Subject: Intermediate Analytics (ALY 6015)

Professor: Richard He

Group members:

* Debpriyo Roy
* Harshal Kumbhar

We are going to use Cardiovascular Disease dataset retrieved from Kaggle.com. The dataset consists of 70 000 observations of patients, with 11 different variables starting with Age, Height, Diastolic, Cholesterol Smoking etc. are one of the few of the parameters considered under the dataset responsible for cardiovascular diseases, depending on which we have we have to select the dependent/target variable, in our case we have” cardio” as target variable. All the data is collected during the medical examination of patients to have precise details of the situation.

We are going to perform different classification methods for analyzing and processing the data with respect to which we wish to predict the presence or absence of cardiovascular diseases in patients’ body on the basis of available parameters.

Kaggle Dataset link provided below for reference:

<https://www.kaggle.com/sulianova/cardiovascular-disease-dataset>