

패턴 분류 문제에서 가중치를 고려한 퍼지 최대-최소 신경망

(A Weighted Fuzzy Min-Max Neural Network for Pattern Classification)

김 호 준 [†] 박 현 정 ^{**}
(Ho-Joon Kim) (Hyun-Jung Park)

요 약 본 연구에서는 패턴 분류문제를 위하여 가중치 개념을 갖는 퍼지 최대-최소 신경망 모델을 제안한다. 제안된 모델은 기존의 FMM신경망 모델에 대하여 학습데이터에 포함되는 특징값의 빈도요소를 효과적으로 반영할 수 있도록 수정한 구조를 갖는다. 본 논문에서는 제안된 모델에 대하여 하이퍼박스 소속함수로 정의되는 새로운 활성화 특성과 학습알고리즘을 정의한다. 학습알고리즘은 하이퍼박스 생성 및 확장, 중첩 테스트, 하이퍼박스 축소의 3단계 과정으로 이루어지며, 각 과정에서 특징값의 빈도요소를 고려하여 가중치값을 갱신하는 규칙이 새롭게 정의된다. 본 연구에서는 또한 제안된 모델의 응용으로서 특징분석 기법을 제안한다. 이를 위하여 특징값, 특징유형, 하이퍼박스, 패턴클래스 상호간 연관도 요소를 4가지 유형의 척도로 정의하여, 주어진 패턴분류 문제에서 각 특징의 상대적 중요도를 평가할 수 있도록 한다. 아이리스 데이터와 클리블랜드 의료데이터에 대한 분류문제에 적용한 실험결과를 통하여 제안된 방법의 타당성을 고찰하였다.

키워드 : 특징 추출, 패턴 인식, 퍼지, 퍼지 최대-최소 신경망, 하이퍼박스

Abstract In this study, a weighted fuzzy min-max (WFMM) neural network model for pattern classification is proposed. The model has a modified structure of FMM neural network in which the weight concept is added to represent the frequency factor of feature values in a learning data set. First we present in this paper a new activation function of the network which is defined as a hyperbox membership function. Then we introduce a new learning algorithm for the model that consists of three kinds of processes: hyperbox creation/expansion, hyperbox overlap test, and hyperbox contraction. A weight adaptation rule considering the frequency factors is defined for the learning process. Finally we describe a feature analysis technique using the proposed model. Four kinds of relevance factors among feature values, feature types, hyperboxes and patterns classes are proposed to analyze relative importance of each feature in a given problem. Two types of practical applications, Fisher's Iris data and Cleveland medical data, have been used for the experiments. Through the experimental results, the effectiveness of the proposed method is discussed.

Key words : Feature selection, Pattern Classification, Fuzzy min-max neural network, Fuzzy, Hyperbox

1. 서 론

일반적으로 패턴 분류 기법의 핵심은 주어진 문제에

서 특징공간을 결정공간으로 구분해 내는 과정이다. 지난 수십 년간 이러한 패턴 분류에 관한 다양한 이론들이 소개된 바 있는데, 최근의 연구에서는 인공 신경망을 사용하여 패턴 분류의 성능을 향상시키는 연구와 분류에 관한 명시적 지식표현을 위한 규칙생성 기법 등이 연구되고 있다[1-10].

퍼지 최대 최소(Fuzzy Min-Max(FMM)) 신경망 은 1992년 Simpson 등에 의해 제안된 패턴분류 모형으로 간결하면서도 강력한 학습기능을 지원한다[1]. 그 후

· 본 연구는 산업자원부 뇌신경정보학 연구사업 및 정보통신부의 유비쿼터스컴퓨팅 및 네트워크 원천기술개발사업의 지원에 의한 것임

[†] 종신회원 : 한동대학교 전산전자공학부 교수

hjkim@handong.edu

^{**} 정 회 원 : 삼성전자 프린팅사업부 연구원

hjpark79@empal.com

논문접수 : 2005년 11월 4일

심사완료 : 2006년 6월 26일

Gabrys등은 FMM 신경망의 내부 연산과 데이터의 표현 형태를 확장하여 일반화된 모델을 제안하였다[4]. 이들 방법에서 각 입력 패턴은 패턴 공간내의 퍼지집합 또는 보통집합(crisp set)의 형태로 정의되는 영역, 이른바 하이퍼박스(hyperbox)로서 정의된다. 그러나 특징값과 하이퍼박스 간의 관계를 항상 균일한 가중치로 유지함으로써 특징값의 발생 빈도와 분포정보를 효과적으로 반영할 수 없다는 점이 약점으로 지적된다. 이는 노이즈 등 비정상적 데이터에 대하여 바람직하지 않은 학습효과를 보일 수 있게 한다.

본 연구에서는 기본적으로 FMM신경망에 가중치 개념을 적용하여 확장된 모델을 고려한다. 하이퍼 퍼지 집합 개념을 이용한 FMM신경망은 내부적으로 최소점과 최대점으로 정의되는 퍼지 구간형태의 멤버쉽함수를 사용하여 다차원 패턴 공간의 한 영역을 하이퍼박스로 정의한다. 제안된 모델의 활성화 특성은 학습데이터 집합으로부터 특징값의 빈도에 따른 상대적인 학습효과의 변화를 반영할 수 있다. 본 논문에서는 이러한 수정된 FMM 모델의 동작특성을 소개하고 이를 패턴분류를 위한 특징분석 과정에 활용하는 방법론을 제안한다. 이와 같이 패턴 분류 문제에서 특징의 종류에 따라 상대적 중요도를 판별하는 일과 개별 특징값이 특정 클래스에 주는 영향을 분석하여 최적의 특징 집합을 선정하는 연구는 실행시간과 분류 성능을 최적화 시키는 관점에서 매우 중요한 의미를 갖는다[11,12].

본 논문의 구성으로서 2절에서는 제안된 모델의 기본이 되고있는 FMM 신경망에 관하여 소개하며, 3절에서는 FMM 신경망에 가중치를 고려한 소속 함수와 학습 알고리즘에 관하여 4절에서는 제안하는 모델의 활용으로써 특징 선정기법에 관하여 구체적으로 기술한다. 5절에는 패턴 분류 문제에서 잘 알려진 아이리스 데이터와 클리브랜드 의료데이터[13]를 이용한 결과를 실었으며, 6절에서 결론을 맺는다.

2. FMM 신경망

2.1 FMM 신경망의 구조

FMM신경망은 하이퍼박스 퍼지 집합으로 구성되는 패턴분류 모델이다[1]. 하이퍼박스란 n 개의 특징에 대하여 완전 클래스 멤버십을 가진 패턴들로 이루어진 n 차원 패턴 공간 영역이다. 최소값과 최대값의 범위과 그 가장자리에서 퍼지 멤버십을 갖는 퍼지 구간(fuzzy interval)들로서 이루어진 하이퍼박스와 그에 상응하는 멤버십 함수는 특징값의 집합과 패턴클래스 간의 사상을 지원한다. FMM 패턴분류 신경망 모델은 그림 1과 같다. 그림에서 보는 바와 같이 FMM 분류기 구현하기 위한 신경망은 3단계 계층으로 구성된다. 그림에서 F_B

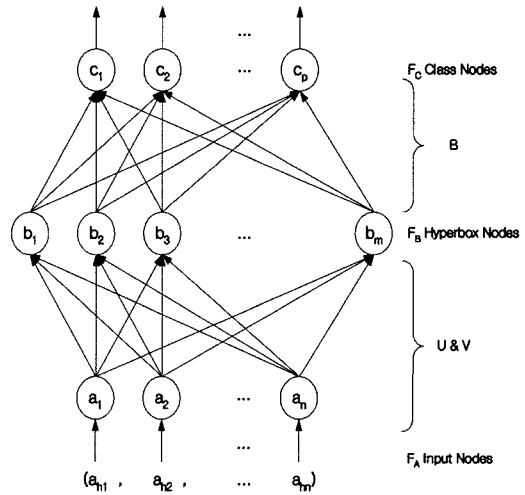


그림 1 FMM 패턴 분류 신경망의 구조

계층의 각 노드는 하이퍼박스 퍼지 집합을 나타내고 있으며 F_A 와 F_B 노드 사이의 연결은 각 하이퍼박스의 최소값과 최대값을 의미한다. 따라서 F_B 계층에 속한 노드의 활성화 특성은 하이퍼박스 소속함수로 정의된다. F_C 계층은 출력 계층으로 각 노드는 각 패턴 클래스와 일치하여 해당 클래스에 적합한 정도를 나타낸다.

2.2 FMM 신경망의 활성화 특성

입력의 하이퍼박스 j 에서 h 번째 입력 패턴 A_h 에 대한 활성화 값은 이 입력패턴에 대한 해당 소속함수 $b_j(A_h)$ ($0 \leq b_j(A_h) \leq 1$)의 값으로 산출된다. 다시 말해, 각 차원별로 특징값이 하이퍼박스의 최소-최대 범위 가장 자리에서 어느 정도 떨어져 있는지와 관련하여 산출된다. $b_j(A_h)$ 가 1에 가까워질 수록 입력 패턴은 하이퍼박스에 속하는 정도가 커지며 소속값이 1로 표현되는 경우 하이퍼박스에 속한다고 할 수 있다. 이러한 모든 조건을 만족하는 하이퍼박스 소속함수는 최소조건 항과 최대조건 항의 보수의 합으로 나타내며 아래와 같이 정의된다.

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

식에서 $A_h = (a_{h1}, a_{h2}, \dots, a_{hn})$ 는 h 번째 입력 패턴으로 n 개의 특징으로 이루어지며 $U_h = (u_{h1}, u_{h2}, \dots, u_{hn})$ 는 b_j 의 최소점을, $V_h = (v_{h1}, v_{h2}, \dots, v_{hn})$ 는 b_j 의 최대점을 의미한다. γ 는 하이퍼박스 특징범위 가장자리에서 퍼지소속함수의 기울기를 결정하는 매개변수이다.

2.3 학습 알고리즘

FMM 신경망 학습 알고리즘은 확장과 축소 처리 과정이다. 학습 패턴 집합 D 는 $\{X_h, d_h\}$ 쌍으로 이루어진 m 개의 원소를 갖는 집합이다. 여기서 $X_h = (x_{h1}, x_{h2}, \dots, x_{hn})$

$\in \mathcal{I}^n$ 는 입력 패턴이며 $d_h \in \{1, 2, \dots, m\}$ 는 m 개 클래스의 색인(index)이다. 학습과정은 학습 패턴 집합 D 에서 학습패턴을 선택하고 그 입력된 패턴을 포함하도록 확장될 수 있는 동일 클래스의 하이퍼박스를 찾아 확장한다. 확장 기준을 만족하는 하이퍼박스가 없는 경우 새로운 하이퍼박스를 신경망에 추가한다. 하이퍼박스 확장시 다른 클래스의 하이퍼박스와 중첩 현상이 발생할 수 있다. 중첩된 부분을 제거하기 위해 하이퍼박스 축소 과정이 필요하다.

요약하면, FMM 신경망 학습 알고리즘은 세가지 과정을 가진다.

- 1) 하이퍼박스 생성/확장 : 학습패턴을 포함하는 하이퍼박스가 있다면 하이퍼박스를 확장
- 2) 중첩 테스트 : 다른 클래스의 하이퍼박스와와의 중첩이 존재하는지 테스트
- 3) 하이퍼박스 축소 : 중첩 부분 제거하기 위한 하이퍼박스 축소

FMM 신경망의 학습알고리즘에 대하여 자세히 설명하겠다. 하이퍼박스의 확장 기준은 아래와 같다.

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

확장기준을 만족하는 하이퍼박스 b_j 가 존재한다면 하이퍼박스의 최대값과 최소값을 아래 수식 (3)를 사용하여 재조정하고 존재하지 않는다면 새로운 하이퍼박스를 생성한다.

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

하이퍼박스 생성/확장 과정 후에 각 차원별로 다른 클래스에 속하는 하이퍼박스와와의 중첩 영역이 발생하였는지 비교하고 중첩 영역이 발생한 경우 이를 제거하기 위해 각 하이퍼박스를 축소하여 패턴 분류의 모호성을 해결한다. 하이퍼박스 B_j 가 확장하여 다른 클래스에 속하는 하이퍼박스 B_k 와 중첩영역이 발생하였다면 가장 적은 중첩이 발생한 차원의 색인(index)를 구하여 축소 과정을 수행함으로 하이퍼박스의 손실을 최소화한다. 중첩은 아래와 같은 4가지 경우로 나누어 고려하였으며 δ 를 유지하여 가장 적은 손실이 발생하는 차원을 구한다. 다시 말해서 δ 는 초기값을 1.0으로 하여 점진적으로 작은 값을 선택하게 되며, 궁극적으로 가장 최소한의 축소 과정으로 하이퍼박스 영역을 재조정하는 방법을 선택하게 된다.

$$\begin{aligned} \text{case1: } & u_{ji} < u_{ki} < v_{ji} < v_{ki} & (4) \\ & \delta^{new} = \min(v_{ji} - u_{ki}, \delta^{old}) \\ \text{case2: } & u_{ki} < u_{ji} < v_{ki} < v_{ji} \\ & \delta^{new} = \min(v_{ki} - u_{ji}, \delta^{old}) \end{aligned}$$

$$\begin{aligned} \text{case3: } & u_{ji} < u_{ki} < v_{ki} < v_{ji} \\ & \delta^{new} = \min(\min(v_{ki} - u_{ji}, v_{ji} - u_{ki}), \delta^{old}) \end{aligned}$$

$$\begin{aligned} \text{case4: } & u_{ki} < u_{ji} < v_{ji} < v_{ki} \\ & \delta^{new} = \min(\min(v_{ji} - u_{ki}, v_{ki} - u_{ji}), \delta^{old}) \end{aligned}$$

위의 4가지 경우 중에 하나라도 만족하는 경우 하이퍼박스 중첩이 발생하였으므로 중첩영역에서의 패턴 분류 모호성을 제거하기 위해 축소과정을 수행해야 한다. 하이퍼박스 축소과정은 위에서 보인 중첩의 4가지 형태에 따라 다르게 적용되며 3번째와 4번째 경우는 중첩영역을 갖는 두 하이퍼박스의 중첩 형태에 따라 2가지 경우로 나누어 적용한다. 이는 아래 식 (5)와 같이 정의되었다.

$$\begin{aligned} \text{case1: } & u_{j\delta} < u_{k\delta} < v_{j\delta} < v_{k\delta} & (5) \\ & v_{j\delta}^{new} = u_{k\delta}^{new} = \frac{v_{j\delta}^{old} + u_{k\delta}^{old}}{2} \end{aligned}$$

$$\begin{aligned} \text{case2: } & u_{k\delta} < u_{j\delta} < v_{k\delta} < v_{j\delta} \\ & v_{k\delta}^{new} = u_{j\delta}^{new} = \frac{v_{k\delta}^{old} + v_{j\delta}^{old}}{2} \end{aligned}$$

$$\begin{aligned} \text{case3a: } & u_{j\delta} < u_{k\delta} < v_{k\delta} < v_{j\delta} \text{ and } (v_{k\delta} - u_{j\delta}) < (v_{j\delta} - u_{k\delta}) \\ & u_{j\delta}^{new} = v_{k\delta}^{old} \end{aligned}$$

$$\begin{aligned} \text{case3b: } & u_{j\delta} < u_{k\delta} < v_{k\delta} < v_{j\delta} \text{ and } (v_{k\delta} - u_{j\delta}) > (v_{j\delta} - u_{k\delta}) \\ & v_{j\delta}^{old} = u_{k\delta}^{old} \end{aligned}$$

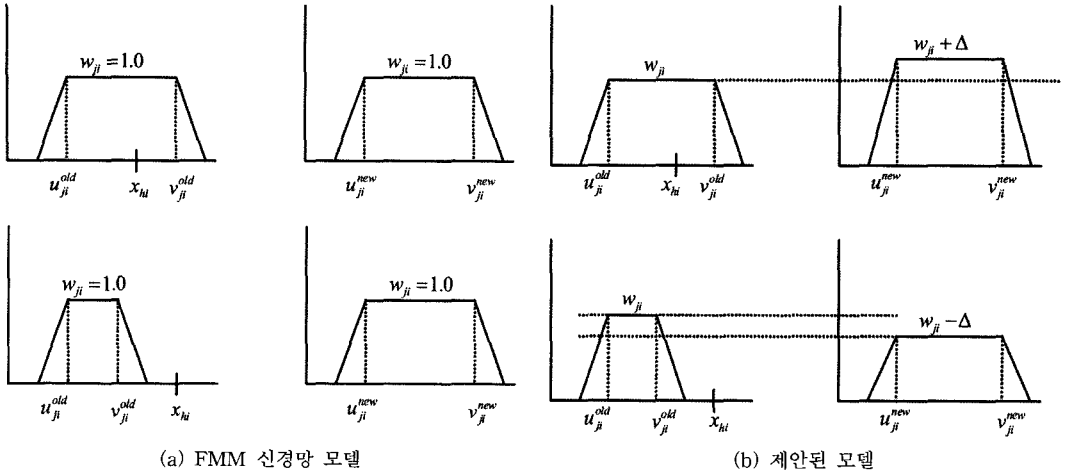
$$\begin{aligned} \text{case4a: } & u_{k\delta} < u_{j\delta} < v_{j\delta} < v_{k\delta} \text{ and } (v_{k\delta} - u_{j\delta}) < (v_{j\delta} - u_{k\delta}) \\ & v_{k\delta}^{new} = u_{j\delta}^{old} \end{aligned}$$

$$\begin{aligned} \text{case4b: } & u_{k\delta} < u_{j\delta} < v_{j\delta} < v_{k\delta} \text{ and } (v_{k\delta} - u_{j\delta}) > (v_{j\delta} - u_{k\delta}) \\ & u_{k\delta}^{new} = v_{j\delta}^{old} \end{aligned}$$

3. 가중치를 고려한 FMM 신경망

앞 절에서 설명한 기존 모델의 소속함수는 학습패턴 집합에서 관찰되는 특징값의 범위만 고려할 뿐 그 빈도를 고려하지 않았다. 이는 극소수의 노이즈나 비정상적인 패턴이 학습패턴에 포함되는 경우에도 민감한 성능 저하를 보일 수 있다. 이를 보완하기 위해 본 연구에서는 빈도요소와 가중치를 고려한 형태로 확장된 모델을 제시한다. 학습 과정에서 특징 값의 발생 빈도에 따라 가중치 값을 하이퍼박스의 최소, 최대 범위와 함께 유지하고 학습한다. 각 가중치 값은 하이퍼박스가 생성되는 시점에 초기값 1.0으로 설정된다. 하이퍼박스 학습과정 중에 가중치를 항상 1.0으로 유지시킨다면 이는 기존 FMM 신경망과 동일한 특성을 갖는다. 즉 기존의 모델은 모든 특징에 대하여 균일한 가중치를 고려하는 개념으로 볼 수 있으므로 제시한 모델은 기존 FMM 모델에 대한 일종의 일반화로 고려할 수 있다. 그림 2는 이러한 가중치 개념의 비교를 보인 것이다.

그림 2에서 보인바와 같이 FMM 신경망 모델은 학습



(a) FMM 신경망 모델 (b) 제한된 모델

그림 2 하이퍼박스 확장에 따른 가중치값 조정 비교

데이터의 발생 빈도나 하이퍼박스의 범위확장에 상관없이 동일한 가중치를 보이고 있는 반면, 제안된 모델은 학습 데이터의 발생 빈도 및 하이퍼박스에 미치는 영향을 고려하여 가중치를 수정하고 있다. 이미 생성된 하이퍼박스의 범위내에서 다시 학습 패턴이 주어지는 경우 가중치를 높여주고, 새로운 학습 데이터에 의해 하이퍼박스가 확장된 경우 하이퍼박스의 최대값과 최소값의 변화에 의해 범위가 넓어지기 때문에 가중치를 낮게 수정하여 준다. 이러한 가중치 개념의 적용을 위하여 수정된 하이퍼박스 소속함수 및 학습방법에 대하여 다음 절에서 설명한다.

3.1 가중치를 고려한 FMM 신경망의 소속함수

가중치를 고려한 FMM 신경망의 소속함수는 아래와 같다.

$$b_j(A_h) = \frac{1}{\sum_{i=1}^n w_{ji}} \cdot \sum_{i=1}^n w_{ji} [\max(0, 1 - \max(0, \gamma \min(1, a_{hi} - v_{ji}))) + \max(0, 1 - \max(0, \gamma \min(1, u_{ji} - a_{hi}))) - 1.0] \quad (6)$$

식에서 w_{ji} 는 j 번째 하이퍼박스와 i 번째 특징사이의 연결 가중치를 의미하며, 기타 표기는 앞절에서의 표기와 동일하다. 식에서 보인 바와 같이 가중치를 고려한 FMM 신경망에서는 하이퍼박스와 특징과의 관계가 w_{ji} 로 각 특징에 가중치로 적용되었다. 학습과정에서 빈도 변수값이 최종적으로 유지되면 이로부터 w_{ji} 값이 아래 수식 (7)에 의해 산출된다. 식에서 f_{ji} 는 하이퍼박스 j 에서 i 번째 특징범위에 대한 발생 빈도를 의미한다.

$$w_{ji} = \frac{\alpha f_{ji}}{R} \quad (7)$$

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

가중치는 식 (7)에서 보인바와 같이 특징의 발생빈도에 비례하고 발생 범위에 반비례한 값으로 결정되며 특징의 범위가 단일 포인트로 나타나거나 범위가 매우 작은 경우에 가중치가 지나치게 큰 값이 되는 것을 방지하기 위해 일정 범위 이하는 지정된 값 s 의 범위로 대치하였다. s 는 문제에서 주어지는 패턴 데이터의 형태에 따라 결정되는 파라미터로 0보다 큰, 0 근처의 값으로 정해준다. 이와 같이 계산되는 가중치 개념은 좁은 특징 범위에 높은 빈도의 학습패턴이 주어질수록 그 특징은 해당 하이퍼박스에 상대적으로 높은 중요도를 가지는 특징이 된다는 점을 반영한다.

3.2 학습 알고리즘

학습알고리즘은 기존의 FMM 신경망에서 이루어지는 하이퍼박스 생성/확장, 중첩 테스트, 하이퍼박스 축소 3단계에 가중치 개념을 도입하여 수정, 개선되었다. 가중치를 고려한 FMM 신경망의 학습 알고리즘은 기존의 FMM 신경망의 학습 과정에 비하여 학습데이터에 포함될 수 있는 비정상적인 패턴이나 노이즈 패턴의 영향을 최소화할 수 있도록 하였다. 또한 FMM 신경망의 하이퍼박스 축소 과정에서 발생하는 하이퍼박스 왜곡 현상을 분석하여 이를 보완할 수 있는 하이퍼박스 축소 방안을 제시하였다. 이는 노이즈 패턴에 의해 발생된 하이퍼박스 확장과정에서 중첩영역 축소로 발생된 분류 오류를 최소화한다.

3.2.1 하이퍼박스 축소기법의 문제점

기존의 FMM 신경망에서는 서로 다른 클래스에 속해 있는 하이퍼박스 간에 중첩 영역이 발생하는 경우 이러한 영역을 제거하기 위한 축소과정이 요구된다. 본 절에서는 FMM 신경망의 하이퍼박스간의 중첩현상과 중첩

영역을 제거하기 위해 수행된 축소 과정에서 발생하는 하이퍼박스 왜곡 현상에 대해 분석하였다.

FMM 신경망은 중첩의 형태에 따라 두 가지 유형으로 구분된다. 두 하이퍼박스 사이의 중첩 영역이 어느 한 가장자리에서 부분적으로 중첩되는 경우와 하나의 하이퍼박스가 다른 하이퍼박스에 완전히 속하는 완전중첩이다. 이들 각각 유형에 대하여 별도의 축소과정이 적용된다. 부분중첩의 경우 중첩된 두 하이퍼박스가 중첩 부분에 대하여 동일한 크기로 축소되며, 완전중첩의 경우는 범위가 작은 하이퍼박스를 그대로 유지하도록 범위가 큰 하이퍼박스가 축소된다. 이러한 하이퍼박스 축소과정이 그림 3에 나타나 있다.

그림에서 이중실선 부분이 중첩 영역이며 축소될 위치를 점선으로 표시하였다. 부분중첩의 경우 두 하이퍼박스 i 와 j 에 나타난 중첩부분을 제거하게 되는데, 이때 축소되는 크기는 각 하이퍼박스에 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. 완전 중첩의 경우 작은 하이퍼박스 정보를 유지하기 위하여 하이퍼박스를 포함하고 있는 큰 하이퍼박스가 중첩영역을 모두 축소할 뿐 아니라 중첩영역이 아닌 범위(그림 3(b)의 빗금친 영역)에 대한 정보를 함께 손실하게 되어 왜곡이 더욱 심해진다.

3.2.2 하이퍼박스 특성 정의 및 학습 방법

본 절에서는 위에서 언급한 기존 모델의 정보 왜곡을 개선하는 방법론을 고찰한다. 특징값의 빈도요소를 고려하고 하이퍼박스 축소과정의 부작용을 개선할 수 있는 새로운 학습 알고리즘을 제안한다. 이를 위하여 하이퍼박스 B_j 의 특성을 아래와 같이 새롭게 정의한다.

$$B_j = \{X, U_j, V_j, F_j, f(X, U_j, V_j, F_j)\} \quad \forall X \in I^m \quad (8)$$

식에서 F_j 는 하이퍼박스 특징 범위내에 나타나는 패턴의 빈도수를 의미하고, f 는 하이퍼박스 활성화 함수를 의미하며, 그 외 표현은 앞절의 표기와 동일하다. 주어진 패턴에 대하여 동일 클래스에 속하는 기존에 생성된 하이퍼박스들 중에 최대의 소속함수값을 보이는 하이퍼박스를 선택하고, 그에 대해 아래 식을 통하여 확장 여부를 결정한다.

$$\forall i=1..n (\max(v_{ji}, x_{hi}) - \min(u_{ji}, x_{hi}) \leq \theta) \quad (9)$$

즉 식 (9)에서 모든 차원에 대하여 θ 이하의 범위를 가질 경우 확장이 이루어 진다. 이는 기존의 확장 조건에 비하여 어느 한 차원에 대해서만 하이퍼박스의 범위가 커지는 것을 방지하여 준다. 확장조건을 만족하지 못하는 경우 해당패턴의 각 특징값을 초기범위로 하는 새로운 하이퍼박스를 생성하고, 만족하는 경우 아래 수식에 의해 기존의 하이퍼박스의 특징범위를 확장한다.

$$\begin{aligned} f_{ji}^{new} &= f_{ji}^{old} + 1 & \forall i=1,2,\dots,n \\ u_{ji}^{new} &= \min(u_{ji}^{old}, x_{hi}) & \forall i=1,2,\dots,n \\ v_{ji}^{new} &= \max(v_{ji}^{old}, x_{hi}) & \forall i=1,2,\dots,n \end{aligned} \quad (10)$$

위와 같이 확장된 하이퍼박스는 기존의 U_{ji} 와 V_{ji} 값으로 하이퍼박스의 범위만을 학습하는 이외에 F_{ji} 를 함께 학습함으로써 이를 반영한 가중치를 산출하고 이로부터 이후의 학습 단계인 축소 과정에서 나타나는 하이퍼박스 왜곡을 최소화 한다. 하이퍼박스 중첩테스트는 기존의 모델과 같이 4가지 경우로 구분하여 수행하며 정보의 손실을 최소화하기 위하여 가장 작은 범위가 중첩되는 차원에서 축소한다.

제안하는 학습 방법에서는 각 하이퍼박스의 특징의 발생 빈도를 고려한 축소과정을 새롭게 정의한다. 축소 방법은 4가지 경우에 대하여 서로 다른 방법으로 이루어지며 이는 아래 수식과 같이 정의된다. 즉 하이퍼박스 중첩이 발견되고 가장 작은 범위의 중첩이 발생한 차원이 Δ 번째 차원이라면 축소과정은 식 (11)과 같이 정의된다.

$$\begin{aligned} \text{Case 1: } & u_{j\Delta} < u_{k\Delta} < v_{j\Delta} < v_{k\Delta} & (11) \\ & \left\{ \begin{aligned} v_{j\Delta}^{new} &= v_{j\Delta}^{old} - \frac{f_{k\Delta}}{f_{j\Delta} + f_{k\Delta}} (v_{j\Delta}^{old} - v_{k\Delta}^{old}) \\ u_{k\Delta}^{new} &= u_{k\Delta}^{old} + \frac{f_{j\Delta}}{f_{j\Delta} + f_{k\Delta}} (v_{j\Delta}^{old} - v_{k\Delta}^{old}) \\ f_{j\Delta}^{new} &= f_{j\Delta}^{old} * \left(\frac{v_{j\Delta}^{new} - u_{j\Delta}^{new}}{v_{j\Delta}^{old} - u_{j\Delta}^{old}} \right) \\ f_{k\Delta}^{new} &= f_{k\Delta}^{old} * \left(\frac{v_{k\Delta}^{new} - u_{k\Delta}^{new}}{v_{k\Delta}^{old} - u_{k\Delta}^{old}} \right) \end{aligned} \right. \end{aligned}$$

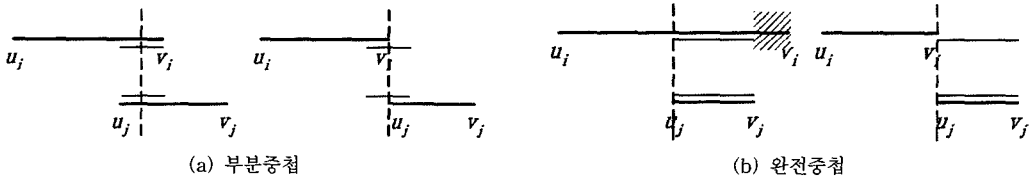


그림 3 FMM 신경망의 축소개념

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

$$\begin{cases} v_{k\Delta}^{new} = v_{k\Delta}^{old} - \frac{f_{j\Delta}}{f_{k\Delta} + f_{j\Delta}} (v_{k\Delta}^{old} - v_{j\Delta}^{old}) \\ u_{j\Delta}^{new} = u_{j\Delta}^{old} + \frac{f_{k\Delta}}{f_{k\Delta} + f_{j\Delta}} (v_{k\Delta}^{old} - v_{j\Delta}^{old}) \\ f_{k\Delta}^{new} = f_{k\Delta}^{old} \cdot \left(\frac{v_{k\Delta}^{new} - u_{k\Delta}^{new}}{v_{k\Delta}^{old} - u_{k\Delta}^{old}} \right) \\ f_{j\Delta}^{new} = f_{j\Delta}^{old} \cdot \left(\frac{v_{j\Delta}^{new} - u_{j\Delta}^{new}}{v_{j\Delta}^{old} - u_{j\Delta}^{old}} \right) \end{cases}$$

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

$$f_{j\Delta}^{new} = f_{j\Delta}^{old} \cdot \left(\frac{v_{j\Delta}^{new} - u_{j\Delta}^{new}}{v_{j\Delta}^{old} - u_{j\Delta}^{old}} \right)$$

$$\begin{cases} u_{j\Delta}^{new} = v_{k\Delta}^{old} & \text{if } (v_{k\Delta} - u_{j\Delta} < v_{j\Delta} - u_{k\Delta}) \\ v_{j\Delta}^{new} = u_{k\Delta}^{old} & \text{otherwise} \end{cases}$$

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

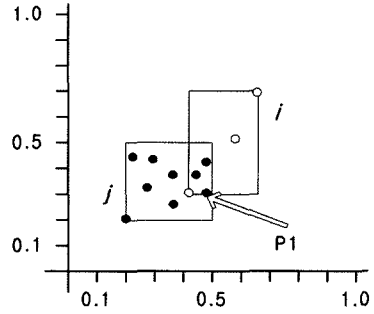
$$f_{k\Delta}^{new} = f_{k\Delta}^{old} \cdot \left(\frac{v_{k\Delta}^{new} - u_{k\Delta}^{new}}{v_{k\Delta}^{old} - u_{k\Delta}^{old}} \right)$$

$$\begin{cases} u_{k\Delta}^{new} = v_{j\Delta}^{old} & \text{if } (v_{j\Delta} - u_{k\Delta} < v_{k\Delta} - u_{j\Delta}) \\ v_{k\Delta}^{new} = u_{j\Delta}^{old} & \text{otherwise} \end{cases}$$

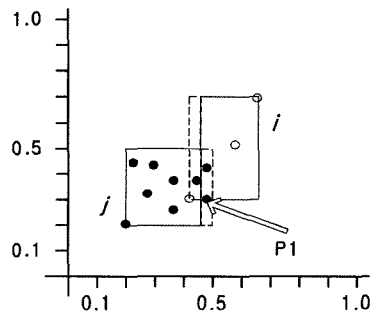
그림 4는 FMM 신경망과 제안된 모델의 축소과정에 대한 비교를 나타낸 것이다. 그림에서 각 사각형은 임의의 정규화된 2차원 특징공간에서의 하이퍼박스를 의미하며 현재 최근 학습된 데이터에 의하여 그림 4(a)와 같이 하이퍼박스 i 와 j 사이에 중첩영역이 발생하였다고 가정한다.

중첩영역을 제거하기 위하여 각 모델은 축소과정을 수행하게 되며 FMM 신경망의 경우 그림 4(b)와 같은 형태로 중첩영역이 제거된다. 그 결과 중첩영역에 나타난 검은점으로 표시된 하이퍼박스 j 에 속한 학습 데이터들이 하이퍼박스 i 에 대하여 반응함으로써 분류 오류가 발생하게 된다. 뿐만 아니라 PI 으로 표시된 학습 패턴의 경우에도 패턴의 발생 빈도를 고려하지 않고 오직 하이퍼박스와와의 거리에 따라 소속함수 값이 결정 되어 하이퍼박스 i 에 반응하게 된다.

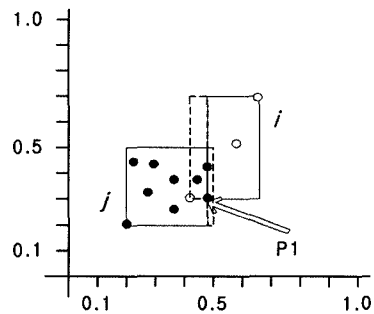
반면 제안된 모델에서는 특정 하이퍼박스 범위에서 나타난 특징의 발생 빈도를 반영하여 축소 하기 때문에 그림 4(c)와 같은 형태로 하이퍼박스 중첩 영역이 제거 됨으로써 학습패턴의 오분류를 최소화 한다. 다시 말해서 PI 과 같이 두 개 이상의 하이퍼박스에 속해있는 패턴의 경우에도 각 하이퍼박스에 대한 특징별 빈도수의



(a) 하이퍼박스 중첩현상



(b) FMM 신경망의 축소



(c) 제안된 모델의 축소

그림 4 FMM 신경망과 제안된 모델의 축소과정 비교

작용, 즉 가중치에 의하여 서로 다른 반응 정도를 산출하므로, 학습과정에서 상대적으로 패턴의 발생빈도가 높은 하이퍼박스에 속하는 것으로 판정하게 된다. 이는 학습패턴에서 각 클래스에 속하는 표준 패턴의 빈도가 균일하게 주어진다고 가정했을 때 노이즈나 비정상적인 패턴으로부터의 영향을 개선할 수 있는 유용한 특징이 된다.

4. 특징 선정 기법

본 절에서는 학습된 신경망으로부터 특징 패턴 클래스에 대하여 각 특징의 연관성 요소를 판별하는 방법을

소개한다. 앞 절에서 제안한 신경망의 학습 방법을 통해 학습된 가중치 값으로부터 각 특징과 클래스 간의 관계를 해석해 낼 수 있으며, 이는 특징 종류에 대한 유용성과 특징값에 대하여 하이퍼박스에 대한 상대적인 연관도를 서로 다른 값으로 판별해 낼 수 있게 한다.

본 연구에서는 그림 5와 같이 특징 집합과 클래스 집합간의 4종류의 관계를 고려하여 특징의 연관도 요소 (Relevance Factor: RF)를 정의한다.

$RF1(x_i, B_j)$: 특징값 x_i 와 하이퍼박스 B_j 사이의 연관도 요소

$RF2(x_j, C_k)$: 특징값 x_j 와 클래스 C_k 사이의 연관도 요소

$RF3(X_i, C_k)$: 특징유형 X_i 와 클래스 C_k 사이의 연관도 요소

$RF4(X_i)$: 주어진 문제에서 특징 X_i 의 중요도 판별위의 RF에서 각 표기는 다음과 같다. X_i 는 i 번째 특징을 의미하며 x_j 는 j 번째 특징값, C_k 는 k 번째 클래스, B_j 는 j 번째 하이퍼박스를 의미한다. N 은 특징의 수를 M 은 클래스의 수를 의미한다.

4.1 특징 값과 패턴 클래스 간의 연관도 요소

특징값과 특정 하이퍼박스간의 상호 연관도는 가중치 값으로 정의될 수 있다. 즉 i 번째 특징과 j 번째 하이퍼박스 간의 상호연관도 요소는 아래와 같이 정의된다.

$$RF1(x_i, B_j) = w_{ij} \tag{12}$$

이러한 연관도를 사용하여 특정 특징값과 임의의 클래스 간의 연관도 요소는 아래 식과 같이 정의할 수 있다.

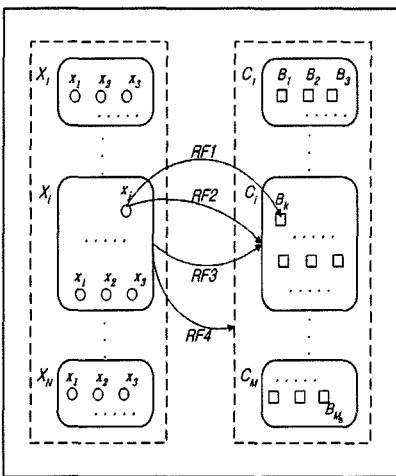


그림 5 특징과 클래스 간의 4가지 연관도 요소

$$RF2(x_i, C_k) = \left(\frac{1}{N_k} \sum_{B_j \in C_k} S(x_i, (u_{ji}, v_{ji})) \cdot w_{ji} \right) - \frac{1}{(N_B - N_k)} \sum_{B_j \notin C_k} S(x_i, (u_{ji}, v_{ji})) \cdot w_{ji} \tag{13}$$

식에서 상수 N_B 는 총 하이퍼박스의 개수를 의미하며 N_k 는 클래스 k 에 속하는 하이퍼박스의 개수이다. 식에서 x_i 는 최종적으로 생성된 하이퍼박스에 나타난 특징 범위로 특징 i 가 속한 하이퍼박스의 최대값을 x_i^L , 최소값을 x_i^U 로 정의하고 두 퍼지 구간의 유사도(Similarity : S)를 아래와 같이 정의한다.

$$S(x_i, (u_i, v_i)) = S((x_i^L, x_i^U), (u_i, v_i)) = \frac{Overlap((x_i^L, x_i^U), (u_i, v_i))}{Max(x_i^U - x_i^L, v_i - u_i)} \tag{14}$$

유사도 수식에서 두 퍼지구간이 모두 한점으로 이루어진 경우 분모 $Max(x_i^U - x_i^L, v_i - u_i)$ 가 0이 되는 경우를 고려하여 아래와 같이 정의한다.

$$S((x_i^L, x_i^U), (u_i, v_i)) = \begin{cases} 1 & \text{if } (x_i^L = x_i^U = v_i = u_i) \\ 0 & \text{otherwise} \end{cases} \tag{15}$$

$Max(x_i^U - x_i^L, v_i - u_i)$ 이 0보다 큰 값일 경우 아래 수식을 적용하여 산출한다.

$$Overlap((x_i^L, x_i^U), (u_i, v_i)) = \begin{cases} x_i^U - u_i & \text{if } (x_i^L \leq u_i \leq x_i^U \leq v_i) \\ v_i - u_i & \text{if } (x_i^L \leq u_i \leq v_i \leq x_i^U) \\ x_i^U - x_i^L & \text{if } (u_i \leq x_i^L \leq x_i^U \leq v_i) \\ v_i - x_i^L & \text{if } (u_i \leq x_i^L \leq v_i \leq x_i^U) \\ 0 & \text{otherwise} \end{cases} \tag{16}$$

만일 $RF2(x_j, C_k)$ 가 양의 값을 가지면, 특징값과 패턴 클래스 사이에 자극성(excitatory) 연관성이 있음을 의미한다. 그러나 음의 값을 가진다면 이들 사이에는 억제성(inhibitory) 연관성이 있음을 의미한다. 주어진 클래스에 대하여 중요한 역할을 하는 주요한 특징들의 목록을 계산된 연관성 요소 값으로 사용하여 선정할 수 있다. 이는 각 특징의 종류에 대하여 모든 값과 모든 하이퍼박스에서 연관성요소의 값을 구분하여 합산하고 이를 정렬하여 출력함으로써 이루어 질 수 있다. 이는 학습패턴으로부터 패턴 분류를 위한 명시적인 지식표현으로서 추출 할 수 있는 가능성을 시사한다.

4.2 특징의 유형과 패턴 클래스간의 연관도

특징유형 X 에 대한 클래스 k 의 연관도($RF3$)는 식 (17)과 같이 $RF2$ 의 향으로 정의될 수 있다.

$$RF3(X_i, C_k) = \frac{1}{L_i} \sum_{x_j \in X_i} RF2(x_j, C_k) \tag{17}$$

식에서 L_i 는 i 번째 특징에 속하는 특징값의 개수를 의미한다. 다시 말해, $RF3$ 는 특정 클래스에 속한 특징 집합들의 상대적 연관도에 대한 평균치를 의미한다. 이는 패턴 분류문제에서 학습이 완료된 신경망으로부터 일종의 지식을 추출할 수 있게 한다. 예를 들어 특정 클래스

를 분류할 때 가장 중요하게 반영되는 특징의 종류를 서열화 또는 선별 선택할 수 있게 한다.

이들 연관도를 활용하여 주어진 문제의 분류 과정에서 특정 특징의 상대적 중요도를 평가할 수 있다. 이는 $RF3$ 를 사용하여 아래와 같이 정의된다.

$$RF4(X_i) = \frac{1}{M} \sum_{j=1}^M RF3(X_i, C_j) \quad (18)$$

즉 특정 특징의 상대적 중요도는 식 (18)과 같이 각 개별 클래스에 대하여 평균적인 중요도를 산출 함으로써 평가된다. 이는 주어진 문제에서 가장 효과적인 특징의 집합을 선별할 수 있게 한다. 따라서 패턴 분류 신경망을 설계하는 과정에서 성능저하를 최소화 하면서도 최적의 입력노드 수를 결정할 수 있게 함으로써 성능과 계산량 및 신경망의 규모를 최적화 할 수 있게 한다. 또한 실제계로부터 추출한 원시 데이터로부터 학습 패턴 집합을 효과적으로 선별 할 수 있게 할 것이다.

5. 실험 및 결과

제한된 모델의 성능을 평가하고 특징의 중요도 판별 기법을 적용하기 위해, 본 실험에서는 패턴 분류 문제에서 잘 알려진 아이리스 데이터와 클리브랜드 의료 진단 데이터를 사용하였다. 아이리스 데이터는 총 3개의 클래스에 각각 50개의 패턴으로 구성되며 각 패턴은 4종류의 특징으로 이루어진다. 클리브랜드 의료 진단 데이터는 총 5개의 클래스에 총 297개의 패턴에 대하여 13개의 특징으로 이루어진다.

FMM모델과 제안된 모델의 성능 비교를 위하여 평가 기준을 분류의 오류 개수와 아래 수식으로 표현되는 오류누적합을 적용하였다.

$$E = \frac{1}{pm} \sum_{i=1}^p \sum_{k=1}^m |c_{ik} - d_{ik}| \quad (19)$$

식에서 p 는 테스트 패턴의 수, m 은 클래스의 수이며 c_{ik} 와 d_{ik} 는 i 번째 입력패턴에 대하여 k 번째 클래스의 실제 출력값과 기대값을 의미한다. 이로써 정확한 분류여부뿐만 아니라 오분류에 대해서도 얼마나 바람직한 값의 변화를 보였는가를 평가하였다.

5.1 가중치를 고려한 FMM 신경망의 성능 평가

5.1.1 성능 평가

첫 번째 실험은 FMM 신경망과 제안된 신경망 모델에 대하여 학습 데이터의 수를 증가시키며 오류 패턴수와 오류 누적합을 비교하였다.

표 1은 아이리스 데이터에 대한 학습패턴수의 증가에 따른 오분류 패턴수와 오류 누적합의 비교이다. 표에 나타난 바와 같이 제안된 모델이 기존의 FMM 신경망에 비해 10~18%가량 개선된 성능을 보였으며, 학습패턴수의 증가에 따라 지속적으로 개선되었다.

이어 실험 데이터를 클리브랜드 의료 데이터를 사용하여 위와 같은 실험을 수행하였다. 표 2는 분류 실험 결과로서 표에서 보인 바와 같이 학습데이터의 수가 증가할수록 제안된 모델이 개선된 성능을 모임을 알 수 있었다.

5.1.2 노이즈/비정상적 데이터에 따른 분류 성능 비교

다음으로 학습데이터에 임의의 노이즈 패턴을 추가하여 분류 성능을 비교하였다. 노이즈 패턴은 학습데이터에서 임의의 1개의 패턴을 의도적으로 바꾸어 다른 클래스에 소속되도록 하였다. 그 결과를 표3에 보였다. 표에 보인바와 같이 주어진 노이즈 패턴이 1개 포함될 경우의 실험에서, 제안된 모델의 분류 성능이 기존의 모델

표 1 학습패턴수 변화에 따른 기존 모델과 제안된 모델의 분류성능비교(Iris 데이터)

학습패턴수	FMM 신경망		제안된모델	
	오류 패턴수	오류 누적합	오류 패턴수	오류 누적합
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

표 2 학습패턴수 변화에 따른 기존 모델과 제안된 모델의 분류성능비교(클리브랜드 의료데이터)

학습패턴수	FMM 신경망		제안된모델	
	오류 패턴수	오류 누적합	오류 패턴수	오류 누적합
50	35	0.785	32	0.684
100	29	0.761	27	0.675
150	21	0.723	19	0.642
200	14	0.654	12	0.580
250	7	0.621	5	0.533

표 3 임의의 노이즈 패턴 1개를 삽입하고 학습패턴수 변화에 따른 분류성능비교(Iris 데이터)

학습패턴수	FMM 신경망		제안된모델	
	오류 패턴수	오류 누적합	오류 패턴수	오류 누적합
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	1	0.00130	1	0.00118

표 4 노이즈 패턴의 증가에 따른 분류 성능 비교(Iris 데이터)

노이즈 패턴수 (학습패턴수)	FMM 신경망		제안된모델	
	오류 패턴수	오류 누적합	오류 패턴수	오류 누적합
1(150)	1	0.00160	1	0.00160
2(150)	4	0.00192	1	0.00156
3(150)	7	0.00227	2	0.00165

보다 개선되었음을 알 수 있었다.

노이즈가 포함된 학습패턴에 대한 분류 성능을 평가한 다른 실험으로 학습패턴 중에 포함되는 노이즈의 패턴수를 1에서 3개로 점진적으로 증가하였을 경우의 분류 성능을 비교하였다. 표 4에 그 결과를 보인바와 같이 제안된 신경망은 기존의 모델보다 노이즈에 강인한 분류 성능을 나타냈으며, 그 개선 효과는 많은 수의 노이즈가 포함 될수록 증가함을 보였다.

5.1.3 하이퍼박스 축소 과정 개선에 따른 성능 비교

본 논문에서는 하이퍼박스 축소 과정에 학습데이터의 특징 빈도수를 고려하도록 개선하였다. 이의 효과를 검증하기 위하여 아이리스 데이터를 사용하여 아래와 같은 실험을 수행하였다.

우선 2차원 특징 공간에 나타나는 각 클래스 별 특징과 하이퍼박스의 분포를 그림 6과 같이 나타내었다. 그림 6에 나타난 것과 같이 *setosa* 클래스는 다른 클래스와의 구별이 명확하여 하이퍼박스 중첩현상이 나타나지 않아 축소 과정에 영향을 미치지 않기 때문에 본 실험에서 제외하여 *versicolor*와 *virginica* 패턴만을 사용하였다.

아이리스 데이터는 각 클래스를 구별하기 위한 4가지 종류의 특징을 가지고 있다. 꽃받침의 길이, 꽃받침의 넓이, 꽃잎의 길이, 꽃잎의 넓이다. 이를 각 $F1, F2, F3, F4$ 로 지정하고 $(F1, F2), (F1, F3), (F1, F4)$ 등으로 총 6가지 특징별 조합을 만들었다.

표 5는 각 특징의 조합별 오류패턴과 오류누적합을 나타낸 것이다. $(F1, F2)$ 의 특징조합을 사용하는 경우 그림 7에 보는 바와 같이 중첩된 영역의 손실이 커서 제안된 모델이 개선된 성능을 보였다. 그러나 $(F3, F4)$ 의 특징 조합을 사용한 경우 그림 8에 보인바와 같이 중첩 영역이 작아 축소에 의한 정보손실이 작기 때문에 기존의 FMM 신경망과 제안된 모델이 거의 유사한 성능을

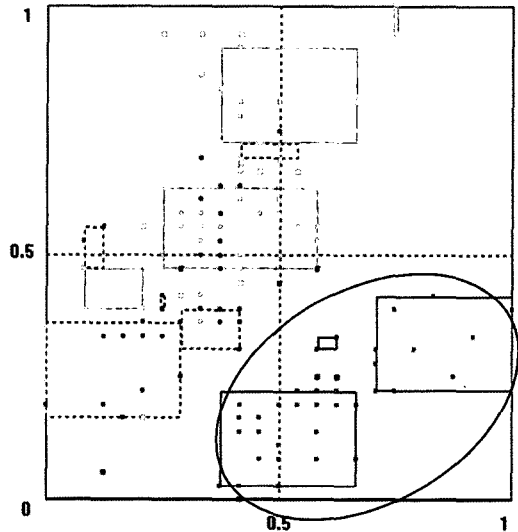


그림 6 2차원 특징공간에서 각 클래스의 특징과 하이퍼박스의 분포

보였다. 이러한 결과에서 축소과정에 의한 성능의 왜곡이 빈도를 고려한 제안된 모델에 의해 보완될 수 있음을 알 수 있다.

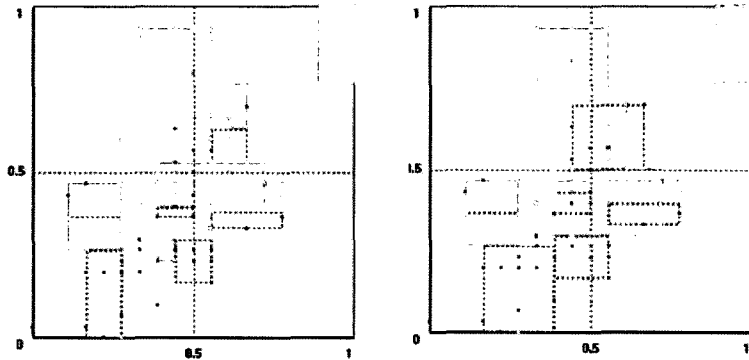
5.2 특징 선정 기법

다음으로 특징의 중요도 판별에 관한 실험을 수행하였다. 아이리스 데이터 분류 문제에서 각 패턴 클래스와 관련된 특징값의 상호연관도 $RF2$ 를 평가하였다. 표에서는 각 클래스 별로 가장 높은 연관도를 갖는 특징을 2개씩 보였다.

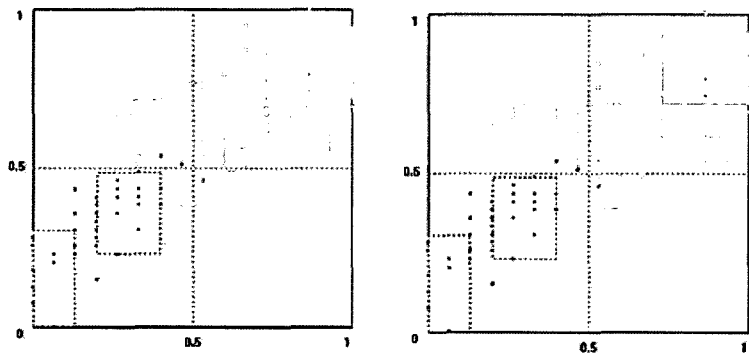
이어서 특징 유형과 클래스와의 관계 $RF3$ 에 대한 실험을 클리브랜드 의료데이터에 대하여 적용하고 그 결과를 표 7에 보였다. 각 클래스에 대하여 연관도가 높은 특징 2개씩을 선택하여 표에 나타내었다. 이는 패턴 분

표 5 특징 조합에 따른 분류 성능 비교(Iris 데이터)

특징 조합	FMM 신경망		제안된 모델	
	오류 패턴수	오류누적합	오류 패턴수	오류누적합
(F1, F2)	32	0.01494	23	0.01051
(F1, F3)	5	0.00539	2	0.00188
(F1, F4)	6	0.00412	6	0.00411
(F2, F3)	9	0.00426	6	0.00294
(F2, F4)	5	0.00283	4	0.00294
(F3, F4)	2	0.00233	2	0.00224



(a) 축소과정 수행 후 (b) 축소과정 제외
그림 7 (F1,F2) 특징집합의 특징 및 하이퍼박스 분포



(a) 축소과정 수행 후 (b) 축소과정 제외
그림 8 (F3,F4) 특징집합의 특징 및 하이퍼박스 분포

표 6 특징값과 특정 클래스 간의 연관도 요소(RF2) (Iris 데이터)

특징값	클래스	RF2
F4 : (0.0, 0.13)	Setosa	0.312
F1: (0.03, 0.22)	Setosa	0.190
F3: (0.51, 0.65)	Versicolor	0.443
F2: (0.13, 0.54)	Versicolor	0.158
F3: (0.65, 0.78)	Virginica	0.272
F2: (0.21, 0.67)	Virginica	0.133

표 7 특징 유형과 특정 클래스간의 연관도 요소(RF3) (클리브랜드 의료데이터)

특징유형	클래스	RF3
F12	0	0.327
F3	0	0.240
F2	1	0.387
F12	1	0.311
F3	2	0.382
F8	2	0.133
F9	3	0.402
F2	3	0.176
F2	4	0.377
F5	4	0.266

류문제에서 특정 클래스의 분류에 가장 영향을 주는 특징의 종류를 구분할 수 있게 함으로써 분류기 설계과정에 유용한 정보로 활용될 수 있을 것이다.

6. 결론 및 향후 과제

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, 주어진 특정 문제에 대하여 특징의 종류별로 상대적 연관도 요소로서 정의되는 중요도를 판별함으로써 좀 더 정제된 지식표현을 얻을 수 있는 가능성을 제시한다. 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. 물론 어떤 문제에서 학습패턴에 대한 사전정보가 없고, 극소수의 빈도로 나타나는 특이한 정상적인 패턴이 존재한다면, 이러한 경우 제안된 모델은 분류성능을 오히려 저하시킬 수 있다. 그러나 일반적인 가정에서는 학습패턴상의 데이터 오류나 노이즈 등의 영향을 개선할 수 있는 유용한 특성이 될 것이다. 또한 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 있을 것이다.

참고 문헌

- [1] P. K. Simpson, "Fuzzy Min-Max neural networks Part 1:classification, IEEE Transaction on Neural Networks," Vol.3, No.5, pp.776-786, 1992.
- [2] 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.
- [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.18, 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, Vo.11, No.3, pp.769-783, 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, No.1, 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.1113-1125, 1998.
- [10] Hung P. Chiu, Din C. Tseng, "Invariant handwritten Chinese 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, pp.997-1006, 1998.
- [12] Jean M. Steppe, Kenneth W. Bauer, Jr. "Improved feature screening in feedforward neural networks," Neurocomputing, Vol.13, pp.47-58, 1996.
- [13] C. L. Blake, and C. J. Merz, UCI Repository of machine learning databases [http://www.ics.uci.edu/~mllearn/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



박 현 정

2003년 한동대학교 전산전자공학부 졸업(학사). 2006년 한동대학교 정보통신학과 졸업(공학석사). 2006년~현재 삼성전자 프린팅사업부 연구원. 관심분야는 인공지능, 패턴분류, 영상처리