MACHINE LEARNING

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?

    a. Regression.

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

    C. Low bias and high variance

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

    B. Predictive modal

7. Lasso and Ridge regression techniques belong to _____?

    D. Regularization

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

    A. Cross-validation

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.

    A. True

11. Pick the feature extraction from below:

    B. Apply PCA to project high-dimensional data

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

13. Explain the term regularization.

- Regularization is one of the most important concepts of machine learning. It is a technique to prevent the model from overfitting by adding extra information to it.
- Sometimes the machine learning model performs well with the training data but does not perform well with the test data. It means the model is not able to predict the output when dealing with unseen data by introducing noise in the output, and hence the model is called overfitting. This problem can be deal with the help of a regularization technique.
- There are two types of regularization as follows:
  - L1 Regularization or Lasso Regularization.
  - L2 Regularization or ridge Regularization.

14. Which particular algorithms are used for regularization?

- There are Two main regularization techniques: -
  - Ridge Regression (L2 Norm)
    - It Modifies the over-fitted or under-fitted models by adding the penalty equivalent to the sum of the squared of the magnitude of coefficients.
    - This means that the mathematical function representing our machine learning model is minimized and coefficients are calculated.
    - The magnitude of coefficients is squared and added. Ridge Regression performs regularization by shrinking the coefficients present.

  - Lasso (L1 Norm)
    - It modifies the over-fitted or under-fitted models by adding the penalty equivalent to the sum of the absolute values of coefficients.
    - Lasso regression also performs coefficient minimization, but instead of squaring the magnitudes of the coefficients, it takes the true values of the coefficients.
    - This means that the coefficient sum can also be 0, because of the presence of negative coefficients.

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

- Linear regression most often uses mean-square error (MSE) to calculate the error of the model MSE is calculated by measuring the distance of the observed y-values from the predicted y-values at each value of x; squaring each of these distances; calculating the mean of the each of the squared distances.