View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Repository@USM

INDIVIDUAL AND ENSEMBLE PATTERN CLASSIFICATION MODELS USING ENHANCED FUZZY MIN-MAX NEURAL NETWORKS

by

MOHAMMED F. M.

Thesis submitted in fulfillment of the

requirements for the degree of

Doctor of Philosophy

February 2014

ACKNOWLEDGMENT

الحمد لله، والصلاة والسلام على رسول الله، وعلى آله وأصحابه أجمعين ومن اهتدى بهداه

I wish to thank and praise, first and foremost, Allah SWT who bestowed upon me by facilitating the successful accomplishment of this research.

During many years of research, I have gained invaluable knowledge that has guided me for my research. Humbly, I acknowledge that my success is not only related to me but its possession to all the people whom guide me, teach me, prayer for me and encourage me. From them, I learned the knowledge, patient, wisdom, humility and also how to gain access to my goals.

It is with immense gratitude that I acknowledge the help of my supervisor Professor Chee Peng Lim who provided me inspiration, enthusiasm, and motivation, as well as for his support, patience, and encouragement throughout my graduate studies. It is not often that one finds an advisor who always finds the time for listening to the little problems and roadblocks that unavoidably crop up in the course of performing research. His technical and editorial advice was essential to the completion of this thesis and has taught me innumerable lessons and insights on the workings of academic research in general. It was an honor to work with him. I would like to express the deepest appreciation also for Assc. Prof. Dr. Umi Kalthum bt Ngah for giving help and advising support.

I would like to express the deepest appreciation to School of Electrical and Electronics, USM for providing me the necessary facilities, equipment, as well as the GA support, which made this research possible. I would like also to thank the School's very helpful staff. I am also grateful to all my friends from Universiti Sains Malaysia, for being the surrogate family during the duration of my research. Also, I am grateful to all my friends in my dear wounded country (Iraq), to all of them, thank you so much for continuously contacting me and give your undivided encouragement.

Last, but not least, I would also like to dedicate my gratitude to the dearest and nearest people to my heart, my father - Falah bin Mohammed, my mother - Sana binti Jamel, my beloved wife-Sura binti Alaa' and my little baby Tasnem, for giving me support, love, and happiness in difficult and critical situations. This thesis is for you. Finally, to all my family members, I am forever indebted for your understanding, support, endless patience and encouragement when it matters the most. You are all like bright candles in the dark days and difficult times, so thank again.

TABLE OF CONTENTS

Acknowledgment ii

Table of Contents iv

List of Tables viii

List of Figures xii

List of Abbreviations xv

Abstrak… xvii

Abstract… xix

CHAPTER 1 - INTRODUCTION

1.1 Background 1

1.2 Computational Intelligence 3

1.3 Problems and Motivations 5

1.4 Research Aim and Objectives 9

1.5 Overview of Methodology 10

1.6 Thesis Outline 14

CHAPTER 2 - LITERATURE REVIEW

2.1 Introduction 16

2.2 Background of the Fuzzy Min-Max Neural Network 16

2.3 Ensemble Methods 24

2.4 Multi Agent Systems 31

2.5 Summary 36

CHAPTER 3 - AN ENHANCED FUZZY MIN-MAX NEURAL NETWORK

3.1 Introduction 37

3.2 The Fuzzy Min-Max Neural Network 38

3.2.1 Dynamics of the Fuzzy Min-Max Neural Network 38

3.2.2 Learning in FMM 41

3.2.3 A Numerical Example 45

3.3 Analysis of the FMM Learning Algorithm 48

3.3.1 Hyperbox Expansion 48

3.3.2 Hyperbox Overlap Test 51

3.3.3 Hyperbox Contraction 53

3.4 The Enhanced FMM Network 53

3.4.1 The Hyperbox Expansion Rule 53

3.4.2 The Hyperbox Overlap Test Rule 56

3.4.3 The Hyperbox Contraction Rule 58

3.5 Performance Evaluation 60

3.5.1 Case Study I 62

3.5.2 Case Study II 65

3.5.3 Case Study III 66

3.6 Summary 70

CHAPTER 4 - FURTHER IMPROVEMENTS OF THE ENHANCED FUZZY MIN-MAX NEURAL NETWORK

4.1 Introduction 71

4.2 Review of the Pruning Strategy 72

4.3 Analysis of the EFMM Learning Algorithm 74

4.3.1 Network Complexity 75

4.3.2 Noise 78

4.4 The Proposed EFMM2 Network 79

4.4.1 The *k*-nearest hyperbox expansion rule 79

4.4.2 Pruning Strategy 83

4.5 Performance evaluation 85

4.5.1 Case Study I 85

4.5.1.1 Different training set sizes 86

4.5.1.2 Adding Different Noise Levels to the Training Sets 91

4.5.2 Case Study II 93

4.5.2.1 Simulations with Varying Hyperbox Sizes 94

4.5.2.2 Comparison with Other Published Results 98

4.6 Summary 105

CHAPTER 5 - A MULTI-AGENT CLASSIFIER SYSTEM WITH ENHANCED FUZZY MIN-MAX NEURAL NETWORKS

5.1 Introduction 106

5.2 Agents in Multi-Agent Systems 107

5.3 Trust in Multi-Agent Systems 108

5.4 The Proposed MACS-CBS Model 111

5.5 Results and Discussion 117

5.5.1 Case Study I 118

5.5.1.1 Circle-in-the-Square 118

5.5.1.2 Four-circle-in-the-square 121

5.5.2 Case Study II 124

5.5.2.1 Simulations with Varying Hyperbox Sizes 125

5.5.2.2 Comparison with Other Published Results 127

5.5.3 Case Study III 133

5.6 Summary 135

CHAPTER 6 - REAL-WORLD MEDICAL AND INDUSTRIAL APPLICATIONS

6.1 Introduction 137

6.2 Real-World Case Studies 137

6.2.1 Case Study I 138

6.2.1.1 FMM and EFMM2 139

6.2.1.2 MACS-CBS, EFMM, and EFMM2 146

6.2.2 Case Study II 147

6.2.2.1 FMM and EFMM2 149

6.2.2.2 MACS-CBS, EFMM, and EFMM2 152

6.3 Summary 154

CHAPTER 7 - CONCLUSION AND FUTURE WORK

7.1 Summary of the Research 155

7.2 Contributions of the Research 158

7.3 Suggestions for Further Work 160

REFERENCES 162

LIST OF PUBLICATIONS 178

LIST OF TABLES

Table 2.1: A summary of the reviewed FMM-based networks 23

Table 2.2: Limitations of the FMM-based networks 24

Table 2.3: A summary of the reviewed ensemble methods 30

Table 2.4: A summary of the reviewed MAS models.....35

Table 3.1: Information of the benchmark data sets.....60

Table 3.2: Comparison between FMM and EFMM in terms of the average number of hyperboxes 65

Table 3.3: Comparison between EFMM and other FMM-related models using the Iris data set. The results (percentages of misclassification rates) of DCFMN, FMCN, FMM, and GFMN are extracted from (Zhang et al., 2011)…. 66

Table 3.4: Performance comparison in percentage between the EFMM and different SVM classifiers as published in Wang et al. (2009) 69

Table 4.1 : Performance for the circle-in-the-square data set using the bootstrap method with Θ =0.05 87

Table 4.2 : Accuracy comparison for different methods with different training sizes for the circle-in-the-square data set using the bootstrap method, with the significance level (α) at 0.05.....87

Table 4.3 : Complexity (hyperbox number) comparison using different training sizes for the circle-in-the-square data set, with Θ =0.05 88

Table 4.4 : Performance for the circle-in-the-square data set (10000 samples) with different noise levels, the significance level (α) at 0.05, and Θ =0.05…. 92

Table 4.5 : Complexity comparison for the circle-in-the-square data set (10000 samples) with different noise levels and Θ =0.05.....93

Table 4.6 : Benchmark Data Sets 93

Table 4.7 : Performance comparison in percentage between EFMM, EFMM2, and different models published in Saez et. al. (2010) for the Iris data set with Θ =0.05 99

Table 4.8 : Performance for the Iris data set, with different noise levels, the significance level (α) at 0.05, and Θ =0.05 99

Table 4.9 : Complexity (hyperbox number) comparison for the Iris data set with different noise levels and Θ =0.05 100

Table 4.10 : Performance comparison in percentage between EFMM, EFMM2, and different models published in Saez et. al. (2010) for the Heart (Statlog) data set with Θ =0.05 101

Table 4.11 : Performance for the Heart (Statlog) data set, with different noise levels, the significance level (α) at 0.05, and Θ =0.05 101

Table 4.12 : Complexity (hyperbox number) comparison for the Heart (Statlog) data set, with different noise levels and Θ =0.05 102

Table 4.13 : Performance comparison in percentage between EFMM, EFMM2, and different models published in Saez et. al. (2010) for the Twonorm data set with Θ =0.6 103

Table 4.14 : Performance for the Twonorm data set, with different noise levels, the significance level (α) at 0.05, and Θ =0.6 103

Table 4.15 : Complexity (hyperbox number) comparison for the Twonorm data set with different noise levels and Θ =0.6 104

Table 5.1 : The MACS-CBS performance comparison in percentage using different C_{bid} values and different Θ size for Heart data set. 115

Table 5.2 : Performance comparison in percentage for the circle-in-the-square data set (10000 samples), with the significance level (α) at 0.05 and Θ =0.05… 120

Table 5.3 : The accuracy comparison in percentage between the MACS-CBS and the weak agents using four-circle-in-the-square data set 123

Table 5.4 : Test accuracy (%) for the four-circle-in-the-square data set, with Θ =0.05…. 123

Table 5.5 : Performance comparison in percentage between EFMM, EFMM2, and different models published in Saez et al. (2010) for the Iris data set with Θ =0.05.....128

Table 5.6 : Performance for the Iris data set, with different noise levels, the significance level (α) at 0.05, and Θ =0.05.....129

Table 5.7 : Performance comparison in percentage between EFMM, EFMM2, and different models published in Saez et al. (2010) for the Heart (Statlog) data set with Θ =0.05 130

Table 5.8 : Performance for the Heart (Statlog) data set, with adding different noise levels, the significance level (α) at 0.05, and Θ =0.05 131

Table 5.9 : Performance comparison in percentage between EFMM, EFMM2, and different models published in Saez et al. (2010) for the Twonorm data set with Θ =0.6 132

Table 5.10 : Performance for the Twonorm data set, with different noise levels, the significance level (α) at 0.05, and Θ =0.6.....133

Table 5.11 : The 95% confidence level of the bootstrapped averages for MACS-CBS using different datasets, with the significance level (α) at 0.05 and Θ =0.05.....134

Table 5.12 : The performance comparison in percentage between MACS-CBS and other ensemble methods in Shahjahan and Murase (2006) 135

Table 6.1 : Sixteen symptoms and features extracted from each patient records for the ACS data set 139

Table 6.2 : The overall results (accuracy in percentage) and average number of hyperboxes of FMM and EFMM2 for the ACS data set 140

Table 6.3 : The Harmonic data set (Yap et al., 2011) 149

Table 6.4 : Comparison of test accuracy (%) among MACS-CBS and other models in Yap et al. (2011) for the Harmonic data set, where Θ =0.05…..153

LIST OF FIGURES

Figure 1.1: Research relationships 11

Figure 1.2: Research methodology 13

Figure 2.1: The FMM learning process 18

Figure 3.1: The FMM network structure 39

Figure 3.2: A three-dimensional hyperbox 40

Figure 3.3: An example of FMM hyperboxes placed along the boundary of a two class problem 44

Figure 3.4: An example illustrating the FMM learning algorithm for a two-class problem 47

Figure 3.5: The FMM expansion process 51

Figure 3.6: Unrecognized overlapping cases 52

Figure 3.7: The modified hyperbox expansion rule 56

Figure 3.8 : The average (bootstrap) test accuracy of EFMM and FMM for different data sets. The error bars indicate the bootstrap 95% confidence intervals modified hyperbox expansion rule 64

Figure 4.1: The EFMM expansion problem 78

Figure 4.2: The EFMM2 expansion process 83

Figure 4.3: Complexity comparison with different training sizes for the circle-in-the-square data set using four models with Θ=0.05 (P.EFMM and P.EFMM2 indicate EFMM and EFMM2 with pruning) 88

Figure 4.4: The EFMM decision boundaries created by different noise-free training samples 90

Figure 4.5: The EFMM2 decision boundaries created by different noise-free training samples 91

Figure 4.6 : Performance comparison in percentage between EFMM and EFMM2 for different data sets 97

Figure 5.1: Overview of the MACS-CBS model 111

Figure 5.2: Overview of the MACS-CBS model 113

Figure 5.3: Performance for the circle-in-the-square data set using the bootstrap method when Θ =0.05 119

Figure 5.4: The circle-in-the-square problem with 10000 samples 120

Figure 5.5: The four-circle-in-the-square problem with 2000 samples 124

Figure 5.6 : Performance comparisons in percentage between weak agents and MACS-CBS for different data sets 127

Figure 6.1 : Comparison between FMM and EFMM2 hyperbox overlapping cases for Θ=0.4 (black and red symbols represent class 1 and class 2 hyperboxes, respectively) 143

Figure 6.2 : Comparison between FMM and EFMM2 hyperbox overlapping cases for Θ=0.5 (black and red symbols represent class 1 and class 2 hyperboxes, respectively) 144

Figure 6.3 : Comparison between FMM and EFMM2 hyperbox overlapping cases for Θ=0.6 (black and red symbols represent class 1 and class 2 hyperboxes, respectively) 145

Figure 6.4 : Performance comparison in percentage among MACS-CBS, EFMM, and EFMM2 for the ACS data set 146

Figure 6.5 : Performance comparison in percentage between FMM and EFMM2 for the Harmonic current data set 150

Figure 6.6 : Comparison of the mean test accuracy rates among FMM, EFMM2, and MACS-CBS for the Harmonic data set.....153

LIST OF ABBREVIATIONS

ACS	Acute Coronary Syndrome
AMGS	Adaptive Merging and Growing Strategy
ANN	Artificial Neural Network
ANR	Automatic Noise Reduction
ARC	Adaptive Resolution Classifier
ART	Adaptive Resonance Theory
BBA	Bucket Brigade Algorithm
CBS	Certified Belief in Strength
CF	Confidence Factor
CI	Computational Intelligence
CNs	Classifying Neurons
CPTS	Correct Predicted Test Samples
CS	Clustering-and-Selection
DCFMN	Data-Core-Based Fuzzy Min–Max Neural Network
DDA	Dynamic Decay Adjustment
DEIR	Distributed Embedded Intelligence Room
DT	Decision Templates
DWM	Dynamic Weighted Majority
EFMM	Enhanced FMM
FAM	Fuzzy ARTMAP
FMCN	Fuzzy Neural Network with Compensatory Neuron
FMM	Fuzzy min-max
GA	Genetic Algorithm

GFMM	General Fuzzy Min-Max
GRFMN	General Reflex Fuzzy Min-Max Neural Network
HA	Hyperbox Accuracy
ICPTS	Incorrect Predicted Test Samples
JADE	Java Agent DEvelopment
KEEL	Knowledge Extraction based on Evolutionary Learning
kNN	k-Nearest Neighbor
LP	Linear Programming
MACS	Multi-Agent Classifier System
MANNFIS	Multi-Agent Neural Net Fuzzy Inference System
MAS	Multi Agent System
MFMM	Modified FMM
MFSs	Multiple Feature Subsets
MLP	Multi-Layer Perceptron
OLNs	Overlapping Neurons
PARC	Pruned Adaptive Resolution Classifier
PWBTS	Performance Weight Based on Tough Samples
RVM	Relevance Vector Machines
SVM	Support Vector Machines
TACF	Threshold-Adjusted Classification Filter
TNC	Trust-Negotiation-Communication
UCI	University of California, Irvine
WAVE	Weight-Adjusted Voting for Ensembles of classifiers
WFMM	Weighted FMM network

MODEL PENGECAMAN CORAK INDIVIDU DAN GABUNGAN DENGAN MENGGUNAKAN RANGKAIAN NEURAL MIN-MAX KABUR YANG DIPERTINGKATKAN

ABSTRAK

Klasifikasi corak adalah salah satu daripada komponen utama untuk rekabentuk dan pembangunan sistem pengecaman corak berkomputer. Tertumpu kepada model kecerdasan berkomputer, tesis ini menerangkan secara mendalam kajian-kajian menerusi dua pendekatan yang berkemungkinan bagi tujuan merekabentuk model klasifikasi corak yang kukuh dan anjal serta berprestasi tinggi. Pertamanya, dengan meningkatkan prestasi pembelajaran rangkaian neural-kabur dan keduanya dengan merangka model gabungan bagi menggabungkan ramalan daripada pelbagai rangkaian neural-kabur menggunakan rangka kerja berasaskan ejen. Disebabkan terdapat beberapa ciri penting termasuk keupayaan pembelajaran secara berperingkat dan wujudnya sempadan keputusan tak linear dengan '*hyperboxes'*, rangkaian Min-Max Kabur (FMM) dipilih sebagai asas bagi merekabentuk model klasifikasi corak yang boleh digunakan dalam kajian ini. Dua varian FMM yang dipertingkatkan, iaitu EFMM dan EFMM2, telah dicadangkan bagi menangani beberapa kelemahan yang terdapat dalam algoritma pembelajaran FMM asal. Untuk EFMM, tiga kaedah heuristik diperkenalkan bagi meningkatkan perkembangan '*hyperbox'*, ujian pertindihan, dan proses-proses penguncupan. Kerumitan rangkaian dan isu-isu toleransi hingar diambil kira dalam EFMM2. Di samping itu, rangka kerja berasaskan ejen digunapakai sebagai model gabungan yang kukuh bagi menempatkan rangkaian berganda berasaskan EFMM. Satu kaedah pengukuran berguna yang dikenali sebagai

Certified Belief in Strength (CBS) telah dibangunkan dan dimasukkan ke dalam model gabungan bagi mengeksploitasi prestasi ramalan rangkaian berasaskan EFMM yang berbeza. Model yang terhasil ini dipanggil 'Multi-Agent Classifier System with Certified Belief in Strength (MACS-CBS)'. Prestasi kedua-dua model berasaskan kedua-dua EFMM iaitu yang tunggal dan gabungan dikaji secara sistematik menggunakan satu siri kajian penanda aras, yang mana keputusan tersebut di analisis dan dibincangkan. Dari sudut penilaian empirikal, model EFMM dan model EFMM2 yang dibangunkan menunjukkan peningkatan prestasi jika dibandingkan dengan rangkaian–rangkaian FMM asal. Keputusan kajian juga menunjukkan bahawa prestasi kedua-dua model adalah setanding atau lebih baik daripada kebanyakan sistem pembelajaran mesin yang telah dilaporkan dalam kajian ilmiah. Tambahan pula, dua aplikasi sebenar yang melibatkan permasalahan dalam bidang perubatan dan industri telah digunakan bagi tujuan penilaian. Keputusan positif yang diperolehi menunjukkan potensi dan keberkesanan model gabungan berasaskan EFMM sekiranya dibandingkan dengan FMM dan model berkaitan yang lain berdasarkan kajian ilmiah dalam menyelesaikan masalah klasifikasi corak dalam persekitaran sebenar.

INDIVIDUAL AND ENSEMBLE PATTERN CLASSIFICATION MODELS USING ENHANCED FUZZY MIN-MAX NEURAL NETWORKS

ABSTRACT

Pattern classification is one of the major components for the design and development of a computerized pattern recognition system. Focused on computational intelligence models, this thesis describes in-depth investigations on two possible directions to design robust and flexible pattern classification models with high performance. Firstly is by enhancing the learning algorithm of a neural-fuzzy network; and secondly by devising an ensemble model to combine the predictions from multiple neural-fuzzy networks using an agent-based framework. Owing to a number of salient features which include the ability of learning incrementally and establishing nonlinear decision boundary with hyperboxes, the Fuzzy Min-Max (FMM) network is selected as the backbone for designing useful and usable pattern classification models in this research. Two enhanced FMM variants, i.e. EFMM and EFMM2, are proposed to address a number of limitations in the original FMM learning algorithm. In EFMM, three heuristic rules are introduced to improve the hyperbox expansion, overlap test, and contraction processes. The network complexity and noise tolerance issues are undertaken in EFMM2. In addition, an agent-based framework is capitalized as a robust ensemble model to house multiple EFMM-based networks. A useful trust measurement method known as Certified Belief in Strength (CBS) is developed and incorporated into the ensemble model for exploiting the predictive performances of different EFMM-based networks. The resulting model is known as

a Multi-Agent Classifier System with Certified Belief in Strength (MACS-CBS). The usefulness of both individual and ensemble EFMM-based models is evaluated systematically using a series of benchmark studies, with the results analyzed and discussed. From the empirical evaluation, the proposed EFMM and EFMM2 models show improved performances as compared with those from the original FMM network. The results are also either comparable with or better than those from many other machine learning systems reported in the literature. Furthermore, two real-world medical and industrial problems are used for evaluation. The outcomes positively demonstrate the potential and efficacy of the proposed ensemble EFMM-based model, as compared with FMM and other related models in the literature, in undertaking pattern classification problems in the real environments.

CHAPTER 1

INTRODUCTION

1.1 Background

Since ancient times, humans are considered as the best pattern recognizer in most instances (Jain et al., 2000). Our brain receives patterns from sensing organs and processes them to become useful information, and subsequently allows us to make appropriate decisions based on the received patterns. The development in sciences and technologies has led humans to research into models and techniques to emulate the functionality of the human brain. One of the attempts is to understand the process and action of pattern recognition by the human brain, and subsequently develop computerized systems to imitate this pattern recognition capability. However, for a computerized system to function as a useful pattern recognizer, it needs to be equipped with robust algorithms in order to be able to extract meaningful features from events or objects, and classify them into different categories. Nowadays, pattern recognition has attracted the attention of many researchers from different fields, and is becoming one of the most important characteristics of intelligent behaviours. In general, the design and development of a computerized pattern recognition system comprises four major components (Rosenfeld and Wechsler, 2000), i.e.:

- *Data Acquisition and Collection*: it is the procedure for finding patterns from physical conditions (events or objects) and changing them from analog to digital values that a computer can process. In other words, a data acquisition

element of a computerized pattern recognition system is considered as the equivalent to the sensing organs in humans.

- *Feature Extraction and Representation*: It is the procedure for transforming raw data (digit values) into a set of specific attributes or features. The input features are then processed by using some form of mathematical function to provide informative and representative measurements for the raw data.
- *Similarity Detection and Pattern Classification*: It is the procedure for categorizing and assigning the input features into one of the target clusters (for unsupervised learning) or target classes (for supervised learning) by applying some form of decision rule.
- *Performance Evaluation*: It is the procedure for applying mathematical measurements to estimate the effectiveness of the computerized pattern recognition system quantitatively, normally based on a set of new and unseen data samples.

The focus of this thesis is on the pattern classification aspect of the overall design and development of a computerized pattern recognition system.

Many methods have been developed for pattern classification. One of the earliest methods for pattern classification was statistical approaches, which were started by designing classical linear discrimination methods proposed in Fisher (1936) and Rao (1948). Then, the Bayesian decision method became one of the most popular statistical approaches for pattern classification (Devijver and Kittler, 1982; Duda and Hart, 1973). However, statistical approaches have difficulties in handling contextual/structural information in patterns, as indicated in Pal and Pal (2002). This

problem was then tackled by using syntactic approaches related to the theory of formal language for pattern classification (Hopcroft and Ullman, 1979). While syntactic approaches work fine for idealized patterns, they are inefficient in handling noisy and distorted patterns (Pal and Pal, 2002). On the other hand, the use of classification trees constitutes another useful method for pattern classification (Breiman et al., 1984). Nevertheless, the method faces the same inefficiency problem as syntactic approaches in dealing with noisy, distorted patterns (Pal and Pal, 2002).

Recently, Computational Intelligence (CI) (Bezdek, 1994) models have emerged as one of the useful methods for pattern classification. CI is a new field that capitalizes on interdisciplinary theories and principles for designing and developing computerized intelligent systems (Jain et al., 2008). In the following sections, a definition of CI is provided. The motivations for developing an ensemble of CI-based systems are given. Then, the research objectives and scope are explained, which is followed by the research methodology. Finally, an overview of the organization of this thesis is presented.

1.2 Computational Intelligence

Imitating human behaviours is the main driving force that inspires researchers to develop CI-based systems. Bezdek (1994) introduced one of the earliest definitions for CI, as follows:

> *"…A system is computationally intelligent when it: deals with only numerical (low-level) data, has pattern recognition components, does not use knowledge in the AI sense; and additionally when it*

(begins to) exhibit i) computational adaptivity, ii) computational fault tolerance, iii) speed approaching human-like tumaround and iv) error rates that approximate human performance."

CI is an emerging field which comprise of a highly interdisciplinary framework that can solve various problems with the use of computers to perform numerical calculations (Rutkowski, 2008). A number of paradigms exist under the umbrella of CI models (Rutkowski, 2008), which include neural networks (Oong and Isa, 2011), fuzzy logic (Chen et al., 2012), evolutionary algorithms (Michalewicz, 1996), rough sets (Pawlak, 1992), and probabilistic methods (Pagan and Ullah, 1999). There are a lot of successful applications of CI-based systems in different areas, which include industrial (West and West, 2000), web intelligence (Cercone et al., 2002), management (Meesad and Yen, 2003), finance and economics (Isasi et al., 2007), medical decision (Papageorgiou, 2009), future power systems (Vale et al., 2011), as well as education (Kim and Cho, 2013).

The focus of this thesis is on Artificial Neural Network (ANN) and other complementary paradigms for developing useful pattern classification systems. An ANN is a mathematical representation model inspired from the biological neural network in the human brain. One of the earlier attempts to understand the organizing principles of the brain was implemented in McCulloch and Pitts (1943). The aim was to imitate the biological neural structure by formulating a mathematical model of biological neurons. The attempt then brought about the development of ANNs. To date, there are a number of different ANN architectures, which include the Multi-Layer Perceptron (MLP) networks (Rumelhart and Zipser, 1986), Hopfield network

(Hopfield, 1982; Hopfield, 1984), and Radial Basis Function (RBF) network (Lowe and Broomhead, 1988).

1.3 Problems and Motivations

ANNs have emerged as one of the popular methods in tackling pattern classification problems. ANNs are useful for handling noisy data collected from real environments. The learning property of ANNs gives them the ability to recognize different types of input patterns. Most of the ANN learning strategies are related to batch or offline learning (Puttige and Anavatti, 2007). However, one of the problems of batch learning in ANN models, such as MLP and RBF, is catastrophic forgetting (Robins, 1993; Ratcliff, 1990; McCloskey and Cohen, 1989). The phenomenon of catastrophic forgetting is concerned with the inability of an ANN to remember what it has previously learned when new information is learned by the ANN (Polikar et al., 2001; Polikar et al., 2000).

There have been many attempts from researchers to solve this catastrophic forgetting problem. One of these attempts was by McCloskey and Cohen (1989), where they used the back-propagation ANN to understand the catastrophic forgetting problem. They found that the network created a new solution based only on the most recent information, when it was given multiple materials to learn (McCloskey and Cohen, 1989). This is obviously different from the functionality of the human brain.

On the other hand, the stability plasticity dilemma in learning systems (Simpson, 1992; Goldberg, 1989; Grossberg, 1980) is also related to the catastrophic

forgetting problem. The dilemma aims to address a number of issues, e.g. how a learning system can remain plastic enough to learn new information, and how to retain previously learned information when new information is provided (Carpenter and Grossberg, 1987; Carpenter and Grossberg, 1988). Solving the stability plasticity dilemma is a crucial issue in ANN learning, especially when the number of data samples increases with time and the ANN has to learn these samples in an autonomous and incremental manner. As a result, Simpson (1992, 1993) proposed two hybrids ANN models, i.e., the Fuzzy min-max (FMM) networks, in an attempt to combat this stability-plasticity dilemma. The first ANN is for pattern classification (Simpson, 1992); while the second is for pattern clustering (Simpson, 1993). In this thesis, the focus is on the supervised FMM (hereafter known as just FMM) network for pattern classification, owing to a number of reasons, as follows.

The FMM network is a supervised CI model which integrates both ANN and fuzzy set theory together in a unified framework. It uses hyperbox fuzzy sets to create and store knowledge (as hidden nodes) in its network structure. Each hyperbox is defined by its minimum (min) and maximum (max) points in an *n*-dimensional pattern space. The fuzzy part in FMM is created by combining the hyperbox min-max points with the fuzzy membership function. The fuzzy membership function determines the degree by which an input pattern belongs to one particular class or another. There are a number of useful properties for FMM to handle pattern classification problems (Simpson, 1992):

a) *Online Learning*: the ability to learn new classes and refine existing classes quickly and without losing old information. This property is important to tackle the stability plasticity dilemma.

b) *Nonlinear Separability*: the ability to build decision regions that separate classes of any shape and size.

c) *Overlapping Classes*: the ability to form a decision boundary to minimize the misclassification rate by eliminating the overlapping regions from different classes.

d) *Training Time*: the ability to learn and revise the decision boundaries of different classes within a short training time and using only one pass learning.

e) So*ft and Hard Decisions*: the ability to provide both soft and hard classification decisions. The hard decision indicates whether a pattern is in a specific class or otherwise (either 0 or 1), while the soft decision describes the degree to which an input pattern fits within a particular class.

All the above salient properties make FMM a unique pattern classifier. However, there are rooms to enhance the FMM learning algorithm. In particular, its expansion, overlapping test, and contraction processes need further improvements. Another shortcoming of FMM is its network complexity and noise tolerance capability. Learning with large data sets increases the FMM network complexity, while learning with noisy data samples results in spurious knowledge stored in its network structure. All these limitations affect the performance of FMM. Therefore, this thesis addresses techniques and strategies to solve these limitations and improve the robustness of FMM for tackling pattern classification problems.

In addition to a robust and efficient learning algorithm, improved performance in CI-based systems for pattern classification can also be achieved by using multiple classifier decisions, instead of using single classifier. In this case, each classifier has to make independent classification errors. By using multiple classifiers with a suitable decision combination scheme, the chance of avoiding classification errors and getting the correct decision for each input pattern can be increased; hence improving the overall classification performance. One of the ways is to use ensemble methods to combine the decisions from a set of classifiers (Dietterich, 2000). In this aspect, the Multi Agent System (MAS) is a viable ensemble method, which has been widely used in different fields, e.g. decision support (Fazlollahi and Vahidov, 2000), industrial steel processing (Gao et al., 2003), robot navigation (Ambastha et al., 2005), power systems (Baxevanos and Labridis, 2007), medical service (Lopez et al., 2008), mobile agent technology (Chen et al., 2009), urban traffic signal control (Balaji and Srinivasan, 2010), military network (Hancock and Lamont, 2011), and energy management (Mets et al., 2012). While the MAS framework can be used for building an ensemble of pattern classifiers, designing an effective MAS model is not as simple as it seems. One of the potential problems is the trust measurement of MAS agents (Yu et al., 2013), i.e. how can one agent trust another agent, and how to measure the trustworthiness of an agent. Therefore, this thesis investigates how to solve these problems in an MAS model, in order to improve the robustness of MAS functioning as a useful ensemble method for tackling pattern classification problems.

1.4 Research Aim and Objectives

The main aim of this research is to investigate the efficacy of the FMM network as a useful and usable pattern classification system. In addition to algorithmic investigations of FMM, the MAS framework is capitalized as a robust ensemble method to devise a multiple classifier model. The rationale is to formulate a useful trust measurement for exploiting the strength of different FMM-based classifiers using the MAS framework. The research objectives are as follows:

1. to propose an Enhanced FMM (EFMM) model by modifying the learning algorithm of FMM through tackling issues related to the expansion, overlapping test, and contraction processes among different classes;
2. to reduce the EFMM network complexity and enhance its noise tolerance ability in handling large and/or noisy data sets;
3. to devise a novel trust measurement scheme that is able to differentiate between good and poor predictions from different classifiers based on performance indicators;
4. to evaluate the usefulness of individual and ensemble EFMM-based models in undertaking pattern classification problems using both benchmark and real-world data sets, quantify their performances using statistical indicators, as well as analyze and compare their effectiveness with different classifiers.

A step-by-step approach is taken in this research to achieve the above objectives. The research methodology is explained in the next section.

1.5 Overview of Methodology

Figure 1.1 depicts an overview of this research. The scope of this research is focused on pattern classification, which is an important part of the overall pattern recognition system as explained in section 1.1. As such, the FMM neural network as single classifier and the MAS framework that supports multiple decisions are discussed, studied, and analysed in this research. The FMM network is selected as the backbone for developing the pattern classifier used in this research, owing to a number of salient features, especially the ability to combat the catastrophic forgetting problem or the stability-plasticity dilemma as explained in section 1.2. In order to enhance the classification performance, two directions are focused: (i) enhancing the learning algorithm of FMM; and (ii) devising a robust MAS-based ensemble framework to combine the predictions from multiple EFMM-based networks.

In this research, in-depth investigations are conducted in two stages, i.e., the learning stage of individual EFMM networks, and the decision combination stage of multiple EFMM networks using the MAS framework. In the learning stage, the challenge is to overcome the expansion, overlapping test, and contraction processes among different classes in FMM and EFMM. In addition, the network complexity and noise tolerance issues resulting from large, noisy data sets are taken into consideration. In the decision combination stage, the challenge is to formulate a useful trust measurement scheme for an ensemble of EFMM-based networks in an MAS framework. This problem is viewed from the angle of "how to measure trust for the predictions from different agents with different (good and poor) performances, and how to arrive at the best final decision". As a result, a number of enhancements to the learning algorithm of FMM are proposed. The EFMM networks

are evaluated thoroughly and systematically, then, the MAS framework is utilized, and a novel trust measurement scheme is proposed. All these investigations aim to improve the classification performance of EFMM-based models, making them a robust and useful pattern classifier for tackling real-world problems. Numerous simulations using benchmark and real data sets are conducted along the course of this research, with the results analyzed, discussed, and compared with those from other related classifiers reported in the literature.

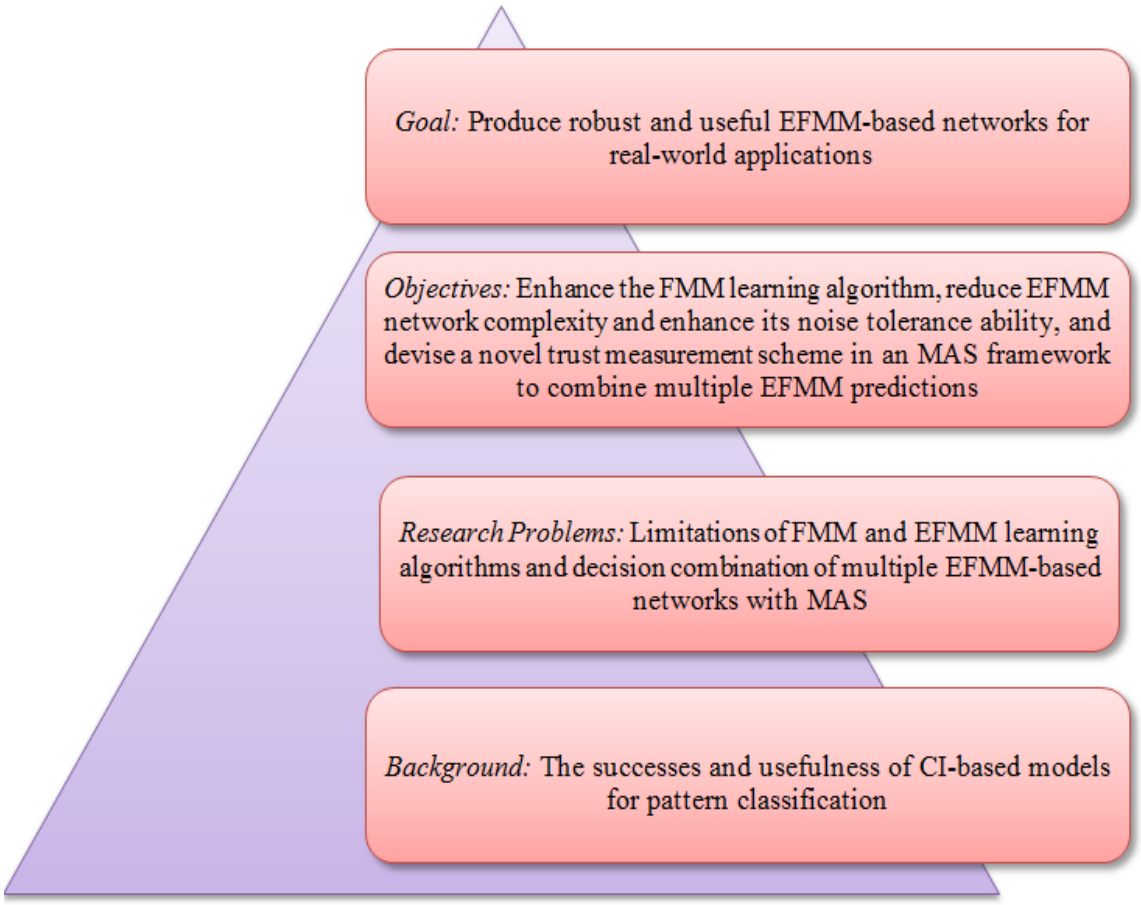

Figure 1.1: Research relationships

A summary of the research methodology is shown in Figure 1.2. The following activities are conducted in order to achieve the overall research objectives.

- *Step 1:* Examining the FMM and EFMM learning algorithms. Three aspects of improvements are identified in order to overcome some of the existing limitations, i.e. the expansion, overlapping test, and contraction processes.
- *Step 2:* Benchmarking the EFMM network with publicly available data sets. The results are analyzed and compared with the original FMM network and those from other methods reported in the literature. This step is necessary to evaluate the effectiveness of the enhanced processes in EFMM.
- *Step 3:* Reducing the network complexity and increasing the noise tolerance capability of the EFMM network. Useful techniques to select the winning hyperbox and to prune hyperboxes with low confidence factors in EFMM are proposed.
- *Step 4:* Benchmarking the second enhanced FMM network (EFMM2) with publicly available noisy and noise-free data sets. The results are analyzed and compared with the previous FMM and EFMM networks, and with those from other methods reported in the literature. This step is necessary to evaluate the effectiveness of the EFMMs performance in terms of network complexity and noise tolerance capability, in addition to classification accuracy.
- *Step 5:* Formulating a trust measurement scheme for a Multi-Agent Classifier System (MACS), with the EFMM-based networks (i.e., EFMM and EFMM2) as its constituent agents. This MACS model is designed to tackle issues related to multiple predictions from an ensemble of EFMM-based networks, each with a different performance indicator. This step aims to devise an MACS framework for combining different EFMM-based networks developed in previous steps.

- *Step 6:* Demonstrating the efficacy of the MACS model with publicly available noisy, noise-free data sets, as well as, real-world data sets. The results are analyzed and compared with individual EFMM-based networks and those from other methods reported in the literature. This step is necessary to ascertain the efficacy of the MACS model in real-world environments.

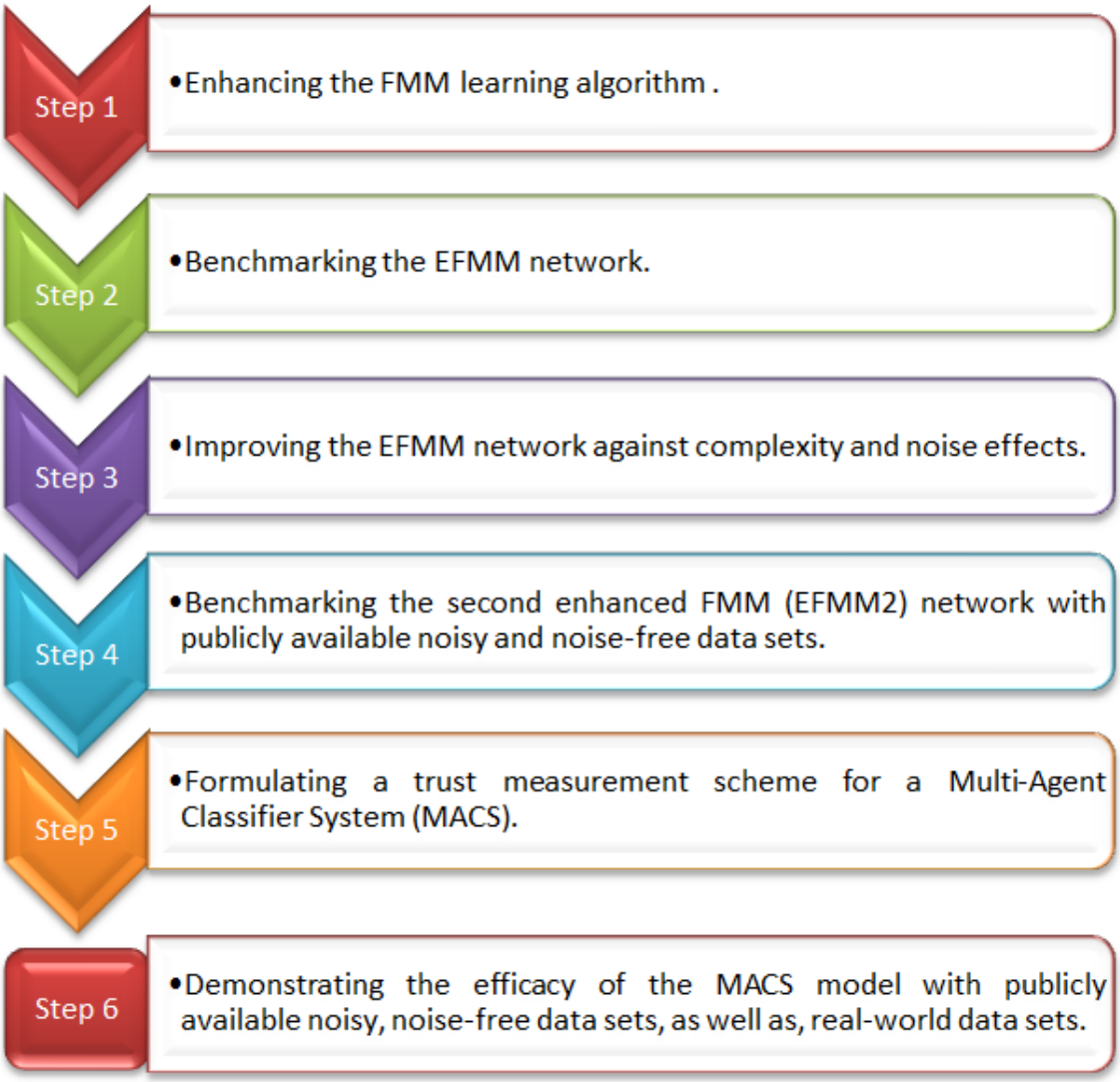

Figure 1.2: Research methodology

1.6 Thesis Outline

This thesis is organized in accordance with the objectives mentioned above. A review of methods and models for improving classification performances based on CI is given in Chapter 2. The review covers the FMM neural network, ensemble methods for pattern classification, as well as multi agent systems as a useful ensemble method.

The learning dynamics of FMM is explained in Chapter 3. A detailed description of the FMM learning algorithm along with a numerical example is presented. Limitations of FMM learning algorithm are highlighted and analyzed. Based on the analysis, novel modifications are proposed to enhance the FMM learning algorithm; hence resulting in the EFMM network. A number of simulations are conducted using benchmark data sets, and the results are compared with those from other methods published in the literature.

Complexity and noise are important issues in pattern classification. As such, a review of existing methods to tackle these issues is presented in Chapter 4. The problems related to network complexity in EFMM are analyzed and discussed. Novel modifications for EFMM are further proposed; hence resulting in EFMM2. Again, a number of simulation studies are conducted using benchmark data sets, and the results are compared with those obtained from other methods.

In Chapter 5, the notion of agent and the MAS model is introduced. A review of trust and its importance in MAS is presented. A novel trust measurement scheme

is proposed, which is known as Certified Belief in Strength (CBS). CBS is incorporated into the MACS model, which comprise EFMM and EFMM2 networks as its agents. The resulting system, MACS-CBS, is evaluated using a number of benchmark data sets, and the results are compared with those from other methods.

To demonstrate the applicability of the MACS-CBS model devised in Chapter 5, and the individual EFMM2 networks proposed in Chapter 4, two real-world case studies are considered in Chapter 6. These data sets are obtained from the medical and power systems domains, as an attempt to ascertain the efficacy of MACS-CBS in real-world environments.

Finally, conclusions are drawn in Chapter 7. Contributions of this research as well as a number of areas to be pursued as further work are presented in Chapter 7 too.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

As explained in Chapter 1, the main focus of this research is to investigate the efficiency of the FMM-based network as a useful and usable pattern classification system. Besides that, the MAS model is capitalized as a robust ensemble framework to devise a multiple classifier system. As such, this chapter presents a review on the FMM network and its variants, ensemble methods, as well as MAS models, whereby all these are utilized in this research. A summary of the review is presented at the end of this chapter.

2.2 Background of the Fuzzy Min-Max Neural Network

An artificial neural network (ANN) is a computational model that consists of an interconnected group of artificial neurons organized in a network structure, which emulates the biological neural system (Li and Ma, 2010; Graupe, 1997). Nowadays, ANNs are widely used in many fields, which include healthcare (Lin et al., 2013; Das and Kundu, 2013), business (Salles et al., 2011), marketing (Abhishek et al., 2012; Azcarraga et al., 2008), financial economics (Li and Ma, 2010), security (Alvarez, 2009; Teoh and Tan, 2010), power (Wei, 2010; Wu and Rastgoufard, 2004), robot programming (Stoica et al., 2010), fault detection (Seera et al., 2012; Seera and Lim, 2013), and airline (Turkmen and Korkmaz, 2010). Among different domains, pattern classification is one of the active areas of ANN applications

(Zhang, 2000). As an example, ANN models have been successfully applied to a variety of real-world classification tasks in industry, business, and science (Zhang, 2000); medical prognosis and diagnosis (Economou et al., 1994), harmonic currents (Yap et al., 2011), as well as industrial fault detection and diagnosis (Quteishat et al., 2009; Seera et al., 2012; Seera and Lim, 2013).

Out of many different types of ANNs, the Fuzzy Min Max (FMM) neural network and its variants have been the focus of many researchers for tackling pattern classification problems (Simpson, 1992; Simpson, 1993; Gabrys and Bargiela, 2000; Nandedkar and Biswas, 2007a; Nandedkar and Biswas, 2007b; Quteishat and Lim, 2008; Zhang et al., 2011; Bargiela et al., 2004; Kim and Yang, 2005). The design of these FMM-based classifiers originates from two FMM models introduced by Simpson (Simpson, 1992; Simpson, 1993). FMM was first presented as a supervised classification neural network (Simpson, 1992), and later as an unsupervised clustering neural network (Simpson, 1993). Both FMM networks combine ANN and fuzzy set theory into a common framework for tackling pattern classification and clustering problems. The FMM structure is built from hyperboxes. A hyperbox is defined by its minimum and maximum points which are encoded from the input patterns. The FMM learning algorithm consists of three steps: expansion, overlapping test, and contraction (Simpson, 1992). The learning procedure in FMM starts by selecting an input pattern and then finding the closest hyperbox that matches the input pattern, as shown in Figure 2.1. The closest hyperbox in the FMM model is found by using a fuzzy membership function, which is defined with respect to hyperbox min–max points (Simpson, 1992).

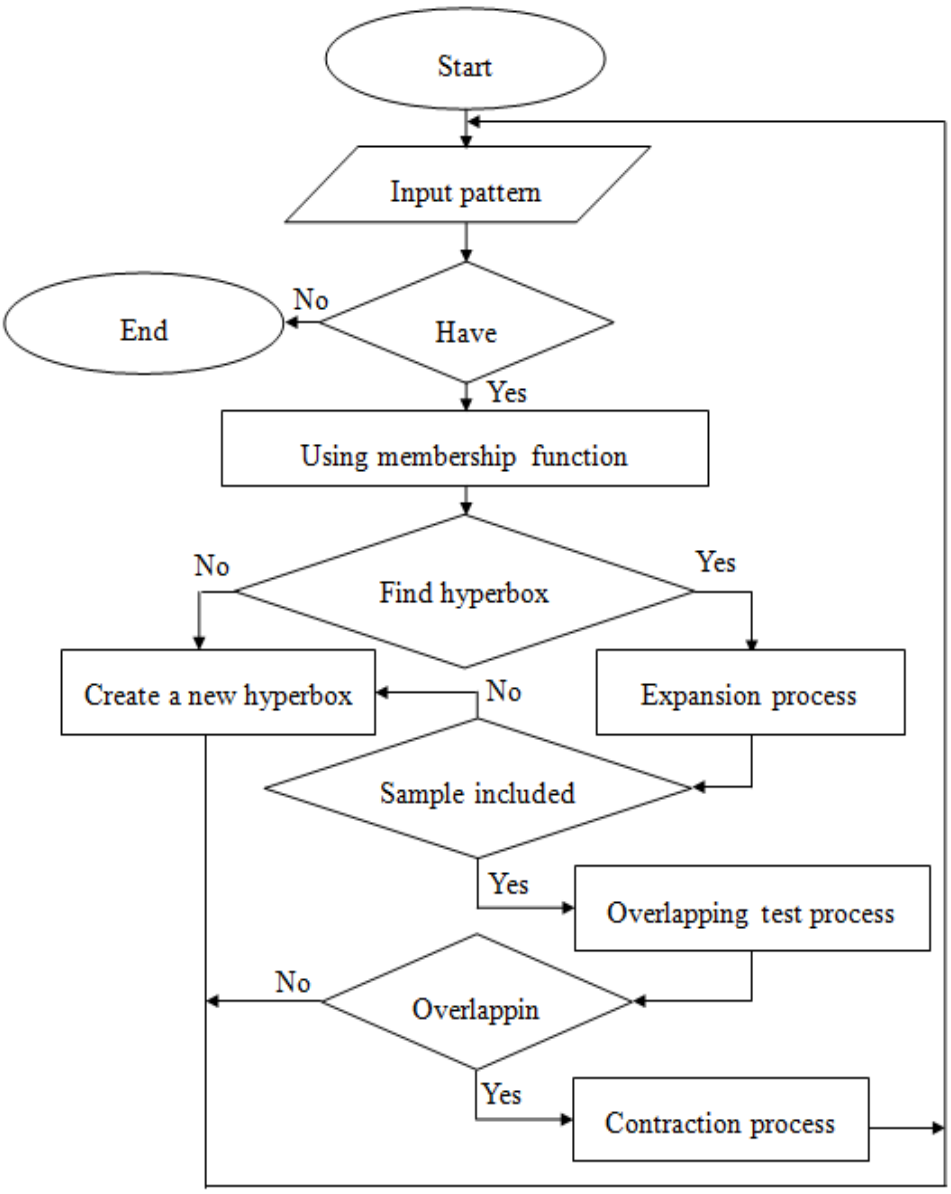

Figure 2.1: The FMM learning process

The membership function represents the degree to which an input pattern fits in the hyperbox, and the membership value ranges between 0 and 1. If the input pattern does not belong to any hyperboxes, even with the expansion process in FMM, a new hyperbox is created to include the input pattern. The overlapping test is carried out to check whether there are any overlaps among hyperboxes from different classes caused by the expansion process, while the contraction process takes place to

eliminate the overlapping area. In other words, FMM entails a dynamic network structure with an online learning capability whereby the number of hyperboxes can be increased when necessary; therefore, avoiding the problem of re-training as faced by many neural network models with an off-line learning capability (Nakashima et al., 2010; Odeh and Khalil, 2011). Further details of FMM can be found in Chapter 3.

A number of FMM variants are available in the literature. An extension of the FMM classification network known as the General Fuzzy Min-Max (GFMM) neural network was proposed by Gabrys and Bargiela (2000). GFMM is established using the expansion and contraction principles, and is able to handle both labelled and unlabelled data simultaneously. The structure of GFMM is similar to that of FMM with three network layer. The advantages of GFMM include the abilities to process input data such as confidence limits, incorporate new information, avoid retraining the network, and combine both supervised and unsupervised learning strategies within a single structure (Gabrys and Bargiela, 2000).

A stochastic FMM network for reinforcement learning was proposed by Likas (2001), which was an extension of the reinforcement FMM model (Likas and Blekas, 1996). It uses the concept of random hyperboxes for reinforcement learning problems. Unlike FMM, the proposed extension (Likas, 2001) uses a stochastic automaton instead of the action label (or class label), where the probability vector of the stochastic automaton determines the corresponding action through random selection. Each hyperbox is associated with a stochastic learning automaton. The location, boundaries of each hyperbox, as well as probability vector of each stochastic automaton are adjusted by the stochastic FMM network.

Inspired by the FMM network, the adaptive resolution Min-Max model was proposed as a new neuro-fuzzy classifier (Rizzi et al., 2002). It employs two algorithms, i.e. the Adaptive Resolution Classifier (ARC) and the Pruned Adaptive Resolution Classifier (PARC). In the proposed model, the hyperbox expansion process is not limited by a fixed maximum size, in order to overcome some undesired properties of the original FMM algorithm. As such, the proposed model possesses a less complex network structure, as compared with original FMM (Rizzi et al., 2002).

In Bargiela et al. (2004), an inclusion/exclusion fuzzy hyperbox classifier was proposed. As its name implies, this model creates two types of hyperboxes, i.e., the inclusion and exclusion hyperboxes. The purpose of the inclusion hyperboxes is to contain the input patterns that belong to the same class. Other overlapped patterns are contained by the exclusion hyperboxes. The use of the exclusion hyperboxes helps reduce the training process from three steps (expansion, overlap test, and contraction) to two (expansion and overlap test). This is achieved through a contentious area of the pattern space to approximate the complex topology of data samples, which helps to solve the overlapping problem in FMM (Bargiela et al., 2004).

A Weighted FMM network (WFMM) was proposed by Kim and Yang (2005). In this model, both the hyperbox contraction and overlap test steps do not restrict hyperbox expansion. A new membership function and a learning method (hyperbox creation, expansion, contraction, and weight update) are defined in WFMM (Kim and Yang, 2005). On the other hand, a new model called FMM with compensatory neuron (i.e., FMCN) was proposed by Nandedkar and Biswas (2007a). Based on the compensatory neuron (CN) architecture, FMCN is used as a supervised

classification model that supports online learning. The CNs are activated when the test sample falls in the overlapped area between two different classes, with the hyperbox contraction process eliminated.

Later, another General Reflex Fuzzy Min-Max Neural Network (GRFMN) was proposed by Nandedkar and Biswas (2007b). GRFMM combines both FMM clustering and classification algorithms as well as the concept of human reflex mechanism to solve the problem of class overlaps (Nandedkar and Biswas, 2007b; Ries et al., 2006; Yuan et al., 2006) into one framework. GRFMN offers on-line training, and exhibits a better ability to identify the underlying data structure as compared with GFMM; hence enhancing classification accuracy (Nandedkar and Biswas, 2007b).

A Modified FMM (MFMM) network that aimed to improve the FMM classification performance was proposed by Quteishat and Lim (2008). MFMM is able to tackle the problem related to a small number of large hyperboxes formed in the network as well as to facilitate rule extraction. In MFMM, the data set is divided into three sub-sets: a training set for learning; a prediction set for pruning and rule extraction; and a test set for performance evaluation. MFMM improves the classification performance with a two-stage process, i.e. (i) network pruning and rule extraction; (ii) prediction by using the Euclidean distance and membership function. Later, MFMM was further enhanced with a Genetic Algorithm (GA)-based rule extractor to form another model called FMM-GA (Quteishat et al., 2010). In the particular study, an application of FMM-GA to medical diagnosis problems has been demonstrated.

A Data-Core-Based Fuzzy Min–Max Neural Network (DCFMN) for pattern classification was recently proposed by Zhang et al. (2011). DCFMN uses new membership functions for two types of neurons, i.e., Classifying Neurons (CNs) and Overlapping Neurons (OLNs). These membership functions take into consideration three factors: noise, the geometric center of the hyperbox, and the data core. The FMM contraction processes are removed in DCFMN. When a new pattern falls in the overlapped area of different classes, DCFMN uses the OLN membership to determine the target class of the new pattern (Zhang et al., 2011).

Table 2.1 shows the characteristics of FMM-based networks discussed in the section. Even though many investigations have been conducted to improve the original FMM neural network, some limitations in the FMM learning algorithm remain unsolved (Hyperbox expansion rule, overlapping test rule, contraction rule, network complexity and noise tolerance capability), which affect its performance. It can be observed that all FMM variants suffer from at least two limitations, as shown in Table 2.2. As such, there are rooms for improving the FMM learning algorithm and performance, and making it a more robust classifier.

In this thesis, two strategies are adopted to enhance FMM. The first is to improve its learning algorithm by solving issues related to overlapping hyperboxes, as detailed in Chapters 3 and 4. Instead of single classifier, the second strategy examines the use of a multiple classifier framework to form an ensemble of classifiers so that the overall performance can be improved, as presented in Chapter 5. The next section reviews some of the ensemble methods that are related to this research.

Table 2.1: A summary of the reviewed FMM-based networks

Model	Characteristics
FMM (Simpson, 1992)	Suitable for supervised learning; combine ANN and fuzzy set theory into a common framework for tackling pattern classification problems; supports on-line adaptation which enables it to avoid the re-training process;
FMM (Simpson, 1993)	Suitable for unsupervised learning; combine ANN and fuzzy set theory into a common framework for tackling pattern clustering problems; supports on-line adaptation.
General Fuzzy Min-Max (GFMM) (Gabrys and Bargiela, 2000)	Handles both labeled data and unlabeled data; suitable for supervised and unsupervised learning.
Stochastic FMM (Likas, 2001)	Uses a stochastic automaton, instead of the action label or class label in original FMM; suitable for reinforcement learning.
Adaptive resolution Min-Max model (Rizzi et al., 2002)	Comprises of the Adaptive Resolution Classifier (ARC) and Pruned Adaptive Resolution Classifier (PARC); the hyperbox expansion process is not limited by a fixed maximum threshold as in the original FMM algorithm; possesses a less complex network structure than that of the original FMM.
Inclusion/Exclusion fuzzy hyperbox classifier (Bargiela et al., 2004)	Uses the inclusion hyperboxes to contain input patterns from the same class; uses the exclusion hyperboxes to contain overlapped patterns.
Weighted FMM network (WFMM) (Kim and Yang, 2005)	Does not restrict hyperbox expansion through the hyperbox contraction and overlap test steps; defines a new membership function; defines a new learning method for hyperbox creation, expansion, contraction, and weight update.
FMM with compensatory neuron (FMCN) (Nandedkar and Biswas, 2007a)	Uses the compensatory neuron (CN) architecture; uses new membership functions for two types of neurons, Classifying Neurons (CNs) and Overlapping Neurons (OLNs).
General Reflex Fuzzy Min-Max Neural Network (GRFMN) (Nandedkar and Biswas, 2007b)	Combines both FMM clustering and classification algorithms; deploys the concept of human reflex mechanism to solve the problem of class overlaps.
Modified FMM (MFMM) (Quteishat and Lim, 2008)	Performs network pruning and rule extraction; uses the Euclidean distance and membership function for prediction.
Modified FMM with Genetic Algorithm (MFMM-GA) (Quteishat et al., 2010)	Performs network pruning and rule extraction; uses the Euclidean distance and membership function for prediction; incorporates a Genetic Algorithm (GA) - based module for rule extraction.
Data-Core-Based Fuzzy Min–Max Neural Network (DCFMN) (Zhang et al., 2011)	Uses membership functions for Classifying Neurons (CNs) and Overlapping Neurons (OLNs); Removing the contraction process.

Table 2.2: Limitations of the FMM-based networks

Limitations / Model Of	Expansion rule	Overlapping test rule	Contraction rule	Complexity	Noise tolerance
Simpson (1992)	√	√	√	√	√
Simpson (1993)	√	√	√	√	√
Gabrys and Bargiela (2000)	√	√	√	√	√
Likas (2001)	√	√	√	X	√
Rizzi et al. (2002)	√	√	√	X	X
Bargiela et al. (2004)	√	√	X	√	√
Kim and Yang (2005)	√	√	√	√	X
Nandedkar and Biswas (2007a)	√	√	X	√	√
Nandedkar and Biswas (2007b)	√	√	√	√	√
Quteishat and Lim (2008)	√	√	√	X	X
Quteishat et al. (2010)	√	√	√	X	X
Zhang et al. (2011)	√	√	X	X	X

√ Have limitation

X Do not have limitation

2.3 Ensemble Methods

As explained in the previous section, FMM is a useful neural network that is able to produce good and accurate results in undertaking pattern classification problems. In FMM (as well as other online ANN models), learning can be unstable whereby any small change in the sequence of the training samples, or parameters of the network could affect the network performance (Navone, 2001). Therefore, it is useful to enhance the learning algorithm and improve its performance. To minimize the classification errors, one useful way is to deploy a group of classifiers for