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:

A) Descriptive model

7. Lasso and Ridge regression techniques belong to _____?

D) Regularization

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

B) Regularization

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

C) Sensitivity and Specificity

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:

B) Apply PCA to project high dimensional data

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) We need to iterate

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

13. Explain the term regularization?

Regularization refers to techniques that are used to calibrate machine learning models in order to minimize the adjusted loss function and prevent overfitting or underfitting.

This is a form of regression, that constrains/ regularizes or shrinks the coefficient estimates towards zero

Using Regularization, we can fit our machine learning model appropriately on a given test set and hence reduce the errors in it.

Regularization Techniques

There are two main types of regularization techniques:

Ridge Regularization

and Lasso Regularization.

14.-Which particular algorithms are used for regularization?

- Ridge Regression.
- LASSO (Least Absolute Shrinkage and Selection Operator) Regression

Ridge Regression

Ridge regression is one of the types of linear regression in which we introduce a small amount of bias, known as **Ridge regression penalty** so that we can get better long-term predictions

In Statistics, it is known as the **L-2 norm**.

Lasso Regression

Lasso regression is another variant of the regularization technique used to reduce the complexity of the model. It stands for Least Absolute and Selection Operator

In statistics, it is known as the **L-1 norm**.

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. As a result of this incomplete relationship, the error term is the amount at which the equation may differ during empirical analysis.

The error term is also known as the **residual**, **disturbance**, or **remainder term,** and is variously represented in models by the letters *e,* ε, or u.

An error term essentially means that the model is not completely accurate and results in differing results during real-world applications.