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: A)

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)

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

A) Positive

B) Negative

C) Zero

D) Undefined

ANS: B)

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)

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)

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)

7. Lasso and Ridge regression techniques belong to _____?

A) Cross validation

B) Removing outliers

C) SMOTE

D) Regularization

ANS: D)

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

A) Cross validation

B) Regularization

C) Kernel

D) SMOTE

ANS: D)

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)

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)

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: B)

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: A) and B)

13. Explain the term regularization?

ANS: Regularization is a method of machine learning modeling to prevent the overfitting of model.Overfitting is when a model works well on a training dataset but it gives poor result on a new or testing dataset.Regularization can reduces the complexity of a model.In machine learning there are two types of regularization:

1)Lasso regression.

2)Ridge regression.

14. Which particular algorithms are used for regularization?

ANS: There are two alogorithm is used for regularization are Ridge regression and Lasso regression.The Ridge regression applies a penalty on the sum of squares of the weights in the model and Lasso regression applies a penalty on sum of weights.These methods prevent overcomplication and overfitting.

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

ANS: In linear regression, the term "error" represents the imbalancing between the observed value and the predicted value by the linear model. Specifically, it refers to the difference between the actual outcome and the predicted outcome for each data point.It is important to refine the regression model using the least square method to minimize the errors.