CS 210 Project Proposal: Group 2

Ahmet Furkan Ün - Efe Tüzün - Ege Demirci - Taner Giray Sönmez

Heart Disease Prediction

Introduction

According to the World Health Organization, 17.9 million people die each year from heart disease around the world[1]. One of the leading causes of mortality among the world's population is heart disease, thus the prediction of cardiovascular diseases is of importance. In the last few years, the global burden of cardiovascular diseases has been quickly increasing. Many studies have been carried out in an attempt to define the most significant factors in heart disease detection and to precisely forecast the overall risk. Heart disease is also referred to as a "silent killer" because it causes death without noticeable symptoms. Early detection of cardiac diseases is critical for implementing lifestyle modifications in high-risk people and, as a result, reducing consequences.

Problem Statement

One of the biggest challenges in heart disease is its detection. While there are existing tools that can predict heart diseases, calculating the probability of heart diseases is either expensive or inefficient. Early detection of heart diseases is extremely important as it can reduce the death rate and complications it will lead to. However, it is not possible to accurately monitor patients every day. Moreover, since this monitoring requires experience, time, and expertise, the patient cannot be consulted by the doctor 24 hours a day. Since we have a good amount of data in today's world, we can come up with a solution to this problem by analyzing the data and using various machine learning techniques. In conclusion, medical data can be used for health diagnosis using appropriate analysis and methods.

[1] "Cardiovascular Diseases," World Health Organization (World Health Organization), accessed April 6, 2022, https://www.who.int/health-topics/cardiovascular-diseases#tab=tab_1.

Methodology

Since heart diseases cannot be noticed easily, there are increasing concerns about their consequences. Therefore, predicting the probability of these deadly diseases will be the most significant goal of this project. Machine Learning techniques will play an important role in this regard. Although heart diseases can present themselves in different ways, they are associated with certain health factors and lifestyle choices. We can draw conclusions by collecting health data from various data sources, classifying them under appropriate attributes, and finally analyzing them to extract the desired data.

Motivation

The key motivation for conducting this research is to develop a machine learning model for predicting the occurrence of heart diseases. In other words, this project is aimed to develop an optimal algorithm for determining the chance of heart disease given a patient. Even though common machine learning algorithms will be used extensively, heart disease prediction is a critical task that requires the highest level of accuracy. Therefore, algorithms are assessed at various levels and using various evaluation methodologies. Researchers and medical professionals will be able to establish a better understanding of the condition as a result of this.

Expected Outcome

At the end of this project, it is expected that the machine learning algorithm, which will be developed using the processed data, will produce a better probability than the random chance of a person being diagnosed with heart disease. In the duration of this project, the main aim is to better understand what conditions are linked to heart diseases and what factors play a role in diagnosing a person with heart disease. We expect as a group to improve our knowledge on visualization techniques, missing value and outlier detection and treatment, feature engineering, statistical analysis tools and machine learning algorithms.

Bibliography

"Cardiovascular Diseases." World Health Organization. World Health Organization. Accessed April 6, 2022. https://www.who.int/health-topics/cardiovascular-diseases#tab=tab_1.