# MACHINE LEARNING

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: B) Predictive modal

7. Lasso and Ridge regression techniques belong to _____?

Ans: B) Removing outliers

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

Ans: B) Regularization

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

Ans: C) Sensitivity and Specificity

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

Ans: A) True

11. Pick the feature extraction from below:

Ans: C) Removing stop words

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

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

13. Explain the term regularization?

Ans: The word regularize means to make things regular or acceptable. This is exactly why we use it for. Regularizations are techniques used to reduce the error by fitting a function appropriately on the given training set and avoid overfitting.

14. Which particular algorithms are used for regularization?

Ans: There are three algorithms are used regularization.

1. Ridge Regression
2. Lasso
3. Dropout

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

Ans: An error term represents the margin of error within a statistical model; it refers to the sum of the deviations within the regression line, which provides an explanation for the difference between the theoretical value of the model and the actual observed results.