**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:

D) All of the above

7. Lasso and Ridge regression techniques belong to \_\_\_\_\_\_\_\_\_?

D) Regularization

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

A) Cross validation

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:

A) Construction bag of words from a email

B) Apply PCA to project high dimensional data

C) Removing stop words

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.

**MACHINE LEARNING**

**13. Explain the term regularization?**

The word regularize means to make things regular or acceptable. This is exactly why we use it for. Regularizations are techniques used to reduce the error by fitting a function appropriately on the given training set and avoid overfitting. this technique discourages learning a more complex or flexible model, so as to avoid the risk of overfitting.

**14. Which particular algorithms are used for regularization?**

The commonly used regularization techniques are:

1. L1 regularization
2. L2 regularization
3. Dropout regularization

A regression model which uses **L1 Regularization**technique is called **LASSO(Least Absolute Shrinkage and Selection Operator)** regression.

**L1 regularization :**  
A regression model that uses **L2 regularization** technique is called **Ridge regression**.

So ridge regression puts constraint on the coefficients (w). The penalty term (lambda) regularizes the coefficients such that if the coefficients take large values the optimization function is penalized. So**, ridge regression shrinks the coefficients and it helps to reduce the model complexity and multi-collinearity.**

**L2 regularization :**

Just like Ridge regression cost function, for lambda =0, the equation above reduces to equation 1.2. The only difference is instead of taking the square of the coefficients, magnitudes are taken into account. This type of regularization (L1) can lead to zero coefficients i.e., some of the features are completely neglected for the evaluation of output. **So, Lasso regression not only helps in reducing over-fitting but it can help us in feature selection.**

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

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.