Rumnaz Binte Islam Sarah

Machine Learning Worksheet 1

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

Ans. D (Both A and B)

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

Ans. A (Linear regression is sensitive to outliers)

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

Ans. A (Positive)

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

Ans. B (Correlation)

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

Ans. C (Low bias and high variance)

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

Ans. B (Predictive modal)

7. Lasso and Ridge regression techniques belong to _____?

Ans. D (Regularization)

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

Ans. B (Regularization)

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

Ans. A (TPR and FPR)

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

Ans. B (False)

11. Pick the feature extraction from below:

Ans. 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?

Ans. A & B (A. We don’t have to choose the learning rate & B. It becomes slow when number of features is very large)

13. Explain the term regularization?

Ans. While training a machine learning model, regularization is used to create a balance in the way in which data is being interpreted. It works to reduce or avoid overfitting and underfitting to provide accurate data. For example, if a feature does not add significant weight to the data, then regularization algorithms will give less or zero importance to it.

14. Which particular algorithms are used for regularization?

Ans. The three main algorithms of regularization are:

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

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

Ans. Error term stands for the difference between observed values and predicted values of data that is being analyzed using linear regression