

# 가중치를 갖는 FMM신경망과 패턴분류를 위한 특징분석 기법

(A Weighted FMM Neural Network and Feature Analysis  
Technique for Pattern Classification)

김 호 준 \* 양 현 승 \*\*

(Ho-Joon Kim) (Hyun-Seung Yang)

**요약** 본 논문에서는 패턴 분류를 위한 수정된 퍼지 최대최소 신경망 모델을 제안하고 그의 유용성을 고찰한다. 이를 위하여 하이퍼박스 내에서 각 특징들에 대하여 가중치 요소를 갖는 새로운 하이퍼큐브 소속함수를 정의한다. 이 가중치 요소는 분류과정에서 임의의 클래스에 대한 각 특징의 상대적인 기여도를 반영한다. 본 연구에서는 이를 위하여 새롭게 정의된 하이퍼박스 생성, 확장 및 축소의 3단계로 이루어지는 학습 방법론을 소개한다. 또한 제안된 모델을 기반으로 하여 학습된 분류기로부터 하이퍼박스 소속함수와 연결가중치를 사용하여 주어진 클래스에 대한 특징의 연관도를 산출하는 형태의 이른바 특징 분석 기법을 제안한다. 이를 위하여 세부적으로 각 특징에 대하여 연관도 척도와 퍼지 소속함수간의 유사도 척도를 정의한다. 또한 실제 패턴 분류문제에 적용한 실험결과를 통하여 제안된 이론의 타당성을 평가한다.

**키워드** : 패턴분류, 퍼지신경망, 규칙추출

**Abstract** In this paper we propose a modified fuzzy min-max neural network model for pattern classification and discuss the usefulness of the model. We define a new hypercube membership function which has a weight factor to each of the feature within a hyperbox. The weight factor makes it possible to consider the degree of relevance of each feature to a class during the classification process. Based on the proposed model, a knowledge extraction method is presented. In this method, a list of relevant features for a given class is extracted from the trained network using the hyperbox membership functions and connection weights. For this purpose we define a Relevance Factor that represents a degree of relevance of a feature to the given class and a similarity measure between fuzzy membership functions of the hyperboxes. Experimental results for the proposed methods and discussions are presented for the evaluation of the effectiveness and feasibility of the proposed methods.

**Key words** : pattern classification, fuzzy neural networks, rule extraction

## 1. 서론

패턴 분류 문제에서 특징의 종류에 따라 상대적 중요도를 판별하는 일과 개별 특징값이 특정 클래스의 판별 과정에 주는 영향을 분석하는 일은 분류기의 성능을 향상시키는 측면에서 매우 중요한 의미를 갖는다. 이를 위하여 많은 이론과 응용시스템을 위한 방법론이 소개된 바 있는데 최근의 연구에서는 인공 신경망을 사용하여

패턴 분류의 성능을 향상시키는 연구와 분류에 관한 명시적 지식표현을 위한 규칙생성 기법 등이 활발하게 연구되고 있다[1-12].

퍼지 최대최소 신경망(Fuzzy Min-Max(FMM) Neural Network)은 다차원 특징공간상의 특정 영역으로서 정의되는 하이퍼박스 퍼지집합 개념을 사용한 패턴분류 모형이다[2]. 이는 Simpson 등에 의해 제안된 신경망 모델로서 구조적 특성이 간결하면서도 강력한 학습기능을 지원한다. 그 후 Gabrys 등은 FMM 신경망의 내부 연산과 데이터의 표현형태를 확장하여 일반화된 모델을 제안하였다[4]. 이들 방법에서는 분류과정에서 입력 특징값 및 내부적으로 연산되는 데이터의 표현형태를 특징공간의 각차원에서 퍼지구간의 형태로 일반화하고, 네

\* 종신회원 : 한동대학교 전산전자공학부 교수  
hjkim@hadong.edu

\*\* 종신회원 : 한국과학기술원 전산학과 교수  
hsyang@cs.kaist.ac.kr

논문접수 : 2004년 8월 18일

심사완료 : 2004년 11월 9일

트위크의 활성화 특성은 이러한 각 구간값의 집합으로 정의되는 하이퍼박스(hyperbox)에 대하여 소속함수(membership function)값으로 산출되도록 하였다.

기존의 FMM 신경망 분류기는 간결한 학습기능을 지원한다는 장점이 있지만 그 활성화 특성과 학습과정의 특성상 주어진 학습패턴에서 특징값의 발생 빈도를 고려하여 반영하기가 어렵다. 예컨대, 많은 수의 표준패턴에 의해 학습된 하이퍼박스의 형태가 단 한개의 노이즈에 의해서도 왜곡될 수 있으며, 분류과정에서 특징값의 중요도를 항상 동일 값으로 고려함으로써 빈도에 따른 상대적 중요도의 차이를 효과적으로 반영할 수 없다. 이에 본 연구에서는 이러한 FMM 신경망의 구조와 동작 특성에 가중치 요소(weight factor)를 추가하여 정의함으로써 학습과정에서 특징값의 발생빈도의 차이를 고려할 수 있도록 하였다. 이를 위하여 새롭게 정의된 하이퍼박스 생성, 확장 및 축소의 3단계로 이루어지는 학습방법론을 소개한다. 또한 제안된 모델을 기반으로 하여 학습된 분류기로부터 하이퍼박스 소속함수와 연결가중치를 사용하여 주어진 클래스에 대한 특징의 연관도를 산출하는 형태의 이른바 특징 분석 기법을 제안한다. 이를 위하여 세부적으로 각 특징에 대하여 연관도 척도와 퍼지 소속함수간의 유사도 척도를 정의한다. 또한 잘 알려진 패턴 분류문제인 아이리스 데이터와 의료진단 데이터 분류문제에 적용한 실험결과를 통하여 제안된 이론의 타당성을 평가한다.

본 논문의 2장에서서는 배경연구로서 기존의 FMM 신경망에 대하여 소개하며 이어서 3장에서는 본 연구에서 제안하는 패턴분류 모델과 학습방법 및 특징 분석 기법에 관해 기술하며 4장에서는 제안된 모델에 대하여 기존의 FMM모델을 사용한 경우와의 성능 비교를 통하여 그 이론의 타당성과 유용성을 평가, 고찰한다.

## 2. 퍼지 최대최소 신경망

FMM신경망은 n차원 특징 공간상에서 하이퍼박스 퍼지 집합으로 구성되는 패턴분류 모델이다. 하이퍼박스를 모든 특징의 종류에 대하여 완전 클래스 멤버십을 가진 패턴들로 이루어진 n차원 패턴 공간 영역이다. FMM 신경망의 활성화 특성은 하이퍼박스를 정의한 특징공간상의 각 차원에 대하여 특징범위의 가장자리에서 퍼지 멤버십을 갖는 이른바 퍼지 구간(fuzzy interval)들로서 정의 된다. FMM 신경망을 이용한 패턴 분류 과정은 그림 1과 같이 이루어진다.

FMM 신경망은 그림에 보인 바와 같이 총 3개의 계층으로 구성되는데, 입력층의 각 노드에 할당되는 특징값으로부터 하이퍼박스 멤버십함수값을 산출하여 중간층 노드의 출력값을 결정한다. 이들 중 최대로 반응한

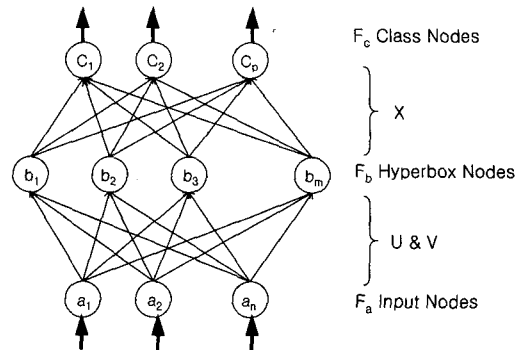


그림 1 FMM 신경망의 구조

하이퍼박스로부터 최종 분류패턴을 판정하게 된다.

임의의 하이퍼박스  $j$ 의 소속함수  $b_j(A_h)$ 는 식 (1)과 같이 산출된다. 이는 학습과정으로 결정된 각 특징범위의 가장자리에서 퍼지 멤버십을 갖는 형태이다.

$$b_j(A_h) = \frac{1}{2^n} \sum \{ \max(0, 1 - \max(0, \gamma \min(1, a_{hi} - v_{ji}))) + \max(0, 1 - \max(0, \gamma \min(1, u_{ji} - a_{ij}))) \} \quad (1)$$

식에서  $A_h = (a_{h1}, a_{h2}, \dots, a_{hn}) \in I^n$ 는  $h$  번째 입력패턴으로 총  $n$ 개의 특징값들로 이루어진다. 또한 각 특징범위의 최소점과 최대점은  $U_j = (u_{j1}, u_{j2}, \dots, u_{jn})$ 와  $V_j = (v_{j1}, v_{j2}, \dots, v_{jn})$ 으로 표현된다.  $\gamma$ 는 특징범위의 가장자리에서 퍼지 소속함수의 기울기를 결정하는 매개변수로 0과 1사이의 값을 갖는다.

FMM 신경망의 학습은 주어지는 학습 패턴에 대하여 하이퍼박스의 생성(creation), 확장(expansion) 및 축소(contraction)의 3가지 과정으로 이루어진다.

새로운 학습패턴에 대하여 기존 하이퍼박스의 확장여부는 식 (2)의 조건에 의해서 결정된다. 즉 새로운 입력에 의해서 평균적으로 확장되는 크기의 범위가 일정 기준보다 작을때에만 확장함을 의미하며 이를 파라미터  $\theta$ 로써 제어한다.

$$n\theta \geq \sum_{i=1}^n (\max(v_{ji}, x_{hi}) - \min(u_{ji}, x_{hi})) \quad (2)$$

하이퍼박스 확장과정에서 새롭게 정의되는 특징범위는 식 (3)과 같이 결정된다.

$$\begin{aligned} u_{ji}^{new} &= \min(u_{ji}^{old}, x_{hi}) & \forall i=1, 2, \dots, n \\ v_{ji}^{new} &= \max(v_{ji}^{old}, x_{hi}) & \forall i=1, 2, \dots, n \end{aligned} \quad (3)$$

식에서  $x_{hi}$ 는 현재의 학습패턴에서 주어지는  $i$ 번째 특징 값을 의미한다. 즉 확장되는 하이퍼박스의 특징범위는 새롭게 주어진 특징값을 포함하여 최대 및 최소치를 재조정하게 된다.

### 3. 가중치를 갖는 FMM 신경망 모델

#### 3.1 개요 및 동작 특성

FMM 신경망은 매우 간결하면서도 강력한 학습기능을 지원한다. 그러나 식 (1)에 정의한 활성화 특성에서 알 수 있듯이 주어진 학습패턴 집합에서 관찰되는 특징값의 범위만을 고려할 뿐 특징값의 발생 빈도를 학습과정에 반영하지 못한다. 이는 노이즈나 비정상적인 패턴이 학습패턴에 포함되는 경우 성능저하를 보일 수 있다. 이에 본 연구에서는 가중치 개념이 추가된 FMM신경망, 소위 WFMM(Weighted Fuzzy Min-Max) 신경망을 제안하고 이의 활용기법을 고찰한다. 이는 하이퍼박스의 소속함수에 가중치 개념을 추가하여 특징값의 발생빈도에 따라 학습되게 함으로써 빈도요소를 반영하는 패턴분류를 가능하게 하며, 나아가 주어진 문제에서 각 특징들의 중요도를 평가하기 위한 기법으로서의 활용될 수 있다. 다시 말해서 기존의 FMM에서 항상 균등하게 고려되었던 특징과 하이퍼박스와의 관계를 학습패턴이 발생 빈도에 따라 가변적으로 고려할 수 있게 하여 특징의 상대적 중요도를 서로 다른 값으로 반영할 수 있게 한다.

제안된 모델에서 하이퍼박스의 활성화 특성은 식 (4)와 같은 멤버십 함수로서 정의된다.

$$b_j(A_h) = \frac{1}{\sum_{i=1}^n w_{ji}} \cdot \sum_{i=1}^n w_{ji} \cdot [\max(0, 1 - \max(0, \gamma \min(1, a_{hi} - v_{ji}))) + \max(0, 1 - \max(0, \gamma \min(1, u_{ji} - a_{hi}))) - 1.0] \quad (4)$$

식에서  $w_{ji}$ 는  $j$  번째 하이퍼박스와  $i$  번째 특징사이의 연결 가중치를 의미하며, 나머지 표기는 앞 절에서 수식 (1)의 경우와 동일하다. 이는 기존 FMM신경망의 특성에 비하여 두 가지 요소가 수정된 것이다. 하나는 가중치 요소가 각 특징값에 반영된다는 점이고 두 번째는 각 하이퍼박스의 출력값 산출 방식을 퍼지범위의 최소 조건 항과 최대조건 항의 보수(complement) 합계에서 1.0을 빼줌으로써 출력값의 범위가 {0,1} 사이가 되도록 조정하였다. 그러므로 평균값 산출을 위한 계수는 멤버십 값을 정규화 하도록 총 가중치합으로 대치하였다.

#### 3.2 학습 알고리즘

제안된 신경망은 학습과정에서 각 특징에 대한 가중치 값이 조정되는데, 이때 기존의 FMM 신경망에서 이루어지는 생성, 확장, 축소의 과정에 대하여도 가중치 요소의 특성을 효과적으로 반영하도록 수정, 개선하였다. 각 학습단계에서, 임의의 하이퍼박스  $j$ 에서 임의의 특징  $i$ 에 대한 가중치 조정은 식 (5)와 (6)의 형태로 이루어진다.

$$w_{ji}^{new} = w_{ji}^{old} + \Delta w_{ji} \quad (5)$$

$$\Delta w_{ji} = \begin{cases} \lambda \\ d \cdot (T - \frac{v^{new} - u^{new}}{v^{old} - u^{old}}) \\ MAX(w_{ji}^{old} \cdot (\frac{v^{old} - u^{old}}{v^{new} - u^{new}} - 1.0), \frac{-w_{ji}^{old}}{2}) \\ \text{if } (v^{new} - u^{new} \leq s) \\ \text{elseif } (\frac{v^{new} - u^{new}}{v^{old} - u^{old}} \leq T) \\ \text{otherwise} \end{cases} \quad (6)$$

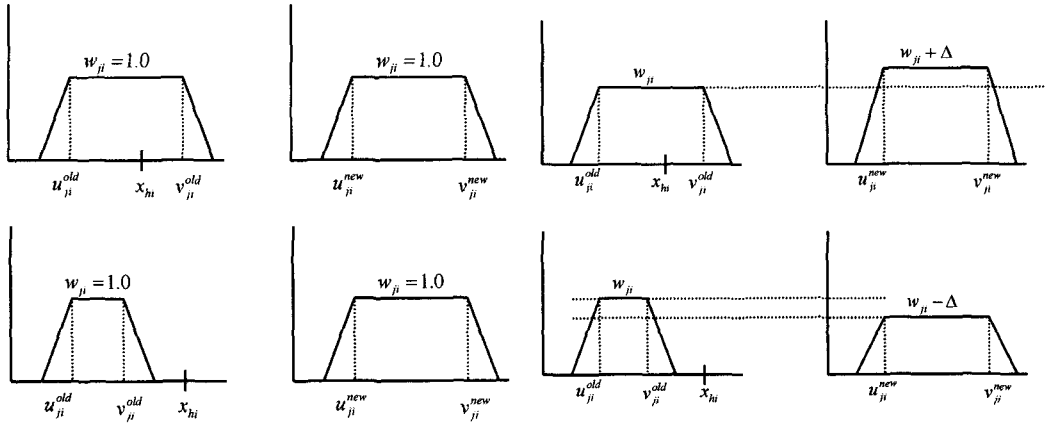
식 (5)와 (6)에 의하면, 새롭게 생성된 하이퍼박스의 최소점과 최대점이 일정값 이내의 작은 범위이거나 이전 하이퍼박스에 비해 크기가 일정크기 이내로 확장된 경우 가중치가 증가되며, 하이퍼박스의 크기가 일정크기 이상으로 확장된 경우 가중치가 감소함을 알 수 있다. 이때 증가되는 양은 확장되는 범위에 반비례한 값으로 결정되게 된다. 수식에서 매개변수  $s$ 는 하이퍼박스의 가중치의 증감을 결정하기 위한 기준값이 된다. 이는 적용되는 응용에서 다수의 실험을 통하여 설정하는 파라미터 값으로, 실험적으로 최적의 성능을 기대하는 하이퍼박스의 기준 크기를 의미한다. 또한 파라미터  $T$ 는 1보다 큰 1근처의 값으로 기준크기 범위의 가장자리에서 퍼지 멤버십의 형태로 점진적으로 가중치 증가분을 감소시키기 위한 범위이다. 파라미터  $d$ 는 학습률(learning rate) 상수인데 특징영역이 확장함에 따라 해당 특징에 대한 가중치가 얼마나 급격히 증가시킬 것인가를 결정하는 요소이다.

각 가중치 값은 하이퍼박스가 생성되는 시점에 초기 값 1.0으로 설정된다. 식 (4)에서 가중치 값이 항상 1.0이라면 이는 기존 FMM의 특성과 유사해 지는데, 이는 제안된 모델이 기존 FMM 모델의 특성을 좀 더 일반화한 개념임을 의미한다.

그림 2는 FMM모델과 제안된 모델의 학습과정에서 하이퍼박스에 대한 확장과정을 비교하여 보인 것이다. 그림에 보인 바와 같이 기존의 FMM신경망에서는 학습패턴 집합에서 동일 특징범위내에 패턴이 몇 개가 나타나는데에 관계없이 동일 가중치로 고려된다. 반면 제안된 모델에서는 특정범위 내에서 패턴이 중복하여 나타나면 가중치를 증가시키며, 확장과정에서 특징데이터에 의하여 범위가 넓어지면 가중치를 감소시킨다.

제안된 모델에서 하이퍼박스 확장과정은 식 (7)과 같은 식으로서 이루어진다. 수식에서  $\theta$ 는 앞서 언급한 바와 같이 하이퍼박스의 확장 또는 생성여부를 결정하는 파라미터이다.

$$n\theta \geq \sum_{i=1}^n (\max(v_{ji}, x_{hi}) - \min(u_{ji}, x_{hi})) \quad (7)$$



(a) FMM 신경망 모델 (b) 제안된 모델  
그림 2 하이퍼박스 확장에 따른 가중치값 조정 비교

이면

$(x_{hi} < u_{ji}^{old})$ 의 경우

$$u_{ji}^{new} = u_{ji}^{old} - \frac{1}{w_{ji}}(u_{ji}^{old} - x_{hi})$$

$(x_{hi} > v_{ji}^{old})$ 의 경우

$$v_{ji}^{new} = v_{ji}^{old} + \frac{1}{w_{ji}}(x_{hi} - v_{ji}^{old})$$

이러한 방법이 기존의 FMM 모델에서의 확장방법에 비해 다른 점은 많은 빈도를 갖는 특징범위를 한개의 관찰된 패턴으로 확장할 때 한꺼번에 전 범위로 확장하지 않고 점진적으로 확장한다는 점이다. 이는 기존의 FMM 모델에 비해 다음과 같이 두 가지 장점을 기대할 수 있게 한다는 것이다.

첫번째로 학습패턴에 포함될 수 있는 노이즈나 비정상적인 과도데이터에 의한 영향을 줄일 수 있다. 데이터 집합내에서 상대적으로 빈도가 적은 패턴에 대해 하이퍼박스 영역이 한꺼번에 확장되는 것을 방지하며, 또한 범위가 늘어나는 경우 가중치 값을 재조정하여 이러한 비정상 패턴에 민감하게 학습되는 현상을 보완한다. 즉 학습과정 중 비정상적인 데이터에 의하여 어떤 특징값의 범위가 큰 범위로 증가하는 경우 한꺼번에 확장하지 않고 점진적으로 범위를 증가시키도록 하는 방법을 채택함으로써 기존의 여러 번의 학습과정을 통해서 조정된 결과가 한 개의 잘못된 비정상적 데이터로 인하여 크게 왜곡되는 것을 방지할 수 있게 한다.

두 번째로 특징에 대한 분석기능을 지원한다는 점이다. 가중치 값으로부터 각 특징과 클래스간의 관계를 해석해 낼 수 있으며, 이는 특징종류에 대한 유용성과 특징값에 대하여 하이퍼박스에 대한 상대적인 연관도를 서로 다른 값으로 판별해 낼 수 있게 한다.

새로운 특징에 의해 하이퍼박스를 확장해 나가는 중

에 두개이상의 하이퍼박스가 상호 중첩될 수 있다. 이 경우 하이퍼박스의 축소(contraction) 연산을 수행함으로써 해결한다. 이러한 축소 연산에 앞서 수행되는 중첩 테스트는 다음과 같은 4가지 경우를 고려한다.

중첩경우 1 :  $u_{j\Delta} < u_{k\Delta} < v_{j\Delta} < v_{k\Delta}$

$$v_{j\Delta}^{new} = u_{k\Delta}^{new} = \frac{v_{j\Delta}^{old} + u_{k\Delta}^{old}}{2}$$

중첩경우 2 :  $u_{k\Delta} < u_{j\Delta} < v_{k\Delta} < v_{j\Delta}$

$$v_{k\Delta}^{new} = u_{j\Delta}^{new} = \frac{v_{k\Delta}^{old} + u_{j\Delta}^{old}}{2}$$

중첩경우 3 :  $u_{j\Delta} < u_{k\Delta} < v_{k\Delta} < v_{j\Delta}$

1)  $(v_{k\Delta} - u_{j\Delta}) < (v_{j\Delta} - u_{k\Delta})$

$$u_{j\Delta}^{new} = v_{k\Delta}^{old}$$

2)  $(v_{k\Delta} - u_{j\Delta}) > (v_{j\Delta} - u_{k\Delta})$

$$v_{k\Delta}^{new} = u_{j\Delta}^{old}$$

중첩경우 4 :  $u_{k\Delta} < u_{j\Delta} < v_{j\Delta} < v_{k\Delta}$

1)  $(v_{k\Delta} - u_{j\Delta}) < (v_{j\Delta} - u_{k\Delta})$

$$v_{k\Delta}^{new} = u_{j\Delta}^{old}$$

2)  $(v_{k\Delta} - u_{j\Delta}) > (v_{j\Delta} - u_{k\Delta})$

$$u_{k\Delta}^{new} = v_{j\Delta}^{old}$$

이러한 4가지 경우 각각에 대하여, 중첩된 하이퍼 박스는 중첩된 영역에 대하여 각 특징공간 별로 중첩범위를 균등하게 분리하여 영역을 축소시킨다.

#### 4. 제안된 신경망을 이용한 특징분석 기법

본 절에서는 제안된 신경망으로부터 특정 패턴 클래스에 대한 각 특징의 상호 연관성 요소를 판별하는 방법을 소개한다. 패턴 분류 응용의 측면에서 볼 때, 이는

두 가지 점에서 유용한 기법이라 할 수 있다. 우선 주어진 문제에서 어떤 특징 종류가 전체적인 분류 과정에 효과적으로 활용될 수 있는가 하는 정보를 제공한다. 두 번째로 어떤 특징값이 특정 클래스에 가장 중요한 값으로 연관되어 있는가를 판별할 수 있다는 점이다.

이를 위하여 본 연구에서는 각 특징과 패턴 클래스에 대하여 연관도(RF:Relevance Factor)를 다음과 같이 정의한다.

각 하이퍼박스 퍼지집합  $B_j = \{X, U_j, V_j, F(X, U_j, V_j)\}$ ,  $\forall X \in I^n$  이라고 정의하자. 그러면 하이퍼박스 퍼지집합의 정의를 사용하여  $k$ 번째 패턴 클래스  $C_k$ 를 다음과 같이 정의할 수 있다.

$$C_k = \bigcup_{j \in K} B_j$$

$K$ 는 클래스  $k$ 와 관련된 하이퍼박스의 색인 집합이다.  $U_j = (u_{j1}, u_{j2}, u_{j3}, \dots, u_{jm})$ 는 주어진 하이퍼박스에서 각 특징에 대한 최소점의 집합이고  $V_j = (v_{j1}, v_{j2}, v_{j3}, \dots, v_{jm})$ 는 최대점의 집합이다. 또한 식에서  $F$ 는 주어진 특징 벡터에 대한 하이퍼박스 멤버쉽함수이다.

학습된 신경망으로부터 주어진 클래스의 연관성 특징 목록을 아래와 같이 산출한다. 즉 클래스  $k$ 와 관련된 특징  $f_i$ 의 연관도(RF)는 식 (8)과 같이 정의된다.

$$RF(f_i, k) = \left( \frac{1}{N_k} \sum_{B_j \in C_k} S(f_i, (u_{jn}, v_{jn})) \cdot w_{jn} - \frac{1}{(N_B - N_k)} \sum_{B_j \in C_k} S(f_i, (u_{jn}, v_{jn})) \cdot w_{jn} \right) / \sum_{B_j \in C_k} w_{jn} \quad (8)$$

식에서 상수  $N_B$ 는 총 하이퍼박스의 개수이며  $N_k$ 는 클래스  $k$ 에 속한 하이퍼박스의 개수이다. 식 (8)에서 특징  $f_i$ 는 최종적으로 생성된 하이퍼박스에 나타나는 특징범위이다. 그러므로 이는 최소값 및 최대값 쌍으로 이루어진 퍼지 구간의 형식으로 나타난다. 임의의 특징  $f_i$ 의 최대값을  $f_i^U$ , 최소값을  $f_i^L$ 로 정의하고, 두 퍼지 구간의 유사도(similarity)  $S$ 를 식 (9)와 같이 정의한다.

$$S(f_i, (u_i, v_i)) = S((f_i^L, f_i^U), (u_i, v_i)) \quad (9)$$

$$= \frac{\text{overlap}((f_i^L, f_i^U), (u_i, v_i))}{\text{Max}(f_i^U - f_i^L, v_i - u_i)}$$

식 (9)에서 두 퍼지구간이 모두 한 점으로 이루어진 데이터라면 분모  $\text{Max}(f_i^U - f_i^L, v_i - u_i)$ 는 0이 되는데 이 경우는 별도로 고려하여 식 (10)을 적용한다. 즉 두 구간이 동일한 한 점일 경우 유사도 값은 1.0이며, 서로 다른 두 점일 경우 0이 된다.

$$S((f_i^L, f_i^U), (u_i, v_i)) = \begin{cases} 1 & \text{if } (f_i^L = f_i^U = u_i = v_i) \\ 0 & \text{Otherwise} \end{cases} \quad (10)$$

만약  $\text{Max}(f_i^U - f_i^L, v_i - u_i)$ 가 0보다 크다면 값은 식

(11)를 적용하여 산출된다.

$$\text{Overlap}((f_i^L, f_i^U), (u_i, v_i)) = \begin{cases} f_i^U - u_i & \text{if } (f_i^L \leq u_i \leq f_i^U \leq v_i) \\ v_i - u_i & \text{if } (f_i^L \leq u_i \leq v_i \leq f_i^U) \\ f_i^U - f_i^L & \text{if } (u_i \leq f_i^L \leq f_i^U \leq v_i) \\ v_i - f_i^L & \text{if } (u_i \leq f_i^L \leq v_i \leq f_i^U) \\ 0 & \text{Otherwise} \end{cases} \quad (11)$$

만약  $RF(f_i, k)$ 가 양의 값을 가지면 이것은 클래스  $k$ 와 특징  $f_i$ 사이의 자극성(excitatory) 연관성이 있음을 의미한다. 그러나  $RF(f_i, k)$ 가 음의 값을 가지면 그들 사이는 억제성(inhibitory) 연관성을 가진다는 것을 의미한다. 또한 주어진 클래스에 대하여 중요한 역할을 하는 흥미있는 특징들의 목록을, 계산된 연관성요소 값을 사용하여 선정할 수 있다. 이는 각 특징의 종류에 대하여 모든 값과 모든 하이퍼박스에서 연관성요소의 값을 구분하여 합산하고, 이를 정렬하여 출력함으로써 이루어질 수 있다.

## 5. 실험 결과 및 고찰

실험은 패턴 분류 문제에서 잘 알려진 아이리스 데이터[2]와, 클리브랜드 대학의 의료진단 데이터[13]를 사용하여 기존의 FMM 신경망을 사용한 경우와 제안된 모델을 사용한 분류결과를 상호 비교하였다. 아이리스 데이터는 총 3개의 클래스에 각각 50개의 패턴으로 구성되며 각 패턴은 4종류의 특징으로 이루어지며, 의료진단 데이터는 총 5개 클래스에 297개 패턴을 대상으로 하였으며, 각 패턴별로 13종류의 특징을 고려하였다.

### 실험 1 : 패턴 분류 성능 평가

본 연구에서는 제안된 모델과 기존의 FMM 모델의 분류성능을 비교하기 위하여 식 (12)와 같은 에러 척도를 사용하였다.

$$E = \frac{1}{pm} \sum_{i=1}^m \sum_{k=1}^m |c_{ik} - d_{ik}| \quad (12)$$

식에서  $p$ 는 테스트 패턴의 개수이며,  $m$ 은 총 클래스 수이다.  $c_{ik}$ 와  $d_{ik}$ 는 각각  $i$ 번째 입력에 대한 클래스  $k$ 의 실제 출력값(actual output)과 기대 출력값(desired output)을 의미한다. 다시 말해서 모든 테스트 패턴에 대하여 기대출력값과 실제 출력값과의 차이에 대한 평균치를 에러척도로서 평가하였다. 그림 3은 이러한 에러 척도를 적용하여 두 모델에 대한 패턴 분류 실험 결과의 비교를  $\theta$  값의 변화에 따라 보인 것이다. 그림에서 보인 바와 같이  $\theta$  값의 전반적인 변화에 관계없이 제안된 방법의 기존의 FMM 모델에 비해 분류율을 개선할 수 있음을 알 수 있다.

표 1은 학습패턴의 수에 따른 분류율의 변화를 비교한 내용이다. 여러번의 튜닝을 통하여 제안된 모델의 파

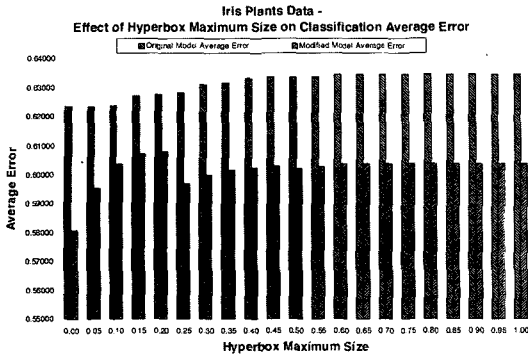


그림 3 아이리스 패턴에 대한 분류 실험결과 비교

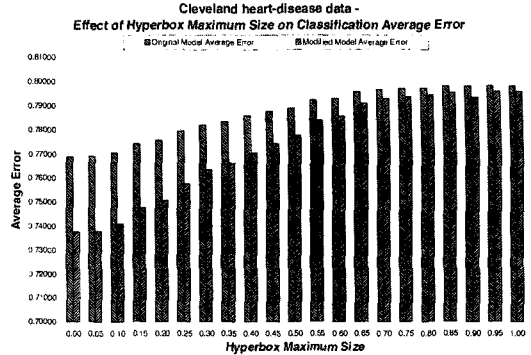


그림 4 클리블랜드 의료 데이터에 대한 분류 실험 결과

라미터 값을 최적화 한 결과  $\theta=0.3$ ,  $\gamma=0.5$ ,  $T=1.2$ ,  $d=0.1$ ,  $s=0.05$ ,  $l=0.1$ 로 설정하였다. 실험 결과는 표에 보인 바와 같이 학습패턴의 수가 증가함에 따라 오분류 패턴 수 및 평균 에러율은 감소하였으며, 제안된 모델의 경우 오분류 패턴 수의 측면에서는 FMM모델과 비교하여 동등한 수준을 보였고, 에러율, 즉 식 (12)에 정의한 척도를 적용한 경우 평균적으로 3~5 %정도 개선된 결과를 보였다.

이어서 동일 실험을 클리블랜드 의료 진단 데이터에 대하여 적용하였다. 그림 4는 두 모델에 대한 패턴 분류 성능을  $\theta$  값의 변화에 따라 비교한 것이다. 그림에 보인 바와 같이  $\theta$  값이 0.3 이내에서 두 모델 모두 성능이 개선됨을 보였으며, 제안된 모델의 성능 개선 효과가 비교적 현저하게 나타났다. 그러나  $\theta$  값을 지나치게 낮추는 경우 하이퍼박스가 지나치게 많이 만들어 지는 현상, 이른바 오버피팅(overfitting)현상이 고려되어야 한다. 따라서 본 실험에서는  $\theta$  값을 0.2로 설정하여 다음 실험에 적용하였다.

이어서 앞서의 실험에서와 같이 학습 패턴수의 변화에 따른 분류 성능의 비교를 표 2에 보였다. 이때 사용된 파라미터 값은  $\theta=0.2$ ,  $\gamma=0.5$ ,  $T=1.2$ ,  $d=30.0$ ,  $s=0.05$ ,  $l=0.1$ ,  $T=1.01$ 로 설정하였다. 실험 결과는 표에 보인 바와 같이 학습패턴의 수가 증가함에 따라 오분류 패턴 수 및 평균 에러율은 감소하였으며, 제안된 모델의 경우 오분류 패턴 수의 측면 및 에러율의 측면 모두에서 기존의 FMM모델과 비교하여 개선된 결과를 보였다.

**실험 2 : 특징 분석 기법 실험**

앞절에서 기술한 특징 분석 기법에 대하여 두 종류의 문제에 적용하였다. 그림 5는 아이리스 데이터에 대한 특징 분석 시뮬레이션 결과 화면의 예이다.

두 개의 패턴 집합에 대하여 특징 분석 실험 결과를 표 3과 표 4에 정리하였다. 표에서 보인 바와 같이 학습된 신경망으로부터 특정 패턴 클래스에 대한 연관 특징(relevant feature)의 목록을 연관도 요소(RF) 및 가중치 요소와 함께 보여준다.

표 1 학습 패턴 수 변화에 따른 분류 성능 변화 비교 (아이리스 데이터)

# of training. patterns	Original FMM		Proposed Model	
	# of Error patterns	error rate	# of error patterns	error rate
40	4	0.635	4	0.604
60	3	0.630	3	0.595
80	1	0.629	1	0.592
100	0	0.624	0	0.592
120	0	0.621	0	0.581

표 2 학습 패턴 수 변화에 따른 분류 성능 변화 비교 (클리블랜드 의료 데이터)

# of training patterns	Original FMM		Proposed Model	
	# of Error patterns	error rate	# of error patterns	error rate
10	35	0.785	32	0.684
20	29	0.761	27	0.675
30	21	0.723	19	0.642
40	14	0.654	12	0.580
50	7	0.621	5	0.533

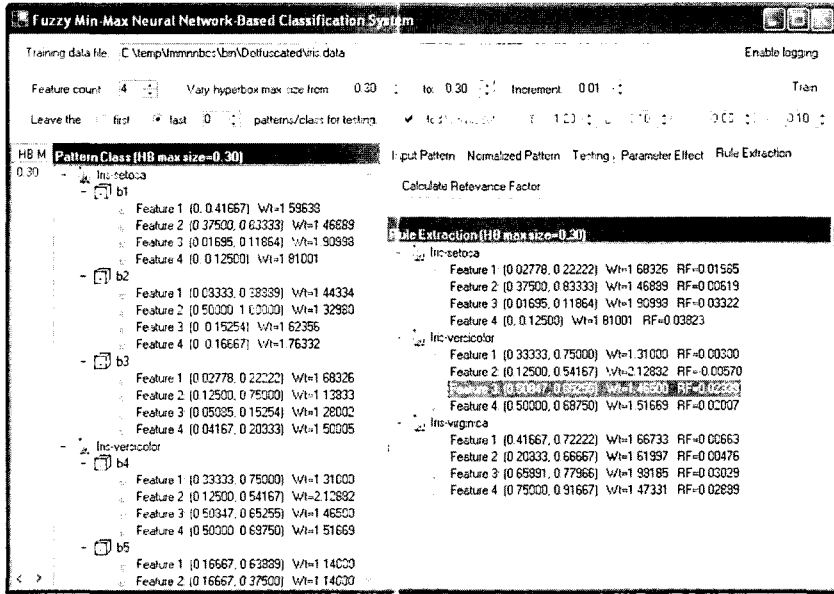


그림 5 특징 분석 시뮬레이션 시스템 출력의 예

표 3 아이리스 데이터 분류문제에 대한 특징 분석 실험 결과

패턴클래스	RF와 가중치를 가진 연관성특징 목록
Setosa	F4:(0, 0.125), RF=0.038, W=1.810, F3:(0.016, 0.118), RF=0.033, W=1.909, F1:(0.027, 0.222), RF=0.019, W=1.683, F2:(0.375, 0.833), RF=0.006, W=1.468
Versicolor	F3:(0.508, 0.652), RF=0.023, W=1.465, F4:(0.5, 0.687), RF=0.020, W=1.516, F2:(0.125, 0.541), RF=0.005, W=2.128, F1:(0.333, 0.75), RF=0.003, W=1.31
Virginica	F3:(0.658, 0.779), RF=0.030, W=1.981, F4:(0.75, 0.916), RF=0.028, W=1.473, F2:(0.208, 0.666), RF=0.004, W=1.619, F1:(0.416, 0.722), RF=0.006, W=1.667

표 3과 4에서 보인 바와 같이 제안된 모델은 패턴 분류 문제에서 특징 클래스와 특징간의 상호 연관도를 분석해 낼 수 있게 한다. 이러한 결과는 실제 응용에서 지식생성 기법이나, 신경망 분류기의 학습효과 분석 등에 사용될 수 있다.

실험결과에 대한 타당성 평가를 위하여, 학습에 사용된 데이터에 대해 특징의 상대적 중요도를 단순화된 방법을 적용한 결과와 비교하였다. 즉 정규화된 특징 공간 상에서 각 클러스터의 중심점으로부터 타 클러스터의 중심점과 상대적 거리의 합을 산출하여 이를 이 특징의 연관도 요소로 고려하였다. 표 5는 이 결과와 본 연구에서 산출한 연관도 값을 비교한 결과의 예이다.

표 4 클리블랜드 의료 데이터 분류문제에 대한 특징 분석 실험 결과

패턴클래스	RF와 가중치를 가진 연관성특징 목록
0	F12:(0, 0), RF=0.004, W=1.8, F9:(0, 0), RF=0.003, W=1.8, F3:(0.666, 0.666), RF=0.003, W=1.5, F2:(1.0, 1.0), RF=0.003, W=1.8
1	F2:(1.0,1.0), RF=0.001, W=1.6, F1:(0.229, 0.604), RF=0.001, W=2.819, F12:(0, 0), RF=0.001, W=1.3, F6:(0, 0), RF=0.001, W=1.6
2	F3:(1.0, 1.0), RF=0.003, W=1.2, F6:(0, 0), RF=0.002, W=1.2, F8:(0.244, 0.541), RF=0.002, W=2.950, F10:(0.225, 0.451), RF=0.001, W=2.399
3	F9:(1.0, 1.0), RF=0.003, W=1.3, F6:(0, 0), RF=0.002, W=1.3, F2:(1.0, 1.0), RF=0.002, W=1.3, F12:(0, 1.0), RF=0.001, W=1.22
4	F2:(1.0, 1.0), RF=0.002, W=1.2, F3:(1.0, 1.0), RF=0.002, W=1.2, F5:(0.182, 0.641), RF=0.001, W=2.057, F13:(1.0, 1.0), RF=0.001, W=1.2

표 5 아이리스 데이터 Setosa 클래스에 대한 특징의 연관도 평가

특징 종류	클래스 중심간 거리합	제안된 모델의 RF값
F1	0.698	0.019
F2	0.455	0.006
F3	1.167	0.033
F4	1.193	0.038

물론 이때 이러한 방법이 주어진 문제에서 궁극적인 특징의 중요도를 정확하게 판정할 수는 없지만, 특징공간상에서 타 클래스와의 거리의 합이 클수록 중요도가 높을 것이라는 가정은 직관적으로 타당하다. 실험결과 표에서 보인 바와 같이 제안된 방법에 의해 산출된 결과가 특징별로 클러스터간 거리로 산출한 단순화된 직관적 방법과 그 상대적인 값의 크기 순서가 유사하게 나타남을 확인함으로써 제안된 이론의 타당성을 뒷받침할 수 있다.

또한 위 두 실험에서 연관도 요소 값이 상대적으로 작은 수치로 산출됨은 가중치 값의 합으로 결정되는 정규화 요소(normalizing factor)가 특징수의 증가와 하이퍼박스 크기증가 및 패턴수의 증가에 따라 커진다는 사실에 기인한다. 제안된 이론에서는 이를 상대적인 값으로만 고려하므로 연관도의 순서판단에는 영향을 주지 않는다. 그러나 향후 이러한 요소가 수치의 크기로서 부여되는 정보가 되기 위해서는 정규화 요소를 재조정할 필요가 있다.

## 5. 결론

기존의 FMM 신경망은 신경망의 학습능력과 퍼지집합이론의 장점을 조화시킨 일종의 퍼지신경망 모델이다. 제안된 모델은 이러한 퍼지 신경망모델의 장점을 활용하면서 보다 확장된 기능을 지원하는 패턴 분류 기법이다. 그러므로 학습을 통한 지식 획득이 가능할 뿐만 아니라, 기존의 순수 신경망에 비해 인간과 같은 논리적인 능력과 전문지식을 결합할 수 있는 가능성을 지원한다. 제안된 신경망의 가중치 요소는 학습패턴에서 특징값의 빈도의 차이에서 오는 영향을 반영한다. 즉 잘 구성된 학습데이터는 각 패턴 클래스에 대하여 균등한 빈도로 나타남이 적당함을 전제로 한다면, 상대적으로 빈도가 적은 비정상 패턴이 분류 시스템의 성능에 영향을 주는 것은 바람직하지 못하다. 이에 본 연구에서 제안하는 모델은 가중치변화 과정에서 이러한 빈도 요소를 반영하게 함으로써 이러한 문제점을 개선하였다. 본 연구에서 제안된 가중치 요소와, 특징과 클래스의 연관도 정의 등은 다른 패턴 분류 방법론과 결합하여 지식분류 및 분석기능과, 분류과정에 대한 시스템 설명기능 등을 구현하는데 활용될 수 있다.

## 참고 문헌

- [1] S. Mitra, and Y. Hayashi, "Neuro-Fuzzy Rule Generation: Survey in Soft Computing Framework," IEEE Transactions on Neural Networks, Vol.11, No.3, pp.748-768, 2000.
- [2] P. K. Simpson, "Fuzzy Min-Max Neural Networks

Part 1:Classification," IEEE Transaction on Neural Networks, Vol.3, No.5, pp.776-786, 1992.

- [3] S. Mitra, R. K. De, and S. K. Pal, "Knowledge-Based Fuzzy MLP for Classification and Rule Generation," IEEE Transactions on Neural Networks, Vol.8, No.6, pp.1338-1350, 1997.
- [4] B. Gabrys, A. Bargiela, "General Fuzzy Min-Max Neural Network for Clustering and Classification," IEEE Transaction on Neural Networks, Vol.11, No.3, 2000.
- [5] P. K. Simpson, "Fuzzy Min-Max Neural Networks Part 2:Clustering," IEEE Transaction on Fuzzy Systems, Vol.1, No.1, pp.32-45, 1993.
- [6] C. Z. Ye, J. Yang, D. Geng, Y. Zhou, N. Y. Chen, "Fuzzy Rules to Predict Degree of Malignancy in Brain Glioma," Medical and Biological Engineering and Computing, Vol.40, pp.145-152, 2002.
- [7] Raphael Feraud, Fabrice Clerot, "A Methodology to Explain Neural Network Classification," Neural Networks, Vol.15, No.1, pp.237-246, 2002.
- [8] Kumar S. Ray and Jayati Ghoshal, "Neuro Fuzzy Approach to Pattern Recognition," Neural Networks, Vol.10, No.1, pp.161-182, 1997.
- [9] Mahn M. Lee, Kuo H. Chen, I. F. Jiang, "A Neural network Classifier with Disjunctive Fuzzy Information," Neural Networks, Vol.11, No.1, pp. 113-1125, 1998.
- [10] Hung P. Chiu, Din C. Tseng, "Invariant Hand-written Chines Character Recognition using Fuzzy Min-Max Neural Networks," Pattern Recognition Letters, Vol.18, No.1, pp.481-491, 1997.
- [11] Jayanta Basak, Rajat K. De, Sankar K. Pal, "Unsupervised Feature Selection using a Neuro-Fuzzy Approach," Pattern Recognition Letters, Vol.19, No.1, pp.997-1006, 1998.
- [12] F. Behloul, B.P.E. Lelieveldt, A. Boudraa, J.H.C. Reiber, "Optimal Design of Radial Basis Function Neural Networks for Fuzzy-Rule Extraction in High Dimensional Data," Pattern Recognition, Vo.35, No.1, pp.659-675, 2002.
- [13] C. L. Blake, and C. J. Merz, "UCI Repository of machine learning databases [http://www.ics.uci.edu/~mlern/MLRepository.html]," Irvine, CA: University of California, Department of Information and Computer Science, 1998.



김 호 준

1987년 경북대학교 전자공학과 졸업. 1987년~1991년 한국원자력연구소 연구원. 1995년 한국과학기술원 전산학과 박사. 1995년 한국과학기술원 인공지능연구센터 연구원. 1996년~현재 한동대학교 전산전자공학부 부교수. 관심분야는 패턴인식, 컴퓨터시각, 신경망, 학습이론 등임





양 현 승

서울대 학사(1976). Purdue University 전자과 석사(1983). Purdue University 전자과 박사 학위 취득(1986). University of IOWA 전자전산과 조교수(1986년~1988년). 인공지능 연구센터 시각 연구실장(1988년~1999년). 한국과학기술원 전산학과 정교수(1988년~현재). 관심분야는 컴퓨터 시각, 로보틱스, 인공지능, 멀티미디어