Machine Learning Assignment

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 - A) Cross validation

**9) The AUC Receiver Operator Characteristic (AUCROC) curve is an evaluation metric for
binary classification problems. It uses _____ to make graph?**
Ans - A) TPR and FPR

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

11) Pick the feature extraction from below:
Ans -

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.
C) We need to iterate.

13) Explain the term regularization?

Ans - Regularization refers to techniques that are used to calibrate machine learning models in order to minimize the adjusted loss function and prevent overfitting or underfitting. Using Regularization, we can fit our machine learning model appropriately on a given test set and hence reduce the errors in it.

14) Which particular algorithms are used for regularization?

Ans -
- Ridge Regression.
- LASSO (Least Absolute Shrinkage and Selection Operator) Regression.
- Elastic-Net Regression.

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

Ans - The standard error of the regression provides the absolute measure of the typical distance that the data points fall from the regression line.

Linear Regression most often uses mean-square error (MSE) to calculate the error of the model.

MSE is calculated by:

1. measuring the distance of the observed y-values from the predicted y-values at each value of x;

2. squaring each of these distances;

3. calculating the mean of each of the squared distances.