IJCTA, 9(34) 2016, pp. 715-736

ANN Training: A Survey of Classical & Soft Computing Approaches

Ashima Kalra#1, Shakti Kumar#2 and Sukhbir Singh Walia#3

ABSTRACT

Artificial neural networks ANNs are inspired by the structure of brain and have the capability to solve various wide varieties of complex problems. Learning in ANNs is a time consuming process. Hence, hunt for lower time complexity and higher accuracy training algorithm is always on. Various learning algorithms are available in the literature to train the neural networks. This paper carries out an extensive survey of the available training approaches. The paper divides the training approaches into two categories. Approaches using exact reasoning are called classical approaches and the training approaches using approximate reasoning for computing have been classified as soft computing based approaches. This paper presents the classical as well as soft computing based search and optimization algorithms available in the literature. The different classical approaches of learning were found to suffer from poor convergence speed and local optima problem. Soft computing based approaches were observed to have better global optima and hence were found to be more suitable for training ANNs. Some hybrid versions of soft computing and classical approaches were found to perform better than standalone versions.

Index terms: Back propagation, artificial neural network, learning algorithm, soft computing.

I. INTRODUCTION

An Artificial neural network (ANN) is a massively parallel, distributed processing system made up of simple processing elements which has the natural ability for storing experiential knowledge and making it available for use when required. ANN is an information processing paradigm that is inspired by the structure of brain. Neural networks have a special ability to derive something meaningful from any complicated data which can be further used to extract patterns and detect trends that are too complex to be noticed by either humans or computers. ANN's are typically applied for pattern classification [1-2] and pattern recognition [3-4]. They have been successfully used for stock market predictions [5], wear and manufacturing processes [6], speech recognition [7], business applications [8], control applications [9], time series modelling and estimation applications [10], medical diagnosis [11-13], aeronautics [14] etc.

ANNs also found their applications as expert system. These were also used to aid fuzzy system design. Gallant S. [15] was the first to describe a system combining the domain expert knowledge with neural training. This connectionist expert system (CES) consists of an expert system implemented throughout a multi-layer perceptron. Various authors [16-17] demonstrated the structural learning with forgetting for training neural networks. Melanie et al. [18] presented a strategy for Modular Integration of Connectionist and Symbolic Processing in knowledge based systems. ANNs also find their application in rule base generation from numerical data for fuzzy systems. Authors [19-20] demonstrated the rules generation from trained network using fast pruning. A good survey has been found on the extraction of rules from the trained neural networks in papers [21-22]. Various authors in their papers [23-42] presented the various different techniques for extracting rules from trained neural networks. Thrun [43] presented extraction of

#1 Research Scholar, Punjab Technical University, Jalandhar, Punjab, INDIA
#3 Punjab Technical University, Jalandhar, Punjab, INDIA
E-mails: [1]uppal.ashi@gmail.com; [3]ahluwaliasukhbir@gmail.com
#2 Computational intelligence lab (CI), Baddi University of Emerging Sciences & Tech, Baddi (HP) INDIA, *E-mail: [2]shaktik@gmail.com*

rules from artificial neural networks with distributed representations. Different types of architectures and corresponding learning algorithms can be found in literature. Some of the widely used architectures along with their learning algorithms and applications are given in Table I. There are basically two types of architectures: Feed forward networks, Feedback /Recurrent networks. Feed forward networks can further be classified as single layer perceptron, multi-layer perceptron (MLP) and radial basis function networks (RBFNs). On the other hand recurrent/ feedback type networks consist of networks like competitive networks, Kohenen's self-organizing Maps (SOM) Hopfield networks (HN) and adaptive resonant theory (ART) models.

Learning is the backbone of any kind of neural network design. Learning is a process by which the free parameters of a neural network are adapted through a process of simulation by the environment in which the network is embedded. There are three types of learning algorithms: Supervised learning, unsupervised learning and hybrid learning algorithms. Supervised learning is the learning with a teacher. In this type of learning, a kind of target or desired output signal is present with which the computed output signal is compared to compute the error signal. Paradigms of supervised learning include Perceptron learning algorithm, error correction learning, stochastic learning, Adaline, Medaline Boltzmann learning algorithm, learning vector quantization etc. An unsupervised learning is a kind of learning without teacher i.e. no desired output signal is available. This type of learning is based on the concept of using self-organization to compute the output. Paradigms of unsupervised learning are Hebbian learning and competitive learning, Kohenen learning (SOM), ART1 and ART2. Hybrid learning is the combination of two types of learning mentioned above. Example of hybrid learning is radial basis learning approach. The essence of a learning algorithm is the learning rule, i.e., a weight-updating rule which determines how connection weights are changed to reduce the error. Different types of learning laws are used to update the synaptic weights like delta law, perceptron law, instar learning, outstar law etc. The training process can be on-line or batch training. In on line training, the weights are adjusted after the processing of a randomly selected training pattern while in batch training; the weights are adjusted after processing.

Since, ANN based systems are highly complex and nonlinear systems, we divide the learning algorithms in the following two classes: a) Classical Learning algorithms b) Soft Computing based algorithms .We define soft computing based algorithms as the one's that uses approximate reasoning in ANN training. These algorithms include fuzzy logic based approaches, swarm based approaches and the other approaches based upon certain other nature inspired computing approaches. Since, ANN training is a time consuming process, research community is continuously forced to search for new, more accurate and less time consuming approaches to ANN training.

This paper consists of V sections. Section II of this paper presents a quick glance on the classical ANN training approaches available in the literature. Section III reviews soft computing based training approaches found in the literature. Section IV of this paper compares both the classes of approaches and Section V concludes the paper.

II. CLASSICAL LEARNING APPROACHES

The main aim of training is to minimize the error between the output and target output For the ANN modelling, it is required to decide the architecture first with consideration to the number of hidden layers and number of neurons in each hidden layer as well. After the Architecture is decided the next step is weight adjustment so as to minimise output error.

For our convenience we define a training approach as a classical approach if it uses exact reasoning o hard computing for its computation. These approaches can further be classified into two classes depending on their basic strategy. First order methods include the computation of error gradient along with simple modifications like the use of a momentum term or the adaptive leaning methods. There may either be one global learning rate for all weights or individual learning rates for each single weight. Second order methods compute a quadratic approximation of the error surface which is then minimized in order to reach the minimum of the actual error function iteratively.

The major classical training approaches available in literature are shown in Table II. Werbos [44] introduced EBP for the first time as the roots of propagation. Rumelhart [45] further elaborated it as the learning representations of ANNs by back propagation of error. This algorithm includes standard or incremental back propagation (IBP) and batch back propagation (BBP). Freeman and Skappura [46] proposed incremental back propagation (IBP) in which the network weights are updated after presenting each pattern from the learning data set, rather than once per iteration. Hagan *et al.* [47] proposed batch back propagation (BBP) in which the network weights are updated once per iteration, while all learning data pattern are processed through the network. Quick propagation (QP), Fahlmann [48] is a heuristic modification of the back propagation algorithm. To improve the convergence speed of EBP, Rumelhart extended his work by introducing the momentum term as given in [49-51]. The various adaptive learning methods like Delta –Bar –Delta (DBD), R A Jacobs, Tollenaere [52] [54], Extended Delta bar delta(EDBD), Minai and Williams [56], SSAB (Super-SAB), Tollenaere [57], resilient propagation(RPROP), Riedmiller and Braun [58] and the generalized no-decrease adaptive method (GNDAM), R. Allard , J. Faubert [59] have been developed to self-adjust the learning rate and to just get rid of the slow convergence problem thereby obtaining the optimal solution. The hybrid learning schemes have been proposed to incorporate second derivative related information in Rprop, such as the QRprop [60], which is the combination of RPROP with the one dimensional secant steps of quick algorithm and the Diagonal Estimation Rprop–DERprop [61], which directly computes the diagonal elements of the Hessian matrix. Also approaches inspired from global optimization theory have been developed to equip Rprop with annealing strategies, such as the Simulated Annealing Rprop–SARprop and the Restart mode Simulated Annealing Rprop i.e., ReSARprop [61] help to escape from shallow local minima. Another improvement was proposed as Improved Rprop (iRprop) algorithm by C Igel and M Husken [62-64] which applies a backtracking strategy (i.e. it decides whether to take back a step along a weight direction or not by means of a heuristic), has shown improved convergence speed when compared against existing Rprop variants, as well as other training methods. Aristoklis D A [65-66] proposed another algorithm G Rprop which was a modification over iRprop. It exhibited better convergence speed and stability than Rprop and iRprop.

The other two second order approaches namely conjugate gradient and Quasi Newton approaches have been reported as the most successfully applied to the training of feed forward neural networks amongst all those using second order information. Conjugate gradient method was first initiated by Hestenes and Stieffel (1952) for linear functions and then based on this work, Fletcher and Reeves (1964) further extended it as conjugate algorithm for nonlinear functions. Afterwards Beale proposed conjugate gradient method with the provision of restarting direction procedure. Navon and Legler [67] has presented a review of various conjugate gradient algorithms for large scale minimization where they covered almost four types of Conjugate Gradient algorithms and compared their advantages as well as shortcomings. Johansson et al. [68] proposed a conjugate gradient with line search (CGL) method where a step size is approximated with line search by avoiding the calculation of Hessian Matrix. Moller [69] proposed a Scaled conjugate gradient (SCG) method for fast supervised learning and was found to be faster than CGL and BP. Barreto, Anderson [70] proposed a restricted Gradient descent (RGD) algorithm for training local RBF networks in the context of reinforcement learning. BFGS Quasi Newton optimization approach with limited memory was first proposed by R. Battiti and F. Masulli [71]. There were basically two update approaches for quasi Newton’s method – BFGS (Broyden, Fletcher, Goldfarb and Shanno) update, Battiti [71-72] and the DFP (named for Davidon, Fletcher and Powell) update, Watterous [73]. J E Dennis and J J More in paper[74] presented the survey with justification of use of Quasi Newton methods over Newton method for general and gradient non linear systems and proved it more computational efficient than Newton method. A Likas, A Stafylopatis [75] presented the training of random neural network using Quasi newton methods.

For fast and efficient training, second order learning algorithms have proved to be very useful. Levenberg Marquardt (LM) algorithm [76], which is a derivative of the Newton method, appears to be one of the most effective algorithm. It is a good combination of Newton's method and steepest descent. The Levenberg-Marquardt algorithm (LM) proposed by M. T. Hagan and M. B. Menhaj [76-77], is also used as another algorithm to increase the convergence speed. But the LM algorithm becomes impractical for large sized neural networks which leads to another modification of LM i.e. called TREAT algorithm, Y Chan [78]. So traditional training algorithms, such EBP and LM have been successfully applied to train neural networks by some authors in papers [79-83]. But still these algorithms require more space complexity, time complexity, and there is always a risk of getting trapped in local minimum, as they are not derivative free. Hence, these appear to be some of the main reasons for a shift towards nature inspired or soft computing based search and optimization approaches used in ANN training.

III. SOFT COMPUTING OPTIMIZATION APPROACHES

Learning most often is modelled as an optimization process wherein the error is minimized as the learning progresses. The heuristic approaches based upon like nature inspired or soft computing based algorithms are much superior in solving complex optimization problems where traditional or classical problem solving methods fail. This is particularly true for NP Hard or NP complete problems. Soft computing based search and optimization approaches can be broadly classified into four- Evolutionary computing (EC), Swarm intelligence (SI), Bio inspired Non-SI and physics or chemistry basedappoaches. Various soft computing based Search and optimizations based approaches available in literature are shown in Table 3 and Table 4. Evolutionary Computing is based on the biological evolution process in nature. Swarm intelligence based algorithms are based upon collective social behaviour of organisms. Bio inspired Non SI optimization algorithms are bio inspired or ecology based but are not inspired by the cooperative behaviour of any organisms. Physics or chemistry based algorithms are actually inspired by certain physical or chemistry laws like electric charges, gravity, theory of universe etc. Literature is also rich with soft computing based search and optimization approaches. ANN learning model, based on evolutionary algorithms (EAs), Yao [84-85], based upon genetic algorithms (GAs), Holland [90-93] , based upon genetic programming (GP), John Koza [97-98], based upon differential evolution(DE), Storn and Price[102] , based upon particle swarm optimization (PSO), Eberhart and Kenedy [113], based upon big bang big crunch (BB-BC) algorithm, Erol Eksin [118-119], ant colony optimization (ACO), Dorigo and Gambardella[122-125], based upon ABC, Karaboga and Basturk [126], based upon firefly mating/foraging behaviour called firefly algorithm (FA) proposed by Yang [129], based upon Fish swarm intelligence (FSA), Li Xiao-lei *et al.* [132], based upon Bacterial foraging optimization algorithm(BFOA), Passino[143], based upon invasive weed optimization (IWO), Mehrabian and Lucas [159-160], biogeography based optimization(BBO), Dan Simon [161],based on artificial immune system (AIS), Dasgupta[166-168], based upon EEIM, Birbil & Fang [188-193] and based upon GSA, Rashiedi *et al.* [197] are available in literature.

These soft computing based approaches have been successfully applied in search and optimization problems. Kumar *et al.* [216] presented ANN model identification for rapid battery Charger using Parallel BB-BC (PB3C) approach. PB3C algorithm is a multi-population algorithm and was first proposed by Kumar *et al.* [218]. Kumar *et al.* [219] proposed fuzzy or ANN model identification in the field of overall rating and evaluation of institutions of higher learning using BB-BC and PB3C algorithms. Similar kinds of techniques are available for fuzzy model identification also. Fusion of Artificial Neural Networks (ANN) and Fuzzy Inference Systems (FIS) have attracted interest of researchers in various scientific and engineering areas due to the growing need of adaptive intelligent systems to solve the real world problems. Neuro Fuzzy (NF) computing is a popular framework for solving complex problems. Various authors [221-225] have discussed the design of neuro fuzzy controller and adaptive neuro fuzzy systems [226-227].

IV. CLASSICAL VS. SOFT COMPUTING APPROACHES

Various authors have compared the soft computing based approaches with classical learning approaches as well as hybrid techniques for NN learning. Various comparisons have been made in literature among classical learning approaches too. So as to speed up the back propagation, Marcus Pfister et al. [228] compared five algorithms i.e., gradient reuse algorithm, DBD, Extended DBD, dynamic adaptation, quick prop and extended quick prop for five different benchmark problems. This paper also concluded that a learning algorithm that may prove fast for one problem, may fail in another case. The results show that Quickprop was the one that performed very well in all the benchmark problems while Extended QP was a big failure. For smaller problems gradient reuse algorithm is faster than BP but even much slower in case of complex problems. Remy Allard *et al.* [229] compared and tested the various adaptive learning methods i.e. MOM (momentum), DBD (Delta Bar Delta), SSAB (Super SAB), Resilent Prop RPROP, and Generalized no-decrease adaptive method GNDAM on four benchmark problems i.e., parity-bit, encoder, texture detection and luminance. It clearly shows that a single AM approach cannot be proved overall best for all the tasks but the results may vary depending upon the task. MOM and DBD had a similar behaviour when they were used on the luminance, encoder and parity task. The only task for which they clearly differed was the texture task where MOM never solved the task as opposed to DBD. RPROP showed a net advantage over SSAB for the parity-bit and luminance tasks detection. A numbers of soft computing based approaches are compared against BP or LM for various bench mark problems in the literature.

Sagar [230] also proposed an EA for Connection Weights in Artificial Neural Networks and compared it with BP algorithm for a X-OR benchmark problem. It was shown that EA-ANN approach gave zero mean square error than the (BP) gradient descent method and the results did not depend on the initial choice of weights. It gives the increased performance of the network in terms of accuracy. Jagtap [231] proposed a quantum based method called QNN method for four well-known benchmark classification problems namely breast cancer, Iris data classification, and heart and diabetes problems. QNN is helpful to provide a set of appropriate weights when evolving the network structure and to alleviate the noisy fitness evaluation problem.

Kitanao [232] compared genetic algorithms (GAs) with BP (back propagation) and presented hybrid approach GA-BP which proved to be faster than GA alone. The author also stated that the GAs is equally efficient to the faster variants of BP in small scale networks but found less efficient in larger networks. Gupta *et al.* [233] compared Standard EBP with GA for optimizing artificial neural networks. The empirical results showed that the GA is superior to BP in effectiveness, ease-of-use and efficiency in training ANNs. Further Zhen Guo Che *et al.* [234] compared BP with GA and drawn conclusion that BP is much superior and having faster training speed than GA with a drawback of having over training which GA didn't have. Paul Batchis [235] compared EA with BP using Weka Knowledge Explorer software package on three classification problems. In this, EA is found to outperform BP method. Asha *et al.* [236] compared ABC with BP for classification task using four benchmark datasets availed from the UCI machine learning. This paper implemented ABC for optimizing the connection weights and concluded that ABC performance was found to be better for the four datasets as compared to BPN performance.

For achieving global optimization, various soft computing based global optimization algorithms can be used in standalone or in a hybrid manner. In a hybrid algorithm, some local search phenomena like BP or LM is hybridized with some soft computing based global optimization algorithm. Such hybrid approaches are also compared with standalone soft computing based approaches. Alba and J. Francisco Chicano [237] proposed training Neural Networks with GA Hybrid Algorithms. They suggested the concept of weak hybridization (just the combination of two algorithms) by introducing and testing GA with the BP algorithm (GABP), and a GA with LM (GALM) algoithm. J Zhang *et al.* [238] proposed a hybrid PSO-BP algorithm where it was shown that the PSO–BP algorithm uses less CPU time to get higher training accuracy than the PSO algorithm as well as the BP algorithm. The hybrid PSO-BP was observed to be better than using BP or PSO alone. Another hybrid learning approach ACO-BP was proposed by L Yan Peng et al. [239] and their results show that the ACO-BP is more effective and efficient than the standalone BP algorithm. It was also concluded that with the variation in number of hidden nodes, the performance of ACO-BP became stable compared to ACO or BP alone. Mavrovouniotis and Yang [240-241] proposed NNACO-BP for different real-world benchmark dataset taken from the UCI repository.This paper compared the performance of ACO and ACO-BP training against: two classical learning approaches i.e. BP and LM, RCH (ACO training without pheromone consideration), a standalone ACO and a hybrid ACO i.e., ACOR and ACOR-BP, respectively, and four soft computing based approaches i.e. GA, PSO, ABC and DE. The author concluded that ACO was a good choice for selecting good values for the BP. The standalone ACO training was outperformed by the standalone ACOR training whereas the hybrid ACO-BP showed superior performance, especially on large problem instances. Secondly, the performance of gradient descent methods is degraded as the problem size increases when compared with the hybrid ACO-BP training algorithm. Third, gradient descent methods usually have better performance than a standalone GA, PSO, ABC, ACO and DE training. ACO has a relatively good performance when compared with other network training algorithms for pattern classification. Huadong Chen *et al.* [242] proposed a hybrid of AFSA-PSO for feed forward Neural Network Training and showed that hybrid FSA-PSO has better global astringency and stability than standalone FSA and standalone PSO. Nandy *et al.* [243] compared the performance of hybrid ABC-BP with hybrid GA-BP on the basis of three parameters i.e. sum of squared error (SSE), convergence speed and stability on optimum solution for four datasets (Iris, Wine, Soyabean and Glass). It showed that ABC-BP is better than GA-BP with increased efficiency.

Various soft computing based approaches for NN learning are also compared with each other. Yun Cai [245] proposed Artificial Fish School Algorithm (AFSA) for Combinatorial Optimization Problem and stated that the algorithm has better convergence performance than GA and ACO. Basturk and Karaboga [246] compared the performance of ABC algorithm with GA, PSO and Particle Swarm Inspired Evolutionary Algorithm (PS-EA). The results showed that ABC outperforms the other algorithms. Basturk and Karaboga further in [247] compared the performance of ABC algorithm with that of DE, PSO and EA for a set of well known test functions. Simulation results show that ABC algorithm performs better than the mentioned algorithms and can be efficiently employed to solve the multimodal engineering problems with high dimensionality. D Karaboga and B Akay [248] compared ABC with GA, PSO, DE and ES for optimizing a large set of numerical test. Results show that the performance of the ABC was better than or similar to those of other population-based algorithms with the advantage of employing fewer control parameters. Saishanmuga Raja et al. [249] compared three optimization techniques GA, ACO and PSO in biomedical application based on processing time, accuracy and time taken to train Neural Networks. The paper concluded that GA outperformed the other two algorithms ACO and PSO and is most suitable for training the neural network with minimum time and minimum mean square error. Srinivasan et al. in paper [250] proposed particle swarm inspired EA (PS-EA) and compared it with Genetic Algorithm (GA) and PSO. It is found that PS-EA is much superior over typical GA and PSO for complex multi-modal functions like Rosenbrock, Schwefel and Rastrigrin functions. Ghaffari *et al.* [251] presented the comparison of five training algorithms-two versions of gradient descent-IBP (incremental), BBP (Batch) and LM, QP, GA with reference to the predicting ability. The convergence speed of BBP is three to four times higher than IBP. The performances in terms of precision of predictive ability were in the order of: IBP, BBP > LM > QP (quick propagation) > GA. Zhang *et al.* [252] presented application of bacteria foraging optimized neural network (BFO NN) for short term electric load forecast. This paper used BFO to find optimized weights of neural network while minimizing the MSE. Simulation results also showed that BFONN converges quickly than Genetic Algorithm optimized neural network (GANN).

Ivona BRAJEVIC *et al.* [253] presented Training Feed-Forward Neural Networks Using Firefly Algorithm (FA) for classification purpose. This paper compares FA with GA and ABC for three well known classification problems (X-OR, three bit parity, four bit encoder /decoder problem) FA using sigmoid transfer function. The parameters used for comparison are MMSE: Mean of Mean Squared Errors, SDMSE: Standard Deviation of Mean Squared Errors, MC: Mean of Cycle Numbers, SDC: Standard Deviation of Cycle Numbers. It shows that FA performs better than GA algorithm, but worse than ABC algorithm for the majority of benchmark problems. It also stated that the choice of transfer functions may strongly influence the performance of neural networks, so it also compared the FA results obtained by using traditional sigmoid transfer function with another by using sine transfer function and showed that FA implemented using sine transfer function is much efficient with fast convergence speed. Ritwick [254] proposed a modified version of Invasive Weed Optimization (MIWO) for training feed-forward Artificial Neural Networks (ANNs) by adjusting the weights and biases of the neural network. In this, Modified IWO was compared with DE, BP, One step secant learning and RPROP based on MSE. Modified IWO performed better than DE and other classical gradient-based optimization algorithms mentioned in terms of learning rate and solution quality. BBO has been further compared with other optimization algorithms like PSO, ACO and was found to be better for detection of abnormal growth of tissues in MRI image segmentation by Kaur et al. [255]. Mirjalili et al. [256] proposed hybrid PSO-GSA for training neural networks and proved it to outperform other optimization based training algorithms such as PSO and ACO in terms of converging speed and local minima avoidance. Sheikhpour [257] proposed a hybrid GSA-GA for neural network training that uses the Gravitational Search Algorithm (GSA) to find global search in the beginning of stage, and then uses the GA to do local search around the global optimum and proved it to be more efficient than standard GSA and back propagation algorithm. Xu et al. [258] proposed an Improved Gravitational Search Algorithm (IGSA) for Dynamic Neural Network Identification. It showed the best performance when compared with the system identification based on gravitational search algorithm neural network (GSANN) and other conventional methods like BPNN and GANN. Ayat et al. [259] compared various ANN learning algorithms in which 12 algorithms concerned with Perceptron multilayer neural networks were studied along with 6 classical learning algorithms (Gdx, cgb, lm, oss, cgf, cgp) .In this paper, conjugate gradient and LM is found to have better efficiency in reaction to the given training as compared to other learning algorithms. The paper concluded that LM is the most convergent and represented better predict of average. Mohammadi et al. [260] compared the PSO with the variants of back propagation techniques (LM, GD, GDM, GDA, GDMA) based on mean square error (MSE) and accuracy. The paper concluded that LM is found to have better performance than other variants of BP but PSO is more superior to LM and other variants of BP. The performance level is found to be PSO>LM>other BP variants.

V. CONCLUSIONS

This paper attempts to present state of the art in ANN training through an exhaustive review of classical as well as the soft computing based approaches available in the literature. In the case of classical learning approaches, it is evident that not a single training algorithm can be proved best for all the test or benchmark problems. In fact it is the problem dependant. It is found that as the classical learning approach such as EBP or LM is having the poor convergence speed. We further observe that the soft computing based approaches perform better as the global optimization approaches. These soft computing based approaches can be used to evolve ANN architecture or the synaptic weights. These approaches can be used standalone or in a hybrid manner with EBP or LM. It is evident that the hybrid techniques are found to be more efficient than standalone soft computing approaches. This survey covers most of the available classical and soft computing approaches used to evolve ANN's. BB BC or parallel BB BC is a new approach among soft computing which could be further utilized broadly for new ANN model identification approaches.

REFERENCES

[1] G. Zhang, "Neural networks for classification: a survey," IEEE Transactions on Systems, Man, and Cybernetics-Part C: Applications and Reviews, vol. 30, no. 4, pp. 451–462, 2000.

[2] Xudong Jiang, Alvin Harvey Kam Siew Wah, "Constructing and training feed forward neural networks for pattern classification", Elsevier, Pattern Recognition 36 (2003), pp 853 – 867.

[3] CHRISTOPHER M. BISHOP, "Neural Networks for Pattern Recognition", CLARENDON PRESS ,OXFORD,1995

[4] CHRISTOPHER M. BISHOP, "Neural Networks: A Pattern Recognition perspective", Neural Computing Research Group, Astan University, UK, January1996.

[5] Dase R.K and Pawar D.D., "Application of Artificial Neural Network for stock market predictions: A review of literature", International Journal of Machine Intelligence, ISSN: 0975–2927, Volume 2, Issue 2, 2010, pp-14-17.

[6] Minodora RIPA and Laurentiu FRANGU. "A SURVEY OF ARTIFICIAL NEURAL NETWORKS APPLICATIONS IN WEAR AND MANUFACTURING PROCESSES", FASCICLE VIII, 2004, ISSN 1221-4590, pp 35-42.

[7] Halima Bahi and Mokhtar Sellami, "A Connectionist Expert Approach for Speech Recognition", The International Arab Journal of Information Technology, Vol. 2, No. 2, April 2005,pp148-153.

[8] Eldon Y. Li, "Applications Artificial neural networks and their business applications", Elsevier, Information & Management 27 (1994), pp 303-313.

[9] António Eduardo de Barros Ruano., "Applications of Neural Networks to Control Systems ", Ph. D Thesis, University Of Wales, Feb 1992.

[10] Juan Peralta Donate , Xiaodong Li, "Time series forecasting by evolving artificial neural networks with genetic algorithms, differential evolution and estimation of distribution algorithm, Neural Comput & Applic (2013) 22,pp 11–20

[11] Filippo Amato, Alberto López, Eladia María Peña-Méndez, Petr Vaòhara, Aleš Hampl, Josef Havel, "Artificial neural networks in medical diagnosis", J Appl Biomed. 11, 2013, pp 47–58.

[12] S. M. Kamruzzaman and Md. Monirul Islam, "An Algorithm to Extract Rules from Artificial Neural Networks for Medical Diagnosis Problems, International Journal of Information Technology, Vol. 12 No. 8, 2006.

[13] Irfan Y. Khan, P.H. Zope, S.R. Suralkar. "Importance of Artificial Neural Network in Medical Diagnosis disease like acute nephritis disease and heart disease", International Journal of Engineering Science and Innovative Technology (IJESIT) Volume 2, Issue 2, March 2013.

[14] A K Ghosh, "Neural networks: Applications and opportunities in Aeronautics", Feb 2006, pp 49-54.

[15] Gallant S. I., "Connectionist Expert Systems," Communications of the ACM, vol. 31, no. 2,1988, pp.152-169.

[16] M. Ishikawa, "Structural learning with forgetting," Neural Networks, vol. 9, pp. 509–521, Apr. 1996.

[17] Damon A. Miller, and Jacek M. Zurada, "A Dynamical System Perspective of Structural Learning with Forgetting", IEEE TRANSACTIONS ON NEURAL NETWORKS, VOL. 9, NO. 3, MAY 1998, pp 508-515.

[18] Melanie Hilario, Christian Pellegrini & Frédéric Alexandre, "Modular Integration of Connectionist and Symbolic Processing in Knowledge-Based Systems", International Symposium on Integrating Knowledge and Neural Heuristics (pp. 123-132). Pensacola, Florida. May, 1994.

[19] Rudy Setiono and Wee Kheng Leow, "Generating rules from trained network using fast pruning", IJCNN' 99.

[20] Alexander N. Gorban, Eugeniy M. Mirkes and Victor G. Tsaregorodtsev, Generation of Explicit Knowledge from Empirical Data through Pruning of Trainable Neural Networks".

[21] R. Andrews, J. Diederich, and A. Tickle, "Survey and critique of techniques for extracting rules from trained artificial neural networks," Knowledge Based Syst., vol. 8, no. 6, pp. 373–389, 1998.

[22] Ashish Darbari, "Rule Extraction from Trained ANN: A Survey," Technical Report, Department of Computer Science, Dresden University of Technology, Dresden, Germany, 2000.

[23] H. Liu and S.T. Tan, "X2R: A fast rule generator," in Proc. of IEEE Int. Conf. on SMC, New York: IEEE Press, 1995, 1631–1635.

[24] R. Setiono and Huan Liu, "Understanding neural networks via rule extraction," Proceedings of the 14th International Joint Conference on Artificial Intelligence, 1995, pp. 480-485.

[25] R. Setiono and H. Liu. "Symbolic representation of neural networks," IEEE Computer, 29(3), 71–77, 1996

[26] R. Setiono, "Extracting M-of-N Rules From Trained Neural Networks'', IEEE Transactions on Neural Networks, Vol. 11, No. 2, 2000, pp512-519.

[27] Olcay Boz, "Knowledge integration and rule extraction in neural networks," EECS Department, Lehigh University, 1995.

[28] R. Setiono, W.K. Leow and J. Zurada, Extraction of rules from artificial neural networks for nonlinear regression, IEEE Transactions on Neural Networks, 13(3), 2002, 564 577.

[29] G. G., Towell and J. W., Shavlik, “Extracting refined rules from knowledge-based system neural networks, “Machine Learning, vol. 13, no. 1, pp. 71-101, 1993.

[30] R. Setiono and H. Liu, “Neural network feature selector,” IEEE Trans. Neural Networks, vol. 8, pp. 654–662, May 1997.

[31] I. Taha and J. Ghosh, “Three techniques for extracting rules from feed forward networks,” Intelligent Engineering Systems Through Artificial Neural Networks, vol. 6, pp. 23-28, ASME Press, St. Louis,1996.

[32] S. M. Kamruzzaman, and Md. Monirul Islam, “Extraction of Symbolic Rules from Artificial Neural Networks”, World Academy of Science, Engineering and Technology 10 2005.

[33] S. M. Kamruzzaman, Ahmed Ryadh Hasan, “Rule Extraction using Artificial Neural Networks”, ICTM 2005.

[34] Zhi-Hua Zhou , “Rule Extraction: Using Neural Networks or For Neural Networks?”,National Laboratory for Novel Software Technology, Nanjing University, Nanjing 210093, China

[35] Lichen Liang, Feng Cai _, Vladimir Cherkassky, “Predictive learning with structured (grouped) data” ,Neural Networks 22 (2009) pp 766-773

[36] J. T. Yao, “Knowledge extracted from trained neural networks — what’s next?”

[37] S. M. Kamruzzaman, “REx: An Efficient Rule Generator”, Proceedings of the 4th International Conference on Electrical Engineering & 2nd Annual Paper Meet 26-28 January, 2006

[38] GG Towell and J W Shivalik, Knowledge-Based Artificial Neural Networks”, Appears in Artificial Intelligence, volume 69 or 70.,Submitted 1/92, Final pre-publication revisions 8/94

[39] W Duch, R Adamczak and K Gra¸bczewski, “Extraction of logical rules from training data using back propagation networks”, Neural Process. Lett. vol. 7, pp. 211 -219, 1996 .

[40] Zhi-Hua Zhou Shi-Fu Chen Zhao-Qian chen, A Statistics based Approach for Extracting Priority Rules from Trained Neural Networks, In: Proceedings of the IEEE-INNS-ENNS International Joint Conference on Neural Networks, Como, Italy, 2000, vol.3, pp.401-406.

[41] Urszula Markowska-Kaczmar and Krystyna Mularczyk, “GA-Based Rule Extraction from Neural Networks for Approximation, Proceedings of the International Multiconference on Computer Science and Information Technology pp. 141–148, 2006.

[42] Hiroshi Tsukimoto, “E x t r a c t i n g Propositions from Trained Neural Networks, Neural Networks, pp 1098-1105.

[43] Thrun S. Extracting rules from artificial neural networks with distributed representations”, Advances in Neural Information Processing Systems 7, Cambridge, MA: MIT Press, 1995, pp.505-512.

[44] D E Rumelhart, G E Hinton and R J Williams “Learning internal representations by error propagation, MIT Press 1986.

[45] Werbos, P.J., “The Roots of Back propagation”, John Wiley k Sons, Inc., New York, 1994.

[46] Freeman, J.A., Skappura, D.M., 1991. Neural Networks Algorithms, Applications and Programming Techniques. Addison-Wesley, Houston, pp. 12– 105.

[47] Hagan, M.T., Demuth, H.B., Beale, M.H., 1996. Neural Network Design. PWS, Boston

[48] Scott E. Fahlman, “An empirical study of learning speed in back propagation Networks”, 1988, CMU-CS-88-162.

[49] Ning Qian, “On the momentum term in gradient descent learning algorithms”, Elsevier, Neural Networks 12 (1999), pp 145–151.

[50] Stavaros J , Karras D A , “An Efficient Constrained Learning algorithm with momentum accerlation”, Elsevier, Neural Networks ,Vol 8, No. 2,1995, pp 237-249.

[51] M Moreira and E Fiesler, “Neural Networks with Adaptive Learning Rate and Momentum Terms” IDIAP, technical report,October1995.

[52] R.A. Jacobs, “Increased rates of convergence through learning rate adaptation”, Neural Networks, 1, pp 295-307, 1988.

[53] Romdhane BenKhalifa, Noureddine BenYahia and Ali Zghal, “Integrated neural networks approach in CAD/CAM environment for automated machine tools selection” Journal of Mechanical Engineering Research Vol. 2(2), pp. 25-38, March 2010.

[54] M Pfister, R Rojas, “speeding up Back propagation: A comparison orthogonal techniques”, 1993, pp 1-8.

[55] Binitha S, S Siva Sathya, A survey of bio inspired optimization algorithms ,IJSCE,Vol 2,Issue 2,May 2012

[56] Ali A Minai and Ronald D Williams, “Acceleration of Back Propagation through Learning Rate and Momentum Adaptation”, Proceedings of the International Joint Conference on Neural Networks, IJCNN, January1990, pp 676-679.

[57] T. Tollenaere, “SuperSAB: Fast adaptive back propagation with good scaling properties”, Neural Networks, 3, pp561-573, 1990.

[58] Martin Riedmiller Heinrich Braun, “A Direct Adaptive Method for Faster Back propagation Learning: The RPROP Algorithm”, IEEE, 1993, pp 586-591.Francisco, CA, 1993, pp. 586-591.

[59] R. Allard, J. Faubert, “The generalized no-decrease adaptive method for large-slow-learning neural networks”, Neural Networks

[60] M. Pfister and R. Rojas, Qrprop-a hybrid learning algorithm which adaptively includes second order information, Proc. 4th Dortmund Fuzzy Days, pp 55–62, 1994.

[61] N. K. Treadgold and T. D. Gedeon, Simulated Annealing and Weight Decay in Adaptive Learning: The SARPROP Algorithm, IEEE Trans. Neural Networks, 9, 4, pp 662–668, 1998.

[62] C. Igel, M. H&usken, Improving the Rprop learning algorithm, in: H. Bothe, R. Rojas (Eds.), Proceedings of the Second International Symposium on Neural Computation, NC 2000, ICSC Academic Press, Canada/Switzerland, 2000, pp. 115–121.

[63] C. Igel, and M. Husken, “Empirical evaluation of the improved Rprop learning algorithms”, Neurocomputing, 50, 2003, pp. 105-123.

[64] Aristoklis D. Anastasiadis et al., “An Efficient Improvement of the Rprop Algorithm

[65] Aristoklis D. Anastasiadis, George D. Magoulas, Michael N. Vrahatis, “A New Learning Rates Adaptation Strategy for the Resilient Propagation Algorithm”, ESANN’2004 proceedings - European Symposium on Artificial Neural Networks Bruges (Belgium), 28-30 April 2004, d-side publi., ISBN 2-930307-04-8, pp. 1-6

[66] Aristoklis D. Anastasiadisa, George D. Magoulasa, Michael N. Vrahatis, “New globally convergent training scheme based on the resilient propagation algorithm”, Elsevier. Neurocomputing 64 (2005), pp 253–270.

[67] Navon and Legler, “Conjugate gradient method for large scale minimization in meterology”,Monthly weather review, Vol 115 ,Aug ,1987,pp 1479-1502

[68] Johansson, E., F. Dowla and D. Goodman (1991). “Back-propagation learning for multilayer feed-forward neural networks using the conjugate gradient method, “International Journal of Neural Systems, vol. 2, pp. 291-302.

[69] M F Moller, “Scaled Conjugate gradient for fast supervised learning”, Elsevier, Neural networks, Volume 6, pp 525-533, 1993.

[70] Andreda Motta Salles Barreto , Charles W. Anderson , “Restricted gradient-descent algorithm for value-function approximation in reinforcement learning, Elsevier, Artificial Intelligence 172 (2008) pp 454–482.

[71] R. Battiti and F. Masulli. Bfgs optimization for faster and automated supervised learning. INNC 90 Paris, International Neural Network Conference,, pages 757–760, 1990.

[72] Roberto Battiti, “Accelerated Backpropagation Learning_ Two Optimization Methods, Complex Systems,volume 3,1989, pp 331-342

[73] Raymond L Watrous “Learning Algorithms for Connectionist Networks_ Applied Gradient Methods of Nonlinear Optimization”,Proceedings of the IEEE,First International Conference on Neural Networks, June1987,volume II,619-627.

[74] J E Dennis,J J More, “Quasi Newton Methods,Motivation and Theory”,SIAM Review,Vol 19.,No.1 9Jan 1977),pp 46-89.

[75] Aristidis Likas , Andreas Stafylopatis, “ Training the random neural network using quasi-Newton methods”,Elsevier, European Journal of Operational Research 126 (2000) , pp 331-339.

[76] Hao Yu, Bogdan M. Wilamowski ,Member IEEE “Levenberg–Marquardt Training”.

[77] Amir Abolfazl Suratgar, Mohammad Bagher Tavakoli, and Abbas Hoseinabadi, “Modified Levenberg-Marquardt Method for Neural Networks Training”,World Academy of Science, Engineering and Technology International Journal of Computer, Information, Systems and Control Engineering Vol:1 No: 6, 2007, pp 1719-1721.

[78] Yixin Chen and Bogdan M. Wilamowski, “TREAT: A Trust-Region-based Error-AggregatedTraining Algorithm for Neural Networks”.

[79] P Sehgal et al., “Minimization of Error in Training a Neural Network Using Gradient Descent Method”, International Journal of Technical Research (IJTR),Vol 1,Issue 1,Mar-Apr 2012,pp 10-12.

[80] M. T. Hagan and M. B. Menhaj, “Training feed forward networks with the Marquardt algorithm,” IEEE Trans. Neural Netw., 1994; vol. 5, no. 6, pp. 989–993.

[81] Bogdan M. Wilamowski, Nicholas Cotton, OkyayKaynak “Neural Network Trainer with Second Order Learning Algorithms” INES 2007 11th International Conference on Intelligent Engineering Systems Budapest, Hungary 2007 IEEE.

[82] Bogdan M. Wilamowski , “Efficient algorithm for training neural networks with one hidden layer”IEEE,1999.

[83] Bogdan M. Wilamowski et al., “An Algorithm for Fast Convergence in Training Neural Networks” ,IEEE,2001.

[84] X. Yao and Y. Liu, “A new evolutionary system for evolving artificial neural networks,”IEEE Trans. Neural Netw., vol. 8, no. 3, pp. 694-713, May 1997.

[85] Yao, “Evolving Artificial Neural Networks”, Proceedings of the IEEE, Vol. 87, No. 9, September 1999, pp 1423-1447.

[86] Back,T. 1996:Evolutionary algorithms in theory and practice. Oxford University Press

[87] A E Eiben, G Rudolph, Theory of Evolutionary algorithms: a bird’s eye view ,theoretical Computer Science 229 (1999),3-9.

[88] Piero P. Bonissone, GE Corporate Research & Development ,Evoultionary algorithms

[89] A.E. Eiben , M. Schoenauer, Evolutionary Computing, Information Processing Letters 82 (2002), 1–6.

[90] J. H. Holland. Adaptation in natural and artificial systems. Ann Arbor: The University of Michigan Press, 1975.

[91] Melanie Mitchell , Introduction to Genetic Algorithms, First MIT Press paperback edition, 1998.

[92] David J. Montana and Lawrence Davis, “Training Feedforward Neural Networks Using Genetic Algorithms”, Machine Learning, pp 762-767.

[93] David J. Montana, “Neural Network Weight Selection Using Genetic Algorithms”, pp 1-17.

[94] J.A. Cabrera *et al.,* Optimal synthesis of mechanisms with genetic algorithms, Mechanism and Machine Theory 37 (2002) 1165–1177

[95] Binitha S, S Siva Sathya, A survey of bio inspired optimization algorithms ,IJSCE,Vol 2,Issue 2,May 2012.

[96] Raymond Chiong, Member, IEEE, and Ooi Koon Beng, Member, IEAust 43, A Comparison between Genetic Algorithms and Evolutionary Programming based on Cutting Stock Problem, Engineering Letters, 14:1, EL_14_1_14 ,Feb 2007

[97] John Koza,1992 .Genetic Programmimg, Cambridge, MIT Press.

[98] Byoung_Tak Zhang et al.., “Evolving Optimal Neural Networks Using Genetic Algorithms with Occam_s Razor”, Complex Systems,7(3),199-220,1993

[99] Hans-George Beyer and Hans-Paul Schwefel ,Evolution Strategies-A cpmprehensive Introduction ,Natural Computing 1: 3–52, 2002

[100] Th. Bäck, H.-P. Schwefel, An overview of evolutionary algorithms for parameter optimization, Evolutionary Comput. 1 (1)(1993) 1–23

[101] Mehrdad Dianati, Insop Song, and Mark Treiber, An Introduction to Genetic Algorithms and Evolution Strategies, University of Waterloo, Ontario, N2L 3G1, Canada

[102] R. Storn, K. Price, Differential evolution – a simple and Efficient heuristic for global optimization over spaces, Continuous Journal of Global Optimization 11 (1997) 341–359.

[103] Kelly Fleetwood, An Introduction to Differential Evolution.

[104] Storn, R.: Differential Evolution Research – Trends and Open Questions, Advances in Differential Evolution, Springer-Verlag Berlin Heidelberg, pp. 1-31 (2008)

[105] Abazar Shamekhi ,Improved DE optimization algorithm, IJRRAS ,Vol 15 Issue 2, May 2013, pp 132-145

[106] Qin, A. K., Huang, V. L. and Suganthan, P. N., “Differential Evolution Algorithm with Strategy Adaptation for Global Numerical Optimization”, IEEE Transactions on Evolutionary Computations, Vol. 13 (2), pp. 398 – 417, 2009.

[107] Ali, M. M. and A. Torn: Population set based global optimization algorithms: Some modifications and numerical studies, Comput. Oper. Res., 31, 10, 1703–1725 (2004).

[108] Radha Thangraj, Millie Pant et al.,”Differential Evolution using a Localized Cauchy Mutation Operator”, Scientific Network for Innovation and Research Excellence, USA

[109] M.G. Epitropakis, V.P. Plagianakos and M.N. Vrahatis, “Higher-Order Neural Networks Training Using Differential Evolution”, ICNAAM – 2006.

[110] Y. Karali, Sibarama Panigrahi, H. S. Behera, “A Novel Differential Evolution Based Algorithm for Higher Order Neural Network Training”, *Journal of Theoretical and Applied Information Technology* 20th October 2013. Vol. 56 No.2 © 2005 - 2013 JATIT & LLS. All rights reserved.

[111] Sibarama Panigrahi, Ashok Kumar Bhoi, Yasobanta Karali, “ A Modified Differential Evolution Algorithm trained Pi-Sigma Neural Network for Pattern Classification “,International Journal of Soft Computing and Engineering (IJSCE) ISSN: 2231-2307, Volume-3, Issue-5, November 2013.

[112] A K Quin, P. N. Suganthan “Self-adaptive Differential Evolution Algorithm for Numerical Optimization”, IEEE Transactions on Evolutionary Computations, Vol.13 (2), pp. 398 – 417, 2009.

[113] Kennedy, J.; Eberhart, R. , “Particle Swarm Optimization”. Proceedings of IEEE International Conference on Neural Networks. ,1995, pp. 1942–1948.

[114] Min-Rong Chen,Xia Li, Xi Zhang, Yong-Zai Lu,”A novel particle swarm optimizer hybridized with extremal optimization”

[115] Simon Garnier · Jacques Gautrais · Guy Theraulaz, “,The biological principles of swarm intelligence”. Springer Science + Business Media, LLC, Swarm Intell (2007) 1 ,pp3–31.

[116] Jaco F. Schutte,”The Particle SwarmOptimization Algorithm”,EGM 6365 – Structural Optimization Fall 2005

[117] Riccardo Poli, James Kennedy, Tim Blackwell: Particle swarm optimization. Swarm Intelligence 1(1),(2007),pp 33-35

[118] Zhi-Hui Zhan, Jun Zhang, Yun Li and Henry Shu-Hung Chung, “Adaptive Particle Swarm Optimization”, IEEE Transactions on Systems, Man, and Cybernetics—Part B: Cybernetics, Vol. 39, No. 6, December 2009, pp -1362-1381.

[119] R. C. Eberhart and Y. Shi, “Comparing inertia weights and constriction factors in particle swarm optimization,” in Proc. IEEE Congr. Evol. Comput., 2000, pp. 84–88.

[120] M Dorigo et al.,” Ant System:Optimization by a colony of cooperating agents “,IEEE transactions on Sysytems,man,Cybernetics-Part B ,26(1),pp 29-41,1996.

[121] Thomas Stutzle , Holger H. Hoos “MAX –MIN Ant System”, Elsevier,Future Generation Computer Systems 16 (2000) 889–914.

[122] M. Dorigo and L. Gambardella, “Ant colony system: A cooperative learning approach to the traveling salesman problem,” IEEE Transactions on Evolutionary Computation, vol. 1, no. 1, pp. 53–66, 1997.

[123] M. Dorigo and T. St ¨utzle, Eds., Ant colony optimization. London, England: MIT Press, 2004.

[124] Christian Blum, “Ant colony optimization: Introduction and recent trends”,Physics of Life Reviews –Vol 2,Science direct-Elsevier, 2005,pp353–373

[125] Marco Dorigo, Mauro Birattari, and Thomas St¨utzle,”Ant Colony Optimization-Artificial Ants as a Computational Intelligence Technique”,IRIDIA – Technical Report Series TR/IRIDIA/2006-023,ISSN 1781-3794,September 2006

[126] D. Karaboga, B. Basturk, “A powerful and efficient algorithm for numerical function optimization: artificial bee colony (ABC) algorithm”, Journal of Global Optimization , Volume 39 Issue 3, November 2007 ,pp 459 – 471.

[127] Yana Mazwin Mohmad Hassim and Rozaida Ghazali, “Training a Functional Link Neural Network Using an Artificial Bee Colony for Solving a Classification Problems”.

[128] Dervis Karaboga · Bahriye Basturk, “A powerful and efficient algorithm for numerical function optimization: artificial bee colony (ABC) algorithm”, Journal Glob Optim (2007) 39, pp 459–471.

[129] X.S Yang, “Fire fly algorithm for multimodal optimization “, proceedings of the stochastic Algorithms. Foundations and Applications(SAGA 109) vol.5792of Lecture notes in Computer Sciences Springer,Oct.2009

[130] S . Nandy et al., “agent based adaptive FA backpropagation neural network traning method for dynamic systems”,IEEE Proceddings,2012.

[131] S Kumar et al.,”Fuzzy Model Identification: A Firefly Optimization Approach”,,International Journal of Computer Applications (0975 – 8887) Volume 58– No.6, November 2012 .

[132] Li, L.X., Shao, Z.J., Qian, J.X.: An Optimizing Method Based on Autonomous Animate: Fish Swarm Algorithm. Proceeding of System Engineering Theory and Practice, 2002,pp. 32–38

[133] Ying Wu*, Xiao-Zhi Gao**, and Kai Zenger, “Knowledge-based artificial fish-swarm algorithm”, Preprints of the 18th IFAC World Congress Milano (Italy) August 28 - September 2, 2011

[134] Reza Azizi, “Empirical Study of Artificial Fish Swarm Algorithm”, International Journal of Computing, Communications and Networking, Volume 3, No.1, January – March 2014.

[135] Eusuff MM and K.E Lansey; Optimization of water distribution network design using SFLA(2003)

[136] Juan Lin, Yiwen Zhong, “Modified Shuffled Frog-leaping Algorithm with Dimension by Dimension Improvement,Journal of Computers”, Vol 9, No 10 (2014), pp 2352-2358

[137] G. G. Samuel and C. C. A. Rajan, “A Modified Shuffled Frog Leaping Algorithm for Long-Term Generation Maintenance Scheduling” Proceedings of the Third International Conference on Soft Computing for Problem Solving, Advances in Intelligent Systems and Computing 258, Springer India 2014

[138] X.S. Yang “ A new metaheuristic bat-inspired algorithm”, Nature Inspired Cooperative Strategies for Optimization (NICSO 2010),pp 65–74, 2010.

[139] Fister Jr., I., Fister, D., and Yang, X. S., (2013). A hybrid bat algorithm, Elekrotehniski Vestnik (English Edition).

[140] Fister, I., Fister Jr., I., Yang, X. S., and Brest, J., (2013). On the representation of individuals using quaternions in swarm intelligence and evolutionary computation, IEEE Trans. Evol. Computation, (2013, submitted).

[141] A. Kaveh and P. Zakian , “Enhanced Bat Algorithm For Optimal Design Of Skeletal Structures”, Asian Journal Of Civil Engineering (BHRC) Vol. 15, No. 2 (2014), 179-212.

[142] Xin-She Yang, Bat algorithm: literature review and applications, Int. J. Bio-Inspired Computation, Vol. 5, No. 3, pp. 141–149 (2013).

[143] Passino KM, “Biomimicry of Bacterial Foraging”, IEEE Control System Magzine, Vol. 22, 2002, pp. 52-67.

[144] Y. liu,K M Passino, “Biomimicry of Social Foraging Bacteria for Distributed Optimization: Models, Principles, and Emergent Behaviors”, Journal of Optimization Theory and Applications: Vol. 115, No. 3, Dec 2002, pp. 603–628.

[145] E M ontes,B H Ocana, “Modified Bacterial Foraging Optimization for Engineering Design”. Intelligent Engineering Systems Through Artificial Neural Networks, Artificial Neural Networks in Engineering Conference -19th,2009, pp357-364

[146] R Majhi,G Panda et al., “ Efficient prediction of stock market indices using adaptive bacterial foraging optimization (ABFO) and BFO based techniques”, Expert Systems with Applications, Volume 36, Issue 6, pp 10097–10104, August 2009.

[147] D.H. Kim et al. A hybrid genetic algorithm and bacterial foraging approach for global optimization,Science direct,/ Information Sciences 177 (2007) 3918–3937

[148] Swagatam Das1, Arijit Biswas1, Sambarta Dasgupta1, and Ajith Abraham2,”Bacterial Foraging Optimization Algorithm: Theoretical Foundations, Analysis, and Applications”, Studies in Computational Intelligence Volume 203, 2009, pp 23-55

[149] S S Patnaik et al., “ A Review of Bacterial Foraging Optimization and Its Applications” National Conference on Future Aspects of Artificial intelligence in Industrial Automation (NCFAAIIA 2012) Proceedings published by International Journal of Computer Applications® (IJCA) pp 9-12.

[150] Glover and laguna, “Tabu Search”, Kluwer Academic Publishers,Tutorial interface,pp 1-18.

[151] Jerzy Balicki , “Tabu Programming for Multiobjective Optimization Problems”, IJCSNS International Journal of Computer Science and Network Security, Vol. 7 No.10, October 2007,pp 44-51

[152] Jaeggi et al., “Multiobjective tabu Search for contrained optimization problems”.

[153] Fred Glover and Belen Melian, “381 –TS-Intro to TS,Chapter Tabu Search .

[154] X. S. Yang and S. Deb, “Cuckoo Search via Levy Flights”, Proceedings of World Congress on Nature & Biologically Inspired Computing, (2009), pp. 210-225.

[155] Ehsan Valian, Shahram Mohanna, “Improved Cuckoo Search Algorithm for Global Optimization”, IJCIT-2011-Vol.1-No.1 Dec. 2011,pp 31-44.

[156] F Ahmadi et al., “Eurygaster Algorithm: A New Approach to Optimization”, International Journal of Computer Applications (0975 – 8887) Volume 57– No.2, November 2012

[157] Amir Hossein Gandomi et al., “Krill herd: A new bio-inspired optimization algorithm” Communications in Nonlinear Science and Numerical Simulation, Volume 17, Issue 12, December 2012, Pages 4831–4845

[158] X.-S. Yang. Engineering optimizations via nature-inspired virtual bee algorithms. volume 3562, pages 317–323, 2005.

[159] A.R. Mehrabian, C. Lucas, A novel Numerical Optimization Algorithm Inspired from Weed Colonization, Ecological Informatics, 2006,vol 1.pp-355-366

[160] Zhi YIN , Mi WEN , Chunming YE, “Improved Invasive Weed Optimization Based on Hybrid Genetic Algorithm”, Journal of Computational Information Systems 8, (2012),pp 3437–3444

[161] Simon, D. 2008. Biogeography-Based Optimization. IEEE Transactions on Evolutionary Computation. 12(6) : 702–713.

[162] MacArthur, R. and Wilson, E. 1967. The theory of island biogeography., Princeton, NJ: Princeton University Press, p 203.

[163] Ammu P K, “Biogeography-Based Optimization - A Survey”, IJECSE, Volume2, Number 1,pp 154-160.

[164] Parvinder Bhalla, “FUZZY MODEL IDENTIFICATION: SOME NEW SOFT COMPUTING BASED APPROACHES “Chapter 4,Proceedings of the faculty Development Program on” Soft Computing & its Applications (SCA-2013)”, December17-19, 2013.

[165] Ala’a Abu-Srhan and Essam Al Daoud, “ A Hybrid Algorithm Using a Genetic Algorithm and Cuckoo Search Algorithm to Solve the Traveling Salesman Problem and its Application to Multiple Sequence Alignment”, International Journal of Advanced Science and Technology Vol.61, (2013), pp.29-38

[166] D. Dasgupta, Artificial Immune Systems and Their Applications, Springer, Berlin, 1999

[167] J.R. Al-Enezi,, "ARTIFICIAL IMMUNE SYSTEMS – MODELS, ALGORITHMS AND APPLICATIONS",IJRRAS 3 (2) , May 2010

[168] Leandro Nunes deCastro lnunes et al., " ARTIFICIAL IMMUNE SYSTEMS-Part 1 basic theory and applications", Technical Report RT – DCA 01/99 December, 1999

[169] Geem, Z. ," Music inspired Harmony search in water pump switching problem. Lecture Notes in Computer Science, 3612:751.

[170] Geem, Z. (2008). Harmony search optimisation to the pump-included water distribution network design. Civil Engineering and Environmental Systems, 99999(1):1{1.

[171] Xin-She Yang1, "Harmony Search as a Metaheuristic Algorithm

[172] P. Chakraborty et al. , "An Improved Harmony Search Algorithm with Differential Mutation Operator", Fundamenta Informaticae 95 (2009) 1–26.

[173] S.He,Q.H.Wu, Senior Member IEEE and J.R Saunders.A novel group search optimizer inspired by Animal Behavioral Ecology;2006,IEEE Congress on Evolutionary Computation,1272-1278

[174] Upeka Premaratne , Jagath Samarabandu, and Tarlochan Sidhu, ¯A New Biologically Inspired Optimization Algorithm ,Fourth International Conference on Industrial and Information Systems, ICIIS 2009,28-31 December 2009, Sri Lanka.

[175] R. G. Reynolds, "Cultural algorithms: theory and applications," in New Ideas in Optimization, D. W. Corne, M. Dorigo, and F. Glover, Eds., pp. 367–378, McGraw-Hill, Maidenhead, UK, 1999.

[176] A. Mucherino and O. Seref, "Monkey search: a novel metaheuristic search for global optimization," in Proceedings of the Conference on Data Mining, Systems Analysis, and Optimization in Biomedicine, pp. 162–173, Gainesville, Fla, USA, March 2007.

[177] X. S. Yang, "Flower pollination algorithm for global optimization," in Unconventional Computation and Natural Computation, vol. 7445 of Lecture Notes in Computer Science, pp. 240– 249, 2012.

[178] Rui Wang and Yongquan Zhou ,Research Article , "Flower Pollination Algorithm with Dimension by Dimension Improvement", Mathematical Problems in Engineering Volume 2014, Article ID 481791, 9 pages.

[179] Murase, H. Finite element inverse analysis using a photosynthetic algorithm. Computers and Electronics in Agriculture, 29 (2000) 115–123.

[180] Yang, X. S., New enzyme algorithm, Tikhonov regulation and inverse parabolic analysis, in Advances in Computational Methods in Science and Engineering, Lecture Series on Computer and Computer Sciences, ICCMSE 2005, Eds. T. Simons and G. Maroulis, 4, 1880-1883 (2005).

[181] Dirk P. Kroese, Reuven Y. Rubinstein, and Peter W. Glynn, "The Cross-Entropy Method for Optimization"

[182] Dirk P. Kroese1, Reuven Y. Rubinstein2, and Peter W. Glynn3 ,"The Cross-Entropy Method for Estimation",Machine Learning:theory and applications,Handbook of Statistics, Vol. 31 ISSN: 0169-7161,Else vier,2013.

[183] Shah-Hosseini, H, " Problem solving by intelligent water drops", Proceedings of IEEE Congress on Evolutionary Computation,2007, pp. 3226-3231, Swissotel The Stamford, Singapore, September.

[184] Shah-Hosseini, H, " Intelligent water drops algorithm: a new optimization method for solving the multiple knapsack problem, Int. Journal of Intelligent Computing and Cybernetics, Vol. 1, No. 2, 2008, pp. 193-212.

[185] Shah-Hosseini, H' "The Intelligent Water Drops algorithm: A nature-inspired swarm-based optimization algorithm. Int. J. Bio-Inspired Computation, Vol. 1, Nos. 1/2, 2009, pp. 71–79.

[186] Hamed Shah-Hosseini, "Optimization with the Nature-Inspired Intelligent Water Drops Algorithm" ,Evolutionary Computation, Book edited by: Wellington Pinheiro dos Santos, ,ISBN 978-953-307-008-7, Chapter 16,pp. 572, October 2009,

[187] Birbil SI, Fang SC (2003). Electromagnetism-like mechanism for global optimization. J. Glob. Optim.,Vol. 25, Issue 3,March 2003,pp 263-282.

[188] Wu, P. T., Yang, W. H., & Wei, N. C. (2004). An electromagnetism algorithm of neural network analysis-an application to textile retail operation. Journal of the Chinese Institute of Industrial Engineers, 21(1), 59–67

[189] Rocha, A. M. A. C., & Fernandes, E. M. G. P. (2008b). On charge effects to the electromagnetism-like algorithm. In Proceedings of 20th international conference/Euro mini conference on continuous optimization and knowledge-based technologies (EurOPT 2008), May 20–23. (pp: 198–203). Neringa, Lithuania.

[190] Rocha, A. M. A. C., & Fernandes, E. M. G. P. (2009a). Self-adaptive penalties in the electromagnetism-like algorithm for constrained global optimization problems. In 8th World congress on structural and multidisciplinary optimization. June 1–5,2009. Lisbon, Portugal

[191] Rocha, A. M. A. C., & Fernandes, E. M. G. P. Implementation of the electromagnetism-like algorithm with a constraint-handling technique for engineering optimization problems,Proceedings of the 8th international conference on hybrid intelligent systems (HIS 2008,2008a, (pp. 690–695).

[192] Ching-Hung Lee et al.,"A Hybrid Algorithm of Electromagnetism-like and for Recurrent Neural Fuzzy Controller Design" Proceedings of the International MultiConference of Engineers and Computer Scientists 2009 Vol I,IMECS 2009, March 18 - 20, 2009, Hong Kong.

[193] Ching-Hung Lee et al.., "An Improved Electromagnetism-like Algorithm for Recurrent Neural Fuzzy Controller Design" International Journal of Fuzzy Systems, Vol. 12, No. 4, December 2010

[194] Ali, M. M., & Golalikhani, M. (2010). An electromagnetism-like method for nonlinearly constrained global optimization. Computers and Mathematics with Applications, 60(8), 2279–2285.

[195] Nai-Chieh Wei, "An electromagnetism-like mechanism for solving cell formation problems", Scientific Research and Essays Vol. 7(9), March, 2012, pp. 1022-1034

[196] Chunjiang Zhang, Xinyu Li, Liang Gao , Qing Wu, "An improved electromagnetism-like mechanism algorithm for constrained optimization" Expert Systems with Applications 40 (2013) 5621–5634

[197] Esmat Rashedi, Hossein Nezamabadi-pour , Saeid Saryazdi , "GSA: A Gravitational Search Algorithm", Information Sciences 179 (2009),pp 2232–2248

[198] Fahlman, S. and C. Lebiere (1990). "The cascade-correlation learning architecture," in Advances in Neural Information Processing Systems 2, D. Touretzky, Ed. San Mateo, CA: Morgan Kaufmann, pp.524-532.

[199] Norlina Mohd Sabri, Mazidah Puteh, and Mohamad Rusop Mahmoo, "A Review of Gravitational Search Algorithm", Int. J. Advance. Soft Comput. Appl., Vol. 5, No. 3, November 2013

[200] J. Zhang, K. Liu, Y. Tan, X. He, Random black hole particle swarm optimization and its application, in: 2008 IEEE International Conference Neural Networks and Signal Processing, ICNNSP, 2008, pp. 359–365.

[201] A. Hatamlou , "Black hole: A new heuristic optimization approach for data clustering", Elsevier, Information Sciences 222 (2013) ,pp 175–184.

[202] S. Kirkpatrick; C. D. Gelatt; M. P. Vecchi ,"Optimization by Simulated Annealing", , Science, New Series, Vol. 220, No. 4598. (May 13, 1983), pp. 671-680.

[203] Sriram G. Sanjeevi, A. Naga Nikhila,Thaseem Khan and G. Sumathi, "Hybrid PSO-SA algorithm for training a Neural Network for Classification", International Journal of Computer Science, Engineering and Applications (IJCSEA) Vol.1, No.6, December 2011

[204] Kenneth D. Boese and Andrew B. Kahng , "Simulated Annealing of Neural Networks: the Cooling Strategy Reconsidered".

[205] Randall S. Sexton *et al.*, "Beyond Backpropagation: Using Simulated Annealing for Training Neural Networks".

[206] Eksin I, Erol OK, "Evolutionary algorithm with modifications in the reproduction phase. IEE Proc Softw are 2001; 148(2):pp 75–80.

[207] Osman K. Erol*, Ibrahim Eksin,"A new optimization method: Big Bang–Big Crunch", Elsevier publications,Advances in Engineering Software 37 (2006) ,pp106–111.

[208] Kaveh,Talatahari,"Size optimization of space trusses using Big Bang-Big Crunch algorithm", Journal Elsevier, Computers and Structures, Volume 87 Issue 17-18, Sept, 2009,pp 1129-1140

[209] Bilal Alatas,"Uniform Big Bang–Chaotic Big Crunch optimization", Communications in Non linear Science and Numerical Simulation,Sept 2011:16(9),pp 3696-3703.

[210] Behrooz Hassani , Mostafa Assari , Morteza Kazemi Torbaghan ,"An Improved Big Bang – Big Crunch Algorithm For Size Optimization of Trusses", 9th International Congress on Civil Engineering, May 8-10, 2012.

[211] Hakki Murat Genc , Osman K. Erol, Ibrahim Eksin, "Big bang-big crunch optimization algorithm with local directional moves" Turkish Journal of Electrical Eng & Computer Science,(2013) 21,pp1359 -1375

[212] Kaveh,Talatahari, "A New Hybrid Meta-heuristic for Optimum Design of Frame Structures" Asian Journal of Civil Engineering (Building and Housing) Vol. 13, No. 6 (2012), pp705-717

[213] K RamaKrishna, G.Ramkumar and R Sambasiva Rao, "Evolution of Mimics of Algorithms of Nature (E-man), Part 5#: Tutorial on Big_Bang–Big_Crunch algorithm", Journal of Applicable Chemistry ,2 (6), pp1413-1458, 2013.

[214] Kaveh,Talatahari,A new hybrid optimization algorithm for recognition of hysteretic non-linear systems, Engineering, July 2013, Volume 17, Issue 5, pp 1099-1108.

[215] A. Kaveh, S. Talatahari and M.T. Alami, A New Hybrid Meta-heuristic for Optimum Design of Frame Structures, Asian *Journal of Civil Engineering (Building and Housing)* Vol. 13, No. 6 (2012), pp 705-717.

[216] Shakti Kumar, Nitika Ohri, Savita Wadhvan , “ANN based design of rapid battery charger”, Trends Of Computational Techniques in Engineering Oct 15-16, , SLIET, Longowal Punjab,2004, pp 129-132.

[217] Shakti kumar,Amar Singh,” Nature Inspired Computing With Emphasis to Fuzzy Model Identification”,Compuational Intelligence Lab.

[218] Shakti Kumar, Sukhbir Singh Walia, Amar Singh, “Parallel Big Bang-Big Crunch Algorithm”, International Journal of Advanced Computing, Vol.46, Issue.3 1, Sept 2013.

[219] Shakti Kumar, Sukhbir Singh Walia, A Kalanidhi Nov 2013 b “Fuzzy Model Identification: A New Parallel BB-BC Optimization Based Approach” International Journal o Electronics and Communication Engineering. Vol 2, Issue 5,Nov. 2013., pp 167-178.

[220] Shakti Kumar, Parvinder Bhalla, AP Singh, “Fuzzy Rulebase Generation from Numerical Data using Big Bang-Big Crunch Optimization”, IE (I)Journal -ET, Volume 91, January 2011 pp 1-8 .

[221] Gurpreet S. Sandhu and Kuldip S. Rattan, “Design of a Neuro Fuzzy Controller”, 2004.

[222] Ajith Abraham, “Neuro Fuzzy Systems: State-of-the-art Modeling Techniques”, School of Computing & Information Technology ,Monash University, Churchill 3842, Australia.

[223] Alfonso Martínez del Hoyo Canterla, “NEURO-FUZZY CONTROL”, Soft Computing, 2005University of Iceland.

[224] S. N. BALAKRISHNAN, R. D. WEIL, “Neurocontrol: A Literature Survey”, Elsevier, Mathl. Comput. Modelling Vol. 23, No. l/2, March1996, pp. 101-117.

[225] Jang et al., “Neuro fuzzy modelling and control”,1995

[226] Azar, Ahmad Taher, “Adaptive Neuro-Fuzzy Systems”, Fuzzy Systems, Book edited by: Ahmad Taher Azar, ISBN 978-953-7619-92-3, pp. 216, February 2010, INTECH, Croatia, pp 85-110.

[227] Jyh-Shing Roger Jang, “ANFIS : Adap tive-Ne twork-Based Fuzzy Inference System, IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS, VOL. 23, NO. 3, MAYIJUNE 1993, pp 665-685.

[228] Marcus Pfister ,Raul Rojas , “Hybrid Learning Algorithms for Neural Networks, The adaptive Inclusion of Second Order Information”,1995

[229] Remy Allard, Jocelyn Faubert, “Neural Networks: Different problems require different learning rate adaptive methods”, Proceedings of SPIE, 2004.

[230] G V.R. Sagar and 2Dr. S. Venkata Chalam , “Evolutionary Algorithm for Connection Weights in Artificial Neural Networks”, International Journal of Electronics and Communication Engineering. ISSN 0974-2166 Volume 4, Number 5 (2011), pp. 517-525

[231] A Jagtap et al., “Efficient Method for Optimizing Artificial Neural Network Using „Quantum-Based Algorithm”, IJARCSSE,Volume 4, Issue 6, June 2014 ,pp 692-700.

[232] Hiroaki Kitano , “Empirical Studies on the Speed of Convergence of Neural Network Training using Genetic Algorithms”,Machime Learning, AAAI-90 Proceedings,pp 789-795.

[233] Jatinder N.D. Gupta, Randall S. Sexton, “Comparing backpropagation with a genetic algorithm for neural network training”,Elsevier, Omega 27 (1999), pp 679-684.

[234] Zhen-Guo Che, Tzu-An Chiang and Zhen-Hua Che, “FEED-FORWARD Neural Networks Training: A Comparison Between Genetic Algorithm and Back-propagation Learning Algorithm” ,International Journal of Innovative Computing, Information and Control, ICIC International 2011 ISSN 1349-4198,Volume 7, Number 10, October 2011, pp. 5839-5850.

[235] Paul Batchis , “An Evolutionary Algorithm for Neural Network Learning using Direct Encoding”

[236] Asha Gowda Karegowda , Darshan M “Optimizing Feed Forward Neural Network Connection Weights Using Artificial Bee Colony Algorithm” ,International Journal of Advanced Research in Computer Science and Software Engineering, Volume 3, Issue 7, July 2013.

[237] Enrique Alba and J. Francisco Chicano, “Training Neural Networks with GA Hybrid Algorithms”

[238] J. Zhang, J. Zhang, T. Lok & M.Lyu, (2007) “ A hybrid particle swarm optimization backpropagation algorithm for feed forward neural network training” , Applied Mathematics and Computation, Vol.185, pp. 1026-1037.

[239] L. Yan-Peng, W. Ming-Guang, and Q. Ji-Xin, “Evolving neural networks using the hybrid of ant colony optimization and bp algorithm,” in Advances in Neural Networks - 3rd International Symposium on Neural Networks, ser. LNCS, vol. 3971. Springer-Verlag, 2006, pp. 714–722.

[240] Michalis Mavrovouniotis, Shengxiang Yang , “Evolving neural networks using ant colony optimization with pheromone trail limits”,IEEE,Computational Intelligence (UKCI), 13th UK Workshop, September 2013.

[241] Michalis Mavrovouniotis, Shengxiang Yang ,"Training neural networks with ant colony optimization algorithms for pattern classification", Centre for Computational Intelligence (CCI),May 2014.

[242] Huadong Chen et al. , "A Hybrid of Artificial Fish Swarm Algorithm and Particle Swarm Optimization for Feedforward Neural Network Training"

[243] Sudarshan Nandy , Partha Pratim Sarkar and Achintya Das, "Training A Feed-forward Neural Network with Artificial Bee Colony Based Backpropagation Method", International Journal of Computer Science & Information Technology (IJCSIT) Vol 4, No 4, August 2012,pp 33-46.

[244] Amir Hatampour, Rasul Razmi and Mohammad Hossein Sedaghat ,"Improving Performance of a Neural Network Model by Artificial Ant Colony Optimization for Predicting Permeability of Petroleum Reservoir Rocks, Middle-East Journal of Scientific Research 13 (Vol 9), ISSN 1990-9233, pp 1217-1223, 2013

[245] Yun Cai, "Artificial Fish School Algorithm Applied in a Combinatorial Optimization Problem", I.J. Intelligent Systems and Applications, 2010, 1, pp 37-43.

[246] B. Basturk, D. Karaboga, An artificial bee colony (ABC) algorithm for numeric function optimization, in: IEEE Swarm Intelligence Symposium 2006, May 12–14, Indianapolis, IN, USA, 2006.

[247] D Karaboga, B Basturk, "On the performance of artificial bee colony (ABC) algorithm", Elesvier,Applied Soft Computing 8 (2008), pp 687–697.

[248] Dervis Karaboga *, Bahriye Akay, "A comparative study of Artificial Bee Colony algorithm", Applied Mathematics and Computation 214 (2009),pp 108–132

[249] V. Saishanmuga Raja and S.P. Rajagopalan, "A Comparative Analysis of Optimization Techniques for Artificial Neural Network in Bio Medical Applications", *Journal of Computer Science* 10 (1): pp 106-114, Science Publications, 2014

[250] D. Srinivasan, T. H. Seow, Particle swarm inspired evolutionary algorithm (PS-EA) for multi-criteria optimization problems,Evolutionary Multiobjective Optimization, Springer Berlin Hei-delberg, 2006, pp.147-165.

[251] A. Ghaffari , H. Abdollahi , M.R. Khoshayand , I. Soltani Bozchalooi ,A. Dadgar , M. Rafiee-Tehrani , "Performance comparison of neural network training algorithms in modeling of bimodal drug delivery", International Journal of Pharmaceutics 327 ,Science direct,Elsevier(2006) 126–138

[252] Zhang Y, Wu L and Wang S, " Bacterial Foraging Optimization Based Neural Network for Short term Load Forecasting", JCIS, Vol. 6(7), 2010, pp. 2099-210.

[253] Ivona Brajevic ,Milan tuba, "Training Feed-Forward Neural Networks Using Firefly Algorithm, Recent Advances in Knowledge Engineering and Systems Science, 12th International Conference on Hybrid Intelligent Systems (HIS),2012,pp156-161.

[254] R Giri.et al.,"A Modified Invasive Weed Optimization Algorithm for Training of Feed-Forward Neural Networks *"Systems Man and Cybernetics (SMC), 2010 IEEE International Conference on,* 10-13 Oct. 2010, pp 3166 – 3173.

[255] Harpreet Kaur , Gaganpreet Kaur, "A Survey on Comparison between Biogeography Based Optimization and Other Optimization Method", Volume 3, Issue 2, February 2013 ISSN: 2277 128X.

[256] S. Mirjalili, S. Z. Mohd Hashim, and H. Moradian Sardroudi, "Training feedforward neural networks using hybrid particle swarm optimization and gravitational search algorithm," Elsevier,Applied Mathematics and Computation, vol. 218, no. 22, (2012), pp. 11125–11137.

[257] Saeide Sheikhpour, Mahdieh Sabouri, and Seyed-Hamid Zahir, " A hybrid Gravitational Search Algorithm–Genetic Algorithm for neural network training,21 st Irananian Conference of Electrical Engineering.

[258] Bao-Chang Xu Ying-Ying Zhang, "An Improved Gravitational Search Algorithm for Dynamic Neural Network Identification, International Journal of Automation and Computing 11(4), August 2014, pp 434-440

[259] Saeed Ayat , Zabihollah Ahmad Pour , "Comparison between Artificial Neural Network Learning Algorithms for Prediction of Student Average considering Effective Factors in Learning and Educational Progress", Journal of mathematics and computer science 8 (2014), 215 – 225.

[260] Naseer mohammadi, Seyed Javad Merabedini, "Comparison of Particle Swarm optimization and back Propagation algorithms for training Feed Forward neural networks", Journal of mathematics and computer science 12 (2014), 113 – 123.

[261] Specht, D. F., (1991). "A General Regression Neural Network," I&E Transactions on Neural Networks, 2(6), 568-576.

[262] Iztok Fister Jr., Xin-She Yang, Iztok Fister, Janez Brest, Dusan Fister, "A Brief Review of Nature-Inspired Algorithms for Optimization", ELEKTROTEHNI SKI VESTNIK, English edition, 80(3): 1–7, 2013.

[263] Hifza Afaq and Sanjay Saini,"Swarm Intelligence based Soft Computing Techniques for the Solutions to Multiobjective Optimization Problems", *IJCSI International Journal of Computer Science Issues*, Vol. 8, Issue 3, No. 2, May 2011.

[264] Yang, X. S., Nature-Inspired Metaheuristic Algorithms,Second edition, Luniver Press (2010).

[265] Yang X. S, "Biology-Derived Algorithms in Engineering Optimization",2005, chapter 32,pp 585-592.

[266] K.M. Saridakis, A.J. Dentsoras , "Soft computing in engineering design – A review",Esevier, Advanced Engineering Informatics 22 (2008) 202–221.

[267] Agnieszka Krok, "The Development of Kalman Filter Learning Technique for Artificial Neural Networks" ,Journal of Telecommunication and information technology,2013,pp 16-21.

[268] S. Haykin (Ed.), Kalman Filtering and Neural Networks. New York:Wiley, 2001.

[269] Amir Beck, Marc Teboulle, "Mirror descent and nonlinear projected subgradient methods for convex optimization", Operations Research Letters 31 (2003) , pp 167 – 175.

[270] R. Rojas, " Neural Networks", Springer-Verlag, Berlin, 1996.

[271] Dan Simon, "Training radial basis neural networks with the extended Kalman filter",Elsevier, Neurocomputing 48 (2002), pp 455–475.

[272] Bernard Widrow, Aaron Greenblatt, Youngsik Kim, Dookun Park, "The No-Prop algorithm: A new learning algorithm for multilayer neural networks", Neural Networks 37 (2013), pp 182–188.

[273] A. Nemirovski, D. Yudin, Problem complexity and Method Effciency in Optimization, Wiley, New York, 1983.

[274] Pieter-Tjerk de Boer *et al.*., " A Tutorial on the Cross-Entropy Method ", Sept, 2003

[275] Nan Du *et al.,* "*An Artificial Fish Swarm Based Supervised Gene Rankaggregation Algorithm For Informative Genes Studies*, Proceedings of the IASTED International Conference November 7-9, 2011 Pittsburgh, USA, Computational Intelligence and Bioinformatics (CIB 2011).

[276] Dennis Weyland, "A Rigorous Analysis of the Harmony Search Algorithm - How the Research Community can be misled by a novel Methodology" International Journal of Applied Metaheuristic Computing,volume 1-2, April-June 2010, pages 50-60

[277] Amir Hatampour, Rasul Razmi and Mohammad Hossein Sedaghat ,"Improving Performance of a Neural Network Model by Artificial Ant Colony Optimization for Predicting Permeability of Petroleum Reservoir Rocks,Middle-East Journal of Scientific Research 13 (Vol 9), ISSN 1990-9233, pp 1217-1223, 2013,

[278] Martin Moller "Efficient Training of FeedForward Neural Networks, PhD dissertation,Dec 1993

[279] [230] G. D. Magoulas, M. N. Vrahatis, and G. S. Androulakis, "Improving the Convergence of the Backpropagation Algorithm Using Learning Rate Adaptation Methods", Neural Computation 11, 1769–1796 (1999).

[280] Hugo Larochelle , Yoshua Bengio et al., "Exploring Strategies for Training Deep Neural Networks", Journal of Machine Learning Research 1 (2009),pp 1-40.

[281] D. Randall Wilson, Tony R. Martinez, "The general inefficiency of batch training for gradient descent learning", Elsevier, Neural Networks 16 (2003) , pp 1429–1451.

[282] Xu Wang, "Method of Steepest Descent and its Applications", Nov 2008, pp 1-3.

[283] Kevin Judd , Leonard Smith , Antje Weisheimer , "Gradient free descent: shadowing, and state estimation using limited derivative information, Physica D 190 (2004), pp 153–166

[284] Howard Hua Yang, "The Efficiency and The Robustness of Natural Gradient Descent Learning Rule",pp 385-391.

[285] Ahmad Hashim Hussein Aal-Yhia, and Ahmad Sharieh , "An Energy Backpropagation Algorithm", Proceedings of the World Congress on Engineering WCE 2007, Vol I,July 2007

[286] Nazri Mohd Nawi, Abdullah Khan, M. Z. Rehman, " A New Levenberg Marquardt Based Back Propagation Algorithm Trained with Cuckoo Search",Elsevier, Procedia Technology 8C (2013) ,pp 18 – 24.

[287] B Widrow, M A Lehr., "30 Years of Adaptive Neural Networks: Perceptron, Madaline, and Backpropagation", Proceedings of The IEEE, Vol. 78, NO. 9, September 1990,pp 1415-1442.

[288] B M Williamowski "Neural Network architectures and learning",Proceedings of IEEE ,2003.

[289] George Nagy, "Neural Networks-Then and How", Letters, IEEE transactions on Neural Networks, Vol 2, No. 2,1991,pp 316-318.

[290] S. N. Balakrishnan, R. D. Weil, "Neurocontrol: A Literature Survey" ,Elsevier, Mathl. Comput. Modelling Vol. 23, No. l/2, March1996, pp. 101-117.

[291] Mike Schuster and Kuldip K. Paliwal, Member, IEEE, "Bidirectional Recurrent Neural Networks", IEEE Transactions On Signal Processing, Vol. 45, NO. 11, November 1997, pp 2673-2681.

[292] J. M. Benýtez, J. L. Castro, and I. Requena , "Are Artificial Neural Networks Black Boxes?", IEEE Transactions On Neural Networks, Vol. 8, NO. 5, September 1997, pp 1156- 1164.

[293] Juan A. Ramírez-Quintana, Mario I. Chacon-Murguia and Jose F. Chacon-Hinojos, "Artificial Neural Image Processing Applications: A Survey", Engineering Letters, 20:1, EL_20_1_09, Feb 2012

[294] K.F. Alvin, A.N. Robertson, G.W. Reich , K.C. Park , "Structural system identification: from reality to models", Elsevier, Computers and Structures 81 (2003),pp 1149–1176.

[295] Identification of Dynamic System Using Neural Network, The scientific journal FACTA UNIVERSITATIS ,Series: Architecture and Civil Engineering Vol.1, No 4, 1997 pp. 525 – 532.

[296] Gábor HORVÁTH, "Neural Networks in System Identification",Chapter 4,NATO ASI presentation (NIMIA) in Crem a Italy, 2002.

[297] Vesna M. Ranković *et al.*, "Identification of Nonlinear Models with Feedforward Neural Network and Digital Recurrent Network", FME Transactions,Belgrade (2008) 36, pp 87-92

[298]]Jyh-Shing Roger Jang, "ANFIS : Adap tive-Ne twork-Based Fuzzy Inference System, IEEE Transactions on Systems, Man, and Cybernetics, Vol. 23, NO. 3, May-June 1993 ,pp 665-685.

[299] Koushal Kumar , Gour Sundar Mitra Thakur , "Advanced Applications of Neural Networks and Artificial Intelligence: A Review", I.J. Information Technology and Computer Science,Vol 6, 2012, pp 57-68.

[300] Yoshihiro Yamamoto, "Identification and Control for Nonlinear Discrete Time Systems Using an Interconnected Neural Network", ICCAS 2005.

APPENDIX

Table 1
Ann Architectures with Learning Algorithms

Paradigm	*learning Rule*	*Architecture*	*Learning Algorithm*	*Application*
Supervised Training	Error Correction	Single layer or Multilayer Perceptron	Perceptron learning Algorithm (LMS), BP, Adaline and Medaline	Function approximation, Prediction and control
	Boltzmann	Recurrent	Boltzmann Learning algorithm	Pattern Classification
	Hebbian learning	Multilayer feed forward	Linear discriminant Analysis	Pattern classification, data Analysis
	Competitive Learning	Competitive	Learning Vector Quantization	within class categorization
		ART networks	ART map	Pattern Classification, within class categorization
Un-Supervised Training	Error Correction	Multilayer Perceptron	Sammon's Projection	Data Analysis
	Hebbian learning	feed forward or Competitive; Hopefield network	Principal Component analysis; Associative Memory Learning	Data Analysis, Data compression; Associative Memory
	Competitive Learning	Competitive	Vector Quantization	Categorization, Data Analysis
		ART networks	ART-I, ART-II	Categorization
Hybrid	Error Correction and Competitive.	RBF Networks	RBF Learning Algorithms	Pattern Classification, Function Approximation, Prediction & Control.

Table II
Classical Learning Approaches Classical Learning Approaches

Algorithm	*Author, Reference*	*Algorithm*	*Author, Reference*
EBP (Error Back Propagation) /Steepest Gradient Method	Werbos, 1974; Rumelhart, 1986[44]	Conjugate Gradient	Beale, 1972[67]
EBP with Momentum (MOM)	Rumelhart, 1986[45]	Quick propagation (QP)	Fahlmann, 1988[48]
Standard or Incremental Backpropagation (IBP),	Freeman and Skappura, 1991[46]	Levenberg-Marquardt algorithm (LM),	M. T. Hagan and M. B. Menhaj, 1994[76]
Batch Backpropagation (BBP)	Hagan et al., 1996 [47]	Modified LM	Bogdan M. Wilamowski, 1999[77]
Delta-Bar-Delta(DBD) (Adaptive learning)	R A Jacobs, 1988 [52][54]	Newton's method	Flectcher, 1975
Extended DBD	Minai and Williams, 1990,[56]	Quasi Newton Method-BFGS	Bryoden, Fletcher, Goldfarb, Shanno, 1970R. Battiti and F. Masulli.,1990 [71][72]
Super-SAB(SSAB) (Adaptive learning)	Tollenaere,1980, 1990[57]	Quasi Newton Method-DFP variant	Davidon,Fletcher, Powell, 1963[73]

contd. Table II

Algorithm	*Author, Reference*	*Algorithm*	*Author, Reference*
Generalized no-decrease adaptive method (GNDAM)	R. Allard, J. Faubert [59]	Scaled Conjugate Gradient(SCG)	M F Moller,1991[69]
Resilent PROP(RPROP)	Riedmiller and Braun, 1993[58]	Restricted Gradient	A M Salles Baretto,CW Anderson, 2008[70]
Qrprop	M. Pfister and R. Rojas, 1994[60]	Conjugate Gradient Algotithm with line search(CGL)	Johannson,Dorwla & Goodman,1990[68]
Diagonal Estimation Rprop–DERprop,	1998 [61],	TREAT algorithm	Y Chen, B M Williamowski[78]
SA-Rprop	N. K. Treadgold and T. D. Gedeon,1998[61]	Improved RPROP (iRPROP)	C Igel and M Husken,2003[62][63][64]
Grprop	Aristoklis D A et al., 2004[65][66]	No-Prop algorithm	Bernard Widrow et al.,2013[272]
mirror descent algorithm (MDA)	Nemirovsky and Yudin, 1983[273]	Entropic MDA (EMDA)	A. Beck, M. Teboulle, 2003 [269]
Cascade Corelation (CC) Learning	Fahlman, S. and C. Lebiere (1990) [198]	EBP-EWLS	
Regression Neural Learning	Specht, D. F., 1991[261]	Extended kalman Filter	Dan Simon,2002[271]

Table III
Soft Computing Based Approaches Evolutionary Algorithms

Algorithm	*Author, Reference*	*Algorithm*	*Author, Reference*
EANN(Evolutionary ANNs)	Yao,1999[84][85]	GA-BP and GA-LM	Enrique Alba and J. Francisco Chicano[237]
GA	Holland,1975[90][91]	Breeder Genetic Programming (BGP)	B T Zang[98]
GP	John Koza,1992[97]	Improved DE	Shamekhi[105]
ES	1965,1975,Richenberg, Schewefel[99][100]	self-adaptive DE (SaDE)	Qin et al. [106]
DE	Storn and Price, 1995 [102]	CMDE-G	Radha Thangraj, Millie Pant *et al.* [108]
QNN method	A Jagtap(2014)[231]	Modified DE	Sibarama Panigrahi *et al.* [111]
PHYSICS OR CHEMISTRY BASED APPROACHES			
Algorithm	Author ,Reference	Algorithm	Author, Reference
IWD	Hamed Shah-hosseini, 2007[183]	Hybrid GSA-GA	Saeide Sheikhpour [257]
IWD-NQ	Shah-Hosseini, 2009 [185]	ImprovedGSA(IGSA)	Bao-Chang Xu,2014 [258]
EEIM	Birbil & Fang, 2003 [187]	BH(Black hole)	Zang Lieu,2008[200]
Hybrid EEIM-GA	Ching-Hung Lee, 2009[192]	SA(Simulated Annealing)	Kirkpatrick, Gelatt and Vecchi,o 1983 [202]
Improved EEIM	Ching-Hung Lee, 2010[193]	Hybrid PSO-SA	Sriram G. Sanjeevi [203]
ICEM	Zang et al., 2013[196]	BB-BC	Erol and Eksin,2006[206]
GSA	Rashiedi et al., 2009[197]	Hybrid BB-BC(HBBBC)	Kaveh,2009[208]
Hybrid PSO-GSA	S. Mirjalili et al., 2012[256]	Uniform Big_Bang–Chaotic Big_Crunch (UBB–CBC)	Alatas,(2011 [209]
BBBC-PSO	Kaveh,2013[214]	Improved BB-BC Algorithm	Behrooz Hassani, 2012[210]
BBBC-PSO-ACO-Harmony Search	Kaveh[215]	BB BC with local Search moves	Genc, Eksin and Erol,2013 [211]
Parallel BB-BC	S Kumar et al.. [218]		

Table IV
Soft Computing Based Approches Swarm Intelligence(si) Based Approaches

Algorithm	*Author, Reference*	*Algorithm*	*Author, Reference*
BFOA	Passino,2002[143]	PSO	Eberhart and kenedy,1995 [113]
GA-BFOA	D H Kim,2007[147]	ACO	Dorigo &Gambardella[122]-[125]
Adaptive BFOA	Majhi,2009[146]	ACO -BP	Liu et al,2006[239]
Modified BFOA	EM Montes,2009[145]	ABC	Karaboga and Basturk ,2007 [126]
BFONN	Zhang et al.[252]	FA	Yang,2009 [129]
Tabu Search	Glover,1977[150]	FSA	Li Xiao-lei et al. (2002)[132]
Cuckoo Search	Yang 2009; Heb 2010[154]	Adaptive FSA	Reza Aziz,2014[134]
Eurygaster Algorithm	Fariborz Ahmadi et al. 2012)[156]	FSA-PSO	Huadong Chen et .al[242]
Krill Herd optimization	A H Gandomi,2012 [157]	SFLA	Muzaffar Eusuff and Kevin Lan sey (2003)[135]
BA	Yang,2010[138]	Modified SFLA(MFLA)	G Samuel and C. C A Raj, 2014 [137]
BA-HS hybrid	Wang and Guo, 2013	BA-DE hybrid	Fister Jr et al. (2013)
Hybrid BA	Fister Jr.[139]	Virtual bee algorithm	X S Yang [158]
Enhanced Bat Algorithm(EBA)	A. Kaveh and P. Zakian[141]	quaternion bat algorithm (QBA)	Fister, 2013[140]
BIO INSPIRED NON-SI APPROACHES			
Algorithm	*Author, Reference*	*Algorithm*	*Author, Reference*
IWO	Mehrabian and Lucas (2006) [159],	Flower pollination algorithm,	(X. S. Yang, 2012)[177]
Modified IWO(MIWO)	Ritwick, 2010[254]	photosynthetic algorithm,	(Murase, 2000) [199]
Improved HGIWO	Zhi YIN,2012 [160]	enzyme algorithm,	(Yang, 2005) [180],
BBO	Dan Simon 2008[161]	cross-entropy algorithm,	Rubinstein (1999; 2001) [181][182].
AIS	Dasgupta,in 1999 [166]	GSO	S.He,Q.H.Wu et al., 2006)[173]
NN Immune System	Leandro Nunes de Castro[168]	cultural algorithms (CA)	R. G. Reynolds,1999) [175]
HS	Geem et al., 2001[169]	monkey search (MS)	Mucherino and O. Seref, 2007) [176]
DHS	Prithwish Chakraborty et al. [172]		