HYBRID MODELS OF FUZZY ARTMAP AND Q-LEARNING FOR PATTERN CLASSIFICATION

FARHAD POURPANAH NAVAN

UNIVERSITI SAINS MALAYSIA

2015

HYBRID MODELS OF FUZZY ARTMAP AND Q-LEARNING FOR PATTERN CLASSIFICATION

By

FARHAD POURPANAH NAVAN

Thesis submitted in fulfilment of the requirements
for the degree of
Doctor of Philosophy

October 2015

ACKNOWLEDGMENTS

All praises are due to the Creator without Whom I wouldn’t have finished this PhD research work.

I would like to express my gratitude to my first main supervisor and current co-supervisor, Professor Lim Chee Peng, whatever I cannot explain with words, for his great advice, patience, warm concern, valuable suggestions, and endless support that has helped me to complete this dissertation. In addition, I would like to thank him for his technical and editorial advices that have been essential to complete this research and has taught me countless lessons in general. It has been an honour to work with him.

I would like to thank to my second main-supervisor, Assoc. Prof. Dr. Umi Kalthum Ngah, for her advice and encouragement and sharing her experiences. Also, I would like to express the deepest appreciation to my current main-supervisor, Assoc. Prof. Dr. Junita Mohamad Saleh, for her inspiration and kindness. I appreciate the time she has spent on reviewing this thesis for final touch-ups.

I would like to acknowledge the School of Electrical and Electronic Engineering, USM for providing me the necessary equipment as well as the Graduate Assistant (GA) support, which has made this research possible. I also would like to thank all staff at the School for their support.

My final thanks go to my family members who have always stood beside me, in particular, my parents, Mr. Heidar Pourpanah Navan and Ms. Raheleh Pourpanah Navan, for their endless love, support and encouragement throughout my life. In addition, I would like to thank my brothers, Mr. Behrad and Mr. Taha, as well as my sister Miss Sepideh. Without their prayers, support, and encouragement, I would have never been able to complete my work. I owe them more than I could ever repay.

TABLE OF CONTENTS

CHAPTER 4 - IMPROVED FUZZY ARTMAP AND REINFORCEMENT LEARNING MODELS

CHAPTER 5 - DEVELOPMENT OF A MULTI-AGENT CLASSIFIER SYSTEM

CHAPTER 6 - REAL-WORLD APPLICATIONS

CHAPTER 7 - CONCLUSION AND FUTURE WORK

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ADaBoost	Adaptive Boosting
ACC	Accuracy
ACLA	Actor-Critic Learning Automata
ACS	Adaptive Classification System
ART	Adaptive Resonance Theory
ARTMAP	Adaptive Resonance Theory Mapping
ANN	Artificial Neural Network
BAR	Bayesian ARTMAP
BNN	Backpropagation Neural Network
CNeT	Competitive Neural Tress
CAFE	Collaborative Agents for Filtering E-mails
CVNN	Complex-Valued Neural Network
CI	Computational intelligence
CAMAS	Context-Aware Multi-Agent System
CF	Crest factor
CFS	Correlation-based Feature Selection
DC	Dendritic Cell
DARL	Direct Reinforcement Adaptive Learning
DDA	Dynamic Decay Adjustment
DPSO	Dynamic Particle Swarm Optimization
DP	Dynamic Programming
EFMM	Enhanced Fuzzy Min-Max
EA	Evolutionary Algorithm

ESPEC	Specificity
FAM	Fuzzy ARTMAP
FI	Fuzzy Integral
FMM	Fuzzy Min-Max
FS	Fuzzy System
GA	Genetic Algorithm
GARE	Generalized Analytic Rule Extraction
GPI	Generalized Policy Iteration
GPM	Global Power Manager
GRG	Greedy Rule Generation
HIF	High Impedance Faults
HHONC	Hybrid Higher Order Neural Classifier
IF	Impulse factor
IRL	Inverse Reinforcement Learning
KGS	Kanas Geological Survey
KLSPI	Kernel-based Least Square Policy Iteration
KNN	K-Nearest Neighbour
KEEL	Knowledge Extraction based on Evolutionary Learning
LF	Latitude factor
LSPI	Least Square Policy Iteration
LP	Linear Programming
LPM	Local Power Manager
LTM	Long Term Memory
MDP	Markov Decision Process

MACS	Multi-Agent Classifier System
MATC	Multi-Agent Text Classification
MAS	Multi-Agent System
MLP	Multi-Layered Perceptron
MFS	Multiple Feature Subsets
MC	Monte Carlo
NNE	Neural Network Ensemble
OELM	Online Extreme Learning Machine
PQ	Power Quality
PNN	Probabilistic Neural Network
POPTVR	Pseudo-Outer Product Truth Value Restriction
QFAM	Q-learning Fuzzy ARTMAP
RBS	Radial Basis Function
RNCL	Regularized Negative Correlation Learning
RL	Reinforcement Learning
SCM	Separability-Correlation Measure
SENS	Sensitivity
SF	Shape factor
SVM	Support Vector Machine
TD	Temporal Difference
TNC	Trust-Negotiation and Communication
VE	Virtual Enterprise
WOL	Web Ontology Language

LIST OF SYMBOLS

α	Choice parameter
β	Learning parameter
ρ	Vigilance parameter
ξ	Learning rate
$r(j)$	Reinforcement signal
λ	Weighting factor
γ	Discount factor
$vig(j)$	Vigilance test value of node *j*
Q_values_j	Q-value of node *j*
$strength(j)$	Strength of node *j*
$V^{\pi}(s,a)$	The value function of action *a* in state *s* under policy π
A	Complement code of input a
B	Complement code of input b
δ	A small positive value
ρ_{ab}	The map-field vigilance parameter
N_a	The number of nodes in ART_a
N_b	The number of nodes in ART_b
$\wedge$	Fuzzy AND operator
$f(s)$	Fitness function
$NCP(s)$	Number of correctly classified patterns
$\lvert S \rvert$	Number of chosen input features
N_{pop}	Number of population size

$P(S)$	Probability of selecting string S
ψ	Population
Q	Quantization level
R_j	Rule number j
CM^k	Confusion matrix of k-th agent
n_{ij}^k	Total number of input samples of class i predicted as class j by k-th agent
e_k	k-*th* agent
$bel(i)$	Belief function of i-th agent
$initial\_trust(i)$	Initial trust of i-th agent
$trust\ (i)$	Trust of i-th agent
$Q\_Value_{C_m}^k$	Mean of all Q-values that are related to class m of agent k
$Q_{C_m z}$	Q-value of the z-*th* prototype of class m

MODEL HIBRID PEMETAAN TEORI RESONAN ADAPTIF KABUR DAN PEMBELAJARAN-Q UNTUK PENGELASAN CORAK

ABSTRAK

Pengelasan corak adalah salah satu isu utama dalam pelbagai tugas pencarian data. Dalam kajian ini, fokus penyelidikan tertumpu kepada reka bentuk dan pembinaan model hibrid yang menggabungkan rangkaian neural Teori Resonan Adaptif (ART) terselia dan model Pembelajaran Pengukuhan (RL) untuk pengelasan corak. Secara khususnya, rangkaian ARTMAP Kabur (FAM) dan Pembelajaran-Q dijadikan sebagai tulang belakang dalam merekabentuk dan membina model-model hibrid. Satu model QFAM baharu terlebih dahulu diperkenalkan bagi menambahbaik prestasi pengelasan rangkaian FAM. Strategi *pruning* dimasukkan bagi mengurangkan kekompleksan QFAM. Bagi mengatasi isu ketidak-telusan, Algoritma Genetik (GA) digunakan bagi mengekstrak hukum kabur *if-then* daripada QFAM. Model yang terhasil iaitu QFAM-GA, dapat memberi ramalan berserta dengan huraian dengan hanya menggunakan bilangan antisiden yang sedikit. Bagi menambahkan lagi kebolehtahanan model-model Q-FAM, penggunaan sistem agen-pelbagai telah dicadangkan. Hasilnya, model gugusan QFAM berasaskan agen dengan ukuran percaya dan kaedah rundingan baharu telah dicadangkan. Pelbagai jenis masalah tanda-aras telah digunakan bagi penilaian model-model gugusan dan individu berasaskan QFAM. Hasilnya telah dianalisa dan dibandingkan dengan FAM serta model-model yang dilaporkan dalam kajian terdahulu. Sebagai tambahan, dua daripada masalah dunia-nyata digunakan bagi menunjukkan kebolehan praktikal model hibrid. Keputusan akhir menunjukkan keberkesanan modul berasaskan QFAM dalam menerajui tugas-tugas pengelasan corak.

HYBRID MODELS OF FUZZY ARTMAP AND Q-LEARNING FOR PATTERN CLASSIFICATION

ABSTRACT

Pattern classification is one of the primary issues in various data mining tasks. In this study, the main research focus is on the design and development of hybrid models, combining the supervised Adaptive Resonance Theory (ART) neural network and Reinforcement Learning (RL) models for pattern classification. Specifically, the Fuzzy ARTMAP (FAM) network and Q-learning are adopted as the backbone for designing and developing the hybrid models. A new QFAM model is first introduced to improve the classification performance of FAM network. A pruning strategy is incorporated to reduce the complexity of QFAM. To overcome the opaqueness issue, a Genetic Algorithm (GA) is used to extract fuzzy *if-then* rules from QFAM. The resulting model, i.e. QFAM-GA, is able to provide predictions with explanations using only a few antecedents. To further improve the robustness of QFAM-based models, the notion of multi agent systems is employed. As a result, an agent-based QFAM ensemble model with a new trust measurement and negotiation method is proposed. A variety of benchmark problems are used for evaluation of individual and ensemble QFAM-based models. The results are analyzed and compared with those from FAM as well as other models reported in the literature. In addition, two real-world problems are used to demonstrate the practicality of the hybrid models. The outcomes indicate the effectiveness of QFAM-based models in tackling pattern classification tasks.

CHAPTER 1
INTRODUCTION

1.1 Background of the study

It is generally recognized that pattern recognition is a basic function of human cognition (Wang, 2008). Since the last few decades, human’s brain has attracted great attention in both experimental and theoretical aspects. The results have demonstrated that the brain has a tremendous parallel architecture that contains many individual neurons with synapses (interconnections). Human’s brain can easily understand a particular situation, recognize face or speech, and also is able to receive patterns from sensing organs and convert them into helpful information to make decisions (Cenggoro et al., 2014). Indeed, humans encounter plenty of recognition tasks daily and make the corresponding decisions unconsciously. By exploiting the technology of digital computers and developing the necessary machine learning and artificial intelligence algorithms, it is now possible to utilize computers to mimic the performance of human’s brain. As a result, many investigations have been conducted to tackle pattern recognition problems.

To solve pattern recognition problems by using a computerized system, it is essential to have appropriate algorithms that are able to exploit proper features from received information or data to recognize patterns. In general, there are four main stages in developing a pattern recognition system. They are (Rosenfeld & Wechsler, 2000): *(i) Data Acquisition and Collection, (ii) Feature Extraction and Representation, (iii) Similarity Detection and Pattern Classification, and (iv) Performance Evaluation.*

Over the years, many methodologies have been proposed for pattern classification. Statistical methods are one of the earliest methodologies for pattern classification. These include the discriminatory analysis proposed by Fisher (1936) and Rao (1948). Bayesian decision theorem is another statistical method that has been extensively used to tackle pattern classification problems (Devijver & Kittler, 1982; Duda & Hart, 1973). Nevertheless, statistical methods are inefficient in handling contextual or structural information of patterns, as explained by Pal and Pal (2002). Syntactic techniques, which are related to the theory of formal languages, have been suggested to overcome this problem (Hopcroft & Ullman, 1979). Nevertheless, syntactic technique does not perform well in the presence of noisy data (Pal & Pal, 2002).

Computational intelligence (CI) (Bezdek, 1992; Marks, 1993) is another useful methodology that has been widely applied to solving a variety of applications, e.g. biomedical (Shi & Eberhart, 1998; Yang et al., 2007), mobile robotics (Wang, 2002), healthcare (Tejima, 2003), Web (Zhang, 2005), games (Lucas, 2009), business (Haider & Nishat, 2009), power system (Venayagamoorthy, 2009), control (Wilamowski, 2010), and wireless (Iram et al., 2011). CI has also been extensively employed to tackle pattern classification problems. Generally, CI contains evolutionary algorithms (EAs), Fuzzy Systems (FSs), artificial neural networks (ANNs), and synthesis of these three models with each other and/or other conventional methods (Shi & Eberhart, 1998; Rutkowski, 2008). The focus of this research is on ANNs and other complementary learning methodologies, which include reinforcement learning (RL) and multi-agent system (MAS), for designing and developing efficient and effective pattern classification systems.

In the following sections, a definition of and an introduction to CI are presented. Then, an introduction to RL is provided. The motivations for developing hybrid ANNs and combining them into a MAS are described. The research scope and objectives are presented. Finally, an overview of the thesis organization is given at the end of the chapter.

1.2 Computational Intelligence

In this information era, besides the dramatic growth of computer technologies, researchers have designed and developed various intelligent systems that are able to mimic human's behaviours. Analysing the collected data samples and translating them into useful information and subsequently making appropriate decisions is one of the major challenges. To cope with such problems, CI-based models have been devised to operate as useful systems with "humanlike" problem-solving capabilities (Rutkowski, 2008). A definition of CI is provided by Bezdek (1994), is as follows:

> "*A system is computationally intelligent when it deals only with numerical (low-level) data, has pattern recognition component, and does not use knowledge in artificial intelligence (AI) sense.*"

Another definition of CI is described by Fogel (1995), is as follows:

> *"These technologies of neural, fuzzy, and evolutionary systems were brought together under the rubric of computational intelligence, a relatively new trend offered to generally describe*

methods of computation that can be used to adapt solutions to new problems and do not rely on explicit human knowledge."

FSs, EAs, and ANNs are a number of paradigms under the umbrella of CI (Rutkowski, 2008). While CI-based systems have been successfully used to solve problems in different domains, which include medicine (Schizas, 1997), power systems (Pahwa et al., 2003), biological systems (Wu et al., 2007), web design (Liu, Khudkhudia, & Ming, 2008), games (Lucas, 2009), business (Wu, 2010), computer security (Perez et al., 2010), education (Venayagamoorthy, 2010), as well as industrial systems (Sariyildiz et al., 2013), each CI paradigm has its advantages and limitations. As such, hybrid CI models, which consist of two or more CI paradigms, have been introduced to harness the merits of the constituents.

An ANN can be viewed as a mathematical model that processes information based on the principle of a biological neural network (Cenggoro et al., 2014). Since the inception of the first mathematical model of an artificial neuron by McCulloch and Pitts (1947, 1943), many different ANN architectures have been proposed, e.g. Multi-Layered Perceptron (MLP) (Rumelhart et al., 1986), Radial Basis Function (RBF) (Moody & Darken, 1989). A detailed review on ANNs is given in Chapter 2.

1.3 Reinforcement Learning (RL)

RL (Barto & Sutton, 1998) is a methodology that learns from experience by interacting with the environment. It is a semi-supervised learning method that has advantages over supervised learning methods under certain conditions. Unlike supervised learning whereby the target output for each input sample is clearly known, only minimal information that indicates the appropriateness of the response pertaining to an input sample is available in RL. As such, it does not require detailed knowledge of the target output. There are two main advantages of RL. Firstly, it has the capability of learning on-line in a search-control-learn mode based on previous experiences (Lee et al., 1998). Secondly, it is an effective method when there is little knowledge about what and how to perform a task (Gullapalli, 1990).

RL has been extensively applied as an effective feedback mechanism to tackle control and decision making problems. Among various successful RL applications include cart-pole balancing (Barto et al., 1983), Backgammon game (Tesauro, 1994), and elevator dispatching problem (Barto & Crites, 1996). RL has also been used to improve the performance of many classifiers (Likas & Blekas, 1996; Likas, 2001; Quah et al., 2005). However, RL is not free from limitations, e.g. the exponential growth of its state-space owing to the curse of dimensionality (Lin & Lee, 1994). Such problems have attracted many researchers to work on RL.

1.4 Problems and Motivations

Many methods have been proposed to solve pattern classification problems, e.g. k-nearest neighbour (Cover & Hart, 1967), naive Bayes classifier (Domingos & Pazzani, 1997), decision tree (Friedl & Brodley, 1997), Support Vector Machine (SVM) (Vapnik, 1995), Artificial Neural Networks (ANNs) (Venkatesan & Balamurugan, 2001). Among them, ANNs have been used as a useful learning model for solving pattern classification tasks (Zhang, 2000), with the capability of handling non-linear as well as noisy data collected from real-world environments.