

Risk Prediction of Heart Disease Based on Swarm Optimized Neural Network

R. Chitra and V. Seenivasagam

Abstract Heart disease (HD) remains the biggest cause of deaths worldwide. This shows the importance of HD prediction at the early stage. In this paper, multi-layer feedforward neural network (MLFFNN) optimized with particle swarm optimization (PSO) is adopted for HD prediction at the early stage using the patient's medical record. The network parameters considered for optimization are the number of hidden neurons, momentum factor, and learning rate. The efficiency of the PSO optimized neural network (PSONN) is calculated using the records collected from standard Cleveland database and Real time clinical dataset. The results show the proposed system can predict the likelihood of HD patients in a more efficient and accurate way.

Keywords Particle swarm optimization • Heart disease prediction • Momentum factor • PSO optimized neural network

1 Introduction

Cardiovascular diseases remain the biggest cause of deaths worldwide over the last two decades. The diagnosis of heart disease (HD) in most cases depends on a complex combination of clinical and pathological data. An intelligent HD prediction system built with the aid of data mining technique like decision trees, naïve

R. Chitra (✉)

Department of Computer Science and Engineering, Noorul Islam Centre for Higher Education, Kanyakumari, India
e-mail: jesi_chit@yahoo.co.in

V. Seenivasagam

Department of Computer Science and Engineering, National Engineering College, Kovilpatti, India
e-mail: yespee1094@yahoo.com

Bayes, and artificial neural network was proposed by Palaniappan and Awang [1]. The result illustrated the peculiar strength of each of the methodologies in comprehending the objectives of the specified mining objectives. Srinivas et al. [2] applied data mining techniques to predict heart attack. Based on the calculated significant weightage, the frequent patterns having value greater than a predefined threshold were chosen for the valuable prediction of heart attack.

Yao and Liu [3] proposed evolutionary feedforward neural network with neuro-genetic approach for optimization of a trading agent. Carlos Ordonez used association rules to improve HD prediction. Association rules were applied on a real dataset, contacting medical records of patient with HD, and the risk factors were identified. Sulaiman et al. [4] proposed hybrid multi-layer feedforward neural network (MLFFNN) technique for predicting the output from a grid-connected photovoltaic system. The classification of biomedical dataset was done with hybrid genetic optimized back-propagation network by Kuok et al. [5] and Karegowda et al. [6] also proposed genetic optimized back-propagation network for classification. According to Song and Gu [7], PSO algorithm has been promised to solve many optimization problems. Lee et al. [8] have used PSO and GA for excess return evaluation in stock market. Based on their experiment, it is proven that PSO algorithm is better compared to GA. In this paper, an intelligent HD prediction system is proposed using an optimized feed forward neural network. In order to improve the performance of the feed forward neural network, its parameters are tuned using PSO algorithm, and it is discussed in the rest of the paper.

2 Materials and Methodology

This section describes the dataset and proposed methodology for HD prediction. The inputs are normalized; min-max normalization techniques are used for normalization, and the normalized attributes are given as the input to the PSNN.

2.1 Heart Disease Dataset

HD dataset is obtained from UCI (University of California, Irvine CA) centre for machine learning and intelligent systems. The data collected from 270 patients are used for proposed work, and it contains 13 risk factors. HD risk factors never occur in isolation, but they are correlated with each other. The digitized data have 150 normal and 120 abnormal cases. In the selected dataset, class 0 specifies the absence of heart attack and class 1 specifies the presence of HD. Normally, direct support clinical decision-making is the intention behind the design of a clinical decision support system, and it presents patient-specific assessments or recommendations produced using the characteristics of individual patients to clinicians for consideration. The 13 attributes considered are age, gender, blood pressure,

cholesterol, fasting blood sugar, chest pain type, maximum heart rate, ST segment slope in ECG, number of major vessels colored in angiogram, and thallium test value.

2.2 Multi-Layer Feed Forward Neural Network

Artificial neural networks (ANN) can be used to tackle the said problem of prediction in medical dataset involving multiple inputs. A MLFFNN is a three-layer network with input unit $x_i = \{x_1, x_2, \dots, x_n\}$, hidden layer $h_j = \{h_1, h_2, \dots, h_n\}$, and output layer $y_k = \{y_1, y_2, \dots, y_n\}$. A MLFFNN consists of a layer of input units, one or more layers of hidden units, and one output layer of units. Each connection between nodes has a weight associated with it. In addition, there is a special weight (w) given as $w_{ij} = \{w_1, w_2, \dots, w_n\}$, where n denotes the number of neurons in the hidden layer, which feeds into every node at the hidden layer and a special weight (z) given as $z_{jk} = \{z_1, z_2, \dots, z_n\}$ that feeds into every node at the output layer. These weights are called the bias and set the thresholding values for the nodes. Each hidden node calculates the weighted sum of its inputs and applies a thresholding function to determine the output of the hidden node. The thresholding function applied at the hidden node is a sigmoid activation function. The back-propagation (BP) algorithm is a commonly used learning algorithm for training ANN. The network is first initialized by setting up all its weights to be small random numbers between $[0, +1]$. The 13 inputs are applied, and the output is calculated. The actual output (t) obtained is compared with the target output (y), and the error E is calculated. This error is then used mathematically to change the weights in such a way that the error will get minimized. The process is repeated again and again until the error is minimal. In the proposed work, mean square error function defined in Eq. (1) is used for training.

$$E = \frac{1}{n} \sum_{i=1}^n (y_i - t_i)^2 \quad (1)$$

The weights are updated using Eq. (2)

$$\Delta w_{ij}(t+1) = -\eta \frac{\partial E}{\partial w_{ij}} + \alpha \Delta w_{ij}(t) \quad (2)$$

In Eq. (2), \hat{l}_{\pm} is the momentum factor and \hat{l} is the learning rate.

The momentum factor \hat{l}_{\pm} allows for momentum in weight adjustment. The magnitude of the persistence is controlled by the momentum factor. Momentum in the BP algorithm can be helpful in speeding in speeding the convergence and avoiding local minima. Learning rate \hat{l} is a training parameter that controls the size of weight and bias changes when learning the selection of a learning rate is of critical importance in finding the true global minimum of the error. Increasing number of hidden

Table 1 Optimizing parameters selected in FFNN

Parameter	Range
Number of hidden neurons in hidden layer	6â€“25
Momentum factor ($\hat{f}\pm$)	0.9â€“0.931
Learning rate ($\hat{f}\cdot$)	0.01â€“0.114

neurons increases processing power and system flexibility. But the complexity and cost of the system also increase depending on the number of hidden neurons. Range of optimizing parameters used is given in Table 1.

2.3 Particle Swarm Optimization

Particle swarm optimization (PSO) is a population-based stochastic optimization technique developed by Eberhart and Kennedy [9]. PSO does not require gradient information of the objective function under consideration. PSO can reach the global optimum value with less iteration. In PSO, each particle in the population has a velocity $v_i(t)$, which enables it to fly through the problem space. Therefore, each particle is represented by a position $x_i(t)$ and a velocity vector. Dimensions of position and velocity vectors are defined by the number of decision variables in the optimization problem. Modification of the position of a particle is performed by using its previous position information and its current velocity.

$$v_i(t + 1) = wv_i(t) + c_1 \text{rand}_1 + c_2 \text{rand}_2(\text{Gbest}_i - x_i(t)) \tag{3}$$

$$x_i(t + 1) = x_i(t) + v_i(t + 1) \tag{4}$$

where $v_i(t)$ is velocity of particle i at iteration t , $x_i(t)$ is current position of the particle i at iteration t , Pbest_i is personal best of particle i , Gbest_i is best position in the neighborhood, rand is random number between 0 and 1, w is weighting function, c_1 is cognition learning rate, and c_2 is social learning rate

The PSONN was mainly developed to search for the optimal training parameters, i.e., the number of neurons in the hidden layer, the learning rate, the momentum rate, the transfer function in the hidden layer, and the learning algorithm. These training parameters are also known as the decision variables for the optimization task. The objective function for the optimization process was to minimize the MSE during training. PSO is chosen and applied in feed forward neural network to enhance the learning process in terms of convergence rate and classification accuracy.

3 Heart Disease Prediction System Using PSO

The input data to PSO need to be normalized as NN only work with data represented by numbers in the range between 0.001 and 0.999. Three-layer PSO was used in this study for model calibration. The number of input neurons depends on the input attributes, and there is only one neuron in the output layer. In the proposed work, 13 input neurons and one output neuron for prediction of HD are used. The optimal training parameters considered are the number of neurons in the hidden layer, momentum rate and the learning rate, the transfer function in the hidden layer, and the learning algorithm.

The PSO algorithm for intelligent heart attack prediction is given below

```

For each particle
  Initialize particle for NN problem
End
Do
  For each particle
    Calculate fitness value (feedforward error or MSE
in If the fitness value is better than the best fitness value
(pBest) in history
  Then set current value as the new pBestEnd
Choose the particle with the best fitness value of all the
particles as the gBest
For each particle
  Calculate particle velocity
  Update particle position
End

```

The PSO parameters and c_1 , c_2 are the acceleration constants which are initially set to 2, and r_1 and r_2 are random integers set to a range between [0, 1]. The initial population is randomly selected, and the population size is set to 25. The total number of iterations is chosen as 200. With these parameters, the number of hidden layer neurons, momentum rate and the learning rate, the transfer function in the hidden layer, and the learning algorithm are optimized. The whole process is repeated until the generation reaches 200 or the minimum fitness is achieved.

4 Experimental Results

In this work, we analyze the use of the PSO algorithm and the cooperative variant with the weight decay mechanism for neural network training, aiming better generalization performances in HD prediction. For evaluating these algorithms, we apply them to benchmark classification problems of the medical field. The training set and testing set accuracy are calculated and analyzed for the standard HD dataset. In the standard dataset, 270 data are used for training and the network is

also tested with same data. The performance metrics accuracy, sensitivity, and specificity of training and testing cases of both dataset are given in Table 2. The performance metrics obtained in the proposed system for standard dataset are compared with SVM classifier and feed forward neural network classifier and are shown in Fig. 1.

ROC curve is a graphical plot created by plotting false positive rate versus true negative rate, and it is used to analyze the performance of the classifier. The ROC curve for SVM, feed forward neural network, and PSO optimized neural network (PSONN) is shown in Fig. 2. From Fig. 1, it is observed that the PSONN has significantly good performance compared with all other classifiers.

The ROC curve depends on the true positive rate and false positive rate. From this curve, it is shown that the performance of the PSONN-based prediction system has shown a significant improvement on classifying the diseased and non-diseased patterns. When analyzing these graphs, the PSONN shows a better performance in terms of accuracy, sensitivity, and specificity. A true positive rate of 0.889 is obtained with a false positive rate of 0.0987 which is achieved by the proposed classifier. The sensitivity and specificity achieved by the PSONN are 88.98 and 90.13 for standard dataset with an overall accuracy of 89.6, which is higher than the other two well-known classifiers. This increase is because of the fact that the neural network structure is optimized with the PSO algorithm. The BP algorithm primarily depends on the initial parameter setting, and the performance of the classifier depends on these parameters. From Fig. 1, it is noticed that FFNN has an improved performance than the SVM; this is due to the fact that SVM does not depend on the

Table 2 Training and testing performance metrics

Performance metric	PSONN	
	Training set	Testing set
Accuracy	88.4	90.8
Sensitivity	87.54	90.42
Specificity	89.62	90.64

Fig. 1 Performance analysis for Cleveland dataset

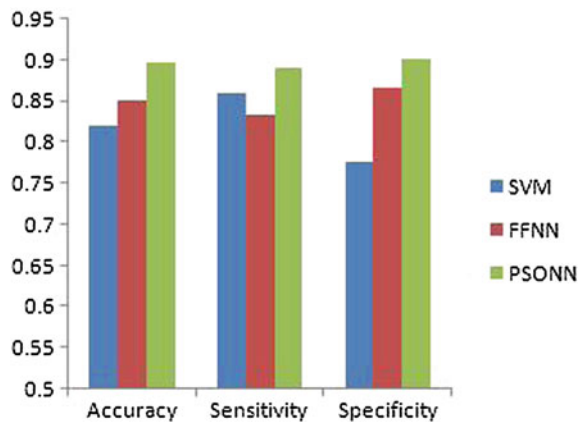
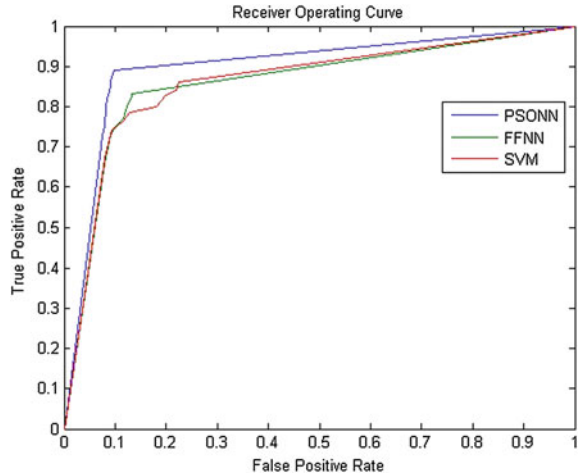


Fig. 2 ROC plot for different classifiers



generalization error. Although FFNN produces a good classification accuracy for prediction of HDs, this classifier selects the initial parameter setting manually, which will take a considerable time to classify the patterns. Hence, FFNN and SVM will not be suitable for the physicians to analyze and predict the diseased and non-diseased patterns. The results of the experiments show that an improvement of 4 % accuracy has been achieved by applying optimized neural network classifiers. This implies that the optimized neural network is one of the desirable classifiers, which can be used as an aid for the physicians to predict the HDs.

5 Conclusion

In this paper, we have proposed an adaptive intelligent mechanism for HD prediction using the profiles collected from the patients. PSONN adopted global and local optimization of the network parameters within the specified range. By exploiting this distinct feature of the PSONN, a computerized prediction algorithm is developed that is not only accurate but is also computationally efficient for heart attack prediction. With the proper optimization using PSO, the method can thus evolve an optimum number of hidden units within an architecture space. The results of proposed system have achieved better accuracy than most of the existing algorithms.

References

1. Palaniappan, S., Awang, R.: Intelligent heart disease prediction system using data mining techniques. In: IEEE, pp. 108–115 (2008)
2. Srinivas, K., Rani, B.K., Govrdhan, A.: Applications of data mining techniques in healthcare and prediction of heart attacks. *Int. J. Comput. Sci. Eng.* **2**(2), 250–255 (2010)

3. Yao, X., Liu, Y.: Towards designing artificial neural networks by evolution. *Appl. Math. Comput.* **91**, 83–90 (1998)
4. Sulaiman, S.I., Rahman, T.K.A., Musirin, I.: A genetic algorithm-based hybrid multi-layer feedforward NeuNetwork for predicting grid-connected photovoltaic system output. In: *IACSIT Hong Kong Conferences, IPCSIT*, vol. 25 (2012)
5. Kuok, K.K., Harun, S., Shamsuddin, S.M.: Particle swarm optimization feedforward neural network for hourly rainfall runoff modeling in Bedup Basin, Malaysia. *Int. J. Civ. Environ. Eng.* **9**(10) 2010
6. Karegowda, A.G., Manjunath, A.S., Jayaram, M.A.: Application of genetic algorithm optimized neural network connection weights for medical Diagnosis of pima Indians diabetes. *Int. J. Soft Comput.* **2**(2), 15–23 (2011)
7. Song, M.P., Gu, G.C.: Research on particle swarm optimization: a review. In: *IEEE*, pp. 2236–2241 (2004)
8. Lee, J.S., Lee, S., Chang, S. Ahn, B.H.: A Comparison of GA and PSO for Excess Return Evaluation in Stock Markets, pp. 221–230. Springer, Berlin (2005)
9. Kennedy, J., Eberhart, R.C.: Particle swarm optimization. In: *Proceedings of IEEE International Conference on Neural Networks*, vol. 4, pp. 1942–1948 (1995)