Coronary Heart Disease predicting using machine learinng

Aladin Kapić   
Faculty of Engineering and   
Natural Sciences  
International Burch University, IBUSarajevo, Bosnia and Herzegovina  
aladin.kapic@stu.ibu.edu.ba

Dr. Jasmin Kevrić  
Faculty of Engineering and   
Natural Sciences  
International Burch University, IBUSarajevo, Bosnia and Herzegovina  
jasmin.kevric@ibu.edu.ba

*Abstract* - TBW

Keywords - CHD, heart disease, artery disease, FHS, heart attack, arrythmia, heart failure

# Introduction

There are many types of heart disease, and each one has its own symptoms and treatment. People started seriously researching causes of them since beginning of twentieth century. Framingham Heart Study (FHS) was launched in 1948 and original goal was to identify common factors or characteristics that contribute to cardiovascular disease. FHS had over 14,000 people from three generation, since its founding. Now, the data collected is used in combination with machine learning algorithms to analyze body condition, predict diseases like CAD, CHD, etc. With that aspect of view, our data is based on Coronary Heart Disease (CHD) and in next chapter we are going to talk about causes, treatments and algorithms that provides either adequate support for doctor’s opinions or exact answers of such disease. Since 1948 till now, it appears to be the most famous and influential investigation in cardiovascular disease epidemiology.

Coronary heart disease (CHD) is a disease in which a waxy substance called plaque builds up inside the coronary arteries. These arteries supply oxygen – rich blood to our heart muscle and if plaque occurs over time, angina or a heart attack can occur. If it’s present, CHD can weaken the heart muscle and lead to heart failure and arrhythmias. If a clot becomes large enough, it can block a coronary artery and lead to heart attack which can be fatal for person. To prevent that, we struggle with understanding exact causes, trying to predict disease based on historical data.

Most significant factors like smoking, high blood pressure, high level of sugar in the blood and high levels of certain fats and cholesterol in the blood can lead to coronary heart disease. Plaque might begin to build up where arteries are damaged. It may start in early childhood, but most often in adults. CHD is leading cause of death for both genders. Since mayor risk factors are smoking habits, blood pressure, cholesterol and diabetes, we are able to pull out information and make prediction of Coronary Heart Disease based on Framingham Heart Study data set, which contains more then enough data to train and test our algorithms.

Many studies have been made since 1948 on this particular topic, including work with all dataset variables and simplified models made by removing some of input data. Some of them are based just on males like [Deconvolution by Weighted Kernels](https://rdrr.io/rforge/DeconWK/) which automatically reduce number of input data and reduces power of algorithm. One interesting work made by Aman Ajmera (which is available on www.kagle.com) , where blood pressure and smoking habit is completely irrelative, gave average results with coefficient of accuracy about 0.75 which represents low accuracy according to model that includes all available inputs, relative to disease, and gave higher accuracy of 0.1 than the first one. Leaving that on side, today, 1,300 scientific papers later, the risk factors behind cardiovascular disease – especially dietary intake – are common knowledge, thanks to the Framingham study. Not only scientists and doctors are up for this study by the time, but most regular people that prefer and practice healthy life.

Main goal of this study is, besides helping doctors reading results, implementation in particular smart device that can process in real time all mayor factors that can and would lead to any sort of heart disease, especially to the Coronary Heart Disease. Real time reading data will process on FPGA device and obtain all results compared to training. According to training results and given real time results, it’ll be possible to diagnose state of patient and prevent any additional problems and health issues. Measuring blood pressure, systolic and diastolic BP, pulse rate and knowing other not fast changing or never changing factors as overweight, age, gender and cholesterol, also using as a part physical condition – is body on move or at state of rest, we should be able to determine possibility of Coronary heart disease.

At this time, we are not able to conclude nature of superior algorithm suitable to give best result but till the end of year, we would determine best way to obtain willing result and integration in device.

Reference

1. National Heart, Lung and blood institute
2. Textbook of Medical Physiology, Arthur C. Guyton and John E. Hall
3. Framingham Heart Study