

# MACHINE LEARNING ASSIGNMENT

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

Ans- Least Square error

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 \_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 & High variance

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

Ans-Predictive model

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

11. Pick the feature extraction from below: Apply PCA to project high dimensional data

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

12. Which of the following is true about Normal Equation used to compute the coefficient of the Linear Regression?

Ans-A) We don't have to choose the learning rate.

B) It becomes slow when number of features is very large.

13. Regularization-The word regularize means to make things regular or acceptable. Regularization is one of the basic and most important concept in the world of machine learning. This technique is used to reduce the error by fitting a function appropriately on the given training set and avoid overfitting.

14. Which particular algorithms are used for regularization?

Ans-Lasso regression

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

Ans-An error term in Linear Regression 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. And it means that the model is not completely accurate and results in differing during realworld applications.