**\*\* MACHINE LEARING WORK SHEET ANSWERS \*\***

* **Objective type questions :-**

1. Which of the following methods do we use to find the best fit line for data in Linear Regression?
   * **Least Square Error.**
2. Which of the following CORRECT statement is true about outliers in linear regression?
   * **Linear regression is sensitive to outliers.**
3. A line falls from left to right if a slope is?

* **Negative.**

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

* **Both of them.**

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

* **Low bias and high variance.**

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

* **Predictive modal.**

1. Lasso and Ridge regression techniques belong to?

* **Regularization.**

1. To overcome with imbalance dataset which technique can be used?
   * **Cross validation.**
2. The AUC Receiver Operator Characteristic (AUCROC) curve is an evaluation metric for binary classification problems. It uses to make graph?
   * **TPR and FPR.**
3. In AUC Receiver Operator Characteristic (AUCROC) curve for the better model area under the curve should be less.

* **False.**

1. Pick the feature extraction from below:

* **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?
   * **We don’t have to choose the learning rate**
   * **It becomes slow when number of features is very large**

# Subjective type questions :-

1. Explain the term regularization?

**Ans :- Regularization is a technique used to reduce the errors by fitting the function appropriately on the given training set and avoid over fitting.**

1. Which particular algorithms are used for regularization?

**Ans :- There are three algorithms are used for regularization.**

1. **Lasso Regression. (L-1)**
2. **Ridge Regression. (L-2)**
3. **Dropout.**
4. Explain the term error present in linear regression equation?

**Ans :- The vertical distance between the data point and the regression line is non as error.**

Mathematical Approach **:**

**Error = Actual Values – Predicted Values.**