# 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?
Ans. A) Least Square Error

2. Which of the following statement is true about outliers in linear regression?
Ans. A) Linear regression is sensitive to outliers

3. A line falls from left to right if a slope is _____?
Ans. B) Negative

4. Which of the following will have symmetric relation between dependent variable and independent variable?
Ans. B) Correlation

5. Which of the following is the reason for over fitting condition?
Ans. C) Low bias and high variance

6. If output involves label then that model is called as:
Ans. A) Descriptive model

7. Lasso and Ridge regression techniques belong to _____?
Ans. D) Regularization

8. To overcome with imbalance dataset which technique can be used?
Ans. D) SMOTE

9. The AUC Receiver Operator Characteristic (AUCROC) curve is an evaluation metric for binary classification problems. It uses _____ to make graph?
Ans. A) TPR and FPR

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

11. Pick the feature extraction from below:
Ans. 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.
D) It does not make use of dependent variable.
Ans. options A,B and D are correct.

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

13. Explain the term regularization?
Ans. Regularization is a technique used in machine learning and statistical modeling to prevent overfitting and improve the generalization of a model. Overfitting occurs when a model learns the training data too well, capturing noise or random fluctuations that may not be present in new, unseen data. Regularization introduces a penalty term to the model's objective function, discouraging overly complex models with large coefficients.

14. Which particular algorithms are used for regularization?
Ans. There are two common types of regularization used:
1.Lasso Regression (L1 Regularization)
2.Ridge Regression (L2 Regularization)

15. Explain the term error present in linear regression equation?
Ans. In linear regression, the term "error" means the difference between the predicted values of the dependent variable based on the linear regression equation and the actual observed values in the dataset.