FUZZY NEURAL NETWORKS WITH GENETIC ALGORITHM-BASED LEARNING METHOD

M. REZA MASHINCHI

UNIVERSITI TEKNOLOGI MALAYSIA

FUZZY NEURAL NETWORKS WITH GENETIC ALGORITHM-BASED LEARNING METHOD

M. REZA MASHINCHI

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

MARCH 2013

To my father, and family members

ACKNOWLEDGEMENT

I am deeply grateful to my supervisor, Prof. Dr. Ali Selamat, for his detailed and constructive comments and for his understanding and important support throughout this work. Also, my sincere thanks are due to the official examiners Assoc. Prof. Dr. Muhammad Suzuri Hitam, and, Prof. Dr. Siti Mariyam Shamsuddin for their detailed review, constructive criticism and excellent advice during the preparation of this thesis. I would also like to thank my father and brother for their constructive academic sharing.

ABSTRACT

This thesis is on the reasoning of artificial neural networks based on granules for both crisp and uncertain data. However, understanding the data in this way is difficult when the data is so complex. Reducing the complexity of the problems that these networks are attempting to learn as well as decreasing the cost of the learning processes are desired for a better prediction. A suitable prediction in artificial neural networks depends on an in-depth understanding of data and fine tracking of relations between data points. Inaccuracies of the prediction are caused by complexity of data set and the complexity is caused by uncertainty and quantity of data. Uncertainties can be represented in granules, and the reasoning based on granules is known as granular computing. This thesis proposed an improvement of granular neural networks to reach an outcome from uncertain and crisp data. Two methods based on genetic algorithms (GAs) are proposed. Firstly, GA-based fuzzy granular neural networks are improved by GA-based fuzzy artificial neural networks. They consist of two parts: granulation using fuzzy c-mean clustering (FCM), and reasoning by GA-based fuzzy artificial neural networks. In order to extract granular rules, a granulation method is proposed. The method has three stages: construction of all possible granular rules, pruning the repetition, and crossing out granular rules. Secondly, the two-phase GA-based fuzzy artificial neural networks are improved by GA-based fuzzy artificial neural networks. They are designed in two phases. In this case, the improvement is based on alpha cuts of fuzzy weight in the network connections. In the first phase, the optimal values of alpha cuts zero and one are obtained to define the place of a fuzzy weight for a network connection. Then, in the second phase, the optimal values of middle alpha cuts are obtained to define the shape of a fuzzy weight. The experiments for the two improved networks are performed in terms of generated error and execution time. The results tested were based on available rule/data sets in University of California Irvine (UCI) machine learning repository. Data sets were used for GA-based fuzzy granular neural networks, and rule sets were used for GA-based fuzzy artificial neural networks. The rule sets used were customer satisfaction, uranium, and the datasets used were wine, iris, servo, concrete compressive strength, and uranium. The results for the two-phase networks revealed the improvements of these methods over the conventional one-phase networks. The two-phase GA-based fuzzy artificial neural networks improved 35% and 98% for execution time, and 27% and 26% for the generated error. The results for GA-based granular neural networks were revealed in comparison with GA-based crisp artificial neural networks. The comparison with other related granular computing methods were done using the iris benchmark data set. The results for these networks showed an average performance of 82.1%. The results from the proposed methods were analyzed in terms of statistical measurements for rule strengths and classifier performance using benchmark medical datasets. Therefore, this thesis has shown GA-based fuzzy granular neural networks, and GA-based fuzzy artificial neural networks are capable of reasoning based on granules for both crisp and uncertain data in artificial neural networks.

ABSTRAK

Tesis ini menyelidik taakulan bagi rangkaian neural buatan berdasarkan granul untuk kedua-dua data jelas dan tidak jelas. Kaedah pemahaman data melalui cara ini adalah sukar apabila kandungan data adalah kompleks.Untuk mengurangkan kekompleksan masalah yang cuba dipelajari oleh rangkaian ini dan juga mengurangkan kos proses pembelajarannya, teknik ramalan yang lebih baik adalah diperlukan. Ramalan yang sesuai dalam rangkaian neural buatan bergantung kepada kebolehan untuk memahami isi kandungan data dengan mendalam dan juga kebolehan untuk mengenal pasti hubungan antara data. Ketakpastian dan kepelbagaian jenis data juga akan menjadikan hasil ramalan yang tidak tepat. Ketakpastian terhadap jenis data disebabkan oleh kekompleksan jenis data tersebut dan juga set data yang mengandungi tahap ketakpastian yang kompleks. Ketakpastian boleh diwakili dengan granul dan taakulan yang dikenali sebagai pengkomputeran granular. Tesis ini menggunakan rangkaian neural granular untuk mencapai hasil daripada data yang tidak pasti dan jelas. Dua kaedah telah diperkembangkan berdasarkan algoritma genetik. Rangkaian granular kabur berasaskan algoritma genetik (GA) telah diperkembangkan menggunakan rangkaian neural buatan kabur berasaskan GA. Rangkaian neural granular kabur berasaskan GA mengandungi dua bahagian: granulasi menggunakan pengelompokan min-c kabur, dan taakulan oleh rangkaian neural buatan kabur berasaskan GA. Untuk mengekstrak peraturan granular kaedah granulasi yang diterokai mengandungi tiga peringkat, iaitu pembinaan semua peraturan granular yang mungkin, pemangkasan data yang berulang dalam set data dan pengurangan peraturan granular yang telah digunakan. Rangkaian neural buatan kabur berasaskan GA berfasa dua telah direka bentuk dalam dua fasa. Dalam keadaan ini peningkatannya berdasarkan kepada nilai *alfa-cut* dalam rangkaian neural. Dalam fasa pertama nilai optimum *alfa-cut* adalah sifar dan boleh diperoleh bagi menentukan pemberat set kabur kepada rangkaian neural. Dalam fasa kedua nilai optimum untuk alfa-cut tengah diperoleh untuk menentukan bentuk set kabur. Uji kaji untuk dua rangkaian neural yang telah dipertingkatkan telah dijalankan berdasarkan kepada jumlah ralat yang dihasilkan dan masa yang diambil bagi melaksanakan uji kaji tersebut. Hasil uji kaji berdasarkan set data mesin pembelajaran repositori di University of California Irvine (UCI). Set data yang digunakan untuk rangkaian neural granular kabur berasaskan GA dan set peraturan digunakan untuk rangkaian neural buatan kabur. Set peraturan yang digunakan adalah kepuasan pelanggan dan uranium manakala set data yang digunakan ialah arak, iris, servo, kekuatan mampat konkrit dan uranium. Hasil untuk rangkaian dua fasa mendedahkan keunggulan kaedah ini berbanding dengan rangkaian konvensional satu fasa. Rangkaian dua fasa telah meningkat sebanyak 35% dan 98% untuk masa pelaksanaan dan 27% dan 26% untuk ralat umum. Hasil untuk rangkaian neural berasaskan GA didedahkan berbanding dengan rangkaian neural buatan jelas berasaskan GA. Sementara itu perbandingan dengan kaedah pengkomputeran granular yang lain yang berkaitan dijalankan menggunakan set data penanda aras iris. Hasil untuk rangkaian ini menunjukkan prestasi purata sebanyak 82.1%. Hasil daripada kaedah yang disarankan telah dianalisis dari segi statistik, kekuatan aturan dan pengkelasan tenaga menggunakan penanda aras set data perubatan. Oleh itu tesis ini menunjukkan bahawa rangkaian neural granular kabur berasaskan GA dan rangkaian neural buatan kabur berasaskan GA mampu sebagai taakulan bagi rangkaian neural buatan berdasarkan granul untuk kedua-dua data jelas dan tidak jelas.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS

$\widetilde{w}$	-	Fuzzy weights of fuzzy artificial neural networks
$\bigcup_{i=0}^{n} i$	-	Union i, where $i = 0\ to\ n$
$\sum_{i=0}^{n} i$	-	Summation i, where $i = 0\ to\ n$
α_l^L	-	Left boundary of α-cut at level l
α_l^R	-	Right boundary of α-cut at level l
$\tilde{b}$	-	Fuzzy valued bias
$\tilde{X}_n$	-	Fuzzy valued input
$\tilde{O}_q$	-	q^{th} fuzzy outcome
$\widetilde{Inter}$	-	Intersection of edges in FANNs
$a\widetilde{g}g$	-	Aggregation of edges in FANNs
F_P	-	Proposed fitness function
F_A	-	Aliev fitness function
D_l^L	-	Left boundaries of α-cuts target values
D_l^R	-	Left boundaries of α-cuts for target values
O_l^L	-	Left boundaries of α-cuts outcome values
O_l^R	-	Left boundaries of α-cuts outcome values
$f_{net}(\tilde{X}_1; \tilde{X}_2)$	-	Actual outcome from input $\tilde{X}_1$, $\tilde{X}_2$

LIST OF ABBREVIATION

ANN	-	Artificial Neural Network
CANN	-	Crisp Artificial Neural Network
PR	-	Probabilistic Reasoning
EC	-	Evolutionary Computing
NC	-	Neuro Computing
FL	-	Fuzzy Logic
FANN	-	Fuzzy Artificial Neural Network
FIG	-	Fuzzy Information Granulation
GC	-	Genetic Computing
GA	-	Genetic Algorithm
GrC	-	Granular Computing
CGNN	-	Crisp Granular Neural Network
FGNN	-	Fuzzy Granular Neural Network
GNN	-	Granular Neural Network
GBLM	-	Genetic Based Learning Method
2P-GBLM	-	2 Phase Genetic Based Learning Method
H	-	High
L	-	Low
TFIG	-	Theory of Fuzzy Information Granulation
TIG	-	Theory of Information Granulation
MSE	-	Mean of Squared Error
SE	-	Squared Error
SC	-	Soft Computing

LIST OF APPENDICES

CHAPTER 1

INTRODUCTION

1.1 Introduction

This chapter presents a brief introduction to the notions that are used to achieve the aim of this thesis. The aim is to use granular computing for problems solving. The introduced notions are granular computing, granular neural networks, fuzzy artificial neural networks, fuzzy information granulation, generic algorithms and fuzzy sets theory. The reason for this study is to investigate the aim of artificial computations that is to solve a problem with the least amount of cost and the best accuracy. Problems become more difficult to be solved when their corresponding data sets are large or contain uncertain information. In the literature of artificial computations, there are many nature inspired computations such as evolutionary computations, artificial neural networks, artificial immune systems, swarm intelligence, etc. (Zomaya, 2006) (Kari, et al., 2008). However, there are two issues behind the proposed algorithms; which are the ability of each algorithm to solve only a particular type of problem; and their performance in solving the problem. To overcome these two issues, this thesis uses granular computing with the aid of learning and optimization mechanisms for an optimal learning from granular rules. Therefore, granular neural networks have been used for learning mechanism; meanwhile fuzzy artificial neural networks are centered in the granular neural networks. Also, the genetic algorithms (GA) are used to increase the performance of fuzzy artificial neural networks. Therefore, GA-based fuzzy artificial neural networks are used in the main part for granular neural networks.

In order to have better performance and wide applicability of computational methods, different individual models need to be unified. This has emerged in granular computing as an inspiration from the human mind. Among natural computations as the source of inspiration for artificial computations, the human mind is known to be superior in problem solving, for example processing large and also uncertain information. The superiority of mind is spotted when an attempt is made to solve a complex problem. A problem is called complex when its components or their relations are difficult to understand. Since uncertainties are considered complexity, a problem consisting of uncertainties becomes complex. Solving complex problems is promised by the granular computing model, which has been inspired from the way that the human mind solves such problems. A major key in granular computing is granulation, which is similar to abstraction in the human mind (Zadeh, 1997). Granulation and granular computing are two similar meanings to information granulations. Historically, the notion of granular computing emerged from the theory of information granulation, which initially was proposed based on fuzziness. Similar to granular computing, the theory of fuzzy information granulation is inspired from how the human mind reacts to solve problems (Zadeh, 1997). It is known as a key in the mind-processing mechanism; a system must consider it in solving a complex problem (Zadeh, 1997).

Therefore, in this thesis, fuzzy modeling has been used for granular computation to solve complex problems. Fuzzy artificial neural networks, which are used in the center of improved granular neural networks, are the combination of fuzzy modeling and artificial neural networks. Fuzzy artificial neural networks are able to predict based on learned instances (Liu, et al., 2004) (Arotaritei, 2011). Their accuracy is highly dependent on tuned weights of network connections, and thus, the connection weights need to be optimally adjusted. This has been done in GA-based fuzzy artificial neural networks, where a genetic algorithm is used to adjust the weights (Aliev, et al., 2001). The genetic algorithms are combined with fuzzy artificial neural networks due to their strong appeal as an optimization technique (Liu, et al., 2004). The optimizations in genetic algorithms are based on the evolutionary process of generations. From an overall view, genetic algorithms are used to identify the best outcome of the results based on the optimization of the fuzzy artificial neural networks.

1.2 Problem background

Artificial neural networks have already been positioned as an important class of non-linear systems. They are highly adaptable systems that have been used for many areas of applications. Crisp artificial neural networks are primarily geared toward the intensive processing of numerical data. However, as the dimensionality of the problems increases, the computational complexity becomes visible because the sizes of the data sets grow up rapidly (Pedrycz, et al., 2001). Therefore, the weak points of crisp artificial neural networks are identified as follows:

(i) Inability to tackle large-scale data, which causes inefficient learning within these networks, and thus predicts unsuitable outcomes.
(ii) Inability to process non-numeric data, such as uncertainties and linguistic variables.

Regarding issue (i), there are a few ways to solve the problem. One solution is modular architectures of neural networks, which helps to break down the problem into a series of simpler subtasks and each of the subtasks is handled independently. The modularization of problems is a viable way of exploring neural architectures in the long run. Breaking down the problem for modulation is a method of simplification, which is similar to the granulation of data. Therefore, granulation is another way to solve issue (i) instead of using the concept of modularization. The granulation in the computation process can be used to solve also issue (ii). There are few granulation approaches; however, fuzzy modeling is the most appealing. The use of fuzzy models in conjunction with artificial neural networks creates granular neural networks, which can solve either issue (i) or (ii). See Figure 1.1 for the location of granular neural networks.

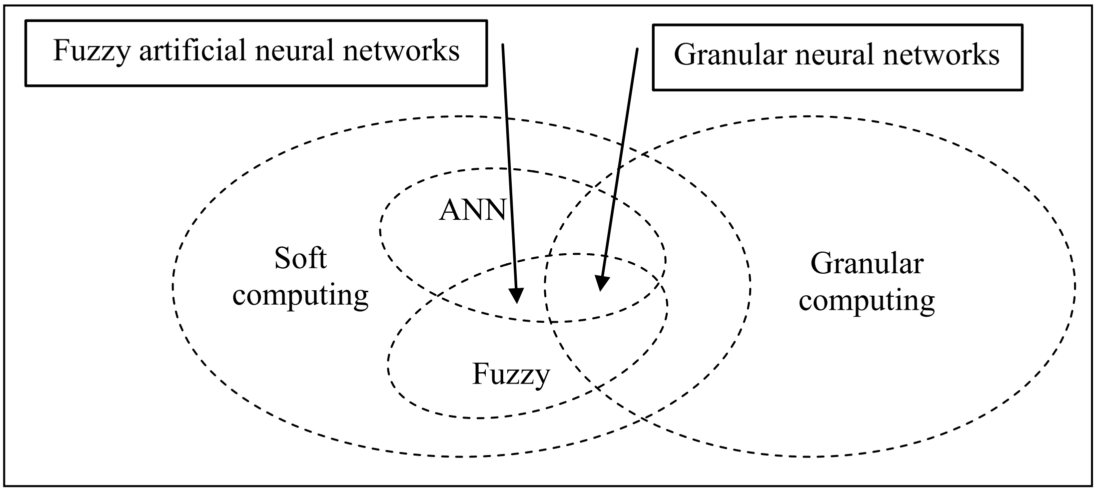

Figure 1.1: Granular neural networks and fuzzy neural networks in conjunction with soft computing and granular computing areas.

A potential solution for solving complex problems is granular computing. Since 1997, the proposal by L.A Zadeh that described granular computing, this method has become an attractive research area. However, most of the studies in this field are theoretic which needs to be studied in actual applications. Table 1.1 presents a few studies on granular computing, and the remaining issues.

Table 1.1: Related researches on granular computing and descriptions of each from the soft computing perspective.

Study	Proposed method	The problems remained unsolved
Information granulation	Proposing the idea of generalizing the information (Zadeh, 1979).	The research remained as an idea.
Reconstruction of information granulation concept	Formalization of information granulation (Zadeh, 1997).	The research does not present implementation of information granulation concept.
Labeling granular computing	Formalization of information granulation (Lin, 1997).	The research does not present implementation of granular computing.
Foundation of granular neural networks	Joint the concepts of information granulation and neural networks (Lin, 1997).	Remained as an idea.
GA-based fuzzy granular neural networks (this study)	Improves the idea of granular neural networks proposed in (Pedrycz, et al., 2001) by this study.	Decreasing the complexity of granulation when dataset is very huge.

Based on Table 1.1, this study improves the granular neural based on GA-based fuzzy artificial neural networks. Furthermore, an information granulation has been investigated to extract the fuzzy granules from the crisp data sets. Therefore, in this thesis, fuzzy sets theory has been used for granular neural networks due to the following reasons:

(i) Fuzzy sets support the modeling of the concepts that exhibit continuous boundaries (Pedrycz, et al., 2001) (Hans-Jurgen, et al., 2012). Continuous boundaries are used to represent fuzzy values on real numbers. Fuzzy values are used to tune fuzzy artificial neural networks by genetic algorithms.

(ii) Fuzzy sets exhibit well-defined semantics and fully meaningful conceptual entities from building modules that are identified in problem solving (Pedrycz, et al., 2001) (Hans-Jurgen, et al., 2012). The originality of the source problem is preserved after modeling it in the fuzzy concept. Meanwhile, it is possible to simplify the fuzzy modeled problem using the concept of granulation. This has been illustrated by two examples in Figure 1.2 and Figure 1.3, respectively with the illustrated modeling of a function in crisp and fuzzy representations. An advantage to fuzzy modeling, as shown in Figure 1.3, is that it does not have the crispness when moving from one segment as a concept to another. This can benefit a rule-based system to decrease the effects from a noisy data.

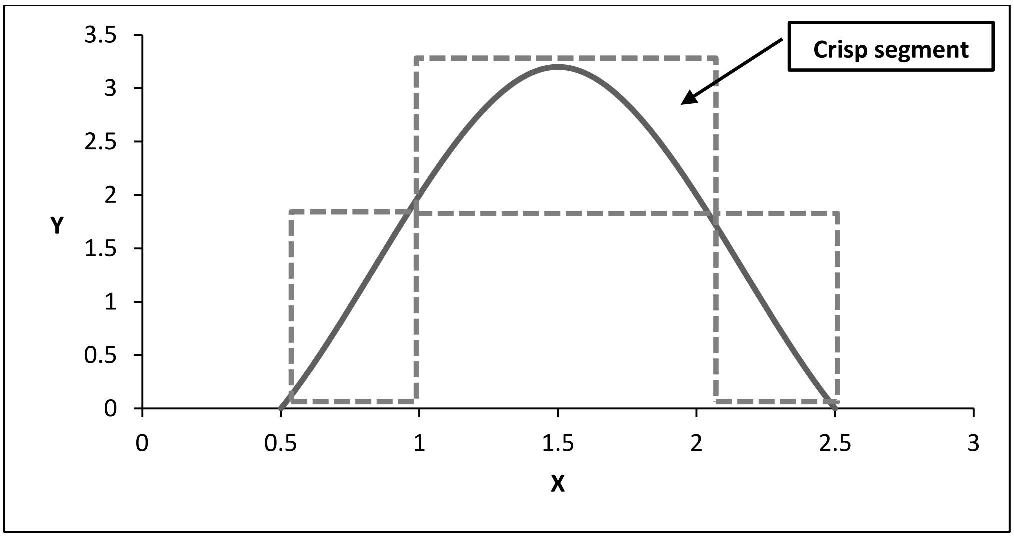

Figure 1.2: An example of crisp modeling in dealing with a function.

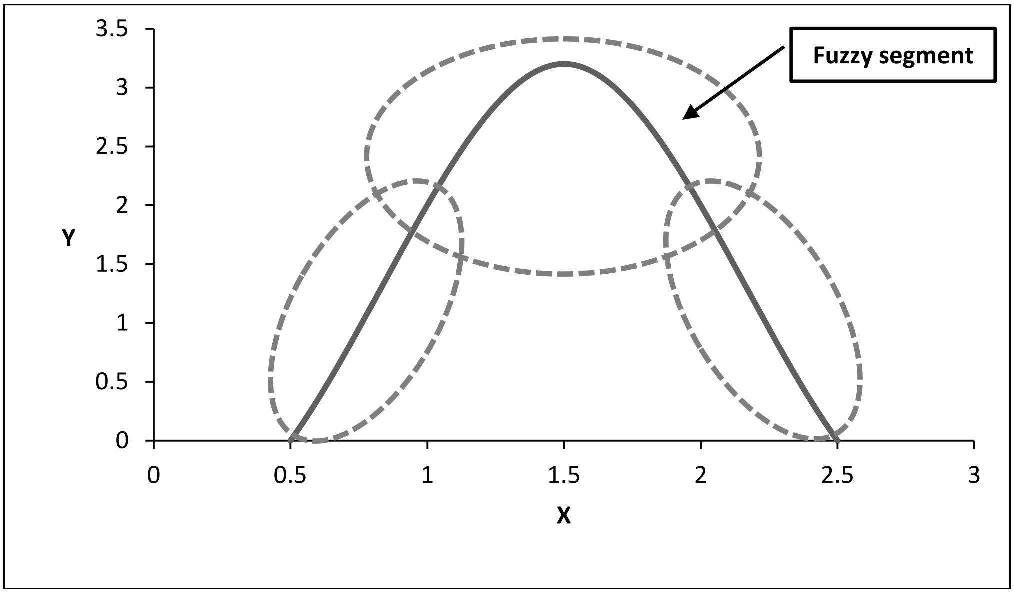

Figure 1.3: An example of fuzzy modeling in dealing with a function.

Due to the above advantages, (i) and (ii), fuzzy modeling has been combined with artificial neural networks to improve the learning mechanism (Zomaya, 2006). This is called as fuzzy artificial neural networks, which are used in this thesis as the main part of granular neural networks. As the fuzzy artificial neural networks are successfully used in learning and approximate reasoning, improvements are still required to increase their performance. There are some issues that need to be considered as follows:

(i) Finding the global solution: The existing methods are usually being trapped into local minima when searching for the optimal network as shown in Figure 1.4. An accurate learning method is needed to avoid local minima and to find an optimal network similar to what is ideal (Arotaritei, 2011). This becomes more notable when the number of alpha cuts is increased to shape the fuzzy numbers. Having more alpha cuts increases the possibility of trapping in the local minima.

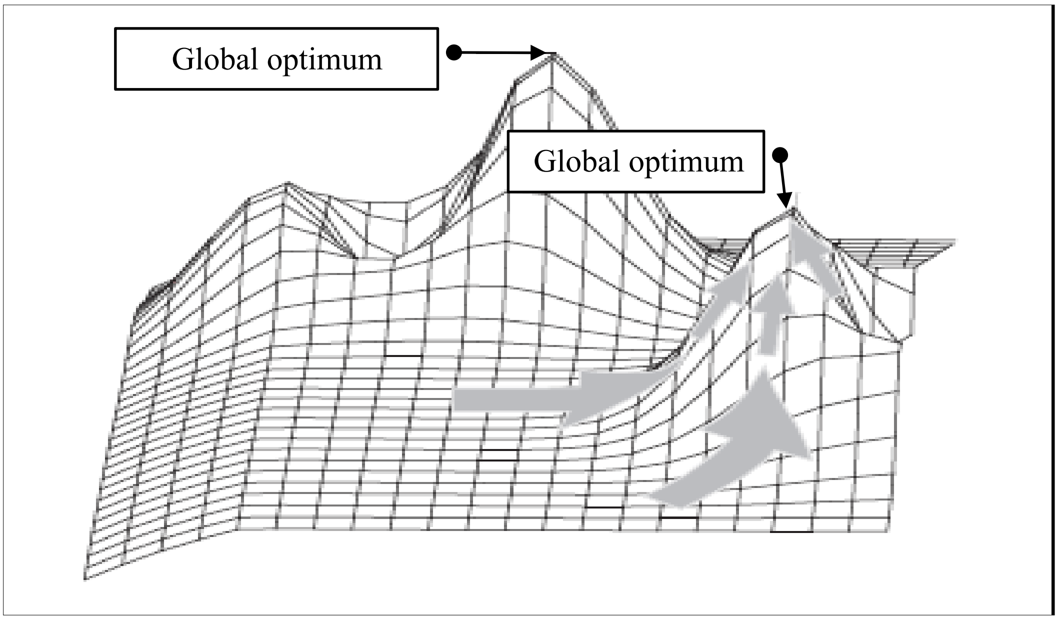

Figure 1.4: Different optimum explorations for fitness landscape (Weise, 2007).

(ii) Increasing the convergence speed: The execution time in finding the optimal network needs to be considered in two cases. First, when learning and reasoning the outcome for an online application. Second, when a well-shaped fuzzy number is needed for a better outcome. Figures 1.5 and 1.6 shows well-shaped fuzzy value with several alpha-cuts and a triangular fuzzy value in two alpha cuts (Lee, 2005) (Hans-Jurgen, et al., 2012). The processing of a well-shaped fuzzy value is more time consuming due to two consequent reasons. First, numerous alpha cuts need to be optimized. Then, they must be optimized within an unconstrained searching space; because the optimal value for each boundary of each alpha-cut can be any real number.

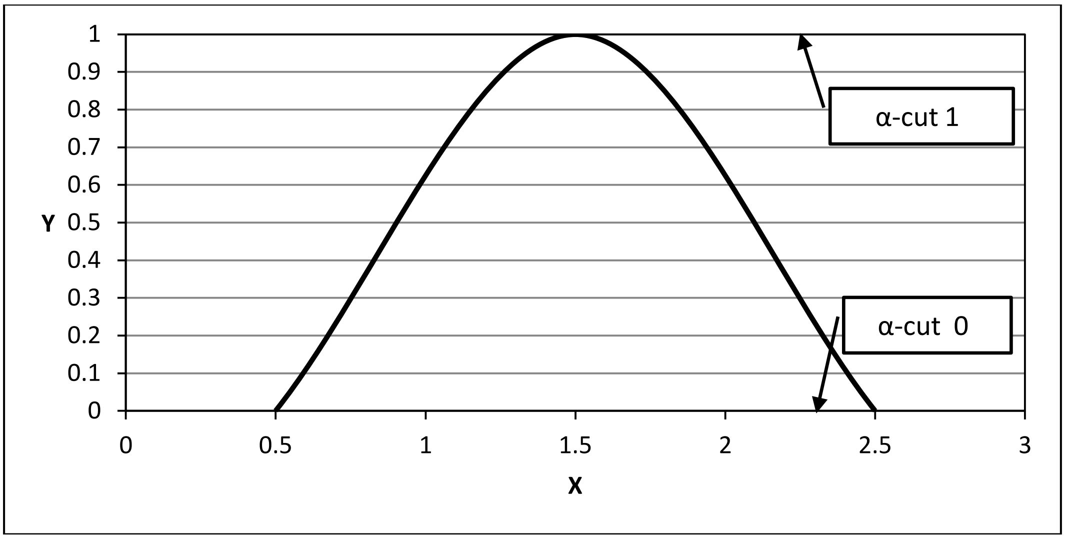

Figure 1.5: A well-shaped fuzzy value represented with eleven α-cuts.

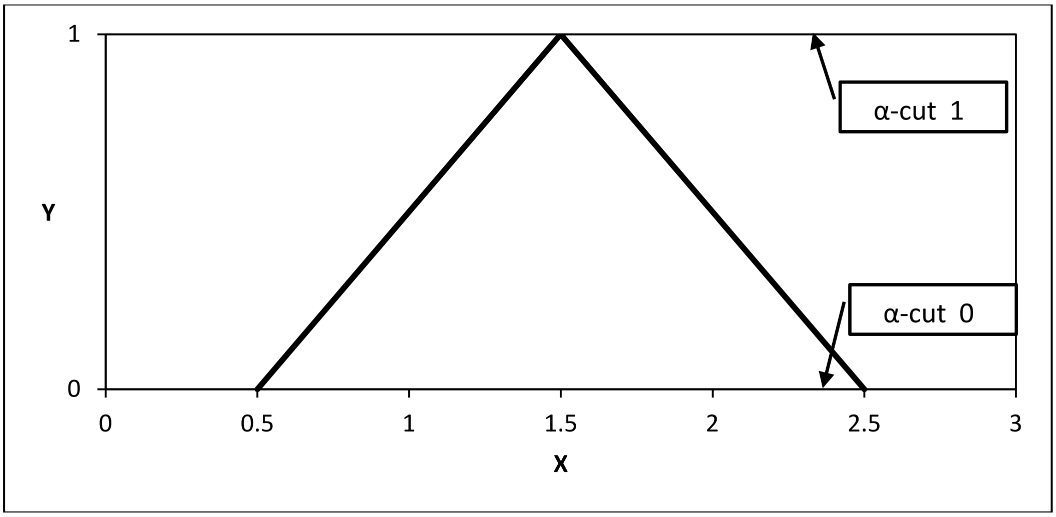

Figure 1.6: A triangular fuzzy value with two α-cuts.

(iv) Dealing with all types of fuzzy values: Most of the existing learning methods for fuzzy artificial neural networks are unable to deal with all kinds of fuzzy numbers (Arotaritei, 2011). This would result in a lack of these networks being applicable to different applications. Therefore, a suitable fuzzy artificial neural network is required to deal with all types of fuzzy numbers as shown in Figure 1.7 (Lee, 2005).

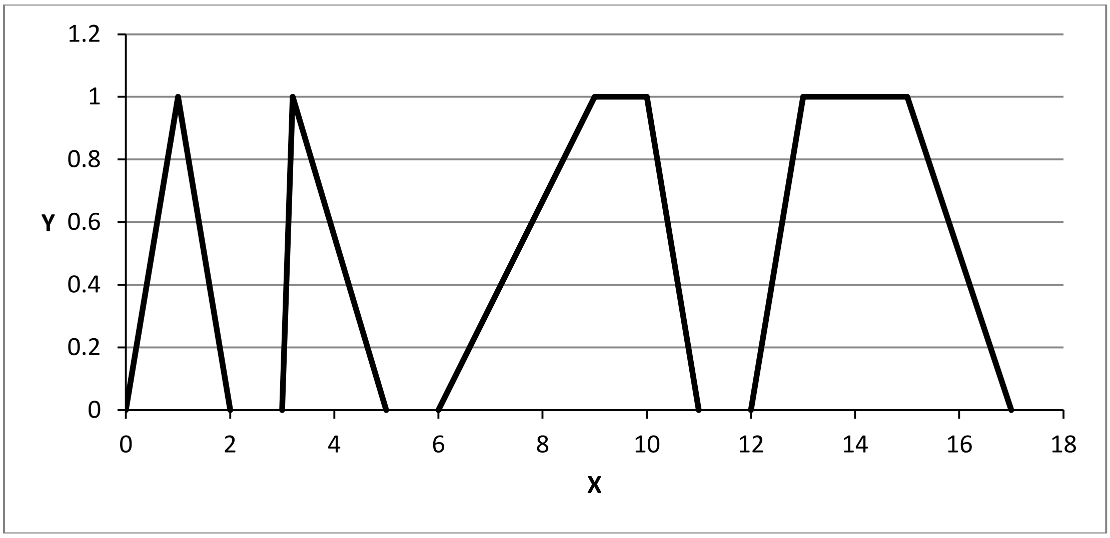

Figure 1.7: Different types of fuzzy numbers, triangular and trapezoidal in symmetrical and asymmetrical representation.

Based on the above-mentioned issues for fuzzy artificial neural networks, there are some major methods that have been proposed in the literatures. First, the direct fuzzification is proposed that fuzzifies the delta rule from the feed-forward artificial neural networks (Buckley, et al., 1992, 1995), (Hayashi, et al., 1993). This method has been rejected from a theoretical point of view (Liu, et al., 2004). Later, some learning approaches for triangular symmetric fuzzy values were proposed (Ishibuchi, et al., 1995, 2001). However, they are not able to deal with other bounded convex types of fuzzy values. Consequently, an approach based on the derivation of the min-max function has been proposed (Zhang, et al., 1996, 1999), (Liu, et al., 2004, 2005) to deal with all types of bounded convex fuzzy values. In all of the above-mentioned methods, avoiding trapping in local minima is not promised. This is due to using the back propagation algorithm, which is a local optimizer.

Another method used in studying fuzzy artificial neural networks is based on genetic algorithms. Buckley et al (Buckley, et al., 1994) introduced the use of genetic algorithms for the first time, and then Aliev et al. (Aliev, et al., 2001) reintroduced it. The reason of these two studies was using the ability of genetic algorithms to improve the fuzzy artificial neural networks. A recent works has been done for these networks by (Arotaritei, 2011). This method has the ability to deal with all types of fuzzy values that are bounded and convexed. However, the speed of learning convergence still needs to be improved. The summary of mentioned methods is given in Table 1.2; also, a description for each method is given in Table 1.3 and Table 1.4 in terms of learning and the speed of convergences.

Table 1.2: List of a few proposed learning methods for fuzzy artificial neural networks.

Researcher	Learning method		
	Name	Advantages	Disadvantages
Buckley and Hayashi (1993)	Direct fuzzification	It proposes fuzzy artificial neural networks for the first time, which fuzzifies delta rule form feed-forward artificial neural networks.	It has been rejected from mathematical aspect.
Zhang, et al. (1996)	Derivation of min-max function	It improves fuzzy artificial neural networks to deal with all types of fuzzy values.	It does not have an adequate learning ability.
Aliev et al. (2001)	GA-based FANNs	It applies the idea of using genetic algorithms for fuzzy artificial neural networks, which were firstly sparked in 1994. It could successfully deal with all types of fuzzy values.	It does not have an adequate speed of learning convergence to be applicable for real applications.
Mashinchi (2007)	Two-phase FANNs based on BP	Proposing FANNs that learn fuzzy rule set in two phases.	Low learning convergence due to using local optimizer in the first stage.
Mashinchi, et al. (2009)	Three-term fuzzy back-propagation	The proposed method enhances the speed of back-propagation based FANNs by adding a fuzzy proportional factor.	High possibility of trapping into local minima for learning convergence process.
Arotaritei (2011)	Local crossover GA-based FANNs	Improves according to feed-forward (FFNR) and fuzzy recurrent networks (FRNN).	Needs to be compared with other methods.
This study	Two-phase GA-based FANNs	It improves the conventional GA-based method by Aliev in terms of generated error and execution time.	It does not have the ability of learning from well-shaped fuzzy values in less time.

Table 1.3: Learning convergences of the methods in Table 1.2.

Method	Learning convergence
Buckley and Hayashi (1993)	It is remained as an idea without implementation.
Derivation of min-max function (Zhang, et al., 1996)	It has an acceptable convergence for triangular fuzzy values; however, it is less promising since it is based on back propagation as a local optimizer. Main studies on this method are done by: (Zhang, et al., 1996, 1999), (Liu, et al., 2004, 2005). The remain issues to be done are as follows: • To guarantee the convergence • To keep the accuracy for well-shaped fuzzy values
GA-based FANNs (Aliev, et al., 2001)	It has acceptable convergence for triangular fuzzy values with more promising since it is based on GA as a global optimizer. Main studies on this method are done by: (Buckley, et al., 1994), (Aliev, et al., 2001). The remain issue to be done is as follows: • To keep the accuracy for well-shaped fuzzy values
Two-phase FANNs based on back-propagation (Mashinchi, 2007)	It has acceptable convergence in compared with its based method, BP; however, it is less promising since it is based on back propagation as a local optimizer. The remain issue to be done is as follows: • To guarantee the convergence • To keep the accuracy for well-shaped fuzzy values
Three-term fuzzy back-propagation (Mashinchi, et al., 2009)	It has acceptable convergence in compared with its based method, BP; however, there is more possibility of trapping into local minima for learning process as BP is a local optimizer. The remain issue to be done is as follows: • To guarantee the convergence To keep the accuracy for well-shaped fuzzy values
Local crossover GA-based FANNs (Arotaritei, 2011)	It has acceptable convergence for triangular fuzzy values with more promising since it is based on GA as a global optimizer. The remain issue to be done is as follows: • To keep the accuracy for well-shaped fuzzy values
Two-phase GA-based FANNs (this study)	It has better convergence with promising results, since it benefits from the optimization features of genetic algorithms. In addition, it keeps the convergence accuracy for well-shaped fuzzy values.

Table 1.4: Speed of convergences of the methods in Table 1.2.

Method	Speed of convergence
Buckley and Hayashi (1993)	It is remained as an idea without implementation.
Derivation of min-max function (Zhang, et al., 1996)	It has acceptable speed of convergence in learning triangular fuzzy values, however, there is a remained issue to be done as follows: • To keep the speed of convergence for well-shaped fuzzy values
GA-based FANNs (Aliev, et al., 2001)	It has acceptable convergence for triangular fuzzy values, however, there is a remained to be done as follows: • To keep the accuracy for well-shaped fuzzy values
Two-phase FANNs based on back-propagation (Mashinchi, 2007)	It has better speed of convergence in learning well-shaped fuzzy values in compared with its base method back-propagation. The remained to be done as follows: • To keep the accuracy for well-shaped fuzzy values
Three-term fuzzy back-propagation (Mashinchi and Shamsuddin, 2009)	It has better speed of convergence in compared with its base method back-propagation due to adding a fuzzy proportional factor. The remained to be done is as follows: • To keep the accuracy for well-shaped fuzzy values
Local crossover GA-based FANNs (Arotaritei, 2011)	It has acceptable convergence for triangular fuzzy values, however, there is a remained to be done as follows: • To keep the accuracy for well-shaped fuzzy values
Two-phase GA-based FANNs (this study)	It has better speed of convergence in learning well-shaped fuzzy values.

In order to improve the performance of fuzzy artificial neural networks, this thesis proposes an improved GA-based fuzzy artificial neural network to overcome the speed and the learning convergence when dealing with well-shaped fuzzy values.

1.3 Problem statement

Complex problems are known to be difficult to solve since understanding them is not easy. An example of complex problems is the unsteadiness of things such as linguistic variables. Involving uncertainties and enlargement of the problems to be tackled cause the problems to be complex. To solve complex problems, simplification of the problems and learning mechanisms can be taken. Respectively, fuzzy granulation and fuzzy artificial neural networks can be used. The combination of these two has emerged in granular neural networks, where fuzzy artificial neural networks can play an important role. If the performance of fuzzy artificial neural networks is increased, therefore, the performance of granular neural networks can be improved. Meanwhile, the complex problem needs to be simplified in the form of a fuzzy rule base to be fed to fuzzy artificial neural networks.

Therefore, the hypothesis of this thesis is stated as follows:

> *"How a complex problem can be solved by granular neural networks and how fuzzy artificial neural networks can collaborate to increase the performance of granular neural networks in solving complex problems."*

1.4 Thesis aim

The aim of this thesis is to solve the complexity and uncertainty of data sets using GA-based fuzzy artificial neural networks, where reasoning based on crisp data sets is considered in granularity. In this regard, granular neural networks aim to be improved by GA-based fuzzy artificial neural networks. In addition, a rule extraction is used to transform the crisp data set into a fuzzy granular rules base. Due to the impact of GA-based fuzzy artificial neural network in learning from the granules, the efficiency of these networks on GA-based fuzzy granular neural networks are investigated. Consequently, enhancing GA-based fuzzy artificial neural networks is considered to improve the performance of granular neural networks. The improvement of GA-based fuzzy artificial neural networks is due to two reasons: the low accuracy of these networks and their centrality in GA-based fuzzy granular neural networks.

1.5 Thesis objectives

The objectives of the thesis are defined as follows:

(i) To propose the reasoning on crisp data sets in fuzzy artificial neural networks using genetic algorithms.

(ii) To propose the transformation of crisp data sets into fuzzy granular rule bases to train fuzzy artificial neural networks.

(iii) To evaluate the performance of proposed fuzzy artificial neural networks based on genetic algorithms with their conventional method and crisp artificial neural networks.

1.6 Thesis scope

The scope of this thesis is defined as follows:

(i) Three standard fuzzy rule sets and six crisp data sets are used to test the improved methods. The rule sets, and also uncertain data sets, are used to test the two-phase GA-based fuzzy artificial neural networks, and the data sets are used to test the fuzzy granular neural networks.

(ii) The rules/data sets are available in University of California Irvine (UCI) machine learning repository, and literatures. Aliev (Aliev, et al., 2001), Liu (Liu, et al., 2005) and customer satisfaction fuzzy rule sets (Fasanghari, et al., 2008), and also a uranium data set with uncertain values (Staudenrausch, et al., 2005), are used to test two-phase GA-based fuzzy artificial neural networks. Meanwhile, the crisp data sets to test GA-based fuzzy granular neural networks are chosen based on size. Wine, servo and iris data sets from UCI repository as well as a uranium data set (Houston, et al., 1987) from the literature are used as small sizes, and concrete compressive strength data set from UCI repository is used as medium size.

(iii) Comparisons for the improved methods are done in terms of generated errorr, execution time, and distribution. To this end, improved two-phase GA-based fuzzy artificial neural networks are compared with corresponding conventional method, and improved GA-based fuzzy granular neural networks are compared with GA-based crisp artificial neural networks.
(iv) Implementations for GA-based fuzzy artificial neural networks are done using Matlab, and the implementations for GA-based fuzzy granular neural networks and GA-based artificial neural networks are done using Microsoft C++.

1.7 Significance of thesis

The significance of this thesis is as follows:

(i) Fuzzy artificial neural networks based on genetic algorithms have been studied and improved using the notion of granular neural networks.
(ii) Fuzzy artificial neural networks can be studied using available data sets for application, and they have been improved in terms of generated error and execution time.
(iii) A granular rule extraction is proposed to simplify a crisp data set and represent it in fuzzy form.

1.8 Contribution of thesis

This thesis contributes to problem solving via soft computing by granulation. A granular rule extraction is given to simplify the problems, and granular learning approaches are improved for reasoning based on these granules. More specifically, three methods have been proposed in this thesis as follows:

(i) Fuzzy artificial neural networks are able to learn from crisp data sets, and conversely, crisp data sets can be processed by fuzzy artificial neural networks.

(ii) GA-based fuzzy artificial neural networks are improved in terms of generated error and execution time. The improved method can learn well-shaped fuzzy values represented in alpha-cuts.

(iii) A granulation method is improved to extract fuzzy granular rules from a crisp data set. It prepares the collaboration of crisp data sets with fuzzy artificial neural networks.

1.9 Thesis plan

The direction of the thesis is taken from research background of the author on solving complex problems by human cognition. Here, granular computing and soft computing are chosen due to their similarities with human cognition. Reasoning based on granules is considered for a crisp data set, and thus a granulation method has been presented. The granulation method uses fuzzy representation, and reasoning is based on fuzzy granules. Respectively, fuzzy granulation and fuzzy artificial neural networks are used, and genetic algorithms aid in improving the learning performance. Drawing these three methods under the model of granular computing constructs GA-based granular neural networks as the main idea. The constructed method is presented to achieve the aim and objectives of this thesis. In addition, GA-based fuzzy artificial neural networks are improved due to their impact on enhancing the GA-based fuzzy granular neural networks. Eventually, contribution of this thesis is carried out for improved methods by comparison.

1.10 Organization of thesis

This chapter presents the framework of this thesis. It introduces the research field and explains the reasons for the study by reviewing the problem background. Five chapters are organized to meet the scopes and objectives. These are the introduction, a literature review, methodology, the results of the proposed fuzzy artificial neural networks, an analysis of the proposed fuzzy granular neural networks and a conclusion.

The organization of individual chapters is as follows:

(i) Chapter 1 presents an overview of the thesis. The problem under study is introduced and the reasons for such study are described. The reasoning process by using granules for crisp data sets is highlighted. The background to the problem is given, and previously utilized approaches published in available literature are reviewed. The objectives and scope of the study are defined along with the aim of thesis. Finally, a synopsis of the likely contribution of the study is given according to the defined aims and objectives.

(ii) Chapter 2 introduces the data sets that are used for the experiments. It also provides a review of soft computing, granular computing and the hybrid techniques that have arisen between them. Fuzzy sets theory and genetic algorithms are discussed with respect to soft computing, and fuzzy information granulation and granular neural networks are discussed with respect to granular computing.

(iii) Chapter 3 discusses the methodology of improved methods that is used to carry out this study. Improved GA-based granular neural networks are presented, including the concepts of granulation and GA-based fuzzy artificial neural networks. The granulation method is presented in order to extract the fuzzy granular rules. The general architecture of fuzzy artificial neural networks and the schema of improved GA-based fuzzy artificial neural networks are also presented.

(iv) Chapter 4 presents the detailed implementation and results of the improved granular neural networks. The constructed algorithms of granulation are detailed, including the fuzzy C-mean, granular rules extraction, granular rule contraction, pruning the granular rules and the defuzzifier.

(iv) The improved method is tested using four standard crisp data sets, which are the wine, servo, iris and concrete compressive strength data sets. The results of improved GA-based granular neural networks are compared with those of other methods in the literature using the benchmarked iris data set. Using the rest of the data sets, they are compared with the results of corresponding GA-based crisp artificial neural networks to contrast granular reasoning against crisp reasoning. This chapter also presents the detailed implantation and results of improved GA-based fuzzy artificial neural networks that are given in Chapter 4.

(v) Chapter 5 reveals the analysis results for distribution, rule strength, and classifier performance. The results are obtained from the proposed methods: two-phase GA-based fuzzy artificial neural networks, granulation, and GA-based fuzzy granular neural networks. The results are carried out in compared with one-phase GA-based fuzzy artificial neural networks and GA-based crisp artificial neural networks, which are based on three data sets: Liver disorder, hepatitis, and diabetes.

(vi) Chapter 6 discusses the work that has been done to complete this study. The proposed approaches are discussed and the work is summarized to conclude the results and analysis of Chapter 4. Finally, the future work is addressed in this chapter.

1.11 Summary

This chapter introduces the thesis by presenting the framework of the study. six chapters are defined to obtain the aims and objectives. Some notions are reviewed, such as granular computing, granular neural networks, fuzzy information granulation, fuzzy artificial neural networks, fuzzy sets theory and genetic algorithms. These notions are discussed according to their relation to the three improved methods of this thesis. The improved methods are two-phase GA-based fuzzy artificial neural networks, GA-based fuzzy granular neural networks and granulation. Particularly, the major studies in the literature for granular neural networks and GA-based fuzzy artificial neural networks are reviewed. Finally, the contributions of this thesis are given according to the aims and objectives, and the organization of the thesis has been presented.

REFERENCES

Aliev R.A., Fazlollahi B. and Vahidov R.M. (2001). Genetic Algorithm-based Learning of Fuzzy Neural Network. Part 1. Feed-forward Fuzzy Neural Networks, Fuzzy Sets and Systems. 118, pp. 351–358.

Angelov P., Xiaowei Z. and Klawonn F. (2007). Evolving Fuzzy Rule-based Classifiers. IEEE Symposium on Computational Intelligence in Image and Signal Processing. pp. 220-225.

Arotaritei D. (2011). Genetic Algorithm for Fuzzy Neural Networks using Locally Crossover. International Journal of Computers, Communications & Control. 5, pp. 8-20, ISSN 1841-9836.

Bargiela A. and Pedrycz W. (2002). Granular Computing: An Introduction. Boston: Kluwer Academic Publishers.

Brereton R.G. (2009). Steepest Ascent, Steepest Descent, and Gradient Methods. Comprehensive Chemometrics. 1, pp. 577-590.

Buckley J.J., Reilly K.D, and Hayashi Y. (1994). Genetic Learning Algorithms for Fuzzy Neural Nets. IEEE World Congress on Computational Intelligence. 3, pp. 1969–1974.

Buckley J.J. and Hayashi Y. (1992). Fuzzy Neural Nets and Applications. Fuzzy Sets and AI. 1, pp. 11–41.

Buckley J.J. and Hayashi Y. (1995). Neural Nets for Fuzzy Systems. Fuzzy Sets and Systems. 71, pp. 265–276.

Chelouah R. and Siarry P. (2003). Genetic and Nelder-Mead Algorithms Hybridized for a More Accurate Global Optimization of Continuous Multiminima Functions. European Journal of Operational Research. 148, pp. 335-348.

Chen M. S., Han J. and Yu P. S. (1996). Data mining: An Overview From Database Perspective. IEEE Transactions on Knowledge and Data Engineering. 8, pp. 866–883.

Collani E. V. (2008). Defining and Modeling Uncertainty. Journal of Uncertain Systems. 2, pp. 202-211.

Daniel F. L., Pyramo C. Jr. and Gomi F. (2009). Evolving Granular Classification Neural Networks. IEEE International Joint Conference on Neural Networks. pp. 1736-1743.

Der-Chiang L., Chiao-Wen L. and Hu S. C. (2010). A Learning Method for the Class Imbalance Problem with Medical Datasets. Computers in Biology and Medicine. 40, pp. 509-518.

Dick S. and Kandel A. (2001). Granular Weights in a Neural Network. IFSA World Congress and 20th NAFIPS International Conference. 3, pp. 1708-1713.

Fasanghari M. and Roudsari F. H. (2008). The Fuzzy Evaluation of E-Commerce Customer Satisfaction. World Applied Sciences Journal. 4, pp. 164-168.

Fasanghari R. M. and Habibipour F. (2008). The Fuzzy Evaluation of E-Commerce Customer Satisfaction. World Applied Sciences Journal. 4(2), pp. 164-168.

Fawcett T. (2006). An Introduction to ROC Analysis. Pattern Recognition Letter. 27, pp. 86-874.

Gabrys B. and Bargiela A. (2000). General Fuzzy Min-Max Neural Network for Clustering and Classification. 11(3), pp. 769-783.

Giunchglia F. and Walsh T. (1992). A Theory of Abstraction. Artificial Intelligence. 323-390 , pp. 56.

Guangfei Y. (2011). A Novel Evolutionary Method to Search Interesting Association Rules by Keywords. Expert Systems with Applications. 38, pp. 13378-13385.

Hamrouni T., Yahia S. B. and Nguifo E. M. (2009). Sweeping the Disjunctive Search Space Towards Mining New Exact Concise Representations of Frequent Itemsets. Data and Knowledge Engineering. 10, pp. 1091–1111.

Hayashi Y., Buckley J.J. and Czogola E. (1993). Fuzzy Neural Networks with Fuzzy Signals and Weights, International Joural of Intelligent Systems. 8, pp. 527-537.

He H. and Garcia A. E. (2009). Learning from Imbalanced Data, IEEE Transactions on knowledge and Data Engineeting. 21(9), pp. 1263-1284.

He H. (2009) Learning from Imbalanced Data. IEEE Transactions on Knowledge and Data Engineeting. 21(9), pp.1263-1284.

Hobbs J.R. Granularity (1985). Proceedings of the Ninth International Joint Conference on Artificial Intelligence. pp. 432-435.

Hoheisela T., Kanzowa C., and Mordukhovich B. S. (2012). Generalized Newton’s Method Based on Graphical Derivatives. Nonlinear Analysis: Theory, Methods & Applications. 75(3), pp. 1324-1340.

Holland H. J. (1975). Adaptation in Natural and Artificial Systems, University of Michigan Press.

Holland H. J. Genetic Algorithms (1992). Scientific America, 267, pp. 66-72.

Houston R.S. and Karlstrom D.E. (198). Application of The Time and Strata Bound Model for The Origin of Uranium Bearing Southeastern Wyoming. Deposits in Proterozoic Quartz-pebble Conglomerates, Report of The Working Group on Uranium Geology Organized By The International Atomic Energy Agency. - USA : IAEA-TECDOC-427.

Hans-Jurgen S. and Antonsson E. K. (2012). Fuzzy Sets in Engineering Design and Configuration. International Series in Intelligent Technologies. 9, ISBN: 978-0-7923-9802-8.

Ishibuchi H. and Nii M. (2001). Numerical Analysis of Learning of Fuzzified Neural Networks from Fuzzy If-Then Rules. Fuzzy Sets and Systems. 120, pp. 281–307.

Ishibuchi H., Kwon K. and Tanakan H. A (1995). Learning Algorithm of Fuzzy Neural Networks with Triangular Fuzzy Weights. Fuzzy Sets and Systems. 71, pp. 277-293.

Jea K. F. and Chang M. Y. (2008). Discovering Frequent Itemsets by Support Approximation and Itemset Clustering. Data and Knowledge Engineering. 65(1), pp. 90–107.

Kari L. and Rozenberg G. (2008). The Many Facets of Natural Computing. 51(10), pp. 72-83

Karmarkar N. (1984). A New Polynomial Time Algorithm for Linear Programming. Combinatorica, 4(4), pp. 373–395.

Karush W. (1939). Minima of Functions of Several Variables with Inequalities as Side Constraints. Chicago : Department of Mathematics, University of Chicago.

Kasabov N. and Woodford B. (1999). Rule insertion and Rule Extraction from Evolving Fuzzy Neural Networks: Algorithms and Applications for Building Adaptive, Intelligent Exert Systems. IEEE International Fuzzy Systems. 3, pp. 1406-1411.

Klee V. and Minty G. J. (1972). How Good Is The Simplex Algorithm?. NewYork: Academic Press, pp. 159–175.

Kuhn H. W. and Tucker A. W. (1951). Nonlinear Programming, Second Berkeley Symposium. Berkeley: University of California Press, pp. 481-492.

Kwak K. and Pedrycz W. (2010). A Design of Genetically Oriented Linguistic Model with the Aid of Fuzzy Granulation. IEEE International Conference on Fuzzy Systems. pp. 1-6.

Lam P. S., Bouzerdoum A. and Hoang N. G. (2009). Learning Pattern Classification Tasks with Imbalanced Data Sets. Pattern recognition, pp. 193-208.

Lee K. H. (2005). First Course on Fuzzy Theory and Applications. Advances in Soft Computing. Springer Berlin / Heidelberg.

Li P. and Li H.X. (2005). Approximation Analysis of Feed Forward Regular Fuzzy Neural Network with Two Hidden Layers. Fuzzy Sets and Systems. 150, pp. 373–396.

Li Y. and Yu F. (2010). Optimized Fuzzy Information Granulation Based Machine Learning Classification. Seventh International Conference on Fuzzy Systems and Knowledge Discovery (FSKD). 1, pp. 259-263.

Lilliefors H. W. (1967). On the Kolmogorov-Smirnov Test for Normality with Mean and Variance Unknown. Journal of the American Statistical Association. 62, pp. 399-402.

Lilliefors H. W. (1969). On the Kolmogorov-Smirnov Test for the Exponential Distribution with Mean Unknown. Journal of the American Statistical Association. 64, pp. 387–389.

Lin T. Y., Yao Y. Y. and Zadeh L. A. (2002). Data Mining, Rough Sets and Granular Computing. Germany: Physica-Verlag GmbH Heidelberg, pp. 110-1537.

Lin T.Y. (1997). From Rough Sets and Neighborhood Systems to Information Granulation and Computing in Words. European Congress on Intelligent Techniques and Soft computing. pp. 1602-1607.

Liu P. and Li H.X. (2005). Approximation Analysis of Feed Forward Regular Fuzzy Neural Network with Two Hidden Layers. Fuzzy Sets and Systems. 150, pp. 373–396.

Liu P. and Li H.X. (2004). Fuzzy Neural Network Theory and Application, World Scientific, River Edge, NJ.

Liua X., Zhaia K. and Pedrycz W. (2012). An Improved Association Rules Mining Method. Expert Systems with Applications. 39(1), pp. 1362-1374.

Mashinchi M. H. (2007). A New Fuzzy Back-propagation Learning Method Based on Derivation of Min-max Function Tuned with Genetic Algorithms. Master Thesis, Universiti Teknologi Malaysia.

Mashinchi M. H. and Shamsuddin S. M. (2009). Three-term Fuzzy Back-Propagation. Foundation of Computational Intelligence. 201, pp 143-158.

Maydeu-Olivares A. and Garca-Forero C. (2012). Goodness-of-Fit Testing. International Encyclopedia of Education. International Encyclopedia of Education Oxford: Elsevier, pp. 190-196.

Nakagawa S., Hashiguchib H. and Naoto N. (2012). A Measure of Skewness for Testing Departures from Normality. Computational Statistics and Data Analysis. To Be Appeared.

Narushima Y. and Yabeb H. (2012). Conjugate Gradient Methods Based on Secant Conditions that Generate Descent Search Directions for Unconstrained Optimization. Journal of Computational and Applied Mathematics. 236(17), pp. 4303-4317.

Panoutsos G. A., Mahfous M., Mills G. H., and Brown B. H. (2010). Generic Framework For Enhancing The Interpretability of Granular Computing-based Information, IEEE International Conference on Digital Object Identifier. pp. 19-24.

Pawlak Z. (1987). Rough logic. Bulletin of Polish Academic Science and Technology. 35, pp. 253-258.

Pedrycz W. and Gacek A. (2002). Temporal Granulation and its Application to Signal Analysis. Information Sciences. 143, pp. 47.

Pedrycz W. and Valente D. O. J. (2008). A Development of Fuzzy Encoding and Decoding Through Fuzzy Clustering. IEEE Transactions on Instrumentation and Measurement. 57(4), pp. 829-837.

Pedrycz W. and Vukovich G. (2001). Granular Neural Networks. Neurocomputing. 36, pp. 205-224.

Pedrycz W. (2001). Granular Computing: An Emerging Paradigm. Heidelberg: Physica-Verlag,.

Raeder T., Forman G. and Chawla N. V. (2012). Learning from Imbalanced Data: Evaluation Matters. Data Mining: Foundations and Intelligent Paradigns. Intelligent Systems Reference Library. 23, pp. 315-331.

Reformat M. and Pedrycz W. (2001). IEEE Evoluionary Optimization of Information Granules. IFSA World Congress and 20th NAFIPS International Conference. 4, pp. 2035-2040.

Sijtsma K. and Emons W. H. M. (2010). Nonparametric Statistical Methods. International Encyclopedia of Education, Elsevier, pp. 347-353.

Simpson P. (1992). Fuzzy Min-Max Neural Networks. Part I: Classification. IEEE Tranactions on Neural Networks. 3-5, pp. 776-786.

Sivanandam S. N. and Deepa S. N. (2008). Introduction to Genetic Algorithms. Springer Berlin Heidelberg.

Staudenrausch S. (2005). Diversity of The Ectomycorrhiza Community at a Uranium Mining Heap, Biol Fertil Soils. 41, pp. 439-446.

Tsoukalas L. H and Uhrig R. E. (1997). Fuzzy and Neural Approaches in Engineering, New York: John Wiely & Sons Inc.

Weise T. (2007). Global Optimization Algorithms – Theory and Application. Thomas Weise, Online available as e-book at http://www.it-weise.de.

Yao Y. Y. and Yao J. T. (2002). Induction of Classification Rules by Granular Computing. Rough Sets and Current Trends in Computing, Proceedings of the Third International Conference on Rough Sets and Current Trend in Computing. pp. 331-338.

Yao Yiyu (2010). Human-Inspired Granular Computing. Novel Development in Granular Computing: Applications for Advanced Human Reasoning and Soft Computation. 212, pp. 15-143.

Yeung D. S. (2010). Sensitivity Analysis for Neural Networks. Verlag Berlin Heidelberg: Springer.

Yu F. and Cai R. (2010). Optimized Fuzzy Information Granulation of Temporal Data. Seventh IEEE Internation Conference on Fuzzy Systems and Knowledge Discovery. 1, pp. 419-423.

Zadeh L.A. (1965). Fuzzy Sets. Information and Control. 8, pp. 338-359.

Zadeh L.A. (1979). Fuzzy Sets and Information Granularity. Advances in Fuzzy Set Theory and Applications. Editors: M. Gupta R. Ragade, R. Yager. Amsterdam: North-Holland, pp. 3-18.

Zadeh L.A. (1973). Outline of a New Approach to the Analysis of Complex System and Decision Processes. IEEE Transactions Systems Man Cybernetics. 3 pp. 28-44.

Zadeh L. A. (1997). Toward a Theory of Fuzzy Information Granulation and its Centrality in Human Reasoning and Fuzzy Logic. Fuzzy Sets and Systems. 90 pp. 111-127.

Zhang X. (1999). An Efficient Computational Algorithm for Min-Max Operations. Fuzzy Sets and Systems. 104, pp. 297–304.

Zhang X., Hang C.C. and Wang P.Z. (1996). The Min-Max Function Differentiation and Training Of Fuzzy Neural Networks. IEEE Transactions on Neural Networks. 7, pp. 1139–1150.

Zhao Y., Yao Y. and Yao J. T. (2006). Level-wise Construction of Decision Trees for Classification. International Journal of Software Engineering and Knowledge Engineering, World Scientific Publishing, 16(1), pp. 103-126.

Zheng R.T. (2005), A Staged Continuous Tabu Search Algorithm for the Global Optimization and its Applications to the Design of Fiber Bragg Gratings, Computational Optimization and Applications. 30, pp. 319–335.

Zhou Z., Jiang Y. and Chen S. (2000). A General Neural Framework for Classification Rule Mining. International Journal of Computers, Systems, and Signals. 1(2), pp. 154-168.

Zomaya A. Y. (2006). Handbook of Nature-inspired and Innovative Computing Integrating Classical Models with Emerging Technologies. Springer, ISBN: 978-0-387-27705-9.