1. Which of the following methods do we use to find the best fit line for data in Linear Regression?

Ans = Least Square Error.

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

Ans = Linear regression is sensitive to outliers.

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

Ans = Negative.

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

Ans = Correlation.

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

Ans = Low bias and high variance.

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

Ans = Predicted Model

7. Lasso and Ridge regression techniques belong to _____?

Ans = Cross validation

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

Ans = SMOTE

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

Ans = TPR and FPR

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

Ans = False

11. Pick the feature extraction from below:

Ans = Apply PCA to project high dimensional data.

12. Which of the following is true about Normal Equation used to compute the coefficient of the Linear Regression?

Ans = 1. We don't have to choose the learning rate.

2. It becomes slow when number of features is very large.

3. We need to iterate.

13. Explain the term regularization?

Ans = While training a machine learning model, 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.

14. Which particular algorithms are used for regularization?

Ans = Lasso regression is a type of linear regression that uses shrinkage. Shrinkage is where data values are shrunk towards a central point, like the mean. The lasso procedure encourages simple, sparse models. This particular type of regression is well-suited for models showing high levels of multicollinearity or when you want to automate certain parts of model selection, like variable selection/parameter elimination.

Ridge regression is a model tuning method that is used to analyse any data that suffers from multicollinearity. This method performs L2 regularization. When the issue of multicollinearity occurs, least-squares are unbiased, and variances are large, this results in predicted values being far away from the actual values.

Elastic net linear regression uses the penalties from both the lasso and ridge techniques to regularize regression models. The technique combines both the lasso and ridge regression methods by learning from their shortcomings to improve the regularization of statistical models.

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 population data. It can also be a variable which represents how a given statistical model differs from reality. The error term is often written.

An error term represents the margin of error within a statistical model; it refers to the sum of the deviations within the regression line, which provides an explanation for the difference between the theoretical value of the model and the actual observed results. The regression line is used as a point of analysis when attempting to determine the correlation between one independent variable and one dependent variable.