

## 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?**

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:- C) Sensitivity and Specificity

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

Ans:- A) True

**11. Pick the feature extraction from below:**

Ans:- D) Forward selection

## MACHINE LEARNING

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?**

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

D) It does not make use of dependent variable.

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

**13. Explain the term regularization?**

Ans:- While training a machine learning model, it's obvious the model can easily be overfitted or under fitted. To avoid this, we use regularization in machine learning to properly fit a model onto our test set. Regularization techniques help reduce the chance of overfitting and help us get an optimal model.

OR

“Regularization are the techniques are used to calibrate/regulate our machine learning models in order to minimize the adjusted loss function and prevent overfitting or underfitting”.

**14. Which particular algorithms are used for regularization?**

Ans:-There are three algorithms, are used for regularization :-

**1.LASSO regression (L1 FORM) :-**

Penalizes the model based on the sum of the magnitude of the coefficients. It will reduce some coefficients exactly to zero when we use a sufficiently large tuning parameter  $\lambda$ . In addition to regularizing ,lasso also performs feature selection.

**2.RIDGE regression (L2 FORM) :-**

Penalizes the model based on the sum of squares of magnitude of the coefficients. It will shrink the coefficients for those predictors which contribute very less in the model but have huge weights, very close to zero but it never makes them exactly zero. Thus, the final model will still contain all those predictors, though with less weights. It will not regularize our model well.

**3.ELASTICNET: -**Penalizes the model use combine penalties of both **LASSO** and **RIDGE** regression.

## MACHINE LEARNING

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

Ans:-An **error term** in statistics is a value which represents how observed data differs from actual data. It can also be a variable which represents how a given statistical model differs from reality. It is the distance between each point and the linear graph is our **error term** (In our graph space between red dots and blue dots).

