MACHINE LEARNING

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?

A) True

11.Pick the feature extraction from below:

A) Construction bag of words from an email

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) We need to iterate.

13. Explain the term regularization?

It is a technique to prevent the model from overfitting by adding extra information to it.

Sometimes the machine learning model performs well with the training data but does not perform well with the test data.

It means the model is not able to predict the output when deals with unseen data by introducing noise in the output, and hence the model is called overfitted.

This problem can be deal with the help of a regularization technique. This technique can be used in such a way that it will allow to maintain all variables or features in the model by reducing the magnitude of the variables.

Hence, it maintains accuracy as well as a generalization of the model. It mainly regularizes or reduces the coefficient of features toward zero.

In simple words, "In regularization technique, we reduce the magnitude of the features by keeping the same number of features."

14.Which particular algorithms are used for regularization?

Lasso regression

Lasso regression imposes a penalty on the coefficients to shrink them to exactly zero. if model think that input variable not giving good output it will remove the variable.

Ridge regression

Ridge regression imposes a penalty on the coefficients to shrink them towards zero, try to reduce the gap between different coefficients but it doesn’t set any coefficients to zero.

Elastic Net regression

Elastic Net regression is a hybrid approach that blends both penalizations of the L2 and L1 regularization of lasso and ridge methods.

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

The error term of a regression equation represents all 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.