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

Answer: A) Least Square Error

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

Answer: A) Linear regression is sensitive to outliners

3. A line falls from left to right if a slope is _____?
   A) Positive B) Negative C) Zero D) Undefined

Answer: 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

Answer: 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

Answer: C) Low 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

Answer: B) Predictive model

7. Lasso and Ridge regression techniques belong to _____?

A) Cross validation B) Removing outliers

C) SMOTE D) Regularization

Answer: A) Cross validation

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

A) Cross validation B) Regularization

C) Kernel D) SMOTE

Answer: 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

Answer: 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

Answer: B) False

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

Answer: B) Apply PCA to project high dimensional data

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.

Answer: B) It becomes slow when number of features is very large & C) We need to iterate

MACHINE LEARNING

13. Explain the term regularization?

Answer: Regularization helps in overcoming the problem of overfitting and also increase the model interpretability. It is one of the most important concepts of ML, as it shrinks the coefficient to zero.

14. Which particular algorithms are used for regularization?

Answer: There are below listed main regularization techniques:

1. Ridge Regression.
2. LASSO (Least Absolute Shrinkage and Selection Operator) Regression.

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

Answer: It defines the difference between the predictive value according to trained data and the actual value which we received.