**Individual Project 1**

**CS430-01**

**Machine Learning on Cloud**

**Fall 2022**

**Exploratory Data Analysis, train the data using classification techniques**

**Goal:** The goal of this assignment is to use Pandas/Matplotlib/Seaborn to explore the dataset, use Sciket-learn libraries do the data preprocessing, split the data and train models with classification techniques.

**The dataset you will have to use :**

Select a relevant dataset from Kaggle or UCI machine learning library.

**You will need to compare four classification techniques (Logistic Regression/ SVC/ DT/ RF) using Randomized search Cross Validation technique or Voting classifier to detect the best model.**

**Prepare a report using the template provided.**

Ensure your notebook is organized and has proper **Markdown comments on data preprocessing, model performed,** etc. You can assume that after someone see the raw notebook, so it should be clear.

1. **Assignment Submission:** Upload a link to your GitHub repository for the project in the area provided in Moodle by the deadline specified.
2. Prepare a PowerPoint on your findings and present it on 10/21/2022 in class.