Assignment 3 for DS2311

MACHINE LEARNING

Question 1. Which of the following methods do we use to find the best fit line for data in Linear Regression?

Answer : D

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

Answer: A

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

Answer : B

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

Answer : B

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

Answer: D

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

Answer: B

7. Lasso and Ridge regression techniques belong to _____?

Answer: D

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

Answer: D

9. The AUC Receiver Operator Characteristic (AUCROC) curve is an evaluation metric for binary
classification problems. It uses _____ to make graph?

Answer : A

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

Answer : B

11. Pick the feature extraction from below:

Answer : D

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?

Answer : A & B

13. Explain the term regularization?

Answer : Regularization is one of the most important concept of machine learning. It is a technique to prevent the model from over fitting by adding extra information to it. Sometimes the machine learning model performs well with the training data but does not perform well with the test data. It means the model is not able to predict the output when deals with unseen data by introducing noise in the output, and hence the model is called over fitted. This problem can be deal with the help of a regularization technique.

14. Which particular algorithms are used for regularization?

Answer: Ridge Regularization & Lasso Regularization

Ridge Regularization is a technique in which a small amount of bias is introduced so that we can get better long-term predictions. It is used to reduce the complexity of the model. It is also called as **L2 regularization**.

Lasso Regularization is a technique that reduces the complexity of the model. It stands for Least Absolute and Selection Operator. It helps in reducing the over fitting of a model as well as feature selection. It is called **L1 Regularization.**

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

Answer: In the context of linear regression, an error refers to the difference between the observed dependent variable value (Y) and the predicted dependent variable value (Ŷ) calculated by the regression model.It does not represent any systematic bias in the regression model.

The term error can also be used to describe the differences between the predicted values (Ŷ) and the actual values of Y in the data. This can help to evaluate the accuracy of the regression model.