EARLY DETECTION OF HEART DISEASE USING BRANN

[1]Archana Lopes, [2]Sangeeta Parshionikar, [3]Heenakausar Pendhari

[1]Assistant Professor, [2]Assistant Professor, [3]Assistant Professor
[1]Department of Electronics,
[1]Fr. C.R.C.E., Mumbai, India

Abstract : **The International Congestive Heart Failure study said that heart related diseases occur a decade early , but the lack of awareness, out of pocket expenditure and most importantly wrong diagnosis of heart disease by doctors increases mortality rate. This paper focuses on early detection of heart disease based on features of ECG signal and neural network algorithm. This paper suggest a second opinion to the doctors so that wrong diagnosis can be avoided.**

IndexTerms **- Heart Failure, Back Propagation Neural Network, Recurrent Neural Network, Bayesian Network.**

I. INTRODUCTION

Heart disease addresses the condition that affects the structuring and functioning of the heart in some way or another. Commonly known heart conditions include heart failure, heart valve disease, congenital heart diseases etc. The different types of heart disease are Angina, Arrhythmia, Atrial Fibrillation, Cardiomyopathy etc. For all the heart diseases periodic medical tests are needed. The best way to detect a problem is to opt for a cardiovascular disease test online.

Computer aided techniques for detection of heart disease was a topic of research for last few years. Current predictive methods like ACC are based on eight risk factors including age, cholesterol level and blood pressure. AI has a crucial role to play in devising of healthcare models for future.

The paper by Michael Auli[1] suggests a method where ANN is trained using genetic algorithm . This paper uses ECG signals for training Artificial Neural Network). After the R peak of the QRS complex is detected, a window containing an ECG period is formed around the R peak. The significant frequency components of the signal in the window are used to form the feature vectors. Genetic algorithm, grow and learn (GAL), multi-layer perceptron (MLP), and Kohonen networks are comparatively investigated to detect seven different ECG waveforms. It is observed that the proposed network results in better classification performance with less number of nodes.

ECG events can be detected using wavelets for feature extraction[2] and classified using ANN. By using the dyadic wavelet transform, the limitations of other methods in detecting ECG features such as QRS complex, the onsets and offsets of P and T waves are overcome. The ECG baseline is approximated using discrete least squares approximation. On classification, two paradigms of learning, supervised and unsupervised, for training the ANN modes are investigated. The Back propagation algorithm and the Kohonen's self-organizing feature map algorithm were used for supervised and unsupervised learning, respectively. The system is evaluated using the MIT-BIH database.

Heart disease can be classified using Recurrent neural networks . [3] Different RNN architectures with various parameter settings were evaluated, including traditional, long short-term memory (LSTM), gated recurrent unit (GRU), unidirectional, and bidirectional networks. Unlike many existing methods, the RNN-based method does not require any feature extraction .

This paper aims at early pre diagnosis of heart disease like arrhythmia using neural networks like Recurrent neural network and back Propagation Neural Network. This paper compares the result of heart disease prediction using Recurrent Neural Network and Back Propagation Neural Network. The prediction is based on the ECG reports of a patient.

II. DATABASE USED

The MIT-BIH database is used for the work. The MIT-BIH Arrhythmia Database contains 48 half-hour excerpts of two channel ambulatory ECG recordings, obtained from 47 subjects studied by the BIH Arrhythmia Laboratory between 1975 and 1979[4]. The recordings were digitized at 360 samples per second per channel with 11-bit resolution over a 10 mV range. Two or more cardiologists independently annotated each record; disagreements were resolved to obtain the computer-readable reference annotations for each beat included with the database. This directory contains the entire MIT-BIH Arrhythmia Database. About half (25 of 48 complete records, and reference annotation files for all 48 records) of this database has been freely available here.

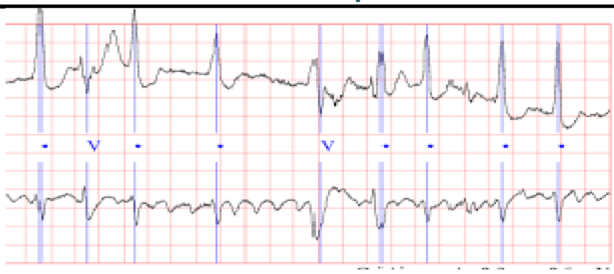

Figure 1: Sample ECG signal from the database

III. PROPOSED ALGORITHM

The work is divided in two phases. In the first phase neural network is trained for the selected training dataset and then the neural network is tested for accuracy. In the second phase , a web application is developed where patient can upload its ECG report and get the prediction of heart disease on line. The web application predicts the heart disease based on the past history of the patient

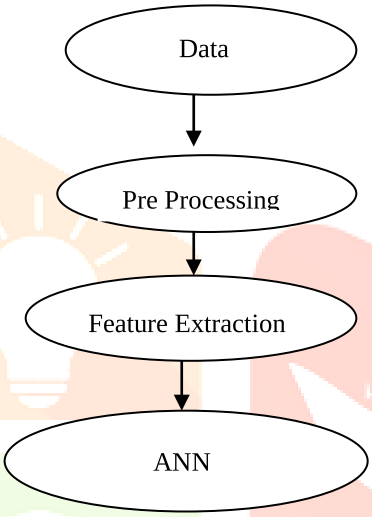

Figure 2: Algorithm

Pre Processing:

Given an ECG Recording, the first step is to segment recording into individual heartbeat waveforms. Since the R peak is the most Prominent peak, it can be used as a marker of a given Heartbeat Waveform. Pre processing can be done using Pan -Tompkins algorithm[5] or by using wavelet transform.

3.1 Pan Tompkins algorithm has four steps : Band Pass Filtering, Differentiation, Squaring and Moving window integration .

3.1.1 Band Pass Filtering: The band pass filter for the QRS detection algorithm reduces noise in the ECG signal by matching the spectrum of the average QRS complex. This attenuates noise due to muscle noise, power line interference, baseline wander, T wave interference.

3.1.2 Differentiation: The derivative procedure suppresses the low frequency components of P and T waves, and provides a large gain to the high-frequency components arising from the high slopes of the QRS Complex.

3.1.3 Squaring: The squaring operation makes the result positive and emphasizes large differences resulting from QRS complexes; the small differences arising from P and T waves are suppressed. The high frequency components in the signal related to the QRS complex are further enhanced.

3.1.4 Moving Window Integration: This integrator sums the area under the squared waveform over a suitable interval advances one sample interval, and integrates the new pre defined interval window. The half width of window has been chosen as 27 to include the time duration of extended abnormal QRS complexes, but short enough that it does not overlap both a QRS complex and a T-wave.

3.2 Wavelet Transform:

Because the frequency band of ECG signals and noises is overlapping, it still needs an in-depth study to isolate completely from noises of ECG signals, so they have some shortcomings for eliminating noise of ECG signals. For above mentioned problems, we adopt Wavelet Transform to filter out the myo-electrical interference, the power frequency interference and the baseline drift[6].

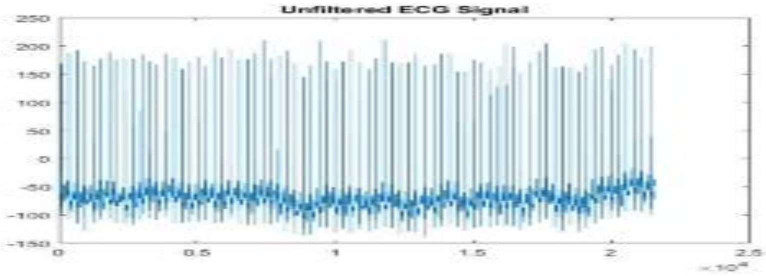

Figure 3: Unfiltered ECG Signal

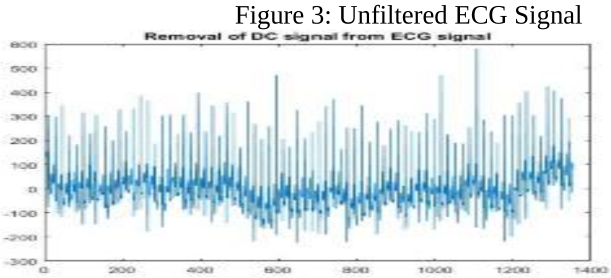

Figure 4: Removal of DC level from ECG signal

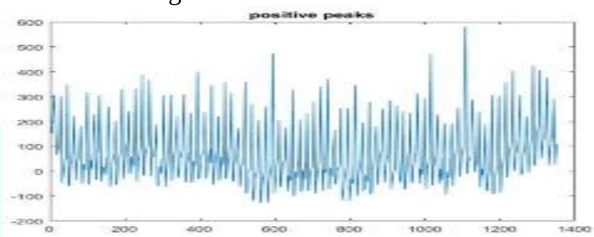

Figure 5: Positive Peaks

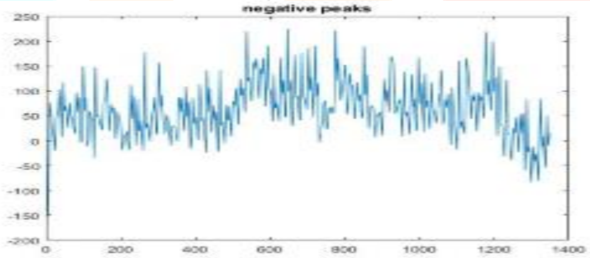

Figure 6: Negative Peaks

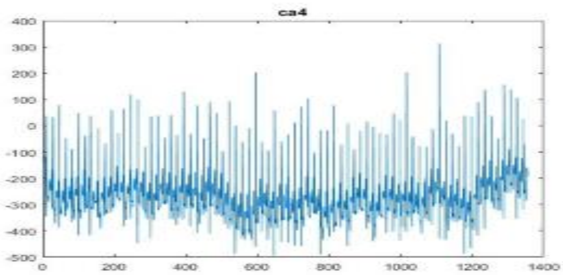

Figure 7: ECG Filtered by wavelet transform

3.3. Feature Extraction:

For designing the neural network features like amplitude of p, q,r, s peaks, duration of p, q, r, s peaks and the average heart rate of the patient whose ECG signal is given as input are taken. The features are extracted using threshold method and Principal Component Analysis.

3.4 Artificial Neural Network:

An ANN also called neural network is a mathematical model which tries to imitate the human brain. Analogically, ANN has mainly three layers, input layer, hidden layer and an output layer. Every layer may have different number of neurons. There can be more than one hidden layers. There are mainly two training algorithms used in ANN, namely Supervised and Unsupervised Learning Algorithms.

Back propagation neural network is a supervised training algorithm which has two phases, namely, forward phase and backward phase. In the forward phase an input is applied at the input layer and is propagated to the output layer. At the output layer, the error is calculated. In the backward phase, the error is propagated back to the input and weights are modified.[8]

Bayesian regularized artificial neural networks (BRANNs)[7] are more robust than standard back-propagation nets and can reduce or eliminate the need for lengthy cross-validation. Bayesian regularization is a mathematical process that converts a nonlinear regression into a "well-posed" statistical problem in the manner of a ridge regression. The advantage of BRANNs is that the models are robust and the validation process, which scales as O(N2) in normal regression methods, such as back propagation, is unnecessary.

A Feed forward neural net is designed for twelve input neurons, ten hidden neurons and four output neurons. The database is divided into training set and testing set. The training data set is used in the first phase of neural net design. The training algorithm used is back propagation neural network algorithm. The training is continued till mean square error is less than set mean square error or the number of epochs are reached. The training algorithm is based on Bayesian regularized neural network. Plots of the training set is shown in the figure below:

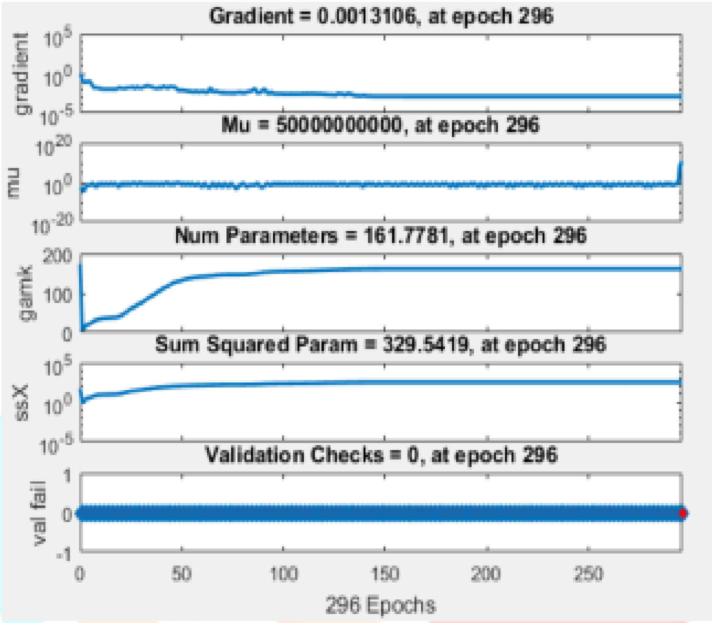

Figure 8: Training state

IV. RESULT

In this paper, the proposed neural network algorithm is trained validated and tested using three hundred and sixty instances of MIT-BIH database. From each data vector twelve features were selected for training the neural net. The accuracy of the proposed algorithm is shown in the plot in Figure 9.

Matlab plots confusion plot to test the result of training. On the confusion matrix plot, the rows correspond to the predicted class (Output Class) and the columns correspond to the true class (Target Class). The diagonal cells correspond to observations that are correctly classified. The off-diagonal cells correspond to incorrectly classified observations. Both the number of observations and the percentage of the total number of observations are shown in each cell. The column on the far right of the plot shows the percentages of all the examples predicted to belong to each class that are correctly and incorrectly classified. These metrics are often called the precision (or positive predictive value) and false discovery rate, respectively. The row at the bottom of the plot shows the percentages of all the examples belonging to each class that are correctly and incorrectly classified. These metrics are often called the recall (or true positive rate) and false negative rate, respectively. The cell in the bottom right of the plot shows the overall accuracy.

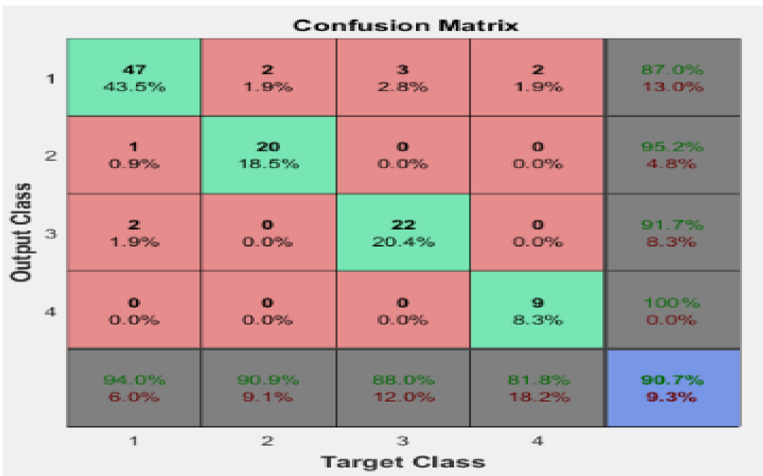

Figure 9: Confusion Plot

As it can be seen the overall accuracy of the system is 91 %.

In the second phase a web page is developed which has user login and admin login. In user login , user can upload ECG signal along with other details. In admin login, new user and new doctor can be added as well as new dataset can be created and selected for training the neural net.

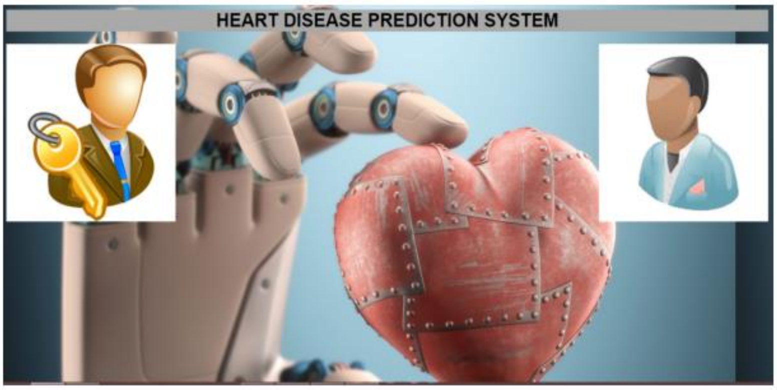

Figure 10: A Web home page

V. CONCLUSION

We proposed a novel predictive framework for heart disease detection using Bayesian regularization method. As compared to back propagation neural network and recurrent network models , this method exhibited superior performance in predicting the heart disease. Future work will include incorporating a mobile application for heart disease prediction and also incorporating expert knowledge into our framework and expanding the additional health care applications

REFERENCES

[1] Michael Auli, Michel Galley, Chris Quirk, and Geoffrey Zweig. Joint language and translation modeling with recurrent neural networks. In EMNLP, pages 1044–1054, 2013.
[2] Pierre Baldi and Gianluca Pollastri. The principled design of large-scale recursive neural network architectures–DAGRNNs
and the protein structure prediction problem. The Journal of Machine Learning Research, 4:575–602,2003.
[3] Richard K. Belew, John McInerney, and Nicol N. Schraudolph. Evolving networks: Using the genetic algorithm with connectionist learning. In In. Citeseer, 1990.
[4] Moody GB, Mark RG. The impact of the MIT-BIH Arrhythmia Database. *IEEE Eng in Med and Biol* 20(3):45-50 (May-June 2001). (PMID: 11446209)
[5] Jiapu Pan, Willis J. Tompkins, A Real Time QRS Detection Algorithm
[6] Yun Fu Tan , Lei Du, Study on Wavelet Transform in the processing of ECG signals.
[7] Burden F, Winkler D, Bayesian regularization of neural networks. Methods Mol Biol. 2008;458:25-44. Review
[8] S.Florence, N.G.Bhuvaneswari Amma, G.Annapoorani, K.Malathi, "Predicting the Risk of Heart Attacks using Neural Network and Decision Tree", International Journal of Innovative Research in Computer and Communication Engineering (An ISO 3297: 2007 Certified Organization) Vol. 2, Issue 11, November 2014