International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Early Prediction of Cardiac Arrythmia

Poornima V[1], Santhrupth Rao[2], Shannmukannanda Redie[3], Venkat Vinay[4], Mrs Deepthi. S[5]

[1,2,3,4]Department of Electronics and Computer Engineering with Specialization in It Infrastructure, Presidency University, Bangalore, India

[5]Professor, Presidency University, Bangalore, India

ABSTRACT:

Heart arrhythmia is a potentially fatal condition that leaves patients with serious health issues. It will be helpful to diagnose arrhythmia problems early on in order to save lives. The Internet of Things (IoT) promises to upgrade the medical field by continuously, remotely, and non-invasively monitoring conditions related to cardiac arrhythmias. An Internet of Things platform for cardiovascular illness prediction employing an IoT-enabled ECG telemetry system obtains and analyses the ECG signal, notifying a clinician of an emergency. An accurate and timely analysis of cardiac disease is beneficial to the doctor. To analyse the ECG signal, we are creating an Internet of Things-enabled ECG monitoring system. The raw ECG signal's statistical characteristics are computed. The Pan Tompkins QRS detection technique is used to analyse the ECG signal in order to extract its dynamic properties. In order to obtain aspects of heart rate variability, the system is utilised to determine the RR intervals from the ECG data. Subsequently, the cardiac arrhythmia condition is classified through the application of statistical and dynamic features. Even at home, people can obtain an ECG signal to assess their cardiac health. Because of its compact size, the system requires fewer upkeep and operating expenses. The doctor can more simply and precisely analyse cardiac problems with this aid.

INTRODUCTION

Cardiac arrhythmia refers to a variety of heart rhythm disorders in which the heartbeat is irregular, rapid, or sluggish. Arrhythmias come in a variety of forms, some of which possess no symptoms Palpitations or a feeling of a halt between heartbeats may be experienced when symptoms are present noticeable. In more extreme instances, lightheadedness, fainting, Breathing, difficulties or pain in the chest might appear. While the majority of arrhythmias are unharmful, some can have catastrophic side effects including cardiac failure or stroke. Some of them might cause cardiac arrest. Arrhythmia affects millions of individuals throughout the world. Cardiovascular disease causes around 15% of all fatalities worldwide, or close to half of all deaths, are caused by sudden cardiac death. Ventricular arrhythmias account for approximately 80% of sudden cardiac death. Arrhythmias can affect people of any age, although they are more frequent as they get older.

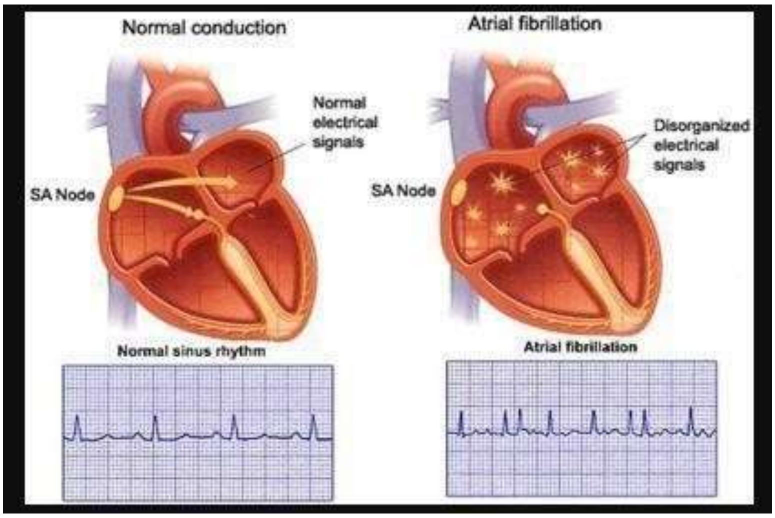

Fig.1.1 Rhythm of Heartbeat

Arrhythmia affects millions of individuals throughout the world. Cardiovascular disease causes around 15% of all fatalities worldwide, or close to half of all deaths, are caused by sudden cardiac death. An interruption creates arrhythmia in the electrical impulses that cause the heart to contract. A healthy heart should beat between 60 and 100 times per minute while the individual is at rest. The lowest potential resting heart rate is desirable.

Because of their highly efficient hearts, Olympic competitors, for instance, be under 60 beats per minute when at rest.

The heart can malfunction due to a variety of circumstances, including:

- Abuse of alcohol
- Abuse of diabetes medications
- coffee drinking that is excessive
- Congestive heart failure is a type of heart disease.
- hypertension is a condition in which the blood pressure
- Thyroid hyperthyroidism

Your heartbeat is pleasant and regular and has just the proper rate when it's normal. A cardiac arrhythmia (abnormal heart rhythm) is one of the most frequent heart illnesses. It occurs when your heartbeat is too rapid, too slow, or beats in an irregular rhythm. In fact, most people experience heart arrhythmias on a regular basis.

Cardiac arrhythmias can range in severity from mild to severe. Most arrhythmias are harmless and insignificant, but some are exceedingly severe and life-threatening. Many of these cause symptoms that might be quite bothersome in your daily life, even if they aren't particularly harmful. The Sino-atrial Node, sometimes referred to as the Sinus Node, also known as the SA node, is a little area of cardiac tissue that generates the electrical impulses that drive the heart to beat in the right atrium. Before activating the atrioventricular (or AV) node, which links the atria and ventricles, the impulse causes both atria to contract (Principal pump chambers). The Bundle of His and Purkinje then conveyed the urge to the opposite edges of the room fibers connect the ventricles, generating a synchronized heart muscle contractions and, as a result, a pulse. The normal resting heart rate of an adult human is 60 to 90 beats per minute. The resting heart rates of children are much greater. On the other hand, athletes the average resting heart rate is 40 beats per minute. while still being deemed normal.

LITERATURE REVIEW

The investigation titled "Early Prediction of Cardiac Arrythmia Using Machine Learning Models" yielded encouraging results by effectively employing machine learning techniques to classify diverse cardiac arrhythmia types with notable accuracy [1] Despite major advancements both in terms of diagnosis and treatment. Improved CVD outcomes require early and precise diagnosis. Cardiovascular imaging is critical in making diagnostic decisions. The majority of current image analysis approaches rely on qualitative picture evaluation and basic quantitative assessments of heart shape and function. In order to maximize cardiac imaging's diagnostic usefulness 5, More improved image analysis techniques are needed to allow for deeper measurement of imaging traits. For a number of essential illnesses including cardiac failure with coronary artery disease (CAD), existing research has already demonstrated the added usefulness of image-based cardiovascular diagnosis using ML (HF). The increased allowing quicker and more precise diagnostic decision-making will lessen the burden of cardiovascular disease. ECG testing, any irregularity in the electrical circuit of the heart may result in cardiac arrhythmia. Symptoms include chest discomfort, shortness of breath, disorientation, and fainting. A quick or slow pulse induced by an abnormal heart rhythm or irregular heartbeat can cause Shortness of breath, fainting, dizziness, and chest discomfort.

DRAWBACKS

The While compared to previous work, these algorithms have the best accuracy rate; nonetheless, the computation time is the most major constraint faced due to the learning necessary when using datasets. Using digitized A convolutional neural network (CNN) was developed and validated using 12-lead ECG data from 2,448 patients with a confirmed diagnosis of HCM and 51,153 non-HCM control individuals who were matched for age and sex. a different dataset with 12,788 controls and 612 HCM persons was used to assess the CNN's capacity to identify HCM. The HCM group's mean age in the pooled datasets was 54.8 15.9 years, compared for the control group, from 57.5 to 15.5 years. After training and validation, the CNN's area under the curve (AUC) in the validation dataset was 0.95 at the ideal probability threshold of having HCM (95 percent confidence interval [CI]: 0.94 to 0.97). In a subgroup study, An AUC of 0.95 (95 percent confidence interval [CI]: 0.94 to 0.97] was found in people whose ECGs showed left ventricular hypertrophy; those with a normal ECG had an AUC of 0.95 (CI]: 0.90 to 1.00). The model functioned superbly with individuals who were younger (sensitivity 95 percent, specificity 92 percent). in both sarcomere-positive and -negative individuals the model-derived median probability for HCM was 97% and 96%, respectively, for mutations.

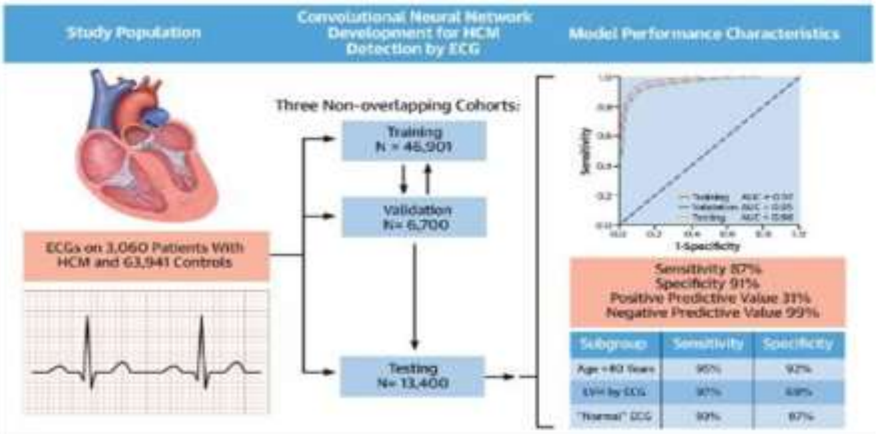

Fig.3.1 CNN Development for HCM Detection by ECG

SYSTEM DESIGN

The Data gathering is the first step in processing the system, and we utilise the UCI repository dataset for this. This dataset has undergone comprehensive verification by a number of researchers and the UCI authority.

In light of the preceding graphic (Figure 4.1), it is explained as follows: Information Extraction: An organised and mathematical dataset was taken from the ECG. Considerations are made for the heart rate, R-R distance, number of deflections, height, gender, and other variables. The UCI Machine Learning Repository provided the dataset used in this work. A csv file is then created and used to hold the data.

Preprocessing:

Due to missing values and inconsistent data, the dataset cannot be handled in the classification process. More variables were deleted since they were the same for each subject.

Invariant qualities are evaluated using the variance or standard deviation value. Average values are used to fill in the remaining missing data.

Feature Extraction:

There are two methods for selecting features: Random Forest and Principal Component Analysis (PCA). The preprocessed data includes a lot of characteristics, and the categorization approach we chose requires a lot of effort. Feature selection is essential to save time.

Classifcation:

The third stage is called the classification stage. So the classification plays an important role in the machine learning model. In the classification stage we have included the five algorithms says: KNN, SVM, Naïve Bayes, Logistic Regression and Random forest. In the previous stage by using feature reduction we have reduced the datasets, so that dataset is stored in a csv file and then using that data the accuracy precision recall F1 score has been calculated.

Evaluation:
The correctness of each method is evaluated and shown, and the selected qualities are then utilised as input for the five classifications that come next.

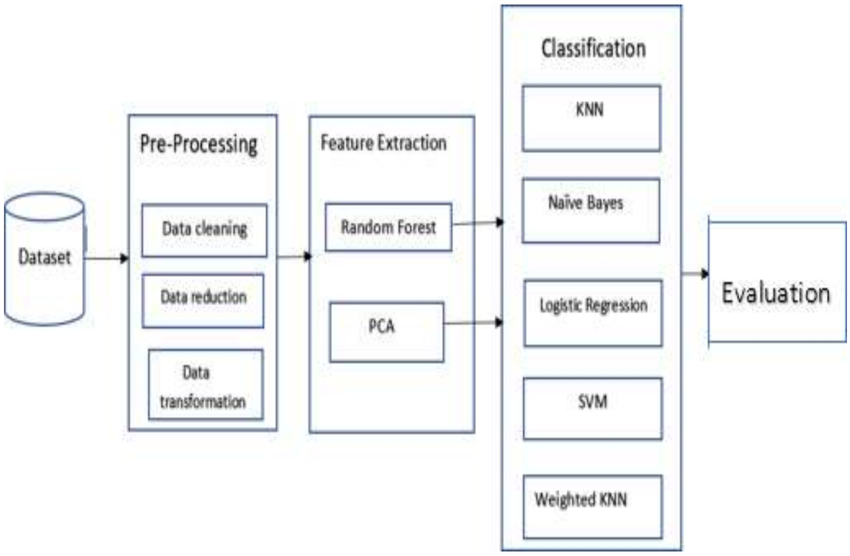

Fig.4.1 General block diagram of system architecture

FLOW DIAGRAM

Data "flow" across an information system is graphically represented by a data flow diagram (DFD), which illustrates various process components. In order to offer a general overview of the system without going into considerable depth, a DFD is frequently utilised as the first phase. Additionally, data processing may be visualised using DFDs (structured design).

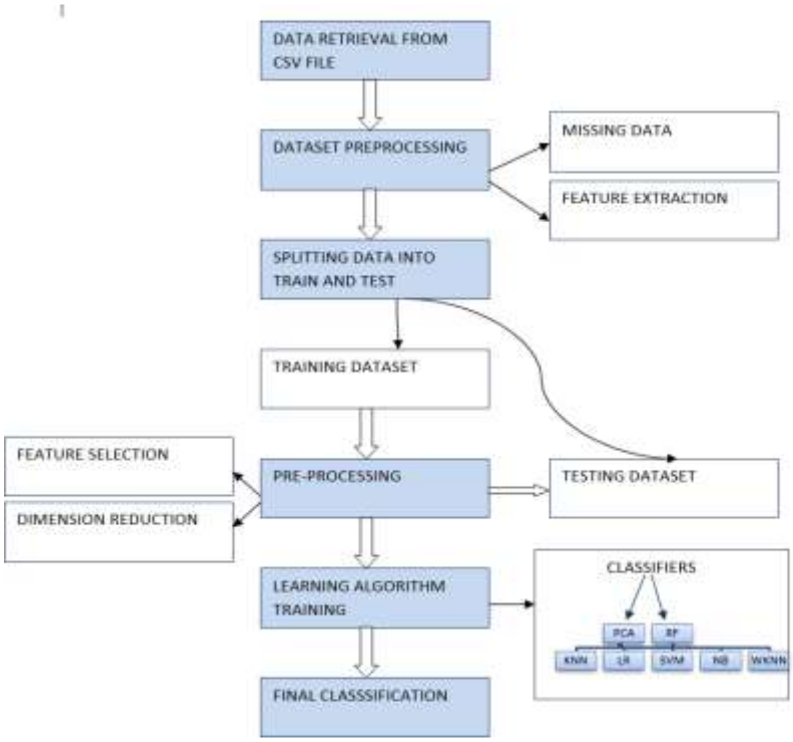

Fig.4.2 Data Flow Diagram

A data flow diagram (DFD) is shown in Figure 6.1 and it indicates the different types of data that data input into the system, data output from the system, data flow through the system, and data storage. It does not include information on process time or whether activities will execute sequentially or simultaneously, in contrast to a normal structured flowchart, which emphasises control flow, and a UML activity workflow diagram, which illustrates both control and data flows as a unified model.

First from csv file the data is taken, and then the dataset is processed. Filtering the measured signals is part of data processing since real ECG data is noisy and tainted with artefacts such as electrocardiography signals caused by breathing and chest movement. It entails missing data, categorical data encoding, and partitioning the data into train and test data.

The model is then trained, features are chosen, and dimensions are reduced. To categorise data, learning techniques are utilised. Dimension reduction is accomplished through the use of Principal Component Analysis (PCA) and Random Forest. KNN, SVM, and NB are utilised to categorise arrhythmia into 13 separate classifications.

IMPLEMENTATION

One of the most effective testing approaches is Machine Learning, which is focused on training and testing. Artificial intelligence (AI) is a vast field of research where robots are made to mimic human skills. Machine learning is a subset of AI. Machine intelligence refers to the mixture of the two technologies. Machine learning systems, on the other hand, are taught how to analyse and utilise data.

In this research, we analyse four algorithms—decision tree, linear regression, k-Neighbour, SVM, and Weighted-KNN—in terms of accuracy utilising biological factors as testing data, such as cholesterol, blood pressure, sex, age, and so on. This is because machine learning, by definition, learns from natural phenomena and things.

METHODOLOGY

Fig.5.1 Methodology

Data Pre-processing:

- Pre-processing is the term for the adjustments we make to our data before sending it to the algorithm, as seen in figure 5. 1. Data Preprocessing is a method for transforming messy data into a tidy collection.
- To put it another way, when data is collected from several sources, it is done so in a raw form that prevents analysis.
- Performing a NaN check
- Checking for NaN is critical during data pre-processing. We were only able to find a few NaNs in this try.
- Changing the value of NaN
- It's critical to get rid of the NaN values. This may be accomplished by:
- removing the whole column having a large number of NaN values
- Method of forward fillna
- Method of backward fillna
- Using the mean technique

Data analysis:

Data analysis is the process of dissecting, sanitising, modifying, and modelling data with the aim of revealing relevant information, guiding deductions, and assisting in decision- making.

Data analysis has many different components and steps, including a wide variety of methods with different names that are applied in a number of business, scientific, and social science fields. Because it helps businesses to operate more efficiently and make more scientific judgments, data analysis is essential in today's business environment.

Feature Extraction:

Feature extraction is the process of converting raw data into numerical traits that may be used while keeping the specifics of the original data set. Compared to just applying machine learning to raw data, it produces superior outcomes.

As a consequence, when training a dataset, it is possible to quantify how much each feature lowers impurity. The greater an attribute's ability to eliminate impurity, the more significant it is. In random forests, the impurity decrease from each feature may be averaged across datasets to determine the variable's final significance.

Train and Test Dataset:

It's time to fit the first machine learning model into your data once you've cleaned it up, visualised it, and learnt more about it. Creating two sets of data: one for training and one for testing.

- Training Dataset: A portion of the data was used to fit the model.
- The test dataset is used to objectively assess the final model's fit to the training dataset.

Prediction and Accuracy:

Stated machine learning algorithms are taught to forecast the customer's smart phone decision. The ability to forecast the customer's choice of smart phone is critical in helping smart phone makers improve their standards by observing what characteristics are important to customers when choosing a smart phone. Simply put, accuracy refers to how well your machine learning model predicts the proper class for a given observation.

RESULTS AND DISCUSSION

The heart disease prediction system's adoption has produced a number of noteworthy results. The most important of these is the development of an accurate prediction model, which is accomplished by applying several machine learning algorithms, such as Decision Tree Classifier, XGB Classifier, KNN, SVC, Naïve Bayes, Random Forest Classifier, and Logistic Regression. This guarantees a thorough and solid model that can produce accurate forecasts for cardiac disease.

Additionally, the system boasts a user-friendly interface, thanks to the incorporation of HTML. This enables users to seamlessly interact with the system, providing datasets and visualizing results with ease. The interface enhances the overall user experience, making the system accessible and intuitive.

Moreover, the implementation leverages Python's capabilities for efficient data processing. With its high-level built-in data structures and dynamic semantics, Python proves to be instrumental in handling and processing the datasets effectively. This efficiency contributes to the system's overall performance and reliability.

COCLUSION

In conclusion, our heart disease prediction system harnesses the power of various machine learning algorithms, achieving an impressive 91.80% accuracy, with Random Forest standing out as a key contributor. The multi-algorithmic approach provides a comprehensive analysis of user-entered parameters. The user-friendly front end, developed with Flask, HTML, and python, mysql, facilitates easy interaction, empowering users to proactively engage with their health data. This system holds immense potential for revolutionizing early diagnosis and intervention in heart disease cases, exemplifying the symbiosis of healthcare and machine learning. Ongoing refinement and exploration of additional features promise to further enhance the system's predictive capabilities, shaping the future of preventive healthcare practices globally

REFERENCES

[1] P. Kola Sujatha and K. Mahalakshmi. Performance evaluation of supervised machine learning algorithms in prediction of heart disease. 2020 IEEE International Conference for Innovation in Technology (INOCON), pages 1– 7, 2020.

[2] Pahulpreet Singh Kohli and Shriya Arora. Application of machine learning in disease prediction. 2018 4th International Conference on Computing Communication and Automation (ICCCA), pages 1–4, 2018.

[3] Abderrahmane Ed-Daoudy and Khalil Maalmi. Real-time machine learning for early detection of heart disease using big data approach. pages 1–5, 04 2019.

[4] A. Lakshmanarao, A. Srisaila, and Srinivasa Tummala. Heart disease prediction using feature selection and ensemble learning techniques. pages 994– 998, 02 2021.

[5] Alperen Erdoğan and Selda Guney. Heart disease prediction by using machine learning algorithms. 2020 28th Signal Processing and Communications Applications Conference (SIU), pages 1–4, 2020.

[6] Shaik Farzana and Duggineni Veeraiah. Dynamic heart disease prediction using multi-machine learning techniques. 2020 5th International Conference on Computing, Communication and Security (ICCCS), pages 1–5, 2020.