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?

(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?

(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.

MACHINE LEARNING

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

13. Explain the term regularization? Regularization is a form of regression, that constrains/ regularizes or shrinks the coefficient estimates towards zero. In other words, this technique discourages learning a more complex or flexible model, so as to avoid the risk of overfitting. In simple terms, Regularization is a technique used for tuning the function by adding an additional penalty term in the error function appropriately on the given training set and avoid overfitting.

14. Which particular algorithms are used for regularization? There are three main regularization techniques, namely: 1. Ridge Regression (L2 Norm) 2. Lasso (L1 Norm) 3. Dropout Ridge and Lasso can be used for any algorithms involving weight parameters, including neural nets. Dropout is primarily used in any kind of neural networks e.g. ANN, DNN, CNN or RNN to moderate the learning.

15. Explain the term error present in linear regression equation? An error term represents the margin of error within a statistical model; it refers to the sum of the deviations within the regression line, which provides an explanation for the difference between the theoretical value of the model and the actual observed results. The regression line is used as a point of analysis when attempting to determine the correlation between one independent variable and one dependent variable. It is often said that the error term in a regression equation represents the effect of the variables that were omitted from the equation