**Tanya Jain – 19A1HP078**

**Learnings from Practitioner session: -**

From any given dataset we expect to identify the underlying pattern in the dataset.We need Dataset, Model And Algorithm to properly understand the pattern. In Dataset we need to give our input in the Dataset. In Model we need to understand the pattern using mathematical expression. In Algorithm we need to derive the pattern. There are 2 types of MLA which are Supervised and Unsupervised. In Supervised, the ground rules to test the Dependent Variable. In Unsupervised, the data scientist needs to understand the outcome. The Gradient Descent Algorithm is the mother of all neural networks and MLA. Its used to optimize the cost function which is the function of Mean Squared Error. Implementation is done through Bias and Weights. There are two ways of scaling the data i.e. Minmax scalar and Standard Scalar. Scikit learn library is used in Dimensionality reduction, Regression, Classification and Clustering problems. The overfitting problem in the train dataset can be removed by using Regularization. These two are the methods of regularization which are adding a penalty to the cost function. In LASSO, summation of the absolute value of the coefficient is taken into consideration. In RIDGE, summation of squared value of coefficient is taken into consideration. Process of Bagging and Boosting was taught by the practitioner. The samples built parallel models and aggregating the results for final prediction is called Bagging. Whereas in Boosting series of Models are built as opposed to bagging.