'National Conference on Advancements in Computer Science and Engineering'
International Journal of Scientific Research in Computer Science,
Engineering and Information Technology | ISSN : 2456-3307 (www.ijsrcseit.com)

Heart Disease Prediction using Artificial Intelligence

Dickshith Raj N, Ganesh, Sachin Y M , M Nirmala*

*Associate Professor, Computer Science and Engineering, New Horizon College of Engineering, Bangalore, Karnataka, India

Department of CSE, New Horizon College of Engineering, Bengaluru, Karnataka, India

ABSTRACT

Clinical decision support systems have made extensive use of artificial intelligence approaches for accurate disease prediction and diagnosis. Due to their capacity to uncover hidden patterns and connections in the medical data supplied by medical practitioners, these classification algorithms are extremely useful when creating clinical support systems. Given that heart disease is one of the top causes of mortality worldwide, diagnosing heart disease is one of the most crucial purposes for such systems. Nearly all cardiac disease prediction systems use clinical datasets with parameters and inputs derived from intricate lab procedures. None of the algorithms can forecast cardiac illnesses based on risk factors such age, family history, diabetes, high blood pressure, high cholesterol, smoking, drinking, obesity, and inactivity, among others. Many of the obvious risk factors that can be used to diagnose heart disease are shared by patients. A system based on such risk factors would benefit medical professionals as well as patients by alerting them to the possibility of heart disease before they enter a hospital or undergo pricey diagnostic tests. So, using various Classifying Algorithms, this paper presents a method for predicting heart disease using key risk factors. The K Neighbours, Support Vector, Decision Tree, and RandomForest algorithms are only a few of the four main classification algorithms used in this method.

Keywords— Heart, AI, K Neighbors, Risk Factors, Algorithms.

I. INTRODUCTION

With 85.6 million Americans believed to have some type of cardiovascular disease (CVD), it is still the primary cause of adult mortality in the United States (US) [1, 2]. Heart and blood vessel illnesses such coronary heart disease, stroke, congestive heart failure, and arrhythmias are referred to as CVDs. In comparison to non-Hispanic whites and other ethnic groups, African Americans, who make up 13.3% of the US population (46.3 million people), had a thrice higher risk of developing a CVD and a twofold higher risk of CVD-related death [3,4,5,6].

The most prevalent kind of CVD and the main factor in one-third of all adult fatalities in the US is coronary heart disease, a condition that results from atherosclerosis [7]. For most racial and ethnic groups, including Hispanics, non-Hispanic Whites, and African Americans, this kind of CVD continues to be the major cause of mortality [5]. The risk of stroke is two to three times higher among African Americans than in non- Hispanic Whites, and the incidence of hypertension is among the highest in the world at 44% [8].

These discrepancies in illness risk and death are caused by a variety of factors. Health behaviors including nutrition, smoking, socioeconomic position, different access to and quality of health care, environmental or neighbourhood impacts, pervasive racial prejudice, and genetic variation have all been hypothesised as factors in African American CVD risk [9,10]. However, not all scenarios are affected by these risk variables, and many of them can be altered with new circumstances or behaviour. To reduce risk, it is necessary to evaluate people's knowledge of and attitudes towards CVD in order to create intervention programmes that will be appealing to the intended audience [6,11].

Previous research on the CVD risk of African Americans included people who lived in the US as a whole or in states in the American South. These findings are constrained because they might unintentionally ignore the distinctive traits of other local populations and cultural variations within communities [6,12]. Although they might not directly apply to other geographical areas or contexts, lifestyle factors in a particular region, like the South, can serve as a jumping off point for new investigations [13]. Regionally specific qualitative insights into African Americans' awareness of CVD risk factors can close research gaps and enable more precise targeting of nutrition and lifestyle-related interventions.

programming that is related [9, 14]. For instance, not all African Americans in or outside of the American South, where 55% of African Americans live, may benefit from dietary interventions focused on modifying traditional Southern or "Soul Food" recipes [15]. African American culture's nutrition and health practices vary due to variances in geographic origin, life experience, social standing, and access to food.

II. PROBLEM STATEMENT

The risk factors that may contribute to the development of heart disease have been identified by recent medical research, but further study is required before this information can be used to lower the incidence of heart disorders. The main risk factors for cardiac illnesses have been identified as diabetes, hypertension, and high blood cholesterol. The main risk factors for heart disease and heart disease are both related to lifestyle risk factors, which include eating habits, physical inactivity, smoking, alcohol consumption, and obesity. According to studies, lowering these risk factors for heart disease can really aid in their prevention. The prevention of heart disease risk has been the subject of numerous studies and investigations.

Based on blood pressure, smoking behaviour, cholesterol and blood pressure levels, and diabetes, data from population studies have assisted in the prediction of cardiac illnesses. Researchers have modified these prediction algorithms to create simplified score sheets that individuals can use to determine their risk of developing heart disease. In algorithms for heart disease prediction, the Framingham Risk Score (FRS) is a well-liked risk prediction criterion. The goal of this work was to create an intelligent system based on classification algorithms for the classification of risk variables in the prediction of heart disease.

III. EXISTING SYSTEM

The current system components use a powerful prediction algorithm to produce comprehensive reports. The patient and the doctor provided the project's input information. Heart illness is then examined utilizing AI algorithms using the doctor's inputs. The results are now compared to those of other models that have been used in the same domain and are found to be improved. The primary goals of the current system are to assess

and compare the results of previous patients with disease to those of new patients with disease in order to predict the likelihood that a certain patient would develop heart disease in the future. By putting the aforementioned model into practice, we can design a system that is more accurate at predicting the likelihood that a new patient would experience a heart attack.
K Neighbours Classifier, Support Vector Classifier, Decision Tree Classifier, and Random Forest Classifier patients collected from the UCI laboratory. Performance and accuracy of the outcomes using various AI algorithms are compared. The model for the Heart Disease Prediction System was developed utilizing a variety of AI approaches and algorithms. But the accuracy is very low when using all the current systems.

IV. LITRATURE REVIEW

People today are so preoccupied with getting what they need and making money that they forget to take care of their physical health. As a result, the food that they eat is different, and their way of life alters. They are more stressed and under a lot of pressure to get money, which leads to blood pressure problems, diabetes, and other ailments at an early age. All ofthose factors lead to a lack of fitness, which raises the risk of heart disease. The heart is the most vital organ in the human body, and if it is damaged, it will also have an impact on the other critical organs. Clinical decisions are frequently decided more on the basis of the physician's experience and intuition than on the knowledge-rich facts concealed in the database. This practice results in unintended biases, errors, and excessive medical costs, which have an impact on the quality of care patients receive. HD is generally diagnosed using the patient's medical history. However, the results of the diagnosis are not an accurate diagnosis of HD. These techniques are also unreliable in terms of computation and accuracy. Numerous articles have put forth various methods for deriving features from heart sounds and categorizing them using neural networks. Clinical choices are frequently made based more on the doctor's knowledgeand experience than on the information-rich facts hidden in the database. This approach has an effect on the standard of care patients receive by causing unintentional biases, errors, and excessive medical costs. The medical history of the patient is typically used to diagnose HD. The results of the diagnostic, though, do not accurately identify HD. Additionally, these methods are inaccurate and unreliable in terms of computation. Various techniques for extracting information from heart sounds and categorizing them using neural networks have been proposed in numerous studies.

V. PROPOSED WORK

The problem with risk factors related to heart disease is that there are many risk factors involved like age, usage of cigarette, blood cholesterol, person's fitness, blood pressure, stress and etc. and understanding and categorizing each one according to its importance is a difficult task. Also, a heart disease is often detected when a patient reaches advanced stage of the disease. Hence the risk factors were analyzed from various sources. The dataset was composed of 12 important risk factors which were sex, age, family history blood pressure, Smoking Habit, alcohol consumption, physical inactivity, diabetes, blood cholesterol, poor diet, obesity. The system indicated whether the patient had risk of heart disease or not.

VI. PROPOSED METHODOLOGY

The data for 50 people was collected from surveys done by the American Heart Association. Most of the heart disease patients had many similarities in the risk factors. The TABLE I below shows the identified important risk factors and the corresponding values and their encoded values in brackets, which were used as input to the system.

TABLE I
RISK FACTORS VALUES AND THEIR ENCODINGS

	Risk Factors	Values
1	Sex	Male (1), Female (0)
2	Age (years)	20-34 (-2), 35-50 (-1), 51-60 (0), 61-79 (1) , >79 (2)
3	Blood Cholesterol	Below 200 mg/dL - Low (-1) 200-239 mg/dL - Normal (0) 240 mg/dL and above - High (1)
4	Blood Pressure	Below 120 mm Hg- Low (-1) 120 to 139 mm Hg- Normal (0) Above 139 mm Hg- High (-1)
5	Hereditary	Family Member diagnosed with HD -Yes (1) Otherwise –No (0)
6	Smoking	Yes (1) or No (0)
7	Alcohol Intake	Yes (1) or No (0)
8	Physical Activity	Low (-1) , Normal (0) or High (-1)
9	Diabetes	Yes (1) or No (0)
10	Diet	Poor (-1), Normal (0) or Good (1)
11	Obesity	Yes (1) or No (0)
12	Stress	Yes (1) or No (0)
Output	Heart Disease	Yes (1) or No (0)

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 incorrectinputs 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 To improve the algorithms' accuracy, we need moredatasets.

Only medical personnel may use the proposed application.Since the suggested application is web- based, mobile devices cannot use it. The effectiveness of the algorithms determines the application's outcome.

Conclusion and further improvement, chapter eight The suggested application employs risk factors, which must be recognized by medical professionals before to usage.

Depending on the identified Risk Factors, the outcome can change. The application could produce inaccurate results if the risk factors it uses are inaccurate or less reliable. Different AI techniques may be used by the application to capture and correct responses based on prior usage.

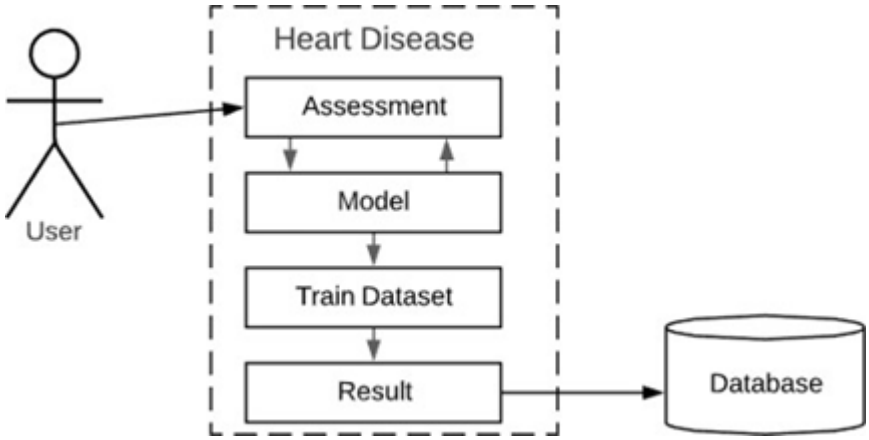

Fig 1. Block Diagram

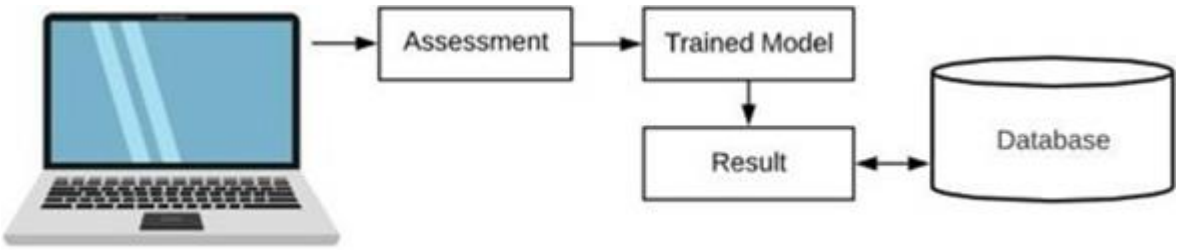

Fig 2. System Architecture Diagram

VII. LIMITATIONS

- Need more datasets, to increase theaccuracy of the algorithms.
- The proposed application can only be used by Medical Personnel.
- The proposed application is Web- based, hence cannot be used in Mobiledevices.

The result of the application depends upon the accuracy of the algorithms.

VIII. 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.

IX. REFERENCES

[1]. World Health Organization, 2011. "Global atlas oncardiovascular disease prevention and control."

[2]. Wilson PW, Kannel WB, Mozaffarian D The risk factors for cardiovascular disease that go beyond known and emerging risk factors are lifestyle factors.2008; Circulation 117: 3031– 3038.

[3]. Poirier P, Healthy lifestyle: high weight carries an elevated risk of acute coronary events even if you are doing everything right. 2008, 117:3057- 3059 copies distributed.

[4]. Wood D. and De Backer's Prevention of Coronary Heart Disease in Clinical Practise: Recommendations of the Second Joint Task Force of European and Other Societies on Coronary Prevention. 199–270 in Atherosclerosis 140, 1998.

[5]. Anderson KM, Odell PM, Risk profiles for cardiovascular disease. Am Heart J 121 (1991): 293-298.

[6]. Anderson KM, Wilson PWF, Odell PM, and Kannel WB. risk profiles for cardiovascular diseases. Am Heart J 121 (1991): 293-298.

[7]. Kannel WB, A study of familial coronary heart disease. Study of Framingham's offspring. 1979, Am J Epidemiol 110: 281–290.

[8]. Effective diagnosis of cardiac disease by neural network ensembles, Expert Systems with Applications, Elsevier, 2009, pp. 7675–7680. [8] R. Das, I. Turkoglu, and A. Sengur.

[9]. Americas Conference on Information Systems, 2009, by T. Porterand B. Green.

[10]. S. Panzarasa et al. (2010) "Data mining techniques for analysing stroke care processes," in Proc. of the 13th World Congress on Medical Informatics.

[11]. T. H. Li, L. Li, Z. Wu, J. Gong, M. Gruidl, J. Zou, M. Tockman, and R. A. Clark, "Data mining techniques for cancer detection using serum proteomic profiling,"

[12]. Elsevier,2004. "Using machine learning algorithms in cardiovascular disease risk evaluation," Journal of Applied ComputerScience and Mathematics, 2009; V. A. Sitar-Tautetal.

[13]. K. Srinivas, B. K. Rani, and A. Govrdhan, "Applications of Data Mining Techniques in Healthcare and Prediction of Heart Attacks," International Journal on Computer Science and Engineering (IJCSE), volume 2, issue 2, pages 250–255, 2010.

[14]. H. Yan et al., "Development of a decision support system for the diagnosis of heart disease using multilayer perceptron," in Proc. of the 2003 International Symposiumon, vol. 5, pp. V-709-V- 712.

[15]. M. C. Tu, D. Shin, and D. Shin, "Effective Diagnosis ofHeart Disease Through Bagging Approach," IEEE, Biomedical Engineering and Informatics, 2009.

[16]. "Neural-based Learning Classifier Systems," IEEE Transactions on Knowledge and Data Engineering, Vol. 20, No. 1, pp. 26–39, 2008. Hai H. Dam, Hussain A. Abbass, and Xin Yao.

[17]. "Intelligent Heart Disease Prediction System Using CANFIS and Genetic Algorithm", International Journal ofBiological and Life Sciences, Vol. 3, No. 3, pp. 157–160,2007. Latha Parthiban and R. Subramanian