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)   B)
   C)   D) Both A and B

2. Which of the following statement is true about outliers in linear regression?
   A) Linear regression is sensitive to outliers   B)
   C)   D)

3. A line falls from left to right if a slope is _____?
   A)   B) Negative
   C)   D)

4. Which of the following will have symmetric relation between dependent variable and independent variable?
   A)   B) Correlation
   C   D)

5. Which of the following is the reason for over fitting condition?
   A)   B)
   C) Low bias and high variance   D)

6. If output involves label then that model is called as:
   A)   B) Predictive modal
   C)   D)

7. Lasso and Ridge regression techniques belong to _____?
   A)   B)
   C)   D) Regularization

8. To overcome with imbalance dataset which technique can be used?
   A)   B)
   C)   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   B)
   C)   D)

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

11. Pick the feature extraction from below:
    A) -
    B) Apply PCA to project high dimensional data
    C) -
    D) -

MACHINE LEARNING

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.
    C) -
    D) -

-

MACHINE LEARNING

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

13. Explain the term regularization?

When we train a model in machine learning, due to its complexity it can be overfitted or underfitted. We use the regularization to properly fit; this means it reduces the overfitting and give us an optimal model.

Low bias low variance can capture data patterns and handle variations in training data, where it can generalize to unseen data and give consistency and accurate predictions.

Regularization techniques, such as Lasso(L1) , Ridge(L2) and Elastic Net =(L1+L2)regularization.

Low bias low variance gives good Balance in the data. If not it balances the bias and variance to get a good fit model.

14. Which particular algorithms are used for regularization?

The most common algorithms that use regularization are:

Ridge Regression (L2): Adds a penalty equal to the sum of the squared values of the coefficients.

Lasso Regression (L1): Adds a penalty equal to the sum of the absolute values of the coefficients.

Elastic Net: Combines both L1 and L2 regularization penalties.

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

The difference in predicted and actual values is the Error. The least error is the best-fit line.

Error effects the accuracy and effectiveness of the model.

$Y=mx+c+e$

X=independent

Y=dependent variable

M=slope

C=intercept

E=error