

Machine Learning assignment

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

A) Least Square Error

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

A) Linear regression is sensitive to outliers

3 - A line falls from left to right if a slope is ____?

B) Negative

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

B) Correlation

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

C) Low bias and high variance

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

B) Predictive modal

7 - Lasso and Ridge regression techniques belong to _____?

D) Regularization

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

D) SMOTE

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

A) TPR and FPR

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

B) False

11. Pick the feature extraction from below:

B) Apply PCA to project high dimensional data

In Q12, more than one options are correct, choose all the correct options:

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

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

C) We need to iterate.

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

13. Explain the term regularization?

Regularization help to reduce the variance without losing any important properties in the data (without substantially increasing the bias)

14. Which particular algorithms are used for regularization?

The particular algorithms used for regularization are (1) Lasso (also called L1) and Ridge Regression (also called L2).

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

The error term in the linear regression is the difference between the expected and observed (predicted) values.