Evaluation on Neural Network and Fuzzy Method - in terms of Learning

Hung-Chang Lee, Tao Jung Wang
Department of Information Management, Tamkang University
Tel: 886-1-2-6215656 ext 645
E_mail: hclee@im1.im.tku.edu.tw

Like a dawn light scattering into the cloud sky of A.I., Neural Network and Fuzzy Logic become state-of-the-art technologies in exploring the intellectual. To make a judgment between both technologies, we propose an evaluation on them in the view point of learning to classification.

Since there are varieties models proposed within both technologies, we focus on most significant model, i.e., Back Propagation Network (BPN) [1] and Wang's fuzzy rule generator [2]. First in the evaluation, we introduce a *Gravity Effect Field* to illustrate these two models' influence under the existence of *one instance*. After that, we virtually construct two classifications problems and discuss the behaviors of both methods through the *Gravity Effect Field*. Finally, we propose another two real examples to demonstrate the results. We conclude that Wang's more suitable for the *piecewise region* classification and need *representative* or *complete* training samples than BPN. BPN is more training data tolerant and less network parameter sensible than that of Wang's fuzzy rule generator. However, basic instinct problems still exist, BPN behaviors more *black box* than fuzzy rule generator.

Key word: Fuzzy Learning, Back Propagation Neural Network

1. Introduction

Neural Network and Fuzzy Logic become state-of-the-art technologies in exploring the intellectual. To make a objective judge between them, we, therefore, need to know the cross effects resulting from both methods under the same testing environment. However, we limit our scope at the most significant one, Back propagation [1] and Wang's fuzzy rule generator [2] in terms of learning to classification. Under this evaluation, we want to know: what kinds of internal differences between them; what are their most applicable situation; what kind of training data quality they need; and how they are affected by their network parameters. To answer all the questions, we need to build a testing environment. Within this environment, we evaluate their behavior from *one instance, regions* to several *real examples.*

We introduce a *Gravity Effect Field* to illustrate these two models' influence under the existence of *one instance*. After that, we virtually construct two classifications problems and discuss the behaviors of both methods through the *Gravity Effect Field.* Finally, we propose another two real examples to demonstrate

0-7803-3687-9/96©1996 IEEE

the results.

We conclude that Wang's more suitable for the *piecewise region* classification and need *representative* or *complete* training samples than BPN. BPN is more training data tolerant and less network parameter sensible than that of Wang's fuzzy rule generator. However, basic instinct problems still exist, BPN behaviors more *black box* than fuzzy rule generator.

2. Evaluation Method and Related parameters

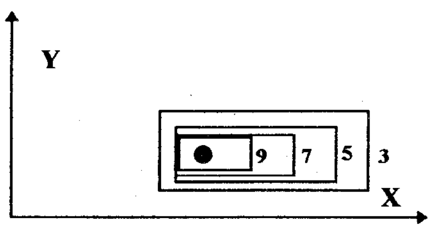
As mentioned above, we focus evaluation on Wang's fuzzy rule generator and Back propagation under the circumstance of learning to calcification. Two kinds of classification problem are feed into two both Wang's and BPN Model. The result are, therefore, generated and analyzed as a function of some important parameters.

Even in a limited technologies such as Wang's and BPN, there are lots of parameters to adjust, e.g., learning rate, montumn constant, ... in B.P[3]. and membership function, accumulation rule in Wang's method. To simplify the evaluation and discuss only the more important parameters. we address only parameters that influence the technologies most. Therefore, we consider the network the topology in BPN, fixed membership and original accumulation methods proposed only in Wang's method.

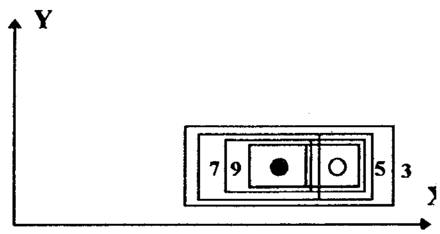
3. Gravity Effect Field and Evaluation Results

3.1 Gravity Effect Field

Before proceed the whole problem, we demonstrate the effect on one instance that both technologies bring, called Gravity Effect Field (like as planet bring effect to the whole university).



(a) GEF under one instance



(b) GEF under two instances ● ○

Fig. 3-1 Gravity Effect Field of Wang's Rule Generators

As for the neural network side, at least two instance are fed to destriminate the gravity Effect Field. As shown in Fig 3-2, there are equal partition on the plane. The place of partition line depends on the parameters that the network own. The instance is a point on a two dimension plane. The Gravity Effect Field (GEF) of the Wang's rule generators in shown in Fig. 3-1 (a) with a fuzzy region of 3, 5,7, 9. For a small number of fuzzy regions, thus the bigger extension of each fuzzy regions, the GEF is towards bigger. The place of the instance in the GEF may change according to it is on the left or right side of the fuzzy regions and this situation change even dramatically at a bigger number of fuzzy regions. Further, two or three instance are analyzed to show the combined effect of two or more GEF. As shown in Fig. 3-1 (b), the two instance in fuzzy regions 3 result in a conflict and solved in higher number of fuzzy regions. This situation comes from the Max dominated parameter and small difference in their Y dimension.

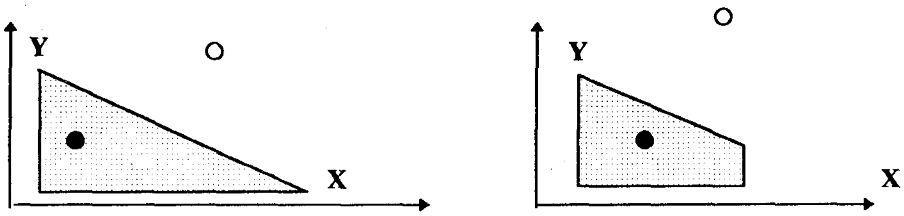
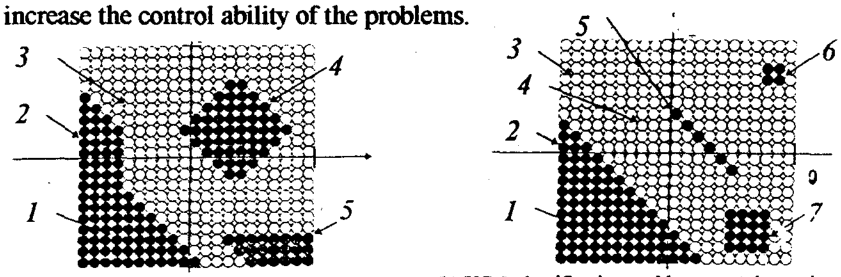


Fig. 3-2 Gravity Effect Field of Back Propagation network

3.2 Virtually Constructed two input variables problems

In this experiment, we use a virtually constructed problems to discuss how well is the training samples influence both models. To accomplish this experiment, we need a well-known problems, and thus, we introduce this kind of problems to increase the control ability of the problems.



(a) VC-1 classification problems contains regions (b) VC-2 classification problems contains regions, line and point

Fig. 3-3. Two Virtually Constructed two input variables classification problem.

We use two kinds of examples to explore the problems more deeper. First problems (called VC-1) contains only well-defined regions, and the other (called VC-2) includes point, line, and regions, shown in Fig. 3-3.

The results of Wang's rule generated with tuned parameters are listed in Table 3-1. In the VC-1 problems, we separate the training set into center or boundary cases, i.e., examples the more represent the region or not. Two kind of classification error ratio (called ER) is discussed; naming ER with no map rule and ER without. As shown in Table 3-1, Wang's rule dramatically change in the ER (9% ~23% to 27%~53%) depending on whether representative of the examples is enough or not. In the case of considering only the mapped rule, ER reduces almost one thirds (e.g., 46% TO 16% in VC-2). VC-2 bring more classification error than VC-1_C (center case) because the *Max dominated* factor in Wang's rule generation.

Table 3-1. Wang's Fuzzy Learning Results of VC-1 (center, border) and VC-2.

	Mapped Number		NoMapped Number	Error Ratio (include noMap)	Error Ratio (Exclude noMap)
	correct	wrong			
VC-1_C	31	3	6	23%	9%
VC-1_B	19	7	14	53%	27%
VC-2	27	5	18	46%	16%

Note: The region number for two input variables are: 7-7 for VC-1 and 7-7 for VC-2.

Table 3-2 shows the results of the BPN with tuned parameters. The BPN react less sensitivity to the representative of the training examples (30% to 45%). This is because the hyperplane can be floated and, thus the representative of the training examples can be complicated with tuned parameters. Another surprising result is that even in VC-2 problems i.e., includes point, line, and region, BPN behavior even better than VC-1. This result comes from that fact the exception case can be more easily eliminated than in Wang's methods.

Table 3-2. Back Propagation Learning Results of VC-1 (center, border) and VC-2.

	Training Set		Error Ratio (in Training set)	Testing Set		Error Ratio (in Testing set)
	correct	wrong		correct	wrong	
VC-1_C	27	5	15%	28	12	30%
VC-1_B	22	4	15%	22	18	45%
VC-2	34	1	4%	37	13	26%

Note: The parameters of BPN is: units in hidden layer is six and RMS is under 0.1.

To know the effect of how the model parameters and the number of training examples influence both technologies, we depict those results in Fig. 3-4. As shown in this figure, Wang's rule generator decrease dramatically with the increase number of training samples while BPN behavior more stable. BPN is less sensible to the models parameters than that of Wang's method.

We conclude from the above experiments that Wang's model need more representative and complete set of the training set than BPN. And if this is the case, Wang's method may overcome BPN.

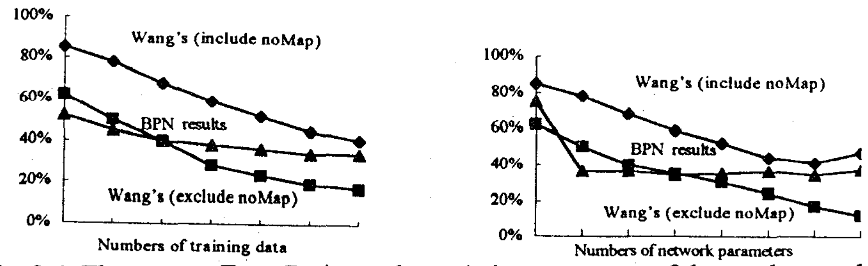


Fig. 3-4. The average Error Ratios under varieties parameters of data and network.

3.3 Real World problems

3.3.1 Road-damage Classification

The procedure to identify the degree of road-damage comes from the some pre-processing: i.e., clustering the pixels in road image into several defined damaged block, e.g., 2x2, 4X4, ... 16X16, which is used as feature vector for the classification. The degree of road-damage is classified into three categories; naming, light, medium, and heavy damage.

We use 24 training set to training, and another 103 images (77-12-14) to testing. The results of Wang's method(include or exclude noMap) and BPN is illustrated in Table 3-3.

Table 3-3. Road-damage Classification Results.

	BPN Classification				Wang's include Nomap				Wang's with all map			
	LD	MD	HD	Error ratio	LD	MD	HD	Error ratio	LD	MD	HD	Error ratio
Light Damage (LD)	65	12	0	15%	62	6	2	20%	70	7	0	9%
Medium Damage (MD)	3	7	2	40%	2	5	1	58%	3	7	2	25%
Heavy Damage (HD)	0	2	12	14%	0	3	9	21%	0	1	13	7%

3.3.2 Violet Flower Classification

The problem is the identify the three species of the violet flower according to the outlook of its calyx and petal; i.e., length and width. This is a typical hyperplane-partitioned problems in a four dimension space. From the experiment shown above, we can obtain the result preliminary; BPN will behavior better than Wang's method. The results as well as the parameters of the classification are depicted in Table 3-4. The average of BPN classification error is 3%; zero error rate in training set and 6% in testing set. Wang's results with a average of 13.5% or 6.5% under incomplete and complete cases: 4% in training set and 23% to 9% in testing set.

Table 3-4. Violet Flower Classification Results.

	Wang's Rule base			Error Ratio (include noMap)	Error Ratio (Exclue noMap)	BPN
	correct	wrong	Nomap			
Violet Flower	39	4	7	22%	9%	6%

4. Conclusion

BPN is more suitable for those hyperplane-partition classification problems than piecewise-connected region classification. Basic instinct problems with BPN is the black box mysterious behavior and maybe little slow upon the raining procedure. Some modified BPN methods to incorporate with the explanation capabilities will certainly be at the sacrifice of the degree of accuracy.

As for the Wang's fuzzy learning methods, it need more critical training pairs, more training data, and is rather sensible to tuning parameters. However, Wang's learning methods provide an production-rule like explanation capabilities which is just the common-use inference procedure with the human being. Some modification comes from three dimensions; adjustable membership function [5], the consequence part (THEN parts) compromise [5], and the unlearned rule inference [4,10,15].

Among those, Sudkamp [4] methods seem obtain good unlearned rule inference; called similarity interpolation. However, as propose in [15] say, too much rules with interpolation act like no rules at all. We think that limited rules with adjustable membership function and consequence part compromise [15] will be the right modification to fuzzy learning methods.

Reference:

[1] J. McClelland, D. Rumelhart, and G.E. Hinton, *Parallel Distributed Processing*, M.I.T. Press, Cambridge, 1986.

[2] L.X. Wang and J.M. Mendel, "Generating Fuzzy Rule From Examples," *IEEE Trans. on Sys. Man, and Cyber*. Vol. 22, No. 6, Dec. 1992, pp. 1414-1427.

[3] Y. C. Yeh, *Neural Network Model and application*, 1993.

[4] T. Sudkamp and R.J. Hammell II, "Interpolation Completion, and Learning Fuzzy Rules," *IEEE Trans. on Sys., Man, and Cyber*., Vol. 24, No.2, Feb. 1994, pp. 332-342.

[5] C. C. Jou, *Developing Decision Support System in Fuzzy Control*, Master Thesis, Dep. of Information Management, Tamkang University, 1994.

[6] G. Lee, *Pattern Recognition in Neural Networks with Road Damage Degree*, Master Thesis, Dep. of Traffic Management, Tamkang University, 1994.

[7] T. Munakata and Y. Jani, "Fuzzy Systems: An Overview," *Communication of ACM*, Vol. 37, No. 3, March 1994, pp. 77-84.

[8] L. A. Zadeh, "Fuzzy Logic, Neural Networks, and soft Computing," *Communication of ACM*, Vol. 37, No. 3, March 1994, pp. 87-105.

[9] Y. Yuan and M.J. Shaw, "Introduction of Fuzzy Decision Trees," *Fuzzy Sets and Systems*, 1995. pp. 125-139.

[10] M.S. Yeh and S. M. Chen, "An Algorithm for Generating Fuzzy Rules from Relational Database System," *Inte'l conference on Information Management*, 1995, pp. 219-226.

[11] C. C. Lee, "Fuzzy Logic in Control System: Part I," *IEEE Trans. on Sys. Man. and Cyber*, Vol. 20, No. 2, pp. 404-418, 1990.

[12] C. C. Lee, "Fuzzy Logic in Control System: Part II," *IEEE Trans. on Sys. Man. and Cyber*, Vol. 20, No. 2, pp. 419-435, 1990.

[13] L. Wang, and J. M. Mendel, "Fuzzy basis functions, universal approximation and orthogonal least-square learning," *IEEE Trans. Neural Networks*, Vol. 3, No. 5, pp. 807-814, 1992.

[14] T. A. Sudkamp, "Similarity, Interpolation, and Fuzzy rule construction," *Fuzzy Sets and Systems*, pp. 73-86, 1993.

[15] T. J. Wang, *Intellectual Information Technologies evaluation in terms of classification*, Master Thesis, Dep. of Information Management, Tamkang University, 1996.