

## 통계적 정보기반 계층적 퍼지-러프 분류기법

# Statistical Information-Based Hierarchical Fuzzy-Rough Classification Approach

손창식<sup>a</sup> · 서석태<sup>b</sup> · 정환목<sup>a</sup> · 권순학<sup>b\*</sup>

Chang S. Son<sup>a</sup>, Suk T. Seo<sup>b</sup>, Hwan M. Chung<sup>a</sup>, Soon H. Kwon<sup>b\*</sup>

<sup>a</sup> 대구가톨릭대학교 컴퓨터정보통신공학부

<sup>b</sup> 영남대학교 전기공학과

### 요 약

본 논문에서는 학습기법을 사용하지 않고 패턴분류의 성능을 최대화하면서 규칙의 수를 줄일 수 있는 통계적 정보기반 계층적 퍼지-러프 분류방법을 제안한다. 제안된 방법에서 통계적 정보는 계층적 퍼지-러프 분류 시스템에서 각 계층의 입력부 퍼지집합의 분할 구간을 추출하기 위해서 사용되었고, 러프집합은 통계적 정보로부터 추출된 분할 구간들과 연관된 퍼지 if-then 규칙의 수를 최소화하기 위해서 사용되었다. 제안된 방법의 효과성을 보이기 위해 Fisher의 IRIS 데이터를 사용한 기존 패턴분류 방법의 분류 정확도와 규칙들의 수를 비교하였다. 그 결과, 제안된 방법은 기존 방법들의 분류 성능과 유사함을 확인할 수 있었다.

키워드 : 패턴분류, 계층적 퍼지 시스템, 분할 구간 선택, 규칙 감축

### Abstract

In this paper, we propose a hierarchical fuzzy-rough classification method based on statistical information for maximizing the performance of pattern classification and reducing the number of rules without learning approaches such as neural network, genetic algorithm. In the proposed method, statistical information is used for extracting the partition intervals of antecedent fuzzy sets at each layer on hierarchical fuzzy-rough classification systems and rough sets are used for minimizing the number of fuzzy if-then rules which are associated with the partition intervals extracted by statistical information. To show the effectiveness of the proposed method, we compared the classification results(e.g. the classification accuracy and the number of rules) of the proposed with those of the conventional methods on the Fisher's IRIS data. From the experimental results, we can confirm the fact that the proposed method considers only statistical information of the given data is similar to the classification performance of the conventional methods.

Key Words : Pattern Classification, Hierarchical Fuzzy System, Partition Interval Selection, Rule Reduction

## 1. 서 론

퍼지 규칙기반 분류 시스템에서 패턴분류의 성능은 퍼지 집합이 정의되는 특징공간상에 분할 구간에 의해서 좌우되고, 규칙의 수는 입력변수들과 퍼지집합들의 수가 증가할 때 지수적으로 증가하게 된다. 그러므로 주어진 입력패턴들로부터 적절한 퍼지분할 구간을 획득하고 그 구간과 연관된 최적의 규칙을 추출하기 위한 다양한 학습방법들이 제안되었다 [1-5]. 그러나 이들 방법들은 패턴분류를 위한 규칙의 수를 최적화하는 데에는 아주 효과적인 방법이지만, 퍼지분할 구간을 결정하는데 있어서 고려되어야 할 요소(즉 학습 파라미

터)들이 많다는 제약점을 가진다. 또한 초기에 정해진 퍼지 집합들을 지속적으로 조정(tuning)하여야 하기 때문에, 최적의 퍼지분할 구간을 획득하는데 다소 많은 시간이 소비된다.

그래서 이러한 제약점을 해결하기 위한 하나의 방법으로 퍼지분할을 계층적으로 분할하는 방법이 제안되었다[9]. 그러나 이 방법 또한 최적의 퍼지분할 구간을 선택하기 위해 정사각형 형태의 고정된 그리드 분할 공간만을 사용하기 때문에 복잡한 비선형 패턴들을 가진 데이터들을 분류하는 것이 다소 어렵다는 제약점을 가진다.

따라서 본 논문에서는 이러한 제약점들을 보다 단순한 방법으로 개선시킬 수 있는 통계적 정보기반 계층적 퍼지-러프 분류기법을 제안한다. 제안된 방법에서 통계적 정보는 계층적 퍼지-러프 분류 시스템에서 각 계층의 입력부 퍼지집합의 분할 구간을 추출하기 위해서 사용되었고, 러프집합은 통계적 정보로부터 추출된 분할 구간에 대응하는 퍼지 if-then 규칙들의 수를 최소화하기 위해서 사용되었다. 그리

접수일자 : 2007년 8월 24일

완료일자 : 2007년 10월 24일

\* Corresponding Author

고 제안된 방법의 타당성을 검증하기 위해 Fisher[8]의 IRIS 데이터를 기반으로 기존의 학습기반 분류방법과 분류 결과들(즉 분류 정확성과 규칙 수)을 비교하였다.

본 논문의 구성은 다음과 같다. 2장에서는 본 논문에서 제안하는 계층적 퍼지-러프 분류기법의 구조와 입력부 퍼지집합의 분할구간의 결정방법에 대해서 논의하고, 패턴분류를 위한 규칙추출과 규칙감축 방법에 대해서 설명한다. 3장에서는 제안된 방법의 타당성을 검증하기 위해서 Fisher의 IRIS 데이터를 기반으로 여러 패턴분류방법과 비교하였고, 4장에서는 결론을 맺는다.

## 2. 계층적 퍼지-러프 분류기법

### 2.1 계층적 퍼지-러프 분류시스템의 구조

다음 그림들은 각각  $k$ 개의 클래스로 레이블된 데이터를 분류하기 위한 계층적 퍼지-러프 분류시스템의 구조와 각 퍼지-러프 분류시스템의 하부 구조를 나타낸다.

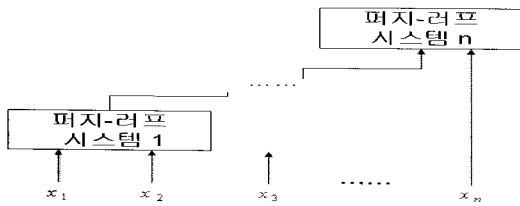


그림 1. 계층적 퍼지-러프 분류시스템의 구조

Fig. 1. The structure of a hierarchical fuzzy-rough classification system

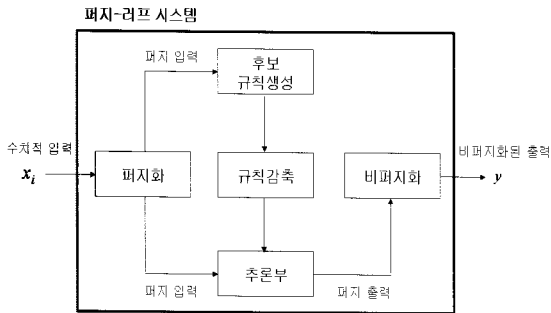


그림 2. 계층적 퍼지-러프 분류시스템의 하부 구조

Fig. 2. Substructure of a hierarchical fuzzy-rough classification system

그림 1은 레이블된 클래스를 가진  $n$ 개의 입력속성,  $x_1, x_2, \dots, x_n$ 에 대한 계층적 퍼지-러프 분류시스템의 구조를 나타내고, 그림 2는 각 계층에서 퍼지-러프 분류시스템의 하부 구조를 나타낸다. 각 계층의 하부 구조는 퍼지화, 후보규칙생성, 규칙감축, 추론부, 그리고 비퍼지화와 같은 5개의 모듈로 구성되어 있으며, 그 모듈들의 기능은 다음과 같다.

- (1) 퍼지화(fuzzyfier)  
주어진 수치적 입력패턴(crisp input pattern)들을 퍼지 입력으로 변환한다.
- (2) 후보규칙생성(generation of candidate rules)

초기에 정의된 각 입력부 퍼지집합의 분할 구간들에 대응하는 규칙만을 추출하여 후보규칙을 구성한다. 제안한 방법에서는 후보규칙들을 추출하기 위해 각 입력 패턴들에 대해서 최대 적합도(maximum compatibility grade)를 가진 언어항(linguistic terms)을 선택하였다.

- (3) 규칙감축(rule reduction)  
후보규칙생성 모듈로부터 추출된 규칙들의 불필요한 속성을 제거함으로써, 규칙을 감축함과 동시에 패턴분류를 위해 반드시 필요한 규칙만을 재추출한다.
- (4) 추론부(Inference)  
규칙감축 모듈로부터 재추출된 규칙들을 기반으로 각 입력패턴의 출력부 클래스를 결정한다.
- (5) 비퍼지화(defuzzification)  
결정된 출력부 클래스들의 비퍼지화된 실수값들을 계산한다. 제안한 방법에서는 실수값들을 얻기 위해 무게 중심법(center of gravity method)을 사용하였다.

### 2.2 퍼지집합의 분할구간의 결정

다음은 패턴분류를 위해 주어진 데이터의 통계적 정보만을 고려하여 입력부 퍼지집합에서 퍼지분할의 수와 그들의 구간을 결정하기 위한 방법을 나타낸다[12].

$$P = \{p_1, p_2, \dots, p_i, \dots, p_n\} \quad (1)$$

$$p_i = \left( \frac{d_i^{\max} - d_i^{\min}}{d_i^{sd}} \right) + 2, \quad d_i \in [\alpha_i^{\min}, \alpha_i^{\max}] \quad (2)$$

$$(\alpha_i^{\min} = d_i^{\min} - d_i^{sd}, \alpha_i^{\max} = d_i^{\max} + d_i^{sd})$$

여기서  $P$ 는 전체 속성에 대한 퍼지분할 정보를 나타내고,  $d_i^{\min}, d_i^{\max}, d_i^{sd}$ 는 그 데이터에서  $i$ 번째 속성의 최소값, 최대값, 그리고 표준편차를 의미한다. 또한  $p_i (i=1, 2, \dots, n)$ 는 전체  $n$ 개의 속성 중에서  $i$ 번째 속성의 퍼지분할의 수를 나타낸다. 즉  $(d_i^{\max} - d_i^{\min})/d_i^{sd}$ 는 각 클래스가 포함하는 최소-최대 구간에서 서로 중첩되는 구간들(overlapped intervals)의 분할수를 의미하고, '2'는 그들을 제외한 나머지 구간에서의 분할수를 나타낸다. 예를 들어, 임의의 한 속성에서 각 클래스간의 서로 중첩되는 구간이 식 (2)에 의해서 3개의 퍼지분할수를 가진다면, 그 부분을 제외한 나머지 양끝 구간에서 각각 1개의 퍼지분할수를 가짐으로, 전체 퍼지분할의 수는 5개가 된다. 식 (2)에서  $\alpha_i^{\min} = d_i^{\min} - d_i^{sd}$ 와  $\alpha_i^{\max} = d_i^{\max} + d_i^{sd}$ 는 새로운 입력 데이터가 들어올 경우, 매핑을 위한 추가적인 공간을 확보하기 위해 표준편차만큼 구간을 확장한 것을 의미한다. 그리고 식 (2)로부터 획득한 퍼지분할의 구간들은 모두 동일한 폭(width)의 크기를 가진다.

### 2.3 후보규칙 추출 및 선택

다음과 같은 퍼지 IF-THEN 규칙이 있다고 가정하자.

$$\text{Rule}^i : \text{IF } x_{i1} \text{ is } A_{i1} \text{ and } \dots \text{ and } x_{in} \text{ is } A_{in} \\ \text{THEN } y_i \text{ is } C_j$$

여기서  $x_{i1}, \dots, x_{in}$ 과  $y_i$ 는 각각 입력변수(antecedent variable)와 출력변수(consequent variable)를 의미하고,  $A_{i1}, \dots, A_{in}$ 과  $C_j (j=1, 2, \dots, k)$ 는 각각 입력부 퍼지집합과 출력부 클래스를 나타낸다.

이때  $m$ 개의 수치적인 입력패턴이 주어졌을 때 후보규칙

의 추출은 다음과 같다.

$$CR_l = \max(\min(\mu_{A_{l1}}(\alpha), \mu_{A_{l2}}(\beta), \dots, \mu_{A_{lm}}(w))) \quad (3)$$

( $l=1,2,\dots,m$ )

여기서  $\mu_{A_{l1}}(\alpha), \mu_{A_{l2}}(\beta), \mu_{A_{lm}}(w)$ 는 각 입력부 퍼지집합에 대응하는 소속함수의 적합도(compatibility grade)를 나타낸다. 즉 각 입력패턴에 대해서 최대 적합도를 가지는 퍼지 집합의 레이블(label)만을 추출하여 규칙을 구성한다.

다음은 식 (3)으로부터 얻은 후보규칙과 적합도를 이용하여 각 클래스와 연관된 규칙을 선택하기 위한 기준(criterion)이다[11].

- 커플링 규칙이 발생하지 않는 경우

추출된 후보규칙의 빈도와 적합도에 관계없이, 각 클래스와 연관된 후보규칙은 해당 클래스의 규칙으로 선택한다.

- 커플링 규칙이 발생하는 경우

$$f_{C_s} = \sum_{i=1}^n CR_i \quad (4)$$

$$f_{C_t} = \sum_{i=1}^n CR_i \quad (5)$$

여기서  $f_{C_s}, f_{C_t}$ 는 각각 전체  $n$ 개의 규칙 중에서  $s$ 와  $t$ 번째 클래스에 대한 후보규칙의 빈도수를 나타낸다.

$$CR_i = \begin{cases} C_s, & f_{C_s} > f_{C_t} \\ C_t, & f_{C_s} < f_{C_t} \end{cases} \quad (6)$$

식 (6)에서  $f_{C_s} > f_{C_t}, f_{C_s} < f_{C_t}$ 는 각각  $s$ 번째 후보규칙의 빈도수가  $t$ 번째 후보규칙의 빈도수보다 큰 경우  $s$ 번째 클래스의 규칙으로 선택하고, 그렇지 않으면  $t$ 번째 클래스의 규칙으로 선택함을 의미한다.

$$CR_i = \begin{cases} C_s, & \max_i(\mu_{C_s}) > \max_i(\mu_{C_t}), \text{ if } f_{C_s} = f_{C_t} \\ C_t, & \max_i(\mu_{C_s}) < \max_i(\mu_{C_t}), \text{ if } f_{C_s} = f_{C_t} \end{cases} \quad (7)$$

식(7)에서는 만약  $s$ 번째 후보규칙의 빈도수와  $t$ 번째 후보규칙의 빈도수가 같다면, 규칙은  $s$ 번째 후보규칙과 대응하는 클래스의 최대-적합도와  $t$ 번째 후보규칙과 대응하는 클래스의 최대-적합도를 비교하여 더 큰 적합도를 가지는 후보규칙의 클래스로 분류하는 것을 나타낸다.

### 2.4 속성감축

2.3절에서 얻은 규칙은 각 입력변수들 간의 상관성만을 고려하여 추출된 규칙이므로, 여전히 그 규칙 내에 불필요한 속성들을 포함하고 있을 수 있다. 그러므로 서로 다른 클래스의 규칙들 간에 상관성을 고려하여 불필요한 속성을 제거하기 위해 러프집합의 코어와 리덕트를 보다 단순한 방법으로 추출할 수 있는 Skowron[7]의 식별가능행렬(discernibility matrix)을 이용하였다.

$$(c_{ij}) = \{a \in A : a(x_i) \neq a(x_j)\}, \text{ for } i, j = 1, 2, \dots, n \quad (8)$$

여기서  $c_{ij}$ 는 속성  $x_i, x_j$ 을 구별하게 하는 모든 속성들의 집합(즉,  $n \times n$  식별가능행렬)을 의미한다. 따라서 식별가능행렬에서 코어는 주어진 속성들 중에서 단일 원소로 이루어진 속성들의 집합이므로 다음과 같이 정의될 수 있다.

$$\text{core}(A) = \{a \in A : c_{ij} = (a), \text{ for some } i, j\} \quad (9)$$

식 (9)로부터 리덕트를 계산하기 위하여 다음과 같이 식별가능 함수  $f(A)$ 을 계산할 수 있다.

$$f(A) = \prod_{(x,y) \in U^2} \{\sum \delta(x,y) : (x,y) \in U^2 \text{ and } \delta(x,y) \neq 0\} \quad (10)$$

식 (10)에 의해서 계산된 리덕트로부터 각 속성  $x$ 에 대한 최종 식별가능 함수  $f^x(A)$ 을 다신 정의할 수 있다.

$$f^x(A) = \prod_{y \in U} \{\delta(x,y) : y \in U \text{ and } \delta(x,y) \neq 0\} \quad (11)$$

이러한 과정으로 얻은 규칙은 1개 이상의 리덕트를 가진 규칙들을 모두 포함하기 때문에, 패턴분류를 위한 규칙의 수가 많다. 따라서 본 논문에서는 보다 적은 수의 규칙을 획득하기 위해서 식별가능행렬로부터 얻은 규칙 중에서 1개의 리덕트만을 가지는 규칙들을 최종 규칙으로 선택하였다.

### 2.5 추론결과로부터의 퍼지분할 구간의 결정

계층적 퍼지-러프 분류기법의 분류성능은 각 계층의 추론 결과에서 분할 구간이 어떻게 선택되느냐?에 따라 다음 계층의 분류 정확도가 결정된다. 그러므로 본 절에서는 추론결과와 퍼지분할 구간을 결정할 수 있는 방법에 대해서 설명한다. 다음은 각 계층의 퍼지-러프 분류시스템으로부터 생성된 결과들의 퍼지분할 구간의 결정 방법이다.

[Step 1] 각 클래스의 비퍼지화된 결과 중에서 오분류 구간(misclassified interval)들을 추출한다. 여기서 오분류 구간은 몇몇 입력패턴에 대해서 잘못 분류된 비퍼지화된 값들의 최소값과 최대값을 의미한다.

[Step 1-1] 만약 각 클래스의 오분류 구간들 간에 중첩 구간이 존재한다면, 그 구간을 퍼지분할 구간으로 결정하고, 식 (2)로부터 획득한 분할 수만큼 그 구간을 나누고 Step4로 이동한다.

[Step 1-2] 각 클래스의 오분류 구간들 간에 중첩 구간이 존재하지 않으면, Step2로 이동한다.

[Step 2] 각 클래스의 비퍼지화된 전체 구간들을 추출한다.

[Step 2-1] 만약 비퍼지화된 구간들 간에 중첩 구간이 존재한다면, 그 구간을 퍼지분할 구간으로 결정하고, 식 (2)로부터 획득한 분할 수만큼 그 구간을 나누고 Step4로 이동한다.

[Step 2-2] 각 클래스의 전체 비퍼지화된 구간들 간에 중첩 구간이 존재하지 않으면, Step3으로 이동한다.

[Step 3] 각 클래스에 대해서 비퍼지화된 결과의 구간  $[\alpha_i^{\min}, \alpha_i^{\max}]$ 에서 퍼지분할을 결정하고 Step4로 이동한다.

[Step 4] 퍼지분할 구간결정을 종료한다.

## 3. 실험결과 및 검토

실험에서는 패턴분류에서 널리 사용되는 Fisher[8]의 IRIS 데이터를 바탕으로 제안한 방법의 효과성을 보이고자 한다. Fisher의 IRIS 데이터의 각 클래스는 3개의 레이블(Setosa, Versicolor, Virginica)로 나뉜다. 또한 각 데이터는 4개의 속성을 가지고 있으며, 각 속성의 레이블은 Sepal Length(SL), Sepal Width(SW), Petal Length(PL), Petal Width(PW)으

로 구성되어 있다.

### 3.1 패턴분류를 위한 입력속성의 결합 순서결정

계층적 퍼지 시스템에서 각 입력속성의 결합 순서에 따라 조합 가능한 계층과 규칙의 수는 다양하다. 그러므로 본 논문에서는 각 계층에서 입력속성의 순서(order)를 결정하기 위해 IRIS 데이터의 각 속성의 상관관계(correlation)를 분석하여 상관성이 높은 속성들을 상위계층으로 할당하였다. 다음은 IRIS 데이터의 상관관계를 분석한 결과를 나타낸다.

표 1. IRIS 데이터의 상관관계  
Table 1. Correlation of IRIS data

상관관계	SL	SW	PL	PW
SL	1	-0.118	0.872	0.8118
SW	-0.118	1	-0.428	-0.366
PL	0.872	-0.428	1	0.963
PW	0.8118	-0.366	0.963	1

표 1로부터 상관관계가 높은 속성들을 추출하여 계층을 구성하면 다음과 같다.

- 계층 1 : Petal Length와 Petal Width
- 계층 2 : 계층 1의 추론결과와 Sepal Length
- 계층 3 : 계층 2의 추론결과와 Sepal Width

### 3.2 입력부 퍼지집합의 퍼지분할 구간

다음은 입력부 퍼지집합의 퍼지분할 구간을 결정하기 위한 IRIS 데이터의 통계적인 정보이다.

표 2. IRIS 데이터의 통계적 정보  
Table 2. Statistical information on IRIS data

속성	Setosa		Versicolor		Virginica		SD
	Min	Max	Min	Max	Min	Max	
SL	4.3	5.8	4.9	7.0	4.9	7.9	0.828
SW	2.3	4.4	2.0	3.4	2.2	3.8	0.436
PL	1.0	1.9	3.0	5.1	4.5	6.9	1.762
PW	0.1	0.6	1.0	1.8	1.4	2.5	0.765

표 2에서 Min과 Max는 주어진 데이터에서 각 속성에 대해서 레이블된 클래스들의 최소값과 최대값을 나타내고, SD는 각 속성의 표준편차를 나타낸다. 표 2의 IRIS 데이터 통계적 정보와 식 (2)로부터 각 속성의 퍼지분할 수를 계산하면 다음과 같다.

속성	Sepal		Petal	
	Length	Width	Length	Width
a	[3.472, 8.728]	[1.564, 4.836]	[0, 8.662]	[0, 3.265]
b	[4.9, 7.0]	[2.2, 3.8]	[4.5, 5.1]	[1.4, 1.8]
c	4개	6개	3개	3개
d	6개	8개	5개	5개

여기서 a : 전체구간, b : 중첩 구간, c : 중첩 구간에서 퍼지분할의 수, d : 총 퍼지분할의 수를 나타낸다.

### 3.3 각 계층의 퍼지-러프 분류결과

본 논문에서는 계층적 퍼지-러프 분류를 위한 추론방법으로 Larsen의 max-product 합성과 무게중심법을 사용하였다.

### 3.3.1 계층 1 : Petal Length와 Petal Width

다음은 3.2절에서 획득한 각 속성의 구간과 퍼지분할의 수를 근거로 2개의 입력부 퍼지집합과 1개의 출력부 퍼지집합을 구성한 것이다.

표 3. 계층 1의 입력부와 출력부 퍼지집합

Table 3. Fuzzy sets for antecedent and consequent parts of layer 1

Petal Length의 소속구간 - 입력부 퍼지집합1	분할 수
(1) PL1 = [0 0 4.2 4.5] (2) PL2 = [4.2 4.5 4.8] (3) PL3 = [4.5 4.8 5.1] (4) PL4 = [4.8 5.1 5.4] (5) PL5 = [5.1 5.4 8.662 8.662]	5개
Petal Width의 소속구간 - 입력부 퍼지집합2	분할 수
(1) PW1 = [0 0 1.2 1.4] (2) PW2 = [1.2 1.4 1.6] (3) PW3 = [1.4 1.6 1.8] (4) PW4 = [1.6 1.8 2.0] (5) PW5 = [1.8 2.0 3.265 3.265]	5개
Y1 - 출력부 퍼지집합	분할 수
(1) Setosa = [0 0.25 0.5] (2) Versicolor = [0.25 0.5 0.75] (3) Virginica = [0.5 0.75 1.0]	3개

표 3에서 PL1, PL5, PW1, 그리고 PW5는 4개의 파라미터를 가지는 사다리꼴 퍼지집합을 나타내고, 나머지 linguistic term들은 3개의 파라미터를 가지는 삼각형 퍼지집합을 나타낸다. 또한 파라미터의 값들 중에서 진하게 표시된 부분은 레이블된 클래스들의 최소-최대 구간이 서로 중첩되는 구간을 의미한다. 아래 표 4는 Petal Length와 Petal Width에서 수치적 입력으로 생성한 각 클래스의 후보규칙 수와 2.4절의 속성감축 후 생성된 각 클래스의 규칙 수이다.

표 4. 속성 감축전과 후의 계층 1의 규칙

Table 4. Number of Layer 1's rules before and after reduction of attributes

속성 감축 전 후보규칙 (수치적인 입력패턴으로 생성)	총 규칙의 수
Setosa : 1개, Versicolor : 9개, Virginica : 9개	19개
속성 감축 후 규칙 (리덕트가 1개인 규칙)	총 규칙의 수
Setosa : 1개, Versicolor : 7개, Virginica : 4개	12개

아래 표 5는 속성감축 후 생성된 규칙을 사용한 Petal Length와 Petal Width에서의 추론결과를 나타낸다.

표 5. 계층 1의 추론결과

Table 5. Inference results on layer 1

클래스	Y1 (계층 1의 추론결과)	
	분류구간	오분류 구간
Setosa	0.2500 (50개)	없음
Versicolor	0.2500 - 0.7500 (25개)	0.2500 - 0.7500 (25개)
Virginica	0.5764 - 0.7500 (47개)	0.5764 - 0.6250 (3개)

표 5의 결과와 2.5절의 퍼지분할 결정방법으로부터 계산된 계층 1의 추론결과와 퍼지분할 정보는 다음과 같다.

a	[0.25-0.217, 0.75+0.217] = [0.0330, 0.9670]
b	[0.5764, 0.6250]
c	2개(min : 0.25, max : 0.75, sd : 0.217)
d	4개

여기서 a:전체구간, b:오분류 구간에서의 중첩구간, c:중첩 구간에서의 퍼지분할 수, d:전체 퍼지분할 수를 나타낸다.

### 3.3.2 계층 2 : 계층 1의 추론결과와 Sepal Length

계층 1의 추론결과에 대한 퍼지분할 정보와 3.2절의 Sepal Length의 퍼지분할 구간을 근거로 다음과 같은 2개의 입력부 퍼지집합과 1개의 출력부 퍼지집합을 생성하였다.

표 6. 계층 2의 입력부와 출력부 퍼지집합  
Table 6. Fuzzy sets for antecedent and consequent parts of layer 2

Y1의 소속구간 - 입력부 퍼지집합1	분할 수
(1) Y11 = [0.033 0.033 0.25 0.5764]	4개
(2) Y12 = [0.25 0.5764 0.625]	
(3) Y13 = [0.5764 0.625 0.75]	
(4) Y14 = [0.625 0.75 0.967 0.967]	
Sepal Length의 소속구간 - 입력부 퍼지집합2	분할 수
(1) SL1 = [3.472 3.472 4.2 4.9]	6개
(2) SL2 = [4.2 4.9 5.6]	
(3) SL3 = [4.9 5.6 6.3]	
(4) SL4 = [5.6 6.3 7.0]	
(5) SL5 = [6.3 7.0 7.7]	
(6) SL6 = [7.0 7.7 8.728 8.728]	
Y2 - 출력부 퍼지집합	분할 수
Y1의 출력부 퍼지집합의 소속구간과 동일	3개

표7은 계층 1의 추론결과와 Sepal Length에서 수치적 입력으로 생성한 각 클래스의 후보규칙 수와 러프집합으로 감축한 후 생성된 각 클래스의 규칙 수이다.

표 7. 속성 감축전과 후의 계층 2의 규칙  
Table 7. Number of Layer 2's rules before and after reduction of attributes

속성 감축 전 후보규칙 (수치적인 입력패턴으로 생성)	총 규칙의 수
Setosa : 2개, Versicolor : 6개, Virginica : 6개	14개
속성 감축 후 규칙 (리덕트가 1개인 규칙)	총 규칙의 수
Setosa : 2개, Versicolor : 6개, Virginica : 3개	11개

표 8은 계층 2에서의 추론결과(계층 1의 추론결과와 Sepal Length를 기반으로 한)를 나타낸다.

표 8. 계층 2의 추론결과  
Table 8. Inference results on layer 2

클래스	Y2 (계층 2의 추론결과)	
	분류구간	오분류 구간
Setosa	0.2500 - 0.5000 (39개)	0.3962 - 0.5000 (11개)
Versicolor	0.2500 - 0.7500 (44개)	0.2500 - 0.7500 (6개)
Virginica	0.7500 (50개)	없음

이전 절의 방법에서처럼 계층 2의 추론결과에 대한 퍼지 분할 정보를 계산한 결과 다음과 같음을 볼 수 있었다.

a	[0.0600, 0.9400]
b	[0.3962, 0.5000]
c	3개 (min : 0.25, max : 0.75, sd : 0.19)
d	5개

여기서 a: 전체구간, b:오분류 구간에서의 중첩 구간, c:중첩 구간에서의 퍼지분할 수, d:전체 퍼지분할 수를 나타낸다.

### 3.3.3 계층 3 : 계층 2의 추론결과와 Sepal Width

계층 2의 추론결과와 3.2절의 Sepal Width의 퍼지분할 구간을 이용하여 다음과 같은 2개의 입력부 퍼지집합과 1개의 출력부 퍼지집합을 생성하였다.

표 9. 계층 3의 입력부와 출력부 퍼지집합  
Table 9. Fuzzy sets for antecedent and consequent parts of layer 3

Y2의 소속구간-입력부 퍼지집합1	분할 수
(1) Y21 = [0.06 0.06 0.25 0.3692]	5개
(2) Y22 = [0.25 0.3962 0.4481]	
(3) Y23 = [0.3962 0.4481 0.5]	
(4) Y24 = [0.4481 0.5 0.75]	
(5) Y25 = [0.5 0.75 0.94 0.94]	
Sepal Width의 소속구간 - 입력부 퍼지집합2	분할 수
(1) SW1 = [1.564 1.564 1.88 2.2]	8개
(2) SW2 = [1.88 2.2 2.52]	
(3) SW3 = [2.2 2.52 2.84]	
(4) SW4 = [2.52 2.84 3.16]	
(5) SW5 = [2.84 3.16 3.48]	
(6) SW6 = [3.16 3.48 3.8]	
(7) SW7 = [3.48 3.8 4.12]	
(8) SW8 = [3.8 4.12 4.836 4.836]	
Y3 - 출력부 퍼지집합	분할 수
Y1의 출력부 퍼지집합의 소속구간과 동일	3개

표 10은 계층 2의 추론결과와 Sepal Width에서 수치적 입력으로 생성한 각 클래스의 규칙 수와 러프집합으로 감축한 후 생성된 각 클래스의 규칙 수이다.

표 10. 속성 감축전과 후의 계층 3의 규칙  
Table 10. Number of Layer 3's rules before and after reduction of attributes

속성 감축 전 후보규칙 (수치적인 입력패턴으로 생성)	총 규칙의 수
Setosa : 12개, Versicolor : 9개, Virginica : 6개	27개
속성 감축 후 규칙 (리덕트가 1개인 규칙)	총 규칙의 수
Setosa : 9개, Versicolor : 9개, Virginica : 1개	19개

표 11은 최종적으로 IRIS 데이터를 분류한 추론결과들을 나타내고, 그 결과로부터 총 150개 데이터들 중에서 1개의 입력패턴에 대해서 Setosa 클래스가 Versicolor 클래스로 잘못 분류되었음을 알 수 있다.

표 11. 계층 3의 추론결과  
Table 11. Inference results on layer 3

클래스	Y3 (계층 3의 추론결과)	
	분류구간	오분류 구간
Setosa	0.2500 - 0.5000 (49개)	0.5 (1개)
Versicolor	0.3657 - 0.6112 (50개)	없음
Virginica	0.74995 - 0.74999 (50개)	없음

3.4 계층적 퍼지-러프 분류결과와 기존의 분류방법들과의 비교

표 12는 계층적 퍼지-러프 분류과정 동안에 각 계층에서의 분류 정확도와 규칙 수의 변화를 나타낸다.

표 12. 각 계층의 분류 정확도와 규칙의 수  
Table 12. Classification accuracy and number of rules on each layer

계층	분류 정확도	규칙의 수	
		감축전	감축후
계층 1	122/150 (81.33%)	19개	12개
계층 2	127/150 (84.67%)	14개	11개
계층 3	149/150 (99.33%)	27개	19개

그림 3에서 4까지는 IRIS 데이터를 최종 분류한 비퍼지화된 추론 패턴들을 나타낸 것이다.

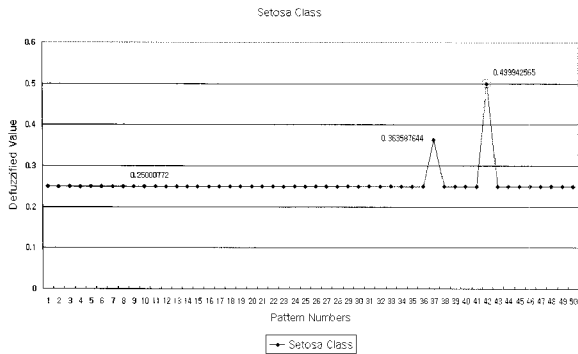


그림 3. Setosa 클래스의 추론 패턴  
Fig. 3. Inference patterns of setosa class

그림에서 'O'는 Setosa 클래스가 Versicolor 클래스로 오분류된 패턴을 나타낸다.

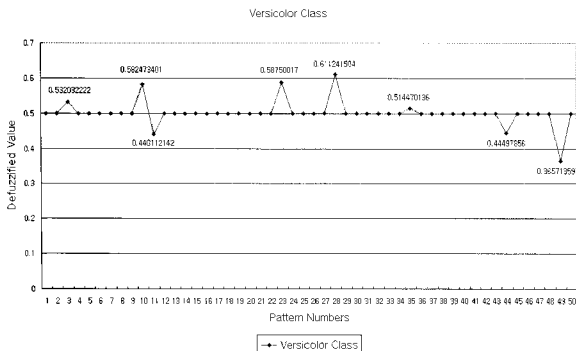


그림 4. Versicolor 클래스의 추론 패턴  
Fig. 4. Inference patterns of versicolor class

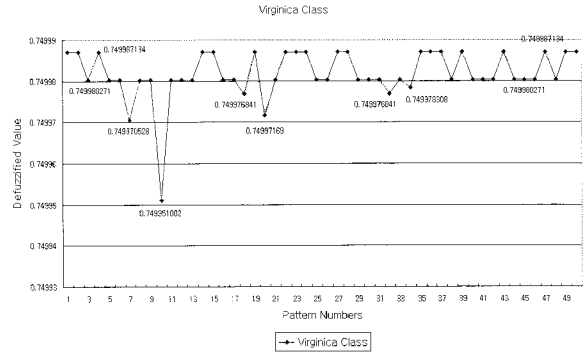


그림 5. Virginica 클래스의 추론 패턴  
Fig. 5. Inference patterns of virginica class

표 13에서는 제안된 방법의 효과성을 보이기 위해 IRIS 데이터를 사용한 여러 패턴분류 방법들과 비교하였다.

표 13. 기존 방법들과의 비교  
Table 13. Comparison the proposed with conventional methods

패턴분류 방법		규칙 수 및 분류 정확도	
Neuro - fuzzy approach	Nauck[1]	7개	145개 (96.6%)
Adjustment of rule weight	Nozaki[2]	6개	142개 (94.8%)
Genetic-algorithm approach	Ishibuchi[3]	6개	149개 (99.3%)
	Ishibuchi[4]	5개	147개 (98.0%)
Entropy based fuzzy-rough	Tsai[5]	16개	149개 (99.3%)
Hierarchical fuzzy-based approach	Guo[10]	분할 수:6 (서브시스템 1:27개, 서브시스템 2:30개)	146개 (97.33%)
	Chen[6]	16개	149개 (99.6%)
The proposed method		19개	149개 (99.3%)

표 13에서 Chen[6]의 분류 정확도는 10번의 실험으로 얻은 결과의 평균치를 나타낸다. 이들 비교 결과들로부터, 제안된 방법은 각 계층의 추론결과 구간들을 다음 계층의 퍼지분할 구간을 결정하는데 사용함으로써, 기존의 패턴분류 방법의 분류 정확성과 비슷한 성능을 보임을 볼 수 있다.

5. 결 론

본 논문에서는 학습기법을 사용하지 않고 패턴분류의 성능을 최대화하면서 규칙의 수를 줄일 수 있는 패턴분류 방법을 제안하였다. 제안된 방법에서는 규칙의 수를 최소화하기 위해서 계층적 퍼지 시스템과 러프집합의 특성을 결합하였고, 분류성능을 극대화하기 위해서 통계적 정보를 활용하여 레이블된 클래스에서 각 속성의 중첩 구간만을 고려하여 분할 구간을 결정하였다. 또한 패턴분류에서 사용되는 규칙들을 추출하기 위해 후보규칙들을 선택할 수 있는 방법에 대해

서도 논의하였다.

제안된 방법의 효과성을 보이기 위해서 Fisher의 IRIS 데이터를 사용하여 기존의 패턴분류 방법의 규칙 수와 분류의 정확성을 비교하였다. 그 결과, 패턴분류를 위한 퍼지분할 구간의 결정에 있어서 학습기법을 사용하지 않고 주어진 데이터의 통계적 정보만을 고려하더라도 유사한 성능을 얻을 수 있음을 확인하였다.

### 참 고 문 헌

- [1] D. Nauck, U. Nauck, and R. Kruse, "Generating classification rules with the neuro-fuzzy system NEFCLASS", *In proceedings of the biennial conference of NAFIPS, Berkeley*, pp. 19-22, 1996.
- [2] S. M. Nozaki, H. Ishibuchi, and H. Tanaka, "Adaptive fuzzy rule based classification systems", *IEEE transactions on fuzzy systems*, vol. 4, no. 3, pp. 238-250, 1996.
- [3] H. Ishibuchi, K. Nozaki, N. Yamamoto, and H. Tanaka, "Construction of fuzzy classification systems with rectangular fuzzy rules using genetic algorithm", *Fuzzy sets and systems*, vol. 65, no. 2-3, pp. 237-253, 1994.
- [4] H. Ishibuchi, 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*, vol. 89, no. 2, pp. 135-150, 1997.
- [5] Y. C. Tsai, C. H. Cheng, and J. R. Chang, "Entropy-based fuzzy rough classification approach for extracting classification rules", *Experts systems with applications*, vol. 31, pp. 436-443, 2006.
- [6] Y. Chen, B. Yang, A. Abraham, and L. Peng, "Automatic design of hierarchical takagi-sugeno type fuzzy systems using evolutionary algorithms", *IEEE transactions on fuzzy systems*, vol. 15, no. 3, pp. 385-397, 2007.
- [7] A. Skowron and C.M. Rauszer, "The Discernibility matrices and functions in information systems", *Institute of computer sciences report 1/91, Technical University of Warsaw*, pp. 1-41, 1991.
- [8] R. A. Fisher, "The use of multiple measurements in taxonomic problems", *In annual eugenics*, vol. 7, no. 2, pp. 179-188, 1936.
- [9] M. A. Kbir, H. Benkirane, K. Maalmi, and R. Benslimane, "Hierarchical fuzzy partition for pattern classification with fuzzy if-then rules", *Pattern recognition letters*, vol. 21, pp. 503-509, 2000.
- [10] N. R. Guo, T.H.S. Li, and C.L. Kuo, "Hierarchical fuzzy model for classification problem", *IECON 2002*, vol. 3, pp. 2096-2101, 2002.
- [11] 손창식, 정환목, 서석태, 권순학, "규칙의 커플링 문제를 최소화하기 위한 퍼지-러프 분류방법", *한국 퍼지 및 지능시스템 학회 논문지*, vol. 17, no. 4, pp. 460-465, 2007.
- [12] C. S. Son, S. T. Seo, H. M. Chung, and Soon, H. Kwon, "Statistical information-based fuzzy rough classification", *Research report # 2007-08*.

### 저 자 소 개

손창식(Chang S. Son)

제17권 제4호 참조

서석태(Suk T. Seo)

제17권 제4호 참조

정환목(Hwan M. Chung)

제17권 제4호 참조

권순학(Soon H. Kwon)

제17권 제4호 참조