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:- B) Removing outliers
8. To overcome with imbalance dataset which technique can be used?

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

   Ans:- D) Recall and precision
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:- A) Construction bag of words from a email
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.

13. Explain the term regularization?

Over fitting is a phenomenon that occurs when a Machine Learning model is constraint to training set and not able to perform well on unseen data. Regulariztion is a techmique used to fitting the model appropriately on given training set to avoid overfitting.

14. Which particular algorithms are used for regularization?

1.L1 Regulrization

2.L2 Regularization

3.Drop out Regularization

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

Ans:-A regression model is always has an error term because in real time dependent variables are not perfect predictors of the independent variable. The line is estimated based on available data set. Linear regression most often uses MSE(mean square error) to calculate the error in the model.