Pergamon

Computers ind. Engng Vol. 35, Nos 3–4, pp. 575–578, 1998

Printed in Great Britain
PII: S0360-8352(98)00162-4
0360-8352/98 $19.00 + 0.00

Performance Evaluation of Fuzzy Rule-Based Classification Systems Obtained by Multi-Objective Genetic Algorithms

Hisao Ishibuchi*, Tadahiko Murata**, and Mitsuo Gen**

*Department of Industrial Engineering, Osaka Prefecture University
Gakuen-cho 1-1, Sakai, Osaka 599-8531, Japan
**Department of Industrial and Systems Engineering, Ashikaga Institute of Technology
Omae-cho 268, Ashikaga, Tochigi 326-8558, Japan
hisaoi@ie.osakafu-u.ac.jp, murata@genlab.ashitech.ac.jp, gen@genlab.ashitech.ac.jp

ABSTRACT

In this paper, we examine the classification performance of fuzzy if-then rules selected by a GA-based multi-objective rule selection method. This rule selection method can be applied to high-dimensional pattern classification problems with many continuous attributes by restricting the number of antecedent conditions of each candidate fuzzy if-then rule. As candidate rules, we only use fuzzy if-then rules with a small number of antecedent conditions. Thus it is easy for human users to understand each rule selected by our method. Our rule selection method has two objectives: to minimize the number of selected fuzzy if-then rules and to maximize the number of correctly classified patterns. In our multi-objective fuzzy rule selection problem, there exist several solutions (*i.e.*, several rule sets) called "non-dominated solutions" because two conflicting objectives are considered. In this paper, we examine the performance of our GA-based rule selection method by computer simulations on a real-world pattern classification problem with many continuous attributes. First we examine the classification performance of our method for training patterns by computer simulations. Next we examine the generalization ability for test patterns. We show that a fuzzy rule-based classification system with an appropriate number of rules has high generalization ability.

KEYWORDS

Fuzzy rule-based system; pattern classification; rule selection; genetic algorithms; knowledge acquisition.

1. INTRODUCTION

Fuzzy rule-based systems were mainly applied to non-linear control problems with a few input variables (see, for example, Sugeno [8]; Lee [2]). For automatically generating fuzzy if-then rules for pattern classification problems, Ishibuchi *et al.* [3] proposed a heuristic rule generation procedure that was based on a simple fuzzy grid. Ishibuchi *et al.* [4] proposed a GA-based fuzzy rule selection method, where a large number of candidate fuzzy if-then rules were first generated, then a GA was employed for selecting a small number of fuzzy if-then rules to construct a compact fuzzy classification system.

In this paper, we examine the performance of a GA-based multi-objective fuzzy rule selection method (Ishibuchi *et al.* [6]). This method can be applied to high-dimensional pattern classification problems with many continuous attributes after slightly modifying its rule generation procedure as in Ishibuchi and Murata [5]. Our rule selection method has two objectives: to minimize the number of selected fuzzy if-then rules and to maximize the number of correctly classified patterns. Because the number of candidate fuzzy if-then rules in the rule selection method exponentially increases as the number of attributes increases, the rule selection method can not handle all the possible fuzzy if-then rules as candidate rules when it is applied to high-dimensional pattern classification problems with many attributes. When we employ the four linguistic values for each axis of the n-dimensional pattern space (see Fig.1), the total number of possible fuzzy if-then rules is 4^n. The relation between the number of attributes (*i.e.*, n) and the number of fuzzy if-then rules (*i.e.*, 4^n) are shown in Table 1. As shown in Table 1, the total number of possible fuzzy if-then rules is too large to be handled by the GA-based rule selection method. Thus we generate only a part of the 4^n fuzzy if-then rules as candidate rules. For this purpose, only fuzzy if-then rules with a small number of antecedent conditions were generated as candidate rules in Ishibuchi and Murata [5]. Let us define the length of a fuzzy if-then rule by the number of its antecedent fuzzy sets except for "*don't care.*" For example, the length of the following fuzzy if-then rule is two.

If x_1 is small & x_2 is don't care & x_3 is don't care & x_4 is don't care & x_5 is large then Class 1, (1)

where there are no conditions on $x_2 \sim x_4$. This fuzzy if-then rule can be rewritten as follows by removing the attributes with "*don't care*":

If x_1 is small & x_5 is large then Class 1. (2)

In Ishibuchi and Murata [5], they reduced the number of candidate fuzzy if-then rules in the GA-based multi-objective rule selection method by the constraint on the length of fuzzy if-then rules. For example, let us consider a 13-dimensional pattern classification problem. As shown in Table 1, the total number of fuzzy if-then rules in this problem is $4^{13} \approx 6.7 \times 10^7$. In Table 2, we show the number of fuzzy if-then rules according to their length. From this table, we can see that the number of candidate fuzzy if-then rules is not large if we only generate fuzzy if-then rules whose length is less than or equal to two. Because we only use fuzzy if-then rules with a small number of antecedent conditions, it is easy to understand fuzzy if-then rules obtained by our method.

In this paper, we examine the performance of our GA-based multi-objective fuzzy rule selection method by computer simulations on wine data (Forina *et al.* [1]). First we examine the classification performance of our method on training patterns. Next we examine the generalization ability for test patterns.

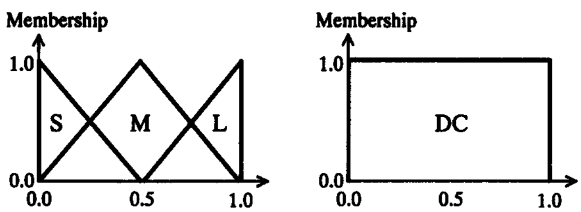

Fig. 1 Four linguistic values for each axis (S: *small*, M: *medium*, L: *large*, and DC: *don't care*).

Table 1 The number of fuzzy if-then rules.

Number of attributes	Number of rules
2	16
4	256
6	4,096
8	65,536
10	1,048,576
13	67,108,863
15	1,073,741,824
20	1,099,511,627,776

Table 2 The number of fuzzy if-then rules.

Length of rules	Number of rules
0	1
1	39
2	702
3	7,722
:	:
12	6,908,733
13	1,594,323
0 ~ 13	67,108,863

2. RULE SELECTION PROBLEM

In general, the performance of a rule set S is high when S has a large number of fuzzy if-then rules. On the other hand, a compact rule set with a small number of fuzzy if-then rules has some advantages. For example, it does not require a lot of storage, and its inference speed for new patterns is high. Based on these discussions, the following two-objective rule selection problem was formulated (Ishibuchi *et al.*, [6]):

[Rule Selection Problem]

Maximize $NCP(S)$ and minimize $|S|$, (3)

where $NCP(S)$ is the number of correctly classified patterns by the rule set S, and $|S|$ is the number of fuzzy if-then rules in S. The performance of the rule set S and its compactness are measured by $NCP(S)$ and $|S|$ in the rule selection problem, respectively. We employ the multi-objective genetic algorithm (MOGA in Murata and Ishibuchi [7]) for selecting fuzzy if-then rules from a large number of candidate rules. By the MOGA, several solutions (*i.e.*, several rule sets) called "non-dominated solutions" are obtained because two conflicting objectives are considered. The classification ability of a compact rule set with a small number of fuzzy if-then rules tends to be inferior to that of a large rule set with many rules. On the other hand, the comprehensibility of a large rule set is not good because human users can not examine each of many rules (*e.g.*, hundreds of fuzzy if-then rules). The decision maker should select a rule set from non-dominated solutions with respect to the compactness and the classification performance.

3. SIMULATION RESULTS

We examine the ability of our GA-based rule selection method by computer simulations on wine data (Forina *et al.* [1]). The wine data consist of 178 samples with 13 continuous attributes from three classes. First we examine the classification performance of our method for training patterns. Next we examine the generalization ability for test patterns.

3.1 Performance for Training Patterns

In our rule selection method, we first generated candidate fuzzy if-then rules whose length was less than or equal to two. The number of the generated fuzzy if-then rules was ${}_{13}C_0 + {}_{13}C_1 \times 3 + {}_{13}C_2 \times 3 \times 3 = 742$. Those fuzzy if-then rules were used as candidate rules in our rule selection method. The MOGA was applied to those candidate rules for finding non-dominated rule sets of the rule selection problem. We used the following parameter specifications in the MOGA:

Population size: 20, **Number of added elite solutions**: 3, **Crossover probability**: 1.0,
Mutation probability: 0.1 for the mutation from 1 to 0,
0.001 for the mutation from 0 to 1,
Stopping condition: 1000 generations.

By the MOGA, eight non-dominated solutions in Table 3 were obtained. For example, the MOGA selected the following seven fuzzy if-then rules that can correctly classify all the training patterns:

If x_{13} is *large* then Class 1 with $CF = 1.00$,
If x_7 is *medium* and x_{13} is *large* then Class 1 with $CF = 0.46$,
If x_{10} is *small* then Class 2 with $CF = 0.51$,
If x_2 is *small* and x_{13} is *small* then Class 2 with $CF = 0.77$,
If x_4 is *large* and x_{10} is *small* then Class 2 with $CF = 0.76$,
If x_7 is *small* and x_{10} is *medium* then Class 3 with $CF = 0.69$,
If x_{11} is *small* and x_{12} is *small* then Class 3 with $CF = 0.93$.

Table 3 Non-dominated solutions obtained by the MOGA for training patterns.

$\|S\|$	0	1	2	3	4	5	6	7
$NCP(S)$	0	71	125	171	174	175	177	178
Rate (%)	0.0	39.9	70.2	96.1	97.8	98.3	99.4	100

Table 4 Generalization ability of non-dominated solutions obtained by the MOGA for test patterns.

$\|S\|$	0	1	2	3	4	5	6	7	8
$NCP(S)$	0	71	126	171	173	173	122	88	35
Rate (%)	0.0	39.9	70.8	96.1	97.2	97.2	98.4	100	100

The consequent class of each fuzzy if-then rule and *CF* (the grade of certainty of each rule) were determined by a heuristic rule generation method in Ishibuchi *et al.* [3]. Because each fuzzy if-then rule has a small number of antecedent conditions, it is easy for human users to understand the fuzzy rule-based classification system.

3.2 Performance for Test Patterns

In order to examine the generalization ability for test patterns, we employed the 10-fold cross validation (10-CV) method. Table 4 shows the simulation results. Note that classification systems with six, seven and eight fuzzy if-then rules were obtained seven, five and two times among the ten trials in the 10-CV method. The other classification systems were obtained in all the ten trials. From Table 4, we can see that fuzzy rule-based classification systems with an appropriate number of rules have high generalization ability.

4. CONCLUSION

In this paper, we examined the performance of a GA-based multi-objective fuzzy rule selection method (Ishibuchi *et al.* [6]) by computer simulations on wine data with 13 continuous attributes. In the computer simulations, we only used fuzzy if-then rules with two or less antecedent conditions as candidate rules. While we used such simple fuzzy if-then rules, a small number of selected fuzzy if-then rules had high classification ability for test patters as well as training patterns.

REFERENCES

[1] Forina, M., *et al.* Wine recognition database, available via anonymous ftp from *ics.uci.edu* in directory */pub/machine-learning-databases/wine* (1992).

[2] Lee, C.C. Fuzzy logic in control systems: fuzzy logic controller. *IEEE Trans. on SMC*, **20**, 404-435 (1990).

[3] Ishibuchi, H., K.Nozaki and H.Tanaka. Distributed representation of fuzzy rules and its application to pattern classification. *Fuzzy Sets and Systems*, **52**, 21-32 (1992).

[4] Ishibuchi, H., K.Nozaki, N.Yamamoto and H.Tanaka. Selecting fuzzy if-then rules for classification problems using genetic algorithms. *IEEE Trans. on Fuzzy Systems*, **3**, 3, 260-270 (1995).

[5] Ishibuchi, H. and T.Murata. Minimizing the fuzzy rule base and its performance by a multi-objective genetic algorithm. *Proc. 6th FUZZ-IEEE*, 259-264 (1997).

[6] Ishibuchi, H., T.Murata and I.B.Turksen. Single-objective and two-objective genetic algorithms for selecting linguistic rules for pattern classification problems, *Fuzzy Sets and Systems*, **89**, 2, 135-150 (1997).

[7] Murata, T. and H.Ishibuchi. MOGA: Multi-objective genetic algorithms. *Proc. 2nd ICEC*, 289-294 (1995).

[8] Sugeno, M. An introductory survey of fuzzy control. *Information Science*, **36**, 59-83 (1985).