

# FMM 신경망에서 하이퍼박스 축소의 부작용을 개선한 학습기법 연구

박현정, 김호준  
한동대학교 전산전자공학부  
e-mail : [hjpark79@empal.com](mailto:hjpark79@empal.com)

## A Learning Method of FMM Neural Network to Reduce the Side Effect of Hyperbox Contraction

Hyun Jung Park, Ho Joon Kim  
School of Computer Science and Electronic Engineering, Handong Global University

### 요 약

본 논문에서는 기존의 FMM 신경망의 학습 단계 중 하이퍼박스 축소 단계에서 발생하는 정보의 왜곡 현상을 보완하여, 패턴분류 성능을 향상시키는 방법론을 제안하고 그 유용성을 고찰한다. 제안된 학습방법에서는 기존 FMM 신경망의 하이퍼박스 생성, 확장, 축소의 3 가지 학습 과정 중 축소과정을 생략하고, 각 특징 범위와 빈도를 고려한 학습으로 대체함으로써 축소과정에서 나타나는 부작용을 보완할 수 있도록 하였다. 이는 가중치 개념을 통해 각 특징값의 상대적 중요도를 서로 다른 값으로 반영할 수 있도록 하며 하이퍼박스 중첩영역에서 발생하는 분류 오류를 최소화한다. 본 연구에서는 개선된 학습기법의 이론을 제시하고, 이를 아이리스 데이터 분류 실험에 적용하여 그 유용성을 고찰하였다.

### 1. 서 론

퍼지 최대최소 신경망(Fuzzy Min-Max(FMM) Neural Network)은 매우 간결하면서도 강력한 학습방법을 갖는 패턴 분류 모델이다[1]. 이는 온라인 단일패스의 학습 알고리즘을 지원하며 새롭게 주어지는 학습패턴에 대하여 점진적인 학습(incremental learning)이 가능하다. 그러나 학습 및 분류과정에서 특징값의 분포 또는 빈도를 고려하지 않음으로써 학습패턴에 포함된 노이즈 패턴이나 비정상적 패턴에 의해 성능저하가 발생할 수 있다[3]. 이에 본 논문에서는 가중치 개념과 특징의 빈도정보를 활용하는 분류 모형과 학습기법을 제안한다. 가중치 값은 각 하이퍼박스의 특징 차원에서 특징범위와 그에 속하는 패턴의 발생빈도에 따라 결정된다. 제안된 모델은 기존의 FMM 신경망에 비하여 다음과 같이 몇 가지 장점을 제공한다. 우선 주어진 학습패턴 집합에서 특정클래스에 속하는 패턴의 수와 특징값의 분포가 불균일한 경우 분류성능의 개선을 기대할 수 있게 한다. 두 번째로 각 특징에 대한 가중치 정보는 주어진 클래스에 대한 해당특정의

상대적 중요도를 반영함으로써 특징의 분석기법에 활용될 수 있다. 또한 특징의 빈도와 가중치 정보는 기존의 FMM 신경망의 학습과정에서 하이퍼박스 축소시 발생할 수 있었던 중첩영역에서의 분류 오류를 개선할 수 있게 한다. 본 논문에서는 제안된 모델의 구조와 동작특성을 소개하고, 새롭게 정의된 학습기법에 관해 기술한다. 기존의 모델에서 하이퍼박스 축소과정에서 발생하는 문제점에 관하여 이론적으로 고찰하고, 이를 개선하기 위한 방법론에 관해 기술한다. 세부적으로 하이퍼박스 생성 및 확장 과정과 이에 수반되는 빈도 및 가중치 조정 방법을 제안한다. 이어서 실제 분류문제를 대상으로 하여 정상패턴과 노이즈 패턴에 대하여 기존의 모델과 분류성능을 비교 하였다. 본 논문의 2 장에서는 기존의 FMM 신경망의 특성에 관해 기술하며, 3 장에서는 기존 모델이 갖는 하이퍼박스 축소과정의 문제점을 분석한다. 이어서 4 장에서는 본 연구에서 제안하는 개선된 FMM 신경망 모델의 구조와 학습방법을 소개하며, 5 장에서 아이리스 데이터 분류 문제에 제안된 이론을 적용하고 제안된 이론의 타당성을 고찰한다.

2. 퍼지 최대 최소 신경망(FMM Neural Network)

FMM 신경망은 하이퍼박스 기반의 패턴 분류 모형으로, 하이퍼박스는 다차원 특징공간에서 개별 특징에 대한 최대값 및 최소값의 쌍으로 표현된다. 주어진 입력패턴에 대하여 하이퍼박스 소속함수 값이 산출되고 이로부터 해당 패턴 클래스가 결정된다. 임의의 패턴  $A_h$  의  $j$  번째 하이퍼박스 퍼지집합에 대한 소속함수  $b_j(A_h)$ 는 0에서 1 사이의 적합도로 표현되며, 이는 식 (1)과 같다.

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

여기서  $A_h = (a_{h1}, a_{h2}, \dots, a_{hn})$ 는  $h$  번째 입력 패턴으로  $n$  개의 특징으로 이루어지며,  $U_j = (u_{j1}, u_{j2}, \dots, u_{jn})$ 는  $b_j$ 의 최소점을,  $V_j = (v_{j1}, v_{j2}, \dots, v_{jn})$ 는 최대점을 각각 의미한다.  $\gamma$ 는 특징범위 가장자리에서 퍼지소속함수의 기울기를 결정하는 매개변수이다.

FMM 신경망은 이러한 하이퍼박스의 최소점과 최대점을 결정하기 위해 3단계의 학습 과정을 거친다. 1 단계는 하이퍼박스 생성 및 확장의 단계로, 학습 패턴이 동일 클래스 내의 기준에 존재하는 하이퍼박스와 평균적으로 일정 거리( $\theta$ )이내에 있는 경우 하이퍼박스를 확장하고 그렇지 않은 경우 새로운 하이퍼박스를 생성한다. 이어서 중첩 테스트의 단계로, 확장된 하이퍼박스가 다른 클래스에 속한 하이퍼박스와 부분중첩 및 완전중첩 여부를 확인하며, 중첩된 경우 마지막 단계인 축소과정을 통해 중첩 부분을 제거한다.

학습과정에 의해 각 하이퍼박스의 특징범위가 결정되면, 식 (1)의 적합도 계산을 통해 임의의 입력 패턴에 대한 분류작업을 수행하게 된다. 이러한 FMM 신경망은 식 (1)에 보인 바와 같이 하이퍼박스 소속함수 값 산출시 특징의 빈도요소를 고려하지 않는다. 또한 학습시 중첩영역을 축소하는 과정에서 기 학습된 분류기능을 왜곡할 가능성을 내재한다.

3. 축소기법의 문제점

FMM 신경망의 학습 과정에서 하이퍼박스 확장이 이루어지면 타 클래스에 속하는 하이퍼박스와 중첩여부를 점검하게 된다. 그 결과 중첩부분이 발생할 경우 축소과정이 이루어진다. 축소과정은 중첩의 형태에 따라 두 가지 유형으로 구분되는데, 부분중첩과 완전중첩이다. 이들 각각 유형에 대하여 별도의 축소과정이 적용된다. 부분중첩의 경우 중첩된 두 하이퍼박스가 중첩 부분에 대하여 동일크기로 축소되며, 완전중첩의 경우는 범위가 작은 하이퍼박스를 그대로 유지하도록 범위가 큰 하이퍼박스를 축소한다. 이러한 하이퍼박스 축소과정을 그림 1에 나타내었다. 그림에서 이중실선 부분이 중첩 영역이며 축소될 위치를 점선으로 표시하였다. 그림에서 보인 바와 같이 부분중첩의 경우 두 하이퍼박스  $i$ 와  $j$ 에 나타난 중첩부분을 제거하게 되는데, 이 때 축소되는 크기는 각 하이퍼박

스에 나타난 특징의 빈도요소와 무관하게 동일크기로 적용된다. 그런데, 만일 이 경우 한 개의 하이퍼박스가 노이즈 패턴에 의해 확장된 경우라면, 기 학습된 하이퍼박스의 특징범위 정보를 왜곡하게 된다.

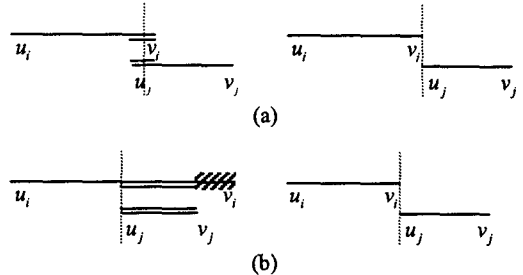


그림 1. FMM 신경망의 축소개념 (a) 부분중첩 (b) 완전중첩

또한 완전 중첩의 경우 위의 문제뿐만 아니라 중첩 부분이 아닌 범위(그림 1(b)에서 빗금친 영역)를 포함하여 축소함으로써 왜곡의 정도가 더욱 심해진다. 이는 작은 하이퍼박스 영역  $j$ 가 큰 하이퍼박스  $i$  영역의 중심점 가까이에 위치하는 경우, 하이퍼박스  $i$ 의 정보가 매우 심하게 왜곡되며, 이는 궁극적으로 패턴 분류 성능을 저하시키는 직접적인 요인이 될 수 있다.

4. 수정된 모델 제안

본 연구에서는 위에서 언급한 기존의 FMM 신경망 모델의 단점을 개선하는 방법론을 고찰한다. 즉 노이즈 패턴에 의해 확장된 하이퍼박스 축소과정에서, 특징 범위 정보를 왜곡하는 부작용을 개선한 새로운 FMM 신경망 모델을 제안한다. 제안된 모델은 기존 학습 과정의 중첩 테스트와 축소 과정을 생략할 수 있게 하며, 중첩문제의 해결은 하이퍼박스 특징 영역에 나타나는 빈도수를 고려하는 학습과정으로 대처한다. 이를 위하여 하이퍼박스  $B_i$ 의 특성을 아래와 같이 새롭게 정의한다.

$$B_j = \{X, U_j, V_j, F_j, f(X, U_j, V_j, F_j)\} \quad \forall X \in I^n$$

여기서  $X$ 는  $\{X_h, d_h\}$ 로  $n$  개의 특징으로 이루어진 패턴  $X_h = \{x_{h1}, x_{h2}, \dots, x_{hn}\}$ 와  $m$  개의 클래스  $d_h = \{1, 2, \dots, m\}$ 로 정의되는 학습 패턴 집합이다.  $U_j$ 와  $V_j$ 는 각각 하이퍼박스의 최소점 및 최대점을 의미한다.  $F_j$ 는 하이퍼박스 특징 범위내에 나타나는 패턴의 빈도수를 의미한다.

제안된 모델의 학습 알고리즘은 다음과 같다. 기존에 생성된 하이퍼박스 중에서, 주어진 패턴에 가장 크게 반응하는 하이퍼박스를 선택하여 확장 여부를 결정한다. 확장여부는 식 (2)의 조건에 따라 결정되는데 확장될 하이퍼박스와 의 최소 평균 거리를 파라미터  $\theta$ 로 제어한다. 식 (2)에서  $n$ 은 특징의 수이고 파라미터  $\theta$ 는  $[0, 1]$  사이의 값을 가진다.

확장과정은 식 (3)과 같이 하이퍼박스의 최대점, 최소점을 재조정하고 특징의 빈도값을 수정한다.

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

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

다시 말해서 식 (2)의 조건을 만족하면 식 (3)을 통해 하이퍼박스의 최대점과 최소점을 재조정하는 과정으로 확장이 이루어지고 특징의 빈도변수  $f_{ji}$ 의 값을 1씩 증가시키게 된다.

이러한 빈도값로부터 특징의 가중치가 산출된다. 본 연구에서는 가중치를 적용하는 형태의 새로운 하이퍼박스 소속함수를 식 (4)와 같이 정의한다.

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

식 (4)에서 나타는  $w_{ji}$ 는  $j$  번째 하이퍼박스와  $i$  번째 특징 간의 연결 가중치를 의미하며, 나머지 표기는 식 (1)과 동일하다. 식 (4)에서 새롭게 정의한 소속함수는 기존 FMM 신경망에서 0.5에서 1 사이의 값으로 나타났던 형태를 0에서 1 사이의 값을 갖도록 조정하였으며, 하이퍼박스와 특징과의 관계  $w_{ji}$ 가 각 특징에 가중치로 적용되도록 하였다.

학습과정에서 빈도변수값이 최종적으로 유지되면 이로부터  $w_{ji}$ 값이 식 (5)에 의해 산출된다.

$$w_{ji} = \frac{\alpha f_{ji}}{R} \quad (5)$$

$$R = \max(s, v_{ji} - u_{ji}) \quad (\text{단, } s > 0)$$

식에서 보인 바와 같이  $w_{ji}$ 는  $j$  번째 하이퍼박스에서  $i$  번째 특징의 발생 빈도에 비례한 값으로 결정되며, 이는 특징의 구간크기에 반비례하여 재조정된다. 단, 특징의 범위가 단일 포인트로 나타나는 경우에 가중치가 지나치게 큰 값이 되는 것을 방지하기 위해 지정된 값  $s$ 의 범위로 대체하여 계산하여준다.  $s$ 는 문제에서 주어지는 패턴 데이터의 형태에 따라 결정되는 파라미터로 0보다 큰, 0 근처의 값으로 정해준다. 이와 같이 계산되는 가중치 개념은 좁은 특징 범위에 많은 빈도의 학습패턴이 주어질수록 그 특징은 해당 하이퍼박스에 상대적으로 높은 중요도를 가지는 특징이 된다는 점을 반영한다.

그림 2는 FMM 신경망과 제안된 모델의 학습과정에 대한 비교이다. 그림에서 각 사각형 영역은 하이퍼박스를 의미하며 현재 최근 한 개의 학습패턴에 의하여 그림 2의 (a)의 형태가 형성되었다고 가정한다. 그림에서 편이상 2차원 공간상에 면적이 큰 하이퍼박스를  $j$ , 작은 하이퍼박스를  $i$ 라고 지정하였다. 그리고 주어진 두 개의 패턴 P1과 P2는 모두 하이퍼박스  $j$ 에 속한 패턴이다.

이 경우 중첩이 발생하였으므로 FMM 신경망은 하이퍼박스 축소과정을 수행하게 되는데, 그 결과 (b)와 같은 형태로 중첩영역이 제거된다. 이 결과로부터 분류실험을 수행하면, 중첩되었던 영역에 정보의 왜곡

이 발생하여 패턴 P2의 경우 하이퍼박스  $i$ 에 대하여 반응함으로써 분류오류가 발생하게 된다. 뿐만 아니라 패턴 P1의 경우에도 하이퍼박스와와의 거리에 따라 소속함수값이 결정되므로 오히려  $i$ 에 대하여 반응하게 된다.

반면 제안된 모델에서는 그림 2의 (b)와 같은 축소가 발생하지 않고, (a)의 원 형태인 상태에서, 각 하이퍼박스의 빈도 변수만을 서로 다른 값으로 유지한다. 그러므로 하이퍼박스 축소에 의한 정보 소실이 없기 때문에, 축소과정의 부작용으로 생기는 위의 문제가 발생하지 않게 된다. 또한 P2와 같이 두 개 이상의 하이퍼박스에 동시에 속해있는 패턴의 경우에도 각 하이퍼박스에 대한 특징별 빈도수의 작용, 즉 가중치에 의하여 서로 다른 반응 정도를 산출하므로, 학습과정에서 상대적으로 패턴의 발생빈도가 높은 하이퍼박스에 속하는 것으로 판정하게 된다. 즉, 학습된 패턴의 빈도를 유지하여 통계적으로 보다 빈번하게 관찰된 하이퍼박스에 속하는 것으로 분류하게 된다. 이는 학습패턴에서 각 클래스에 속하는 표준 패턴의 빈도가 균일하게 주어진다고 가정했을 때 노이즈나 비정상적인 패턴으로부터의 영향을 개선할 수 있는 바람직한 특성이 된다.

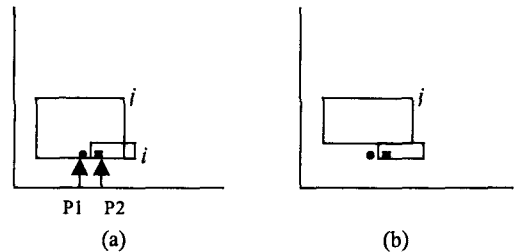


그림 2. FMM 신경망에서 하이퍼박스 축소과정의 문제점, (a) 중첩의 형태, (b) 하이퍼박스 축소 결과

### 5. 실험 결과 및 고찰

FMM 신경망과 제안된 모델의 성능 비교를 위하여 학습된 모델에 동일 패턴집합을 입력하여 분류과정을 수행하였다. 실험결과와 분석은 오류 패턴의 개수에 따라 식 (6)으로 계산된 오류율을 비교, 분석하였다.

$$E = \frac{1}{pm} \sum_{i=1}^p \sum_{k=1}^m |c_{ik} - d_{ik}| \quad (6)$$

여기서  $p$ 는 테스트 패턴의 수이며,  $m$ 은 클래스의 수,  $c_{ik}$ 와  $d_{ik}$ 는  $i$ 번째 입력패턴에 대한  $k$ 번째 클래스의 실제 출력값과 기대 출력값을 각각 의미한다. 다시 말해서 정확한 분류여부뿐만 아니라 오분류에 대해서도 얼마나 올바른 값의 변화를 보였는가를 평가하고자 함이다. 실험은 다음과 같이 세가지 경우에 대하여 기존의 FMM 신경망을 적용한 경우와 상호 비교, 평가하였다.

첫 번째로 패턴분류 문제로 잘 알려진 Iris 데이터를 적용, 분류실험을 수행하였다. 아이리스 데이터는 총 3개의 클래스에 각 50개씩 총 150개의 데이터

패턴으로 구성되어 있다. 실험에 사용된 파라미터 값들은 여러 회의 튜닝을 통해,  $\theta = 0.255$ ,  $\gamma = 0.2$ ,  $s = 0.001$ ,  $\alpha = 1.0$ 으로 설정하였다.

FMM 신경망과 제안된 신경망 모델에 대하여 학습데이터의 수를 증가시키며 분류성능 및 에러율을 비교하였다. 표 1은 학습패턴수의 증가에 따른 오분류 패턴수 및 식 (6)에 정의한 에러율의 비교이다. 표 1에 나타난 바와 같이 제안된 모델이 기존의 FMM 신경망에 비해 10%에서 18%가량 개선된 성능을 보였으며, 학습패턴수의 증가에 따라 지속적으로 개선되었다.

표 1. Iris 데이터에 대한 실험결과

학습패턴수	FMM 신경망		제안된 모델	
	에러패턴	에러율	에러패턴	에러율
30	7	0.01190	5	0.01084
60	4	0.00386	4	0.00386
90	5	0.00329	3	0.00284
120	3	0.00215	1	0.00182
150	3	0.00165	1	0.00140

이어서 제안된 모델이 노이즈나 왜곡된 패턴의 영향개선 효과에 대한 실험을 수행하였다. 이를 위해, 학습 패턴에서 임의로 1개의 왜곡된 패턴을 의도적으로 삽입, 학습한 결과를 비교하였다. 그 결과를 표 2에 나타내었다. 표에서 보인 바와 같이 주어진 학습 패턴중에 1개의 노이즈 패턴을 포함할 때, 성능저하가 FMM 신경망에 비하여 개선됨을 알 수 있다. 이는 제안된 모델이 노이즈나 비정상 패턴에 대해 보다 강한 분류성능을 가질 수 있음을 시사한다.

표 2. 노이즈 패턴에 의한 성능저하의 개선효과 실험 결과 비교

학습패턴수	FMM 신경망		제안된 모델	
	에러패턴	에러율	에러패턴	에러율
30	6	0.00893	6	0.00885
60	4	0.00385	3	0.00355
90	5	0.00189	4	0.00174
120	2	0.00158	1	0.00144
150	2	0.00130	1	0.00118

또 다른 실험으로 학습 패턴중에 포함되는 노이즈 패턴수를 점진적으로 증가시켰을 때의 분류 성능의 변화를 비교하였다. 표 3은 노이즈 패턴 수를 1에서 3개까지 증가시킨 실험 결과이다. 이 결과에서 보인 바와 같이 제안된 신경망은 노이즈에 강한 분류 성능을 보였으며, 그 개선 효과는 많은 수의 노이즈가 포함될수록 증가함을 보였다.

6. 결론

본 연구에서는 기존의 FMM 신경망 모델의 학습 알고리즘의 특성에서, 하이퍼박스 축소 과정으로 인해

발생하는 정보의 손실을 최소화하기 위하여 새로운 모델을 제안하였다.

표 3. 노이즈 증가에 대한 실험결과

노이즈 패턴수	FMM 신경망		제안된 모델	
	에러패턴	에러율	에러패턴	에러율
1	1	0.00160	1	0.00160
2	4	0.00192	1	0.00156
3	7	0.00227	2	0.00165

이를 위하여 가중치 요소가 반영되는 하이퍼박스 소속함수를 정의하였으며, 가중치에 대한 학습과정을 새롭게 제안하였다. 기존의 FMM 신경망이 갖는 하이퍼박스 확장 및 생성, 중첩테스트, 축소의 과정을 변형하여, 중첩테스트, 축소의 과정을 생략할 수 있게 함으로써 축소에 의한 부작용을 개선하였다. 이를 위해 학습과정에서 임의의 하이퍼박스에 대한 특징별 빈도수 정보를 특징범위의 최소점, 최대점과 함께 유지하도록 하이퍼박스 특성을 새롭게 정의 하였다. 또한 특징별 중요도에 따라 분류과정에서 서로 다른 값이 적용되도록 소속 함수에 가중치 개념을 도입하였다. 실제 데이터를 사용한 분류 실험에서 제안된 모델은 기존의 모델보다 개선된 분류 성능을 보였으며, 비정상적인 학습패턴의 영향에 대한 실험과 패턴수의 변화에 따른 분류성능의 평가에서도 기존 모델보다 향상된 분류 성능을 보였다. 향후 연구로는 빈도요소 외에 특징값의 무게중심점을 학습과정에서 유지하고, 이를 하이퍼박스 확장 및 생성과정에서 조정되도록 하여 학습패턴집합에서 특징값의 분포를 고려한 분류 기능을 지원하는 방법에 관해 연구할 예정이다.

\* 이 연구는 과학기술부 뇌과학 연구개발사업으로 수행되었음

참고문헌

- [1] P. Simpson, "Fuzzy Min-Max Neural Networks-Part 1:Classification,"IEEE Transactions on Neural Networks, Vol.3, No.5,pp.776-786,1992.
- [2] B. Gabrys and A. Bargiela, "General Fuzzy Min-Max Neural Network for Clustering and Classification,"IEEE Transactions on Neural Networks, Vol.11, No.3, 2000.
- [3] Ho J. Kim, Tae W. Ryu, Thai T. Nguyen, Joon S. Lim, S. Gupta, "A Weighted Fuzzy Min-Max Neural Network for Pattern Classification and Feature Extraction," Lecture Notes in Computer Science, LNCS 3046, pp.791-798, 2004.
- [3] 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.
- [4] Hung-Pin Chiu, Din-Chang Tseung, "Invariant Handwritten Chinese Character Recognition using Fuzzy Min-Max Neural Networks," Pattern Recognition Letters, Vol.18, pp.481-491, 1997.
- [5] Jayanta Basak, Rajat K. De, Sankar K. Pal, "Unsupervised Feature Selection using a Neuro-Fuzzy Approach," Pattern Recognition Letters, Vol.19, pp.997-1006, 1998.
- [6] 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.