LEARNING ENHANCEMENT OF THREE-TERM BACKPROPAGATION NETWORK BASED ON ELITIST MULTI-OBJECTIVE EVOLUTIONARY ALGORITHMS

ASHRAF OSMAN IBRAHIM ELSAYED

A thesis submitted in fulfilment of the
requirements for the award of the degree of
Doctor of Philosophy (Computer Science)

Faculty of Computing
Universiti Teknologi Malaysia

JUNE 2015

To my father, late mother and my late son (KARIM)

To my beloved wife, son, brothers, sisters and friends

ACKNOWLEDGEMENT

In the name of Allah, Most Gracious and Most Merciful
All praise and thanks be to Allah, peace and blessings be upon his messenger, Muhammad (S.A.W).

I thank Allah (S.W.T), for granting me perseverance and strength that was needed to complete this thesis.

I would like to express my great thanks and appreciations to my supervisor, Prof. Dr.Siti Mariyam Shamsuddin, for her encouragement, guidance, advice and support throughout my study. I am also thankful to Dr. Sultan for his assistance and advice.

My sincere appreciation also goes to all UTM staff and colleagues in the Faculty of Computing. I would also like to express my gratitude to Soft Computing Research Group (SCRG), UTM big data center and all of my friends for their continuous help and support.

This work was partially supported by the Ministry of Higher Education (MOHE) under FRGS GRANT R.J130000.7828.4F347 - NEW ROTATIONAL MOMENT INVARIANTS).

Finally, I am highly indebted to my father, brothers, sisters and all my family members for their support and prayers which without, this thesis would have not been completed successfully. It is also my wish to thank my wife for her patience, encouragement, support and understanding.

ABSTRACT

The pattern classification problem in machine learning algorithms is the task of assigning objects to one of a different predefined group of categories related to that object. Among the successful machine learning methods are Artificial Neural Networks (ANNs), which aim to minimize the error rate of the training data and generate a simple network architecture to obtain a high classification accuracy. However, designing the ANN architecture is difficult due to the complexity of the structure, such as the network structure, number of hidden nodes and adjustment of weights. Therefore, a number of Evolutionary Algorithms (EAs) has been proposed to improve these network complexities. These algorithms are meant to optimize the connection weight, network structure, network error rate and classification accuracy. Nevertheless, these algorithms are implemented to optimize only one objective, despite the importance of executing many objectives simultaneously. Therefore, this study proposes simultaneous learning and structure optimization for designing a Three-term Backpropagation (TBP) network with four variants of Elitist Multi-objective Evolutionary Algorithms (EMOEAs). These include the Elitist Multi-objective Genetic Algorithm (EMOGA), Hybrid Elitist Multi-objective Genetic Algorithm (HEMOGA), Memetic Adaptive Elitist Multi-objective Genetic Algorithm (MAEMOGA) and the Elitist Multi-objective Differential Evolution (EMODE). The proposed methods are developed to evolve towards a Pareto-optimal set that is defined by multi-objective optimization consisting of connection weight, error rate and structural complexity of the network. The proposed methods are tested on binary and multi-class pattern classification problems. The results show that the proposed MAEMOGA and EMODE are better than EMOGA and HEMOGA in obtaining simple network structure and classification accuracy..

ABSTRAK

Masalah pengkelasan pola dalam algoritma pembelajaran mesin merupakan suatu tugas pengkelasan objek kepada salah satu kategori kumpulan yang berkaitan dengan objek itu. Rangkaian Neural Buatan (ANN) merupakan salah satu kaedah pembelajaran mesin yang berjaya mengurangkan kadar ralat data pengujian dan menjana senibina rangkaian mudah untuk menghasilkan kadar ketepatan pengkelasan yang tinggi. Walau bagaimanapun, merekabentuk suatu senibina ANN adalah rumit kerana ia melibatkan penentuan struktur seperti struktur rangkaian, bilangan nod tersembunyi dan pelarasan pemberat. Sehubungan dengan itu, beberapa Algoritma Evolusi (EA) telah dicadangkan bagi menambahbaik penyelesaian kepada kerumitan rangkaian ini. Algoritma ini adalah bertujuan untuk mengoptimumkan pemberat hubungan, struktur rangkaian, kadar ralat rangkaian dan ketepatan pengkelasan. Walau bagaimanapun, algoritma ini umumnya dilaksanakan untuk mengoptimumkan satu fungsi objektif sahaja, walaupun ia berkepentingan dalam melaksanakan kesemua objektif secara serentak. Oleh itu, kajian ini mencadangkan pembelajaran serentak dan pengoptimuman struktur untuk merekabentuk rangkaian Tiga Istilah Perambatan Balik (TBP) dengan empat varian algoritma-Algoritma Evolusi Elitis Multi-objektif (EMOEAs). Ini termasuk Algoritma Genetik Elitis Multi-objektif (EMOGA), Algoritma Genetik Hibrid Elitis Multi-objektif (HEMOGA), Algoritma Evolusi Penyesuai *Memetic* Elitis Multi-objektif (MAEMOGA) dan Pembezaan Evolusi Elitis Multi-objektif (EMODE). Kaedah yang dicadangkan telah dibangunkan untuk mengevolusi set Pareto yang optimum yang ditakrifkan pengoptimuman multi-objektif yang terdiri daripada pemberat penghubung, kadar ralat dan kerumitan struktur rangkaian. Kaedah yang dicadangkan telah diuji ke atas masalah pengkelasan pola binari dan pelbagai. Keputusan menunjukkan bahawa teknik MAEMOGA dan EMODE yang dicadangkan adalah lebih baik daripada EMOGA dan HEMOGA dalam memperoleh struktur rangkaian yang mudah dan ketepatan pengkelasan.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATION

AMGA	-	Archive-Based Micro GA
AMGA2	-	Enhanced Archive-Based Micro GA
ANNs	-	Artificial Neural Networks
BP	-	Backpropagation
DE	-	Differential Evolution
DEGL	-	DE Local And Global
EANNs	-	Evolutionary Artificial Neural Networks
EAs	-	Evolutionary Algorithms
EMODE	-	Elitist Multi-Objective Differential Evolution
EMOEA	-	Elitist Multi-Objective Evolutionary Algorithm
EMOGA	-	Elitist Multi-Objective Evolutionary Algorithm
ES	-	Evolution Strategy
FN	-	False Negative
FP	-	False Positive
FWNN	-	Fuzzy Wavelet Neural Networks
GA	-	Genetic Algorithm
GD	-	Gradient Descent
GDE	-	Generalized DE
GDE2	-	Extension Generalized De
GDE3	-	Improved Version Of GDE
GMLP	-	Generalized Multilayer Perceptron
HEMOGA	-	Hybrid Elitist Multi-Objective Genetic Algorithm
HPDENN	-	Hybrid Pareto DE Neural Network
LS	-	Local Search
MA	-	Memetic Algorithms
MAEMOGA	-	Memetic Adaptive Elitist Multi-bjective Genetic Algorithm

MLP	-	Multilayer Perceptron
MM	-	Modified Cost Function
MODE	-	Multi-Objective Differential Evolution
MOEANNs	-	Multi-Objective Evolutionary Artificial Neural Networks
MOEAs	-	Multi-Objective Evolutionary Algorithms
MOGA	-	Multi-Objective Genetic Algorithm
MOO	-	Multi-Objective Optimization
MOOPs	-	Multi-Objective Optimization Problems
MPANN	-	Memetic Multi-Objective Pareto Artificial Neural Network
MPDENN	-	Memetic Pareto De Neural Network
MSE	-	Mean Square Error
NSDE	-	Non-Dominated Sorting Differential Evolution
NSGA	-	Non-Dominated Sorting Genetic Algorithm
NSGA-II	-	Non-Dominated Sorting Genetic Algorithm-II
OW-MOSaDE	-	Multi-Objective Self-Adaptive Differential Evolution Algorithm With Objective-Wise Learning Strategies
PAES	-	Pareto Archive Evolution Strategy
PDE	-	Pareto-Based DE
PF	-	Proportional Factor
PSO	-	Particle Swarm Optimization
QAC	-	Qualitative Analytical Chemistry
RBF	-	Radial Basis Function
RNN	-	Recurrent Neural Network
SA	-	Simulated Annealing
SA-SBX	-	Self-Adaptive Simulated Binary Crossover
SBX	-	Simulated Binary Crossover
SD	-	Standard Deviation
SDLS	-	Local Search Based On Spatial Distribution
SPDE	-	Self-Adaptive Pareto Differential Evolution
SPEA	-	Strength Pareto EA
SVM	-	Support Vector Machine

TBP	-	Three-term Backpropagation
TN	-	True Negative
TP	-	True Positive
VEDE	-	Vector Evaluated De
VEGA	-	Vector Evaluated GA

CHAPTER 1

INTRODUCTION

1.1 Overview

Machine learning is an important sub field of artificial intelligence (AI) that is applied in the development of computational algorithms and allows computers to include patterns and rules from a priori data. It involves adaptive mechanisms. Hence, it enables computers to learn by example and from experience. The machine learning can be accomplished by supervised or unsupervised learning.

Among the common machine learning methods are artificial neural networks (ANNs). Recently, ANNs have been widely used in different areas with different applications (Cheok *et al.*, 2012; Khosrowshahi, 2011; Kuo and Lin, 2010; Melesse *et al.*, 2011; Yaghini *et al.*, 2012). ANNs are considered to be information processing systems that are largely influenced by the way in which biological neurons process information in the brain. The processes present the form of learning that enables the brain to think and learn something. When the signals received are strong enough (surpasses a certain threshold), the neuron is activated and emits a signal though the axon. This signal might be sent to another synapse, and might activate other neurons. The phenomenon of learning that happens in the brain has inspired ANNs to adapt the learning concept and translate it into a mathematical model of human cognition. It clearly happens since the learning concept becomes the major concern of ANN in order to generate the intelligent system that can learn the pattern and generate the desired output.

The backpropagation algorithm (BP) was introduced by Rumelhart *et al.* (1986), as an ANN, it is also known as the two-term BP algorithm. BP is a supervised learning algorithm used by multilayered neural networks for learning purposes. Moreover, the BP algorithm uses a gradient descent (GD) technique (gradient search based learning algorithm), which calculates the error and propagates it back to the weights of the connections of the network. It is the most famous training algorithm for multilayer perceptions, and it is the most widely used to train the feed forward ANN (Cui *et al.*, 2012; Ding *et al.*, 2011a; España-Boquera *et al.*, 2007; Melesse *et al.*, 2011; Miguez *et al.*, 2014; Nawi *et al.*, 2013; Wu *et al.*, 2010; Xiao *et al.*, 2009). On the other hand, the three-term backpropagation (TBP) was proposed by (Zweiri *et al.*, 2003). The TBP network introduced a third parameter called the proportional factor (PF), which has proven successful in fastening the weight adjustment process through the increase in the convergence rate of the algorithm and reduction in learning stalls.

Recently, there has been a remarkable increase in the use of evolutionary algorithms (EAs) for solving optimization problems as well as optimizing the ANN learning (Dragoni *et al.*, 2013). The design and optimization of ANNs is considered to be one of the most important problems that need to be solved using these kinds of algorithm. The earlier approaches tackled the single objective optimization problems in some of the previous works, such as particle swarm optimization (PSO) (Zhang *et al.*, 2000), genetic algorithms (GA) (Ding *et al.*, 2011a) and differential evolution (DE) (Ilonen *et al.*, 2003), and others, that were considered for optimizing ANNs.

These EAs are population-based algorithms, which allow for simultaneous exploration of different parts in the Pareto-optimal set. As an alternative to dealing with a single optimal solution, a set of optimal solutions called Pareto-optimal sets exist for such problems. The corresponding objective functions, whose non-dominated solutions in the Pareto-optimal set are called a Pareto front, in which each of the Pareto-optimal solutions signifies a different balance between the objectives, and with a lack of preferred information, none of them can be considered better than the others. Therefore, EAs are good candidates for multi-objective optimization problems (MOOPs) because of their ability to search for multiple Pareto-optimal solutions and better performance in global search space. Therefore, Pareto-optimal solutions are used to evolve ANNs that are optimal both with respect to classification accuracy and architecture complexity.

1.2 Problem Background

The backpropagation (BP) algorithm is one of the most popular ANNs. It has good self-learning, self-adapting, robustness and generalization ability. Despite the general success of BP in learning, it has major limitations, such as slow convergence speed, long training time and is easily trapped at a local minima. In addition, the choice of a proper network structure (number of hidden nodes) and design of a proper network are considered to be among the most important problems of the BP algorithm. Hence, there is a real necessity to develop solutions to overcome these problems, and several major deficiencies still need to be solved or improved (Chun-Dong *et al.*, 2012; Miguez *et al.*, 2014; Tang *et al.*, 2011; Xue and Ma, 2011). Recently, many methods have been tried to overcome the slow convergence (Ren *et al.*, 2012; Yu and Peng, 2012), while other methods tried to avoid the local minima problems (Bari *et al.*, 2011; Burse *et al.*, 2011; Hamid, 2012; Yi *et al.*, 2014). Some studies have tried designing ANNs by determining the optimal structure (number of hidden nodes) and the optimal connection weights and architecture design (Qasem and Shamsuddin, 2011; Sagar *et al.*, 2011; Yu and Peng, 2012).

The TBP network is one of the improved BP algorithms. It was proposed to speed up the weight adjusting process and has outperformed BP in terms of convergence speed and the ability to escape from local minima. According to Zweiri (2007) optimization is required to facilitate the application of the TBP network. Recently, there have been many studies in the literature associated with TBP network learning for different applications. These include, TBP network for moisture prediction (Abdulkadir *et al.*, 2012a; Abdulkadir *et al.*, 2012b), classification problems (Mashinchi and Shamsuddin, 2009; Saman, 2006; Shamsuddin *et al.*, 2009), XOR and parity problems (Burse *et al.*, 2011), which have been applied to the problem of service selection in ubiquitous computing (Cai *et al.*, 2006). Despite the success of these applications, improvement is still required to the network topology and accuracy results. Furthermore, it is an ANN and also needs optimization and design for final error output and good architecture. Therefore, a good network design needs to be tried for a simple network and better performance.

The performance of ANNs is sensitive to the number of the hidden nodes, in that a network with less hidden nodes gives poor approximation, while a network with more hidden nodes may contribute to over-fitting problems. In addition, it may perform better in training data, but it may not give a good generalization on testing data. However, the success of ANNs mostly depends on the network design. Therefore, the design of an ANN is a difficult task as it depends on human experience (Ding *et al.*, 2013; Garro *et al.*, 2010). Therefore, many researchers are concerned about the problems of determining the optimal architecture design of the ANNs and improving the generalization of the network. These problems have been addressed using evolutionary algorithms (EAs) (Chang *et al.*, 2012; Han *et al.*, 2011; Irani *et al.*, 2011; Tang *et al.*, 2011; Wang *et al.*, 2011; Wang and Qian; Yi *et al.*, 2014; Yu and Peng, 2012). They have proved that these kinds of algorithm are feasible and effective for this task. This is because the evolutionary algorithms (EAs) provide a robust and efficient approach to explore a massive search space. However, these optimization techniques only optimize one factor, such as hidden nodes, connection weights or network error rates. The limitation of these kinds of algorithm is that they can only produce a single optimal solution. In ANNs optimization, there is more than one parameter that needs to be optimized. In this case, a set of optimal solutions or Pareto-optimal solutions are required for such problems. Therefore, multi-objective optimization algorithms are preferred because of their ability to optimize more than one objective simultaneously.

However, the trend of implementing multi-objective evolutionary algorithms (MOEAs) for optimizing ANN network structures has increased in recent years. MOEAs, also known as multi-objective optimization (MOO), is the process of simultaneously optimizing two or more conflicting objectives subject to certain constraints (Cruz-Ramírez *et al.*, 2012b). Therefore, MOEAs are suitable for producing and designing appropriate and accurate ANNs with the optimization of two conflicting objectives, namely, the minimization of ANNs structure complexity and the maximization of network capacity. Hence, recently, MOEAs have been applied successfully to optimize the ANNs. It has been demonstrated that MOEAs have a significant advantage over the conventional BP method because of their low computational requirement when searching in a large solution space (Fernández *et al.*, 2012; Qasem *et al.*, 2013). MOEAs for the learning problem were applied to improve the generalization of the training and unseen data. These kinds of algorithm have been used to evolve ANNs for different kinds of problem, such as classification problems (Ou and Murphey, 2007; Qasem and Shamsuddin, 2011; Qasem *et al.*, 2013), some of its key exponents being Abbass (2002a) and Jin and Sendhoff (2008).

Various methods and techniques have been developed to find better approaches to evolve ANNs in trying to design networks with good generalization capability. In the same way, the issue of finding a good ANN architecture has also been debated in the field of ANNs. In addition, some works have used multi-objective genetic algorithms (MOGAs) for optimizing ANNs. One of the most successful applications in this area, is a hybrid method that uses ANNs with evolutionary Pareto-based algorithms (Jin and Sendhoff, 2008). In Pettersson *et al.* (2007) multi-objective genetic algorithm optimization for training a feed forward neural network was effectively constructed by minimizing the training error and the network size using noisy data from an industrial iron blast furnace. The work presented by Liu and Kadirkamanathan (1999) highlighted the benefits of multi-objective genetic algorithms for the selection and identification of nonlinear systems, while optimizing the size of neural networks. (Garcıa-Pedrajas *et al.*, 2004) presented a method based on the generalized multilayer perceptron (GMLP) with two hidden layers, which improved the performance of the evolutionary model for real world classification problems. Another major study by Delgado *et al.* (2008) proposed a hybrid MOGA method based on the SPEA2 and NSGA2 algorithms to optimize the training and topology of the recurrent neural network (RNN) simultaneously for time-series prediction problems. Fernandez Caballero *et al.* (2010), introduced multi-objective and considered a memetic Pareto evolutionary approach based on the NSGA2 evolutionary ANN algorithm to optimize two conflicting main objectives: a high correct classification rate and a high classification rate for each class. The recent work by Ak *et al.* (2013) used a non-dominated sorting genetic algorithm-II (NSGA-II) to train the neural network and optimize their weights and biases with respect to maximum accuracy and minimum dimension to provide the prediction intervals of the scale deposition rate.

In addition, there are a limited number of studies using multi-objective differential evolution (MODE) algorithms to train a population of multi-objective ANNs, which are commonly used to minimize the error in the training set and the complexity of the network. One of the first works in this field, by Abbass and Sarker (2001), presented a multi-objective method that includes the PDE algorithm to train the ANN and to optimize the number of hidden nodes and connection weights simultaneously. Moreover, Abbass offered various works using the multi-objective idea for design and training ANNs using accuracy and complexity as objectives (Abbass, 2002a; Abbass, 2003; Abbass *et al.*, 2001). Another study, by Ilonen *et al.* (2003), analysed DE as a candidate global optimization technique for feed-forward neural networks as compared to gradient approaches, and designed ANNs using the mean square error as the objective function. Likewise, Fieldsend and Singh (2005) used the Pareto-optimal approaches to train a multilayer perceptron network. They achieved a Pareto-optimal evolutionary neural network as a parallel evolution of a population and considered multiple error measures as objectives. Similarly, (Fernández *et al.*, 2009) suggested MODE based on the Pareto dominance concept and multilayer perceptron (MLP) for multi-classification problems using models. The hybrid local search algorithm also offered to optimize two conflicting objectives. The work in Cruz-Ramírez *et al.* (2010) presents the optimization technique for two objectives to determine the growth limits of two pathogens simultaneously.

However, many works concerning optimization and design of ANNs have been conducted (Ak *et al.*, 2012; Cruz-Ramírez *et al.*, 2012a; Cruz-Ramírez *et al.*, 2010; Qasem *et al.*, 2011; Ramesh *et al.*, 2011), which demonstrated that EAs, such as the genetic algorithm and its upgraded derivatives are feasible for optimal design. The main advantage of the evolutionary approach over traditional learning algorithms, like BP, is its ability to escape a local optimum. Its robustness and its ability to adapt itself to a changing environment (Cruz-Ramírez *et al.*, 2012b; Fernandez Caballero *et al.*, 2010; Qasem *et al.*, 2013; Qasem *et al.*, 2011). On the other hand, the main disadvantage of the evolutionary approach is the high computation, as the evolutionary approach is usually slow. To overcome the slow convergence of the EAs, hybrid techniques have been used to speed up convergence by enhancing EAs with a local search algorithm, such as BP (Fernandez Caballero *et al.*, 2010; Yan *et al.*, 1997).

An additional possible advantage of the Pareto-based learning approach is that using multi-objective techniques may help the learning algorithm to escape from local optima, thus improving the accuracy of the learning model. Therefore, this study, proposes elitist multi-objective evolutionary algorithm (EMOEA) methods, which include the elitist multi-objective genetic algorithm (EMOGA), the hybrid elitist multi-objective genetic algorithm (HEMOGA), memetic adaptive elitist multi-objective genetic algorithm (MAEMOGA) and the elitist multi-objective differential evolution (EMODE) to optimize the TBP network structure, error rates and connection weight of the network simultaneously.

1.3 Problem Statement

From the problem background, it can be claimed that further works are still required to develop new methods of BP network with multi-objective evolutionary algorithms (MOEAs), such as an elitist multi-objective genetic algorithm (EMOGA), hybrid elitist multi-objective genetic algorithm (HEMOGA), memetic adaptive elitist multi-objective genetic algorithm (MAEMOGA) and elitist multi-objective differential evolution (EMODE) to optimize the TBP network parameters simultaneously. The proposed methods aim to achieve a better network performance and network architecture simultaneously. In other words, the intention is to design an appropriate and accurate TBP network and enhancement of the learning process simultaneously.

This study raises several challenges, which include developing a TBP network using EMOEAs to reduce the network complexity in terms of the number of hidden nodes and weights of the TBP network. In addition, it aims to design an appropriate and accurate TBP network and enhance the learning process simultaneously.

Based on the above issues, the main research question is:

Are the proposed methods, which include EMOGA, HEMOGA, MAEMOGA and EMODE, beneficial and efficient for evolving TBP network learning?

Thus, these challenges will be addressed by providing answers to the following questions:

- Is the proposed hybrid scheme of EMOEA capable of optimizing the TBP network?
- Are the proposed methods capable of optimizing the TBP network structure (reduce the Complexity) in terms of the number of hidden nodes and weights?
- Are the proposed improved methods able to achieve a better network performance and network architecture simultaneously?
- Are the proposed methods able to improve the classification accuracy in the classification problems?

1.4 Goal of the Study

The aim of this research is to improve the weight and structure of the TBP network simultaneously, and to achieve a better optimized network performance, optimal architecture and simple and accurate TBP network using an elitist multi-objective genetic algorithm (EMOGA), hybrid elitist multi-objective genetic algorithm (HEMOGA), memetic adaptive elitist multi-objective genetic algorithm (MAEMOGA) and elitist multi-objective differential evolution (EMODEN).

1.5 Objective of the study

To achieve the aim of this study, the objectives of this research are stated as follows:

1. To propose a hybrid scheme of TBP network with elitist multi-objective evolutionary algorithms (EMOEAs) for optimizing the network structure, connection weights and error rate simultaneously.

2. To improve the generalization and network accuracy of the proposed hybrid scheme of TBP network with an elitist multi-objective genetic algorithm (EMOGA) technique. These include:

    i. Hybrid elitist multi-objective genetic algorithm (HEMOGA) to enhance the proposed EMOGA.

    ii. Memetic adaptive elitist multi-objective genetic algorithm (MAEMOGA) to enhance the proposed HEMOGA.

3. To propose a hybrid scheme using a TBP network with an elitist multi-objective differential evolution algorithm (EMODE), to achieve a simple structure and more accurate classification results.

1.6 Scope of the Study

To achieve the above objectives of this research, the scope of this study is:

- Binary, multi-class and complex real problem datasets for classification tasks to validate the proposed methods.

- A focus on multi-objective optimization and the TBP network with MOEA methods (EMOGA, HEMOGA, MAEMOGA and EMODE) for the training and testing in pattern classification problems.

- Performance is measured based on convergence towards error, the structure of the network, 10-fold cross validation, sensitivity, specificity, classification accuracy and statistical test.

- The programs are customized, developed and applied to the TBP network and MOEAs using Microsoft Visual C++ 2010.

1.7 Significance of the study

The significance of this research is to optimize the structure (reduce the Complexity) in terms of the number of hidden nodes and weights of the TBP network using EMOEAs methods, for better accuracy in classification problems, and accelerate the artificial neural network. The proposed methods are investigated using various parameter measurements. These include the number of hidden nodes, MSE, sensitivity, specificity and classification accuracy.

1.8 Organization of Thesis

This section describes the organization of the thesis. There are six chapters in this thesis, as follows:

Chapter 1, *Introduction*: this chapter presents a general introduction to the topic of the proposed research work. Brief overviews of some of the issues concerning the research are also mentioned in this chapter. Besides the problem background, this chapter also includes the problem statement, objectives of study, research scope, significance of the study and the expected contribution.

Chapter 2, *Literature Review*: in this chapter, we explain some principles of artificial neural networks (ANNs) and multi-objective evolutionary algorithms (MOEAs). The relevant works of artificial neural networks (ANNs) are elucidated. Since this research proposes a multi-objective evolutionary algorithm based solution, the chapter also reviews the types and approaches of EMOEAs. Moreover, it reviews and discusses the state of the art related works on EMOEAs and ANNs. Next, we highlight the many studies in the literature that have designed artificial neural networks using evolutionary algorithms. Finally, we clarify the concept of multi-objective optimization techniques that relate to this study, such as EMOGA, HEMOGA, MAEMOGA and EMODE.

Chapter 3, *Research Methodology*: this chapter illustrates the methodology adopted in this research to achieve the study objectives. A methodology is generally a guideline for solving a research problem. This includes discussion on the research components, such as the phases, techniques and describes the overall solving-tools adopted.

Chapter 4, *three-term backpropagation (TBP) network optimizing by elitist multi-objective genetic algorithms (EMOGAs):* this chapter presents the TBP network based on elitist multi-objective evolutionary algorithms. The EMOEA algorithms used in this chapter include an elitist multi-objective genetic algorithm (EMOGA), HEMOGA by hybrid local search algorithm and (MAEMOGA) utilized self-adaptive simulated binary crossover. The methods are compared with each other and against "state of the art" methods that are similar systems based on GA.

Chapter 5, *Hybrid TBP Network and elitist multi-objective evolutionary approach of differential evolution algorithm (DE)* based on elitist non-dominated sorting differential evolution (NSDE): the main goal of this chapter is to improve the TBP network based on EMOEA, which is a DE algorithm for optimizing the network to achieve a simple structure and more accurate classification results.

Chapter 6, *Conclusion and Future Work:* this chapter concludes the research work and attempts to give an overall discussion regarding all the contributions presented in this research, and, finally, it presents recommendations and suggestions for future work.

REFERENCES

Abbass, H. A. (2001). A memetic pareto evolutionary approach to artificial neural networks *AI 2001: Advances in Artificial Intelligence* 1-12,Springer.

Abbass, H. A. (2002a). An evolutionary artificial neural networks approach for breast cancer diagnosis. *Artificial Intelligence in Medicine*. 25(3), 265-281.

Abbass, H. A. (2002b). The self-adaptive pareto differential evolution algorithm. In *Proceedings Congress on Evolutionary Computation, 2002. CEC'02.* IEEE, 831-836.

Abbass, H. A. (2003). Speeding up backpropagation using multiobjective evolutionary algorithms. *Neural Computation*. 15(11), 2705-2726.

Abbass, H. A. and Sarker, R. (2001). Simultaneous evolution of architectures and connection weights in ANNs. In *Proceedings of the 2001 Proceedings of Artificial Neural Networks and Expert System Conference*, 16-21.

Abbass, H. A. and Sarker, R. (2002). The Pareto differential evolution algorithm. *International Journal on Artificial Intelligence Tools*. 11(04), 531-552.

Abbass, H. A., Sarker, R. and Newton, C. (2001). PDE: a Pareto-frontier differential evolution approach for multi-objective optimization problems. In *Proceedings of the 2001 Congress on Evolutionary Computation*: IEEE, 971-978.

Abdul-Kader, H. (2009). Neural networks training based on differential evolution algorithm compared with other architectures for weather forecasting 34. *Int J Comput Sci Netw Secur*. 9(3), 92-99.

Abdulkadir, S. J., Shamsuddin, S. and Sallehuddin, R. (2012a). Moisture prediction in maize using three term back propagation neural network. *International Journal of Environmental Science and Development*. 3(2), 199-204.

Abdulkadir, S. J., Shamsuddin, S. M. and Sallehuddin, R. (2012b). Three Term Back Propagation Network for Moisture Prediction.In *Proceedings of the 2012b International Conference on Clean and Green Energy*.103-107.

Abraham, A. and Nath, B. (2000). Optimal design of neural nets using hybrid algorithms. *PRICAI 2000 Topics in Artificial Intelligence*. 510-520.

Ahmad, F., Isa, N. A. M., Hussain, Z. and Sulaiman, S. N. (2012). A genetic algorithm-based multi-objective optimization of an artificial neural network classifier for breast cancer diagnosis. *Neural Computing and Applications*. 1-9.

Ak, R., Li, Y., Vitelli, V., Zio, E., LóPez Droguett, E. and Magno Couto Jacinto, C. (2013). NSGA-II-trained neural network approach to the estimation of prediction intervals of scale deposition rate in oil & gas equipment. *Expert Systems with Applications*. 40(4), 1205-1212.

Ali, M., Siarry, P. and Pant, M. (2012). An efficient differential evolution based algorithm for solving multi-objective optimization problems. *European journal of operational research*. 217(2), 404-416.

Almeida, L. M. and Ludermir, T. B. (2010). A multi-objective memetic and hybrid methodology for optimizing the parameters and performance of artificial neural networks. *Neurocomputing*. 73(7), 1438-1450.

Babu, B. (2007). Improved differential evolution for single-and multiobjective optimization: MDE, MODE, NSDE, and MNSDE. In Deb, K., et al. (Ed.) *Advances in Computational Optimization and its Applications* (24–30)Universities Press, Hyderabad.

Babu, B. and Gujarathi, A. M. (2007). Elitist-multi-objective differential evolution (E-MODE) algorithm for multi-objective optimization.In *Proceedings of the 2007 Proc. of 3rd Indian International Conference on Artificial Intelligence (IICAI-2007)*, 441-449.

Babu, B. and Jehan, M. M. L. (2003). Differential evolution for multi-objective optimization.In *Proceedings of the 2003 Evolutionary Computation, 2003. CEC'03. The 2003 Congress on*: IEEE, 2696-2703.

Bari, A., Bhasin, K. and Karnawat, D. N. (2011). Introduction to Neural Network and Improved Algorithm to Avoid Local Minima and Faster Convergence. *Computational Intelligence and Information Technology*. 396-400. Springer Berlin Heidelberg.

Bazoobandi, H. and Eftekhari, M. (2014). A Differential Evolution and Spatial Distribution based Local Search for Training Fuzzy Wavelet Neural Network.

International Journal of Engineering-Transactions B: Applications. 27(8), 1185.

Burse, K., Manoria, M. and Kirar, V. P. S. (2011). Improved back propagation algorithm to avoid local minima in multiplicative neuron model *Information Technology and Mobile Communication* . 67-73. Springer.

Cai, H., Sun, D., Cao, Q. and Pu, F. (2006). A Novel BP Algorithm Based on Three-term and Application in Service Selection of Ubiquitous Computing.In *Proceedings of the 2006 Control, Automation, Robotics and Vision, 2006. ICARCV'06. 9th International Conference on*: IEEE, 1-6.

Chang, Y. T., Lin, J., Shieh, J. S. and Abbod, M. F. (2012). Optimization the initial weights of artificial neural networks via genetic algorithm applied to hip bone fracture prediction. *Advances in Fuzzy Systems*. 2012, 6.

Chen, Y. and Yang, H. (2012). Multiscale recurrence analysis of long-term nonlinear and nonstationary time series. *Chaos, Solitons & Fractals*. 45(7), 978-987.

Cheok, C. Y., Chin, N. L., Yusof, Y. A., Talib, R. A. and Law, C. L. (2012). Optimization of total phenolic content extracted from Garcinia mangostana Linn. hull using response surface methodology versus artificial neural network. *Industrial Crops and Products*. 40, 247-253.

Chun-Dong, X., Wei, L., Mu-Gui, Z. and Jin-Gao, L. (2012). A BP Neural Network Activation Function Used in Exchange Rate Forecasting. *Information Technology and Agricultural Engineering*. 69-76. Springer Berlin Heidelberg.

Coello, C. A. (2000). An updated survey of GA-based multiobjective optimization techniques. *ACM Computing Surveys (CSUR)*. 32(2), 109-143.

Cruz-Ramírez, M., Hervás-Martínez, C., Fernández, J. C., Briceño, J. and de la Mata, M. (2012a). Multi-Objective Evolutionary Algorithm for Donor-Recipient Decision System in Liver Transplants. *European Journal of Operational Research*, 222(2), 317-327.

Cruz-Ramírez, M., Hervás-Martínez, C., Gutiérrez, P. A., Briceño, J. and de la Mata, M. (2011). Memetic pareto differential evolutionary neural network for donor-recipient matching in liver transplantation *Advances in Computational Intelligence*. 129-136. Springer.

Cruz-Ramírez, M., Hervás-Martínez, C., Gutiérrez, P. A., Pérez-Ortiz, M., Briceño, J. and de la Mata, M. (2013). Memetic Pareto differential evolutionary neural network used to solve an unbalanced liver transplantation problem. *Soft Computing*. 17(2), 275-284.

Cruz-Ramírez, M., Hervás-Martínez, C., Jurado-Expósito, M. and López-Granados, F. (2012b). A multi-objective neural network based method for cover crop identification from remote sensed data. *Expert Systems with Applications*. 39(11), 10038-10048.

Cruz-Ramírez, M., Sánchez-Monedero, J., Fernández-Navarro, F., Fernández, J. and Hervás-Martínez, C. (2010). Memetic pareto differential evolutionary artificial neural networks to determine growth multi-classes in predictive microbiology. *Evolutionary intelligence*. 3(3-4), 187-199.

Cui, Y., Xiong, H., Zheng, K. and Chen, J. (2012). On the application of BP neural network based on Levenberg-Marquardt algorithm in the diagnosis of mental disorders.In *Proceedings of the 2012*: IEEE, 1940-1943.

Deb, K. (1999). An introduction to genetic algorithms. *Proceedings of the 1999 Sadhana (Academy Proceedings in Engineering Sciences)*: Indian Academy of Sciences, 293-315.

Deb, K. (2001). Multi-objective optimization. *Multi-objective optimization using evolutionary algorithms*. 13-46.

Deb, K. and Agrawal, R. B. (1994). Simulated binary crossover for continuous search space. *Complex Systems*. 9, 1-34.

Deb, K. and Agrawal, R. B. (1995). Simulated binary crossover for continuous search space. *Complex systems*. 9(2), 115-148.

Deb, K., Agrawal, S., Pratap, A. and Meyarivan, T. (2000). A fast elitist non-dominated sorting genetic algorithm for multi-objective optimization: NSGA-II. *Lecture notes in computer science*. 1917, 849-858.

Deb, K. and Goldberg, D. E. (1989). An investigation of niche and species formation in genetic function optimization. In *Proceedings of the 1989 Proceedings of the 3rd International Conference on Genetic Algorithms*: Morgan Kaufmann Publishers Inc., 42-50.

Deb, K., Pratap, A., Agarwal, S. and Meyarivan, T. (2002). A fast and elitist multiobjective genetic algorithm: NSGA-II., *IEEE Transactions on Evolutionary Computation*. 6(2), 182-197.

Deb, K., Sindhya, K. and Okabe, T. (2007). Self-adaptive simulated binary crossover for real-parameter optimization. *Proceedings of the 2007 Proceedings of the*

Genetic and Evolutionary Computation Conference (GECCO-2007), UCL London: In Proceedings of the Genetic and Evolutionary Computation Conference (GECCO-2007), UCL London, 1187-1194.

Delgado, M., Cuéllar, M. P. and Pegalajar, M. C. (2008). Multiobjective hybrid optimization and training of recurrent neural networks. *Systems, Man, and Cybernetics, Part B:, IEEE Transactions on Cybernetics*. 38(2), 381-403.

Ding, S., Li, H., Su, C., Yu, J. and Jin, F. (2013). Evolutionary artificial neural networks: a review. *Artificial Intelligence Review*. 39(3), 251-260.

Ding, S., Su, C. and Yu, J. (2011a). An optimizing BP neural network algorithm based on genetic algorithm. *Artificial Intelligence Review*. 36(2), 153-162.

Ding, S., Xu, X., Zhu, H., Wang, J. and Jin, F. (2011b). Studies on Optimization Algorithms for Some Artificial Neural Networks Based on Genetic Algorithm (GA). *Journal of Computers*. 6(5), 939-946.

Dragoni, M., Azzini, A. and Tettamanzi, A. G. (2013). SimBa: A novel similarity-based crossover for neuro-evolution. *Neurocomputing*, 130, 108-122.

Dufo-López, R., Bernal-Agustín, J. L., Yusta-Loyo, J. M., Domínguez-Navarro, J. A., Ramírez-Rosado, I. J., Lujano, J. and Aso, I. (2011). Multi-objective optimization minimizing cost and life cycle emissions of stand-alone PV–wind–diesel systems with batteries storage. *Applied Energy*. 88(11), 4033-4041.

España-Boquera, S., Zamora-Martínez, F., Castro-Bleda, M. and Gorbe-Moya, J. (2007). Efficient BP algorithms for general feedforward neural networks. *Bio-inspired Modeling of Cognitive Tasks*. 327-336.

Falas, T. and Stafylopatis, A. G. (1999). The impact of the error function selection in neural network-based classifiers. *Proceedings of the 1999 Neural Networks, 1999. IJCNN'99. International Joint Conference on*: IEEE, 1799-1804.

Fawcett, T. (2006). An introduction to ROC analysis. *Pattern recognition letters*. 27(8), 861-874.

Feng, H., Tang, M. and Qi, J. (2011). A back-propagation neural network based on a hybrid genetic algorithm and particle swarm optimization for image compression.In *Proceedings of the 2011*: IEEE, 1315-1318.

Fernandez Caballero, J. C., Martínez, F. J., Hervás, C. and Gutiérrez, P. A. (2010). Sensitivity versus accuracy in multiclass problems using memetic Pareto evolutionary neural networks. *IEEE Transactions on Neural Networks*. 21(5), 750-770.

Fernández, J. C., Hervás, C., Martínez, F. J., Gutiérrez, P. A. and Cruz, M. (2009). Memetic Pareto differential evolution for designing artificial neural networks in multiclassification problems using cross-entropy versus sensitivity *Hybrid artificial intelligence systems*. 433-441. Springer.

Fieldsend, J. E. and Singh, S. (2005). Pareto evolutionary neural networks. *IEEE Transactions on Neural Networks*. 16(2), 338-354.

Fiszelew, A., Britos, P., Ochoa, A., Merlino, H., Fernández, E. and García-Martínez, R. (2007). Finding optimal neural network architecture using genetic algorithms. *Research in Computing Science Journal*. 27, 15-24.

Fonseca, C. M. and Fleming, P. J. (1993). Genetic algorithms for multiobjective optimization: Formulation, discussion and generalization. In *Proceedings of the 1993 Proceedings of the fifth international conference on genetic algorithms*: San Mateo, California, 416.

García-Pedrajas, N., Hervás-Martínez, C. and Muñoz-Pérez, J. (2002). Multi-objective cooperative coevolution of artificial neural networks (multi-objective cooperative networks). *Neural networks*. 15(10), 1259-1278.

García-Pedrajas, N., Hervás-Martínez, C. and Ortiz-Boyer, D. (2005). Cooperative coevolution of artificial neural network ensembles for pattern classification. *IEEE Transactions on Evolutionary Computation*. 9(3), 271-302.

Garcıa-Pedrajas, N., Ortiz-Boyer, D. and Hervás-Martınez, C. (2004). Cooperative coevolution of generalized multi-layer perceptrons. *Neurocomputing*. 56, 257-283.

García-Pedrajas, N., Sanz-Tapia, E., Ortiz-Boyer, D. and Hervás-Martínez, C. (2001). Introducing multi-objective optimization in cooperative coevolution of neural networks. *Connectionist Models of Neurons, Learning Processes, and Artificial Intelligence*. 645-652.

Garro, B. A., Sossa, H. and Vázquez, R. A. (2010). Design of artificial neural networks using differential evolution algorithm *Neural Information Processing. Models and Applications*. 201-208. Springer.

Goh, C.-K., Teoh, E.-J. and Tan, K. C. (2008). Hybrid multiobjective evolutionary design for artificial neural networks. *IEEE Transactions on Neural Networks*. 19(9), 1531-1548.

Goldberg, D. E. (1989). *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison-Wesley Longman Publishing Co., Inc.

Goldberg, D. E. and Richardson, J. (1987). Genetic algorithms with sharing for multimodal function optimization.In *Proceedings of the 1987 Proceedings of the Second International Conference on Genetic Algorithms on Genetic algorithms and their application*: L. Erlbaum Associates Inc., 41-49.

Hamid, N. A. (2012). Automated web-based user interfaces for novice programmers. In *Proceedings of the 50th Annual Southeast Regional Conference.* ACM 2012, 42-47.

Han, F., Gu, T. Y. and Ju, S. G. (2011). An improved hybrid algorithm based on PSO and BP for feedforward neural networks. *JDCTA: International Journal of Digital Content Technology and its Applications*. 5(2), 106-115.

Hervás, C., Silva, M., Gutiérrez, P. A. and Serrano, A. (2008). Multilogistic regression by evolutionary neural network as a classification tool to discriminate highly overlapping signals: Qualitative investigation of volatile organic compounds in polluted waters by using headspace-mass spectrometric analysis. *Chemometrics and Intelligent Laboratory Systems*. 92(2), 179-185.

Holland, J. H. (1975). *Adaptation in natural and artificial systems: An introductory analysis with applications to biology, control, and artificial intelligence*. U Michigan Press.

Huang, V. L., Zhao, S. Z., Mallipeddi, R. and Suganthan, P. N. (2009). Multi-objective optimization using self-adaptive differential evolution algorithm.In *Proceedings of the 2009 Evolutionary Computation, 2009. CEC'09. IEEE Congress on*: IEEE, 190-194.

Ilonen, J., Kamarainen, J.-K. and Lampinen, J. (2003). Differential evolution training algorithm for feed-forward neural networks. *Neural Processing Letters*. 17(1), 93-105.

Iorio, A. and Li, X. (2005). Solving rotated multi-objective optimization problems using differential evolution. *AI 2004: Advances in Artificial Intelligence*. 861-872.

Irani, R., Shahbazian, M. and Nasimi, R. (2011). Evolving neural network using real coded genetic algorithm for permeability estimation of the reservoir. *Expert Systems with Applications,* 38(8), 9862-9866.

Jiang, M., Deng, B., Wang, B. and Zhang, B. (2003). A fast learning algorithm of neural networks by changing error functions. *Proceedings of the 2003 Neural Networks and Signal Processing, 2003. Proceedings of the 2003 International Conference on*: IEEE, 249-252.

Jin, Y., Okabe, T. and Sendhoff, B. (2004). Neural network regularization and ensembling using multi-objective evolutionary algorithms. *Proceedings of the 2004 Evolutionary Computation, 2004. CEC2004. Congress on*: IEEE, 1-8.

Jin, Y. and Sendhoff, B. (2008). Pareto-based multiobjective machine learning: An overview and case studies. *Systems, Man, and Cybernetics, Part C: IEEE Transactions on Applications and Reviews*. 38(3), 397-415.

Jin, Y., Sendhoff, B. and Körner, E. (2005). Evolutionary multi-objective optimization for simultaneous generation of signal-type and symbol-type representations. *Proceedings of the 2005 Evolutionary Multi-Criterion Optimization*: Springer, 752-766.

Jin, Y., Sendhoff, B. and Körner, E. (2006). Simultaneous generation of accurate and interpretable neural network classifiers *Multi-Objective Machine Learning*. 291-312. Springer.

Kakde, M. M. R. O. G. (2004, December). Survey on multiobjective evolutionary and real coded genetic algorithms. In *Proceedings of the 8th Asia Pacific symposium on intelligent and evolutionary systems* 150-161.

Karegowda, A. G., Manjunath, A. and Jayaram, M. (2011). Application of genetic algorithm optimized neural network connection weights for medical diagnosis of pima Indians diabetes. *International Journal on Soft Computing*. 2(2), 15-23.

Khosrowshahi, F. (2011). Innovation in artificial neural network learning: Learn-On-Demand methodology. *Automation in Construction*, 20(8), 1204-1210.

Knowles, J. D. and Corne, D. W. (2000). Approximating the nondominated front using the pareto archived evolution strategy. *Evolutionary computation*. 8(2), 149-172.

Kukkonen, S. and Lampinen, J. (2004). An extension of generalized differential evolution for multi-objective optimization with constraints. In *Parallel Problem Solving from Nature-PPSN VIII*. 752-761. Springer Berlin Heidelberg.

Kukkonen, S. and Lampinen, J. (2005). GDE3: The third evolution step of generalized differential evolution.In *Proceedings of the 2005 Evolutionary Computation, 2005. The 2005 IEEE Congress on*: IEEE, 443-450.

Kuo, R. and Lin, L. (2010). Application of a hybrid of genetic algorithm and particle swarm optimization algorithm for order clustering. *Decision Support Systems*. 49(4), 451-462.

Lara, A., Sanchez, G., Coello Coello, C. A. and Schutze, O. (2010). HCS: a new local search strategy for memetic multiobjective evolutionary algorithms. *IEEE Transactions on Evolutionary Computation*. 14(1), 112-132.

Liu, G. and Kadirkamanathan, V. (1999). Multiobjective criteria for neural network structure selection and identification of nonlinear systems using genetic algorithms. *IEE Proceedings-Control Theory and Applications*. 146(5), 373-382.

Luna, J. M., Romero, J. R. and Ventura, S. (2013). Grammar-based multi-objective algorithms for mining association rules. *Data & Knowledge Engineering*. 86, 19-37.

Madavan, N. K. (2002). Multiobjective optimization using a Pareto differential evolution approach. *Proceedings of the 2002 Evolutionary Computation, 2002. CEC'02. Proceedings of the 2002 Congress on*: IEEE, 1145-1150.

Mashinchi, M. H. and Shamsuddin, S. M. H. (2009). Three-term fuzzy back-propagation *Foundations of Computational, Intelligence Volume 1*.143-158. Springer.

McLachlan, G. J., Do, K. A. and Ambroise, C. (2004). *Analyzing microarray gene expression data*. (Vol. 422) John Wiley & Sons.

Melesse, A., Ahmad, S., McClain, M., Wang, X. and Lim, Y. (2011). Suspended sediment load prediction of river systems: An artificial neural network approach. *Agricultural Water Management*, 98(5), 855-866.

Mezura-Montes, E., Reyes-Sierra, M. and Coello, C. A. C. (2008). Multi-objective optimization using differential evolution: a survey of the state-of-the-art *Advances in differential evolution* . 173-196. Springer.

Miguez, G., Xavier, A. E. and Maculan, N. (2014). An evaluation of the bihyperbolic function in the optimization of the backpropagation algorithm. *International Transactions in Operational Research*. 21(5), 835-854.

Mlakar, M., Petelin, D., Tušar, T. and Filipič, B. (2014). GP-DEMO: Differential Evolution for Multiobjective Optimization based on Gaussian Process models. *European journal of operational research*, 243 (2), 347‑361.

Nawi, N. M., Khan, A. and Rehman, M. Z. (2013). A new back-propagation neural network optimized with cuckoo search algorithm *Computational Science and Its Applications–ICCSA 2013* . 413-426. Springer.

Neri, F. and Cotta, C. (2012). Memetic algorithms and memetic computing optimization: A literature review. *Swarm and Evolutionary Computation*. 2, 1-14.

Ning, D., Zhang, W. and Li, B. (2008). Differential evolution based particle swarm optimizer for neural network learning. *Proceedings of the 2008*: IEEE, 4444-4447.

Ou, G. and Murphey, Y. L. (2007). Multi-class pattern classification using neural networks. *Pattern Recognition*. 40(1), 4-18.

Parsopoulos, K., Tasoulis, D., Pavlidis, N., Plagianakos, V. and Vrahatis, M. (2004). Vector evaluated differential evolution for multiobjective optimization. In *Proceedings of the 2004 Evolutionary Computation, 2004. CEC2004. Congress on*: IEEE, 204-211.

Peng, C. C. and Magoulas, G. D. (2011). Nonmonotone Levenberg–Marquardt training of recurrent neural architectures for processing symbolic sequences. *Neural Computing & Applications*. 20(6), 897-908.

Pettersson, F., Chakraborti, N. and Saxén, H. (2007). A genetic algorithms based multi-objective neural net applied to noisy blast furnace data. *Applied Soft Computing*. 7(1), 387-397.

Piotrowski, A. P. (2014). Differential Evolution algorithms applied to Neural Network training suffer from stagnation. *Applied Soft Computing*. 21, 382-406.

Qasem, S. N. and Shamsuddin, S. M. (2011). Memetic elitist pareto differential evolution algorithm based radial basis function networks for classification problems. *Applied Soft Computing*. 11(8), 5565-5581.

Qasem, S. N., Shamsuddin, S. M., Hashim, S. Z. M., Darus, M. and Al-Shammari, E. (2013). Memetic multiobjective particle swarm optimization-based radial basis function network for classification problems. *Information Sciences*. 239, 165-190.

Qasem, S. N., Shamsuddin, S. M. and Zain, A. M. (2011). Multi-objective hybrid evolutionary algorithms for radial basis function neural network design. *Knowledge-Based Systems,* 27, 475-497.

Ramesh, S., Kannan, S. and Baskar, S. (2011). Application of modified NSGA-II algorithm to multi-objective reactive power planning. *Applied Soft Computing*, 12(2), 741-753.

Reddy, M. J. and Kumar, D. N. (2006). Optimal reservoir operation using multi-objective evolutionary algorithm. *Water resources management*. 20(6), 861-878.

Ren, H., Ma, Y. and Dong, B. (2012). The Application of Improved GA-BP Algorithms Used in Timber Price Prediction. *Mechanical Engineering and Technology*. 803-810.

Rimer, M. and Martinez, T. (2006). CB3: an adaptive error function for backpropagation training. *Neural processing letters*. 24(1), 81-92.

Robič, T. and Filipič, B. (2005). DEMO: Differential evolution for multiobjective optimization. *Proceedings of the 2005 Evolutionary Multi-Criterion Optimization*: Springer, 520-533.

Rudolph, G. (1996). Convergence of evolutionary algorithms in general search spaces. *Proceedings of the 1996 In Proceedings of the Third IEEE Conference on Evolutionary Computation*: Citeseer,

Rumelhart, D., Hinton, G. and Williams, R. (1986). Learning Internal Representations by Error Propagation, Parallel Distributed Processing, Explorations in the Microstructure of Cognition, ed. DE Rumelhart and J. McClelland. Vol. 1. 1986. MIT Press, Cambridge, MA.

Sagar, G., Chalam, S. V. and Singh, M. K. (2011). Evolutionary Algorithm for Optimal Connection Weights in Artificial Neural Networks. *International Journal of Engineering (IJE)*. 5(5), 333.

Saman, F. I. (2006). *Three-term backpropagation algorithm for classification problem*, (Doctoral dissertation, Universiti Teknologi Malaysia, Faculty of Computer Science and Information System).

Schaffer, J. D. (1985). Multiple objective optimization with vector evaluated genetic algorithms. In *Proceedings of the 1985 Proceedings of the 1st international Conference on Genetic Algorithms*: L. Erlbaum Associates Inc., 93-100.

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

Shamsuddin, S. M., Sulaiman, M. N. and Darus, M. (2001). An improved error signal for the backpropagation model for classification problems. *International Journal of Computer Mathematics*. 76(3), 297-305.

Shim, V. A., Tan, K. C., & Tang, H. (2015). Adaptive memetic computing for evolutionary multiobjective optimization. *IEEE Transactions on Cybernetics*,45(4), 610-621.

Si, T., Hazra, S. and Jana, N. (2012). Artificial Neural Network Training Using Differential Evolutionary Algorithm for Classification.In *Proceedings of the 2012*: Springer, 769-778.

Slowik, A. and Bialko, M. (2008). Training of artificial neural networks using differential evolution algorithm. *Proceedings of the 2008 Human System Interactions, 2008 Conference on*: IEEE, 60-65.

Srinivas, N. and Deb, K. (1994). Muiltiobjective optimization using nondominated sorting in genetic algorithms. *Evolutionary computation*. 2(3), 221-248.

Storn, R. and Price, K. (1997). Differential evolution–a simple and efficient heuristic for global optimization over continuous spaces. *Journal of global optimization*. 11(4), 341-359.

Subudhi, B. and Jena, D. (2009). An improved differential evolution trained neural network scheme for nonlinear system identification. *International Journal of Automation and Computing*. 6(2), 137-144.

Tang, C., He, Y. and Yuan, L. (2011). A Fault Diagnosis Method of Switch Current Based on Genetic Algorithm to Optimize the BP Neural Network. *Electrical Power Systems and Computers*. 943-950.

Telfer, B. A. and Szu, H. H. (1994). Energy functions for minimizing misclassification error with minimum-complexity networks. *Neural networks*. 7(5), 809-817.

Tiwari, S., Fadel, G. and Deb, K. (2011). AMGA2: improving the performance of the archive-based micro-genetic algorithm for multi-objective optimization. *Engineering Optimization*. 43(4), 377-401.

Tiwari, S., Koch, P., Fadel, G. and Deb, K. (2008). AMGA: an archive-based micro genetic algorithm for multi-objective optimization. *Proceedings of the 2008 Proceedings of Genetic and Evolutionary Computation conference (GECCO-2008), Atlanta, USA*: Proceedings of Genetic and Evolutionary Computation conference (GECCO-2008), Atlanta, USA, 729-736.

Wang, H., Li, D., Zhao, Z., Qi, H. and Liu, L. (2011). The application of BP Neural Network based on improved PSO in BF temperature forecast. *Proceedings of the 2011*: IEEE, 2626-2629.

Wang, X., Tang, Z., Tamura, H. and Ishii, M. (2004). A modified error function for the backpropagation algorithm. *Neurocomputing*. 57, 477-484.

Wang, Y. and Qian, J. (2012). Measuring the uncertainty of RFID data based on particle filter and particle swarm optimization. *Wireless Networks*. 18(3), 307-318.

Wdaa, I. and Sttar, A. (2008). *Differential evolution for neural networks learning enhancement*, (Doctoral dissertation, Universiti Teknologi Malaysia, Faculty of Computer Science and Information System).

Wiegand, S., Igel, C. and Handmann, U. (2004). Evolutionary multi-objective optimisation of neural networks for face detection. International Journal of Computational Intelligence and Applications. 4(03), 237-253.

Wolpert, D. H. and Macready, W. G. (1997). No free lunch theorems for optimization. IEEE Transactions on Evolutionary Computation. 1(1), 67-82.

Wu, W., Liu, B. and Chen, T. (2010). Analysis of firing behaviors in networks of pulse-coupled oscillators with delayed excitatory coupling. Neural Networks. 23(7), 783-788.

Xiao, Z., Ye, S.-J., Zhong, B. and Sun, C.-X. (2009). BP neural network with rough set for short term load forecasting. *Expert Systems with Applications*. 36(1), 273-279.

Xiaoyuan, L., Bin, Q. and Lu, W. (2009). A New Improved BP Neural Network Algorithm. *Proceedings of the 2009 Intelligent Computation Technology and Automation, 2009. ICICTA'09. Second International Conference on*: IEEE, 19-22.

Xue, F., Sanderson, A. C. and Graves, R. J. (2003). Pareto-based multi-objective differential evolution. *Proceedings of the 2003 Evolutionary Computation, 2003. CEC'03. The 2003 Congress on*: IEEE, 862-869.

Xue, H. Y. and Ma, W. L. (2011). Research on Image Restoration Algorithm Base on ACO-BP Neural Network. *Key Engineering Materials*. 460, 136-141.

Yaghini, M., Khoshraftar, M. M. and Seyedabadi, M. (2012). Railway passenger train delay prediction via neural network model. *Journal of Advanced Transportation,* 47(3), 355-368.

Yan, W., Zhu, Z. and Hu, R. (1997). A hybrid genetic/BP algorithm and its application for radar target classification. *Proceedings of the 1997 Aerospace and Electronics Conference, 1997. NAECON 1997., Proceedings of the IEEE 1997 National*: IEEE, 981-984.

Yang, H., Mathew, J. and Ma, L. (2007). Basis pursuit-based intelligent diagnosis of bearing faults. *Journal of Quality in Maintenance Engineering*. 13(2), 152-162.

Yao, X. (1999). Evolving artificial neural networks. *Proceedings of the IEEE*. 87(9), 1423-1447.

Yi, J.-h., Xu, W.-h. and Chen, Y.-t. (2014). Novel Back Propagation Optimization by Cuckoo Search Algorithm. *The Scientific World Journal*. 2014.

Yu, Q. and Peng, J. (2012). Music Category Based on Adaptive Mutation Particle Swarm Optimization BP Neural Network. *Advances in Computer, Communication, Control and Automation*. 657-663.

Zeng, F., Low, M. Y. H., Decraene, J., Zhou, S. and Cai, W. (2010). Self-adaptive mechanism for multi-objective evolutionary algorithms. In *Proceedings of the International MultiConference of Engineers and Computer Scientist 2010 Vol 1*, IMECS 2010 March 17-19, 2010, Hong Kong.

Zhang, C., Ren, M. and Zhang, B. (2013). A self-adaptive multi-objective genetic algorithm for the QoS based routing and wavelength allocation problem in WDM network. *Optik-International Journal for Light and Electron Optics,* 124(20), 4571-4575.

Zhang, C., Shao, H. and Li, Y. (2000). Particle swarm optimisation for evolving artificial neural network. *Proceedings of the 2000 Systems, Man, and Cybernetics, 2000 IEEE International Conference on*: IEEE, 2487-2490.

Zitzler, E., Laumanns, M. and Thiele, L. (2001). SPEA2: Improving the strength Pareto evolutionary algorithm. Eidgenössische Technische Hochschule Zürich (ETH), Institut für Technische Informatik und Kommunikationsnetze (TIK).

Zitzler, E. and Thiele, L. (1999). Multiobjective evolutionary algorithms: A comparative case study and the strength pareto approach. *IEEE Transactions on Evolutionary Computation*. 3(4), 257-271.

Zweiri, Y. H. (2006). Optimization of a three-term backpropagation algorithm used for neural network learning. *International Journal of Computational Intelligence*. 3(4), 322-327.

Zweiri, Y. H. (2007). Optimization of a three-term backpropagation algorithm used for neural network learning. *Int J Comput Intell*. 3, 322-327.

Zweiri, Y. H., Seneviratne, L. D. and Althoefer, K. (2005). Stability analysis of a three-term backpropagation algorithm. *Neural Networks*. 18(10), 1341-1347.

Zweiri, Y. H., Whidborne, J. F., Althoefer, K. and Seneviratne, L. D. (2002). A new three-term backpropagation algorithm with convergence analysis. *Proceedings of the 2002 2002 IEEE International Conference on Robotics and Automation, May 11, 2002 - May 15, 2002* Washington, DC, United states: Institute of Electrical and Electronics Engineers Inc., 3882-3887.

Zweiri, Y. H., Whidborne, J. F. and Seneviratne, L. D. (2003). A three-term backpropagation algorithm. *Neurocomputing*. 50, 305-318.