MACHINE LEARNING

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

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

A) Linear regression is sensitive to outliers

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

B) Negative

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

B) Correlation

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

C) Low bias and high variance

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

D) All of the above

7. Lasso and Ridge regression techniques belong to _____?

D) Regularization

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

A) Cross validation

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

C) Sensitivity and Specificity

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

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 stops words

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?

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

C) We need to iterate.

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

13. Explain the term regularization?

Ans: Regularization is the concept of machine learning. It is a technique to prevent the model from over fitting. One reduces the magnitude of the features by keeping the same number of features.
The technique used can be Ridge and Lasso technique.

14. Which particular algorithms are used for regularization?

Ans: Following algorithms can be used:

- Ridge Regression: scale down the coefficient by adding penalty to the loss
- LASSO (Least Absolute Shrinkage and Selection Operator) Regression: scale down the coefficient and it can reduce to zero, it means the label can be removed if contribution is zero.
- Elastic-Net Regression: It is the technique which uses both Ridge or lasso technique, depending on the data set.

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

Ans: Error Term: It is the difference between the actual value and the predicted value.
Linear regression often uses mean-square error (MSE) to calculate the error of the model.
Linear regression fits a line to the data by finding the regression coefficient that result in the smallest MSE.