# Machine Learning - Worksheet 1

**Q1 to Q11, only one option is correct, choose the correct option:**

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

   **Ans.** Lease square method.

2. Which of the following statement is true about outliers in linear regression?

   **Ans.** Linear regression is sensitive to outliers.

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

   **Ans.** Positive.

4. Which of the following will have symmetric relation between dependent variable & independent variable?

   **Ans.** Correlation.

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

   **Ans.** None of the above**.**

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

   **Ans.** Predictive Model.

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.** TPR and FPR.

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

    **Ans.** False

11. Pick the feature extraction from below:

    **Ans.** A) Construction bag of words from email B) Apply PCA to project high dimensional data C) Removing stop words.

**In Q12, more than one options are correct, choose all the correct options:**

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

**Ans.** A and B

**Q13 and Q15 are subjective answer type questions, Answer them briefly.**

**13.** Explain the term regularization?

**Ans.** In Machine Learning Regularization helps to solve overfitting problems by reducing the variance/complexity of the model under consideration and shrink the coefficients of the independent features. It converts a complex model into a simpler one to avoid the risk of overfitting. It is very important concept of Machine learning.

**14.** Which particular algorithms are used for regularization?

**Ans.** There are three types of Regularization algorithms:

- Ridge Regression – Ridge regression is used for analysing data that suffer from multi-collinearity. Ridge regressions adds a penalty L2 to the loss function that is equivalent to the square of the magnitude of the coefficients.
- Lasso Regression- Lasso is a regression analysis method that performs both features selection and regularization in order to enhance the prediction accuracy of the model. Lasso regression adds a penalty L1 to the loss function that is equivalent to the magnitude of coefficient.
- Elastic Net Regression- Elastic Net is regularized regression that linearly combines the L1 and L2 penalties of Lasso and Ridge methods.

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

**Ans.**