

MACHINE LEARNING

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

Ans: A) Least Square Error

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: B) Negative

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: D) SMOTE

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

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

ASSIGNMENT – 39

MACHINE LEARNING

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

13. Explain the term regularization?

Ans: Regularizations are techniques 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: Ridge Regression (L2 Norm), Lasso (L1 Norm) , Dropout

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

Ans: 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.