Heart Disease Prediction System using Electrocardiogram

[1] Hasnain Tahir, [2] Mazhar Ellahi Qazi, [3] Hamza Asghar, [4] Adnan Kiani
[1,2,3,4] Department of Computer Science, National University of Modern Languages, Rawalpindi, Pakistan.
Email: [1] hasnain.tahir@numl.edu.pk,[2]qazimazhar57@gmail.com, [3] hamza.asghar728@gmail.com,
[4] adnankiani44@gmail.com

ABSTRACT

Heart disease is a major problem in the society. Among numerous deadly diseases, heart disease has gathered much attention in medical research. According to WHO (World Health Organization), millions of people die due to different heart diseases every year. Predictions of heart infirmity is considered as one of the key matter in clinical information analysis. Its diagnosis is very challenging because of its complexity and if the disease is falsely predicted, the treatment provides no cure. The prediction of cardiovascular illness is commonly formed on signs, symptoms and physical check-up of the patient.

In this research, we proposed a new method to identify weather the person has a heart disease or not using an Electrocardiogram report in an easy way. The person only has to calculate the QRS value (Q wave, R wave, S wave) and HR value (Heart Rate) from Electrocardiogram report and enter it in the application. The output will be displayed according to the given input values of the user.

This technique is easy to use and it displays the result immediately. It doesn't contain any heavy or huge algorithm which requires time for execution. It simply asks user for input, and against that input the result is calculated and displayed. This is time saving and user-friendly such that every single person can use it and check whether they might have a disease or not. The technique works on QRS value and HR value present on the Electrocardiogram report and displays result according to that report.

1. INTRODUCTION

The cardiovascular disorder ordinarily refers to conditions comprehending small or blocked arteries that can lead to cardiac attack, chest throb (angina), or stroke. Other heart conditions, such as those affecting your heart muscle, valves, or rhythm, are also considered to be a form of heart disease. Cardiac illness is one of the major origin of illness and death among people in the globe. Predicting heart disorder is considered one of the most primary studies in the clinical data survey phase.

A key obstacle faced by health caution firms, such as hospitals and medical facilities, is the allocation of quality aid at nominal costs [1][2]. The quality aid makes out detecting patients properly and supervising effectual cures.There are number of reasons that escalate the chance of heart illness, in particular smoking cessation, body cholesterol levels, and family background of heart issue, obesity, high blood pressure, and absence of exercise.

The heart is the organ that circulates blood into different pieces of the body through the vessels with an appropriate measure of oxygen and other fundamental supplements. The endurance of any life form

depends exclusively on the appropriate working of the heart and, in case the heart's siphoning activity is problematic, the body's principle bodies like mind and kidneys will go through unfriendly impacts [8].

The most common and mandatory test required for predicting a heart disease is Electrocardiogram (ECG) test. Using that Electrocardiogram (ECG) test we are implementing a system in which the system will predict weather the person is having a heart disease or not. The system will take the input from the user and perform processing on it and then it will show the result whether the user is in danger stage or safe stage.

The main goal of our project is to:

- Predict whether the patient is in safe stage or danger stage.
- Detect heart disease in case the patient is in a danger stage.

2. LITERATURE REVIEW

Some of the previously developed techniques for heart disease prediction are discussed below:

I. CANFIS (Coactive Neuro Fuzzy Inference systems)
II. Data Mining Techniques (Neural Network, Naive Bayes and Decision Tree)
III. Weight Associated Classifier
IV. Data Mining Techniques (Classification and Regression Tree, decision table and Iterative Dichotomized 3)
V. ANFIS (Adaptive Neuro Fuzzy Interface System)
VI. Multilayer Perceptron Neural Network
VII. Hidden Native Bayes
VIII. Random Forest and Evolutionary Method (random forest algorithm, chi-square and genetic algorithm)
IX. Machine Learning Algorithm

CANFIS (Coactive Neuro Fuzzy Inference systems)

CANFIS is a combination of neural network flexibility skills and a logical cognitive approach. It is then combined with a genetic algorithm to test for heart disease. The main unit of CANFIS is the complex axon. Puts in membership methods to entries. The outcome of this axon is calculated by the rule: fj (x, w) = min∀i (MF (xi, wij)). This program can be seen as a three-phase feed-up neural network. The first layer contains input variables, the middle layer or hidden layer has ambiguous rules and the third layer contains output variables. The effectiveness of the CANFIS model was evaluated for the effectiveness of the training and phase adjustment [1].

Data Mining Techniques (Neural Network, Naive Bayes and Decision Tree)

Decision Trees, Naïve Bayes and Neural Network each process has a special ability to capture the meaning of the explained mining objectives. The technique uses the (data mining technique) approach to create mining models, which have six main categories: data comprehension, business understanding, modeling, data preparation, dispatch and testing. Data Mining Extension is the expression used for model training, model prediction, model creation and access to model content. It uses medical features such as age, sex, blood sugar and blood pressure to forecast whether a patient has heart disorder or not [2].

Weight Associated Classifier

This classifier uses Weighted Association Rule. To extract Association rule from data repository, Weighted Support and Confidence Framework is used rather than traditional support and confidence, which helps in mining and classification. At the start the data warehouse of heart disease will be preprocessed for mining process. Every attribute is assigned with a weight value from 0 to 1 depending on their importance in the prediction model. High weight (approx 0.9) will be assigned to attributes have greater impact and (approx 0.1) low weight to the attributes having smaller impact. After the preprocessing the Weighted Association Rule Mining (WARM) algorithm is used to initiate patterns. Rules that are initiated in this process are called Classification Association Rule (CAR). These Rules will be saved in the Rule Base and whenever a new patient enters his specification, the Classification Association (CAR) comes in action and predicts [3].

Data Mining Techniques (Classification and Regression Tree, decision table, Iterative Dichotomized 3)

The three most famous data mining algorithms (Classification and Regression Tree) CART, (decision table) DT and ID3 (Iterative Dichotomized 3) which is drawn out from a rule-based or decision tree classifier to develop prediction models using huge datasets. They also measured the unbiased estimate by using 10-fold cross-validation method. In this system ID3, CART and decision table used to forecast attributes such as age, sex, blood sugar and blood pressure to check whether the patient has a heart disease or not. WEKA ("Waikato Environment for Knowledge Analysis") tool is used for data analysis and implementation. WEKA consists of a set of machine learning algorithms used for data mining problems. Data mining techniques are helpful in carrying out hidden information from large data. Then 10-folds cross-validation method is used to reduce any bias in the procedure and enhance the efficiency of the procedure [4].

ANFIS (Adaptive Neuro Fuzzy Interface System)

This technique represents Takagi-Sugeno-type fuzzy interface system. ANFIS trains initial membership functions by training itself by given dataset, after that it adjusts the membership functions by using hybrid learning algorithm or back propagation algorithm to reduce the error measure. In this network premise parameter and consequent parameters are used. In this ANFIS model, hybrid learning algorithm is used to identify the parameters. In this algorithm there are two types of passes (Forward pass and Backward pass). In forward pass the consequent parameters are computed by LSE (least square estimate) then the error measure is calculated. In backward pass the error signals are distributed backwards to upgrade the premise parameter. For input in this prediction model, seven variables are used. To check the capability of this ANFIS model for diagnosis of heart disease, k-fold cross validation method was used [5].

Multilayer Perceptron Neural Network

The technique works on two steps; first it accepts 13 clinical attributes as an input then the network is trained by training dataset through back-propagation learning algorithm. Multilayer Perceptron Neural Network consists of multiple layers (inner layer, hidden layer and output layer). It is mostly used classification of input patterns, recognition of patterns, approximation and prediction based input information. Back-propagation algorithm is popular for neural network trainings. It generally trains multilayer perceptrons and many other neural networks. In back-propagation, the calculated output is compared with the expected output and then error is calculated. This error is given back to the neural network, weight are adjusted so that the result can become closer to the expected output. This process is repeated for many times till the output comes closer to the required output [6].

Hidden Native Bayes

Hidden Native Bayes (HNB) classifier technique is a more effective classifier as compared to native bayes classifier, because of its attributes dependencies. Hidden Native Bayes needs more time for training, as it is a structure-extension-based algorithm. In this system the data set of heart disease will be an input. Then the HBN algorithm will classify and the output will tell whether the person has a heart disease or not which shows an effective result [7].

Random Forest and Evolutionary Method (random forest algorithm, chi-square and genetic algorithm)

Random forest algorithm is used as a classifier and chi square, genetic algorithm as feature selection measures. Random forest algorithm is a most precise learning algorithm, which is effective for medical applications. Chi square feature selection measure is used for analyses between variables. It also decides whether they are associated or not. Random forest algorithm will be applied on the heart disease dataset and heart stalog dataset. Chi square and genetic algorithm will identify the attributes which are useful in the prediction of the heart disease. Attributes like age, sex, height, weight, BMI etc are taken for heart disease dataset and for stalog; age, sex, blood pressure, chest pain, slop, class etc attributes are taken. The results taken out have shown that the classification done was much accurate and was very much effective in the diagnosis of heart disease [8].

Machine Learning Algorithm

In this technique, Decision tree classifier is used to carry out the required attribute from the provided symptom and datasets. Then support vector machine algorithm (SVM) is used to classify the datasets. These datasets are classified on the basis of kernel and by using hyperplane it groups the dataset. In this system they have also used certain libraries of PYTHON. PANDAS library is used to extract the dataset. SCIKIT LEARN library is used for the implementation of machine learning algorithms (decision tree and SVM) and MATPLOTLIB library is used to check the outcome of the inputs. To implement data operation there are some attributes like age, chest pain, cholesterol, blood pressure etc in our dataset. The decision tree algorithm as well as the Support vector machine both were implemented and then compared. The comparison showed that the decision tree algorithm was effective for the provided datasets [9].

No.	Paper	Techniques										
		CANFIS	Decision tree	Naive Bayes	Neural Network	WAC	CART, DT, ID3	ANFIS	Multilayer Perceptron Neural Network	HNB	random forest algorithm	Machine learning algorithm, SVM
1	[1]	**T**										
2	[2]		**T**	**T**	**T**							
3	[3]					**T**						
4	[4]						**T**					
5	[5]							**T**				
6	[6]								**T**			
7	[7]									**T**		
8	[8]										**T**	
9	[9]											**T**

Table 1: Comparative Analysis of Previous Techniques.

In contrast to the previously developed techniques, our proposed technique is unique in such a way that it diagnoses the human heart health based on the input parameters which are QRS value (Q wave, R wave, S wave) and HR value (Heart Rate) from an Electrocardiogram (ECG) report. These values are a standard and are present in every electrocardiogram (ECG) report. All the patient needs to do is to enter the mentioned values in an easy-to-use application with a very simplified interface. The fuzzy Logic Algorithms helps to match the input from the datasets. Using this methodology, our proposed technique will be able to determine whether the patient has a good heart health or if the heart health of a patient is not good and the patient needs consultation from a cardiologist for further diagnosis.

In rare cases, the patient may end up with a Halt screen output for which there may be various reasons. Some of the reasons associated with a halt screen output are the crash of the application or the breakage of the software tunnel (expo tunnel). One other reason can be that the application didn't find a match of the value with dataset at that time. The application has been tested rigorously by entering data of various patients with different heart health conditions. We tested it on 15 people, after testing them the application showed approx. 95% to 98% accuracy, which is a significant improvement over the previously mentioned techniques.

For further simplicity, the application has a built-in help and a startup guide which can guide the patient on how to enter the desired values in the application in order to determine the patient's heart health condition. The application also has a reference range chart which may be helpful for the patients that find it difficult to enter the values manually in the application. Using the reference range chart, the patient is able to manually see the reference safe ranges for the QRS value and the HR value and compare these values with his/her own values that are present on his/her electrocardiogram (ECG) report. This is a feature that is curently not available on any of the previously developed applications and will greatly enhance the efficiency of the application is being beneficial for the layman patients who find it difficult to enter these values manually.

3. METHODOLOGY

The most common and mandatory test required for predicting a heart disease is Electrocardiogram (ECG) test. Using that ECG test we are implementing a system in which the system will predict weather the person is having a heart disease or not. The system will take the input from the user and perform processing on it and then it will show the result whether the user is in danger stage or safe stage.

We have tested the system by inserting the QRS value and HR value. The values are then matched with the ECG (electrocardiogram) dataset and gives result accordingly. The fuzzy Logic Algorithms helps to match the input from the datasets. There is also a guide provided in the dashboard to calculate QRS value and HR value.

While using are Heart Disease Application, two outputs were planned (.i.e. Good Heart Health!!! No need to worry or Warning!!! You might have disease go and consult the doctor). But while using the application another output (Halt Screen) had occurred. So while using this application we have three outputs.

Figure1 shows the output that the person has no disease. It shows that the person heart health is good.

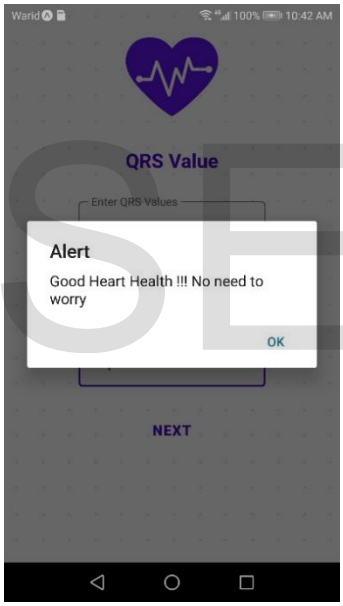

Figure 1 Negative Output

Figure 2 shows the output that the person might have a disease. It shows that the person might have a disease and he must see a doctor.

Figure 2 Positive Output

Figure 3 shows the output of a Halt Screen. This screen can occur due to several reasons.

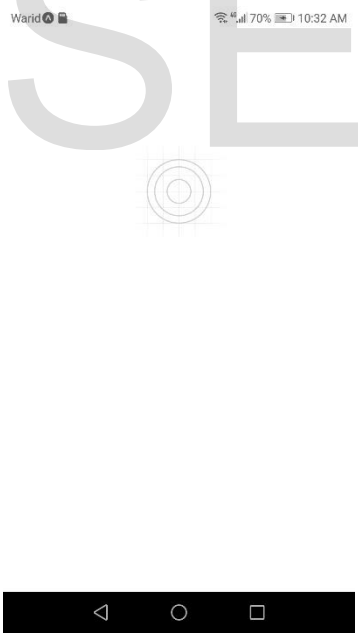

Figure 3 Halt Screen

4. RESULTS

After completing our Heart Disease Prediction System Application, we tested it on different people. Some of them were have a disease and some were not. We tested it on 15 people, after testing them the application showed approx. 95% to 98% accuracy, which is quite good.

Table 2 shows the result of the people having good heart health.

USER	AGE	ORIGNAL REPORT	QRS VALUE	HR VALUE	APPLICATION OUTPUT
1	32	No Disease	85.5	80	Good Heart Health!!! No need to worry
2	39	No Disease	88	84	Good Heart Health!!! No need to worry
3	47	No Disease	94.5	80	Good Heart Health!!! No need to worry
4	33	No Disease	81	75.5	Good Heart Health!!! No need to worry
5	43	No Disease	79.5	80	Good Heart Health!!! No need to worry
6	52	No Disease	89.75	85.5	Good Heart Health!!! No need to worry
7	40	No Disease	83	78	Good Heart Health!!! No need to worry

Table 2 Good Heart Heath Users

Table 3 shows the result of the people having a disease or bad ECG report.

USER	AGE	ORIGNAL REPORT	QRS VALUE	HR VALUE	APPLICATION OUTPUT
1	58	Has Disease	90.5	105	Warning!!! You might have disease go and consult the doctor
2	55	Has Disease	110	104	Warning!!! You might have disease go and consult the doctor
3	48	Has Disease	89.5	57	Warning!!! You might have disease go and consult the doctor
4	57	Has Disease	112	85.5	Warning!!! You might have disease go and consult the doctor
5	47	Has Disease	88	107	Warning!!! You might have disease go and consult the doctor

Table 3 Bad Heart Health Users

Table 4 shows the result of the people which didn't get any result.

USER	AGE	ORIGNAL REPORT	QRS VALUE	HR VALUE	APPLICATION OUTPUT
1	37	Has Disease	88.75	58	Halt Screen
2	55	No Disease	90	86.5	Halt Screen
3	48	Has Disease	89.5	57.5	Halt Screen

Table 4 Users with no Result

The output of halt screen might occur due to the crash of the application or the breakage of the software tunnel (expo tunnel). One other reason can be that the application didn't find a match of the value with dataset at that time.

5. CONCLUSION

The application (Heart Disease Prediction System using Electrocardiogram) is designed and developed to forecast weather the person could have a disease or not in an easiest way. There can be a situation where a patient couldn't understand the Electrocardiogram (ECG) report or the person is not correctly informed

about his/her heart condition. So to overcome these situations, this application was developed. The interface of this application was designed attractive and user friendly so everyone can use it. To get started, the user only needs to create an account by email, name and password. After that he/she can proceed further for disease prediction. There is also a guide available in which the user is helped how to calculate and insert the values. After that the output is shown to the patient whether he has a disease or not. The system accuracy is approx. 95% to 98%. The bottom line is that the Heart Disease Prediction System Using Electrocardiogram (ECG) is an easy to use application, fast and mostly accurate to predict the heart health of a person.

6. FUTURE WORK

In the application that we developed, the user needs to calculate the QRS value and HR value and then feed it in to the system but in future we can design a system where user can upload the image of his/her Electrocardiogram report where the system will directly read the graph and display the result according to it. It will result in reducing the complexity associated with entering the QRS value and HR value and will significantly improve the overall efficiency of the system. Furthermore, in the future, many more suitable features can be added in the application which may be helpful for the patient in maintaining heart health such as the diet recommendation, sleep recommendation, water intake recommendation, dos and don'ts, workout recommendation etc. Another useful feature that may be added in the future is the recommendation of medications based on the heart health of the user which will significantly improve the overall satisfaction of the users that are using the application.

REFERENCES

[1] Latha Parthiban, R.Subramanian, "Intelligent Heart Disease Prediction System using CANFIS and Genetic Algorithm", International Journal of Medical and Health Sciences Vol. 1, No. 5, 2007.

[2] S. Palaniappan and R. Awang, "Intelligent heart disease prediction system using data mining techniques," 2008 IEEE/ACS International Conference on Computer Systems and Applications, Doha, 2008, pp. 108-115.

[3] JyotiSoni, Uzma Ansari, Dipesh Sharma, SunitaSoni, "Intelligent and Effective Heart Disease Prediction System using Weighted Associative Classifiers", International Journal on Computer Science and Engineering (IJCSE) Vol. 3 No. 6 June 2011

[4] Chaitrali S. Dangare, Sulabha S. Apte, PhD., "Improved Study of Heart Disease Prediction System using Data Mining Classification Techniques", International Journal of Computer Applications (0975 – 888) Volume 47– No.10, June 2012.

[5] Negar Ziasabounchi, Iman Askerzade, “ANFIS Based Classification Model For Heart Disease Prediction”, International Journal of Engineering & Computer Science IJECS-IJENS Vol:14 No:02, April 2014.

[6] J. S. Sonawane and D. R. Patil, "Prediction of heart disease using multilayer perceptron neural network," International Conference on Information Communication and Embedded Systems (ICICES2014), Chennai, 2014, pp. 1-6.

[7] M. A. Jabbar and S. Samreen, "Heart disease prediction system based on hidden naïve bayes classifier," 2016 International Conference on Circuits, Controls, Communications and Computing (I4C), Bangalore, 2016, pp. 1-5.

[8] M.A. Jabbar, B.L. Deekshatulu, Priti Chandra, "Intelligent heart disease prediction system using random forest and evolutionary approach", Journal of Network and Innovative Computing. ISSN 2160-2174 Volume 4 (2016) pp. 175-184

[9] Praveen Kumar Reddy M, T Sunil Kumar Reddy, S.Balakrishnan, Syed Muzamil Basha, Ravi Kumar Poluru, "Heart Disease Prediction Using Machine Learning Algorithm", International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-8 Issue-10, August 2019.