

MACHINE LEARNING

n Q1 to Q11, only one option is correct, choose the correct option:

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?

A) Regression

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

A) High bias and high variance

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

D) All of the above

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 B) Sensitivity and precision

C) Sensitivity and Specificity D) Recall and precision

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

A) True B) False

11. Pick the feature extraction from below:

A) Construction bag of words from a email

B) Apply PCA to project high dimensional data

C) Removing stop words

D) Forward selection

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?

Ans: Regularization is used to prevent data over fitting.

14. Which particular algorithms are used for regularization?

Ans: LASSO and RIDGE regression algorithm are used in regularization.

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

Ans: When the model is not sure about the relation between the input and the output variables, the error is created.