Machin 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:- C) Both of them
5. Which of the following is the reason for over fitting condition?
   Ans:- C)) Low bias and high variance
6. If output involves label then that model is called as
   Ans:- B)) Predictive modal
7. Lasso and Ridge regression techniques belong to_____?
   Ans:- D) Regularization
8. To overcome with imbalance dataset which technique can be used?
   Ans:- D) 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:- A) true
11. Pick the feature extraction from below:
    Ans:- 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?
    Ans:- A) B) C) all are correct.
13. Explain the term regularization?
    Ans:- 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 : L1 regularization. L2 regularization
14. Which algorithms are used for regularization?
15. Ans:-**1 :Lasso(L1 regularization).**
    2:Ridge (L2 regularization)

    L1 regularization adds a penalty that is equal to the absolute value of the magnitude of the coefficient. This regularization type can result in sparse models with few coefficients. Some coefficients might become zero and get eliminated from the model. Larger penalties result in coefficient values that are closer to zero (ideal for producing simpler models). On the other hand, L2 regularization does not result in any elimination of sparse models or coefficients. Thus, Lasso Regression is easier to interpret as compared to the Ridge

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

Ans:- 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 observed results. The regression line is used as a point of analysis when attempting to determine the correlation between one independent variable and one dependent variable.

Formula:- $Y=\alpha X+\beta\rho+\epsilon$

where:

α,β=Constant parameters

X,ρ=Independent variables

ϵ=Error term