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: Both of them (Regression & 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: Predictive modal

7. Lasso and Ridge regression techniques belong to _____?

Ans: Regularization

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: True

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: 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.

13. Explain the term regularization?

Ans: The word regularize means to make things regular or acceptable. It is a technique used to reduce the errors by fitting the function appropriately on the given training set and avoid overfitting.

13. Which particular algorithms are used for regularization?

Ans: **Ridge Regression:** It is a method of analyzing data that suffer from multi-collinearity.

LASSO Regression: It is a method that performs both feature selection and regularization in order to enhance the prediction accuracy of the model. It is also known as the L1 Regularization.

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

Ans: The error term of a regression equation represents all of the variation in the dependent variable not explained by the weighted independent variables.

A regression equation is the formula for a straight line — in this case, the best-fit line through a scatterplot of data. If there were no error, all the data points would be located on the regression line; to the extent they are not represents error; this is what the error term summarizes.