MACHINE LEARNING ASSIGNMENT – 39

1. Which of the following methods do we use to find the best fit line for data in Linear Regression?  
   Ans: D) Both A and B
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

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

Ans: A) Regression

1. Which of the following is the reason for over fitting condition?

Ans: A) High bias and high variance

1. If output involves label then that model is called as:

Ans: B) Predictive modal

1. Lasso and Ridge regression techniques belong to \_\_\_\_\_\_\_\_\_?

Ans: D) Regularization

1. To overcome with imbalance dataset which technique can be used?

Ans: A) Cross validation

1. 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

1. In AUC Receiver Operator Characteristic (AUCROC) curve for the better model area under the curve should be less.

Ans: A) True

1. Pick the feature extraction from below:

Ans: B) Apply PCA to project high dimensional data

1. Which of the following is true about Normal Equation used to compute the coefficient of the Linear Regression?

B) It becomes slow when number of features is very large.

C) We need to iterate.

1. Explain the term regularization?

It is a technique to prevent the model from overfitting by adding extra information to it. Sometimes the [machine learning](https://www.javatpoint.com/machine-learning) model performs well with the training data but does not perform well with the test data. It means the model is not able to predict the output when deals with unseen data by introducing noise in the output, and hence the model is called overfitted. This problem can be deal with the help of a regularization technique.

This technique can be used in such a way that it will allow to maintain all variables or features in the model by reducing the magnitude of the variables. Hence, it maintains accuracy as well as a generalization of the model.

It mainly regularizes or reduces the coefficient of features toward zero. In simple words, "*In regularization technique, we reduce the magnitude of the features by keeping the same number of features."*

1. Which particular algorithms are used for regularization?

Ridge regression is a regularization technique, which is used to reduce the complexity of the model. It is also called as L2 regularization.

Lasso regression is another regularization technique to reduce the complexity of the model. It stands for Least Absolute and Selection Operator.

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

It is the difference between actual and predicted label value.