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?

A) Regression

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 modal

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:

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?

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?

Regularization is one of the most important concepts of machine learning. 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 test data. It means the model is not able to predict the output data when deals with unseen data by introducing noise in the output, and hence the model is called overfitted and to deal with this overfitting problem we use a technique called REGULARIZATION.

In simple words, this technique converts a complex model into a simpler one, so as to avoid the risk of overfitting and shrinks the coefficients, for lesser computational cost.

14. Which particular algorithms are used for regularization?

1. RIDGE
2. LASSO
3. ELASTIC-NET-REGRESSION

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

An error term is a residual variable produced by a statistical or mathematical model, which is created when the model does not fully represent the actual relationship between the independent variables and the dependent variables.

The error term is the difference between what the model is predicting and the actual value. Basically, error term indicates the uncertainty in the model.