MACHINE LEARNING

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: D) Both A and B

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: C) Low bias and high variance

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

A) Descriptive model   B) Predictive model

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

Ans: A) Construction bag of words from a email

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 A & B

13. Explain the term regularization?

Ans: Regularization is a technique used in machine learning models to prevent overfitting, which is a scenario where the model performs well on training data but not on unseen data.

14. Which particular algorithms are used for regularization?

Ans: Lasso Regression and Ridge Regression.

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

Ans: the term "error" refers to the difference between the observed value and the predicted value for a given observation.