MACHINE LEARNING

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

   Answer: **Least Square Error**

2. Which of the following statement is true about the outliers in linear regression?

   Answer: **Linear regression is sensitive to outliers**

3. A line falls from left to right if a slope is _____?

   Answer: **Negative**

4. Which of the following will have symmetric relation between dependent variable and independent variable?

   Answer: **Corelation**

5. Which of the following is the reason for over fitting condition?

   Answer: **Low bias high variance**

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

   Answer: **Predictive modal**

7. Lasso and Ridge regression techniques belong to _____?

   Answer: **Regularization**

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

   Answer: **SMOTE**

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

   Answer: **TPR and FPR**

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

    Answer: **False**

11. Pick the feature extraction from below:

    Answer: **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?

    Answer: **A) We don't have to chose learning rate**

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

    C) We need to iterate

13. Explain the term regularization?

Answer: Regularization is technique used to minimise error by fitting the function properly by using training data set and avoid overfitting. Commonly used regularization techniques are L1: Regularization and L2: Regularization.

14. Which particular algorithms are used for regularization?

    Answer: **There are three main Regularization techniques are used, namely,**

    1. **Ridge Regression (L2 Norm)**
    2. **Lasso (L1 Norm)**
    3. **Dropout**

    Ridge and Lasso can be used for any algorithms involving weight parameters, including neural test. Dropout is primarily used in any kind of neural networks e.g. ANN, DNN, CNN, or RNN to moderate the learning.

15. Explain the term error present in linear regression equation?
    Answer**: Error is difference between expected value from a linear regression equation and actual value. The error terms stand for any influence impacted the variable.**

Name: Humayun Kabir
Batch No.: Internship-25.
Project: Machine Learning Worksheet – 1.