TELKOMNIKA, Vol.11, No.3, September 2013, pp. 611~616
ISSN: 1693-6930, accredited A by DIKTI, Decree No: 58/DIKTI/Kep/2013
DOI: 10.12928/TELKOMNIKA.v11i3.1271

Balanced the Trade-offs Problem of ANFIS using Particle Swarm Optimization

Dian Palupi Rini[1], Siti Mariyam Shamsuddin[2], Siti Sophiayati Yuhaniz[2]
[1]Faculty of Computer Science, Sriwijaya University, Indonesia
[2]Soft Computing Research Group UTM, Skudai Johor bahru Malaysia
*Corresponding author, e-mail: dian_rini@unsri.ac.id, {mariyam, sophia}@utm.my

Abstrak

Peningkatan nilai perkiraan akurasi dan interpretabilitas pada sebuah sistem samar adalah persoalan penting baik pada teori sistem samar ataupun pada aplikasinya. Telah diketahui bahwa optimisasi secara silmultan pada kedua persoalan tersebut adalah saling seimbang, namun hal ini akan meningkatkan pencapaian system dan menghindari pelatihan berlebihan. Particle Swarm Optimisasi (PSO) adalah bagian dari algoritma evolusioner yang merupakan calon algoritma yang baik untuk memecahkan masalah tersebut dan memiliki ruang pencarian global yang lebih baik. Tulisan ini mengenalkan sebuah integrasi antara PSO dan ANFIS untuk optimisasi pembelajarannya terutama pada penyesuaian nilai parameter fungsi kepemilikan dan menentukan jumlah aturan yang optimal untuk memperoleh nilai klasifikasi yang lebih baik. Usulan algoritma ini telah dites pada 4(empat) dataset standar dari mesin pembelajaran UCI, yaitu: dataset Iris Flower, Haberman's Survival Data, Balloon dan Thyroid. Hasil menunjukkan bahwa nilai klasifikasi yang lebih baik pada algoritma PSO-ANFIS yang diusulkan dan kompleksitas waktunya menurun bersesuaian.

Kata kunci: *ANFIS, interpretabilitas, akurasi, algoritma evolusioner, particle swarm optimisasi*

Abstract

Improving the approximation accuracy and interpretability of fuzzy systems is an important issue either in fuzzy systems theory or in its applications . It is known that simultaneous optimization both issues was the trade-offs problem, but it will improve performance of the system and avoid overtraining of data. Particle swarm optimization (PSO) is part of evolutionary algorithm that is good candidate algorithms to solve multiple optimal solution and better global search space. This paper introduces an integration of PSO dan ANFIS for optimise its learning especially for tuning membership function parameters and finding the optimal rule for better classification. The proposed method has been tested on four standard dataset from UCI machine learning i.e. Iris Flower, Haberman's Survival Data, Balloon and Thyroid dataset. The results have shown better classification using the proposed PSO-ANFIS and the time complexity has reduced accordingly.

Keywords: *ANFIS, interpretability, accuracy, evolutionary algorithms, particle swarm optimization*

1. Introduction

The strength of neuro-fuzzy systems involves two contradictory requirements in fuzzy modelling: interpretability and accuracy. Improving the approximation accuracy and interpretability of fuzzy systems is an important issue either in fuzzy systems theory or in its applications [1]. An adaptive neuro-fuzzy inference system (ANFIS) based on TSK model is a specific approach of neuro-fuzzy that has shown significant results in classification problem. The structure of adaptive neuro-fuzzy system (ANFIS) is similar with general neuro-fuzzy system. It learns features in the data set, and adjusts the system parameters according to a given error criterion [2]. But in the rule layer, it gives a number of nodes that represents a self generating all possible fuzzy rules in neuro-fuzzy structure. The self generating rules give a chance to produce effective or ineffective rules. So, a simultaneously technique that generates the ANFIS that has good accuracy and has effective rules is necessary in this research.

Particle swarm optimization (PSO) is one of evolutionary algorithms (EAs) techniques that is widely used and rapidly developed by researchers, due to its easy implementation and few particles required to be tuned. Furthermore PSO is a very simple concept and paradigms which can be implemented in a few lines of computer code. It requires only primitive

Received April 25, 2013; Revised July 4, 2013; Accepted July 18, 2013

mathematical operators, and is computationally inexpensive in terms of both memory requirements and speed [3-4]. The research on conflicting fuzzy neural network (FNN) problem by using PSO has been publish by [5], but in this research, the rules is generates by an expert.

Based on the advantages of PSO and the needed of ANFIS to stabilise the conflicting criteria in neuro-fuzzy problem, this study will examined particle swarm optimization to balance accuracy and interpretability trade-offs in ANFIS structure.

The organisation of the paper is as detailed. Section 2 describes the adaptive neuro-fuzzy concepts. Section 3 describes the proposed modified neuro-fuzzy classifier using PSO in detail. Section 4 describes the performance evaluation procedure for the classifier. The conclusions are presented in Section 5.

2. Adaptive Neuro Fuzzy System

As the general fuzzy modelling structure, ANFIS is mainly characterised by two features that assess the quality of the obtained fuzzy models: interpretability and accuracy [6]. Interpretability refers to the ability of fuzzy models to represent the habitual of its systems. Some of researchers agreed that the interpretability covers several issues, such as the model structure, the number of input variables, and the number of fuzzy rules, the number of linguistic terms, and the shape of the fuzzy sets. The interpretability is important issues in ANFIS process because it is affecting the complexity and processing time of the system. Based on [1], interpretability can be improved by fine-tuning the fuzzy rules with regularisation such as growing and pruning fuzzy rule number to find the effective one from all possible fuzzy rules in neuro-fuzzy structure.

The accuracy has straightforward definition. It refers to the capability of the fuzzy model to faithfully represent the modelled system. The closer the model to the system, the accuracy closeness is higher as the similarity between the responses of the real system and the fuzzy model is understood. One of the neuro-fuzzy advantages is neuro-fuzzy can be designed based solely on approximation and the linguistic information. Therefore, based on [7] the satisfactory level of accuracy can be achieved by tuning the network structure and parameter learning of neuro-fuzzy and based on [8], there are a high relationship between membership function parameter and accuracy. This statement has led to ideas on how to tuning membership functions to improve the accuracy of neuro-fuzzy systems.

Based on the two contradictory requirements in fuzzy modelling, ANFIS can be formulated with two objectives that will simultaneously optimised, i.e.

1. Tuning the parameter learning of ANFIS to obtain a good performance (Accuracy) of fuzzy modelling based on Mean Squared Error (MSE). It is given as:

$$f_1 = \min\left\{\frac{1}{N}\sum_{j=1}^{N}\left(y_j - yd_j\right)^2\right\} \tag{1}$$

   Where y_j and yd_j are the network output and the desired output, respectively, and N is the number of data.

2. Growing and pruning fuzzy rule number to obtain a good Interpretability of fuzzy modelling. It is given as:

$$f_2 = \min\left\{\sum_{r=1}^{R} \mathrm{O}_r\right\} \tag{2}$$

Where R is the maximum number of rule nodes, and $\mathrm{O}_r \epsilon\, \mathrm{O}$ is a binary value used to indicate whether the rule node r exists or not. It works as a switch to turn a rule node on or off.

3. Modified Adaptive Neuro Fuzzy System using PSO

Particle swarm optimization (PSO) is part of evolutionary algorithm that will use by ANFIS as a learning method to optimise both conflicting criteria. A definition of PSO based on describes PSO as a swarm of particles, where particle represent a potential solution.

Particle swarm optimization consists of particle where position of the particle is influenced by velocity. Let $\boldsymbol{x_i(t)}$ denote the position of particle $\boldsymbol{i}$ in the search space at time step $\boldsymbol{t}$; unless otherwise stated, t denotes discrete time steps. The position of the particle is changed by adding a velocity, $\boldsymbol{v_i(t)}$ to the current position:

$$x_i(t) = x_i(t) + v_i(t) \tag{3}$$
$$v_i(t) = c_1 r_1\big(pbest(t) - x_i(t)\big) + c_2 r_2\big(gbest(t) - x_i(t)\big) \tag{4}$$

Where $\boldsymbol{c_1}$ and $\boldsymbol{c_2}$ are acceleration coefficient, $\boldsymbol{r_1}$ and $\boldsymbol{r_2}$ are random vector and $\boldsymbol{pbest}$ and $\boldsymbol{gbest}$ local best and global best respectively.

PSO is a potential technique to solve the ANFIS problems. In the context of PSO, ANFIS is considered as one particle and parameters that influence the ANFIS process is considered as a dimension of the particle. While in the PSO, there are some particles; means there are some ANFIS processes that compete to achieve the potential solution of objective function of ANFIS.

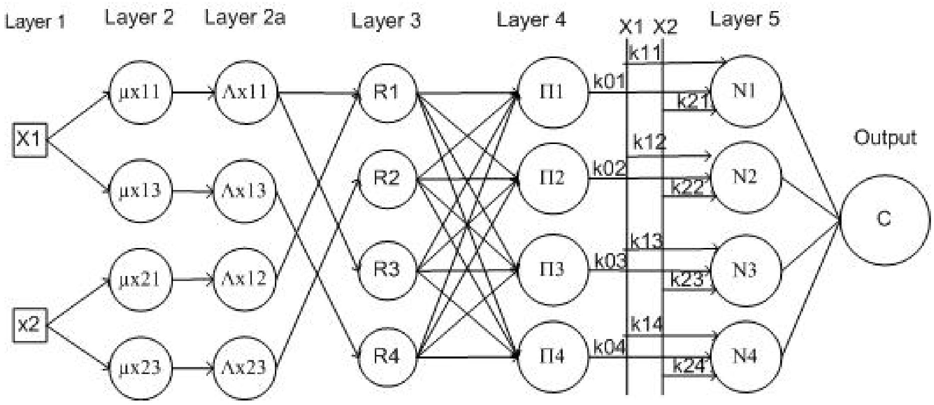

Figure 1. The adaptive neuro-fuzzy architecture.

Layer 1 is input layer, while $xi\ (i = 1 \dots n)$ are input signals; Layer 2 is fuzzification process of antecedent parameter namely membership function (μ), that each node connected with single node of layer 2a. The connections present modify the membership function value; Layer 3 is the fuzzy rule base layer; while layer 4 is the normalization layer. In this layer, the optimized fuzzy rule (O) will selected. Layer 5 is the defuzzification layer while the layer is affected by consequent parameter (k), and C is output classification.

As illustrated in Fig. 1, each node in layer 2 has single connectivity with node in layer 2a, means each membership function parameter in layer 2 will modified in layer 2a to obtain the appropriate membership which will used to measure the output and get significant minimum error. Further, each node in layer 3 is connected with each node in layer 4, where each of the connection represents a fuzzy rule. The optimal fuzzy rule will selected based on how importance of each fuzzy rule in the corresponding system.

In the proposed PSO-ANFIS, PSO will used to tune all the parameters learning of ANFIS and simultaneously will growing and pruning fuzzy rule number in ANFIS structure to get the best value of parameters learning and fuzzy rule number, respectively. The propose PSO-ANFIS will used to design the ANFIS with a small number of fuzzy rules with high performance accuracy. This task is performed through maximizing the accuracy, minimizing the number of selected rules.

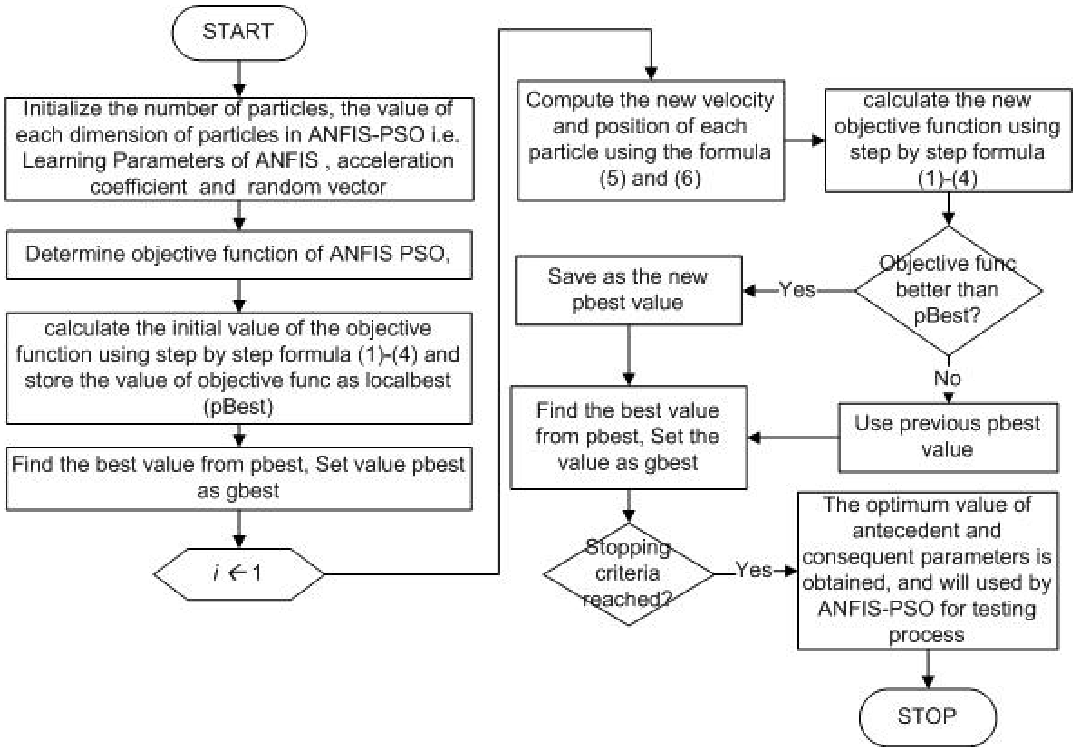

Figure 2. An overview of the ANFIS-PSO process

Figure 2 presents the process of ANFIS PSO. Below is illustration of PSO used to tune all the parameters learning of ANFIS and simultaneously will grow and prune fuzzy rule number in ANFIS structure:

1. Initialize particle position (X_d) using equation (16) and velocity (v_d) with d number of dimensions.
2. Initialize fitness function (f_i) for ANFIS-PSO. Fitness function of ANFIS-PSO is the objective function of the ANFIS i.e. equation (1) and (2).
3. Find objective function of ANFIS using equation (5) – (10). Based on the fitness function find particle's position (x_d) in each local best $(pbest_d)$. If fitness (x_d) is better than fitness $(pbest_d)$ then $pbest_d$=x_d.
4. Find best value of $pbest_p$. Set best of $pbest_p$ as $gBest$
5. Update velocity and position using equation (3) and (4)
6. For each particle, find new fitness function(f_{x_i}). Check $errorfunc$ as fitness function based on step 3) and find the gbest value based on step 4)
7. Check whether the value has convergence then stop, otherwise back to 5). Check, if gbest_fitness better that stopping criteria then process stop, else goto step 5)

4. Experimental studies

To evaluate performance of the proposed algorithms, several experiments are conducted on four real-world datasets from UCI machine learning (http://archive.ics.uci.edu/ml/datasets.html): Iris Flower, Balloon, Haberman's Survival Data and Thyroid. Table 2 is summarises the characteristics of the datasets used in this experiments.

Table 2. Characteristics of datasets

Dataset	Samples	Input No.	Output No.	Class	No. of Instance in Each Class
Iris Flower	150	4	1	3	C1= 50 inst, C2 =50, C3=50
Haberman's	306	3	1	2	C1= 255 inst, C2 = 81
Balloon	20	4	1	2	C1= 8 inst, C2 = 12
Thyroid	215	5	1	2	C1= 150 inst, C2=35, C3=30

To ensure the consistency of the data, values of the datasets are normalised in the range of [0, 1] using normalisation formula. Then hold-out cross validation is used to test the performance of the system. Hence, the datasets are partitioned into two sets: a training set and a testing set. The training set is used to train the network in order to get the ANFIS learning while the testing set is used to test the generalisation performance of ANFIS and is not seen during the training process. These datasets are partitioned randomly i.e. 80% of data are used for the training set and the rest 20% for the testing set.

Table 3. Specification of proposed method

PSO Parameter	Value
Number of particle	50
Number of linguistic fuzzy set	3
Number of iterations	1000
Obj. Function 1 ($f1$)	Mean Square Error (MSE)
Obj. Function 2 ($f2$)	Optimal Number of rule
acceleration coefficient	$c_1 = 0{,}5 \;\; c_2 = 1$
random vector r_1 and r_2	random

Initial values that required in ANFIS-PSO process are described in table 3. The experiments of ANFIS-PSO are conducted based on ten runs on each dataset. The mean and SD (indicates the mean value and standard deviation, respectively) result about the MSE (satisfy accuracy), the number of rule and time consume are calculated and reported

Table 4 shows the mean and standard deviation of ANFIS learning based PSO algorithms. The table shows that the best error rate on the training process is Balloon and on the testing process is Iris Flower. The result indicates that the error rate might not be influenced by the number of input parameters and samples but it might be due to the distributions of the datasets itself. For example, although Habermans's datasets have three input variables, but the distribution of its classes is extremely imbalanced (there are 255 instances for class 1 and 81 instances for class 2), thus it be seen that the significant error of its value is not so good compared to the balloon and iris data. Balloon has fewer instances then others and the distribution data is rather normal, but there is a significance difference of error results between training and testing. On the contrary, Iris data have more variable than Haberman's and more instances than balloon, but it has normal distribution. The output shows minimal error value either in training or testing data. Furthermore, thyroid data obtained the worse results in both set of data in which it has more variable than others and has abnormal distribution data. So, it might be concluded that the distribution of each class gives a large effect to the error rate value.

Table 4. Result of ANFIS PSO

Dataset	Experiments	Training	Testing	No of Optimal Rule	Time (s)
		MSE ($f1$)	Error Rate ($f1$)	($f2$)	
Iris Flower	Mean	0.151	0.155	30.2	80.2
	SD	0.057	0.103	9.64	25.5
Haberman' s	Mean	0.163	0.195	16.1	38.9
	SD	0.005	0.004	2.85	0.57
Balloon	Mean	0.116	0.241	12.5	13.6
	SD	0.055	0.127	8.86	0.52
Thyroid	Mean	0.189	0.580	115.4	147.4
	SD	0.021	0.063	40.313	2.675

In number of optimal rule ($f2$) column, the smallest number of rule is obtained by balloon, while thyroid has the most one. From the table 3, it is known that Balloon data has the smallest number of data and thyroid has the most one. It seems that there is a correlation between number of input and number of optimal rule. However, Haberman's has less input variable but the optimal number of rule is more than balloon. If seeing from number of sample (instance), the Haberman's has more number of instances than Balloon. While Iris data has more optimal numbers due to it has more input than Haberman's and more instances than balloon. So, it might be concluded that number of input and instances give a substantial contribution to find optimal number of rules.

Table 4 shows that the time process is balanced with the number of rule. The more the number of rule the more the time process is obtained. For an example, Balloon data have less number of optimal rules compared to time spent less than others. Thyroid has much number of rule than others, then it also spend more time in it process. So, it can be concluded that the time complexity will reduced while the optimal number of rule is obtained.

Table 5. Classification Measurement of proposed method

Dataset	Sensitivity	Specificity	Accuracy
Iris Flower	0.891155	0.029563	0.941667
Haberman	0.794521	0.24	0.790984
Balloon	1	0	1
Thyroid	0.63081	0.086873	0.854651

In the context of evaluation of the classification measurement of PSO-ANFIS, an average of sensitivity, specificity, and predicting accuracy of model was performed in Table 5. Based on the table 5, when the sensitivity is high and the specificity gives low values, the accuracy is better due to the correlation between these two measurements. In this result, the highest classification accuracy is obtained by balloon dataset (Accuracy equals to one means that all the datasets have been classified precisely), while Iris flower and Haberman’s is in the second and third ranking score, respectively and the worst result is obtained by thyroid. By observing from the behaviour of data, Iris flower, Balloon and thyroid have more input variable and have less number of instance compare with Haberman’s, but the classification accuracy is bigger than Haberman’s. However, the distribution of each class in Haberman’s dataset is most un-normal compared to the other. So in this context, the distribution data of each class might result to a large effect in classification measurement besides inconsistent data class of dataset. However, for over all the results given feasible accuracy in classification for all data sets which in all datasets accuracy more than 0.75 is obtained.

5. Conclusion

In this paper, an approach multiple solutions based on PSO is proposed and applied to develop generalisation and classification accuracy of several objectives for adaptive neuro-fuzzy system (ANFIS). This is done by simultaneously optimising the ANFIS architecture based on two criteria: enhance the accuracy and reduce the complexity based on interpretability. The findings indicated that the proposed method provides promising accuracy which could reduce time complexity.

References

[1] Paiva, R.P. and A. Dourado, Interpretability and learning in neuro-fuzzy systems. Elsevier. *Fuzzy Sets and Systems*. 2004. 147: 17-38.

[2] Negnevitsky, M., Artificial Inteligence: A guide to intelligent systems. second edition ed. 2005, England: Pearson Education Limited. 415.

[3] Bai, Q., Analysis of Particle Swarm Optimization Algorithm. *Computer and Information Science*. 2010; 3(1): 180-184.

[4] Engelbrecht, A.P., Fundamental of Computational Swarm Inteligent. First ed. 2005, The atrium, Southern Gate, Chichester, West Sussex PO19 8SQ, England: John Wiley & Sons Ltd.

[5] Ma, M., et al., Fuzzy Neural Network Optimization by a Particle Swarm Optimization Algorithm, in Advances in Neural Networks - ISNN 2006, J. Wang, et al., Editors. Springer Berlin / Heidelberg. 2006: 752-761.

[6] Di Nuovo, A.G. and V. Catania. *Linguistic Modifiers to Improve the Accuracy-Interpretability Trade-Off in Multi-Objective Genetic Design of Fuzzy Rule Based Classifier Systems*. in Intelligent Systems Design and Applications, 2009. ISDA '09. Ninth International Conference on. 2009.

[7] Lee, C.-H. and C.-C. Teng, Fine Tuning Of Membership Functions For Fuzzy Neural System. *Asian Journal of Control*. 2001; 3(3): 216-225.

[8] Zeng, X.-J. and M.G. Singh. A Relationship Between Membership Functions and Approximation Accuracy in Fuzzy Systems. *IEEE Transactions On Systems, Man, And Cybernetics-Part B: Cybernetics*, 1996. 26(1): 176-180.