**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?

**C) Low bias and high variance**

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

**B) Predictive model**

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

**D) 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.

**B) False**

11. Pick the feature extraction from below:

**A) Construction bag of words from an email**

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?

**Both A & B**

**A) We don’t have to choose the learning rate.**

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

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

13. Explain the term regularization?

**In Machine Learning, Regularization is a technique used to prevent models from overfitting by penalizing the complexity of the model. The goal of regularization is to encourage simpler models that generalize well to unseen data. Regularization helps to balance model complexity and performance, contributing to more reliable predictions across different datasets.**

14. Which particular algorithms are used for regularization?

**Most common regularization techniques are L1 (Lasso), L2 (Ridge), and Elastic Net regularization.**

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

**The term "error" refers to the discrepancy between the predicted values of the dependent variable (outcome variable) and the actual observed values. This discrepancy arises because the linear regression model attempts to approximate the relationship between the independent variables (predictors) and the dependent variable, but it cannot perfectly capture all the variability in the data.**