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 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

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 modal

C) Reinforcement learning D) All of the above

Ans:- B) Predictive modal

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:- A) We don't have to choose the learning rate.

B) It becomes slow when number of features is very large.

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

13. Explain the term regularization?

Ans:- regularization resolve the over fitting problem by reducing the coefficients magnitude basically regularization technique is used for to decrese the complexity of model which is come after the training phase or the output of the training phase and to shrink the magnitude of of regression coefficient is called regularization.

14. Which particular algorithms are used for regularization?

Ans:- there are few algorithms are used for regularization or to reduce the complexity of ml model and to resolve overfitting the problem just by reducing the magnitude of coefficient . the common algorithms used for regularization in machine learning include Lasso (L1 regularization), Ridge (L2 regularization), and Elastic Net (combination of L1 and L2 regularization). These techniques help prevent overfitting by adding a penalty term to the loss function, encouraging simpler models. Regularization is especially useful when dealing with high-dimensional datasets to improve model generalization and reduce model complexity.

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

Ans:- the term error in linear regression equation is noyhing but the difference between the actual value and the predicted value its done by the line of linear regression in brief linear regression, the term "error" refers to the difference between the predicted values and the actual values of the dependent variable. These errors represent the extent to which the model fails to capture the true relationship between the independent and dependent variables. Minimizing these errors is the key objective in linear regression to obtain an accurate predictive model.