

Machine Learning Assignment

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

Answer: A) Least Square Error

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

Answer: A) Linear regression is sensitive to outliers

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

Answer: B) Negative

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

A) Regression

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

Answer: C) Low bias and high variance

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

Answer: B) Predictive model

7. Lasso and Ridge regression techniques belong to _____?

Answer: D) Regularization

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

Answer: D) SMOTE

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

Answer: A) TPR and FPR

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

Answer: B) False

11. Pick the feature extraction from below:

Answer: 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?

Answers:

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?

Answer: In machine learning, overfitting is a frequently occurring problem. Regularization is one of the ways to overcome this problem. It reduces the model variance by fitting the function appropriately with the training dataset and avoid overfitting. What it basically does is that it reduces the degree of freedom of the function by adding a shrinkage penalty.

14. Which particular algorithms are used for regularization?

Answer: Mainly two algorithms are used for regularization:

1. **L1 Regularization** also called LASSO(Least Absolute Shrinkage and Selection Operator) regression.
2. **L2 regularization** also called Ridge regression.

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

Answer: Error in a linear regression equation means the average distance of actual values from the predicted values of the equation. It gives us an overall idea of how wrong the linear regression model is.