

A Construction of Fuzzy Model for Data Mining

Do Wan Kim*, Young Hoon Joo**, and Jin Bae Park*

* Department of Electrical and Electronic Engineering, Yonsei University

** School of Electronic and Information Engineering, Kunsan National University

Abstract

A new GA-based methodology using information granules is suggested for the construction of fuzzy classifiers. The proposed scheme consists of three steps: selection of information granules, construction of the associated fuzzy sets, and tuning of the fuzzy rules. First, the genetic algorithm (GA) is applied to the development of the adequate information granules. The fuzzy sets are then constructed from the analysis of the developed information granules. An interpretable fuzzy classifier is designed by using the constructed fuzzy sets. Finally, the GA are utilized for tuning of the fuzzy rules, which can enhance the classification performance on the misclassified data (e.g., data with the strange pattern or on the boundaries of the classes). To show the effectiveness of the proposed method, an example, the classification of the Iris data, is provided.

Key Words : Fuzzy classifier, data mining, fuzzy set, information granules, genetic algorithm.

1. Introduction

Since Zadeh suggested the fuzzy set in 1965 [3], the fuzzy theory has been regarded as a resolution of a mathematical modelling of the intuitive human knowledge [1,2]. With this technical point of view, various construction methods of fuzzy classifiers have been proposed [4-10]. A table look up scheme was studied to generate the fuzzy rules for the fuzzy classifier directly [5], or by using the genetic algorithm (GA) [6]. Abe *et al.* designed several fuzzy classifiers from the various information granules to express the characteristics of the fuzzy classifiers by considering the shapes of the information granules [4,8].

Generally, the construction process of the the fuzzy classifier consists of the following three stages: (i) selecting the information granules, (ii) constructing the fuzzy sets associated with the information granules, and (iii) tuning the fuzzy rules. In Stage (i), it is important to optimally select the information granules from the feature spaces because the performance of the fuzzy classifier highly depends on the choice of the information granules. Unfortunately, dividing the feature space into the optimal information granules is generally known to be a complex and mutually associated problem. An approach to resolve this difficulty is to generate the activation and the inhibition hyperboxes, recursively [8]. However, it generates too many ones to analyze information granules easily as well as to show the pattern

visually. In Stage (ii), the fuzzy sets produced by information granules should fulfill two requirements: the fuzzy sets should exhibit a large degree of overlapping, and the fuzzy sets should describe the context of the information granule. Wu *et al.* proposed a construction technique of the fuzzy sets via the α -cut and the similarity degree [10]. However, when all the fuzzy sets for a feature are overlapped, some of the fuzzy sets are merged into one. Consequently, they may be irrelevant to the patten classification because all classes are represented by only one fuzzy set. In Stage (iii), the tuning of the fuzzy rules constructed from the previous stages is necessary to appropriately deal with the misclassified data. Although Abe *et al.* proposed a tuning method of the slope of the membership function [4], their technique neither consider the overfitting nor the outlier problem.

Motivated by the above observations, this paper aims at improving the performance of the fuzzy classifier by resolving the above-mentioned problems. To this end, we propose a GA-based construction method of the fuzzy classifier. In order to determine optimal information granules, the GA, which has been shown to be a flexible and robust optimization tool [11], are used in Stage (i). An efficient fuzzy sets constructing technique based on the α -cut and the similarity degree is then proposed in Stage (ii) to extract, merge, and reduce the fuzzy sets from the developed information granules, which produces a set of human interpretable classification rules in the form of the multi-inputs and multi-outputs (MIMO) Takagi-Sugeno (T-S) fuzzy model. In Stage (iii), in order to decrease the number of the misclassified data, the GA is also used to tune parameters of the obtained fuzzy sets and to newly generate additional

접수일자 : 2002년 10월 18일

완료일자 : 2003년 02월 01일

fuzzy rules.

2. Preliminaries

2.1 Fuzzy Classifier Model

To perform the classification, the MIMO T-S fuzzy system is designed by

$$R_i \text{ IF } x_1 \text{ is } A_{i1} \text{ and } \dots \text{ and } x_N \text{ is } A_{iN} \text{ THEN the class is } i \text{ (} y_{i1} = \xi_{i1}, \dots, y_{iL} = \xi_{iL} \text{)} \quad (1)$$

where $x_j, j \in I_N = \{1, 2, \dots, M\}$, is the j th feature; $R_i, i \in I_L = \{1, 2, \dots, L\}$, is the i th fuzzy rule; $y_{ih}, (i, h) \in I_L \times I_L$, is the h th singleton output in the i th rule; $\xi_{ih} \in [0, 1]$ is a real constant value; and $A_{ij}, (i, j) \in I_L \times I_N$ is the fuzzy set to be determined from the information granules.

Remark 1 Here, we initially equate the number of the classes to the one of the fuzzy rules, and set the i th output of the i th fuzzy rule R_i , y_{ii} as a maximum so that the fuzzy rule R_i can well describe the i th class.

By using product inference engine, singleton fuzzifier, and center average defuzzifier, the global output of the h th class inferred from the T-S fuzzy system (1) is represented by

$$y_h = \sum_{i=1}^L \theta_{A_i}(x) y_{ih}$$

where $\theta_{A_i}(x) = \frac{\prod_{j=1}^N A_{ij}(x_j)}{\sum_{i=1}^L \prod_{j=1}^N A_{ij}(x_j)}$ and $A_{ij}(x_j) \in [0, 1]$ is

the membership function value of the linguistic variable x_j on the fuzzy set A_{ij} . The global output y_h implies the confidence measure of each class. The predicted class is computed as

$$y = \max_{h \in I_L} y_h$$

2.2 Information Granules

Information granules are a tool that can effectively describe the input-output pattern from the given data. Figure 1 shows an example of information granule in 2 dimensional case, where '*' denotes given data and the rectangle means the information granule or the hyperbox. Most of the existing clustering techniques operate on the numeric objects and produce the representatives that are again entirely numeric, whereas the information granules technique operates on the visual objects and produces the representatives that are again visual [12].

Remark 2 To determine the optimal fuzzy region in the supervised learning, the information granules should satisfy the following requirements as much as possible:

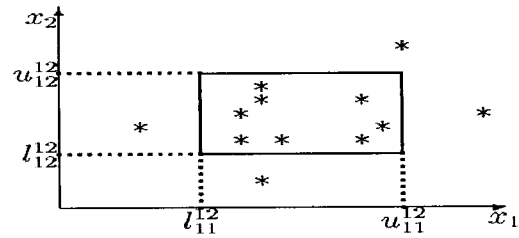


Fig. 1. Example of information granules in 2 dimensional case. *: the given data; the rectangle: the information granule

- (i) The information granules should cover as many data as possible so that the density of the data within the information granules is made as high as possible [12].
- (ii) There should be the minimum overlaps among the information granules because each information granule should describe different characteristics in the given data.
- (iii) One information granule should cover one class within the admissible limits so that the fuzzy sets extracted from an information granule describe a class on several feature spaces.

3. Classification Using Information Granules

3.1 Optimization of Information Granules Using the GA

This section discusses the parameters optimization of information granules using the GA. In order to obtain the fuzzy classifier with low complexity, we concentrate on determining the information granules subjected to the number of the class regions without generating the additional information granules.

Obviously according to Remark 2, the information granules should be developed such that the following objectives are fulfilled:

$$\begin{aligned}
 & \text{Maximize}_{1 \leq m < n \leq N} J_1^{mn} = \frac{\Lambda^{mn}}{D} \\
 & \text{Minimize}_{1 \leq m < n \leq N} J_2^{mn} = \frac{\sum_{i=1}^L \Omega_i^{mn}}{\Lambda^{mn}} \\
 & \text{Maximize}_{1 \leq m < n \leq N} J_3^{mn} = \prod_{i=1}^N \frac{\Psi_i^{mn}}{\Omega_i^{mn}} \\
 & \text{Maximize}_{1 \leq m < n \leq N} J_4^{mn} = \sum_{i=1}^L \frac{\Omega_i^{mn}}{\text{area}(l_{ij}^{mn}, u_{ij}^{mn})}
 \end{aligned} \quad (2)$$

where the superscript 'mn' denotes a 2 dimensional feature space with x_m and x_n . Throughout this paper, it fulfills $1 \leq m < n \leq N$. J_1^{mn} is the covering rate; J_2^{mn} is the overlapping rate; J_3^{mn} is the classification performance; J_4^{mn} is the density rate, and D is the number of the

given data; A^{mn} is the number of the data covered by all information granules in the mn th feature space; Ω_i^{mn} is the number of the data covered by the i th information granule in the mn th feature space; Ψ_i^{mn} is the largest data number of the intersections among the class regions and the i th information granule and $area(l_{ij}^{mn}, u_{ij}^{mn})$ is the area of the i th information granule to be searched, where l_{ij}^{mn} and u_{ij}^{mn} are the left-lower and the right-upper vertex of the i th hyperbox in the mn th feature space, respectively. The multi objective function for the mn th feature space can be written as

$$\text{Minimize}_{1 \leq m < n \leq N} J^{mn} = \frac{J_2^{mn}}{J_1^{mn} J_3^{mn}} + \frac{\nu_1}{J_4^{mn}} \quad (3)$$

where ν_1 is a positive scalar to adjust the weight between the two terms in (3).

The GA represents the searching variables of the given optimization problem (3) as the chromosome that contains one or more sub-strings. In this case, the searching variables are l_{ij}^{mn} and u_{ij}^{mn} . A convenient way to convey the searching variables into a chromosome is to gather all searching variables associated with all information granules in the mn th feature space, and to concatenate the strings as follows:

$$G^{mn} = \{(l_1^{mn}, u_1^{mn}), \dots, (l_N^{mn}, u_N^{mn})\} \\ G = \{G^{12}, G^{13}, G^{23}, \dots, G^{(n-2)(n-1)}, G^{1n}, \\ \dots, G^{(n-1)n}, G^{1(n+1)}, \dots, G^{(N-1)N}\} \quad (4)$$

where G^{mn} is the parameter substring of the mn th feature space in a chromosome and G denotes a chromosome.

Initial population is made up with initial individuals to the extent of the population size. To efficiently perform the optimization, the search spaces in GA should be reduced as much as possible. In this case, we utilize the following constraints on the search spaces

$$S(G^{mn}) \subset F^{mn} \quad (5)$$

$$S(l_{ij}^{mn}, u_{ij}^{mn}) \subset C_i^{mn} \quad (6)$$

where $S(G^{mn})$ and $S(l_{ij}^{mn}, u_{ij}^{mn})$ are the search space for the sub-string and the i th information granule, respectively, F^{mn} implies the mn th feature space, and C_i^{mn} denotes the i th class region.

Since the GA originally searches the optimal solution so that the fitness function value is maximized, it is necessary to map the objective function (3) to the fitness function formed by

$$f(J^{mn}) = \frac{J_1^{mn} J_3^{mn}}{J_2^{mn}} + \frac{J_4^{mn}}{\nu_1} \quad (7)$$

The best chromosome $(G_t)^*$ is composed of the best sub-chromosomes $(G_t^{mn})^*$ with the highest fitness

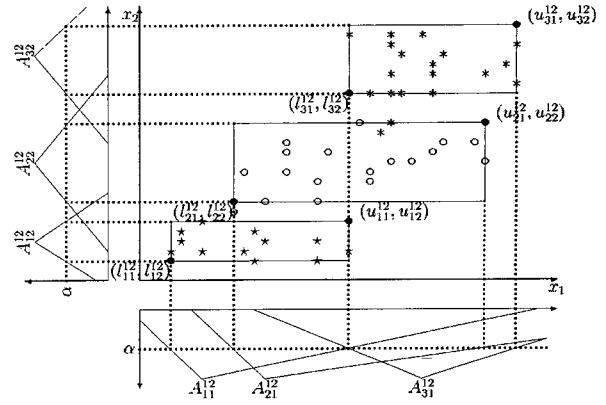


Fig. 2. Example of the fuzzy sets extraction with $L=3$ and $N=2$.

values $(f_t^{mn})^*$ at the t th generation as follows:

$$(G_t)^* = \{(G_t^{12})^*, (G_t^{13})^*, (G_t^{23})^*, \dots, (G_t^{(n-2)(n-1)})^*, \\ (G_t^{(n-2)(n-1)})^*, (G_t^{1n})^*, \dots, (G_t^{(n-1)n})^*, \\ (G_t^{(n-1)n})^*, (G_t^{1(n+1)})^*, \dots, (G_t^{(N-1)N})^*\}$$

where $t \in I_T = \{1, 2, \dots, T\}$ is the genetic process generation number. Then, the best overall fitness value in the m th generation can be written as

$$(f_t)^* = \sum_{1 \leq m < n \leq N} (f_t^{mn})^*$$

3.2 Construction of Fuzzy Sets

This subsection describes how to construct the fuzzy sets from the information granules developed in the previous subsection. The α -cut operation and the similarity measurement are utilized for extracting, merging, and removing the fuzzy sets. In the mn th feature space, A_{ij}^{mn} extracted from L information granules are represented by:

$$A_{ij}^{mn}(x_j) = \begin{cases} \frac{x_j - a_{ij}^{mn}}{|b_{ij}^{mn} - a_{ij}^{mn}|} & \text{if } a_{ij}^{mn} \leq x_j < b_{ij}^{mn} \\ \frac{c_{ij}^{mn} - x_j}{|c_{ij}^{mn} - b_{ij}^{mn}|} & \text{if } b_{ij}^{mn} \leq x_j < c_{ij}^{mn} \\ 0 & \text{otherwise} \end{cases} \quad (8)$$

where l_{ij}^{mn} and u_{ij}^{mn} denote the left-lower and the right-upper vertices of the i th information granule on the horizontal x_j -axis, and $b_{ij}^{mn} = \frac{l_{ij}^{mn} + u_{ij}^{mn}}{2}$, respectively. The parameters a_{ij}^{mn} and c_{ij}^{mn} in (8) are computed from the α -cut operation $(A_{ij}^{mn})_\alpha = \{x_j \in U_j | A_{ij}^{mn}(x_j) \geq \alpha\}$, $U_j = [x_{j_{\min}}, x_{j_{\max}}]$, $\alpha \in [0, 1]$ of a fuzzy set A_{ij}^{mn} as follows:

$$a_{ij}^{mn} = b_{ij}^{mn} - \frac{b_{ij}^{mn} - l_{ij}^{mn}}{1 - \alpha} \\ c_{ij}^{mn} = b_{ij}^{mn} + \frac{u_{ij}^{mn} - b_{ij}^{mn}}{1 - \alpha}$$

Figure 2 shows an example of extracting the fuzzy sets with $L=3$ and $N=2$. A_{ij}^{mn} extracted from the mn th feature spaces are merged into one A_{ij} with the following parameters:

$$\begin{aligned} a_{ij} &= b_{ij} - \frac{b_{ij} - l_{ij}}{1 - \alpha} \\ c_{ij} &= b_{ij} + \frac{u_{ij} - b_{ij}}{1 - \alpha} \end{aligned} \quad (9)$$

where $b_{ij} = \frac{1}{n-1} \sum_{1 \leq m < n \leq N} b_{ij}^{mn}$, $l_{ij} = \frac{1}{n-1} \sum_{1 \leq m < n \leq N} l_{ij}^{mn}$, and $u_{ij} = \frac{1}{n-1} \sum_{1 \leq m < n \leq N} u_{ij}^{mn}$. There are a lot of methods to measure the similarity between two distinct fuzzy sets [13,14]. In this study, the set-theoretic operation-based similarity measurement [10,18]

$$S((\cdot)_a, (\cdot)_a) = \frac{|(\cdot)_a \cap (\cdot)_a|}{|(\cdot)_a \cup (\cdot)_a|}$$

is applied to removing the similar fuzzy sets, where $|\cdot|$ denotes the cardinality of a set. Therefore, if A_{ij} satisfies the following condition inequality, it is removed.

$$\sum_{i=1, i \neq j}^N S((A_{ij})_a, (A_{ij})_a) \geq \beta \quad (10)$$

where $\beta \in [0, 1]$ denotes a real constant value.

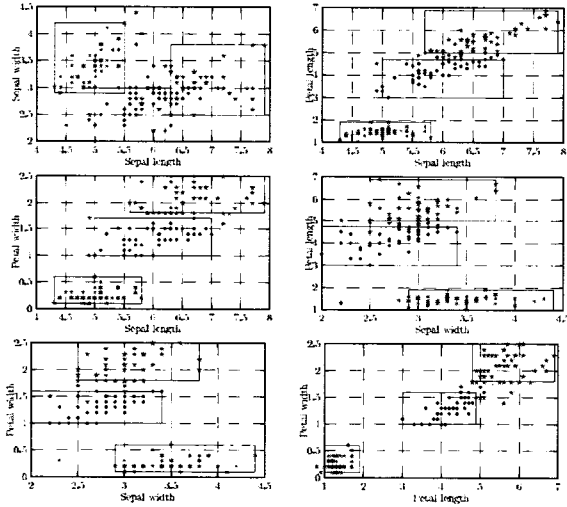


Fig. 3. Information granules for 150 training data.
*: Iris setosa; ◦: Iris versicolor; ★: Iris virginica.

3.3 GA-based Management of Misclassification

It is necessary to tune the fuzzy classifiers constructed in the previous subsection for the classification performance improvement on the data that lie into the overlaps of the classes, and the strange pattern unlike the well classified data such as outliers. This subsection proposes a GA-based management technique for misclassification, which

consists of the tuning process of the obtained fuzzy rules and the generation process of additional fuzzy rules. To decrease risk due to overfitting, the following constraints are imported:

$$\begin{aligned} A_{ij}(x_j) &= \begin{cases} \Delta A_{ij}(x_j; \Delta a_{ij}, b_{ij}, \Delta c_{ij}) & \text{if overlaps exist in } C^i \\ A_{ij}(x_j; a_{ij}, b_{ij}, c_{ij}) & \text{otherwise} \end{cases} \\ y_{ih} &= \begin{cases} \Delta y_{ih} & \text{if overlaps exist in } C^i \\ y_{ih} & \text{otherwise} \end{cases} \\ L' &\in [0, \bar{L}] \end{aligned} \quad (11)$$

where $\bar{L} \in Z^+$ and L' means the number of fuzzy rules to be newly generated, C^i denotes the i th class region, and $\Delta(\cdot)$ symbolizes the tuning of (\cdot) to be determined from the GA-based technique that will be discussed in this subsection. Considering (8) and (11), and tuning its cutting level α_{ij} for A_{ij} produce

$$A_{ij}(x_j) = \begin{cases} \frac{x_j - \Delta a_{ij}}{|b_{ij} - \Delta a_{ij}|} & \text{if } \Delta a_{ij} \leq x_j < b_{ij} \\ \frac{\Delta c_{ij} - x_j}{|\Delta c_{ij} - b_{ij}|} & \text{if } b_{ij} \leq x_j < \Delta c_{ij} \\ 0 & \text{otherwise} \end{cases} \quad (12)$$

where $\Delta a_{ij} = b_{ij} - \frac{b_{ij} - l_{ij}}{1 - \Delta \alpha_{ij}}$, $\Delta c_{ij} = b_{ij} + \frac{u_{ij} - b_{ij}}{1 - \Delta \alpha_{ij}}$.

Now, we can design the final fuzzy system by adding the following supplementary fuzzy rules to the fuzzy classifier (1).

$$\begin{aligned} R_i: & \text{ IF } x_1 \text{ is } A_{i1} \text{ and } \dots \text{ and } x_N \text{ is } A_{iN} \\ & \text{ THEN the class is misclassification.} \\ & (y_{i1} = \xi_{i1}, \dots, y_{iL} = \xi_{iL}) \end{aligned} \quad (13)$$

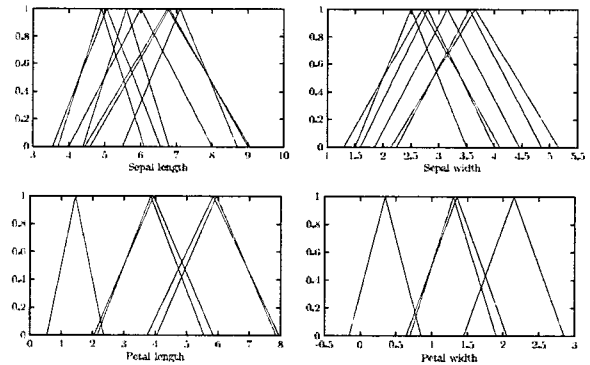


Fig. 4. Initial fuzzy sets extracted from information granules.

where $\xi_{ih} \in [-1, 1]$. By using product inference engine, singleton fuzzifier, and center average defuzzifier, the h th global output of the final fuzzy classifier (1) and (13) is represented as

$$y_h = \sum_{i=1}^I \theta_{A_i}(x) y_{ih} + \sum_{i=1}^{L'} \theta_{A_i}(x) y_{ih} \quad (14)$$

where,

$$\theta_{A_i}(x) = \frac{\prod_{j=1}^N A_{ij}(x_j)}{\sum_{i=1}^N \prod_{j=1}^N A_{ij}(x_j) + \sum_{i=1}^N \prod_{j=1}^N A_{i\bar{j}}(x_j)}$$

$$\theta_{A_i}(x) = \frac{\prod_{j=1}^N A_{i\bar{j}}(x_j)}{\sum_{i=1}^N \prod_{j=1}^N A_{ij}(x_j) + \sum_{i=1}^N \prod_{j=1}^N A_{i\bar{j}}(x_j)}$$

The predicted class \hat{y} is decided by the maximum of (14).

Naturally, the objective of (14) is that the classification performance should be as high as possible. Furthermore, it is desired to reduce the number of the additional fuzzy rules from the viewpoints of hardware implementation and computation resource. The following objective functions should be maximized and minimized, respectively:

$$\text{Maximize } J_5 = \frac{1}{D} \sum_{d=1}^D \delta(\hat{y}(x_d) - (\text{class of } x_d))$$

$$\text{Minimize } J_6 = L$$

where $\delta(\cdot)$ is the Dirac delta function. The fitness function to be maximized via the GA is then

$$f(J_5, J_6) = J_5 + \nu_2 \frac{1}{J_6 + \lambda}$$

where λ is the coefficient to prevent the large fitness function value.

4. An example: Iris data

The Iris data [15] is a common benchmark in the classification and the pattern recognition studies [8,16,17,19]. It has four continuous features: x_1 -sepal length, x_2 -sepal width, x_3 -petal length, and x_4 -petal width and consists of 150 instances: 50 for each class (Iris setosa, Iris versicolor, and Iris virginica). To examine the effectiveness of the proposed method, we provide two simulations for Iris data-Simulation 1: all 150 instances are selected as training data, and Simulation 2: one half of 150 instances are randomly selected as the training data and the other half are used as the test data.

The GA-based method is applied to search the optimal information granules in all $6 = {}_4C_2$ featurespaces. Figure 3 shows the information granules determined via the proposed scheme. From the parameters obtained by the information granules, the fuzzy sets are constructed by the proposed procedures: extracting, merging, and removing the fuzzy sets. Figure 4 shows the initial fuzzy sets extracted from the information granules. By means of the condition (10), the similar fuzzy sets are removed as shown in Fig. 5. Because all fuzzy sets for sepal

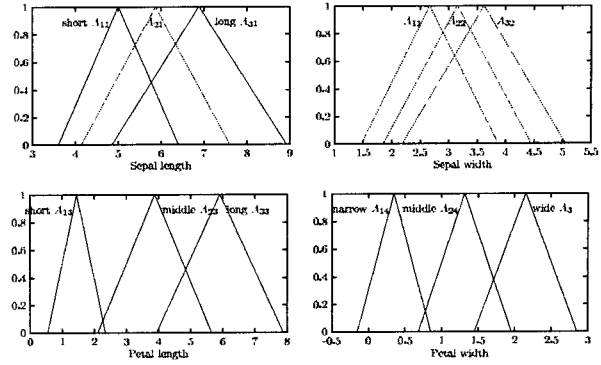


Fig. 5. Fuzzy sets merged from the initial fuzzy sets: the removed fuzzy sets (dotted line) had the final fuzzy sets (solid line).

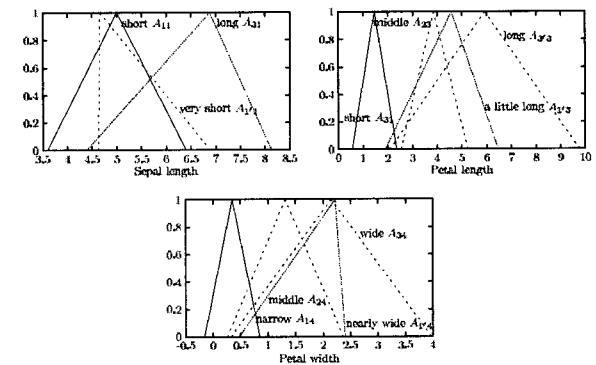


Fig. 6. Antecedent fuzzy sets in the fuzzy classifier (16): the given fuzzy sets (solid line), the tuned fuzzy sets (dash-dotted line), and the generated fuzzy sets(dotted line).

width are removed, three features (i.e., sepal length, petal length, and petal width) are selected for the pattern classification. A_{21} for the sepal length is also removed. Based on the constructed fuzzy sets in Fig. 5, the obtained fuzzy classifier is

- R_1 : IF x_1 is short and x_3 is short and x_4 is narrow THEN the class is Iris setosa ($y_{11}=1, y_{12}=0.5, y_{13}=0.5$)
- R_2 : IF x_3 is middle and x_4 is middle THEN the class is Iris versicolor ($y_{21}=0.5, y_{22}=1, y_{23}=0.5$)
- R_3 : IF x_1 is long and x_3 is long and x_4 is wide THEN the class is Iris virginica ($y_{31}=0.5, y_{32}=0.5, y_{33}=1$)

The GA-based method for the misclassification is applied to (15). Because there are the overlaps among the class region 2 and the class region 3, the antecedent fuzzy sets of R_2 and R_3 and y_{22} and y_{33} are tuned by using the GA method. As a result, the final fuzzy classifier is

- R_1 : IF x_1 is short and x_3 is short and x_4 is narrow
 THEN the class is *Iris setosa*
 ($y_{11}=1, y_{12}=0.5, y_{13}=0.5$)
 R_2 : IF x_3 is middle and x_4 is middle
 THEN the class is *Iris versicolor*
 ($y_{21}=0.5, y_{22}=0.6959, y_{23}=0.5$)
 R_3 : IF x_1 is long and x_3 is long and x_4 is wide (16)
 THEN the class is *Iris virginica*
 ($y_{31}=0.5, y_{32}=0.5, y_{33}=0.5267$)
 R_4 : IF x_1 is very short and x_3 is a little long
 and x_4 is nearly wide
 THEN the class is *Iris virginica*
 ($y_{41}=0.2373, y_{42}=0.6690, y_{43}=0.7894$)

Table 1. Classification results of Iris data (150 training data)

Ref.	# of rules	Accuracy rate(%)
[19]	4	98%
[16]	7	96%
ours	3	96.67%
	4	98.67%

Table 2. Classification results of Iris data (75 test data)

Ref.	# of rules	average accuracy rate(%)
[8]	5	92%
[17]	6.21	95.57%
ours	3.5	97.33%

where R_4 is generated for the management of misclassification. Its antecedent fuzzy sets are depicted in Fig. 6. The classification performance of the fuzzy classifier (15) is the 96.67% accuracy rate with 5 misclassifications on 150 training data. On the other hand, by using the GA-based management method for the misclassification, the classification performance of the final fuzzy classifier (16) is improved to the classification performance 98.67% with 2 misclassifications on 150 training data. In Table 1, the simulation result is compared with the other methods in terms of the rule number and the classification performance. Moreover, we tried the other simulation for the Iris data: one half of 150 instances are randomly selected as the training data and the other half are used as the test data and vice versa. The average accuracy rate of the final fuzzy classifier is 99.33% on the training data and 97.33% on the test data. Finally we examine the average performance of two simulations: training with one half of the given data and testing with the other half, vice versa. The result is in Table 2, which also dictates our approach is superior to the compared methods.

5. Conclusions

In this paper, the GA-based method for constructing the fuzzy classifier has been proposed. The advantages of the proposed method are threefold: First, although the number of the information granules equals to the number of the classes, the information granules developed by the GA accomplish the satisfactory fuzzy region. Second, the procedure of constructing the fuzzy sets from the information granules provides an effective tool for the feature selection and the pattern classification. Finally, as we additionally generate the fuzzy rules for the misclassification management, the final fuzzy classifier can describe the misclassified data. Therefore, the proposed method provides the selection of the information granules as well as the solution to the two major problems: the feature selection and the pattern classification. The simulation results have highly visualized that the proposed method has the effectiveness for the classification.

References

- [1] Y. H. Joo, H. S. Hwang, K. B. Kim, and K. B. Woo, "Linguistic model identification for fuzzy system," *Electron. Letter*, vol. 31, no. 4, pp. 330-331, 1995.
- [2] Y. H. Joo, H. S. Hwang, K. B. Kim, and K. B. Woo, "Fuzzy system modeling by fuzzy partition and GA hybrid schemes," *Fuzzy Sets Syst.*, vol. 86, no. 3, pp. 279-288, 1997.
- [3] L. A. Zadeh, "Fuzzy sets," *Informat. Control*, vol. 8, pp. 338-353, 1965.
- [4] S. Abe and R. Thawonmas, "A fuzzy classifier with ellipsoidal regions," *IEEE Trans. Fuzzy Systems*, vol. 5, no. 3, pp. 358-368, 1997.
- [5] L. X. Wang and J. M. Mendel, "Generating fuzzy rules by learning from examples," *IEEE Trans. Syst. Man, Cybern. B.*, vol. 22, no. 6, pp. 1414-1427, 1992.
- [6] H. Ishibuchi, K. Nozaki, N. Yamamoto, and H. Tanaka, "Selecting fuzzy if-then rules for classification problems using genetic algorithms," *IEEE Trans. Fuzzy Systems*, vol. 3, no. 3, pp. 260-270, 1995.
- [7] R. Thawonmas and S. Abe, "A novel approach to feature selection based on analysis of class regions," *IEEE Trans. Syst. Man, Cybern. B.*, vol. 27, no. 2, pp. 196-207, 1997.
- [8] S. Abe and M. S. Lan, "A method for fuzzy rules extraction directly from numerical data and its application to pattern classification," *IEEE Trans. Fuzzy Systems*, vol. 5, no. 1, pp. 358-368, 1995.
- [9] H. M. Lee, C. M. Chen, J. M. Chen, and Y. L. Jou, "An efficient fuzzy classifier with feature

저 자 소 개

- selection based on fuzzy entropy," *IEEE Trans. Syst. Man, Cybern. B.*, vol. 3, no. 3, pp. 426-432, 1997.
- [10] T. P. Wu and S. M. Chen, "A new method for constructing membership functions and fuzzy rules from training examples," *IEEE Trans. Syst. Man, Cybern. B.*, vol. 29, no. 1, pp. 25-40, 1999.
- [11] D. E. Goldberg, *Genetic algorithms in search, optimization, and machine learning*. Addison-Wesley publishing company, Inc., 1989.
- [12] W. Pedrycz and A. Bargiela, "Granular clustering: a granular signature of data," *IEEE Trans. Syst. Man, Cybern. B.*, vol. 32, no. 2, pp. 212-224, 2002.
- [13] S. M. Chen, M. S. Yeh, and P. Y. Hsiao, "A comparison of similarity measures of fuzzy values," *Fuzzy Sets Syst.*, vol. 72, no. 1, pp. 79-89, 1995.
- [14] V. R. Young, "Fuzzy subethood," *Fuzzy Sets Syst.*, vol. 77, no. 3, pp. 371-384, 1996.
- [15] R. A. Fisher, "The use of multiple measurements in taxonomic problems," *Ann. Eugenics.*, vol. 7, no. 2, pp. 179-188, 1936.
- [16] S. Halgamuge and M. Glesner, "Neural networks in designing fuzzy systems for real world applications," *Fuzzy Sets Syst.*, vol. 65, pp. 1-12, 1994.
- [17] T. P. Hong and C. Y. Lee, "Induction of fuzzy rules and membership functions from training examples," *Fuzzy Sets Syst.*, vol. 84, no. 3, pp. 33-47, 1996.
- [18] H. Roubos and M. Setnes "Compact transparent fuzzy models and classifiers through iterative complexity reduction," *IEEE Trans. Fuzzy Systems*, vol. 9, no. 4, pp. 516-524, 2001.
- [19] Y. Shi, R. Eberhart, and Y. Chen, "Implementation of evolutionary fuzzy systems," *IEEE Trans. Fuzzy Systems*, vol. 7, no. 2, pp. 109-119, 1999.



김도완(Do Wan Kim)

제 13권 2호(2003년 4월호) 참조



주영훈(Young Hoon Joo)

제 13권 2호(2003년 4월호) 참조



박진배(Jin Bae Park)

제 13권 2호(2003년 4월호) 참조