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

Ans : A) Least Square Method

2. 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 : A) high 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 model

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

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 : both B) and C)

MACHINE LEARNING

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

13. Explain the term regularization?

Ans : Regularisation techniques are used to solve the problem of overfitting or underfitting of models by penalising the features causing the overfitting or underfitting. This is done by Lasso regularisation and Ridge regularisation

14. Which particular algorithms are used for regularization?

Ans : lasso regularization and Ridge regularization are regularisation techniques. In lasso regularization the value of the differences is considered whereas in ridge regularization the sum of the squares of the differences are considered.

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

Ans : The model predicting the output might be predicting close to the actual values. The difference between the predicted value and the actual value is called the error in linear regression equation. The average of the sum of squares of the differences gives the average error.