

# **Memetic Elitist Pareto Evolutionary Algorithm of Three-Term Backpropagation Network for Classification Problems**

Ashraf Osman Ibrahim<sup>1</sup>, Shafaatunnur Hasan<sup>1</sup>, and Sultan Noman<sup>2</sup>

<sup>1</sup>UTM Big Data Centre, Universiti Teknologi Malaysia, Skudai, Johor, Malaysia  
e-mail: ashrafosman2@gmail.com; shafaatunnur@utm.my

<sup>2</sup>Computer Science Department, College of Computer and Information Sciences,  
Al Imam Mohammad Ibn Saud Islamic University (IMSIU), Riyadh, Saudi Arabia

## **Abstract**

*Evolutionary Algorithms (EAs) are population based algorithms, which allow for simultaneous exploration of different parts in the Pareto optimal set. This paper presents Memetic Elitist Pareto Evolutionary Algorithm of Three-Term Backpropagation Network for Classification Problems. This memetic elitist Pareto evolutionary algorithm is called METPB and used to evolve Three-term Backpropagation (TBP) network, which are optimal with respect to connection weight, error rates and architecture complexity simultaneously. METPB is based on NSGA-II benefit from the local search algorithm that used to enhance the individuals in the population of the algorithm. The numerical results of METPB show the advantages of the combination of the local search algorithm, and it is able to obtain a TBP network with better classification accuracy and simpler structure when compared with a multiobjective genetic algorithm based TBP network (MOGATBP) and some methods found in the literature, the results indicate that the proposed method is a potentially useful classifier for enhancing classification process ability.*

**Keywords:** Artificial Neural Network, Hybridization, Genetic algorithm, NSGA-II, Multiobjective optimization.

## **1 Introduction**

In recent years, Artificial Neural Network (ANN) has become the substrate of soft computing methods, which are used for solving different problems successfully. ANN is the computing model that mimics the way neuron system of human brain works. For this reason, numerous techniques and methods have been used to

support the computational systems. Specifically, ANNs are most commonly used classifiers due to their high ability in prediction and adaptability [1]. However, more intensive works are needed to design and develop the ANNs classifier for the classification problems. Meanwhile, the use of a computational method has been growing gradually. Nowadays, continuous research efforts to employ soft computing techniques for the classification problems have been in focus since last decade.

Evolutionary Algorithm (EAs), are good candidates for Multi objective optimization problems (MOOPs) [2]. This is because of their abilities to search for multiple Pareto optimal solutions and they perform better in global search space. Multiobjective evolutionary algorithms (MOEAs) research area has become one of the hottest areas in the field of evolutionary computation [3]. They are suitable to produce and design the appropriate and accurate ANNs with the optimization of two conflicting objectives, namely; minimization of ANNs structure complexity and maximization of network capacity. Hence, recently MOEAs have been successfully applied to optimize both the structure, connection weights and network training simultaneously [4-6].

The proposal presented in this paper is applied Hybrid pareto optimal algorithm for optimizing TBP network to improve the generalization of the training and unseen data. In the proposed method, NSGA-II hybrid with local search algorithm was applied to optimize three objectives which are: connection weights, error rates and complexity of the network simultaneously, to solve pattern classification problems. The local search algorithm used to enhance all individuals in the population and increase the quality of the Pareto optimal solutions.

The rest of the paper is organized as follows: In Section 2, the related works. In Section 3, presents background materials. We describe the proposed method in Section 4 and the experimental Study, experimental settings, data set besides obtained results and discussion shown in in Section 5. Finally, in Section 6 the conclusions are drawn.

## 2 Related Works

Several studies have used the Pareto optimal concept in classification problems using multiobjective optimization techniques [7-9]. Recently, multiobjective evolutionary algorithms (MOEAs) have been used to produce and optimize the ANNs parameters with the optimization of two or more conflicting objectives. These kind of algorithms is applied to improve the generalization of the training and unseen data in the network. Moreover, MOEAs are convenient to produce and design the appropriate and accurate ANNs with the optimization of two or more conflicting objectives simultaneously. Therefore, MOEAs have been applied successfully to optimize the network structure, connection weights and train the network, due to their ability to improve structural performance.

The work in [10] provided a general framework employing GAs for using EAs for evolving ANNs, consider as a one of the most successful applications of the MOEAs used for evolving ANNs. Study introduced by [9] implemented in which a Memetic Pareto evolutionary Neural Networks technique was used to solve two class and multiclass classification problem called MPENSGA2E and MPENSGA2S, by using a multilayer perceptron neural network hybrid with the NSGA2 algorithm. Similarly, a study given by [4] used hybrid multiobjective evolutionary method and artificial neural networks based on a micro-hybrid genetic algorithm for classification of medical data and other data. A multiobjective GA using Pareto optimal optimization of the ANN for classification of the breast cancer diagnosis problems presented by [11]. In addition, hybrid model using Genetic algorithm (GA) and Backpropagation (BP) networks for the diagnosis of diabetes diseases used GA to optimize the network connection weights which were introduced by [12]. As an instance, multiobjective genetic algorithm optimization was used by [13] 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 [14]. The authors in [5] 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. Another approach presented by [15] used to generalized multi-layer perceptrons (MLP) improved the performance of the evolutionary model.

Hybridization is an important feature in the area of EAs which has received relative attention over the past few years. Hybrid techniques are used to enhance the performance of ANNs. Therefore, many studies in the literature focused on the hybrid algorithms that combines ANNs with other techniques such as PSO [16], GA[17], DE[18] and Artificial Bee Colony Algorithm [19]. As such, hybrid global and local search algorithms have been one of the new research areas, known as Memetic Algorithm (MA). Due to the phenomenal success of this kind of algorithms, MAs, have been successfully used in different application to solve a lot of problems. One of the most successful domains to apply MAs is multiobjective optimization problems [3, 20]. In contrast, several previous studies have presented prospective achievements by using ANNs, MOEAs and local optimizers to speed up the convergence [21-23]. In addition, [21] studied the advantages of hybridizing Pareto differential evolution with the BP algorithm as a local search algorithm for a training method to speed up convergence and long training time. Furthermore, one of the most famous works in this area in [8] concludes that his approach memetic Pareto artificial neural network (MPANN), which is based on a Pareto optimal solution, has better generalization and positive results were obtained. For example, [24] introduced a multiobjective evolutionary learning algorithm using an improved version of the NSGA-II algorithm called MPE NSGA-II hybridized with a local search algorithm for training ANNs with

generalized radial basis functions. This current work has developed ANN by optimizing connection weights, structure of the network and error rates simultaneously, based on memetic elitist pareto evolutionary algorithm of TBP network for solving pattern classification problems.

## **3 Background Materials**

### **3.1 Three-Term Backpropagation Algorithm (TBP)**

The Three Term Backpropagation proposed by Zweiri in [25] 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.

### **3.2 NSGA-II**

The non-dominated sorting genetic algorithm-II (NSGA-II) proposed in [26, 27], it is upgraded version of NSGA [28]. Moreover, NSGA-II is a fast and elitist multiobjective genetic algorithm, which was used to obtain the set of Pareto optimal front solutions. For it's a good performance of 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: the fast non-dominated sorting approach and the crowded comparison operator. It is well known that NSGA-II is one of the most famous Pareto optimal solution algorithms as it requires the simultaneous minimization or maximization of two or more objective functions. So far, there are many works about optimization and design which have been done [24, 29-31]. And all these studies demonstrated that the genetic algorithm and its upgraded derivatives are feasible for optimal design.

NSGA-II algorithm beginning by generating a random population of chromosomes or solutions of size  $N$ . Firstly, 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 the offspring population only. Subsequently, 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.

### **3.3 Local Search Algorithm**

Local search algorithms are widely used for several problems in different areas, but it has received more attention in computer science and engineering, particularly artificial intelligence applications. It is known that the local methods

are able to find the local optimum when searching in a small area of space. Hybrid global and local search algorithms have been one of the new research areas, this type of hybrid algorithm is known as Memetic algorithms (MAs). Due to the phenomenal success of this kind of algorithms, MAs, have been successfully used in different application to solve a lot of problems. MAs are able to provide not only the best speed of convergence to the evolutionary approach, but also the best accuracy for the final solutions [32]. In this study, we have used a classical BP algorithm as a local search method.

## **4 The proposed METBP method**

The proposed method is adapted the non-dominated sorting genetic algorithm (NSGA-II) hybrid with Backpropagation (BP) algorithm as a local search algorithm to enhance all individuals in the population; this process will be a good option to improve the performance of the network. The hybrid non-dominated sorting genetic algorithm based TBP network is implemented. The proposed algorithm was evolved the network architecture and accuracy simultaneously with each individual being a fully specified TBP network. This algorithm has been proposed to determine the best performance and the corresponding architecture of the TBP network.

However, the proposed method begins with the first step which is collected, normalizing and reading the dataset, followed by, dividing the data set into training data and testing data. Then the minimum, maximum 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 are the generation and initialization of a population of the NSGA-II. 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. To evaluate the TBP network performance of the proposed method, three objective functions were used in this study. The first was fitness function which is the performance of the network (Accuracy) based on the Mean Square Error (MSE) on the training set, while the second fitness function is the Complexity of the network based on the number of hidden nodes in the hidden layer of the TBP network.

## **5 Experimental Study**

### **5.1 Experimental settings**

In this section, experiments are conducted to validate the effectiveness of the proposed method. The details that required to adapting and adjusting to find the optimal parameters of the NSGA-II combinations are determined by depending to the previous studies in the literature that were used and applied NSGA-II algorithm [6, 33] . The proposed method is used for training the TBP network

based on hybrid NSGA-II for all datasets with the same parameters. The population size of NSGA-II is set to 100, crossover rate used is 0.90 and mutation rate is  $1/N$ , where “N” refers to the dimension of individual; while the maximum number of iterations is 1000. Regarding to the local search algorithm which is the BP algorithm, we set the learning rate to 0.01.

## 5.2 Data set

This section presents the experimental of the study on hybrid NSGA-II for TBP network. For the experimental design, we considered 11 different data set were used to validate the proposed algorithm. It contains binary class and multiclass data set. The breast cancer, heart, hepatitis, liver and diabetes datasets are example for the binary class. While, iris, lung cancer, QAC, segment, wine, and yeast data sets are multi class data for classification problems. The data sets were used in this study are obtained from the UCI machine learning repository [34]. Furthermore, for the preprocessing all the dataset values are normalized in the range of [0,1].

## 5.3 Results and Discussion

As can be seen in Table 1, the training and testing error rates results show the generalization error of the proposed METBP and our previous method MOGATBP [35]. Moreover, Table 1 reflects the promising results in performance (training and testing error) in all datasets. Additionally, the training and testing error are the average of the errors obtained in a single run of the hybrid NSGA-II for TBP network and they are reasonable error values.

Table 1: Comparison of training and testing error on all data

Data Set	METBP		MOGATBP		
	Training Error	Testing Error	Training Error	Testing Error	
Breast Cancer	Mean	0.0185	0.0235	0.0186	0.0241
	SD	0.0011	0.0090	0.0016	0.0082
Diabetes	Mean	0.1655	0.1701	0.1696	0.1719
	SD	0.0070	0.0153	0.0152	0.0157
Heart	Mean	0.1148	0.1199	0.1192	0.1219
	SD	0.0050	0.0320	0.0042	0.0255
Hepatitis	Mean	0.1280	0.1354	0.1280	0.1311
	SD	0.0170	0.0290	0.0080	0.0251
Iris	Mean	0.1158	0.1179	0.1180	0.1196
	SD	0.0139	0.0148	0.0110	0.0114
Liver	Mean	0.2258	0.2336	0.2167	0.2212

Lung Cancer	SD	0.0128	0.0151	0.0094	0.0128
	Mean	0.1866	0.1977	0.1866	0.1987
QAC	SD	0.0130	0.0280	0.0130	0.0291
	Mean	0.1209	0.1205	0.1172	0.1181
Segment	SD	0.0035	0.0036	0.0074	0.0066
	Mean	0.11817	0.1244	0.1287	0.1309
Wine	SD	0.0057	0.0071	0.0116	0.0144
	Mean	0.1537	0.1531	0.1555	0.1563
Yeast	SD	0.0262	0.0214	0.0301	0.0242
	Mean	0.0816	0.0821	0.0816	0.0816
	SD	0.0088	0.0104	0.0088	0.0088



Fig 1: Comparison average error obtained by training and testing on all data

Table 2: Comparison of Hidden nodes on all data

Data set		METBP	MOGATBP
Breast Cancer	Mean	4.10	4.70
	SD	1.10	1.64
Diabetes	Mean	5.10	5.60
	SD	2.33	2.72
Heart	Mean	4.40	4.60
	SD	0.84	1.26
Hepatitis	Mean	4.90	5.20
	SD	1.85	2.53
Iris	Mean	4.60	5.20

	SD	1.51	2.82
Liver	Mean	5.10	4.60
	SD	2.08	1.70
Lung Cancer	Mean	2.00	2.00
	SD	0.00	0.00
QAC	Mean	3.90	4.10
	SD	0.99	1.20
Segment	Mean	4.50	4.80
	SD	0.85	1.14
Wine	Mean	4.50	4.90
	SD	1.72	2.18
Yeast	Mean	3.60	3.50
	SD	1.07	0.85

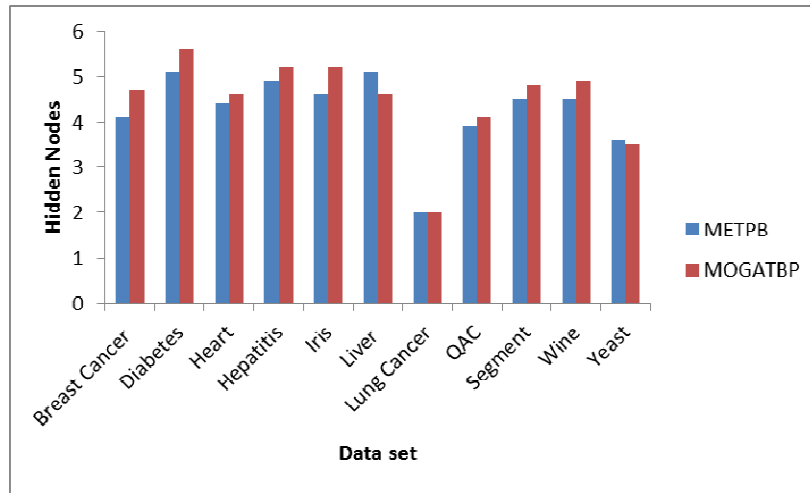


Fig 2: Comparison averages of hidden nodes of the METPB and MOGA on all data

Table 3: Comparison of Training and Testing accuracy on all data

Dataset		METBP		MOGATBP	
		Training	Testing	Training	Testing
Breast Cancer	Mean	98.00	97.07	97.65	96.69
	SD	0.22	1.19	0.39	1.01
Diabetes	Mean	75.65	74.57	75.93	73.98
	SD	1.29	3.91	2.13	3.46



Heart	Mean	84.32	83.83	83.46	82.78
	SD	0.60	6.18	1.02	3.81
Hepatitis	Mean	82.24	81.18	81.79	80.82
	SD	3.394	5.23	2.61	3.73
Iris	Mean	82.84	82.63	79.71	79.26
	SD	3.923	3.953	5.14	5.80
Liver	Mean	62.80	61.46	62.01	61.35
	SD	5.18	5.17	4.18	4.40
Lung Cancer	Mean	69.78	69.26	69.78	69.26
	SD	8.84	9.18	8.83	9.17
QAC	Mean	85.71	85.71	85.79	85.78
	SD	0.00	0.00	0.24	0.20
Segment	Mean	85.64	85.78	85.71	85.71
	SD	0.24	0.21	0.00	0.00
Wine	Mean	77.64	76.86	74.60	73.74
	SD	8.45	7.22	2.21	1.34
Yeast	Mean	90.00	90.02	90.01	90.01
	SD	0.00	0.05	0.03	0.03

The evaluation measures of the results, which are: accuracy in Table 3, sensitivity and specificity in Table 4, which were obtained using 10-fold cross validation by the proposed algorithm, for training and testing data. We can see the calculation of those evaluation measures in Equations 1- 3 as follows:

$$Sensitivity = \frac{TP}{TP + FN} \quad (1)$$

$$Specificity = \frac{TN}{TN + FP} \quad (2)$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (3)$$

Where, TP is true positive, FP is false positive, TN is true negative and FN is false negative.

The results of the training and testing accuracy are shown by details in Table 4. As well known, if the training set is performed effectively and accurately, the result of the final result of classification should be accurate. In general, it has generated a high the training and testing classification result, some data have a high classification accuracy such as, breast cancer, yeast and wine. Comparing with MOGATBP from same table, the proposed method outperforms the MOGATBP in training accuracy of all data set, except in QAC data the MOGATBP is obtained better result.

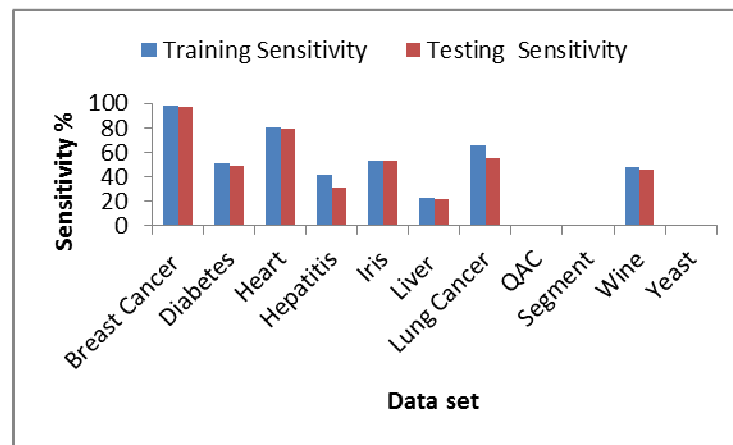


Fig 3: Average the training and testing sensitivity of the METPB of all data set

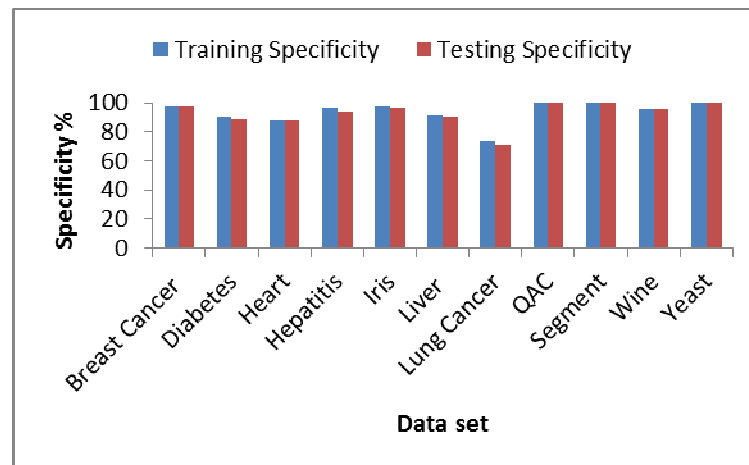


Fig 4: Average the training and testing specificity of the METPB on all data set.

Table 4: Comparison of Sensitivity and Specificity for Training and Testing data

Data Set	METBP				MOGATBP				
	Training		Testing		Training		Testing		
	sensitivity	Specificity	sensitivity	Specificity	Sensitivity	Specificity	Sensitivity	Specificity	
Breast Cancer	Mean	98.74	97.60	96.67	97.30	98.47	97.55	97.43	98.43
	SD	0.49	0.13	2.64	1.77	0.62	0.11	2.10	1.83
Diabetes	Mean	51.28	90.24	48.43	89.20	50.18	89.14	48.10	88.72
	SD	4.21	1.88	10.92	4.34	4.22	2.39	12.00	4.71
Heart	Mean	80.13	87.92	78.85	88.13	84.01	84.81	79.90	83.35
	SD	1.88	2.00	7.94	8.04	2.04	1.37	7.95	4.32
Hepatitis	Mean	41.76	96.57	30.83	94.23	27.24	97.48	30.00	96.67
	SD	19.46	2.20	25.47	6.84	15.37	2.28	29.19	4.30
Iris	Mean	53.11	97.70	52.67	97.00	42.49	99.14	40.00	99.01
	SD	14.65	2.57	15.54	5.76	14.31	1.42	9.94	1.69
Liver	Mean	22.99	91.67	21.38	90.50	22.99	91.67	21.38	90.50
	SD	20.80	6.84	19.95	9.85	20.80	6.84	19.95	9.85
Lung Cancer	Mean	66.06	73.55	55.00	71.11	66.06	73.55	55.00	71.11
	SD	15.65	9.37	27.27	14.30	15.65	9.37	27.27	14.30
QAC	Mean	0.00	100.00	0.00	100.00	2.87	99.64	3.93	99.21
	SD	0.00	0.00	0.00	0.00	6.20	0.80	12.42	1.68
Segment	Mean	0.00	100.00	0.00	100.00	1.24	99.84	1.10	99.65
	SD	0.00	0.00	0.00	0.00	2.74	0.38	2.35	0.70
Wine	Mean	48.28	95.80	45.96	96.15	48.28	98.15	46.73	98.70
	SD	25.75	6.40	22.12	6.00	25.75	2.09	23.34	1.86
Yeast	Mean	0.00	100.00	0.00	100.00	0.00	100.00	0.00	100.00
	SD	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

For the sensitivity and Specificity, Table 4, the proposed method has achieved 52.6667% for iris, 54.9008% for wine and 0.00 % for yeast dataset. The sensitivity of the QAC, Segment and yeast data set is very difficult, due to their unbalanced data. The Same Table is also shows the specificity for all datasets, we can note that the specificity rate was achieved good result by the proposed method.

Table 5: Comparison of the hidden nodes of TBP network obtained by the METPB and other methods

Data set	Methods			
	METBP	MOGATBP	MEPGANf1f2	MEPGANf1-f3
Breast Cancer	4.10	4.70	6.60	5.40
Diabetes	5.10	5.60	5.40	5.80
Iris	4.60	5.20	5.50	5.60
Heart	4.40	4.60	6.20	5.60
Hepatitis	4.90	5.20	5.10	5.80
Liver	5.10	4.60	6.60	4.60
QAC	3.90	4.10	9.10	4.50
Yeast	3.60	3.50	8.20	6.90
Lung Cancer	2.00	2.00	2.00	4.60
Segment	4.50	4.80	10.00	10.00
Wine	4.50	4.90	6.5	6.00

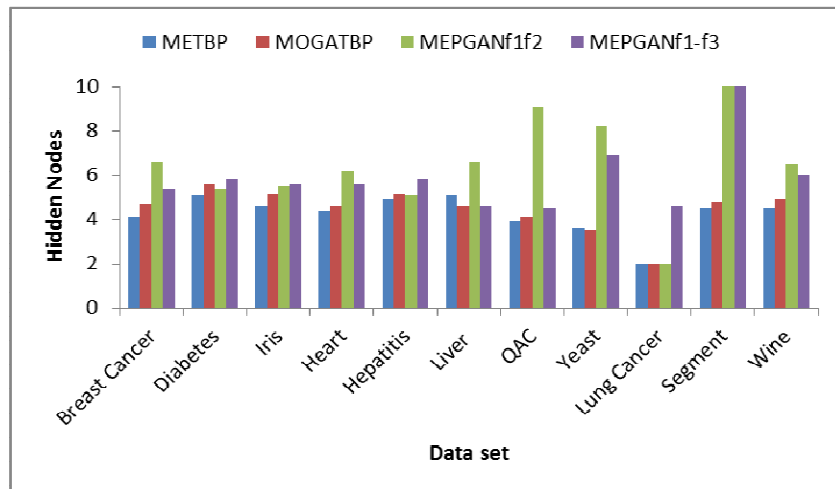


Fig 5: Comparison of the hidden nodes of the METPB and other methods of all data

Table 6: Comparison of the testing accuracy of the METPB and other methods

Data set	Methods								
	METBP	MOGA TBP	MEPGAN f1f2	MEPGAN f1-f3	SVM	HMOE N L2	HMOE N HN	MPEN SGA2E	MPENS GA2S
Breast Cancer	97.07	96.69	96.78	97.80	96.49	96.26	96.82	95.87	95.60
Diabetes	74.57	73.98	72.78	68.35	65.10	78.48	75.36	78.99	76.96
Iris	82.63	79.26	83.78	84.44	96.67	98.00	91.03	97.18	96.50
Heart	83.38	82.78	79.07	80.79	54.88	79.69	81.06	-	-
Hepatitis	81.18	80.82	80.04	79.38	79.36	80.30	75.51	-	-
Liver	61.46	61.35	62.63	63.50	59.42	68.00	68.94	-	-
QAC	85.71	85.78	85.71	85.71	80.95	-	-	85.56	85.40
Yeast	90.02	90.01	90.00	90.01	43.26	-	-	59.91	53.21
Lung Cancer	69.26	69.26	66.67	66.67	50.00	-	-	-	-
Segment	85.78	85.71	86.90	86.22	65.37	-	-	-	-
Wine	76.86	73.74	72.18	72.04	44.38	-	-	-	-

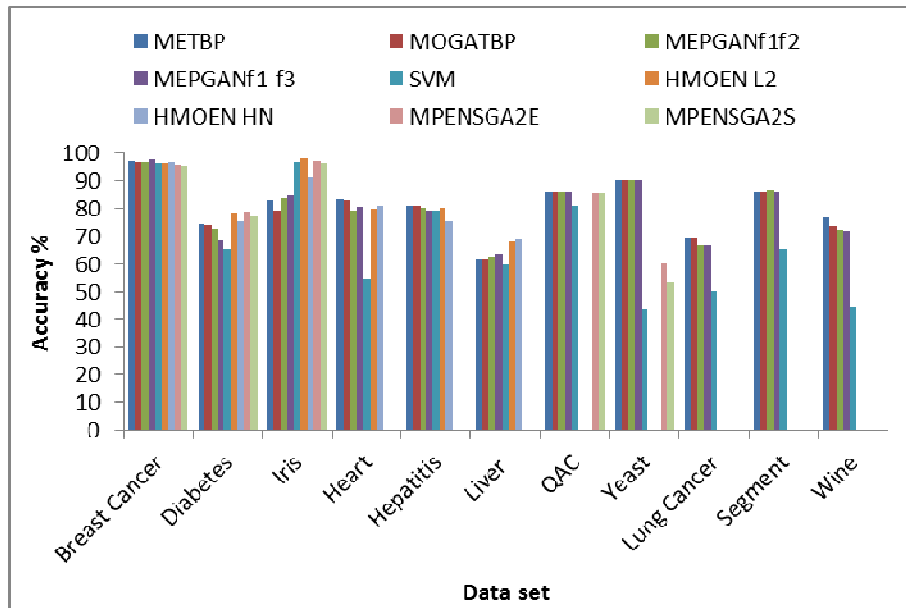


Fig 6: Comparison of the testing accuracy of the METPB and other methods of all data

An analysis of the accuracy and the number of hidden nodes are compared to MOGATBP, MEPGANf1f2 and MEPGANf1-f3 [6] was found in the literature. From Table 5 and Fig 5, it is clearly seen that the results are improved compared

to the other methods. Specially, in the iris and wine data, but for yeast data the MOGATBP is perform better hidden nodes than proposed method.

In accuracy, Table 6, show the testing accuracy results for proposed method compared to other mentioned methods. We can observe clearly the proposed method has achieved better than other methods in hart, hepatitis, yeast, lung cancer and wine data. On the other hand, the proposed method failed to outperform others in breast cancer, iris, segment, liver and QAC. Moreover, Fig. 6, show the compression of the accuracy for proposed method and other methods.

## 6 Conclusion

This paper introduced Memetic Elitist Pareto Evolutionary Algorithm, is used for training the TBP network and is proposed to achieve optimization of three objectives. A hybrid proposed method called METBP. All three objectives have been evaluated using three types of performance evaluation indicators to assess the effect of the proposed method. The proposed method was applied to solve multiclass pattern classification problem. Finally, experimental results indicating the efficiency of the METBP as a multiobjective evolutionary neural network. More precisely, the numerical results of METPB show the advantages of the combination of the local search algorithm, and it is able to obtain a TBP network with better classification accuracy and simpler structure when compared with MOGATTBP and other algorithms. In our future work, we will propose some adaptive method to improve the efficiency of METBP.

### ACKNOWLEDGEMENTS

The authors would like to thank UTM Big Data Centre and Soft Computing Research Group (SCRG) Universiti Teknologi Malaysia (UTM) for their continuous support and motivation. This work is partially supported by Flagship Gran - Predictive Analytics Framework for E-Learning Big Data Computing(Q.J130000.2428.02G38).

## 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] Zitzler, E., M. Laumanns, and S. Bleuler, *A tutorial on evolutionary multiobjective optimization*, in *Metaheuristics for multiobjective optimisation*. 2004, Springer. p. 3-37.
- [3] 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.

- [4] 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.
- [5] 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.
- [6] Qasem, S.N. and S.M. Shamsuddin, *Memetic elitist pareto differential evolution algorithm based radial basis function networks for classification problems*. Applied Soft Computing, 2011. **11**(8): p. 5565-5581.
- [7] 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.
- [8] Abbass, H.A., *An evolutionary artificial neural networks approach for breast cancer diagnosis*. Artificial Intelligence in Medicine, 2002. **25**(3): p. 265-281.
- [9] 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.
- [10] Yao, X., *Evolving artificial neural networks*. Proceedings of the IEEE, 1999. **87**(9): p. 1423-1447.
- [11] 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.
- [12] 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.
- [13] 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.
- [14] 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.
- [15] Garcia-Pedrajas, N., D. Ortiz-Boyer, and C. Hervás-Martinez, *Cooperative coevolution of generalized multi-layer perceptrons*. Neurocomputing, 2004. **56**: p. 257-283.
- [16] Zhang, C., H. Shao, and Y. Li. *Particle swarm optimisation for evolving artificial neural network*. in *Systems, Man, and Cybernetics, 2000 IEEE International Conference on*. 2000. IEEE.

- [17] Ding, S., C. Su, and J. Yu, *An optimizing BP neural network algorithm based on genetic algorithm*. Artificial Intelligence Review, 2011. **36**(2): p. 153-162.
- [18] Ilonen, J., J.-K. Kamarainen, and J. Lampinen, *Differential evolution training algorithm for feed-forward neural networks*. Neural Processing Letters, 2003. **17**(1): p. 93-105.
- [19] Qiongshuai, L. and W. Shiqing. *A hybrid model of neural network and classification in wine*. in *Computer Research and Development (ICCRD), 2011 3rd International Conference on*. 2011. IEEE.
- [20] Ishibuchi, H., et al., *Use of biased neighborhood structures in multiobjective memetic algorithms*. Soft Computing, 2009. **13**(8-9): p. 795-810.
- [21] Abbass, H.A., *Speeding up backpropagation using multiobjective evolutionary algorithms*. Neural Computation, 2003. **15**(11): p. 2705-2726.
- [22] Jin, Y., B. Sendhoff, and E. Körner, *Simultaneous generation of accurate and interpretable neural network classifiers*, in *Multi-Objective Machine Learning*. 2006, Springer. p. 291-312.
- [23] Wiegand, S., C. Igel, and U. Handmann, *Evolutionary multi-objective optimisation of neural networks for face detection*. International Journal of Computational Intelligence and Applications, 2004. **4**(03): p. 237-253.
- [24] Cruz-Ramírez, M., et al., *Multi-Objective Evolutionary Algorithm for Donor-Recipient Decision System in Liver Transplants*. European Journal of Operational Research, 2012.
- [25] Zweiri, Y., J. Whidborne, and L. Seneviratne, *A three-term backpropagation algorithm*. Neurocomputing, 2003. **50**: p. 305-318.
- [26] Deb, K., et al., *A fast and elitist multiobjective genetic algorithm: NSGA-II*. Evolutionary Computation, IEEE Transactions on, 2002. **6**(2): p. 182-197.
- [27] Deb, K., et al., *A fast elitist non-dominated sorting genetic algorithm for multi-objective optimization: NSGA-II*. Lecture notes in computer science, 2000. **1917**: p. 849-858.
- [28] Srinivas, N. and K. Deb, *Multiobjective optimization using nondominated sorting in genetic algorithms*. Evolutionary computation, 1994. **2**(3): p. 221-248.
- [29] 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.
- [30] Ramesh, S., S. Kannan, and S. Baskar, *Application of modified NSGA-II algorithm to multi-objective reactive power planning*. Applied Soft Computing, 2011.



- [31] 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.
- [32] Lara, A., et al., *HCS: A new local search strategy for memetic multiobjective evolutionary algorithms*. Evolutionary Computation, IEEE Transactions on, 2010. **14**(1): p. 112-132.
- [33] Qasem, S.N., et al., *Memetic multiobjective particle swarm optimization-based radial basis function network for classification problems*. Information Sciences, 2013. **239**: p. 165-190.
- [34] A. Asuncion and D.J. Newman, *UCI Machine Learning Repository*. <http://www.ics.uci.edu/~mllearn/MLRrepository.html>, 2007.
- [35] Ibrahim, A.O., et al., *Three-Term Backpropagation Network Based On Elitist Multiobjective Genetic Algorithm for Medical Diseases Diagnosis Classification*. Life Science Journal, 2013. **10**(4).