ECMM 6014 – DATA MINING

Project Topic

Assessment of Obesity Levels Using Physical Health Condition and Eating Patterns from Dataset

Group Members

Mohammad Shahedul Islam, B00807172
Kazi Alimul Alam, B01025571

Project Topic:

Assessment of Obesity Levels Using Physical Health Condition and Eating Patterns from Dataset

Dataset Source:

https://www.kaggle.com/datasets/berkayuras/obesity-prediction-classification-and-pca

Dataset Description:

This dataset contains information for estimating the prevalence of obesity in people from Mexico, Peru, and Colombia based on their physical characteristics and eating patterns.

Insufficient Weight, Normal Weight, Overweight Level I, Overweight Level II, Obesity Type I, Obesity Type II, and Obesity Type III are the values that can be used to classify the data, which consists of 2111 records and 17 attributes. The records are labeled with the class variable NObesity (Obesity Level).

Project Goal:

- Compare obesity level between Male and female based on Age factor and their use of technology (insights)
- More insights from data such as scatter plots, bar graphs(visualize the dataset), distributions of multiple factor relate to obese.
- Build an algorithm that can predict whether a patient has Obese, based on their health data.
- Find relationship between drinking more family history, alcohol, smoke and being Obese by performing Correlation and logistic regression.

In Scope:

Learn and explore data mining technique and work with this health data set and accomplish project goal in timely manner.

Team Member:

Mohammad Shahedul Islam, Kazi Alimul Alam

Assumptions:

After completing this project, we will be able to properly grasp the necessary data mining skill with hand-on project experience and will be confident to perform this technique in any raw datasets with proper technique.

Date: Feb 26th, 2025