International Journal of Advanced Research in Arts, Science, Engineering & Management (IJARASEM)

| ISSN: 2395-7852 | www.ijarasem.com | Bimonthly, Peer Reviewed & Referred Journal|

| Volume 3, Issue 5, September 2016 |

A Novel Approach for Heart Disease Detection and Classification

Triveni Dharmraj Nagpure

Department of Computer Engineering, AISSMS COE, Pune, India

ABSTRACT: Heart disease is a common condition that may be fatal in older people and those who don't lead healthy lifestyles. They may somewhat avoid it with regular diagnostic tests, proper eating habits, and regular checkups. Hospitals generate a large amount of patient data, such as x-rays, lung tests, heart pain tests, chest pain tests, personal health records (PHRs), etc. A decision tree classifier is built using the symptoms, or more specifically, the characteristics required for prediction. We can identify particular traits that are the best and lead to a better prediction of the datasets using the decision tree approach. The utilization of hospital data gathering is ineffective. While some uses of these technologies are forbidden, others are used to retrieve data from the database used to diagnose heart disease. This study determines whether or not individuals have cardiac ailments based on the data in their records by using a variety of optimization techniques machine learning and Deep Learning(mRNN) algorithms, and health care data. To determine if the patient has heart illness, try using the data as a model.

KEYWORDS*:* Monitoring System, Performance evaluation, Machine learning, heart diseases

I. INTRODUCTION

A healthcare system should provide better healthcare facilities in an accessible & patient-friendly way to people all over the globe. Recently, the healthcare system has shifted from an old-fashioned approach to a patient-centered one. In an old-fashioned way, doctors play an essential role. To deliver accurate diagnoses and suggestions, they must visit the patients. The two primary issues with this approach are that it requires the necessary doctors to be on-site the whole time the patient is there and that the patient must be in the hospital for an extended period with linked biomedical equipment wired to the bedside. I came up with a patient-friendly answer to these two issues. Patients are aware of the abilities and information necessary to participate more actively in this area of sickness diagnosis and prevention. This second technique must include a reliable patient monitoring system that is simple to use (PMS). The most challenging problem for individuals worldwide is their health. The World Health Organization states that everyone has a fundamental right to the highest level of fitness (WHO). By being active, one may safeguard their lifetime income. Healthy people may also reduce the strain on networks, governmental or non-governmental organizations, and charities dedicated to public safety, relieving pressure on already overworked hospitals, clinics, and medical professionals. A healthy person needs a cutting-edge, readily accessible healthcare system. Recent technological improvements have had a tremendous impact on the patient monitoring system. We calculate patient parameters (ECG, temperature, heart rate, pulse, blood pressure, and so on) in our system using various sensor data. According to the most recent research, the death rate from heart disease is increasing. An intelligent heart disease prediction system is also necessary to lower mortality. Numerous things, such as changing behaviors and rising stress levels, might contribute to heart disease. Predicting cardiac disease is thus a crucial component of life. According to the literature, many data mining techniques have been used to predict heart illness. The experiment parameters include body temperature, pulse rate, blood pressure, and heart rate. Every day, enormous amounts of medical data are produced, making it challenging to extract any valuable information since the heart is the most vital organ for human life, and excellent health results from its proper operation.

II. LITERATURESURVEY

According to [1] the categorization method explains how to use supervised learning to tackle a specific issue. There are typically certain measures that must be taken. The sort of training example must be decided first. The training set must next be compiled, either by human specialists or by measures. It must also be indicative of how the function is used in the actual world. The input feature's representation must then be determined. The representation should be detailed enough to reliably anticipate the outcome. After that, you must choose a learning algorithm. We execute the learning algorithm on the obtained training data when we finish the design. Some supervised learning techniques are needed at this step from the user to set specific control parameters, particularly for prediction problems.

In AsmaBaccoucheet. al. [2] proposed an ensemble-learning framework based on unidirectional and bidirectional BiLSTM or BiGRU model with a CNN and achieved the accuracy of 91% for different types of heart disease. A data preprocessing with feature selection is implemented to improve the classifier performance.

| e-ISSN: 2319-8753, p-ISSN: 2320-6710| www.ijirset.com | Impact Factor: 8.118|

||Volume 11, Issue 8, August 2022||

| DOI:10.15680/IJIRSET.2022.1108039 |

In Li Yang et. al.[3] several methods were used to build prediction model. Consistent follow-up was managed using electronic health record system. They Provided a three years risk assessment prediction model based on huge population with high risk in eastern China for CVD (Cardio Vascular Disease).

In YounessKhourdifiet. al. [4] researchers improved the heart disease classifier by filtering redundant features using the Fast Correlation-Based Feature Selection (FCBF). Then, they performed a classification based on different classification algorithms.

Researchers employed the Rapid Miner tool and other ML techniques to Fahd SalehAlotaibi [5] to increase the accuracy score and forecast heart disease. We examined the UCI heart disease dataset. The suggested work raised the accuracy rating previously achieved.

Lewlyn L. R. Rodrigues [6] proposed use of partial least squares for structural equation modeling while analyzing data. Using machine learning, they looked at the relationship between body mass index, age, systolic and diastolic blood pressure, the number of cigarettes smoked each day, the amount of alcohol consumed each week, and hypertension or coronary heart disease. They found that all of the characteristics, with the exception of age, SBP, and BMI, were significantly positively associated with coronary heart disease (CHD) and hypertension. These findings helped ML academics and practitioners who are trying to find the correlations between these variables.

Mohd Ashraf et al. Deep Neural Network approach was suggested by researchers al. [7] to develop an automated method for heart attack detection. Multiple datasets were used to assess the accuracy of ML approaches. The suggested solution brought an automatic preprocessing approach to the data and eliminated systemic irregularities.

In Sumit Sharma, Mahesh Parmar [8] researchers proposed Talos Hyper-parameter optimization model for prediction of cardiac and heart disease. Heart disease is a key area to use Deep Neural Network which boost overall heart classification consistency Disease. Classification performed differently using SVM, Naïve Bayes , Random Forest. UCI heart attack Dataset to display the Talos Hyper-parameter Optimization performed better than other classification algorithms mentioned.

In AsmaBaccoucheet. al. [9] proposed an ensemble-learning framework based on unidirectional and bidirectional BiLSTM or BiGRU model with a CNN and achieved the accuracy of 91% for different types of heart disease. A data preprocessing with feature selection is implemented to improve the classifier performance.

Chengjin Yu. et al. [10] developed an adversarial training strategy to estimate multi-type Cardiac Indices in MRI and CT using multitask learning. These task dependencies are shared and taught via multitask learning networks. Finally, they used CT to transfer characteristics learnt from MRI. The scientists conducted a series of studies in which they first improved system efficiency using ten-fold cross-validation on 2900 cardiac MRI images. A different set of 2360 cardiac CT images was used to test the network after that. A number of cardiac indices can be calculated using the suggested reverse mapping, according to all of the studies.

III. PROPOSED METHODOLOGY

In this project we design the system which is useful for the heath care center as well as for hospital. The temperature, blood pressure, and pulse rate sensor values should be used to detect patient body parameters. A variety of efficient hybrid machine learning techniques will be used to create the graphical user interface.Training and testing are the two basic elements of the suggested system. This study suggests using deep learning algorithms to model diseases effectively. The execution as a whole is crucial for achieving the classification accuracy, and the dataset is crucial.

| e-ISSN: 2319-8753, p-ISSN: 2320-6710| www.ijirset.com | Impact Factor: 8.118|
||Volume 11, Issue 8, August 2022||
| DOI:10.15680/IJIRSET.2022.1108039 |

A. Architecture

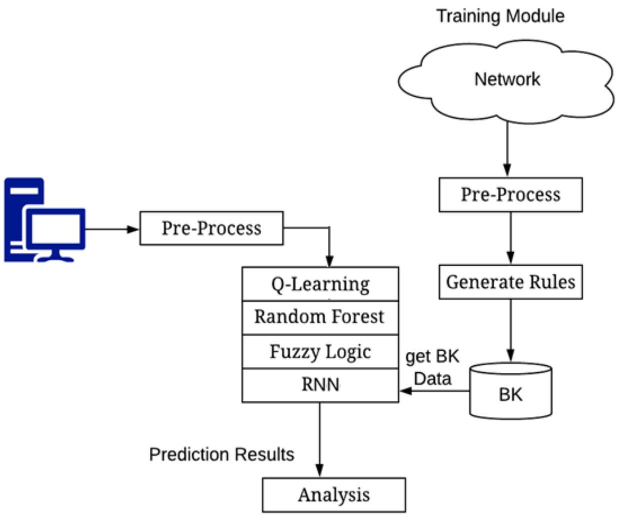

Figure1.Proposed System Architecture

The proposed system is collaboration of various modules such IoT modules, database modules and GUI modules. In below section we described each module in detail

1) Training:

- Collect data.
- Apply data mining approaches.
- Data is been saved into the database called as background knowledge, which is used at the time of testing.

2) Testing:

- The system uses artificial and real-time patient data input over the internet to predict the likelihood of a disease based on training modules.
- All of the acquired data is kept in a global database using a link-oriented architecture.
- All training and testing data are read concurrently during testing.
- Apply the machine's classification and anticipate how the decision-making process will be used in the future.
- Lastly, show the study's consistency with the system's real (positive) and misleading results (negative)

B. Algorithms

Algorithm 1: Proposed modified Recurrent Neural Network Algorithm (mRNN)

Input: Train_Feature set [], // Set of training dataset
Test_Feature set [] //Set of test dataset
Threshold denominator Th
Collection List cL

Output: Generate class label for all test instances based on classification results.

Step 1: Utilize the following function to read all properties from the Testing dataset.

$$\text{Test\_Feature} = \sum_{j=1}^{n} (\mathrm{T}[\mathrm{j}])$$

Step 2: Utilize the following function to read all attributes from the training dataset.

$$\text{Train\_Feature} = \sum_{k=1}^{m} (\mathrm{T}[\mathrm{k}])$$

Step 3: Use the function below to read the total characteristics from the train instances.

Step 4: Create weights for both feature sets and calculate the similarity index.

| e-ISSN: 2319-8753, p-ISSN: 2320-6710| www.ijirset.com | Impact Factor: 8.118|

||Volume 11, Issue 8, August 2022||

| DOI:10.15680/IJIRSET.2022.1108039 |

$$Weight = classifyInstance(\text{Train\_Feature}, \text{Test\_Feature})$$

Step 5: Verify with Th

optimized_Instance_result = Weight >Th ?1 : 0;

Add each optimized_instance to cL, when instances = null

Step 6: Return cL

Machine Learning Algorithm

Q- Learning Algorithm

Input: Threshold group TMin[1...n] and TMax[1...n] for all sensors, Desired Threshold Threshold group inp[1.....n] all input parameters produced by sensors.

Output: Trigger executed for output device as lable.

Step 1 : Read every data file (R into DB)

Step 2: Parts [] ← Split(R)

Step 3:

$$\text{CVal} = \sum_{k=0}^{n} \textbf{Parts[k]}$$

Step 4: check (Cval and Respective TMin[1...n] and TMax[1...n] Thresholds)

Step 5: if (Cval> Threshold)

Read all measure of for penalty TP and reward FN

Else continue. Tot++

Step 6: calculate penalty score = (TP *100 / Tot)

Step 7: if (score >= Th)

Generate event

End for

Fuzzy Logic Algorithm

Input: User-inputted data record with patient id, timestamp, and values for all body parameters and sensor parameters.

Output: Classified label

Step 1: Read all of R's properties from the current parameters.

Step 2: Detailed train map with samples for each feature.

Step 3 :calculate average weight of train DB with same evidences

$$Avg\_Score = \sum_{k=0}^{n} (Sc)$$

Step 4 :evaluate Avg_Score>th

Step 5: ReturnAvg_Score

Random Forest

Input : Selected data feature of all test instances D[i….n], Training database policies {T[1]…………T[n]}

Output : No. of probable classified trees with weight and label.

| e-ISSN: 2319-8753, p-ISSN: 2320-6710| www.ijirset.com | Impact Factor: 8.118|

||Volume 11, Issue 8, August 2022||

| DOI:10.15680/IJIRSET.2022.1108039 |

Step 1: Read (D into D[i])

V ←Extract features (D)

Step 2: N ←CountFeatures(D)

Step 3: for each(c into TrainDB)

Step 4: Nc[i] ←ExtFeatures(c)

Step 5: select relevant features of w= {Nc[i], N}

Step 6: Statement (w>t)

Step 9: Return Tree Insatnce{ Nc[i], N, w, label}

IV. RESULT AND DISCUSSIONS

The relationship between clinical ratings and quantitative assessments makes sense and has the ability to solve a variety of decision-making issues. Six machine learning models were used with the created training repository to identify patterns of frequent, dubious, and risky behaviors. Utilizing the behavior categorization, training-database, the 3-fold, 5-fold, & 10-fold cross-validation model was used to test and evaluate the machine learning methodologies. Figure 2 below illustrates the 3-fold classification method applied to all parameters and demonstrates the consistency of all implementations.

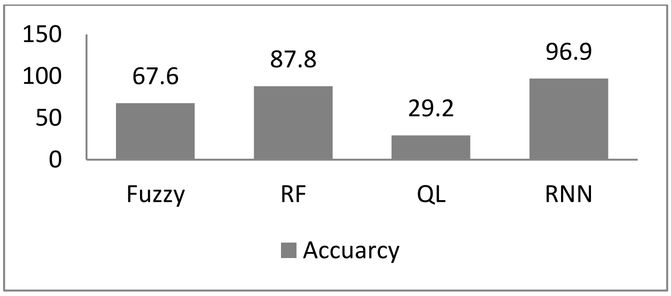

Figure 2: Accuracy of various ML and RNN classification

The overall accuracy of each method, including the suggested RNN, is depicted in Figure 2. It has a 96.9 % accuracy rate. Random Forest's minimum accuracy is 87.8%, which is higher than that of other methods..

The above figure 2, with its seven separate algorithms—Q-Learning, RF, Fuzzy, and Recommended Perceptron Algorithm—improves the significance of diverse experimental study focused on various statistical tests. The classifier employed the mRNN classification method for data management. Each model's neural network was presented and discussed. The system accuracy for correctly identifying, incorrectly classifying, recession, & device recall is displayed by both uncertainty metrics.

V. CONCLUSIONS

Because it offers the typical person in many industrial locations a basic network they can afford, the Internet of Things Design is a well-functioning technology. Healthcare research is an essential and inescapable component of our daily life in the area mentioned above. A more reliable platform for compiling sensory data in the medical field & integrating it into devices is provided by the Internet of Things (IoT).Super brilliance offers the best supervision for the poor. This is the most fundamental degree of intellectual thought in android devices, which can also be referred to as smart gadgets. In the traditional approach, most tests are invasive, causing patient's anxiety and causing frustration or carelessness in their health. It's tough for them to deal with those circumstances. In this scenario, patients can use internet technology to contact the doctor 24 hours a day, seven days a week, and be informed.

| e-ISSN: 2319-8753, p-ISSN: 2320-6710| www.ijirset.com | Impact Factor: 8.118|

||Volume 11, Issue 8, August 2022||

| DOI:10.15680/IJIRSET.2022.1108039 |

REFERENCES

[1] PROTOPAPADAKIS, Eftychios, et al. Stacked autoencoders driven by semi-supervised learning for building extraction from near infrared remote sensing imagery. Remote Sensing, 2021, 13.3: 371.

[2] AsmaBaccouche, Begonya Garcia-Zapirain ,Cristian Castillo Olea and Adel Elmaghraby, "Ensemble Deep Learning Models for Heart Disease Classification: A Case Study from Mexico", Information 2020, 11, 207; doi:10.3390/info11040207.

[3] Li Yang, Haibin Wu, Xiaoqing Jin, and Pinpin Zheng4 et al., "Study of cardiovascular disease prediction model based on random forest in eastern China ",Scientific Reports (2020) 10:5245 https://doi.org/10.1038/s41598-020-62133-5.

[4] YounessKhourdifi and Mohamed Bahajm," Heart Disease Prediction and Classification Using Machine Learning Algorithms Optimized by Particle Swarm Optimization and Ant Colony Optimization", International Journal of Intelligent Engineering and Systems, Vol.12, No.1, 2019 DOI: 10.22266/ijies2019.0228.24.

[5] Fahd SalehAlotaibi," Implementation of Machine Learning Model to Predict Heart Failure Disease", (IJACSA) International Journal of Advanced Computer Science and Applications,Vol. 10, No. 6, 2019.

[6] Lewlyn L. R. Rodrigues, Dasharathraj K Shetty, NitheshNaik , ChetanaBalakrishnaMaddodi and et al., "Machine learning in coronary heart disease prediction: Structural equation modelling approach ", Rodrigues et al., Cogent Engineering (2020), 7: 1723198.

[7] Mohd Ashraf, M. A. Rizvi and Himanshu Sharma, "Improved Heart Disease Prediction Using Deep Neural Network", Asian Journal of Computer Science and Technology , ISSN: 2249-0701 Vol.8 No.2, 2019, pp. 49-54 © The Research Publication.

[8] Sumit Sharma, Mahesh Parmar , "Heart Diseases Prediction using Deep Learning Neural Network Model ", International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-9 Issue-3, January 2020.

[9] AsmaBaccouche, Begonya Garcia-Zapirain ,Cristian Castillo Olea and Adel Elmaghraby, "Ensemble Deep Learning Models for Heart Disease Classification: A Case Study from Mexico", Information 2020, 11, 207; doi:10.3390/info11040207.

[10] C. Yu et al., "Multitask Learning for Estimating Multitype Cardiac Indices in MRI and CT Based on Adversarial Reverse Mapping," in IEEE Transactions on Neural Networks and Learning Systems, doi: 10.1109/TNNLS.2020.2984955