

MACHINE LEARNING

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?

Answer : Least Square Error

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

Answer: Linear regression is sensitive to outliers

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

Answer: Negative

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

Answer: B) Correlation

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 modal

7. Lasso and Ridge regression techniques belong to Regularization?

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.

B) False

11. Pick the feature extraction from below:

Answer: 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?

Answers:

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

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

MACHINE LEARNING

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

13. Explain the term regularization?

Answer 13:

- Regularization is a way to prevent overfitting in machine learning models by adding a penalty for complexity. This helps the model to generalize better to new, unseen data, avoiding a situation where it becomes too closely fitted to the training set.
- Two common regularization methods are
 - Lasso, which uses L1 regularization, and
 - Ridge, which uses L2 regularization.
- Each adds a different type of penalty to the model's loss function, helping to keep the model simpler and more robust.

14. Which particular algorithms are used for regularization?

Answer 14:

- Common algorithms used for regularization:
 - Lasso Regression (L1 Regularization): Adds a penalty equal to the absolute value of the magnitude of coefficients. It can shrink some coefficients to zero, effectively performing feature selection.
 - Ridge Regression (L2 Regularization): Adds a penalty equal to the square of the magnitude of coefficients. It helps in reducing the model complexity but doesn't shrink coefficients to zero.
 - Elastic Net: Combines both L1 and L2 regularization. It balances the benefits of both Lasso and Ridge regression.

These techniques help in creating simpler models that generalize better to new data.

MACHINE LEARNING

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

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

Answer 15:

- Error in Linear Regression Equation:
 - Residuals: The difference between the observed values and the predicted values by the model.
 - Mean Squared Error (MSE): The average of the squared differences between the observed and predicted values.
 - Root Mean Squared Error (RMSE): The square root of the MSE, providing a measure of the average magnitude of the errors.
 - Mean Absolute Error (MAE): The average of the absolute differences between the observed and predicted values.

These errors help in evaluating the performance of the linear regression model
