Three-Term Backpropagation Network Based On Elitist Multiobjective Genetic Algorithm for Medical Diseases Diagnosis Classification

Ashraf Osman Ibrahim[1, 2], Siti Mariyam Shamsuddin[1], Nor Bahiah Ahmad[1], Sultan Noman Qasem[3]

[1.] Soft Computing Research Group, Faculty of Computing, Universiti Teknologi Malaysia, 81310 Skudai, Johor, Malaysia.
[2.] Faculty of computer and technology, Alzaiem Alazhari University, Khartoum North 13311, Sudan.
[3.] College of Computer and Information Sciences, Al-Imam Muhammad Ibn Saud Islamic University, Riyadh, Saudi Arabia.
ashrafosman2@gmail.com

Abstract: Recently, the problems related to intelligent medical disease diagnosis classification have become one of the important areas of study. Therefore, this paper proposes a new intelligent classifier approach, by using the Three-Term Backpropagation (TBP) network based on the Elitist Multiobjective Genetic Algorithm (MOGA). One of the recent MOGAs is a Non-dominated Sorting Genetic Algorithm II (NSGA-II), which is used to reduce or optimize the error rate and network structure of TBP simultaneously to achieve more accurate classification results. In addition accuracy, sensitivity, specificity and 10-fold cross validation are used as performance evaluation indicators to evaluate the outcome of the proposed method. The proposed intelligent methodology is applied in four kinds of standard medical diseases datasets, obtained from the University of California at Irvine (UCI) repository. The results illustrate that our approach is viable in medical diseases diagnosis classification when compared with some other methods found in the literature.

[Ashraf Osman Ibrahim, Siti Mariyam Shamsuddin, Nor Bahiah Ahmad, Sultan Noman Qasem. **Three-Term Backpropagation Network Based On Elitist Multiobjective Genetic Algorithm for Medical Diseases Diagnosis Classification.** *Life Sci J* 2013;10(4):1815-1822]. (ISSN:1097-8135). http://www.lifesciencesite.com. 238

Keywords: Artificial Neural Networks; Medical Diseases Diagnosis; Multiobjective Evolutionary; Three Term Backpropagation; NSGA-II.

1. Introduction

Artificial Neural Network (ANN) is an architecture that uses mathematical models. ANN methods have become important and widely used in medical computations, particularly when dealing with huge and high dimensional medical data. However, the use of a computational method in medicine has been growing gradually.

Currently, medical computational systems and mathematical models have become the substrate of medical technology, which are used to diagnose the type of diseases in medical data. For this purpose, numerous mathematical and computational methods have been used to support the medical disease diagnosis classification[1, 2]. On the other hand, intelligence techniques are commonly used in medical disease diagnosis classification. Specifically, ANNs are one of the most commonly used classifiers due to their high ability in prediction and adaptability. But despite that, there is still more work needed to design and develop the ANNs classifier for the medical disease diagnosis model.

The design and development of a classifier method involves approximating an unknown input and the output mapping function from available data. Moreover, the classifier can be used to predict the class labels that correspond to unseen data. Hence, the objective of developing a good classifier is to ensure high prediction accuracy of the specific problem, especially in important and sensitive data such as medical diagnosis. One of the solutions for designing a good structure of ANNs for classification tasks is an Evolutionary Algorithm (EAs). They are good candidates for Multiobjective Optimization Problems (MOOPs) because of their abilities to search for multiple Pareto optimal solutions and they perform better in global search space. Many optimization problems in the world involve numerous incompatible objectives. As an alternative to dealing with a single optimal solution, a set of optimal solutions called a Pareto optimal set exists for such problems. The corresponding objective functions, whose non-dominated solutions are in the Pareto optimal set, are called a Pareto front. Each of the Pareto optimal solutions signifies a different balance between the objectives and in the absence of preference information; none of them can be supposed to be better than others. In addition, Pareto optimal solutions are used to evolve artificial neural networks (ANNs) which are optimal both with respect to classification accuracy and architecture complexity [3]. These EAs are population-based algorithms, which allow for simultaneous exploration of different parts in the Pareto optimal set. There are

many studies that have used the Pareto optimal notion in medical disease diagnosis classification using multiobjective optimization techniques [3-5]. Moreover, the research that used soft computing and artificial intelligence methods in the medical area has attracted a lot of attention.

Evolutionary Algorithms (EAs) are increasing used for the optimization problem. One of the most successful applications of the EAs was used for evolving ANNs, as in [6] who provided a general framework for using EAs for evolving ANNs; his method employed GAs for optimizing ANNs. Furthermore, Multiobjective evolutionary Algorithms MOEAs research is one of the hottest areas in the field of evolutionary computation [7]. So, there are various methods that have used MOOP to solve medical classification problems and other problems. These include [3] who introduced hybrid learning of RBF network with the multiobjective particle swarm optimization MOPSO to improve classification accuracy for medical disease diagnosis problems. [4] applied MPANN based on a Pareto optimal solution to breast cancer, through which positive results were obtained. [5] applied a Multiobjective Evolutionary Neural Networks method to solve medical data classification and multiclass problems called MPENSGA2E and MPENSGA2S, by using a multilayer perceptron neural network hybrid with the NSGA2 algorithm. Also, [8] used hybrid method multiobjective evolutionary and artificial neural networks, based on a micro-hybrid genetic algorithm for medical classification data and other data. [9] used a hybrid data mining approach integrated statistical method and Discrete PSO to classify a breast cancer dataset. [10] proposed an intelligent liver diagnosis model (ILDM) using ANN to examine the patients that suffer from liver disease and also to determine the types of the liver disease. [11] presented a medical decision support system based on the multilayer perceptron (MLP) neural network architecture for heart disease diagnosis. They used an improved BP to train their model. [12] presents a multiobjective GA using Pareto-optima optimization of the ANN for classification of the breast cancer diagnosis problem. [13] introduced a hybrid model using GA and BP networks for the diagnosis of Pima Indians' diabetes; they used GA to optimize the network connection weights.

As an instance, multiobjective genetic algorithm optimization was used by [14] for training a feed forward neural network, number of nodes, the architecture, as well as the weights, and a Pareto front was effectively constructed by minimizing the training error and the network size using noisy data. Also, a general framework using GA for designing neural network ensembles was presented in [15]. Another method used by generalized multi-layer perceptrons (MLP) improved the performance of the evolutionary model [16]. The authors in [17] proposed a hybrid MOGA method based on the SPEA2 and NSGA2 algorithms to optimize the training and the topology of the Recurrent Neural Network (RNN) simultaneously in time-series prediction problems. In addition, [18] studied the benefits of hybridizing Pareto differential evolution with the BP as a local search algorithm for a training method to speed up convergence and long training time. Thoroughly, the optimization of the structure is carried out by minimizing the number of network connections, even though numerous studies offered reasonable solutions for feed-forward ANNs. Also, [19] introduced a multiobjective evolutionary learning algorithm using an improved version of the NSGA2 algorithm called MPENSGA2 hybridized with a local search algorithm for training ANNs with generalized radial basis functions.

However, in this paper we proposed an intelligent medical disease diagnosis classifier. Multiobjective Evolutionary Algorithms (MOEAs) are applied to improve the generalization of the training and unseen data in the network. MOEAs are suitable to produce and design the appropriate and accurate ANNs with the optimization of two or more conflict objectives simultaneously. Therefore, NSGA-II is applied to optimize two objectives simultaneously: the number of hidden nodes in the hidden layer and errors of the network, to solve the medical disease diagnosis classification problem. Our intelligent classifier is compared with many algorithms in the literature.

The rest of this paper is organized as follows: Section 2 provides a description of Preliminaries related to this study. Section 3 presents the proposed MOGATBP network. Results and discussion reported in section 4. Finally, concludes the paper.

2. Preliminaries

This section provides a brief explanation of Multiobjective Evolutionary Algorithms, Multiobjective learning problem, Three Term Backpropagation Algorithm and NSGA-II Algorithm along with some of the key basic concepts.

2.1 Multiobjective Evolutionary Algorithms

A multiobjective evolutionary algorithm (MOEAs), also known as multiobjective optimization algorithms (MOOAs), is the process of simultaneously optimizing two or more conflicting objectives subject to certain constraints; they are a population based search. Hence, in a single run it can get many of Pareto optimal sets (solutions) and that are attractive of this kind of algorithms. Pareto

optimal solutions are used to evolve artificial neural networks (ANNs) which are optimal both with respect to architecture complexity and classification accuracy. A multiobjective optimization problem (MOOP) can be defined as follows:

Definition: MOOP contains a set of *n* decision variables, a set of *k* objective functions, and a set of *m* constraints. Objective functions and constraints are functions of the decision variables. The optimization goal is to Maximize:

$$y = f(x) = (f_1(x), f_2(x), ..., f_K(x))$$

Subject to:

$$e(x) = (e_1(x), e_2(x), ..., e_m(x)) \leq 0$$

Where,

$$x = (x_1, x_2, ..., x_n) \in X$$

$$y = (y_1, y_2, ..., y_k) \in Y$$

And *x* is the decision vector, *y* is the objective vector, *X* is indicated in the decision space, and *Y* is the objective space. The constraints $e(x) \leq 0$ determine the set of feasible solutions.

The recent researches focusing on the application of multiobjective evolutionary algorithms to solve multi objective optimization problems in different fields [19-22]. To evolve ANNs there are various methods and techniques that have been developed to identify better approaches [17], by attempting to design networks with good generalization capability. However, developing a good ANN architecture has also been discussed in the ANNs researches. The main advantages of the evolutionary approach to ANN training are its ability to escape a local optimum, its robustness and its ability to adapt itself to a changing environment [19-21].

2.2. Multiobjective learning problem

Multiobjective learning problem usually has to achieve many objectives simultaneously, which are often conflicting with each other. In ANNs, minimizing the network complexity and the maximizing network capacity are conflicting objectives. In addition, network selection has to deal with the trade-off between network complexity and classification accuracy. Therefore, multiobjective learning problem in this paper can be formulated as two objective functions which are used to evaluate the TBP performance for all algorithms as follows:

1. The performance of the network (Accuracy) is based on the Mean Square Error (MSE) on the training set, this performance as a first objective function is given in equation 1:

$$f_1 = \frac{1}{N}\sum_{j=1}^{N}(t_j - o_j)^2 \qquad (1)$$

Where, o_j Network error at output unit, t_j target value of output, *N* number of samples.

2. The complexity of the network is based on the number of hidden nodes in the hidden layer of TBP, as a second objective function as in equation 2:

$$f_2 = \sum_{h=1}^{H}\rho_h \qquad (2)$$

Where, $\rho_h \in \rho$, vector ρ is the dimension of maximum number of hidden nodes H of the network, and ρ is binary value used to refer to the hidden node if it exists in the network or not. It works as a switch to turn a hidden unit ON or OFF and is the maximum hidden nodes of TBP.

2.3. Three Term Backpropagation Algorithm (TBP)

The Three Term Backpropagation proposed by Zweiri in [23] employs the standard architecture and procedure of the standard backpropagation algorithm. However, in addition to learning rate and momentum parameters, the third parameter, called proportional factor (PF), is introduced. This is proven to be successful in improving the convergence rate of the algorithm and speeding up the weight adjusting process.

2.4. NSGA-II Algorithm

The genetic algorithm (GA) is based on simulating the biological evolution of the search space in the searching process automatically and it is a parallel global search method [24]. The non-dominated sorting genetic algorithm-II (NSGA-II) was proposed by [25], as it has a good performance in global searching. A non-dominated sorting multiobjective optimization genetic algorithm becomes a preferred method of optimization algorithm. It proposes a new method and a new arithmetic operator by improving the first version of the NSGA [21]: the fast non-dominated sorting approach and the crowded comparison operator. So far, there are many studies on optimization and design that have been conducted [19, 26-29]. And all these studies prove that the genetic algorithm and its upgraded derivatives are feasible for optimal design.

NSGA-II algorithm beginning by generates random population of chromosomes or solutions of size N. Referring to Figure 1, first both the parent population and offspring population are combined to form a combined population of size 2N instead of

finding the non-dominated fronts of offspring population only. Then, the non-dominated sorting procedure is performed on the entire. This procedure allows a global non-domination check between the offspring and parent solutions, and improves NSGA-II to converge faster.

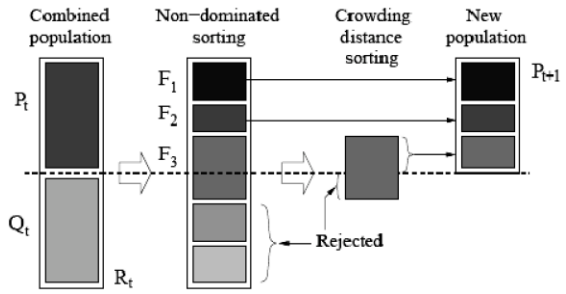

Figure 1. Schematic of NSGA-II algorithm.

The crowding distance is used in the selection of parents for a new individual and the selection of a new population based on comparison of the congestion around a solution. A greater crowding distance is preferred in order to maintain the diversity of the solutions.

3. The Proposed MOGATBP Network

The non-dominated sorting genetic algorithm (NSGA-II) is based on [25] of the TBP network, which is implemented for solving medical disease diagnosis classification problems. The non-dominated sorting genetic algorithm based TBP network (MOGATBP) is implemented. The network architecture and accuracy are evolved simultaneously with each individual being a fully specified TBP network. In this study, MOGATBP network has been proposed to determine the best performance and the corresponding architecture of the TBP network. To assist TBP design, GA and MOO are combined as a rank-density based GA to carry out fitness evaluation and mating selection schemes. Similarly, MOGATBP network begins by collecting, normalizing and reading the dataset, dividing the data set into training data and testing data. Then the number of hidden nodes and maximum number of iterations is set. Also, the individual length is computed. Furthermore, the parameters of TBP network are determined by the traditional algorithms. Then there is the generation and initialization of a population of TBP network. Every individual is evaluated for every iteration based on objective functions. After the maximum iterations are reached the proposed method stops and outputs a set of non-dominated TBP networks.

3.1. Parameter Setting

The NSGA-II is used for training the TBP network for all datasets with the same parameters. The population size is 100 in the progress, crossover rate used is 0.90. The mutation rate is 1/N, where "N" refers to the dimension of individual; while the maximum number of iterations is 1000. The fitness values are the hidden nodes and network training error or performance of the network. The training set is used to train the TBP network in order to obtain the Pareto optimal solutions; while the testing set is used to test the generalization performance of Pareto TBP network.

3.2. Dataset Description

For the experimental design we consider four datasets from the medical field (disease diagnosis classification problem) listed in Table 1, all data sets obtained from the UCI repository [30]. The datasets that are used in this study are partitioned into two sets: training and testing data. In addition, all the dataset values are normalized in the range of [0,1].

The "Breast Cancer Wisconsin" [31] data set has 699 patterns; 458 (65.5%) of the patterns in the datasets are benign, while 241 (34.5%) of the patterns are malignant. There are nine attributes/inputs (clump thickness, uniformity of cell size and shape, marginal adhesion, single epithelial cell size, bare nuclei, bland chromatin, normal nucleoli and mitoses) and two output classes (benign or malignant).

The diabetes problem is to diagnose a Pima Indian individual based on personal data and medical examination. There are eight attributes/inputs (no. of times pregnant, plasma glucose concentration, diastolic blood pressure, triceps skin fold thickness, serum insulin, Body Mass Index (BMI), Diabetes pedigree function and age) and two output classes (diabetes positive or diabetes negative).

The hepatitis problem is a complex and noisy data as it contains a large number of missing data (there are 167 missing values in total in this dataset). The learning task is to predict whether a patient with hepatitis will live or die. There are nineteen attributes/inputs (age, sex, steroid, antivirals, fatigue, malaise, anorexia, liver big, liver film, spleen palpable, spiders, ascites, varices, bilirubin, alk. phosphate, SGOT, albumin, protime and histology) and two output classes (live or die).

The purpose of the heart dataset is to predict the presence or absence of heart diseases given the results of various medical tests carried out on a patient. It contains 303 patterns of which 139 are positive instances and 164 are negative instances. There are thirteen attributes/inputs, which have been extracted from a larger set of 75 and two output classes (heart positive or heart negative).

Table 1. Summary of data sets used in the experiments

Dataset	features	classes	Instances
Wisconsin breast cancer	9	2	699
Pima Indians Diabetes	8	2	768
Heart	13	2	279
Hepatitis	19	2	155

4. Results and Discussions

According to Table 1, the experiments are conducted by using four medical datasets applied to test the efficiency of the proposed classifier.

The NSGA-II is used for training the TBP network for all datasets with same parameter values that were mentioned earlier.

Sensitivity is the measure of the classifier's ability to identify the correct positive samples, and it depends on the number of true positives and false negatives as in Equation 3. While the Specificity is a measure of the classifier to predict the correct negative samples; this depends on the number of true negatives and false positives (see Equation 4). Also, the accuracy is a measure of the classifier's ability to produce the level of accurate diagnosis; Equation 5 shows the accuracy formula.

All datasets used in this study are two output classes. Therefore, the sensitivity and specificity are shown in Equations 3 and 4. True positive (TP) means that diagnosis is correctly classified as sick and false positive (FP) means diagnosis is incorrectly classified as sick. While, true negative (TN) means that diagnosis is correctly classified as healthy; also the false negative (FN) means the diagnosis is incorrectly classified as healthy. The classification of the diagnosis used in the datasets is benign or malignant in breast cancer, for diabetes it is positive or negative, live or die from hepatitis and for the heart is a positive or negative heart.

$$\text{Sensitivity} = \frac{\text{True Positive (TP)}}{\text{True Positive (TP)} + \text{False Negative (FN)}} \% \quad (3)$$

$$\text{Specificity} = \frac{\text{True Negative (TN)}}{\text{True Negative (TN)} + \text{False Positive (FP)}} \% \quad (4)$$

$$\text{Accuracy} = \frac{\text{True Positive (TP)} + \text{True Negative (TN)}}{\text{True Positive (TP)} + \text{True Negative (TN)} + \text{False Positive (FP)} + \text{False Negative (FN)}} \% \quad (5)$$

Table 2.The average and standard deviations of training and testing accuracy.

Dataset		Training Set			Testing Set		
		Sensitivity (%)	Specificity (%)	Classification (%)	Sensitivity (%)	Specificity (%)	Classification (%)
Wisconsin breast cancer	Mean	98.47	97.55	97.65	96.01	97.08	96.97
	STD	0.62	0.11	0.389	2.53	1.843	1.09
Pima Indians Diabetes	Mean	49.88	88.95	76.65	46.73	88.00	74.99
	STD	4.47	2.80	1.40	11.72	4.52	4.40
Heart	Mean	83.78	83.61	83.69	79.57	83.13	82.83
	STD	2.21	3.23	1.19	9.18	5.15	4.03
Hepatitis	Mean	17.24	98.56	81.79	20.00	98.33	81.02
	STD	18.61	1.86	2.61	28.11	3.51	3.87

From Table 2 we clearly notice the statistical results for sensitivity, specificity and accuracy of the proposed method; the average and standard deviation are shown for 10-fold runs on the datasets. All the results are a Pareto optimal solution to improve the generalization of the network. In terms of accuracy for all datasets the accuracy rates are very good, especially in Wisconsin breast cancer and the Heart dataset. As we can see in Table 2, they achieved 96.97% and 82.83% respectively. While for the others, Pima Indians' Diabetes and Hepatitis, the accuracy rate is acceptable with 74.99% and 81.02% respectively.

For the sensitivity, Wisconsin breast cancer achieves the highest sensitivity rate 96.01%, followed by the Heart dataset 79.57%. For the Hepatitis, Pima Indians' Diabetes datasets, the improvement in sensitivity are very difficult to interpret, due to the difficult classification problems of these datasets, as they are very unbalanced. So, this means that these datasets have lower sensitivity. For the specificity, all of the used datasets achieved high specificity rates. Furthermore, Figure 2 shows the comparison of statistical results obtained for sensitivity, specificity and the accuracy of the proposed method.

The results in Table 3 demonstrate the performance of the proposed method (Training error and testing error) for the used datasets. Error rates for all results are shown in Table 3. The results show the generalization error of the proposed method. From Table 3, we can observe that in all datasets on the mean rows, the proposed method is giving promising results in performance (training and testing error), especially the Wisconsin breast cancer data, which has a lowest error. Furthermore, the training and testing error that is shown in the same table is the average of the errors obtained in a single run of the MOGA to TBP and they are reasonable error values. Moreover, Figure 3 shows the comparison of all errors obtained in the training and testing set using the proposed method. From the same figure axis Y plots the MSE, while axis X plots the data sets used in this study, we can see that the yeast dataset has a lower error rate than other datasets, while the others are quite similar.

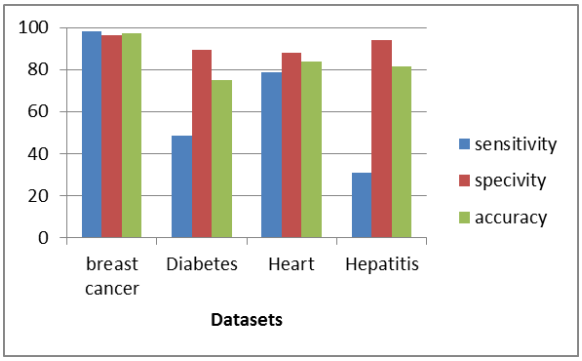

Figure 2. The average of accuracy, sensitivity and specificity of the datasets

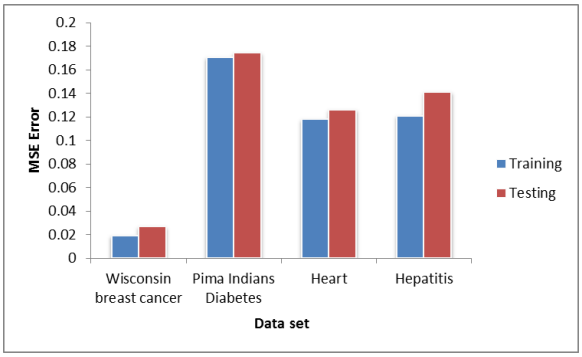

Figure 3. The MSE average error of the datasets

Table 3. The performance of the proposed method

Dataset		Training Error	Testing Error	Hidden Nodes
Wisconsin breast cancer	Mean	0.0189	0.0266	4.7
	STD	0.0009	0.0152	1.6
Pima Indians Diabetes	Mean	0.1705	0.1741	5.6
	STD	0.0140	0.0130	2.7
Heart	Mean	0.1180	0.1260	4.6
	STD	0.0052	0.0363	1.2
Hepatitis	Mean	0.1210	0.1410	5.1
	STD	0.0120	0.0327	2.6

Table 4. Classification accuracies of the proposed method and some methods in literatures

Method/Reference	Breast cancer	Diabetes	Heart	Hepatitis
MOGATBP	96.97	74.99	82.84	81.02
RBFN-TVMOPSO [3]	96.53	78.02	-	82.26
RBFN-MOPSO [3]	90.95	70.58	-	82.32
RBFN-NSGA-II [3]	87.30	69.59	-	83.78
C4.5 [32]	94.71	74.21	80.74	79.25
MPANN [4]	98.10	74.90	-	-
HMOEN L2 [8]	96.26	78.45	79.69	80.30
HMOEN HN [8]	96.82	75.36	81.06	75.51
MPENSGA2E [5]	95.87	78.99	-	-
MPENSGA2S [5]	95.60	76.96	-	-

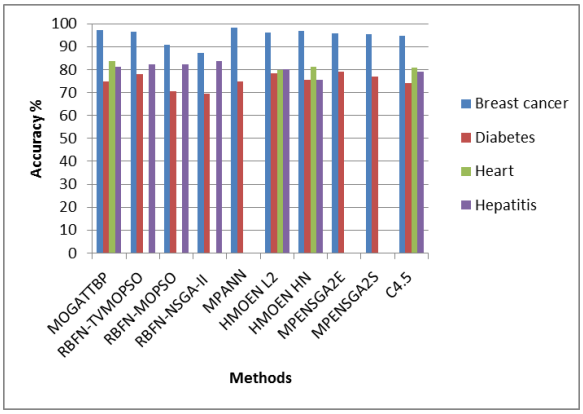

Figure 4. The classification accuracies of the proposed method and some methods in literatures

The performance of the proposed method compares with other methods in the literature using the same data sets. These include some of multiobjective evolutionary ANN algorithms such as (RBFN-NSGA-II [3], RBFN-TVMOPSO [3], RBFN-MOPSO [3], MPANN [4], HMOEN L2 [8], HMOEN HN [8], MPENSGA2E [5], MPENSGA2S [5]) and C4.5 [32]. Table 4 and Figure 4 show the summary of the results. Our method is the best classifier of all methods reported in Table 4 in the Heart dataset. Also, in the Breast cancer data it is the best method of all methods except MPANN [4]. For the Diabetes and Hepatitis our method showed acceptable and modest results, being neither worse nor better than others. From Table 4, all methods that we compared with our method, are using local search approaches, except one algorithm which is C4.5 [32]. While our method using multiobjective evolutionary algorithm without local search algorithm. As will known, the local search process improves a solution and achieves better accuracy of the final result. In this case local search algorithms improve all individuals in the population. Thus enabled these methods to improve their algorithms to achieve good results.

Conclusion

In this paper a new intelligent classifier by using hybrid Three Term BP based on a Multiobjective Genetic Algorithm was introduced and successfully applied for the medical disease diagnosis classification problem; it is effective as a classifier with good performance and a high accuracy rate. The results indicate that our proposed method demonstrated effectiveness in dealing with the medical disease diagnosis classification problems. Also, NSGA-II was used to optimize the MSE error and hidden nodes to develop a simple and accurate TBP network. The advantages of our method are that it is simple, and easy to implement and use. For future work, we will enhance the proposed method to achieve more accuracy and robustness in the results by using one of the local search algorithms to enhance all individuals in the population; this will be a good option to improve the performance of TBP network.

Acknowledgements:

The authors would like to thank Soft Computing Research Group (SCRG) (UTM) and IDF for their continuous support and motivation. This work is supported by (UTM) under GUP -Tier2 Research Grant Scheme (Q.J130000.2628.09J07).

Corresponding Author:
Ashraf Osman Ibrahim,
Soft Computing Research Group,
Faculty of Computing,
Universiti Teknologi Malaysia (UTM), Skudai, 81310, Johor, Malaysia.
E-mail: ashrafosman2@gmail.com

References

1. Fan, C.-Y., et al., A hybrid model combining case-based reasoning and fuzzy decision tree for medical data classification. Applied Soft Computing, 2011. 11(1): p. 632-644.
2. Talhaa, M. and A. Al-Elaiwi, Enhancement and Classification of Mammographic Images for Breast Cancer Diagnosis using Statistical Algorithms. Life Science Journal, 2013. 10(2).
3. Qasem, S.N. and S.M. Shamsuddin, Radial basis function network based on time variant multi-objective particle swarm optimization for medical diseases diagnosis. Applied Soft Computing, 2011. 11(1): p. 1427-1438.
4. Abbass, H.A., An evolutionary artificial neural networks approach for breast cancer diagnosis. Artificial Intelligence in Medicine, 2002. 25(3): p. 265-281.
5. Fernandez Caballero, J.C., et al., Sensitivity versus accuracy in multiclass problems using memetic Pareto evolutionary neural networks. Neural Networks, IEEE Transactions on, 2010. 21(5): p. 750-770.
6. Yao, X., Evolving artificial neural networks. Proceedings of the IEEE, 1999. 87(9): p. 1423-1447.
7. Zhou, A., et al., Multiobjective evolutionary algorithms: A survey of the state of the art. Swarm and Evolutionary Computation, 2011. 1(1): p. 32-49.
8. Goh, C.-K., E.-J. Teoh, and K.C. Tan, Hybrid multiobjective evolutionary design for artificial neural networks. Neural Networks, IEEE Transactions on, 2008. 19(9): p. 1531-1548.
9. Yeh, W.-C., W.-W. Chang, and Y.Y. Chung, A new hybrid approach for mining breast cancer

pattern using discrete particle swarm optimization and statistical method. Expert Systems with Applications, 2009. 36(4): p. 8204-8211.
10. Lin, R.-H. and C.-L. Chuang, A hybrid diagnosis model for determining the types of the liver disease. computers in Biology and Medicine, 2010. 40(7): p. 665.
11. Yan, H., et al., A multilayer perceptron-based medical decision support system for heart disease diagnosis. Expert Systems with Applications, 2006. 30(2): p. 272-281.
12. Ahmad, F., et al., A genetic algorithm-based multi-objective optimization of an artificial neural network classifier for breast cancer diagnosis. Neural Computing and Applications, 2012: p. 1-9.
13. Karegowda, A.G., A. Manjunath, and M. Jayaram, Application of genetic algorithm optimized neural network connection weights for medical diagnosis of pima Indians diabetes. International Journal on Soft Computing, 2011. 2(2): p. 15-23.
14. Pettersson, F., N. Chakraborti, and H. Saxén, A genetic algorithms based multi-objective neural net applied to noisy blast furnace data. Applied Soft Computing, 2007. 7(1): p. 387-397.
15. García-Pedrajas, N., C. Hervás-Martínez, and D. Ortiz-Boyer, Cooperative coevolution of artificial neural network ensembles for pattern classification. Evolutionary Computation, IEEE Transactions on, 2005. 9(3): p. 271-302.
16. Garcıa-Pedrajas, N., D. Ortiz-Boyer, and C. Hervás-Martınez, Cooperative coevolution of generalized multi-layer perceptrons. Neurocomputing, 2004. 56: p. 257-283.
17. Delgado, M., M.P. Cuéllar, and M.C. Pegalajar, Multiobjective hybrid optimization and training of recurrent neural networks. Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on, 2008. 38(2): p. 381-403.
18. Abbass, H.A., Speeding up backpropagation using multiobjective evolutionary algorithms. Neural Computation, 2003. 15(11): p. 2705-2726.
19. Cruz-Ramírez, M., et al., Multi-Objective Evolutionary Algorithm for Donor-Recipient Decision System in Liver Transplants. European Journal of Operational Research, 2012.
20. Schaffer, J.D. Multiple objective optimization with vector evaluated genetic algorithms. in Proceedings of the 1st international Conference on Genetic Algorithms. 1985. L. Erlbaum Associates Inc.
21. Srinivas, N. and K. Deb, Muiltiobjective optimization using nondominated sorting in genetic algorithms. Evolutionary computation, 1994. 2(3): p. 221-248.
22. Fonseca, C.M. and P.J. Fleming, Multiobjective optimization and multiple constraint handling with evolutionary algorithms. II. Application example. Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on, 1998. 28(1): p. 38-47.
23. Zweiri, Y., J. Whidborne, and L. Seneviratne, A three-term backpropagation algorithm. Neurocomputing, 2003. 50: p. 305-318.
24. Deb, K. An introduction to genetic algorithms. in Sadhana (Academy Proceedings in Engineering Sciences). 1999. Indian Academy of Sciences.
25. Deb, K., et al., A fast and elitist multiobjective genetic algorithm: NSGA-II. Evolutionary Computation, IEEE Transactions on, 2002. 6(2): p. 182-197.
26. Ak, R., et al., NSGA-II-Trained Neural Network Approach to the Estimation of Prediction Intervals of Scale Deposition Rate in Oil & Gas Equipment. Expert Systems with Applications, 2012.
27. Ramesh, S., S. Kannan, and S. Baskar, Application of modified NSGA-II algorithm to multi-objective reactive power planning. Applied Soft Computing, 2011.
28. Qasem, S.N., S.M. Shamsuddin, and A.M. Zain, Multi-objective hybrid evolutionary algorithms for radial basis function neural network design. Knowledge-Based Systems, 2011.
29. Alhamdy, S.A.S., et al., A neural network-based multi-objective genetic algorithm for designing cells incrementally in a dynamic environment. Life Science Journal, 2013. 10(2).
30. A. Asuncion and D.J. Newman, UCI Machine Learning Repository. http://www.ics.uci.edu/$\sim$mlearn/{MLR}epository.html, 2007.
31. Wolberg, W.H. and O.L. Mangasarian, Multisurface method of pattern separation for medical diagnosis applied to breast cytology. Proceedings of the national academy of sciences, 1990. 87(23): p. 9193-9196.
32. Antonie, M.-L., O.R. Zaiane, and R.C. Holte. Learning to use a learned model: A two-stage approach to classification. in Data Mining, 2006. ICDM'06. Sixth International Conference on. 2006. IEEE.

7/1/2013