MACHINE LEARNING

(Suvarna Kumar)

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
   A) Descriptive 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?
   C) Sensitivity and Specificity

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 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.

13. Explain the term regularization?
    It means making something regular or uniform or normal. Meaning in machine learning if something is over then one must make it normal is known as Regularization. So, it is a technique to prevent the model from overfitting with the help of adding extra data to it.
    Examples of regularization, K-means, Random Forest.

14. Which algorithms are used for regularization?
    There are basically two algorithms used for regularization.
    a. Ridge Regression
    b. LASSO (Least absolute Shrinkage and Selection Operator) Regression

15. Explain the term error present in linear regression equation?
    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 results. Linear regression most often uses mean-square error (MSE) to calculate the error of the model. It 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 each of the squared distances.