Q1. Which of the following methods do we use to find the best fit line for data in Linear Regression?

Ans: Least Square Error

Q2. Which of the following statement is true about outliers in linear regression? Ans: Linear regression is sensitive to outliers

Q3. A line falls from left to right if a slope is _____?

Ans :Negative

Q4. Which of the following will have symmetric relation between dependent variable and independent variable?

Ans: Both of them

Q5. 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: Sensitivity and precision

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

curve should be less.

Ans: False

Q11. 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 & 14. Regularizations are techniques used to reduce the error by fitting a function appropriately on the given training set and avoid overfitting.

The commonly used regularization techniques are :

1. L1 regularization also know as Lasso regression used to reduce the complexity of the model. It stands for Least Absolute and Selection Operator.

$$Loss = \sum_{i=1}^{n} (y_i - (w_i x_i + c))^2 + \lambda \sum_{i=1}^{n} |w_i|$$

2. L2 regularization also know as Ridge regression we introduce a small amount of bias, known as Ridge regression penalty so that we can get better long-term predictions.

$$Loss = \sum_{i=1}^{n} (y_i - (w_i x_i + c))^2 + \lambda \sum_{i=1}^{n} w_i^2$$

3. Dropout regularization: Dropout is a regularization technique for reducing overfitting in neural networks by preventing complex co-adaptations on training data. It is a very efficient way of performing model averaging with neural networks. The term "dropout" refers to dropping out units (both hidden and visible) in a neural network.

4.Elastic-Net Regression is regularized regression method that linearly combines the L1 and L2 penalties of the LASSO and Ridge methods respectively.

$$Loss = \sum_{i=0}^{n} (y_i - (w_i x_i + c))^2 + \lambda_1 \sum_{i=0}^{n} |w_i| + \lambda_2 \sum_{i=0}^{n} w_i^2$$

15. Error 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 ε.