

In 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
- B) Maximum Likelihood
- C) Logarithmic Loss
- D) Both A and B

ANS D) Both A and B

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

- A) Linear regression is sensitive to outliers
- B) linear regression is not sensitive to outliers
- C) Can't say
- D) none of these

Ans A) Linear regression is sensitive to outliers

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

- A) Positive
- B) Negative
- C) Zero
- D) Undefined

Ans B) Negative

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

- A) Regression
- B) Correlation
- C) Both of them
- D) None of these

Ans B) Correlation

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

- A) High bias and high variance
- B) Low bias and low variance
- C) Low bias and high variance
- D) none of these

Ans C) Low bias and high variance

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

- A) Descriptive model
- B) Predictive modal
- C) Reinforcement learning
- D) All of the above

Ans B) Predictive modal

7. Lasso and Ridge regression techniques belong to_____?

- A) Cross validation
- B) Removing outliers
- C) SMOTE
- D) Regularization

Ans D) Regularization

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

A) Cross validation

B) Regularization

C) Kernel

D) SMOTE

Ans 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

Ans A) TPR and FPR

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

A) True

B) False

Ans 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

Ans 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?

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

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

C) We need to iterate.

D) It does not make use of dependent variable.

Ans A & B

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

13. Explain the term regularization?

Regularization is a way to stop a machine learning model from becoming too complicated. When a model becomes too good at learning from the training data, it might also learn unnecessary details that don't help with new data. Regularization adds a small penalty to the model, so it stays simple and doesn't focus too much on any one part of the data. This helps the model work better on new data and prevents mistakes

14. Which particular algorithms are used for regularization?

1 Generalized Linear Models

2 Decision Trees and Ensemble Methods

3 Neural Networks

4 SVMs

5 linear and Logistic Regression

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

Ans In a linear regression equation, "error" is the difference between the actual values and the predicted values. It shows how well the model fits the data. When we draw a line to represent the data, some points will be above or below the line, and those differences are called errors or residuals. These errors can happen for many reasons, like random noise, missing important information, or mistakes in measuring. By looking at these errors, we can improve the model and make better predictions