Healthy Lifestyle Recommendation System

Nia Smith

Obesity is a continuously growing chronic health problem in the United States and the world. In 2022, 1 in 8 people were living with obesity. In 2022, 43% of adults aged 18 years and over were overweight and 16% were living with obesity. In 2022, 37 million children under the age of 5 were overweight. These statistics clearly are not good and need improvement.

Obesity and being overweight comes from an imbalance of energy intake through diet and energy expenditure through physical activity. Multiple factors can influence obesity such as environmental factors, psychosocial factors, and genetics. The environment heightens the likelihood of obesity in individuals and populations due to the lack of safe and easy physical mobility and the lack of available healthy food options.

Even with prior knowledge on ways to combat obesity, there is still no true recommendation. In this project, I will create a recommendation system that will help people identify distinctive options for a healthier lifestyle. Many people would reap the benefits of being to remain fit and healthy. The entire population of people regardless of gender, race, and age can benefit from this new recommendation system.

For this data, I gathered data from a dataset on Kaggle.com that shows the obesity levels based on eating habits and physical activity. 77% of the data was generated synthetically using the SMOTE filter and the WEKA tool while the remaining 23% of the data was collected directly from users through a web platform. The goal that I plan to achieve is to be able to recommend healthier methods to improve lifestyle. This data is a good dataset to use for prediction among people in Peru, Mexico, and Columbia due to 17 attributes and 2111 records.

The data is prone to discrepancies due to varying estimations. The class variable NObesity level is important in visualizing where exactly an individual stands with the classification of their health. Values of the NObesity include Insufficient Weight, Normal Weight, Overweight Level I or II, and Obesity Type I, II, or III.

In the Obesity levels notebook, I showed the daily meals intake compared to the weight of the individual. The daily water intake and the height were also shown. Encoding is a required pre-processing step that is necessary prior to fitting and evaluating a model. Ordinal Encoder was used for this data, but OneHot Encoder could also be used. Random Forest regressor was explored in this notebook. Extraneous information that is not necessary for the recommendation system can be deleted.
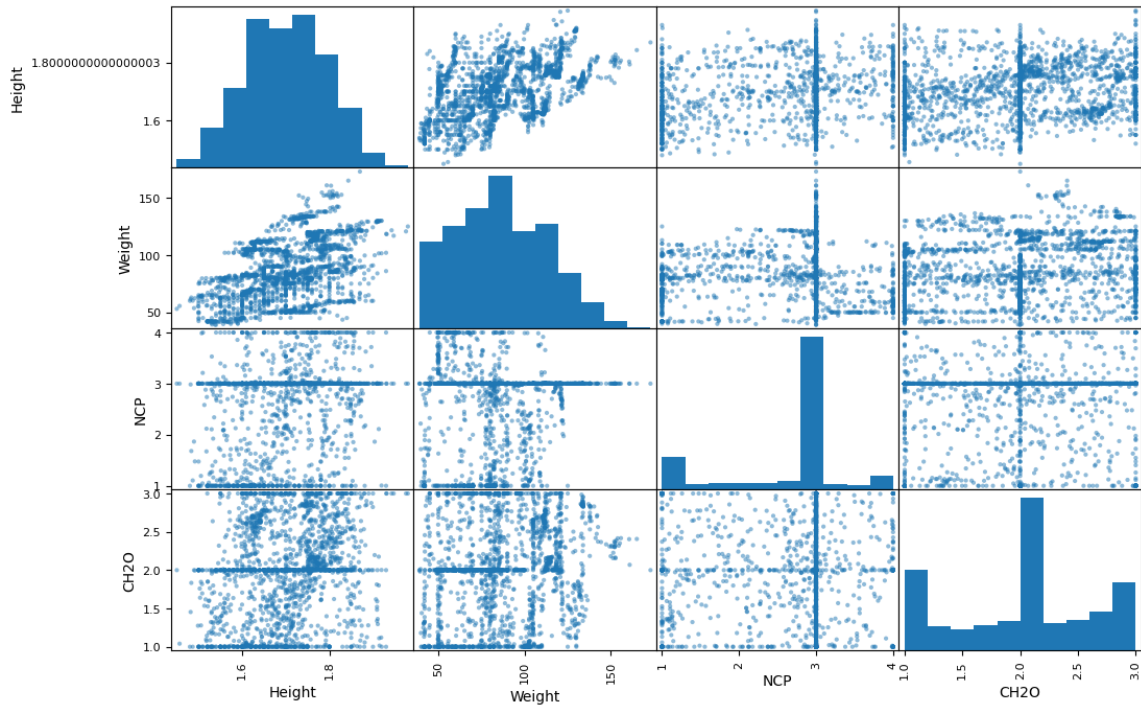
The graph above shows the correlation between the weight, height, the daily amount of water, and the

Three recommendations that are used today are evidence-based method, collaborative based method, and a combination method. The evidence-based method is strictly following and adhering to methods that were previously shown to work for maintaining healthy weight and not becoming obese or overweight. The collaborative based method is when methods are suggested by others who may have heard about success stories but are not proven to work. A combination method is using a mixture of both methods to ensure that weight is maintained at a healthy weight and not too much or too little.

The best method that I chose to work with was the combination method because people can choose how much of one method to adhere to and when. Unfortunately, exact foods were not described in how to improve the obesity level, therefore, evidence-based method by itself could be deemed as challenging. In the future, I would like to work with new novel combination methods.