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: A) Regression.

5.Which of the following is the reason for over fitting condition?

Ans: A) High Variance and High bias.

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

Ans: B) Predictive model.

7. Lasso and Ridge regression techniques belong to _____?

Ans: D) Regularization

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

Ans: C) 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: True

11. Pick the feature extraction from below:
A) Construction bag of words from a email
B) Apply PCA to project high dimensional data
C) Removing stop words
D) Forward selection

Ans: A, B, C.

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

Ans:1,2,3.

13. REGULARIZATION:

Regularization is a technique used to reduce the errors by fitting the function appropriately on the given training set and avoid overfitting.
The commonly used regularization techniques are :

1. L1 regularization
2. L2 regularization
3. Dropout regularization

14. Which particular algorithms are used for regularization?

A regression model which uses **L1 Regularization** technique is called **LASSO(Least Absolute Shrinkage and Selection Operator)** regression.

A regression model that uses **L2 regularization** technique is called **Ridge regression**.
Lasso Regression adds *"absolute value of magnitude"* of coefficient as penalty term to the loss function(L).

Ridge regression adds "*squared magnitude*" of coefficient as penalty term to the loss function(L).

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

The standard error of the regression (S), also known as the standard error of the estimate, represents the average distance that the observed values fall from the regression line. Conveniently, it tells you how wrong the regression model is on average using the units of the response variable. Smaller values are better because it indicates that the observations are closer to the fitted line.