MACHINE LEARNING-Answers

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?

Ans (A) Least Square Error

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

Ans (A) Linear regression is sensitive to outliers

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

Ans (B) Negative

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

Ans (B) Correlation

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

Ans (C) Low bias and high variance

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

Ans (B) Predictive modal

7. Lasso and Ridge regression techniques belong to _____?

Ans (D) Regularization

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

Ans (D) SMOTE

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

Ans (A) TPR and FPR

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

Ans (B) False

11. Pick the feature extraction from below:

Ans (B) Apply PCA to project high dimensional data

In Q12, more than one options are correct, choose all the correct options:

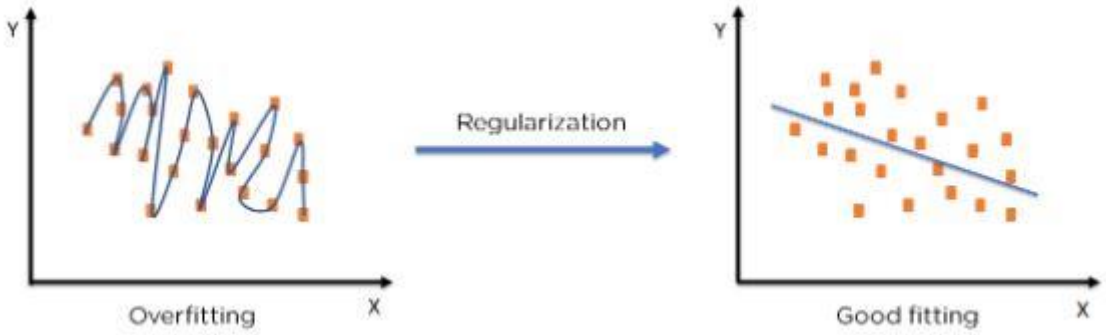
12. Which of the following is true about Normal Equation used to compute the coefficient of the Linear Regression?
A) We don't have to choose the learning rate.
B) It becomes slow when number of features is very large.
C) We need to iterate.
D) It does not make use of dependent variable.

Ans (a) and (b)

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

13. Explain the term regularization?

Ans- Regularization refers to techniques that are used to calibrate machine learning models in order to minimize the adjusted loss function and prevent overfitting or underfitting.



Regularization on an over-fitted model

Using Regularization, we can fit our machine learning model appropriately on a given test set and hence reduce the errors in it.

The commonly used regularization techniques are :

1. L1 regularization
2. L2 regularization

A regression model which uses **L1 Regularization** technique is called **LASSO(Least Absolute Shrinkage and Selection Operator)** regression.
A regression model that uses **L2 regularization** technique is called **Ridge regression**.

14. Which particular algorithms are used for regularization?

Ans There are mainly two types of regularization techniques:-

1) Ridge Regularization
2) Lasso Regularizatiion

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

Ans An error term is a residual variable produced by a statistical or mathematical model, which is created when the model does not fully represent the actual relationship between the independent variables and the dependent variables. As a result of this incomplete relationship, the error term is the amount at which the equation may differ during empirical analysis.

The error term is also known as the residual , disturbance or remainder term and is variously represented in models by the letters e, ε, or u.