# Machine Learning Techniques Applied for Exploring Heart Disease and Classifying Stages through ECG Signal

**K. Babu[1*], Dr.P. Marikkannu[2]**

## Abstract

Electrocardiogram (ECG) has given crucial data concerning various heart attack criterion of the human heart this examination can be the key goal of the research for the detection and prevention of cardiac circumstances frighten. The proposed method using machine learning techniques for classifying and analyzing ECG signal processing and this research mainly developed for early discovery of heart diseases and also the prediction of stages level. The dataset was utilized as a person ECG signal of Heart Database which was taken from the UCI repository of Machine learning dataset vault. In this paper, a simple algorithm is presented for to discover R-peaks automatically from a single lead digital ECG data. The proposed method detecting the time interval of the ECG signal from the R-peaks level next level with the double squared difference signal is used to localize the region of QRS which is the time interval between the binary data. this method consists of different stages of sorting from the raw data for reducing nosier signal, threshold a difference signal of ECG by analyzing the time interval of QRS, and finally a comparison of relative magnitude to detect the region of interval processing to analyze accuracy result. The proposed research novel machine learning techniques of the multi-module neural network system (MMNNS) is used to analyze the imbalance problem form the ECG signal classification if the wave was abnormal then the user of dataset patients will be affected by heart diseases. Using the time interval varies from the range of QRS then it analyzes the abnormal of ECG by MMNNS algorithm to define the classification result and finally analyze the stages. If the patients were presented an abnormal ECG signal compared with a normal ECG signal wave graph then they were affected by heart diseases if whether they patients were affected then finally classify the stages of heart diseases whether they are predictable/unpredictable stages. The examination result is completed among two strategies on the better accuracy premise and effective training time of the process.



**Key Words:** Machine Learning, ECG Signal, MMNNS Algorithm, Interval Processing, Detecting QRS Region, Stages Classification, Heart Disease Prediction.

**DOI Number: 10.14704/nq.2022.20.5.NQ22197** **NeuroQuantology 2022; 20(5):473-482**

## Introduction

Heart diseases can also be known as Cardiovascular Diseases (CVDs) that mainly caused a high level of mortality rate. CVDs can rise with a blood deficiency at the coronary artery, in which blood is supplied all through the heart vice versa. Heart disease results produce irregular beats that are called arrhythmia and according to the cruelty condition of arrhythmia can occur sudden death. ECG illustrated a human heart's electrical motion and morphologies of signal of ECG have given better information concerning numerous kinds of arrhythmia according to the various cardiac condition.

**Corresponding author:** K. Babu
**Address:** [1*]Research Scholar, Department of Information Technology, Anna University, Regional Campus, Coimbatore, Tamil Nadu, India; [2]Assistant Professor, Department of Information Technology, Anna University, Regional Campus, Coimbatore, Tamil Nadu, India.
E-mail: [1*]babu123phd@gmail.com; [2]pmarikkannu@gmail.com
**Relevant conflicts of interest/financial disclosures:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.
**Received:** 10 April 2022 **Accepted:** 10 May 2022

By using the ECG wave graph, precise and quick heart attack recognition will save numerous lives and provisions of health care throughout the world. This can be optimally optimized to execute a complete evaluation of ECG examination and proposed in the appearance of stages-based classification model for the clarification of run of each stage of ECG signal examination. With the massive efficient heart, the signal of ECG analyzed and offered a public economy and a healthy way.

The pre-processing step is used to neglect noisy data and also reduce the size by removing unwanted data of the heart disease dataset. The time-frequency will define the range of time interval between each peak up and down of the signal. The proposed model enormously described the traditional time/frequency-domain of ECG signal with the advanced machine learning method used to detect heart diseases and also the stages by classifying real-time dataset of person ECG.

## Literature Survey

Heart attack can kill a person in 3 attempts but this literature survey demonstrated the dangers of the first attempt of a heart attack and this can be since of their age, a reduced amount of physical activity, diet, and various other factors. The proposed survey [1] has demonstrated to interface a microcontroller sensor and permitted for checking heart rate readings and that will be transmitted over the Internet. This system using two circuits: the first one could be the circuit broadcasting for the patient and the second circuit has been utilized as receiver circuit which was controlled with the help of doctors or nurses. The limit was set by the sensor users from high-level as well as low-level of heartbeat then the system has began observing and signals for lesser heartbeats. To discover the current heartbeat level system makes consumption of a heartbeat sensor and displays the results on an LCD screen. The heart disease through ECC classification technique by Machine Learning (ML) depending on the numerous features of ECG was investigated in [2]. The technique has been executed by means of ML-libs and Apache Spark has been utilized throughout the ML library. The important challenges in ECG categorization were for the abnormalities discovery in the ECG signals which might be very significant to distinguish the position of a patient. The literature results show that the approach reached a random forest 97.98% using binary classification and the overall accuracy of the GDB Tree algorithm gives 96.75% accuracy. Random Forest has taken 98.03% accurateness with multi-class classification.

Heart disease is necessary to have a better framework that can effectively recognize the frequency of heart disease in a huge amount of random set data instantly. The probable of nine classification procedures were estimated in [3] for predicting heart disease such as neural network (NN), Support Vector Machine (SVM), naive Bayesian (NB), Artificial Neural Network (ANN), decision tree (DT), and KNN (K-Nearest Neighbour). Consequently, an algorithm of SVM and the A priori algorithm have been presented for predicting heart disease. By employing the medical profiles such as, sex, age, blood pressure, fasting blood sugar, and chest pain type according to this, type a medical society has taken a part concern in perceiving and avoiding heart problems. A framework was proposed in [4], in which the pre-processing dataset, the gap Vague set, Fuzzy involvement, and rule mining Fuzzy Correlation were combined for the better decision construction process. Along with the doctors, the decision-making problem is solved by using an Interval vague set regarding heart issues amongst the patient who can be in a state of uncertain. Based on the selected criteria using these mining of rule process, the generated rules were applied to evaluate heart diseases. Since the heart diseases affecting a foremost difficulty and they became the major caused worldwide of death as of these diseases hard for discovering according to the symptoms for disease diagnosis. By exploiting various data mining methods [5] by evidently achieved an analysis of medicine. This process is achieved by detecting the health condition of the heart by formulating questions that are given related to heart diseases. Gathering evidence regarding and heart diseases by utilizing machine techniques for identifying diseases and analyzing proper medicines. By evidently if the disease is not cured by medicine analysis then the user should analyze the best drugs for advising each patient.



## Proposed Methodology

### *A) ECG Data Source*

The dataset of person heart ECG was predicting a heart disease which was taken from the repository of UCI of ML. UCI database collections are used to implement advanced machine learning algorithms. A real-time dataset of ECG Signals was taken by the parameter of binary values. Binary data of 0 and 1

were presented in the programming language where ECG was presented by the interval of time. In the examination, classification system has contributed high efficiency and achieves high percentage of accurateness than normal ECG methods. Four models are used to design in the research: acquisition of data, data preprocessing, feature extraction, and stages categorization002E.

### *B) Acquisition of Data & Transformation*

The ECG Signal dataset analysis for classification and feature extraction according to the various cardiac conditions, the dataset must be original raw data that motivates the heart disease prediction. The attribute of binary data is recorded from the ECG signal that helps to decide a feature extraction to explore a process. This phase covers various sources of ECG acquisition as an input to the heart disease prediction and stages classification with a unique prominence on the source of data. The dataset will be raw data that acquisition and transformation and processing a supported high-level programming language for changing a high level should follow the below factors.

**Table 1.** Overview of heart illness factors

| **Types of Cardiovascular Disease** | **Heart Attack Symptoms** | **Heart Disease Risk Factor** |
|---|---|---|
| • Coronary artery disease<br>• Cardiac arrest<br>• Clogged heart stoppage<br>• Stroke, and more | • Shortness of breath<br>• Pain and discomfort in the chest area<br>• Neck, back, jaw, or stomach Fatigue<br>• Pain might be extended to the left or right arm<br>• Heartburn or abnormal pain<br>• Cold sweat and unsteadiness<br>• Rapid heartbeat or irregular heartbeat | • Smoking<br>• High blood pressure<br>• More fat<br>• Ancestors history of coronary sickness<br>• High Cholesterol<br>• Diabetics<br>• Consuming too much alcohol |

The ECG could be recorded as an electrical motion which was taking place from the cycle of cardiac in the heart part. It will be confined on a wave graph illustrated in figure 1 (2 signal cycles of ECG were displayed in this figure 1). The ECG signal's electrical motion can be in the form of little probable made through tissues of heart that are picked up throughout the of the ECG's electrodes.

### *i) Data Preprocessing*

Signal of ECG recording can be contaminated frequently through various kinds of artifacts and noises that will affect experiment of subsequent and classification of final results. It is significant to the implementation of a reasonable process to pre-processing the original raw data ECG signal without losing any useful information. Even though the ECG signals were obtained from a public dataset that does not contain as much noise directly from the patients, only there are some noises where presented from the spectrum of signal overlapped with useful information. The user should utilize a practical method to preprocessing the original ECG signal without losing beneficial information from the dataset.



### *ii) Denoising ECG Signal*

ECG signal analyzing and classification required real-time or prerecorded signals of ECG since main input data. ECG signal data acquisition in both cases has been accomplished with the help of sensors that lead to the individual human. The noise could also been captured along with the inventive signal in the process ECG signal acquisition since the noise can affect the signal value and ECG Signal categorization.

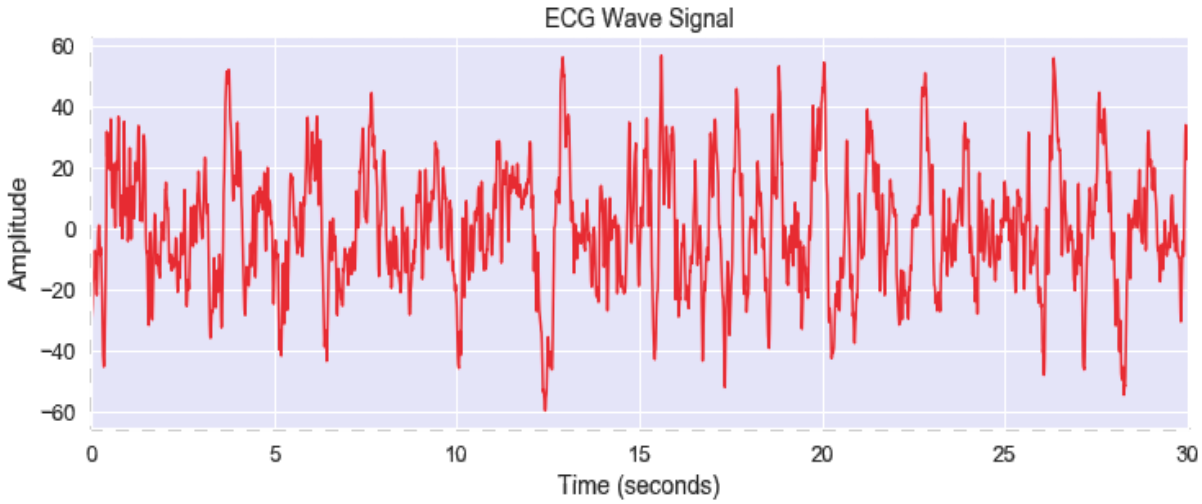


**Figure 1.** ECG Signal Processing

Removal of noise from the ECG signal is called denoising; accurately identifying different anomalies it has been a topmost concern of researchers to remove noise data from the ECG signal. Neural network methods denoise the ECG signals that include applying a band-pass filter (0.05-45Hz) with sample entropy to verify the data quality. Noise can sometimes cause false alarms that are critical to measuring health status. In ECG analysis, noise is usually removed after acquiring data from data sources.

### *C) Interval Processing of QRS Region*

The various ECG signal waves have been termed in alphabetic order is known as P, QRS, and T-U waves. Amplitude, shape, and period can provide significant data concerning wellbeing and the status of heart problems. The wave of P represented that reproduces of atrial depolarization. Then reflects of ventricular depolarization is defined as QRS complex. By employing the T-U wave, depolarization of ventricles will be reflected. For an ECG lead, the electrocardiograph can observe a positive wave while a current of depolarization extends toward the relevant lead's positive pole. On the other hand, a negative wave is appeared in the case while the current transmits away from the pole.

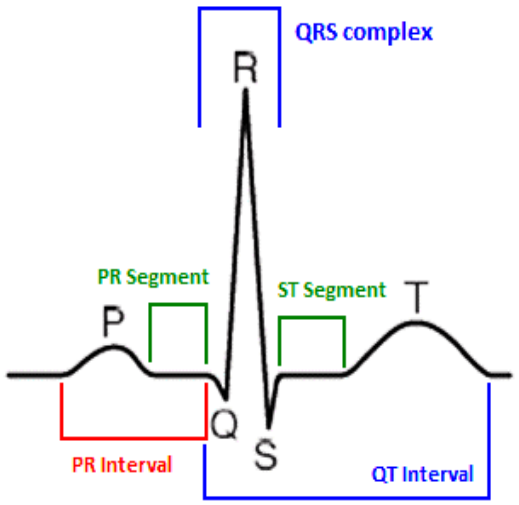


**Figure 2.** QRS time interval classifying ECG Curve

$$I_{(x,y)} = \int_{+\infty}^{-\infty} f(t)\psi_{(x,y)}^{*}(t)dt \quad (1)$$

Where, I denote a time interval and (x,y) between the poles of time taken for signal the 't' time taken for processing, distributed time process of the heuristic method. An ECG's abnormal signal means numerous belongings at times an ECG irregularity can be a heart's beat standard dissimilarity of that has not distressed human health. An ECG's abnormal signal provides signal a medical emergency like heart attack.

**Algorithm 1: QRS interval extraction**

**Data:** Output of Moving window integrator

**Result:** QRS interval

```
initialization;
while / do
        read current
 if y> AND > T then
                start count,
                        if R-peak matched then
                           QRS-interval count;
                           Reset com
                           Delay 100 m
                 end
        end
end
```

It is pointed out low-frequency noise, generally upcoming from baseline wander can be detached by medium filter efficiently. Use a 200 ms width mean filter to take away QRS complex wave and P wave, then uses a 600 ms width medium strain to remove T wave, and afterward take away the filtered signal from the unique signal of ECG to obtain baseline enhancement signal. The 12-tap low-pass strain is used to eliminate power-line inquiring and high-frequency data noise.

### *D) MMNNS Machine Learning Algorithm*

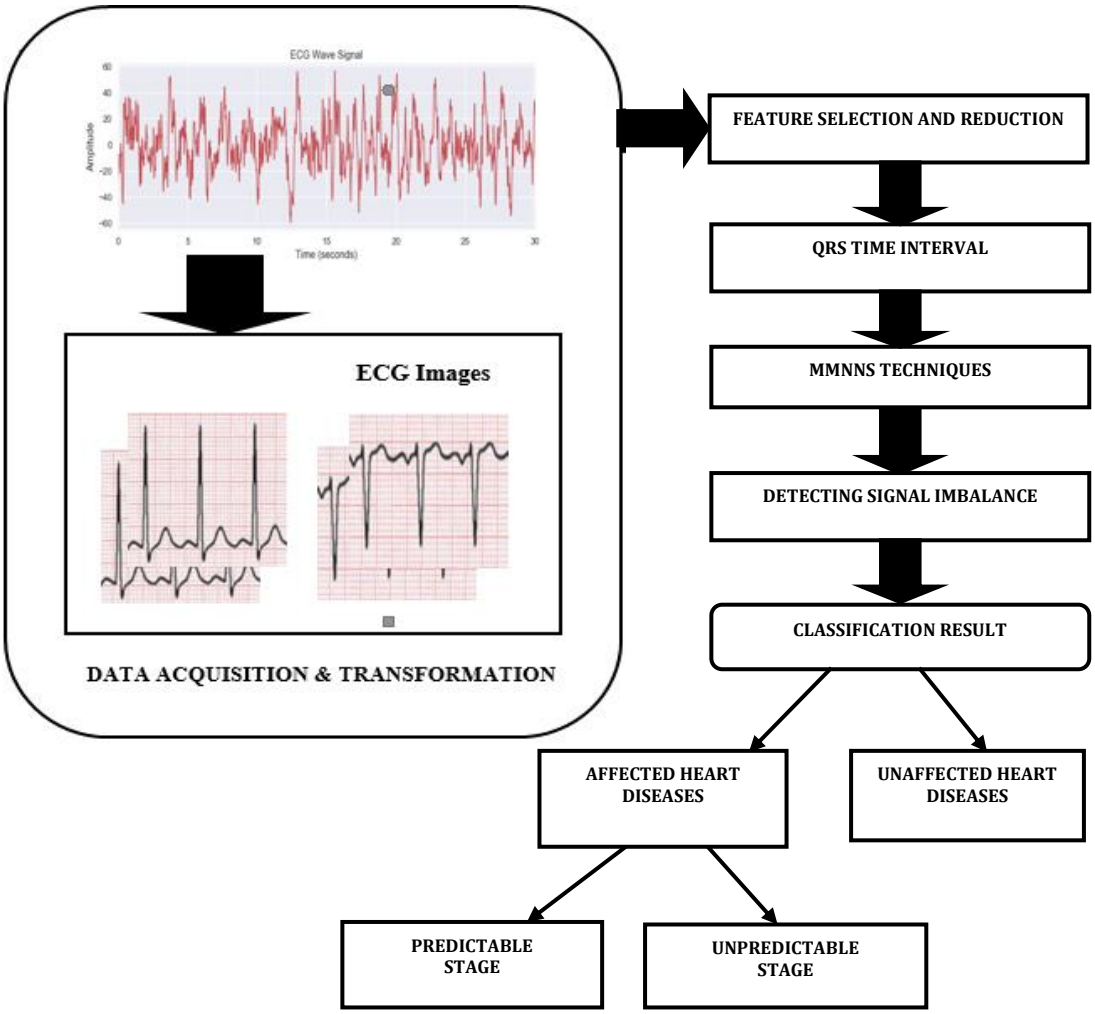


**Figure 3.** Overall Proposed System

The research analyzes a heart disease through the ECG signal of a person the dataset should be in the binary form of 0 and 1 of programming range value. And the dataset will be acquisition and transformation from ECG, data preprocessing, and features are classified throughout the signal by its range. The advanced machine learning techniques used to analyze by the comparison of the wave graph of ECG that will vary by its smoothening and level of normal heartbeat range will vary throughout the signal process.

The binary dataset will vary by the time interval between the P to T and the peaks of time interval QRS of the imbalanced signal will classify by using the MMNNS algorithm of machine learning techniques. And finally, the stages also classified into two categories: Predictable and unpredictable stages are analyzed. Using the comparison of the ECG dataset of a binary data signal with the proper range of ECG signal suppose if the range varies in the ECG wavelet then they are affected by Heart diseases otherwise they are unaffected by these diseases. The research contributes to better achievement of high accuracy, effective classification results, and taken to analyze possible minimum time.

## Comparative Analysis

The presented method is estimated that has been improved with the classification of the ML algorithm which was compared with proposed QRS and MMNNS techniques with the existing method. The presented heart disease detection method was compared with the existing techniques in the term of Specificity, Sensitivity, Accuracy, classification accuracy, and time complexities.

### *Sensitivity, Specificity, and Accuracy*

The performance of the presented improved ML technique MMNNS classifier is compared with existing Principal component analysis (PCA) and Gradient Boosting & AdaBoost classifier in terms of specificity, sensitivity, and accuracy of the heart disease prediction and they have been considered by the following equations (2-4).

The arithmetical estimations that might be measured are sensitivity, specificity, and accuracy.

$$\text{Sensitivity} = \frac{TP}{TP+FN} * 100 \quad (2)$$

$$\text{Specificity} = \frac{TN}{TN+FP} * 100 \quad (3)$$

$$Accuracy = \frac{TP+TN}{TP+FN+TN+FP} * 100 \quad (4)$$

In the above equations (2-4), TP specifies the True Positive, FP represents the False Positive, TN can indicate the True Negative, and FN can be the False Negative.

The proposed algorithm multi-module neural network system (MMNNS) algorithm of machine learning where compared with the PCA and Gradient Boosting & AdaBoost for given parameters of these three sectors.

**Table 2.** Performance analysis of proposed and existing ML methods

| Parameters | MMNNS (%) | PCA (%) | Gradient Boosting & AdaBoost (%) |
|---|---|---|---|
| Sensitivity | 91 | 82 | 78.6 |
| Specificity | 93.5 | 84.9 | 86.4 |
| Accuracy | 98.6 | 88.1 | 82.8 |

The performance comparison has been demonstrated in Table 2. From the above table 2, the MMNNS has provided 91% in sensitivity, 89% in specificity and 98.2% in the accuracy compared to existing techniques RNN and ANN classifier.

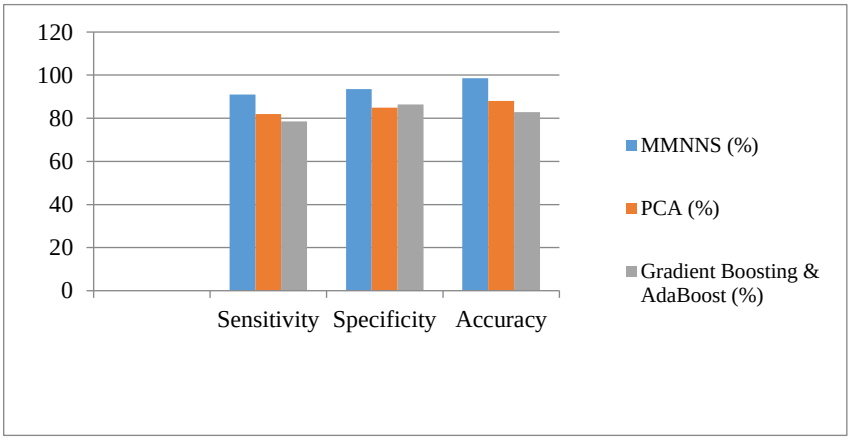


**Figure 4.** Classification of Heart Disease

The above chart demonstrated the color classification techniques the proposed method Hist multi-module neural network system (MMNNS) is used. By the above comparison, the show the accuracy classification rate compared with existing techniques.

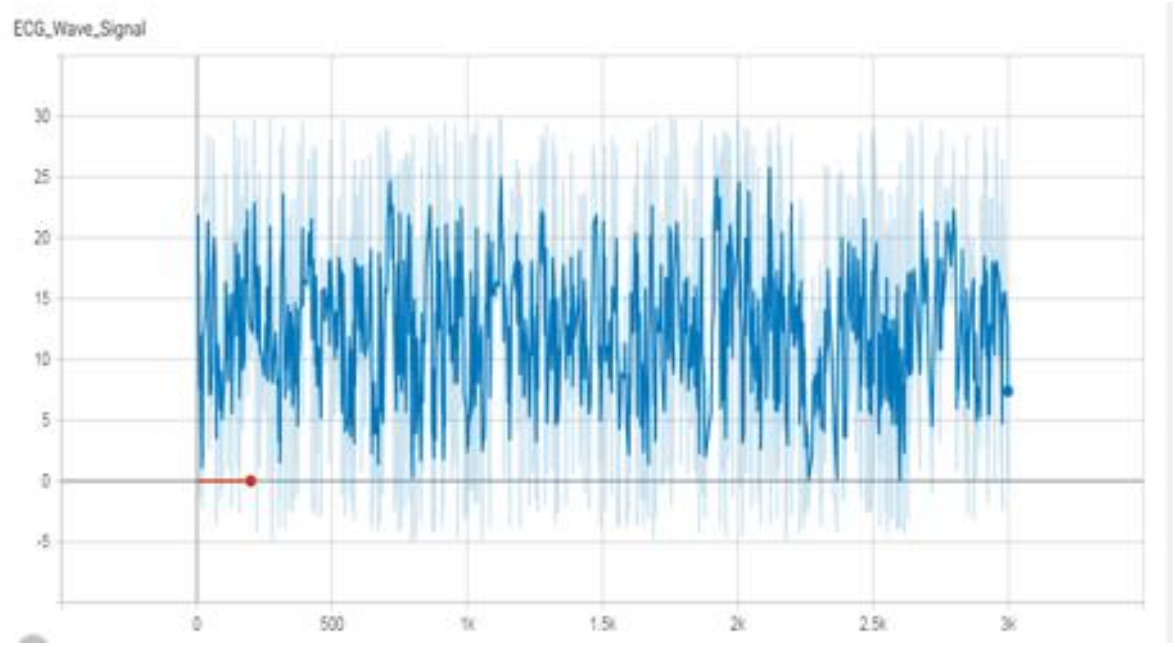


**Figure 5.** Normal ECG signal



The above signals of ECG can be normal heart rate of wave graph which was taken by the range value. This normal signal of ECG was contrasted with the known input signal of ECG.

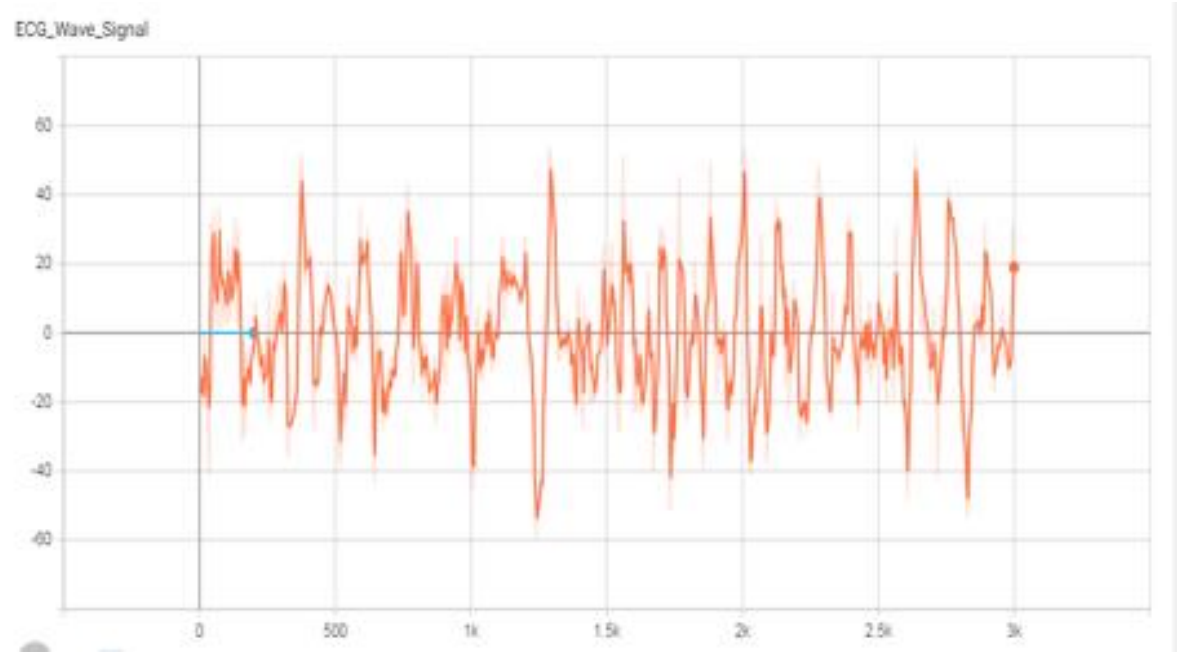


**Figure 6.** Abnormal ECG Signal

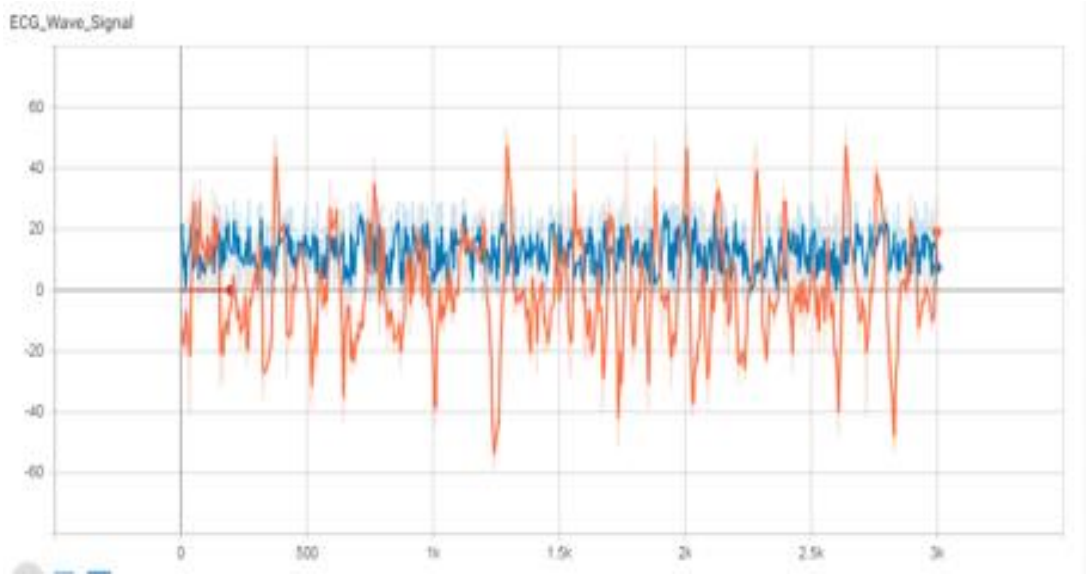


**Figure 7.** Comparison of Normal Vs Abnormal ECG signal

A normal dataset of heart disease compared with a given dataset illustrated in figure 7. By comparing these datasets can see the difference between the normal signal the present dataset shows the abnormal wave graph. Then the given dataset was affected by heart disease and finally can analyze the stages of heart diseases.

A result was provided by the output system as a prediction when the person contains heart issues, in the terms of class label 'Yes' or 'No'. In this system, an idea was provided about the heart condition directing heart attack before. When the person disease stage to contain heart problems after that the consequences were attained could be criterion of 'Yes' and vice versa.

**Table 3.** Calculations of a classification report

| Class Label | Precision (%) | Recall (%) | Support (%) |
|---|---|---|---|
| Yes | 93 | 91 | 93.9 |
| No | 92 | 92 | 91 |
| **Average** | 92.5 | 91.5 | 92.5 |



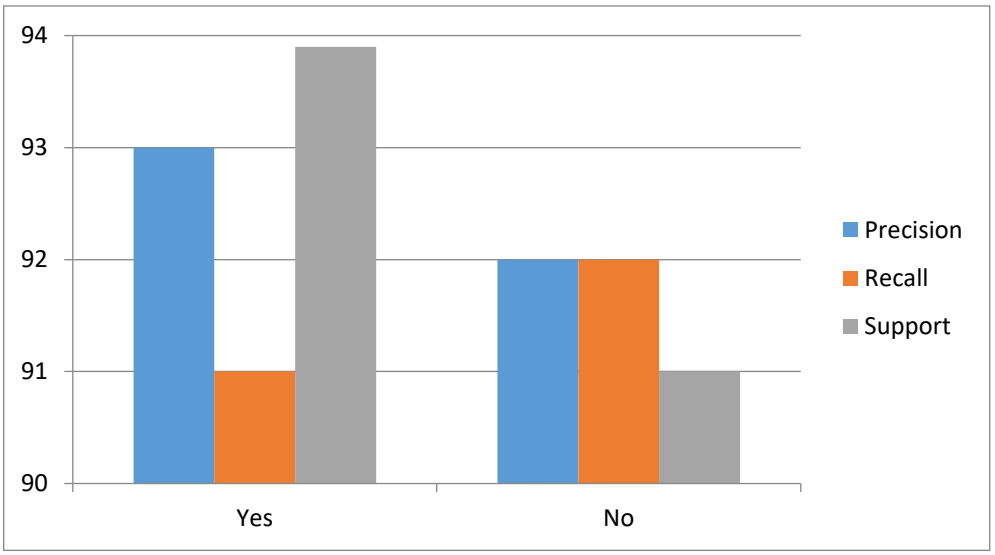


**Figure 8.** Graphical Representation of the results

### *The Time Complexity for Heart Disease Detection*

The heart disease detection will be applied by applying the proposed method of a time interval and MMNNS algorithms this has been implemented with the time by execution milliseconds by comparing with the existing classification this is established under the comparison results in figure 9.

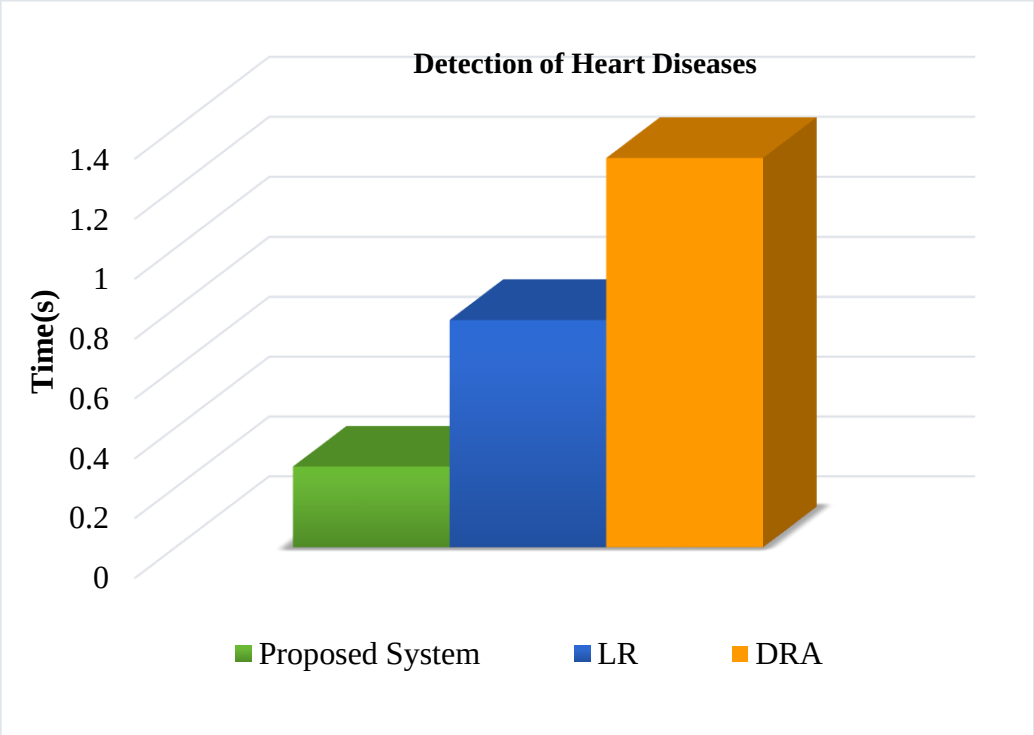


**Figure 9.** Comparison of Time complexity

The above comparison graph shows the proposed method of the QRS and MMNNS algorithm has taken minimum time for analyzing heart diseases. The existing method of Linear Regression (LR) and Dimensionality Reduction Algorithms (DRA) was compared with the proposed system. The experiment results established an efficient accurate result and minimum time taken for efficient heart disease prediction.

## Conclusion

ECG Signal is a significant tool and was employed to recognize difficulities of the heart function. Thus, an early detection of heart issues could give better person lives with a difficult task, other than with ML methods, automated analysis of heart diseases has been attained with ECG Signal examination and categorization. In this paper, MMNNS is presented in each stage of ECG signal examination, expressly for the ECG categorization process. For the ECG signal detection, QRS complexes of time intervals were discovered in this research. Investigated a stages-based heart disease approach for ECG signal examination in this paper wherever the mass of several ECG Signal of individual persons was categorized into stages of the prediction model. The MMNNS algorithm is used to detect an abnormal ECG signal by comparing it with normal ECG signals. The experiment results demonstrated an efficient and accurate result by computing sensitivity and Specificity.



## Bibliography

Dr.A.A. Gurjar, and Neha A. Sarnaik, "Heart Attack Detection by Heartbeat Sensing using Internet of Things: IoT" Volume: 05 Issue: 03 | Mar-2018, www.irjet.net.

Fajr Ibrahem Alarsan, and Mamoon Younes, "Analysis and classification of heart diseases using heartbeat features and machine learning algorithms" *Springer: journal of big data.* https://doi.org/10.1186/s40537-019-0244-x.

C. Sowmiya, and P. Sumitra, "Analytical study of heart disease diagnosis using classification techniques", IEEE Xplore: 01 March 2018, DOI: 10.1109/ITCOSP.2017.8303115.

P. Umasankar, and V. Thiagarasu, "Decision Support System for Heart Disease Diagnosis Using Interval Vague Set and Fuzzy Association Rule Mining", *4th International Conference on Devices, Circuits and Systems (ICDCS).* DOI: 10.1109/ICDCSyst.2018.8605065.

P. Sudeshna, S. Bhanumathi, and M.R. Anish Hamlin, "Identifying symptoms and treatment for heart disease from biomedical literature using text data mining", IEEE Xplore: 15 February 2018. DOI: 10.1109/ICCPEIC.2017.8290359.

Rairikar, A., Kulkarni, V., Sabale, V., Kale, H., & Lamgunde, A. (2017, June). "Heart disease prediction using data mining techniques". *In International Conference on Intelligent Computing and Control (I2C2)* (pp. 1-8). IEEE.

Gandhi, Monika, and Shailendra Narayan Singh. "Predictions in heart disease using techniques of data mining." *In International Conference on Futuristic Trends on*

*Computational Analysis and Knowledge Management (ABLAZE),* pp. 520-525. IEEE, 2015.

C. Ye, B.V.K.V. Kumar, and M.T. Coimbra, "Heartbeat classification using morphological and dynamic features of ECG signals," *IEEE Transactions on Biomedical Engineering,* vol. 59, no. 10, pp. 2930–2941, Oct. 2012.

A. Aldallal, and Al-Moosa, A.A.A. (2018, September). "Using Data Mining Techniques to Predict Diabetes and Heart Diseases". *In 4th International Conference on Frontiers of Signal Processing (ICFSP)* (pp. 150- 154). IEEE.

Sultana, Marjia, Afrin Haider, and Mohammad Shorif Uddin. "Analysis of data mining techniques for heart disease prediction." *In 3rd International Conference on Electrical Engineering and Information Communication Technology (ICEEICT),* pp. 1-5. IEEE, 2016.

Shetty, Deeraj, Kishor Rit, Sohail Shaikh, and Nikita Patil. "Diabetes disease prediction using data mining." *In International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS),* pp. 1-5. IEEE, 2017.

U.R. Acharya, H. Fujita, O.S. Lih, Y. Hagiwara, J.H. Tan, and M. Adam, "Automated detection of arrhythmias using different intervals of tachycardia ECG segments with a convolutional neural network," *Information sciences,* vol. 405, pp. 81–90, 2017.

Dewan, Ankita, and Meghna Sharma. "Prediction of heart disease using a hybrid technique in data mining classification." *In 2nd International Conference on Computing for Sustainable Global Development (INDIACom),* pp. 704-706. IEEE, 2015.

A.B. De Luna, P. Coumel, and J.F. Leclercq, "Ambulatory sudden cardiac death: Mechanisms of production of fatal arrhythmia based on data from 157 cases," *American Heart Journal,* vol. 117, no. 1, pp. 151–159, 1989.

T. Ince, S. Kiranyaz, and M. Gabbouj, "A generic and robust system for automated patient-specific classification of ECG signals," *IEEE Transactions on Biomedical Engineering,* vol. 56, no. 5, pp. 1415–1426, 2009.

F. Agrafioti and D. Hatzinakos, "Fusion of ECG sources for human identification," *In 3rd International Symposium on Communications, Control and Signal Processing,* 2008, pp. 1542–1547.

Y.H. Hu, S. Palreddy, and W.J. Tompkins, "A patient-adaptable ECG beat classifier using a mixture of experts approach," *IEEE Transactions on Biomedical Engineering,* vol. 44, no. 9, pp. 891–900, 1997.

J.H. Tan, Y. Hagiwara, W. Pang, I. Lim, O.S. Lih, M. Adam, R.S. Tan, M. Chen, and U. R. Acharya, "Application of stacked convolutional and long short-term memory network for accurate identification of CAD ECG signals," *Computers in Biology and Medicine,* vol. 94, pp. 19–26, 2018.

U.R. Acharya, H. Fujita, O.S. Lih, Y. Hagiwara, J.H. Tan, and M. Adam, "Application of deep convolutional neural network for automated detection of myocardial infarction using ECG signals," Information sciences, vol. 415, pp. 190–198, 2017.

D.M.B. et al: "Microvolt T-wave alternans and the risk of death or sustained ventricular arrhythmias in patients with left ventricular dysfunction," Journal of the American College of Cardiology, vol. 47, no. 2, pp. 456–463, 2006.

## Authors Profile

**Dr.P. Marikkannu** is currently working as Assistant Professor of IT. He was awarded the Gold medal for academic excellence towards his M.Tech degree by College of Engineering, Guindy, Anna University, and Chennai. He received his Ph.D degree in the year 2012 from Anna University, Chennai. He was instrumental in the installation of Linux server in the department. He has served as Director – publications and Assistance COE-Directorate of Distance Education in erstwhile AUT, Coimbatore. He is a life member of ISTE.

**Mr.K. Babu** is currently doing Ph.D in Anna University Regional Campus, Coimbatore. He received his M.Tech (Multimedia Technology) degree in the year 2013 from Anna University, Regional Campus, and Coimbatore. He received his B.Tech (Information Technology) degree in the year 2011 from E.S College of Engineering and Technology.