

A Cooperative Learning Model for the Fuzzy ARTMAP-Dynamic Decay Adjustment Network with the Genetic Algorithm



Shing Chiang Tan¹, M.V.C. Rao², Chee Peng Lim³

Faculty of Information Science & Technology¹, Faculty of Engineering & Technology² Multimedia University, Malaysia

School of Electrical & Electronic Engineering³
University of Science Malaysia, Malaysia

Abstract

In this paper, combination between a Fuzzy ARTMAP-based artificial neural network (ANN) model and the genetic algorithm (GA) for performing cooperative learning is described. In our previous work, we have proposed a hybrid network integrating the Fuzzy ARTMAP (FAM) network with the Dynamic Decay Adjustment (DDA) algorithm (known as FAMDDA) for tackling pattern classification tasks. In this work, the FAMDDA network is employed as the platform for the GA to perform weight reinforcement. The performance of the proposed system (FAMDDA-GA) is assessed by means of generalization on unseen data from three benchmark problems. The results obtained are analyzed, discussed, and compared with those from FAM-GA. The results reveal that FAMDDA-GA performs better than FAM-GA in terms of test accuracy in the three benchmark problems.

Presentation Overview

- Introduction
- Fuzzy ARTMAP-Based Networks
- Reinforcement Learning of FAMDDA with GA
- Experimental Results and Discussion
- Summary

Introduction

- Combination between artificial neural networks (ANNs) and evolutionary algorithms (EAs) has attracted a lot of attention. The main purpose is to improve accuracy rate through hybridization of ANNs and EAs.
- ANNs and EAs cross-fertilize each other; ANNs provide a framework of accurate and exact computation whereas EAs provide a robust and efficient approach for undertaking complex optimization problems.

Introduction

- Among the existing EAs, the most well-known model perhaps is the Genetic Algorithms (GAs), which is essentially a mechanism of natural selection, genetic, and evolution.
- GAs are general-purpose optimization methods that can find near-optimum solution of a given problem by evaluating many points in the search space simultaneously.

Objective of the Work

- In this paper, GAs are used to assist in machine learning of ANNs and to optimize the weights of ANNs.
- Main objective of our work - to improve the learning process of ANNs by searching for the weights that would eventually lead to a better network generalization performance.

Use of GAs with ANNs

- Multi-Layer Perceptron (MLP) with backpropagation + GAs (Lam and Leung 2004; Tsai et al. 2006). GAs were used to determine the number of hidden nodes and the network weights.
- Limitation of a standard MLP network - difficulty in employing an incremental learning scheme in its structure.
- Fuzzy ARTMAP (FAM) (Carpenter et al. 1992) - a supervised model of Adaptive Resonance Theory (ART) (Carpenter and Grossberg 1987) - surpass the limitation by overcoming the plasticity-stability dilemma.

Use of GAs with ANNs

- Combination between GA and FAM has been reported by Palaniappan and Raveendran (2002) and Palaniappan et al. (2002). However, the work leans towards the use of GA in selecting relevant features for assisting the learning of FAM.
- Relatively little work in the literature that reports fusion between GA and FAM in machine learning and optimization.

- Unlike MLP, FAM performs incrementally learning for which it does not require handcrafting of network architecture prior to the training process. Nevertheless, like MLP, FAM might likely to produce sub-optimal solutions upon the completion of the training process.
- FAM learns incrementally the information by means of prototypes and the resulting templates are employed by the GA as a set of sub-optimal weights that would guide the future search and adaptation. The cooperative learning between FAM and GAs can improve the search efficiency of the hybrid system for obtaining good solutions.

Fuzzy ARTMAP-Based Networks

- FAM (standard supervised ART network (Carpenter et al. 1992))
- FAM with Dynamic Decay Adjustment (FAMDDA) – A FAM-based network that comprises a learning scheme that can resolve conflicts resulting from overlapping among the prototypes of different classes (Tan et al. 2004a).
- FAMDDA-GA (*proposed network*) – weight reinforcement of the FAMDDA network with GA (Baskar et al. 2001).

Reinforcement Learning of FAMDDA with GA

- 1 Weight initialization and nodes self-organization using FAMDDA.
- 2 Generate chromosomes (Pittsburgh Approach).

Repeat

- 3 Compute fitness value of each chromosome.
 - 4 Apply Roulette-Wheel selection.
 - 5 Generate new generation through crossover and mutation.
- Until terminating condition has been satisfied.

Fig. 1. The overall training procedure of FAMDDA-GA.

(For further detail, please refer to section 3 of the paper)

Experiments

- Experiments were conducted with 3 benchmark datasets - Pima Indian diabetes (PID), Australian credit approval (AUS), and heart (HEA), from the UCI machine-learning repository (Blake and Merz 1998).
- Each dataset was divided into the training and test sets according to a ratio 50:50.
- A bootstrap hypothesis test was used to compare the performances of the proposed method (FAMDDA-GA) and FAM-GA; FAMDDA-GA and FAMDDA; and, FAM-GA and FAM.

Results

Table 1 Performance comparison between FAM-GA and FAMDDA-GA, in terms of test accuracy and number of nodes. The results are averages of eight runs, and their respective standard deviation is given in parenthesis. The p -values are the results from the bootstrap hypothesis tests between FAM-GA and FAMDDA-GA.

Dataset	Accuracy (%)			Size (Nodes)		
	FAMDDA-GA	FAM-GA	p -value	FAMDDA-GA	FAM-GA	p -value
PID	72.55 (2.24)	68.53 (2.45)	0.006	15.6 (1.4)	13.9 (1.9)	0.03
AUS	78.89 (2.57)	72.61 (2.27)	0	15.2 (3.1)	13.0 (2.6)	0.081
HEA	77.02 (2.33)	71.50 (4.63)	0.010	9.3 (1.5)	8.5 (1.4)	0.151

Analysis

- The null hypothesis states no difference between the test accuracy rate of FAM-GA and FAMDDA-GA whereas the alternative hypothesis claims that the test accuracy rate of FAM-GA is lower than that of FAMDDA-GA.
- FAMDDA-GA produces higher accuracy rates than FAM-GA in the case studies.
- By running bootstrap hypothesis tests (Efron, 1979) with a significance level of 0.05, all p -values of the test for the accuracy between FAMDDA-GA and FAM-GA in the respective case studies are less than 0.05. This indicates that the classification performance of FAM-GA is statistically lower than that of FAMDDA-GA.

Results

Table 2 Performance comparison between the FAM and FAMDDA classifiers, in terms of test accuracy and number of nodes. The results are averages of eight runs, and their respective standard deviation is given in parenthesis. The p -values are the results from the bootstrap hypothesis tests between the classifier and its GA-based network (a denotes FAM vs. FAM-GA; and b denotes FAMDDA vs. FAMDDA-GA).

Dataset	Accuracy (%)			Size (Nodes)		
	FAMDDA	FAM	p -value	FAMDDA	FAM	p -value
PID	70.11	64.36	0.008 ^a	16.2	13.9	0.172 ^a
	(2.26)	(3.35)	0.007 ^b	(2.6)	(1.9)	0.403 ^b
AUS	72.25	67.61	0.006 ^a	15.2	13.0	0.513 ^a
	(2.42)	(3.58)	0 ^b	(3.1)	(2.6)	0.496 ^b
HEA	69.28	62.97	0.003 ^a	8.7	9.0	0.685 ^a
	(5.33)	(4.46)	0.018 ^b	(1.9)	(2.0)	0.271 ^b

Analysis

- All p -values of the test for the accuracy between FAMDDA (FAM) and FAMDDA-GA (FAM-GA) are smaller than 0.05 whereas all p -values of the test for comparing network size FAMDDA (FAM) and FAMDDA-GA (FAM-GA) are greater than 0.05.
- These results indicate that the generalization performances of individual classifiers are, respectively, lower than their GA-version counterparts, and the difference in the network sizes is statistically insignificant.

Summary

- A neuro-genetic system which integrates FAMDDA and the GA (FAMDDA-GA) for performing classification is proposed.
- The performance of the proposed FAMDDA-GA was assessed with three benchmark datasets. Performance comparison between FAMDDA-GA and FAM-GA and with individual classifiers was made statistically.
- The results reveal that the generalization performance of FAMDDA-GA is better than that of FAM-GA.

Further work

- Investigation on the feasibility of combining FAMDDA with Differential Evolution (Storn and Price 1997).
- Investigation on the possibility of including a local search mechanism into the existing framework for refining the solutions that have been obtained via global search.

References

- Baskar, S., Subraraj, P., Rao, M.V.C. (2001), "Performance of hybrid real coded genetic algorithms," *International Journal of Computational Engineering Science*, vol. 2, pp. 583-601.
- Blake, C. and Merz, C. (1998), UCI Repository of Machine Learning Databases, URL <http://www.ics.uci.edu/~mlearn/MLRepository.html>
- Carpenter, G. A. and Grossberg, S. (1987), "A massively parallel architecture for a self-organizing neural pattern recognition machine," *Computer Vision, Graphics and Image Processing*, vol. 37, pp. 54-115.
- Carpenter, G. A., Grossberg, S., Markuzon, N., Reynolds, J., and Rosen, D. (1992), "Fuzzy ARTMAP: A neural network architecture for incremental learning of analog multidimensional maps," *IEEE Trans. Neural Networks*, vol. 3, pp. 698-713.
- Efron, B. (1979), "Bootstrap methods: another look at the jackknife," *The Annals of Statistics*, vol. 7, pp. 1-26.

References

- Lam, H.K., and Leung, F.H.F. (2004), “Digit and command interpretation for electronic book using neural network and genetic algorithm,” *IEEE Trans. Systems, Man, and Cybernetics – Part B: Cybernetics*, vol. 34, pp. 2273 – 2283.
- Palaniappan, R., and Raveendran, P. (2002), “Individual identification technique using visual evoked potential signals,” *Electronics Letters*, vol. 28, pp. 1634-1635.
- Palaniappan, R., Raveendran, P., and Omatu, S. (2002), “VEP Optimal Channel Selection Using Genetic Algorithm for Neural Network Classification of Alcoholics,” *IEEE Trans. Neural Networks*, vol. 13, pp. 486-491.
- Storn, R., and Price, K. (1997), “Differential evolution—A simple and efficient heuristic for global optimization over continuous spaces,” *J. Glob. Optim.*, vol. 11, no. 4, pp. 341–359.