

# 가중 퍼지 소속함수 기반 신경망을 이용한 Wisconsin Breast Cancer 예측 퍼지규칙의 추출

임 준 식\*

요 약

본 논문은 가중 퍼지소속함수 기반 신경망(Neural Network with Weighted Fuzzy Membership Functions, NNWFM)을 이용하여 Wisconsin breast cancer의 예측을 수행하는 퍼지규칙을 추출하고 있다. NNWFM는 자기적응적(self adaptive)가중 퍼지소속함수를 가지고 주어진 입력 데이터로부터 학습하여 퍼지규칙을 생성하고 이를 기반으로 예측을 수행한다. 신경망 구조의 중간 부분인 하이퍼박스(hyperbox)들은  $n$  개의 대, 중, 소의 가중 퍼지소속함수 집합으로 구성되며, 학습 후 각 집합은 퍼지집합의 bounded sum을 사용하여 다시 하나의 가중 퍼지소속함수로 합성된다.  $n$  개의 특징입력(feature input)은 학습된 모든 하이퍼박스에 연결되어 예측 작업을 수행한다. NNWFM으로 추출된 2개의 퍼지규칙은 99.41%의 예측 인식율을 가지며 이는 퍼지규칙의 수와 인식율에 있어 현재 발표된 논문의 결과보다 우수함을 보여준다.

## Extracting Wisconsin Breast Cancer Prediction Fuzzy Rules Using Neural Network with Weighted Fuzzy Membership Functions

Joon Shik Lim\*

ABSTRACT

This paper presents fuzzy rules to predict diagnosis of Wisconsin breast cancer using neural network with weighted fuzzy membership functions (NNWFM). NNWFM is capable of self-adapting weighted membership functions to enhance accuracy in prediction from the given clinical training data.  $n$  set of small, medium, and large weighted triangular membership functions in a hyperbox are used for representing  $n$  set of featured input. The membership functions are randomly distributed and weighted initially, and then their positions and weights are adjusted during learning. After learning, prediction rules are extracted directly from the enhanced bounded sums of  $n$  set of weighted fuzzy membership functions. Two number of prediction rules extracted from NNWFM outperforms to the current published results in number of rules and accuracy with 99.41%.

**키워드 :** 퍼지 신경망(Fuzzy Neural Network), 규칙 추출(Rule Extraction), 가중 퍼지 소속함수(Weighted Fuzzy Membership Function), Wisconsin Breast Cancer

### 1. 서 론

패턴 분류나 진단, 예측 등을 위해 신경망 및 퍼지집합 이론을 결합한 적응적 의사결정지원 툴(adaptive decision support tool)인 퍼지신경망(Fuzzy Neural Network, FNN)이 제안되어 왔다[1, 4, 7-9, 12, 15]. 여러 가지 구조의 FNN이 학습(learning), 적응(adaptation), 규칙추출(rule extraction)을 위한 알고리즘과 더불어 제시된 바 있다[3, 6, 10, 11].

if-then 퍼지규칙의 추출 기법이 의사결정을 위한 FNN의 장점 중 하나이다. 이는 분류나 예측에 있어서 보편적인 패

턴을 if-then 퍼지규칙과 같은 단순한 형태의 지식 표현으로 나타내는 것이다. 주어진 일련의 학습 데이터로부터 지식 추출을 위하여 자기 조직화 시스템(self-organizing system) 기반 퍼지신경망이 [5, 14, 16]이 개발되었다. Setnes[11]는 유전자 알고리즘을 이용하여 간단하고 정확도가 높은 퍼지규칙 기반 모델을 얻는데 초점을 맞추었다.

본 논문은 가중 퍼지소속함수 기반 신경망(Neural Network with Weighted Fuzzy Membership Functions, NNWFM)을 이용하여 Wisconsin breast cancer의 진단 예측을 수행하는 퍼지규칙을 추출하고 있다. NNWFM는 3개 계층 즉 입력 계층(input layer), 하이퍼박스 계층(hyperbox layer), 클래스 계층(class layer)으로 구성된다. 각 하이퍼박스 노

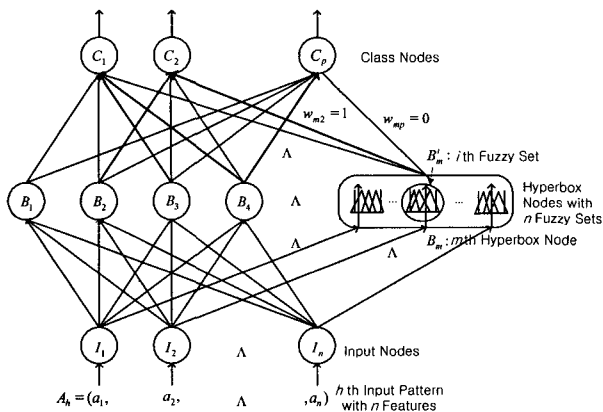
\* 정 회 원 : 경원대학교 소프트웨어학부 교수  
논문접수 : 2004년 4월 30일, 심사완료 : 2004년 8월 17일

드는  $n$  개(입력패턴수)의 퍼지집합으로 구성되는데 이것은 분류를 위한 클래스 노드와 연결된다. 각 퍼지집합은 3개의 가중 퍼지소속함수를 가지게 된다. NNWFM은 하이퍼박스에 있는 가중 퍼지소속함수를  $n$  개의 특징 입력패턴에 따라서 학습시킨다. 학습이 끝나면 각 퍼지집합에 있는 3개의 가중 퍼지소속함수는 bounded sum에 의해 하나의 가중 퍼지소속함수로 합성되어 예측용규칙을 만들게 된다.

**2. 가중 퍼지소속함수 기반 신경망(Neural Network with Weighted Fuzzy Membership Function, NNWFM)**

**2.1 NNWFM의 구조**

NNWFM의 구조는 (그림 1)에 나타나 있다. NNWFM는 3개의 계층 즉 입력 계층, 하이퍼박스 계층, 클래스 계층으로 구성되어 있다. 입력 계층에는  $n$  개의 입력 노드들이 포함되어 있는데 각 노드는  $n$  개의 특징 패턴을 입력하게 된다. 하이퍼박스 계층은  $m$  개의 하이퍼박스 노드로 구성된다. 하이퍼박스 노드  $B_i$  은 단 하나의 클래스 노드에 연결되는데  $n$  개의 입력 노드를 위한  $n$  개 퍼지집합을 가진다.  $B_i$  의  $i$  번째 퍼지집합은  $B_i^j$  로 표시되는데 (그림 2)에서 보는 바와 같이 3개의 가중 소속함수를 가진다. 출력 계층은  $p$  클래스 노드들로 구성된다. 각기 클래스 노드는 하나 혹은 하나 이상의 하이퍼박스 노드에 연결되어 있다.  $h$  번째 입력 패턴은  $I_h = \{A_h = (a_1, a_2, \dots, a_n), class\}$  로 기록된다. 여기서  $class$  는 진단 결과이고  $A_h$  는 상이한  $n$  개 특징에 대한 패턴이다. 미상의 특징은 NULL로 표기된다.



(그림 1) 가중 퍼지 소속함수 신경망의 구조

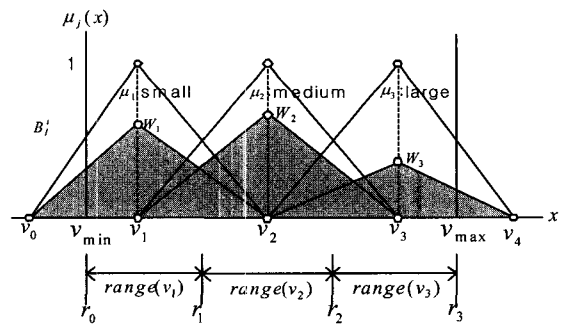
**2.2 정의 및 오퍼레이션(Definitions and Operations)**

①  $w_{ii}$  : 하이퍼박스 노드  $B_i$  과 클래스 노드  $C_i$  사이의 연결 가중치는 초기값이 0으로 설정된  $w_{ii}$ 로 표시된다. 만

일 하이퍼박스 노드  $B_i$  과 클래스 노드  $C_i$  가 연결되어 있다면  $w_{ii}$ 는 0에서 1로 설정된다.  $C_i$  는 하이퍼박스 노드로부터 하나 이상의 연결을 가질 수 있는 반면  $B_i$  은 클래스 노드에 하나만의 연결을 가지도록 제한 되어 있다.

- ②  $v_i$  :  $v_1, v_2, v_3$  는 (그림 2)에서와 같이 대, 중, 소의 소속함수의 중심점을 각각 나타낸다. 중심점들은 학습 수행 중에 조정되며 중심점  $v_0$  와  $v_4$  는 고정되어 있다. 입력  $a_i$  는 (그림 2)의  $v_{min} \sim v_{max}$  범위 내로 한정한다.
- ③  $\mu_j$  및  $W_j$  :  $\mu_j$  는 하이퍼박스 노드  $B_i$  의  $i$  번째 집합의 소속함수인데  $j = 1, 2, 3$  은 대, 중, 소를 각각 나타낸다. 각 소속함수  $\mu_j$  의 모양은 3개의 정점( $v_{j-1}, v_j, v_{j+1}$ )을 가지는 삼각형으로 되어 있는데 학습에 의한 소속함수의 강도(strength)를 나타내는 소속함수 가중치  $W_j$  ( $0 \leq W_j \leq 1$ , 초기값  $0.45 \leq W_j \leq 0.55$ )를 갖는다. (그림 2)에서 가중 퍼지소속함수인 회색 부분의 삼각형은  $v_{j-1}, W_j, v_{j+1}$  로 그려진다.
- ④  $Len(\mu_j)$  :  $\mu_j$  의 크기를 측정하기 위한 길이는  $Len[8]$  으로 정의된다.

$$Len(\mu_j) = ((v_{j+1} - v_{j-1})/2)W_j \quad (1)$$



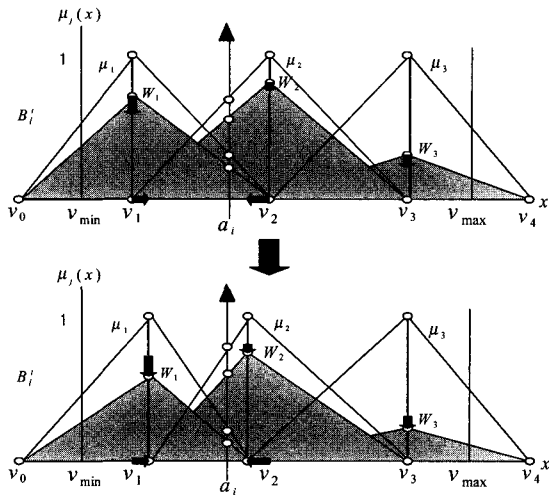
(그림 2)  $B_i$  의  $i$  번째 가중 퍼지 소속함수 집합

⑤  $Adjust(B_i)$  : 이 오퍼레이션은 입력  $A_h = (a_1, a_2, \dots, a_n)$  에 의해 하이퍼박스 노드  $B_i$  의 퍼지집합의 소속함수와 그 가중치를 조정한다. 입력  $a_i$  를 갖는  $i$  번째 소속함수와  $v_j$  와  $W_j$  의 소속함수는  $\mu_j(a_i)$  값에 의하여 조정된다( $j = 1, 2, 3$ ).  $Adjust(B_i)$  오퍼레이션 결과 새로운 정점  $new(v_j)$ 와 새로운 가중치  $new(W_j)$ 가 다음과 같은 표현으로 설정된다.  $\alpha$  와  $\beta$  는 0~1사이의 학습율범위이고 변수  $E_j$  는  $v_j$  와 입력  $a_i$  간의 차이를 나타낸다. 만일  $E_j$  가 인접  $E_{j+1}$  보다 크면 적은 것이 선택된다.

$$new(v_j) = v_j \pm \alpha E_j \mu_j(a_i) W_j \quad (2)$$

$$new(W_j) = W_j + \beta (\mu_j(a_i) - W_j) \quad (3)$$

자세한  $Adjust(B_i)$ 의 오퍼레이션과정은 다음장의 NN-WFM 알고리즘에서 기술한다. (그림 3)은 입력  $a_i$ 와  $B_i$ 의  $i$ 번째 집합의 가중 퍼지소속함수를 위한  $Adjust(B_i)$  오퍼레이션의 예를 보여주고 있다.



(그림 3)  $Adjust(B_i)$  연산 전후의 예

⑥  $Random(B_i)$ : 이 오퍼레이션으로 하이퍼박스  $B_i$ 이 만들어 지는데 이 것은 0.45~0.55 범위의 임의의 가중치를 가진 함수로서  $n$ (입력 수)개의 소속함수를 가진다. 임의의 중심점  $v_i$ 는 아래에서와 같은 범위 내이어야 한다.

$$r_{i-1} \leq v_i < r_i, \text{ 여기서 } i = 1, 2, 3$$

한편  $v_0$ 와  $v_4$ 는 (그림 2)에서와 같이 고정되어 있다. 또한 연결 가중치  $w_{ii}$ 는 0으로 설정되어 있다.

⑦  $Output(B_i)$ : 하이퍼박스  $B_i$ 에  $n$ 개 특징을 가진  $h$ 번째 입력  $A_h = (a_1, a_2, \dots, a_n)$ 에 의하여  $B_i$ 의 출력은 다음 식과 같이 계산된다.

$$Output(B_i) = \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^3 B_i^j(\mu_j(a_i)) W_j. \quad (4)$$

### 2.3 NNWFM 학습을 위한 알고리즘

이 장에서는 NNWFM이라는 학습을 위한 알고리즘을 기술한다. 알고리즘은  $Learning(B_i, C_i)$ 프로시저를 이용하여 정점 및 가중치의 위치를 조정하고 하이퍼박스 노드를 클래스 노드에 연결한다.

#### Algorithm NNWFM ;

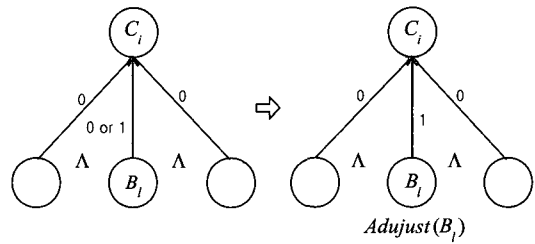
```

1 While (result is satisfied)
1.1 for  $l=1$  to  $m$  //  $m$  is number of hyperboxes, usually start from number input
1.1.1  $Random(B_l)$  ;
1.1.2 for  $j=1$  to  $p$  //  $p$  is number of class nodes
1.1.2.1  $w_{ij} = 0$  ; // initial connection weight between  $B_l$  and  $C_j$ 
1.2 for  $k=1$  to  $h$  //  $h$  is number of input patterns
1.2.1 find  $B_l$  that has the maximum value of  $EnhOutput(B_l)$  among  $m$  hyperbox nodes from the input  $A_k$  ;
// input vector :  $I_k = \{A_k = (a_1, a_2, \dots, a_n), diagnosis\}$ 
1.2.2  $Learning(B_l, c_i)$  ; //  $C_i$  is a diagnosis in  $I_k$ 
    
```

#### Procedure Learning( $B_l, C_i$ ) ;

```

//  $m$  is number of hyperboxes
1 Case 1 :  $\forall m, w_{mi} = 0$ , where  $m \neq l$  ;
1.1  $w_{ii} = 1$  ;
1.2  $Adjust(B_l)$  ;
    
```

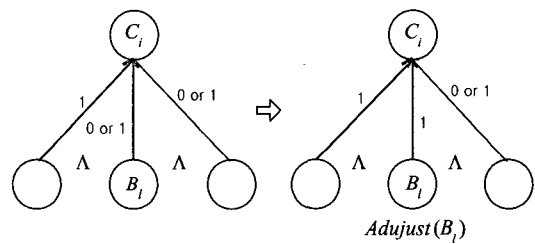


(그림 4) Case 1 :  $Adjust$

#### 2 Case 2 : $\exists m$ satisfying $w_{mi} = 1$ , where $m \neq l$

```

2.1  $w_{ii} = 1$  ;
2.2  $Adjust(B_l)$  ;
    
```



(그림 5) Case 2 :  $Adjust$

#### Procedure Adjust( $B_l$ ) ;

```

1 for  $i=1$  to  $n$  // for each  $i$ th set of membership function in  $B_l$ 
1.1 for  $j=1$  to 3 // for each membership function
1.1.1 if  $v_{j-1} \leq a_j < v_j$  // for left side of  $\mu_j$ 
1.1.1.1  $E_j = \min(|v_j - a_j|, |v_{j-1} - a_j|)$  ;
1.1.1.2  $new(v_j) = v_j - \alpha E_j \mu_j(a_i) W_j$  ;
1.1.2 else  $v_j \leq a_j < v_{j+1}$  // for right side of  $\mu_j$ 
1.1.2.1  $E_j = \min(|v_j - a_j|, |v_{j+1} - a_j|)$  ;
1.1.2.2  $new(v_j) = v_j + \alpha E_j \mu_j(a_i) W_j$  ;
1.1.3  $new(W_j) = W_j + \beta (\mu_j(a_i) - W_j)$  ;
    
```

2.4 퍼지규칙 추출(Fuzzy Rule Extraction)

학습을 거친 NNWFM은 입력패턴을 분류하기 위한 if-then 형식의 퍼지규칙 추출에 사용될 수 있다. 학습을 거친 후에는 하이퍼박스 노드  $B_i$ 의 각 퍼지 집합에는 3개의 가중 퍼지소속함수(weighted fuzzy memberships, WFM, (그림 6)의 회색 부분의 소속함수)가 포함된다. 규칙들은 WFM으로부터 직접 추출될 수 있다. 규칙 추출 전략(rule extraction strategy)을 아래와 같이 제안한다.

- ①  $\mu_b^i(x)$ (그림의 굵은 선)로 표시된 WFM의 bounded sum은 다음과 같이 정의된다. WFM의 bounded sum은 (그림 6)에서와 같이 3개의 WFM의 퍼지 특성을 통합한다.

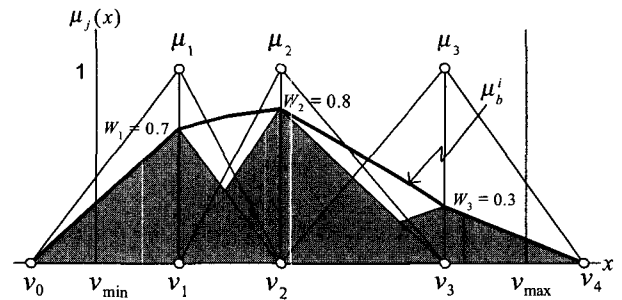
$$\mu_b^i(x) = \sum_{j=1}^3 B_j^i(\mu_j(x)). \quad (5)$$

- ②  $Output(B_i)$  값을 얻기 위해 수행하는 추론 메커니즘은 Takagi & Sugeno 모델[13]을 사용한다. 이에 더하

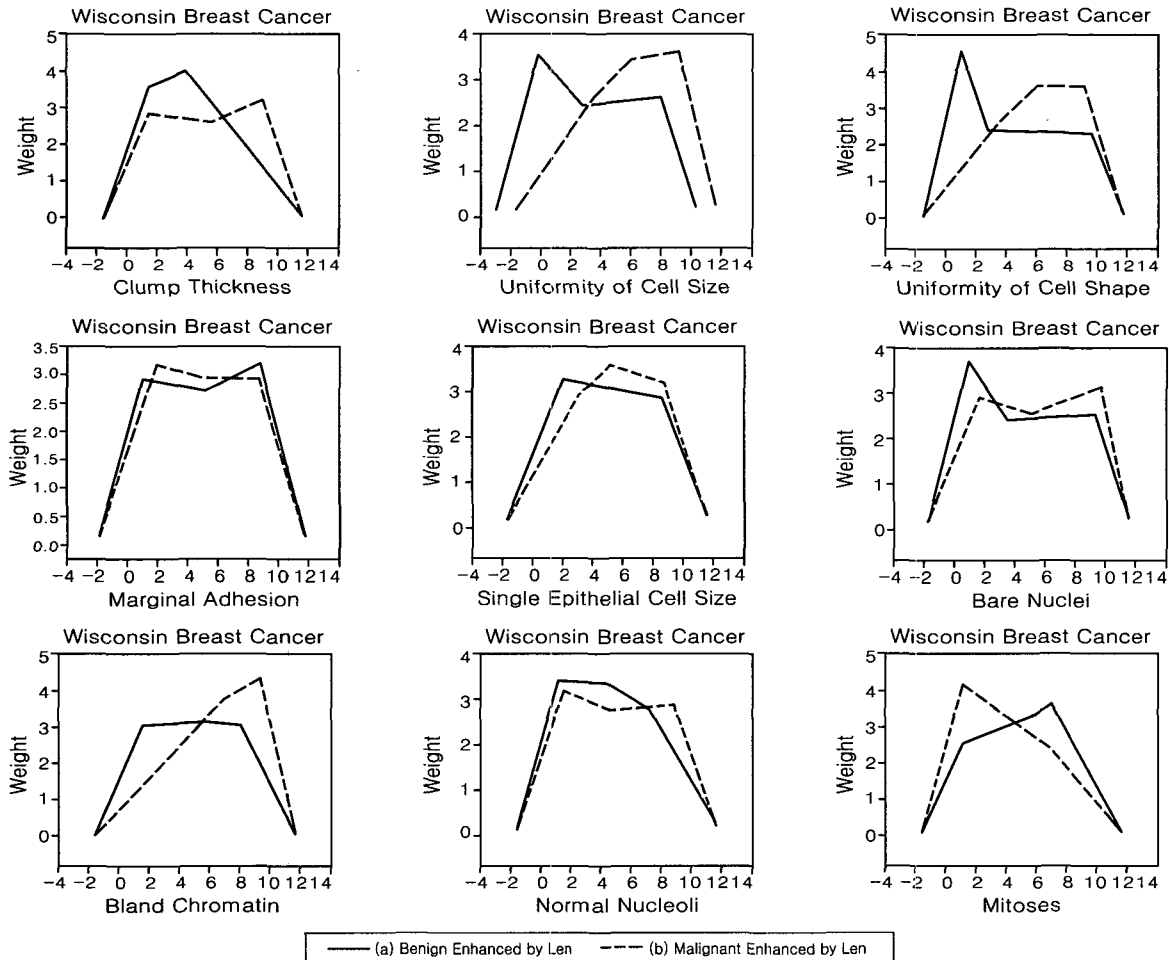
여 추론 성능을 강화하기 위하여  $Len(\mu_j)$  휴리스틱을 사용한  $EnhOutput(B_i)$ 을 다음과 같이 제안한다.

$$EnhOutput(B_i) = \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^3 B_j^i(\mu_j(a_i)) W_j / Len(\mu_j) \quad (6)$$

(그림 6)의 굵은 선들은 식 (6)을 이용한 Enhanced BS-WFM이다.  $Len(\mu_j)$  휴리스틱은  $Len(\mu_j)$ 이 적을수록 보다 많은 집합 정보를 가진다는 개념에 기초를 둔 것이다.



(그림 6) 3개 가중 소속함수의 Bounded Sum 결과 예(굵은 선)



(그림 7) Wisconsin Breast Cancer의 Enhanced BSWFMs

<표 1>은 (그림 7)의 enhanced BSWFMs로부터 얻어낸 WBC 분류를 위한 퍼지규칙을 보여준다. <표 2>는 WBC분류에 있어서 본 논문에서 제시한 퍼지규칙과 그 밖의 분류 기법[10, 11]과의 비교 결과를 보여주고 있다.

<표 1> Wisconsin Breast Cancer 규칙

	CT	CSi	CSh	MA	SECS	BN	BC	NN	Mi	Class
rule 1	(a)	(a)	(a)	(a)	(a)	(a)	(a)	(a)	(a)	Benign
rule 2	(b)	(b)	(b)	(b)	(b)	(b)	(b)	(b)	(b)	Malignant

(a)s and (b)s are the Enhanced BSWFMs in (Figure 8).

<표 2> Wisconsin Breast Cancer 인식 결과 비교

Algorithm	Recognition Rate	No. of Rules
Nauck [10]	96.5%	4
Gomez [2]	99.12%	11
NNWFM	99.41%	2

### 3. 실험 결과(Experimental Results)

이번 장에서는 NNWFM에 의해 추출된 Wisconsin breast cancer 진단 퍼지규칙을 보여주고 이의 성능을 인식율을 통해 평가한다. NNWFM으로 추출된 2개의 퍼지규칙은 99.41%의 인식율을 가지며 이는 퍼지규칙의 수와 인식율에 있어 지금까지의 다른 논문의 결과보다 우수함을 보여주고 있다.

#### 3.1 Wisconsin Breast Cancer 분류

Wisconsin breast cancer 퍼지규칙 인식을 실험에서는 위스칸신 대학병원에서 수집한 Wisconsin breast cancer(WBC)의 데이터베이스[17]를 사용하였다. 이 자료는 WBC 진단을 위한 9개의 특징값과 양성 혹은 악성의 진단 클래스로 구성되어 있다. 이 데이터베이스는 699개의 레코드를 갖고 있지만 16개 레코드는 특징값이 몇 개 빠져 있어 이를 제외한 683개 레코드를 사용하였다. (그림 7)은 NNWFM에 의한 학습 과정을 거친 강화된 가중 퍼지소속함수(enhanced BSWFMs)이다.

### 4. 결론(Concluding Remarks)

본 논문은 NNWFM을 이용하여 Wisconsin breast cancer 진단을 위한 퍼지규칙을 제안하고 있다. 추출된 퍼지규칙은 그 인식율에 있어서 타 논문에서 보여준 인식율 보다 높은 결과인 99.41%를 보여주고 있으며 퍼지규칙의 수도 가장 적은 2개로 수행한다. 이러한 결과는 학습방법에 있어 하이퍼파스 노드  $B_i$ 의 퍼지집합은 대, 중, 소의 가중 퍼지소속함수를 사용하고, 학습 후 BSWFM에 의해 통합된 하나의 가중 퍼지소속함수로 입력패턴의 특징을 표현하기 때문이다. 이 BSWFM은

퍼지 소속함수의 범위인 0~1이 아니라 각 특징의 중요도에 따라 0에서 1 이하인 값을 가짐으로써 퍼지 소속함수에 가중치를 부여하였다는 점이 분류에 대한 분별력을 강화하였다. 또한 enhanced BSWFM은 입력패턴의 특징이 집중된 BSWFM을 증폭하고 특징이 분산된 BSWFM을 약화시킴으로써 인식율을 더욱 향상 시키는 결과를 주고 있다.

### 참 고 문 헌

- [1] G. A. Carpenter, S. Grossberg and J. H. Reynolds, "ART-MAP : Supervised real-time learning and classification of nonstationary data by a self-organizing neural network," Neural Networks, Vol.4, pp.565-588, 1991.
- [2] A. F. Gomez-Skarmeta, M. V. F. Jimenez, J. G. Marin-Blazques, "Approximative Fuzzy Rules Approaches for Classification with Hybrid-GA Techniques," Information Sciences, Vol.136, pp.193-214, 2001.
- [3] H. Ishibuchi and T. Nakashima, "Voting in Fuzzy Rule-Based Systems for Pattern Classification Problems," Fuzzy Sets and Systems, Vol.103, pp.223-238, 1999.
- [4] R. Jang, "ANFIS : Adaptive network-based fuzzy inference system," IEEE Trans. Syst., Man, Cybern., Vol.23, pp.665-685, May-June, 1993.
- [5] C. F. Juang and C. T. Lin, "An On-Line Self-Constructing Neural Fuzzy Inference Network and Its Applications," IEEE Trans. Fuzzy Systems, Vol.6, No.1, pp.12-32, 1998.
- [6] N. Kasabov, Foundation of Neural Networks, Fuzzy Systems and Knowledge Engineering, The MIT Press, Cambridge, MA, 1996.
- [7] T. Kasuba, "Simplified Fuzzy ARTMAP," IEEE AI Expert, pp.19-25, Nov., 1993.
- [8] H.-M. Lee, K.-H. Chen and I.-F. Jiang, "A Neural Networks with Disjunctive Fuzzy Information," Neural Networks, Vol.11, pp.1113-1125, 1998.
- [9] C. T. Lin and C. S. George Lee, "Neural-network-based fuzzy logic control and decision system," IEEE Trans. Computers, Vol.40, No.12, Dec., 1991.
- [10] D. Nauck and R. Kruse, "A Neuro-Fuzzy Method to Learn Fuzzy Classification Rules from Data," Fuzzy Sets and Systems, Vol.89, pp.277-288, 1997.
- [11] M. Setnes and H. Roubos, "GA-Fuzzy Modeling and Classification : Complexity and Performance," IEEE Trans., Fuzzy Systems, Vol.8, No.5, pp.509-522, 2000.
- [12] P. Simpson, "Fuzzy min-max neural networks-Part 1 : Classification," IEEE Trans., Neural Networks, Vol.3, pp. 776-786, 1992.
- [13] T. Takagi and M. Sugeno, "Fuzzy Identification of Systems and Its Application to Modeling and Control," IEEE Trans., Syst., Man, Cybern., Vol.15, pp.116-132, 1985.
- [14] K. Tanaka, M. Sano and H. Watanabe, "Modeling and Control of Carbon Monoxide Concentration Using a Neuro-Fuzzy technique," IEEE Trans. Fuzzy Systems, Vol.3, pp.271-279, June, 1995.

- [15] C. Z. Ye, J. Yang, D. Y. Geng, Y. Zhou, N. Y. Chen, Fuzzy Rules to Predict Degree of Malignancy in Brain Glioma, Medical and Biological Engineering and Computing, Vol.40, 2002.
- [16] J. S. Wang and C. S. G. Lee, "Self-Adaptive Neuro-Fuzzy Inference System for Classification Applications," IEEE Trans., Fuzzy Systems, Vol.10, No.6, pp.790-802, 2002.
- [17] W. Wolberg, O. Mangasarian, "Multisurface Method of Pattern for Medical Diagnosis Applied to Breast Cytology," Proc. National Academy of Sciences, Vol.87, pp.9193-9166, 1990.



### 임준식

e-mail : js.im@kyungwon.ac.kr

1986년 인하대학교 전자계산학과(학사)

1989년 University of Alabama at Birmingham, Dept. of Computer Science(M.S)

1994년 Louisiana State University, Dept. of Computer Science(Ph.D.)

1995년~현재 경원대학교 소프트웨어학부 부교수

관심분야 : neuro-fuzzy systems, biomedical prediction systems, human-centered systems