Machine Learning Assignment

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

2. Which of the following statement is true about outliers in linear regression?
    a) Linear regression is sensitive to outliers.

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

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

5. Which of the following is the reason for over fitting condition?
    d) Low bias high variance

6. If output involves label then that model is called as:
    a) Descriptive model.

7. Lasso and Ridge regression techniques belong to _____?
    d)Regularization

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

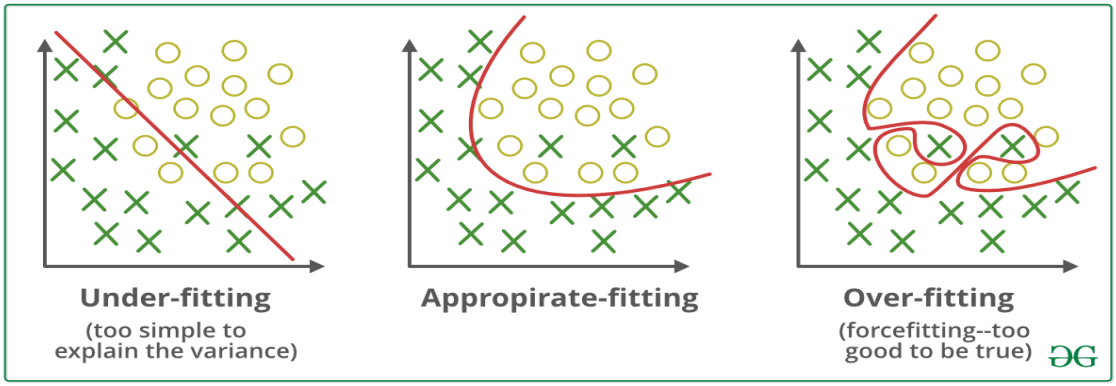
9. The AUC Receiver Operator Characteristic (AUCROC) curve is an evaluation metric for binary classification problems. It uses _____ to make graph?
    a. TPR and FPR

10. In AUC Receiver Operator Characteristic (AUCROC) curve for the better model area under the curve should be less.
    c) False
11. . Pick the feature extraction from below:
    b) 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.
    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 ?

Regularization consists of different techniques and methods used to address the issue of over-fitting by reducing the generalization error without affecting the training error much. Choosing overly complex models for the training data points can often lead to overfitting.



Regularization is a technique used to reduce the errors by fitting the function appropriately on the given training set and avoid overfitting.
The commonly used regularization techniques are:

1. L1 regularization
2. L2 regularization
3. Dropout regularization

This article focus on L1 and L2 regularization.

A regression model which uses L1 Regularization technique is called LASSO(Least Absolute Shrinkage and Selection Operator) regression.
A regression model that uses L2 regularization technique is called Ridge regression.
Lasso Regression adds *"absolute value of magnitude"* of coefficient as penalty term to the loss function(L).

13. Which particular algorithms are used for regularization?

    A regression model which uses L1 Regularization technique is called LASSO(Least Absolute Shrinkage and Selection Operator) regression.

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

    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.

- Mean Absolute Error.
- Mean Square Error.
- Mean Absolute Percentage Error.
- Mean Percentage Error.