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?

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?

D) none of these

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

B) Predictive model

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.

A) True

11. Pick the feature extraction from below:

A) Construction bag of words from an 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?

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

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

MACHINE LEARNING

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

13. Explain the term regularization?

A. Regularization is a strategy for reducing mistakes and avoiding overfitting by fitting the function suitably on the supplied training set.

The following are some of the most often used regularisation techniques:

- L1 regularization
- L2 regularization
- Dropout regularization

14. Which particular algorithms are used for regularization?

- Ridge Regression (L2): Ridge regression is a model tuning technique that may be used to analyze data with multicollinearity. L2 regularization is achieved using this approach. When there is a problem with multicollinearity, least-squares is unbiased, and variances are significant, resulting in projected values that are distant from the actual values.
- Lasso (L1): Lasso regression is a type of regularization. For a more accurate forecast, it is preferred over regression approaches. Shrinkage is used in this model. Data values are shrunk towards a central point known as the mean in shrinkage. Simple, sparse models are encouraged by the lasso approach.

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

A. The total of the deviations inside the regression line, which gives an explanation for the discrepancy between the theoretical value of the model and the actual observed results, is referred to as the error term in a statistical model. When attempting to find the correlation between one independent variable and one dependent variable, the regression line is utilized as a point of analysis.

The error term is the difference between the projected price at a given moment and the price that was actually observed in a linear regression model tracking a stock's price over time. The price will fall on the trend line and the error term will be 0 in cases where the price is precisely what was expected at a given period.