term error present in the linear regression equation?

- **Answer:**
  - The error in a linear regression equation is the difference between the actual values and the predicted values made by the model. This error is often called "residuals" and is crucial for understanding how well the model fits the data. In mathematical terms, it can be expressed as:
    Error (Residual)=Actual Value−Predicted Value\text{Error (Residual)} = \text{Actual Value} - \text{Predicted Value}Error (Residual)=Actual Value−Predicted Value
  - The goal of linear regression is to minimize these errors, typically using techniques like the least squares method, where the sum of the squared differences between actual and predicted values is minimized. Errors can be further categorized into:
    - **Bias Error:** The error due to the model's assumptions being too simplistic, leading to a failure in capturing the true relationship.
    - **Variance Error:** The error resulting from the model being too complex and sensitive to fluctuations in the training data.

- Together, bias and variance contribute to the overall error in the model's predictions. Balancing these errors is key to building effective models.