

# 러프-신경망과 $\chi^2$ 검정에 의한 효율적인 의사결정지원시스템

정 환 목<sup>†</sup> · 피 수 영<sup>††</sup> · 최 경 옥<sup>†††</sup>

## 요 약

의사결정에 있어 정보란 의사결정자가 의사결정을 하는데 사용하도록 의미있고 유용한 형태로 처리된 데이터이다. 이러한 정보들에 있어서 불필요한 속성들을 제거하여 처리함으로써 의사결정의 효율을 높일 수 있다. 러프 집합 이론은 불필요한 속성을 제거하고 분류화 하는데 뛰어난 능력을 가지고 있으나 속성 감축시 속성 수와 튜플 수에 따라 복잡한 계산을 요구한다. 따라서 속성들 사이의 상호연관성을 나타내는 척도로서, 두 변수간의 독립성에 관한 검정방법인  $\chi^2$ 와 러프 집합의 종속성을 이용하여 속성들을 감축하고 이를 신경망의 입력 유니트로 사용함으로써 기존의 모든 속성을 입력으로 하여 신경망을 구성하는 것보다 간단하며 학습효율의 향상 및 정확한 의사결정을 지원할 수 있다.

## Efficient Decision Making Support System by Rough-Neural Network and $\chi^2$

Hwan-Mook Chung<sup>†</sup> · Su-Young Pi<sup>††</sup> · Kyoung-Oak Choi<sup>†††</sup>

## ABSTRACT

In decision-making, information is the thing manufactured as the useful type for decision-making. We can improve the efficiency of decision-making by elimination of unnecessary information. Rough set is the theory that can classify and reduce the unnecessary attributes. But the reduction process of rough set becomes more complex according to the number of attribute and tuple. After eliminating of the dispensable attributes using  $\chi^2$  and rough set, the indispensable attributes are used for the units of input layers in neural network. This rough-neural network can support more correct decision-making of neural network.

### 1. 서 론

의사결정이란 어떠한 목표의 설정과 달성을 위해 결정대안을 합리적으로 선택하고 결정하는 제반활동과 행동을 나타내는 것이다. 의사결정에 있어서 발생하는

모든 문제는 대안과 관련된 여러 가지 정보들을 필요로 하게 된다. 의사결정에 있어 정보란 의사결정자가 의사결정을 하는데 사용하도록 의미있고 유용한 형태로 처리된 데이터이다. 현대의 그룹은 대부분 어떠한 형태로든 정보시스템을 가지고 있다. 문제는 이를 활용하여 자료를 축적하고 생성하여 의사결정에 필요한 적절한 형태의 데이터를 어떠한 방식으로 지원하는가이다. 즉 불필요한 정보를 제거하여 의사결정의 효율을 높이는 것이다[4].

† 정 회 원 : 대구효성가톨릭대학교 공과대학 전자정보공학부 교수, 한국퍼지 및 지능시스템 학회 부회장 겸 대구·경북 지회장

†† 비 회 원 : 대구영진전문대학 컴퓨터정보기술계열 겸임교수

††† 준 회 원 : 대구미래대학, 서라벌대학 강사

논문접수 : 1998년 12월 14일, 심사완료 : 1999년 8월 5일

Pawlak에 의하여 도입된 러프 집합 이론은 데이터 처리 측면에서 현재 광범위하게 이용되고 있다[11]. 이 이론은 불충분하거나 일관성이 없는 정보로부터 체계적이고 자동적인 방법으로 핵심적인 정보를 추출해 내며, 특히 방대하고 불분명한 자료 및 정보를 해석하는데 있어서 여러 가지 속성을 이용한 분류화 및 근사화를 효과적으로 제공하는 능력이 있다.

그러나 러프 집합의 속성감축 방법은 정보시스템의 조건속성 수 및 튜플 수에 따른 비교 계산을 필요로 하기 때문에, 큰 데이터베이스에 러프 집합을 그대로 적용한다는 것은 많은 시간 복잡도를 요구하게 된다[13].

따라서 두 변수간의 독립성에 관한 검정방법인  $\chi^2$  적합도검정을 이용하여 구한 값을 러프 집합의 종속성을 이용한 속성 감축에 적용하여 러프 집합의 속성감축을 위한 시간 복잡도를 줄일 수 있다.

본 논문에서는  $\chi^2$ 와 러프 진처리 과정을 통해 불필요한 속성들을 제거시킨 후 이를 신경망의 입력 유니트로 사용하는 러프-신경망 모델을 의사결정에 적용함으로써 기존의 모든 속성을 입력으로 하여 신경망을 구성하는 것보다 훨씬 간단하며 학습효율의 향상 및 정확한 의사결정을 지원할 수 있다.

## 2. 러프 집합과 지식의 표현

### 2.1 러프 집합

러프 집합 이론은 객체들의 속성값들의 특징을 분석하여 객체들을 군집화하는, 모호성과 불확실성에 대한 새로운 수학적 접근 방법이다. 러프 집합을 통해 불완전하거나 일관성이 없는 정보로부터 체계적인 정보를 가려냄으로써 여러 속성 값들로부터 구분화 및 근사화할 수 있다[1, 2].

$KB=(U, R)$ 가 지식베이스라 할 때,  $U(U \neq \emptyset)$ 는 논의영역이라 불리는 유한 집합이고,  $R$ 은  $U$ 에서의 동치관계들의 모임이다.

#### [정의 1] Indiscernibility Relation

$$[X]_{IND(P)} = \bigcap_{R \in P} \{X\}_R, \quad (P \subseteq R, P \neq \emptyset)$$

#### [정의 2] Lower Approximation

$R$ 의 동치류들 중에서  $X$ 에 포함되는 모든 동치류들의 합집합을  $X$ 의  $R$ -하한근사라 한다.

$$RX = \bigcup \{Y \in U/R : Y \subseteq X\}$$

#### [정의 3] Upper Approximation

$R$ 의 동치류들 중에서  $X$ 와의 교집합이 공집합이 아닌 동치류들의 합집합을  $X$ 의  $R$  상한근사라 한다.

$$\bar{R}X = \bigcup \{Y \in U/R : Y \cap X \neq \emptyset\}$$

#### [정의 4] Positive, Negative, and Borderline region

$POS_R(X) = RX$ 는  $X$ 의  $R$ -긍정영역

$NEG_R(X) = U - \bar{R}X$ 는  $X$ 의  $R$ -부정영역

$BN_R(X) = \bar{R}X - RX$ 는  $X$ 의  $R$ -경계영역

#### [정의 5] Partial Dependency

지식베이스  $KB=(U, R)$ 이고  $P, Q \subseteq R$ 일때 아래의 수식에 의한  $k$ 값에 의하여 지식  $Q$ 는 지식  $P$ 에  $k(0 \leq k \leq 1)$ 만큼 종속된다고 한다.

$$k(P, Q) = \frac{\text{card } POS_P(Q)}{\text{card } U}, \quad (\text{card: 원소의 개수})$$

### 2.2 러프 집합의 속성 감축

연산을 위해서는 지식의 구문론적 표현을 필요로 한다. 결국 컴퓨터 처리에 적합한 기호 형태로 동치관계들 또는 분할을 표기하기 위해서 사용되는 특정한 형식언어인 표 형태의 지식 표현을 사용한다. 이런 데이터 표를 지식표현시스템(Knowledge Re-representation System : KRS)이라 한다.

#### [정의 6] Knowledge Representation System

지식표현 시스템  $S = (U, A)$

$U$ : 전체집합 ( $U \neq \emptyset$ )

$A$ : 기본속성의 유한집합 ( $A \neq \emptyset$ )

다음의 지식표현시스템의 예를 통해 러프 집합의 속성감축의 계산복잡도를 볼 수 있다.

$$U = \{1, 2, 3, 4, 5\}$$

$$A = \{a, b, c, d\}$$

$$V = V_a = V_b = V_c = V_d = \{0, 1, 2\}$$

이 지식표현 시스템은 다음 <표 1>과 같이 표현한다.

<표 1> 지식표현시스템

U	a	b	c	d
1	1	0	2	0
2	2	1	0	1
3	2	1	2	2
4	1	2	2	0
5	1	2	0	2

조건속성들의 집합  $C = \{a, b, c\}$ 와 결정속성의 집합  $D = \{d\}$ 로 구성되어 있는 <표 1>의 지식표현시스템에서 집합  $C$ 와 집합  $D$  사이의 종속성의 정도는 다음과 같다.

$$U/IND(D) = \{\{1, 4\}, \{2\}, \{3, 5\}\} = \{X_1, X_2, X_3\}$$

$$U/IND(C) = \{\{1\}, \{2\}, \{3\}, \{4\}, \{5\}\}$$

$$= \{Y_1, Y_2, Y_3, Y_4, Y_5\}$$

$$CX_1 = Y_1 \cup Y_4 = \{1, 4\}$$

$$CX_2 = Y_2 = \{2\}$$

$$CX_3 = Y_3 \cup Y_5 = \{3, 5\} \text{ 이므로}$$

$$POS_C(D) = Y_1 \cup Y_2 \cup Y_3 \cup Y_4 \cup Y_5$$

$$= \{1, 2, 3, 4, 5\}$$

그러므로  $C$ 와  $D$ 의 종속성의 정도는

$$k(C, D) = 5/5 = 1$$

여기서, 각 속성의 indispensable 여부는 다음과 같다. 먼저 속성  $a$ 에 관한 indispensable 여부를 계산한다.

$$U/IND(D) = \{\{1, 4\}, \{2\}, \{3, 5\}\} = \{X_1, X_2, X_3\}$$

$$U/IND(C - \{a\}) = \{\{1\}, \{2\}, \{3\}, \{4\}, \{5\}\} \text{에서}$$

$$C - \{a\}X_1 = Y_1 \cup Y_4 = \{1, 4\}$$

$$C - \{a\}X_2 = Y_2 = \{2\}$$

$$C - \{a\}X_3 = Y_3 \cup Y_5 = \{3, 5\} \text{ 이므로}$$

$$POS_{(C - \{a\})}(D) = Y_1 \cup Y_2 \cup Y_3 \cup Y_4 \cup Y_5$$

$$= \{1, 2, 3, 4, 5\}$$

$$k(C - \{a\}, D) = k(C, D) = \frac{5}{5} = 1.$$

그러므로 속성  $a$ 는 dispensable이다.

동일한 방법으로 속성  $b$ 와 속성  $c$ 에 대해서 구하면 속성  $b$ 는 dispensable이고, 속성  $c$ 는 indispensable이 된다.

여기서 속성과 튜플사이의 모순여부를 계산하면,  $(\{a\}, \{d\})$ 에 대해 아래와 같이 모순이 생긴다.

$$(\{a\}, \{d\}) = \{ a_1 \rightarrow d_0, a_2 \rightarrow d_1, a_2 \rightarrow d_2, a_1 \rightarrow d_0, a_1 \rightarrow d_2 \}$$

같은 방법으로  $(\{b\}, \{d\}), (\{c\}, \{d\})$  모두 불일치이며, 속성  $b$ 와  $c$ 가  $(\{b, c\}, \{d\})$ 에서,  $a$ 와  $c$ 가  $(\{a, c\}, \{d\})$ 에서 모순이 생기지 않는다.

$$(\{b, c\}, \{d\}) = \{ b_0 c_2 \rightarrow d_0, b_1 c_0 \rightarrow d_1, b_1 c_2 \rightarrow d_2, b_2 c_2 \rightarrow d_0, b_2 c_0 \rightarrow d_2 \}$$

$$(\{a, c\}, \{d\}) = \{ a_1 c_2 \rightarrow d_0, a_2 c_0 \rightarrow d_1, a_2 c_2 \rightarrow d_2, a_1 c_0 \rightarrow d_2 \}$$

$$(\{a, c\}, \{d\}) = \{ a_1 c_2 \rightarrow d_0, a_2 c_0 \rightarrow d_1, a_2 c_2 \rightarrow d_2, a_1 c_0 \rightarrow d_2 \}$$

$$a_2 c_2 \rightarrow d_2, a_1 c_2 \rightarrow d_0, a_1 c_0 \rightarrow d_2 \}$$

따라서 <표 1>은 <표 2>와 <표 3>으로 감축된다.

<표 2> 지식표현시스템

U	b	c	d
1	0	2	0
2	1	0	1
3	1	2	2
4	2	2	0
5	2	0	2

<표 3> 지식표현시스템

U	a	c	d
1	1	2	0
2	2	0	1
3	2	2	2
4	1	2	0
5	1	0	2

이처럼 러프집합은 주어진 조건속성에 대한 감축 결과를 구하기 위해 이들의 조건속성 수에 따라 부분집합을 형성하고, 또한 이들 집합에 대한 튜플들 간의 비교를 필요로 하게 된다. 즉 조건속성에 따른 계산 ( $O(2^n)$ )과 튜플수에 따른 비교 계산 유도를 필요로 한다. 러프집합의 이와 같은 성질로 인하여 많은 속성 수와 많은 튜플을 가지는 큰 데이터베이스에 러프집합을 그대로 적용한다는 것은 복잡한 계산유도를 필요로 하므로 본 논문에서는  $\chi^2$ 를 이용하여 속성의 중요도를 계산하고 이 값에 따라 러프집합의 종속성을 적용하였다.

### 3. $\chi^2$ 와 종속성에 의한 속성감축 알고리즘

$\chi^2$ 검정법은 실제 관측된 도수분포와 어떤 가정하에서 기대되는 이론 분포의 차를 검정하고, 양자의 일치 정도에 의해 그 가정을 기각할 것인가, 채택할 것인가를 판정하는 척도를 구하는 방법이다. 또한 독립성이라는 귀무가설(null hypothesis)에 대하여 통상적으로 많이 사용되는 검정방법이다.

$$\chi^2 = \sum \frac{(f_0 - f_e)^2}{f_e}, (f_0: \text{관찰도수}, f_e: \text{기대도수})$$

지식 표현시스템의 속성감축을 위해  $\chi^2$ 를 이용하여

관련이 많은 속성부터 차례대로 정렬한 후 이들을 러프 집합의 종속성의 성질을 이용하여 속성감축을 유도한다.

먼저  $\chi^2$  값에 의해 정렬된 속성들을 하나씩 부분집합 SM에 삽입시켜 종속성을 계산하고 이를 모든 조건속성을 포함한 조건속성의 집합 C와 결정속성의 집합 D 사이의 종속성과 비교하여 그 값이 같아질 때까지 부분집합 SM을 증가시켜 간다. 이때 계산된 집합 SM에서 다시 속성들을 하나씩 제거하였을 때의 종속성을 계산하여 이 값이 모든 속성을 포함한 속성집합의 종속성과 같으면 그 속성을 제거시키고, 그렇지 않으면 집합 SM에 다시 삽입시킨다. 이러한 방법으로 최종적으로 감축된 속성들로 구성된 집합 SM을 계산한다.

[속성감축을 위한 알고리즘]

$$A = \{a_1, a_2, \dots, a_n\}$$

$$\chi^2 = \sum_{i=1}^n \sum_{j=1}^k \frac{(f_{ij} - e_{ij})^2}{e_{ij}}$$

$$k(C, D) = \frac{\text{card}(\text{POS}(C, D))}{\text{card}(U)}$$

for  $i=1$  to  $n$

$$\chi^2(a_i)$$

endfor

sort  $\chi^2(a_i)$

while  $k(SM, D) \neq k(C, D)$  do

$$SM \leftarrow SM \cup a_i$$

$$k(SM, D)$$

endwhile

for  $i=0$  to  $n-1$  do

$$SM \leftarrow SM - a_i$$

$$k(SM, D)$$

if  $K(SM, D) \neq K(C, D)$  then

$$SM \leftarrow SM \cup a_i$$

endif

endfor

4. 러프-신경망에 의한 의사결정지원시스템 설계와 구현

러프 집합의 속성감축계산을 줄이기 위해  $\chi^2$ 에 따라서 속성을 정렬하고, 이를 러프 이론의 종속성에 따라 속성을 감축하는 알고리즘을 유도한다. 이렇게 하여 감축된 속성을 입력유니트로 하는 다층 신경회로망을 구성하였다. 이때 다층 신경회로망은 각 유니트의

연결강도를 변경하는 학습에 일반적으로 널리 사용되는 백프로파게이션(Backpropagation) 학습알고리즘을 적용하였다[7].

4.1 콘택트렌즈 의사결정 사례

본 논문에서 제안한 러프-신경망 모델을 콘택트렌즈 의사결정 사례에 적용하였다. 전문가에 의뢰하여 9개의 속성과 45개의 튜플을 가지는 데이터를 얻었다. 환자의 증상에 따라 콘택트렌즈의 사용여부와 콘택트렌즈의 종류를 결정하는 의사결정으로서, 이때 사용된 속성은 <표 4>와 같다.

<표 4> 콘택트렌즈 의사결정의 속성

속성 \ 속성치	1	2	3
a (성별)	남자	여자	
b (나이)	청소년	중년	노년
c (원근시)	근시	원시	
d (난시유무)	없음	있음	
e (눈물생성율)	감소	보통	
f (부동시정도)	2디옵터 이하	2~4 디옵터	4디옵터 이상
g (직업)	학생	컴퓨터 관련 직업	그 외 직업
h (시력)	매우 나쁨	중간	좋음
i (각막염증정도)	거의 없음	가끔 있음	많음
j (렌즈선택)	하드렌즈	소프트렌즈	렌즈착용안함

<표 5> 모든 조건속성을 포함한 정보시스템

	조건속성									결정속성
	a	b	c	d	e	f	g	h	i	
1	1	1	1	2	2	2	2	1	2	1
2	2	2	1	2	2	2	1	1	1	1
3	2	1	1	2	2	2	3	2	1	1
4	1	2	2	1	2	1	3	2	1	2
5	1	2	1	1	2	1	3	1	1	2
6	2	2	1	1	1	1	3	2	1	3
7	2	2	1	1	2	2	3	1	1	2
8	1	3	2	1	2	1	2	2	1	3
9	2	3	2	1	1	2	3	2	3	3
10	2	2	1	1	2	2	1	1	1	2
11	2	1	2	1	2	1	1	1	2	1
12	1	2	1	2	2	2	2	1	3	3
13	2	1	1	2	2	3	3	1	2	1
14	2	1	2	1	2	1	1	1	1	2
15	2	2	1	1	1	2	3	1	1	3
16	1	3	2	1	2	1	3	2	1	2
17	1	2	1	2	2	2	1	2	2	1
18	1	2	1	1	2	1	3	3	1	3
19	1	1	2	2	2	3	2	1	3	3
20	2	2	2	2	1	1	2	2	3	3

IND(C) = {{1}, {2}, {3}, {4}, {5}, {6}, {7}, {8}, {9},  
 {10}, {11}, {12}, {13}, {14}, {15}, {16},  
 {17}, {18}, {19}, {20}}

IND(D) = {{1, 2, 3, 11, 13, 17}, {4, 5, 7, 10, 14, 16},  
 {6, 8, 9, 12, 15, 18, 19, 20}}

POS<sub>C</sub>(D) = {1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15,  
 16, 17, 18, 19, 20}

$$k(C, D) = \frac{20}{20} = 1$$

▶ 각각의 조건속성과 결정속성 간의  $\chi^2$ 는 다음과 같다.

① a (성별)

a \ j	1	2	3	total
1	2	3	4	9
2	4	3	4	11
total	6	6	8	20

$$\chi^2(a) = 0.4714$$

<표 6>는 <표 5>에서의 속성에 대한  $\chi^2$ 를 나타낸다. 이때 조건속성은  $\chi^2$ 에 따라 관련성이 높은 속성부터 내림차순으로 정렬한다.

<표 6> 각 속성의  $\chi^2$

속 성	$\chi^2$
i (각막염증 정도)	17.7778
d (난시유무)	8.7153
g (직업)	8.0000
e (눈물생성율)	7.5000
b (나이)	6.0480
f (부등시정도)	3.5648
h (시력)	2.5758
c (원근시)	1.9444
a (성별)	0.4714

▶  $\chi^2$  값이 IND(D) = {{1, 2, 3, 11, 13, 17}, {4, 5, 7, 10, 14, 16}}, 높은 순서에 따라 속성을 부분집합 SM에 넣으면서 러프 집합의 종속도를 비교한다.

① SM에 i 삽입, SM = {i}일 경우

IND(SM) = {{2, 3, 4, 5, 6, 7, 8, 10, 14, 15, 16, 18}, {1, 11, 13, 17}, {9, 12, 19, 20}}

IND(D) = {{1, 2, 3, 4, 11, 13, 17}, {4, 5, 7, 10, 14, 16},  
 {6, 8, 9, 12, 15, 18, 19, 20}}

POS<sub>SM</sub>(D) = {1, 11, 13, 17, 9, 12, 19, 20}

$$k(SM, D) = \frac{8}{20} = 0.4, \quad k(SM, D) \neq k(C, D)$$

$k(SM, D) = k(C, D)$ 일 때까지 속성을 SM에 삽입한다. 부분집합 SM = {i, d, g, e, b, f, h} 을 얻을 수 있다. 이렇게 얻어진 SM을 가지고 다음의 계산을 수행한다.

▶ SM에서 속성을 한 개씩 제거한 후,  $k(SM, D) \neq k(C, D)$ 이면 그 속성을 다시 삽입하고, 만약  $k(SM, D) = k(C, D)$ 이면 그 속성은 제거시킨다.

① SM에서 i 제거, SM = {d, g, e, b, f, h}일 경우

IND(SM) = {{1}, {2}, {3}, {4}, {5}, {6}, {7}, {8}, {9},  
 {10}, {11, 14}, {12}, {13},

IND(D) = {{1, 2, 3, 11, 13, 17}, {4, 5, 7, 10, 14, 16},

{15}, {16}, {17}, {18}, {19}, {20}}

{6, 8, 9, 12, 15, 18, 19, 20}}

POS<sub>SM</sub>(D) = {1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 13, 15, 16, 17,  
 18, 19, 20}

$$k(SM, D) = \frac{18}{20} = 0.9, \quad k(SM, D) \neq k(C, D)$$

그러므로 속성 i 는 제거할 수 없다.

이렇게 하여 최종적으로 SM = {i, d, g, e, h}을 구할 수 있다. <표 7>의 정보시스템은 <표 5>의 정보시스템의 감축결과이다.

<표 7> 속성을 감축한 정보시스템

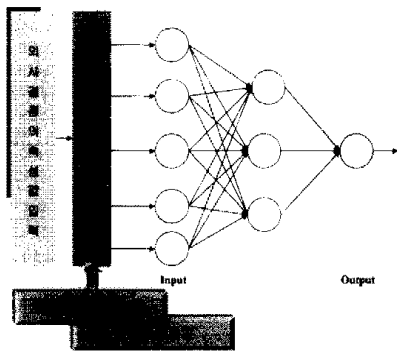
	조건속성					결정속성
	d	e	g	h	i	j
1	2	2	2	1	2	1
2	2	2	1	1	1	1
3	2	2	3	2	1	1
4	1	2	3	2	1	2
5	1	2	3	1	1	2
6	1	1	3	2	1	3
7	1	2	2	2	1	3
8	1	1	3	2	3	3
9	1	2	1	1	1	2
10	1	2	1	1	2	1
11	2	2	2	1	3	3
12	2	2	3	1	2	1
13	1	1	3	1	1	3
14	2	2	1	2	2	1
15	1	2	3	3	1	3
16	2	1	2	2	3	3

#### 4.2 러프-신경망 의사결정지원 시스템

본 논문에서 제안하는 러프-신경망 시스템에서는  $\chi^2$ 와 러프 집합을 이용하여 의사결정과정에서의 입력값들 중 불필요한 속성을 제거하고 제거된 속성을 입력 유니

프로 하인 신경망을 구성하였다. 감축된 속성을 입력 유니트로 사용함으로써 기존의 모든 속성을 입력으로 하여 신경망을 구성하는 것보다 훨씬 간단하며 학습 성능 효율이 높아져 의사결정의 정확성을 높일 수 있었다.

본 논문의 신경망은 러프집합을 통해 감축된 속성을 신경망의 입력 유니트로 받아 은닉층으로 전달, 최종적으로 출력층으로 전달하여, 이때 출력값과 목표값을 비교하여 그 차이를 감소시키는 방향으로 연결강도를 수정하는 BP 학습알고리즘을 사용하였다. (그림 1)은 본 논문에서 제안한  $\chi^2$ 와 러프 집합에 의한 전처리과정을 가지는 러프-신경망 모델이다.



(그림 1) 러프-신경망 모델

러프 집합을 이용한 전처리 과정을 통해 얻어진 속성들을 신경망의 입력층 유니트로 사용하였다. 이때 9개의 속성 중  $\chi^2$ 와 러프 집합의 종속성 정도에 따라 4개의 속성을 감축하고 나머지 5개의 속성을 신경망의 입력 유니트로 사용하였다. 이때 은닉층 유니트의 수는 2개로 하였을 때 보다 3개로 할 때가 더 좋은 성능을 보였다. 45개의 데이터 중 20개의 데이터를 학습 데이터로 사용하고 나머지 25개의 데이터를 테스트 데이터로 사용하였다.

학습회수에 따른 신경망의 학습율을 <표 8>에 나타내었다. 모든 조건을 사용했을 때의 학습율보다 전처리과정을 통해 감축된 속성을 가지고 신경망을 학습시켰을 때의 학습율이 훨씬 높아졌음을 알 수 있다

<표 8> 러프-신경망 학습결과

학습횟수 조건속성	2000회	3000회	4000회	5000회
모든 조건속성 사용시의 학습율	67.80%	68.63%	69.00%	69.53%
감축된 조건속성 사용시의 학습율	96.82%	97.53%	97.89%	98.12%

### 5. 결 론

본 논문에서는 러프 집합을 이용하여 신경망 학습의 효율을 높이는 의사결정지원시스템의 새로운 모델을 제안하였다. 제안된 모델은  $\chi^2$ 를 계산하고 여기에 러프 집합의 종속성을 적용하여 불필요한 속성을 제거함으로써 러프 집합의 계산복잡도를 감축하였다. 이렇게 하여 감축된 속성을 신경망의 입력 유니트로 사용하는 의사결정지원시스템의 새로운 모델을 제안하였다.

제안된 모델을 이용하여 환자의 여러 가지 증상에 따라 콘택트 렌즈를 선택하는 의사결정지원시스템을 구현하였다.

구현 결과 모든 속성을 입력 유니트로 사용하는 신경망보다 러프-신경망 모델의 학습효율이 68.74%에서 97.59%로 증가했음을 알 수 있었다. 따라서 러프-신경망 모델을 사용함으로써 속성감축의 복잡도를 단순화시킬 수 있었으며 보다 높은 학습성능의 효율로 인해 정확한 의사결정을 지원할 수 있었다.

### 참 고 문 헌

- [1] 전상화·정환목, "러프 집합을 이용한 규칙베이스와 사례베이스의 통합 추론 모델에 관한 연구", 한국정보처리학회 논문집, 제 5권 1호, pp.103-110, 1998.
- [2] Z. Pawlak, "Rough Sets-Theoretical Aspects of Reasoning about Data," Kluwer, 1991.
- [3] Roman Slowinski, "Intelligent decision support : handbook of applications and advances of the rough sets theory," Dordrecht : Kluwer Academic, 1992.
- [4] N. DeClaris, M. C. Su, "A Neural Network Based Approach to Knowledge Acquisition and Expert Systems," 1993.
- [5] E. Khan, "NeuFuz : An Intelligent Combination of Fuzzy Logic with Neural Nets," proceeding of 1993 International Joint Conference on Neural Networks pp.2945-2950, 1993.
- [6] S. H. Kim, K. S. Park, K. C. Jeong, "A Knowledge-Based system Using Neural Network for Management Evaluation and It's Support," Journal of the Korean OR/MS Society Vol.19, No.2, August 1994.

[7] LiMin Fu, "Neural Network in Computer Intelligence," McGraw Hill, 1994.

[8] Y. Yoon, T. Guimaraes, G. Swales, "Integrating artificial neural networks with rule-based expert systems." Decision Support System, pp.497-507, 1994.

[9] L. P. J. Veclenturf "Analysis and application of artificial neural networks," Prentice Hall, 1995.

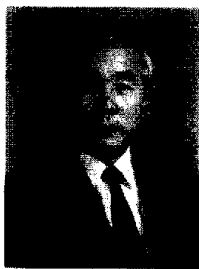
[10] S. Mitra, M. Banerjee, "knowledge-Based Neural Net with Rough Sets," Methodologies for the Conception Design and Application of Intelligent System proceeding of IIZUKA'96, pp.213-216, 1996.

[11] T. S. Quah, C. L. Tan, K. S. Raman, B. Srinivasan, "Toward integrating rule-based expert systems and neural networks," Decision Support Systems, pp.99-118, 1996

[12] Z. Pawlak, "Rough Logic," Bulletin of the Polish Academy of Sciences, Technical Science 35, 1997.

[13] H. Chung, K. O. Choi, H. M. Chung, "Integrated Method for Knowledge Discovery in Databases," The Third Asian Fuzzy Systems Symposium, June pp.18-21, 1998.

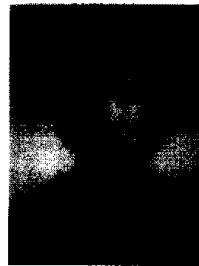
[14] N. Cercone, H. Hamilton, X. Hu, N. Shan, "Data Mining Using Attribute-Oriented Generalization and Information Reduction".



**정 환 목**

e-mail : hmchung@cuth.cataegu.ac.kr  
 1972년 한양대학교 전자공학과 졸업(학사)  
 1982년 인하대학교 대학원 졸업(이학석사)  
 1987년 인하대학교 대학원 졸업(이학박사)

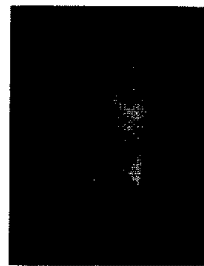
1984년~현재 대구효성가톨릭대학교 공과대학 전자정보공학부 교수  
 1998년~현재 한국퍼지 및 지능시스템 학회 부회장 겸 대구·경북 지회장  
 관심분야 : 다치이론, 신경망, 러프, 전자상거래, 에이전트 등



**피 수 영**

e-mail : syipi@yeungjin-c.ac.kr  
 1987년 대구효성가톨릭대학교 전산통계학과 졸업(학사)  
 1989년 대구효성가톨릭대학교 대학원 전산통계학과 졸업(석사)

1996~현재 대구효성가톨릭대학교 대학원 전산통계학과 전산전공 (박사과정 수료)  
 1995~현재 대구영진전문대학 컴퓨터정보기술계열 겸임교수  
 관심분야 : 신경망, 러프, 전자상거래, 에이전트 등



**최 경 옥**

e-mail : g7521003@cuth.cataegu.ac.kr  
 1997년 대구효성가톨릭대학교 전자계산학과 졸업(학사)  
 1999년 대구효성가톨릭대학교 대학원 전산통계학과 졸업(석사)  
 1999년~현재 대구미래대학, 서라벌대학 강사

관심분야 : 신경망, 러프집합, 데이터마이닝 등