MACHINE LEARNING

1)Which of the following methods do we use to find the best fit line for data in Linear Regression?

Ans: Both A and B

2) Which of the following statement is true about outliers in linear regression

Ans: Linear regression is sensitive to outliers

3) A line falls from left to right if a slope is _____?

Ans: Negative

4) Which of the following will have symmetric relation between dependent variable and independent variable?

Ans: Correlation

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

Ans: Low bias and high variance

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

Ans: redictive modal

7) Lasso and Ridge regression techniques belong to _____?

Ans: Regularization

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

Ans: SMOTE

9) The AUC Receiver Operator Characteristic (AUCROC) curve is an evaluation metric for binary classification problems. It uses _____ to make graph?

Ans: TPR and FPR

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

Ans: False

11) Pick the feature extraction from below:

Ans: 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) It does not make use of dependent variable.

13) Explain the term regularization?

Ans) Sometimes while training a model, overfitting and underfitting of data occur. This leads to the inaccuracy of the trained model.Regularization is the technique used to reduce the error by fitting a function appropriately on the given training set and avoid overfitting andunderfitting.Regularization techniques are used to calibrate the linear regression models in order to minimize the adjusted loss function and prevent overfitting or underfitting. There are types of regularization methods: **Ridge Regression**: It modifies the overfitted and underfitted models by adding the penalty equivalent to the sum of the squares of the magnitude of coefficients.
Lasso Regression: It modifies the overfitted and underfitted models by adding the penalty equivalent to the sum of the absolute values of coefficients.

14) Which particular algorithms are used for regularization?

Ans: algorithms are used for regularization

i.Decision Trees with Pruning: Decision trees can be regularized through pruning. Pruning removes unnecessary branches and nodes, simplifying the tree and preventing overfitting.

ii. Lasso Regression: Similar to Ridge Regression, Lasso adds a penalty, but it has an additional benefit of encouraging some coefficients to become exactly zero. This leads to feature selection, as less important features are excluded from the model.

iii. Ridge Regression: It adds a penalty to the model to discourage large values for the coefficients. This helps in controlling the complexity of the model and prevents overfitting.

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

Ans: The error term of a regression equation represents all of the variation in the dependent variable *not* explained by the weighted independent variables.A regression equation is the formula for a straight line — in this case, the best-fit line through a scatterplot of data. If there were no error, all the data points would be located *on* the regression line; to the extent they are not represents error