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?
   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:
   B) Predictive 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?
   A) TPR and FPR

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

11. Pick the feature extraction from below:
    A)
    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.

MACHINE LEARNING

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

13. Explain the term regularization?

Ans) Regularization are those techniques in machine Learning to avoid Over fitting and Under fitting by balancing Bias and Variance and maximize the accuracy. When the model has too much noise then it becomes more flexible to the risk of overfitting and underfitting. There are 2 types of regularization- Lasso and Ridge Regression which controls the coefficients. Also ElasticNet.

14. Which particular algorithms are used for regularization?

Ans) There are 3 types of Regularization Techniques

1. Lasso Regression(L1)- The Lasso Regression will internally control the coefficients and omits those variable or columns on its own if it consider that column has no effect on the output or target. It will make it completely zero
2. Ridge Regularization(L2)- It reduces the large gap between the coefficient values. If there is a high negative value and high positive value then Ridge regression try to reduce those values and make the gap lesser.
3. ElasticNet- It uses both Lasso and ridge together.

   All three techniques use Alpha parameter which will tune the variable to adjust accordingly. It should be a very fine value.

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

Ans) The distance of the actual point to the best fit line is called the error. To find out error, we take the square of difference between the actual value and the predicted value and then taking the aquare root.

$$e = \sqrt{(y - \hat{y})^2}$$