

Heart Disease Prediction Using Artificial Intelligence

Isha Jagat
Department of Biomedical Engineering
National Institute of Technology, Raipur
Email: ishajagat1999@gmail.com

November 2021

1 Abstract

Artificial intelligence approaches have been widely employed in clinical decision support systems to accurately anticipate and diagnose a variety of diseases. Because of their capacity to uncover hidden patterns and relationships in medical data provided by medical practitioners, these classification approaches are highly useful in the development of clinical support systems. Because heart disease is one of the leading causes of death worldwide, one of the most important applications of such systems is in the diagnosis of heart disease. Almost all systems that forecast cardiac illnesses rely on clinical datasets that include parameters and inputs from lab testing. None of the algorithms can predict heart disease based on risk variables such as age, family history, diabetes, hypertension, high cholesterol, cigarette use, alcohol consumption, obesity, and physical inactivity. Many of those apparent risk factors are shared by heart disease patients, and they can be utilised to diagnose the disease quite well. A system based on such risk factors would not only assist medical experts, but it would also alert patients to the possibility of heart disease before they visit a hospital or undergo costly medical examinations. As a result, employing several Classifying Algorithms, this research proposes a technique for predicting heart disease using main risk factors. K Neighbors, Support Vector, Decision Tree, and Random Forest methods are among the four major classification algorithms used in this technique.

2 Introduction

The diagnosis of heart diseases because it is one of the leading causes of deaths all over the world. Almost all systems that predict heart diseases using clinical dataset having parameters and inputs from complex tests conducted in labs. None of the systems predicts heart diseases supporting risk factors like age, case history, diabetes, hypertension, high cholesterol, tobacco smoking, alcohol intake, obesity or physical inactivity, etc. Heart disease patients have many of those visible risk factors in common which may be used very effectively for diagnosis. A system based on such risk factors would not only help medical professionals but it would give patients a warning about the probable presence of heart disease even before the patient visits a hospital or goes for costly medical check-ups. Hence this paper presents a technique for prediction of heart disease using major risk factors with help of different Classifying Algorithms. This technique involves four major classification algorithms such as K Neighbors, Support Vector, Decision Tree, Random Forest algorithms. The survey contains the data from the prescribed patients from 5 different cities of India i.e. Delhi, Chennai, Bangalore, Kolkata and Hyderabad. The attributes that is defined as the features for the model are the different demographic details of the patients like Age and Gender along with the different Risk Factors which we have defined previously. Here the predictor variable is suffering from Heart Problem or Not. The main intension of this paper is to help in the decision making of a doctor for detecting the possibility or identifying the patient's suffering from any heart attack. Apart from the above mentioned, this method should diminish the False Negative Rate of the prediction. It is the number of the actual positives which is negative through

the prediction to the total negatives. In statistical hypothesis testing, this ratio is represented by the letter α . The reason is when a patient has a certain rate to suffer from the heart any heart attack and the model predicts “no problem”, is which the model should be trained to diminish and that makes this product different from others in the market. In the following sections we will discuss the different terminologies and factors related to this project and the methodology of HPPS, which can be a partner of the doctor in the decision making of whether the patient is going to suffer from any heart attack or not. In the next section we will discuss about the factors that we have taken for the survey and their correlations with the predictor output, followed by the proposed model and scenarios and lastly with the results for the selection of the algorithms.

3 About CVD

Cardiovascular disease (CVD) remains the leading cause of death for adults in the United States (US) with an estimated. 85.6 million Americans experiencing some form of CVD. The term CVD is used to describe disorders of the heart and blood vessels such as coronary heart disease, stroke, congestive heart failure, and arrhythmias. African Americans comprise 13.3% of the US population (46.3 million people) yet have a three-fold greater risk of developing CVD and a two-fold greater risk of CVD related mortality than that of non-Hispanic whites and other ethnic groups.

4 Application of AI in CVD

AI technologies have been applied in cardiovascular medicine including precision medicine, clinical prediction, cardiac imaging analysis and intelligent robots. There are optimistic prospects of the use of AI in cardiovascular medicine. Clinical prediction: Through machine learning and big data analytics, AI can help clinicians to make more accurate predictions for patients. Research from Dawes TJW suggests that AI can predict possible time periods of death for heart disease patients. In their research, AI software recorded the results of cardiac magnetic resonance imaging (MRI) scans and blood tests of 256 heart disease patients. The software measured the movement of 30,000 points that are marked on the heart structures in each heartbeat. By combining these data with the patients’ eight-year health records, AI could predict the abnormal conditions that will lead to patient death. Additionally, their software was able to predict the survival rates of patients for the next five years, and the prediction accuracy of the next year survival of patients could easily reach 80%. However, the prediction accuracy of the clinicians was only 60%. Furthermore, Motwani M and his colleagues established a predictive model through the use of deep learning, in order to evaluate the risk of death for the next five years, for 10,030 suspected coronary heart disease (CHD) patients. Their results indicated that the risk assessment based on AI is superior to traditional clinical judgement and coronary computed tomography angiography.

Cardiac imaging analysis: AI can also be used to analyze echocardiographic images, including automatic measurement of the size of each chamber and assessment of left ventricular function. Moreover, it can be used to assess structural diseases, such as valvular disease, to help determine the classification and staging of the disease.

convolution neural networks: by using the echocardiographies of 267 randomized patients (age range: 20–96 years) between 2000 and 2017 from the university medical center. From the perspective, 223,000 images were divided into fifteen categories. Furthermore, this classification algorithm has outperformed the human cardiovascular physicians in the classification competition of cardiac ultrasound images.

Intelligent robots: AI can perform cardiac interventional operations, such as percutaneous coronary intervention (PCI) operations and catheter ablations of atrial fibrillation, on patients; which will reduce the radiation exposure for the clinicians from the use of digital subtraction angiography. With the use of reinforcement learning, the ability of AI will be far superior to that of a human being, especially for uses in repetition drills. Thus, the AI will learn how to more quickly perform operations than human clinicians. Altogether, the combined use of AI and surgical robots will promote the revolution of traditional medicine.

5 Dataset

The numerous demographic information of the patients, such as age and gender, as well as the distinct Risk Factors that we have described previously, are defined as features for the model. The predictor variable here is whether or not the person has a heart problem. As a result, there are numerous terminology that must be defined. The following are a few of them:

1. Atherosclerosis-related heart disease: In this situation, the walls of the arteries stiffen or harden as a result of fatty deposits known as plaques in medical terms.
2. Cerebrovascular disease: This is caused by a blockage in the blood vessels that supply blood to the brain.
3. Ischemic heart disease: This is caused by cholesterol deposits on the inner walls of the arteries. This is similar to the water flow in the pipes, with the heart acting as the water pump.
4. Hypertensive heart disease (HHD): This is caused primarily by excessive blood pressure. The heart is one of the key organs that helps in the transfer of oxygenated blood and nutrients as well as the removal of wastes from the body, thus there are many more than the ones listed above. We assigned a value of 1 to heart-related difficulties and a value of 0 to no heart-related problems in the expected value. This challenge has now been transformed into a binary classification problem.

6 Risk Factor

Family History, Smoking, Hypertension, Dyslipidemia, Fasting Glucose, Obesity, Life Style, CABG, and High Serum in Blood are some of the risk variables considered in this study. Aside from the aforementioned risk factors, the patients' demographic information is also recorded. The most crucial thing that each diagnostic should avoid is exposing a healthy human body to radioactive radiation from a CT scan.

7 Work

The trouble with heart disease risk factors is that there are so many of them, such as age, cigarette use, blood cholesterol, fitness, blood pressure, stress, and so on, that identifying and categorising each one according to its value is difficult. In addition, heart disease is frequently identified only after a patient has progressed to an advanced stage of the condition. As a result, risk factors were examined from a variety of sources. Sex, age, family history, blood pressure, smoking habit, alcohol use, physical inactivity, diabetes, blood cholesterol, poor diet, and obesity were all included in the dataset. The system will indicate weather the patient is in risk or not.

8 Methodology

The identified important risk factors and the corresponding values and their encoded values in brackets, which were used as input to the system.

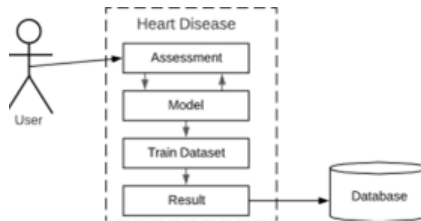


Figure 1: Block Diagram

Data analysis has been carried out in order to transform data into useful form, for this the values were encoded mostly between a range $[-1, 1]$. Data analysis also removed the inconsistency and anomalies in the data. This was needed. Data analysis was needed for correct data preprocessing. The removal of missing and incorrect inputs will help the neural network to generalize well. The

proposed application is developed using python and is capable of identifying if a patient has heart disease or not. There are number of factors which increased risk of heart diseases, like family history of heart disease, smoking, cholesterol, high blood pressure, obesity, lack of physical exercise etc. Heart disease is a major health problem in today's time. Thus, there is necessity to develop a system which will predict the heart disease using Artificial intelligence. In this project we have implemented AI algorithms such as

1. K Neighbours Classifier
2. Support Vector Classifier
3. Decision Tree Classifier
4. Random Forest Classifier

Which can predict heart disease So in this project we are first taking input from doctor about heart related information that is smoking, cholesterol, high blood pressure, or whether the patient has Diabetes etc. and based on the factors our system will predict the heart disease from given algorithms and it will generate a detailed report of the heart disease So from this we can define that which algorithm is best for prediction of Heart disease.



Figure 2: System Architecture Diagram

9 Conclusion and future Enhancement

The proposed application uses Risk Factors, which need to be identified by Medical Professionals before using the application. The result may vary based on the identified Risk Factors. If the Risk Factors identified are less accurate or wrong, the application may give wrong results. The application may use different AI techniques to capture and correct response based on past experiences. The result of the application depends on the accuracy of the Classification Algorithms. If the accuracy is low, the result generated may be wrong or less accurate. Increasing the dataset, may result in more accurate results. We can build an intelligent system which could predict the disease using risk factors hence saving cost and time to undergo medical tests and check-ups and ensuring that the patient can monitor his health on his own and plan preventive measures and treatment at the early stages of the diseases.

10 References

- [1] "Global atlas on cardiovascular disease prevention and control", WHO, 2011.
- [2] Mozaffarian D, Wilson PW, Kannel WB, Beyond established and novel risk factors: lifestyle risk factors for cardiovascular disease. *Circulation* 117: 3031–3038, 2008.
- [3] Poirier P, Healthy lifestyle: even if you are doing everything right, extra weight carries an excess risk of acute coronary events. *Circulation* 117:3057–3059, 2008.
- [4] Wood D, De Backer, Prevention of coronary heart disease in clinical practice: recommendations of the Second Joint Task Force of European and other Societies on Coronary Prevention. *Atherosclerosis* 140: 199–270, 1998.
- [5] Anderson KM, Odell PM, Cardiovascular disease risk profiles. *Am Heart J* 121: 293–298. 1991.
- [6] Rodgers, Anthony, et al. "Blood pressure and risk of stroke in patients with cerebrovascular disease." *Bmj* 313.7050 (1996): 147.
- [7] Gertler, Menard M., et al. "Ischemic heart disease." *Circulation* 46.1 (1972): 103-111.
- [8] Diamond, Joseph A., and Robert A. Phillips. "Hypertensive heart disease." *Hypertension research* 28.3 (2005): 191-202.

[9] Leander, Karin, et al. "Family history of coronary heart disease, a strong risk factor for myocardial infarction interacting with other cardiovascular risk factors: results from the Stockholm Heart Epidemiology Program (SHEEP)." *Epidemiology* 12.2 (2001): 215-221.

[10] US Department of Health and Human Services. "The health consequences of smoking: a report of the Surgeon General." (2004): 62.