Heart Disease Prediction

Project Proposal

Prepared For:
Selim Balcısoy

Prepared By:
Hüseyin Alper Karadeniz (28029)
Ahmet Bilal Yıldız (27925)
Safa Abdullah Söğütlügil (29214)
Pelinsu Saraç (28820)

Overview

Heart diseases have serious effects on human life, and they are still one of the leading causes of death around the globe despite all advancements that took place in the healthcare industry.[1,2] The part which makes things harder to solve is that there are many types of heart diseases, which have different impacts on blood vessels and heart, and have several causes from a wide range. Also, different manifestations of heart disease can make the diagnosis process take longer. It is now a common fact that high blood pressure, smoking, high alcohol intake, diabetes, high stress level, genetic factors and many other parameters have a significant effect on heart diseases.[3]

In our project, we propose the development of a machine learning model which predicts the risk of heart disease of a person by looking at some related health parameters. This will be achieved by analyzing various types of past records of patients with heart disease from different datasets, that include possible variables that contribute to the occurrence of heart diseases with these patients.

Objectives

- To investigate the relationships between the occurrence of heart disease and several health parameters such as resting blood pressure, fasting blood sugar, cholesterol level, exercise angina, chest pain type, age, sex, ST slope, and more.
- To develop a machine learning prediction model which predicts the occurrence risk of heart diseases by using the investigated health parameters.
- To visualize the processed data with different types of data visualization techniques such as histogram, scatter plot, bar chart, boxplot etc.
- To raise awareness about the risk factors of heart disease.

[1] https://www.cdc.gov/heartdisease/index.htm
[2] https://www.who.int/health-topics/cardiovascular-diseases#tab=tab_1
[3] https://www.cdc.gov/heartdisease/about.htm

Requirements

The project has some requirements to be carried out. One of the most crucial parts of them is the data, because the outcomes of the project significantly depend on the inputs. The needed health data will be collected from various datasets from reliable resources; thus, there will be no need to conduct a data collection process for the health records that will be analyzed. Some additional parameters related to the topic can be identified and added to the dataset such as HbA1c levels of diabetics. Also, existing parameters can be gathered under different subgroups such as changeable and unchangeable risk factors to have a better understanding of the data. In order to do so, some research on the topic might also be needed.

Data will be analyzed in Python, which will be the main programming language of the project, so the knowledge of Python programming language and a computer that is sufficient to handle programs written in Python are some fundamental needs for the project. In order to provide a more efficient and a more convenient way for the data visualization and data analysis parts, additional Python libraries such as NumPy, Pandas, Matplotlib, and more will be used. Also, no budget needed for the project.

Significance

Heart disease is a common cause of morbidity and mortality around the globe. By foreseeing potential cardiovascular diseases, patients can be referred to a physician even before they develop any symptoms. This can decrease the rates of morbidity and mortality, prevent unnecessary hospitalization, and allow people to try lifestyle changes before having to use medication, which could save them from lifelong costs and potential side effects of these medications. This project can also be expanded as an early warning service in terms of health issues for everyone, but especially useful for the people who don’t have enough time or money to spend on their health-related issues.