A HYBRID BP AND HSA FOR ENHANCING A MULTILAYER PERCEPTRON LEARNING

ABDIRASHID SALAD NUR

UNIVERSITI TEKNOLOGI MALAYSIA

A HYBRID BP AND HSA FOR ENHANCING A MULTILAYER PERCEPTRON LEARNING

ABDIRASHID SALAD NUR

A dissertation report submitted in partial fulfillment of the
requirements for the award of
the degree of Master of Science (Computer Science)

Faculty of Computing
Universiti Teknologi Malaysia

March, 2014

To my beloved Family and Friends

To my respected supervisor

ACKNOWLEDGEMENT

First and foremost, all praises are due to ALLAH (SWT) whose mercy and favours to me are countless; secondly I praise the prophet (PBUH) whose guidance will guide me in the right path and safeguards me from the hell fire. Subsequently, I wish to express my sincere appreciation to my supervisors, Prof. Dr. Siti Mariyam Shamsuddin and Dr. Nor Haizan Mohd Radzi for their valuable assistance, encouragement, beneficial condemnations, knowledge, and motivation throughout the period of this research. Without their continual support and concentration, this thesis would not have been the same as presented here. I would like to send my gratitude to all who supported me directly or indirectly in this study.

Also, I would like to send my grateful appreciations to my family for all their tireless support, extraordinary patience and encouragement they gave me. I thank to all my friends whose supports have morally encouraged me.

ABSTRACT

Training neural networks is a great significance of a difficult task in the field of supervised learning because; its performance depends on underlying training algorithm as well as the achievement of the training process. In this study, three training algorithms namely Back-Propagation algorithm, Harmony Search Algorithm (HSA) and hybrid BP and HSA called BPHSA are employed for the supervised training of MLP feed forward type of NNs by giving special attention to hybrid BPHSA. A suitable structure for data representation of NNs is implemented to BPHSA, HSA and BP. The proposed model is empirically tested and verified by using five benchmark classification problems namely Iris, Glass, Cancer, Wine and thyroid datasets on training NNs. The MSE, training time, classification accuracy of hybrid BPHSA are compared with the standard BP and meta-heuristic HSA. The experiments showed that proposed model (BPHSA) has better results in terms of convergence error and classification accuracy compared to BP and HSA and this makes the BPHSA look as promising algorithm for neural network training.

ABSTRAK

Latihan rangkaian neural (RN) adalah satu tugas yang sukar dan penting di dalam bidang pembelajaran diselia. Prestasi NN bergantung kepada algoritma latihan serta pencapaian proses latihan. Satu algoritma baru telah dibangunkan di dalam penyelidikan ini untuk memperbaiki penumpuan ralat di dalam algoritma pembelajaran Rambatan Balik (RB) dengan memperbaiki nilai pemberat neuron menggunakan algoritma Gelintaran Harmoni (GH). Algorithm hibrid ini dikenali sebagai RBGH. Satu siri eksperimen dilaksana untuk menguji, menentusahkan dan mengukur prestasi algoritma yang dibangunkan dengan algoritma piawai Rambatan Balik dan Gelintaran Harmoni. Ketiga-tiga algoritma ini digunakan untuk menyelesaikan masalah pengkelasan di dalam pembelajaran terselia rangkaian neural multiaras suapan hadapan. Set data pengkelasan yang di gunakan di dalam eksperimen ini terdiri daripada set data *Iris, Glass, Cancer, Wain* dan tiroid. Stuktur rangkaian neural yang bersesuaian dibangunkan untuk setiap set data. Nilai min ralat kuasa dua (MSE), masa latihan, kejituan pengkelasan yang dihasilkan oleh alogritma hibrid RBGH dibandingkan dengan nilai MSE yang dihasilkan oleh algoritma piawai BP dan HSA. Hasil eksperimen menunjukkan nilai penumpuan ralat, kejituan pengkelasan yang dihasilkan oleh RBGH lebih baik berbanding RB dan GH. Hasil kajian ini menunjukkan algoritma hibrid RBGH mampu meningkatkan prestasi latihan rangkaian neural.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATION

ACO	Ant Colony Optimization
ANNs	Artificial Neural Networks
BP	Back Propagation
BP-GA	Back Propagation-Genetic Algorithm
BPHSA	Back Propagation & Harmony Search Algorithm
BP-PSO	Back Propagation-Particle Swarm Optimization
EA	Evolutionary Algorithm
FFANN	Feed Forward Artificial Neural Network
FFNN	Feed Forward Neural Network
GA	Genetic Algorithm
HS	Harmony Search
HSA	Harmony Search Algorithm
IGHS	Intelligent Global Harmony Search
MLP	Multi-Layer Perceptron
MLPNN	Multi-Layer Perceptron Neural Network
MSE	Mean Square Error
NN	Neural Network
PSO	Particle Swarm Optimization
SA	Simulated Annealing
SGHS	Self-adaptive Global Harmony Search
SL	Supervised Learning
TS	Tabu Search

LIST OF APPENDICES

CHAPTER 1

OVERVIEW

1.1 Introduction

Artificial Neural Networks (ANNs) also known as Neural Networks are computational models based on the human brain microstructure which consists of billions of neurons which act as processors. NNs have been widely applied to solve a large number of complex problems (Neruda, 2007) including prediction of future events, classification, noise filtering, and pattern recognition (Narayan *et al.*, 1996).

A single - layer network is one of the easiest and simplest form of feed forward NNs. This architecture consists of input and output layers and can only classify linearly separable data (Noriega, 2005). The training method used for this architecture is known as a perceptron learning rule. The weight and biases for this perceptron are trained to generate a target vector properly.

A Multi-Layer Perceptron (MLP) is another form of feed forward NN where one or more additional layer called hidden layer is inserted between the existing input and output layers. Besides the ability to distinguish between linearly separable data, MLP could also be used to classify the nonlinear separable data (Sunila and Nirmal, 2011).

Learning is the process of training the NN before it can be used to solve a problem. Generally during this process, a weight is initially given at random to every neuron. The weight is then modified according to algorithm used to achieve the correct results. A training dataset is presented to its input to match specific output data or determine the correct outputs after being well trained and tested with similar related data. This process lasts until the desired or a near optimal result (output) is achieved (Kattan *et al.*, 2010). Learning algorithm for NNs can be divided into three approaches namely: supervised learning, unsupervised learning and reinforced learning.

NN learns from the provided input and output in supervised learning (SL) approach. The network learns through processing the input and compares the resulted output with the desired output. The error between the output and the desired output is propagated and the weight of the network is continuously adjusted until it reaches the optimal or near optimal error (certain optimal criteria is given). Stochastic and error correction gradient descent are the most essential for weight training algorithms in supervised learning. In stochastic, input weights are adjusted probabilistically while in error correction gradient-descent, errors are minimized in terms of weights and the activation function of the network. Supervised learning is appropriate when the class memberships of training data are known (Kulluk *et al.*, 2012).

Unsupervised learning is the scenario where there is no controller (teacher) that presents the network's expected output. The system itself learns by familiarizing and learning the structural features in the output pattern. This learning method is appropriate for the cases that the class memberships of training data are not known, or there is no desired output or the output cannot be predicted (Htike and Khalifa, 2010).

Reinforced learning is not as popular as its counterparts supervised learning and unsupervised learning. It is output based technique where there is a watcher (controller or teacher) which only controls the designated output whether is correct or not. On the other hand, the desired output is not presented therefore; the information provided to the network supports it in its learning process. So, a reward is given to the right answer and a penalty for a wrong answer (da Motta Salles and Anderson, 2008). However, the training of NN is basically succeeded in two ways: (1) adjusting the connection weights when the issue is predefined of ANN structure such as the number of neurons and their connections, the number of hidden layers and finally (2) the activation function parameters (Zamani and Sadeghian, 2010).

One of the most widely used algorithms for supervised training is the Back-Propagation (BP). It produces instances of the inputs and outputs to compute the network's output and reduce the mean square error (MSE) between the actual output and the required output by adjusting weights accordingly (Kulluk *et al.*, 2012). However, the BP algorithm experiences some drawbacks: (1) its learning and adoption are very slow in finding for the global minimum of the search space and (2) always is trapped at local minimum and (3) it requires a differentiable neuron transfer function (Castro and Von Zuben, 2011; Er and Liu, 2009; Kattan *et al.*, 2010; Kulluk *et al.*, 2012).

Development of learning algorithm for MLP has continuously attracted the focus of the researchers. As a result, many algorithms have been used by the researchers to train the MLP. The algorithms for training MLP include those inspired from biological processes like evolutionary (EA) as well as non-biological processes based algorithms such as: HSA.

EA is member of meta-heuristic algorithms and it has adopted to train feed forward NN (Correa and Gonzalez, 2011; El-Henawy *et al.*, 2010; Gao, 2008; Malllikarjuna *et al.*, 2011). Genetic Algorithm which can be considered within EA category has been used by (Correa and Gonzalez, 2011; Er and Liu, 2009; Ramos-Pollán *et al.*, 2011; Zhu and Wang, 2010) to train NN. Meanwhile, swarm intelligence-based algorithm called Particle Swarm Optimization has been applied to train NN by (Abdull Hamed *et al.*, 2012; Carvalho and Ludermir, 2007; Correa and Gonzalez, 2011; Zamani and Sadeghian, 2010) and Ant Colony (Gao, 2008; Yoshikawa and Otani, 2010).

Harmony Search algorithm (HSA) which is inspired from musician improvisation has also been used to train ANNs (Kattan *et al.*, 2010; Kulluk *et al.*, 2012; Tavakoli *et al.*, 2012). HSA has outperformed than BP, GA and PSO in terms of convergence rate speed, accuracy and towards reaching the global optimal solution (Geem *et al.*, 2007; Kattan *et al.*, 2010; Kulluk *et al.*, 2012; Soltani *et al.*, 2011; Tavakoli *et al.*, 2012). HSA has been revealed in various applications including engineering optimization, water distribution networks, vehicle routing and many more (Fetanat *et al.*, 2011; Yang, 2009). In this study, HSA was selected to adjust the weights and biases of BP-MLP whenever BP fails.

1.2 Problem Background

Traditionally training process of MLP NNs is divided into two phases. The first phase involves with determining the structure of hidden layers, hidden neurons and connection scheme. While the second phase, is involved with adjustment of the connection weights. The adjustment of connection weights is used by Back-Propagation (BP) algorithm (Malllikarjuna *et al.*, 2011; Teixeira *et al.*, 2008). BP algorithm was used to solve many real world problems using the concept of Multilayer Perceptron. Although many efforts had been done to speed up the BP's convergence rate still its convergence rate quite slows and always is trapped in local minima (Malllikarjuna *et al.*, 2011).

Therefore, the researchers such as (Abdull Hamed *et al.*, 2012; Carvalho and Ludermir, 2007; Zamani and Sadeghian, 2010) proposed a PSO technique to overcome the learning problem of MLP back-propagation. Some others such as (El-Henawy *et al.*, 2010; Zanchettin *et al.*, 2011) while (Er and Liu, 2009) used GA to train MLP. The author (Gao, 2008) employed ACO for MLP training. Further more, researchers such as (Kattan *et al.*, 2010; Kulluk *et al.*, 2012; Tavakoli *et al.*, 2012; Zinati and Razfar, 2012) proposed to train the MLP using HSA to optimize the weights and biases; since weight adjustments and correctly determining the ANN parameters are leading successful learning achievements to gain optimum result (Khorani *et al.*, 2011).

GA was used to initialize the weights of BP in order to train MLP feed forward NN and avoid both local minimum and low convergence rate of BP. GA has ability to escape the local minimum of BP and can successfully increase the possibility of finding a global solution instead of local one. Especially when the search space is large and has no obvious global minimum, at this time GA provides the best solution where most of other techniques fail (El-Henawy *et al.*, 2010; Yang, 2009). Its power to seek many local minimums made it so popular despite the major disadvantage of GA are that it demands more time during mutation (number of iterations) in the search process to the best solution as compared to PSO (El-Henawy *et al.*, 2010). Hence, other evolutionary algorithms have been given to consideration such as PSO

PSO was developed by Kennedy and Elberhart in 1995 (Sedighizadeh and Masehian, 2009). Kole and Halder (2010) have applied PSO to train MLP and found that it has performed better than GA. PSO performance in traing of MLP is also compared to Harmony Search Algorith (HSA) (Zamani and Sadeghian, 2010). PSO consumes much more time calling the fitness function by the number of its population size while HSA calls the fitness function only once in each iteration (Zamani and Sadeghian, 2010). Moreover, PSO is found trapped into a local minimum due to improper values assigned to parameters (Li *et al.,* 2009).

Ant Colony Optimization (ACO) is an optimization algorithm that derived from the behavior of ant colonies and proposed by Italy scholar M. Dorigo in 1990's and was intended to solve complicated combinatorial optimization problems (Gao, 2008). ACO generates simple agents in iterative process which repeatedly constructs candidate solutions those guided by heuristic information on given problem scenario along with a shared memory that contains previous knowledge collected by the ants in earlier repetitions; it has been applied to wide range of complex computational problems (Al Salami, 2009). Regarding the experomantal results done by Gao (2008) discovered that ACO can improve the NN (MLP) learning with efficiency and also tested that it is better than BP network of MLP however it was verified by typical XOR problem and it has been reported that XOR has no local minima (Kattan *et al.*, 2010).

Though HSA is a relatively new meta-heuristic algorithm, its effectiveness and advantages have been demonstrated in various applications (Ayachi *et al.*, 2013). Since its first appearance in 2001, it has been applied to solve many optimization problems mostly in engineering and industry, and the last ten years HSA has observed in applying many IT related applications such as robotics, web page clustering and classification problems. The possibility of combining harmony search with other algorithms such as Particle Swarm Optimization (PSO) has also been investigated (Soltani *et al.*, 2011). HSA is very successful in a wide variety of optimization problems. It also presents several advantages with respect to conventional optimization techniques. HSA does not require initial values for the decision variables and it imposes fewer mathematical requirements (Aungkulanon *et al.*, 2011). Furthermore, HSA is faster than PSO and it has a significant convergence rate to reach the optimal solution (Soltani *et al.*, 2011). HSA has been used to optimize feed forward MLPNN (Kattan *et al.*, 2010; Kulluk *et al.*, 2012; Tavakoli *et al.*, 2012; Zinati and Razfar, 2012). The experiments conducted by aforementioned researchers demonstrated that the performance of HSA is a good in terms of accuracy, speed (fast) and optimality besides, its ability to escape from local minimum.

According to the studies conducted by (Lu *et al.*, 2000; Shi *et al.*, 2009; Sun *et al.*, 2011; Wang *et al.*, 2008; Zhang *et al.*, 2007) and (Shayeghi *et al.*, 2010), BP has been combined with other evolutionary algorithms such as GA and PSO. The hybridizing models have shown achievements in the learning process of MLP BP aiming to avoid the BP's problem of getting into local minimum and slow convergence rate to perform better and increase the accuracy of the classifier. Therefore, there is a need for an enhanced model to enhance the convergence rate and avoid local minima of BP, since BP has the local search ability and HSA has global search ability, we combine their advantages and eradicate the disadvantages of both techniques.

In this study, a hybrid BP and HSA is developed to optimize and train the MLPNN to produce better performance. HSA is used to initialize weights and bias of MLP when BP failed. This is done for the cause of BP's ineffective attempts after the initialized weights are far from a good solution or near to poor local optimum which will take a lot of time to train since many iteration steps are required that cause the network not to converge to a satisfactory solution in the end. If the BP is not approaching the training goal or surely can be trapped in local minima the HSA will take the weight and bias adjustment using early stopping mechanism named “steady state’ in order to prevent the network being trapped into local minima or less accurate results.

1.3 Problem Statement

The MLP learning is influenced by many factors including the local minimum, learning rate, minimum error; number of input, hidden and output neurons as well as the activation function used. These factors can affect the convergence efficiency of MLP learning. Several meta-heuristic algorithms such as GA, PSO, ACO and HSA had been used by the researchers to determine parameters of MLP that aim for learning improvement.

In this study, a hybrid Harmony Search Algorithm and Back-Propagation Algorithm is employed to enhance the MLP learning. The performance of hybrid BPHSA-MLP is assessed in terms of speed convergence rate and avoiding local minima. HSA is a good at identifying the high performance regions of the solution space at a reasonable time. Therefore; this study will walk around the significance of implementing hybrid BP and HSA.

The research main question can be stated as:

How can the hybrid BPHSA-MLP escape the local minimum and slow convergence rate of BP-MLP to enhance the MLP learning?

The research sub questions are stated below:

1) How BPHSA can affect the MLP learning rate?
2) How can BPHSA avoid the BP's problem of local minimum and reach local optimal as well as global one?
3) How does BPHSA improve the convergence rate problem in MLP learning and increase the efficiency and the effectiveness of traditional machine learning (MLP training)?

1.4 Dissertation Aim

The aim of this research is to identify the efficiency of hybrid BPHSA compared to both standard BP and HSA in terms of classification accuracy and convergence rate in their application of artificial neural networks, especially MLP. The study will improve the learning of MLP using the benefits of both algorithms (HSA's fastness and the ability to escape local minimum plus the local search ability of BP).

1.5 Dissertation Objectives

The following objectives are proposed:

1. To propose BPHSA to enhance the learning performance of MLP.
2. To compare the proposed method with BP-MLP, HSA-MLP, in terms of classification accuracy, ability of escaping local minimum and error convergence.

1.6 Scope of the Study

A) This study will consider five different datasets and they are: Iris, Cancer, Wine, Thyroid and Glass. These mentioned datasets are used to test the performance of the proposed scheme (BPHSA).
B) Matlab is used to implement MLP FFANN, BP, HSA and BPHSA algorithms.
C) Mean Square Error (MSE) is used as a fitness or objective function.

1.7 Significance of the Study

The performance of HSA and BP algorithms in enhancing the MLP supervised training is investigated using the BPHSA integration. How BPHSA can avoid local minima and increase the speed of convergence rate is also studied, thus it is possible to determine which method is better to employ for MLP learning.

To identify the suitable and most appropriate technique for MLPNN training in terms of efficiency, accuracy, less time and human capital efforts as well as less economical cost is more important for future study and can be implemented in real world applications.

1.8 Dissertation Organization

This research currently consists of the following chapters

1. Chapter one (introduction) presents introduction of the study, problem background, problem statement, objectives and the scope of the research.
2. Chapter two (literature Review) reviews studies on machine learning techniques, their strengths and weaknesses and lastly solutions had been achieved since now and current issues on NNs learning methods.
3. Chapter three (research methodology) discusses the framework of the study; datasets that will be used to train the network as well as the techniques (algorithms) will be conducted in this research. This chapter focuses on the design of the network architecture according to different datasets used in this study as well as the proposed methodology of the study.
4. Chapter four (results) concentrates on the results obtained from the experiment and compares the outcome of the traditional MLP learning algorithm (BP) and recently SC introduced HSA and finally compares the results of hybrid form with the results obtained from both algorithms according to the datasets used as well as the accuracy and performance of the classifier.
5. Chapter Five (conclusion and future work) in finally the conclusion and recommendation for future work will be explained and discussed

REFERENCES

Abdull Hamed, H. N., Shamsuddin, S. M. and Salim, N. 2012. Particle Swarm Optimization For Neural Network Learning Enhancement. Jurnal Teknologi. 49(1), 13–26.

Al-Betar, M. A., Khader, A. T. and Zaman, M. 2012. University course timetabling using a hybrid harmony search metaheuristic algorithm. Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on. 42(5), 664-681.

Aungkulanon, P., Phruksaphanrat, B. and Luangpaiboon, P. 2011. Various hybridizations of harmony search algorithm for fuzzy programming approach to aggregate production planning. Proceedings of the 2011 Proceedings of the International Multi Conference of Engineers and Computer Scientists,

Ayachi, I., Kammarti, R., Ksouri, M., Borne, P. and de Lille, L. E. C. 2013. Harmony search algorithm for the container storage problem. arXiv preprint arXiv:1306.0090.

Carvalho, M. and Ludermir, T. B. 2007. Particle swarm optimization of neural network architectures andweights. Proceedings of the 2007 Hybrid Intelligent Systems, 2007. HIS 2007. 7th International Conference on: IEEE, 336-339.

Castro, P. A. and Von Zuben, F. J. 2011. Training multilayer perceptrons with a Gaussian artificial immune system. Proceedings of the 2011 Evolutionary Computation (CEC), 2011 IEEE Congress on: IEEE, 1250-1257.

Correa, B. and Gonzalez, A. M. 2011. Evolutionary algorithms for selecting the architecture of a MLP neural network: a credit scoring case. Proceedings of the 2011 Data Mining Workshops (ICDMW), 2011 IEEE 11th International Conference on: IEEE, 725-732.

da Motta Salles Barreto, A. and Anderson, C. W. 2008. Restricted gradient-descent algorithm for value-function approximation in reinforcement learning. Artificial Intelligence. 172(4), 454-482.

El-Henawy, I., Kamal, A., Abdelbary, H. and Abas, A. 2010. Predicting stock index using neural network combined with evolutionary computation methods. Proceedings of the 2010 Informatics and Systems (INFOS), 2010 The 7th International Conference on: IEEE, 1-6.

Er, M. J. and Liu, F. 2009. Genetic algorithms for MLP neural network parameters optimization. Proceedings of the 2009 Control and Decision Conference, 2009. CCDC'09. Chinese: IEEE, 3653-3658.

Fetanat, A., Shafipour, G. and Ghanatir, F. 2011. Box-Muller harmony search algorithm for optimal coordination of directional overcurrent relays in power system. Scientific Research and Essays. 6(19), 4079-4090.

Gao, W. 2008. Evolutionary neural network based on new ant colony algorithm. Proceedings of the 2008 Computational Intelligence and Design, 2008. ISCID'08. International Symposium on: IEEE, 318-321.

Geem, Z. W., Tseng, C.-L., Kim, J. and Bae, C. 2007. Trenchless water pipe condition assessment using artificial neural network. Proceedings of the 2007 The ASCE International Conference on Pipeline Engineering and Construction, Boston, USA.[doi: 10.1061/40934 (252) 26],

He, X. and Xu, S. 2009. Process Neural Networks: Theory and Applications. Springer.

Htike, K. K. and Khalifa, O. O. 2010. Comparison of supervised and unsupervised learning classifiers for human posture recognition. Proceedings of the 2010 Computer and Communication Engineering (ICCCE), 2010 International Conference on: IEEE, 1-6.

Kattan, A. and Abdullah, R. 2013. Training feed-forward artificial neural networks for pattern-classification using the harmony search algorithm. Proceedings of the 2013 The Second International Conference on Digital Enterprise and Information Systems (DEIS2013): The Society of Digital Information and Wireless Communication, 84-97.

Kattan, A., Abdullah, R. and Salam, R. A. 2010. Harmony search based supervised training of artificial neural networks. Proceedings of the 2010 Intelligent

Systems, Modelling and Simulation (ISMS), 2010 International Conference on: IEEE, 105-110.

Khorani, V., Forouzideh, N. and Nasrabadi, A. M. 2011. Artificial neural network weights optimization using ICA, GA, ICA-GA and R-ICA-GA: Comparing performances. Proceedings of the 2011 Hybrid Intelligent Models And Applications (HIMA), 2011 IEEE Workshop On: IEEE, 61-67.

Kole, D. K. and Halder, A. 2010. An efficient dynamic Image Segmentation algorithm using a hybrid technique based on Particle Swarm Optimization and Genetic Algorithm. Proceedings of the 2010 Advances in Computer Engineering (ACE), 2010 International Conference on: IEEE, 252-255.

Kulluk, S., Ozbakir, L. and Baykasoglu, A. 2012. Training neural networks with harmony search algorithms for classification problems. Engineering Applications of Artificial Intelligence. 25(1), 11-19.

Lee, K. S. and Geem, Z. W. 2005. A new meta-heuristic algorithm for continuous engineering optimization: harmony search theory and practice. Computer methods in applied mechanics and engineering. 194(36), 3902-3933.

Li, L., Lu, S.-b., Chu, X.-s. and Yu, G.-m. 2009. A Combinatorial Search Method Based on Harmony Search Algorithm and Particle Swarm Optimization in Slope Stability Analysis. Proceedings of the 2009 Computational Intelligence and Software Engineering, 2009. CiSE 2009. International Conference on: IEEE, 1-4.

Lu, C., Shi, B. and Chen, L. 2000. Hybrid BP-GA for multilayer feedforward neural networks. Proceedings of the 2000 Electronics, Circuits and Systems, 2000. ICECS 2000. The 7th IEEE International Conference on: IEEE, 958-961.

Malllikarjuna, K., Sarma, G., Subramanyam, M. and Prasad, K. 2011. EBP based GK-LM method for neural network training. Proceedings of the 2011 Computer Research and Development (ICCRD), 2011 3rd International Conference on: IEEE, 504-507.

Mitchell, T. M. 2006. The discipline of machine learning. Carnegie Mellon University, School of Computer Science, Machine Learning Department.

Narayan, S., Tagliarini, G. A. and Page, E. W. 1996. Enhancing MLP networks using a distributed data representation. Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on. 26(1), 143-149.

Neruda, R. 2007. Hybrid evolutionary algorithm for multilayer perceptron networks with competitive performance. Proceedings of the 2007 Evolutionary Computation, 2007. CEC 2007. IEEE Congress on: IEEE, 1620-1627.

Noriega, L. 2005. Multilayer perceptron tutorial. School of Computing. Staffordshire University.

Ramos-Pollán, R., Guevara López, M. A. and Naimy González de Posada, N. 2011. Optimizing the area under the ROC curve in multilayer perceptron-based classifiers. Proceedings of the 2011 FUTURE COMPUTING 2011, The Third International Conference on Future Computational Technologies and Applications, 75-81.

Romero, V. I. M., Tomes, L. L. and Yusiong, J. P. T. 2011. Tetris Agent Optimization Using Harmony Search Algorithm. International Journal of Computer Science Issues. 8(1), 22-31.

Sedighizadeh, D. and Masehian, E. 2009. Particle swarm optimization methods, taxonomy and applications. international journal of computer theory and engineering. 1(5), 486-502.

Shamsuddin, S. M., Alwee, R., Kuppusamy, P. and Darus, M. 2009. Study of cost functions in Three Term Backpropagation for classification problems. Proceedings of the 2009 Nature & Biologically Inspired Computing, 2009. NaBIC 2009. World Congress on: IEEE, 564-570.

Shayeghi, H., Shayanfar, H. and Azimi, G. 2010. A Hybrid Particle Swarm Optimization Back Propagation Algorithm for Short Term Load Forecasting. International Journal on Technical and Physical Problems of Engineering. 1(3), 12-22.

Shi, Q., Xu, Y. and Chen, S. 2009. Green Construction Alternatives Evaluation Using GA-BP Hybrid Algorithm. Proceedings of the 2009 Management and Service Science, 2009. MASS'09. International Conference on: IEEE, 1-5.

Soltani, M. A. Z., Haghighat, A. T., Rashidi, H. and Chegini, T. G. 2011. A Couple of Algorithms for K-Coverage Problem in Visual Sensor Networks. IACSIT. 96-101.

Sun, X., Liu, Q. and Zhang, L. 2011. A BP neural network model based on genetic algorithm for comprehensive evaluation. Proceedings of the 2011 Circuits, Communications and System (PACCS), 2011 Third Pacific-Asia Conference on: IEEE, 1-5.

Sunila, G. and Nirmal 2011. Intelligent and Effective Decision Support System Using Multilayer Perceptron. International Journal of Engineering Research and Applications (IJERA). 1(3), 513-518.

Tang, H., Tan, K. C. and Yi, Z. 2007. Neural networks: computational models and applications. (Vol. 53) Springer.

Tavakoli, S., Valian, E. and Mohanna, S. 2012. Feedforward neural network training using intelligent global harmony search. Evolving Systems. 3(2), 125-131.

Teixeira, L. A., Oliveira, F. T. and Oliveira, A. L. 2008. Adjusting Weights and Architecture of Neural Networks through PSO with Time-Varying Parameters and Early Stopping. Proceedings of the 2008 Neural Networks, 2008. SBRN'08. 10th Brazilian Symposium on: IEEE, 33-38.

Wang, S., Yin, S. and Jiang, M. 2008. Hybrid neural network based on GA-BP for personal credit scoring. Proceedings of the 2008 Natural Computation, 2008. ICNC'08. Fourth International Conference on: IEEE, 209-214.

Yang, X.-S. 2009. Harmony search as a metaheuristic algorithm Music-inspired harmony search algorithm (pp. 1-14)Springer.

Yoshikawa, M. and Otani, K. 2010. Ant Colony Optimization Routing Algorithm with Tabu Search. Proceedings of the 2010 proceedings of the International Multiconference of Engineers and Computer Scientists, 17-19.

Zamani, M. and Sadeghian, A. 2010. A variation of particle swarm optimization for training of artificial neural networks. Computational Intelligence and Modern Heuristics, Al-Dahoud Ali.

Zanchettin, C., Ludermir, T. B. and Almeida, L. M. 2011. Hybrid training method for mlp: optimization of architecture and training. Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on. 41(4), 1097-1109.

Zhang, J.-R., Zhang, J., Lok, T.-M. and Lyu, M. R. 2007. A hybrid particle swarm optimization–back-propagation algorithm for feedforward neural network training. Applied Mathematics and Computation. 185(2), 1026-1037.

Zhu, M. and Wang, L. 2010. Intelligent trading using support vector regression and multilayer perceptrons optimized with genetic algorithms. Proceedings of the 2010 Neural Networks (IJCNN), The 2010 International Joint Conference on: IEEE, 1-5.

Zinati, R. F. and Razfar, M. R. 2012. Constrained optimum surface roughness prediction in turning of X20Cr13 by coupling novel modified harmony search-based neural network and modified harmony search algorithm. The International Journal of Advanced Manufacturing Technology. 58(1-4), 93-107.