

Remote Heart Risk Monitoring System based on Efficient Neural Network and Evolutionary Algorithm

S. Radhimeenakshi^{1*} and G. M. Nasira²

¹Department of Computer Science, Tiruppur Kumaran College for Women, Bharathiar University, Tirupur - 641687, Tamil Nadu, India; radhimeenakshisenthilkumar@yahoo.in

²Department of Computer Applications, Chikkanna Government College, Tirupur - 641602, Tamil Nadu, India; nasiragm99@yahoo.com

Abstract

Objective: The objective of this paper is to predict the risk level of Heart Disease by applying Probabilistic Neural Network trained with Particle Swarm Optimization in case of Remote Health Monitoring. **Methods:** In order to achieve the aim of the activity, we propose hybrid model of Particle Swarm Optimization (PSO) and Probabilistic Neural Network (PNN). PSO is a population based meta-heuristic Evolutionary Algorithm (EA) whose goal is to explore the search space in order to find near – optimal solutions for feature selection. The optimal features selected can be used for prediction system to develop a classification model using probabilistic Neural Network. **Results:** First, we quantify the clinical data set from the UCI machine learning repository and measured the complexity. There are 13 attributes are used such as the age which identifies the age of the person, chest pain type has 4 values, serum cholesterol level, blood sugar, resting ECG results, serum cholesterol level, amount of heart rate achieved, x old peak, number of major vessels colored by fluoroscopy, slope of the peak exercise ST segment, thal, sex, height, weight and additional factor smoking. It has been shown that the time complexity of hybridizing PSO and PNN obtained the promising results compared to other two algorithms such as regression tree and PSO optimization. We also proposed the data mining process to deal with complexity, missing values and high dimensionality followed by incorporating the data mining functionalities like characterization, discrimination, association, classification, prediction and evolution analysis. The experiment carried out in Java on stat log heart disease data set performs better in all noise conditions. **Conclusion:** The performance was evaluated in terms of time complexity, accuracy, sensitivity and specificity and it proved that the hybrid model of PSO and PNN outperformed the Regression tree and PSO.

Keywords: Expectation Maximization (EM), Heart Disease, Particle Swarm Optimization (PSO), Probabilistic Neural Network (PNN), Remote Heart Risk Monitoring System (RHRMS)

1. Introduction

Among the countries of whole universe, Heart Disease and Cardiovascular disorders are the main threat to the humanity. Even if there are prominent curable for this disease it is still a major disease that affects the human beings. According to a report published by the Associated Chambers of Commerce and Industry of India (ASSOCHAM), India has a considerable increase

in the number of Heart Disease cases due to unhealthy life style, diet and lesser level of physical activity. Mr. K.K. Aggarwal president Health Care Foundation of India has asserted that 2.4 million Indians die due to Heart disease every Year .This will continue to grow and there should be a substantial treatment and preventive measures to save the lives. As a part of it, a development of an automated system that predicts the risk level of heart disease has been done and researched by many authors.

* Author for correspondence

In many scientific disciplines, studies that predict or forecast what will happen in the future have contributed to our understanding of the world. The value of scientific studies that provide models to inform strategies that can modify and possibly mitigate future events is of importance to society. These prediction models have been accepted as valuable tools by scientists and have provided critical information for the development of strategies to modify predicted trends.

This proposed method is examined by utilizing the European stat log data gathered from the UCI repository. Data Mining is an important extraction of hidden, unknown and potential helpful information about data. Data mining gives a set of technique to find hidden or unknown pattern from data. Heart Risk (HR) prediction is important and complicated task that is essential to be executed efficiently and accurately diagnosing the heart problem based on doctor's knowledge and experience.

There are numerous tools available for a data mining approach such as various types of regression, decision trees and neural networks¹. The neural network is trained from the various historical data with the trust that it will find hidden conditions and that it will have the capacity to utilize them for prediction. In this paper, PSO algorithm² is utilized for prediction of heart failure and optimizes the input and hidden neurons, number of layers, the sort of transfer functions etc. Expectation Maximization (EM) Algorithm is an iterative optimization approach to estimate some unknown Heart disease parameters. In PSO, the positions of the heart disease particles are optimized using particle velocity in D-dimensional search space³ and then the Probabilistic Neural Network (PNN) used for prediction process. The PNN-based classifier is implemented to achieve the prediction tasks.

Health Care Monitoring (HCM) system supports to finding the patients with Chronic Diseases (CD). Ambient Intelligence (AmI) for personalized healthcare monitoring is a promising solution that gives the medical services in minimum healthcare cost⁴. The author summarizes recent evolution in the field of AmI for HCM. They focused the concept of Artificial Intelligence (AI) and AmI and the data mining approach utilized in AmI application that enable wearable sensor technology to monitor adults, older and patient with CD. Wireless Sensor Networks (WSNs) is used for research communication technology which is using wearable system for healthcare monitoring. The Computational Intelligence (CI) has utilized in

different kind of medical application. Among them biomedical engineering has progressed into a one of the growing application part of computational intelligence. The author reviewed briefly about computational intelligence in biomedical application and also demonstrated the efficiency of the CI for medical diagnosis⁵.

The classification method evaluates the cardiovascular pathology based on Electrocardiographic (ECG) and clinical Holter data. So the author proposed a new method for predicting a coronary stenosis difficult revascularization in patients. A solution of this issue set in context of rough set theory and methods. From the rough set theory, the classification based on decision tree is used for local discretization method calculation⁶. In this approach the decision tree creating approach emphasizes the visibility of the object appropriate decision classes specified by human experts. This method may be utilized to evaluate the coronary revascularization.

The wearable ECG sensors are very popular in medical research, these sensors run on batteries which are used to monitor the fetal health named as FECG sensors. Thus the author proposes the cellular connection based optimized signal transfer to central server. In this research, the battery life increases and the load on buffer memory decreases. Sensors are utilized as FECG transmission unit and data collection. FECG sensor data transmission is taken small amount of energy. The FECG signal optimization using Genetic Algorithm (GA) to minimize its size with no loss of feature and reduce the data transmission process⁷. GA is a robust and effective optimization technique which is based on Fitness function. Fitness function is utilized for ECG data compression.

The authors propose an intelligent Adaptive Filter for effective feature extraction⁸. Noise reduction technique is utilizing Particle Swarm Optimization (PSO) based on conventional FIR adaptive filtering. Though SVM is a well proven technique for classification, ANN performs well on suspicious records⁹. The authors propose a demonstration of clinical support system for diabetes mellitus by clustering and classification¹² and has recommended for extension in other health care systems.

2. Process

Many people now-a-days undergone with heart surgery or affected with severe heart issues. Therefore, monitoring the heart beat from their residential becomes essential in the

field of medicine. Hence a device is needed to monitor the heart function remotely and it must be reliable and robust in critical condition. The data is encrypted automatically from the transmitter and send to the central database. This transmission takes place by the Global System for Mobile devices (GSM) technology. In the central data base, the encrypted data after transform into the readable form send to the medical expert side who can verify the report transmitted and take measure accordingly. The tool detects the abnormalities in the heart rate in the early stage and alerts the medical experts immediately. These remote monitoring reasonably decreases the burden of patients and doctors precious time, regular follow-ups.

2.1 Preprocessing and Feature Extraction

The current medical world presents the equipped technology which gives an automated health prediction. Remote health care monitoring power resources are limited. So evolved in the low-complex delineation (feature extraction) algorithms¹⁰. ECG intervals are optimized from the extracted features using PSO optimization and then classify the data as normal or abnormal using PNN classifier.

3. Particle Swarm Optimization

The main aim of integrating Particle Swarm Optimization (PSO) and Expectation Maximization (EM) algorithm is used for particles location improvement. The main advantage of this integration is that both the algorithms compensate their disadvantages. Such as, the EM algorithm is very efficient but exceptionally sensitive to the initial estimation. The low initial point is getting the computational difficulties. Poor starting point might make the premature termination of the EM algorithm. PSO is a global search procedure, which requires abundant computational effort than traditional search procedures.

In this paper, the PSO is used to search the high value and also employ it as a starting idea to explore EM algorithm which upgrade the motif location. The EM scoring methods, a particle $P = p_i$ is defined as a vector, the velocity vector $V = v_i$ is maintained and the current vector velocity of the particle in the given sequence is v_i . In the proposed system, the system asks the user for the data. Then the standard implementation of PSO is used where the number of particles is represented with the random

v_i and p_i values. The individual factor, social factor and the maximum velocity of the particles are defined and are permitted to fly. In the end of each step, the EM score value is examined for each particle and the gbest and pbest values are updated simultaneously till the particles fail to upgrade the gbest score values which is the termination condition. The P matrix for the last gbest score is ultimately used as a seed for the EM algorithm when the termination condition met. Then the process continues iteratively till no improvement met. The utilization of PSO to seed the EM algorithm also leads to the frequent termination in the local solution. Hence in the proposed system, the PSO is reinitialized along with the use of re-run algorithm to process until no improvement is found. Proposed heart monitoring system is shown in Figure 1.

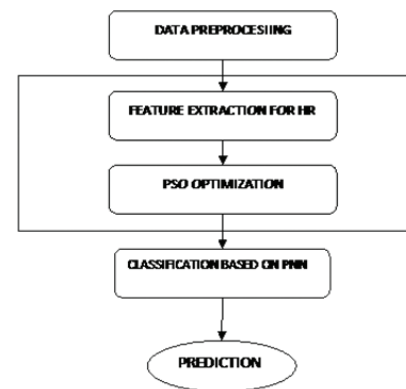


Figure 1. Heart monitoring system.

4. Probabilistic Neural Network (PNN) Classification

The Heart Risk (HR) variability is considered by the quantitated amplitude is shown in Table 1.

Table 1. Quantitated Amplitude

Absent	Undetectable
Minimal	Greater than undetectable and less than or equal to 5 bpm
Moderate	6 bpm-25-bpm
Marked	Greater than 25 bpm

Probabilistic Neural Network (PNN) consists of input, hidden, summation, and output layer is presented in Figure 2. The input layer $\Phi = [\phi_{11}, \dots, \phi_{1m1}, \dots, \phi_{j1}, \dots, \phi_{jmg}, \dots, \phi_{p1}, \dots, \phi_{pMP}]$ is linked to the PNN, and inputs are the heart

disease attributes. The amount of hidden nodes H_k ($k = 1, 2, 3, \dots, K$) is identical to the amount of training data, even though the number of output nodes O_t ($t = 1, 2, 3, \dots, m$) and summation nodes S_t equals to the classified types. Connecting input node and hidden node as denotes the weights w_{ki}^{IH} are determined by K output and input training pairs. The final output node is defined by eq. (1)

$$H_k = \exp \left[- \sum_{k=1}^K \frac{(\varphi_{ji} - w_{ki}^{IH})^2}{2\sigma_k^2} \right] \quad (1)$$

$$O_t = \frac{\sum_{k=1}^K w_{tk}^{HS} H_k}{\sum_{k=1}^K H_k} \quad t = 1, 2, 3, \dots, m \quad (2)$$

The eq. (2) is planned to minimize the anticipated squared slip capacity

The optimal σ_k is projected to minimize the squared error function $e_t(\varphi, T) = [T_t - O_t(\varphi)]^2$ where T_t is the preferred output for input vector Φ . The optimization method is utilized to σ_k with iteration process, as in eq. (3)

$$\sigma_k(ltr + 1) = \sigma_k(ltr) + \eta \frac{\partial e_t(\varphi, T)}{\partial \sigma_k} \quad (3)$$

Where ltr is iteration number and η is the learning rate. The PNN algorithm contains two stages: "Recalling Stage" and "Learning Stage", as explained below.

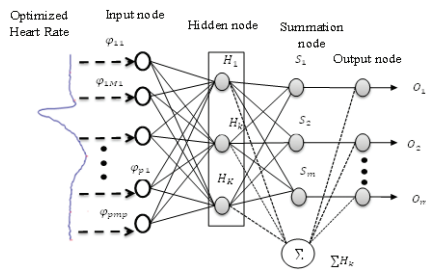


Figure 2. PNN classification.

In Learning Stage each training data

$\Phi(K) = [\varphi_{11}(K), \dots, \varphi_{1m1}(K), \dots, \varphi_{j1}(K), \dots, \varphi_{jmj}(K), \dots, \varphi_{p1}(K), \dots, \varphi_{pmp}(K)]$, $k = 1, 2, 3, \dots, K$, $I = 1, 2, 3, \dots, M_j$, $j = 1, 2, 3, \dots, P$ create weights w_{ki}^{IH} between hidden node and input node by using eq. (9)

$$w_{ki}^{IH} = \Phi(K) \quad (4)$$

Create w_{tk}^{HS} weight between summation nodes S_t and hidden node H_k by

$$w_{tk}^{HS} = \begin{cases} 1 & (t=1, 2, 3, \dots, m) \\ 0 & \end{cases} \quad (5)$$

Where the values of w_{tk}^{HS} are the recognized outputs linked with each pattern w_{ki}^{IH} . Connection weights from hidden nodes H_k .

In recalling stage, first get the weights such as w_{ki}^{IH} and w_{tk}^{HS} then employ the testing vector to the $X[x_{11}, \dots, x_{1m1}, \dots, x_{ji}, \dots, x_{jmj}, \dots, x_{p1}, \dots, x_{pmp}]$ transformation function. Compute the optimized features by

$$\varphi_{ji} = c_j n' + d_j x_{ji} + f_j + g_j \sin\left(\frac{\pi n'}{D}\right) + h_j \sin\left(\frac{\pi n'}{D}\right) \quad (6)$$

$$\Phi(\varphi_{ij}) = \bigcup_{j=1}^P \varphi_{ij}, i=1, 2, 3, \dots, M_j \quad (7)$$

Finally, compute the output from hidden node H_k , $k = 1, 2, 3, \dots, K$ by using Gaussian activation function.

$$H_k = \exp \left[- \sum_{k=1}^K \frac{(\varphi_{ji} - w_{ki}^{IH})^2}{2\sigma_k^2} \right] \quad (8)$$

Where $\sigma_1 = \sigma_2 = \dots, \sigma_k = \sigma$ and the optimal value can be acquired by utilizing optimization method. The outputs of node O_t is computed by utilizing the eq.2

4.1 PNN Classification Algorithm

Step 1: Pre-Processing

Acquire raw data from remote monitoring system; employ the low-complex delineation (feature extraction) algorithms.

Step 2: Optimization

In PSO consider the problematic area having the maximum information according to the EM scoring methods, a particle $P = p_i$ is defined as a vector which consists of first character location of the ECG waves in the given sequence s_i . The velocity vector $V - v_i$ is maintained and the current vector velocity of the particle in the given sequence is v_i .

The EM score value is examined for each particle and the pbest and gbest values are updated simultaneously, finally, get optimized output.

Step 3: Training of PNN

This process is executed for the purpose of training the PNN and the data are taken from the optimized attributes to train PNN.

Transform the data into the format of PNN

$\Phi(K)=[\varphi_{11}(K),\dots,\varphi_{1m1}(K),\dots,\varphi_{j1}(K),\dots,\varphi_{jmi}(K),\dots,\varphi_{p1}(K),\dots,\varphi_{pmp}(K)]$

The number of samples is equal to number of training instances. The instances k groups where each group comprises one class of vector.

Step 4: Gaussian Function

Each feature vector in set k describe the summed Gaussian output function

$$H_k = \exp \left[- \sum_{k=1}^K \frac{(\varphi_{ji} - w_{ki}^{IH})^2}{2\sigma_k^2} \right]$$

Each Gaussian function in each classified it to read as input test vector and calculate the Gaussian function for each hidden layer which is as the inputs to output layer.

Step 5: Testing of PNN Initiate with the first record of the best values. For each testing instance, testing the QRS region belongs to the non-QRS region using

$$O_t = \frac{\sum_{k=1}^K w_{tk}^{HS} H_k}{\sum_{k=1}^K H_k} \quad t = 1, 2, 3, \dots, m$$

5. Results and Discussion

5.1 Metric

Evaluated Performance of the proposed PNN with PSO algorithm utilizes the following metrics: accuracy, specificity and sensitivity. These measures are defined using True Negative (TN), True Positive (TP), False Positive (FP) and False Negative (FN).

1. Accuracy- It refers to the total number of records that are correctly classified by the classifier

$$\text{Accuracy} = \frac{(TP + TN)}{N} \quad (9)$$

2. Sensitivity- It is the proportion of people who are identified as having disease

$$\text{Sensitivity} = \frac{TP}{(TP + FN)} \quad (10)$$

3. Specificity- It is the proportion of healthy people who have no diseases.

$$\text{Specificity} = \frac{TN}{(FP + TN)} \quad (11)$$

4. Time complexity- It quantifies the amount of time taken by the algorithm. This is illustrated in Figure 3.

5.2 Results

The database is collected from European Stat Log project¹¹ that contains analysis the performances of machine learning, EM and PNN algorithms on information sets from medical data. There are 13 attributes are used such as the age which identifies the age of the person, chest pain type has 4 values, serum cholesterol level, fasting blood sugar, resting blood sugar, resting ECG results have three dissimilar values, serum cholesterol level, maximum amount of heart rate achieved, x old peak, number of major vessels colored by fluoroscopy, slope of the peak exercise ST segment, thal and sex signifies the gender.

The STATLOG attribute information as shown in Table 2. From these attributes, proposed system focused on Heart rate, sex, age, blood sugar, BP, cholesterol, and additional factor smoking. Figure 3 shows that the time complexity of different types of algorithm which results show that proposed approach obtained the promising results compared to other two algorithms such as regression tree and PSO optimization.

Table 2. STATLOG Information

No	Attribute
1	Age
2	Sex
3	chest pain type (4 values)
4	resting blood pressure
5	serum cholesterol in mg/dl
6	fasting blood sugar > 120 mg/dl
7	resting electrocardiographic results (values 0,1,2)
8	maximum heart rate achieved
9	exercise induced angina
10	old peak = ST depression induced by exercise relative to rest
11	the slope of the peak exercise ST segment
12	number of major vessels (0-3) colored by fluoroscopy
13	thal: 3 = normal; 6 = fixed defect; 7 = reversible defect

Figure 4 gives classification results in terms of specificity, accuracy, and sensitivity for three various systems such as regression tree, PSO optimization, and proposed PNN with PSO. Overall training and testing was very good for almost all network models. But the proposed model performs better in all noise conditions when compared with other two models.

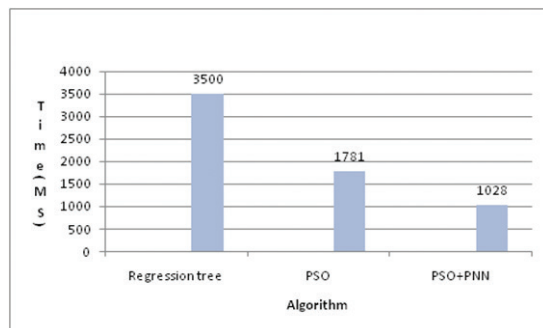


Figure 3. Time Complexity.

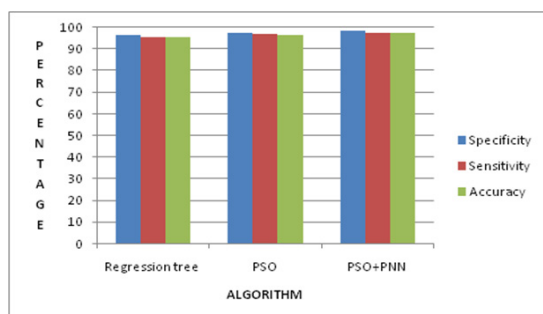


Figure 4. Testing Accuracy, Sensitivity and Specificity comparison.

6. Conclusion

Heart Risk (HR) prediction is an important and complicated task that essentials to be executed efficiently and accurately diagnosing the heart problem based on doctor's knowledge and experience. The proposed Remote monitoring system emphasized employing Probabilistic Neural Network (PNN) trained with Particle Swarm Optimization (PSO) for the prediction of heart failure. The experimental result is carried out using Java in Stat log Heart dataset. The simulation results show that Probabilistic Neural Network (PNN) outperformed the PSO optimization and regression tree algorithm based on time complexity, accuracy, specificity and sensitivity.

7. References

1. Nasira GM, Radhimeenakshi S. A study on prediction of cardiovascular victimization data processing techniques. International Journal of computer and Organization Trends. 2014 Jun; 9(1):32–5.

2. Nasira GM, Radhimeenakshi S. Heart failure prediction using particle swarm optimization in case of remote health monitoring. International Journal of Advanced Research in Computer Science and Software Engineering. 2014 Jun; 4(6):486–90.
3. Nasira GM, Radhimeenakshi S. Particle swarm optimization - a review on algorithmic rule, applications and scope. International Journal of Science and Research. 2014 Jun; 3(8):1720–2.
4. Salih ASM, Abraham A. A review of ambient intelligence assisted healthcare monitoring. International Journal of Computer Information Systems and Industrial Management Applications. 2013; 5:741–50. ISSN: 2150–7988.
5. Karamzadeh K, Askarifar H, Moharrami H. Biomedical engineering via computational intelligence. International Journal of Advancements in Research and Technology. 2013; 2(7):47. ISSN: 2278–7763.
6. Bazan JG, Socha SB. Predicting the presence of serious coronary artery disease based on 24 hour Holter ECG monitoring. The Federated Conference on Computer Science and Information Systems; IEEE; 2012. p. 279–86..
7. Chaudhary S, Kaur R. Fetal ECG signal optimization on signal obtained from fecg sensor for remote areas with lower signal strength for its smooth propagation to medical databases. International Journal of Science and Research (IJSR). 2012. ISSN: 2319–7064.
8. Jatoth RK, Anoop SSVKK, Prabhu CHM. Biologically inspired evolutionary computing tools for the extraction of fetal electrocardiogram. WSEAS Transactions On Signal Processing. 2009; 5(3):106–15. ISSN:1790–5052.
9. Sundar C, Chitradevi M, Geetharamani G. Incapable of identifying suspicious recored in CTG data using ANN based machine learning techniques. Journal of scientific and Industrial Research. 2014; 73:510–6.
10. Marco LYD, Chiari L. A wavelet-based ECG delineation algorithm for 32-bit integer online processing. BioMedical Engineering OnLine [Internet]. 2011 [updated 3 Apr 2011]; 10:23.
11. Available from: <http://archive.ics.uci.edu/ml/datasets/Statlog+Heart%29>
12. Nagarajan S, Chandrasekaran RM. Design and implementation of expert clinical system for diagnosing diabetes using data mining techniques. Indian Journal of science and Technology. 2015 Apr; 8(8):771–6.