수정된 퍼지 최대최소 신경망을 이용한 패턴분류

최형수[0], 정경훈, 김호준
한동대학교 전산전자공학부
coldputer[0]@lycos.co.kr, khjung@handong.edu, hjkim@handong.edu

A Modified Fuzzy Min-Max Neural Network for Pattern Classification

Hyoung Soo Choi[0], Kyeong Hoon Jung, Ho Joon Kim
School of Computer Science and Electronic Engineering, Handong Global University

요 약

본 연구에서는 효과적인 패턴 분류를 위한 방법론으로서 수정된 퍼지 최대최소 신경망 모델을 제안하고 그 유용성을 고찰한다. 제안된 모델에서 각 하이퍼박스는 다차원의 특징공간상에서 한 영역으로 정의되며, 각 특징에 대하여 가중치 개념이 추가된 소속함수를 갖는다. 이는 기존의 FMM 신경망에서 모든 특징에 대하여 균일하게 고려되었던 특징의 상대적 중요도를 서로 다른 값으로 반영할 수 있게 한다. 본 연구에서는 제안된 모델의 동작특성 및 학습방법을 소개하며, 실제 패턴 분류문제에 적용한 실험결과를 통하여 제안된 이론의 타당성을 평가한다.

1. 서 론

일반적으로 패턴분류 기법의 핵심은 주어진 문제에서 특징공간(feature space)을 결정공간(decision space)으로 구분해 내는 과정이다. 지난 수십년 간 이러한 패턴 분류에 관한 다양한 이론들이 소개된 바 있는데, 최근에 활발하게 연구되고 있는 방법론 중의 하나로 신경망 및 퍼지집합 이론을 들 수 있다[1-5].

퍼지 최대최소 신경망(Fuzzy Min-Max(FMM) Neural Network)은 1992년 Simpson등에 의하여 최초로 제안된 이후에 내부연산 및 데이터의 표현형태에 대하여 일반화된 모델이 소개된 바 있으며, 그 후 패턴분류 및 인식에 관한 많은 응용분야에 적용하는 연구가 활발하게 이루어 지고 있다[2].

FMM 신경망은 다차원 특징공간상의 특정 영역으로서 정의되는 하이퍼박스(hyperbox) 퍼지집합 개념을 사용한 패턴분류 모형이다. 이는 패턴분류 문제에서 매우 간결하면서도 강력한 학습기능을 지원한다[1]. 그러나 퍼지 소속함수는 학습패턴집합에서 관찰되는 특징값의 범위만을 고려할 뿐 그 빈도를 고려하지 않는다. 이는 노이즈나 비정상적인 패턴이 학습패턴에 포함되는 경우 성능저하를 보일 수 있다.

이에 본 연구에서는 하이퍼박스의 소속함수에 가중치 개념을 추가하여 특징값의 발생빈도에 따라 서로 다른 값으로 학습되게 함으로써 성능개선 및 특징분석 기능 등 여러 가지 이점을 얻을 수 있게 한다. 본 논문에서는 이를 위하여 새로운 소속함수를 제안하며, 학습과정에서 가중치 변화를 위한 방법론과, 수정된 형태의 하이퍼박스 생성(creation), 확장(expansion) 및 축소(contraction) 기법을 소개한다.

2. 퍼지 최대최소 신경망

FMM신경망은 하이퍼박스 퍼지집합으로 이루어지는데 각 하이퍼박스는 다차원 패턴공간상에서 특정 클래스에 포함되는 한 영역으로 정의되고, 개별 특징에 대하여 각각 최대 및 최소값 쌍으로 표현되는 퍼지 구간을 유지한다[1]. 각 학습단계에서 하이퍼박스 영역의 확장, 각 클래스간의 하이퍼박스 중첩테스트, 축소 등의 과정이 이전에 학습된 하이퍼박스의 파괴를 최소화하며 이루어 진다. FMM 신경망에서 임의의 하이퍼박스 j의 소속함수 $b_j(A_h)$는 다음과 같이 정의된다.

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

수식에서 $A_h = (a_{h1}, a_{h2}, \ldots, a_{hn})$은 입력패턴으로 총 n개의 특징값으로 이루어진다. 또한 각 특징에 대한 최소점과 최대점은 $U_j = (u_{j1}, u_{j2}, \cdots, u_{jn})$ 와 $V_j = (v_{j1}, v_{j2}, \cdots, v_{jn})$로 표현된다. γ는 특징범위의 가장자리에서 퍼지 소속함수의 기울기를 결정하는 매개변수로 0과 1사이의 값을 갖는다.

FMM에서 학습과정은 각 하이퍼박스에서 각 특징에 대한 최대 및 최소점을 결정하는 과정으로 하이퍼박스의 생성, 확장, 축소의 3가지 과정으로 이루어 진다. 본 연구에서 제안하는 모델은 식 (1)의 특성에 추가로 가중치 요소를 고려하게 되는데, 이와 더불어 이들 3가지 과정에 대해 수정되는 내용을 다음 절에서 소개한다.

3. 가중치를 고려한 FMM 모델

앞 절에서 언급한 바와 같이 본 연구에서 제안하는

모델은 가중치를 고려한 하이퍼박스 소속함수를 갖는다. 학습과정에서 각 특징에 대한 가중치 값이 조정되는데, 이때 기존의 FMM 신경망에서 이루어지는 생성, 확장 및 축소의 과정도 가중치를 사용하여 수정, 개선하였다. 식 (2)는 제안된 모델의 하이퍼박스 소속함수를 나타낸다. 식에서 W_{ji}는 i번째 특징과 j번째 하이퍼박스 간의 연결 가중치이고 나머지 표기는 식 (1)과 동일하다.

$$b_j(A_h) = \frac{1}{\sum_{i=1}^{n} w_{ji}} \bullet \sum_{i=1}^{n} w_{ji}[\max(0, 1-\max(0, \gamma \min(1, a_{hi} - v_{ji}))) + \max(0, 1-\max(0, \gamma \min(1, u_{ji} - a_{hi}))) - 1.0] \quad (2)$$

식에서 보인바와 같이 가중치를 추가한 것 외에도 함수의 형식을 변형하였다. 다시 말해서 가중치는 특정 하이퍼박스에 대하여 각 특징값의 상대적인 중요도를 서로 다른 값으로서 표현할 수 있게 한다. 또한 기존 FMM 모델에서는 식 (1)에 보인 바와 같이 각 하이퍼박스가 0.5와 1사이의 값을 출력하게 되는데, 이를 좀 더 일반적인 모델의 특성을 갖도록 하기 위하여 0와 1사이의 값을 갖도록 식의 형태를 변형하였다.

학습과정에서 가중치 변화는 생성되는 하이퍼박스의 형태에 따라 결정된다. 즉 각 학습단계에서 임의의 가중치 W_{ji}는 식(3)과 (4)에 의해서 조정된다.

$$w_{ji}^{new} = w_{ji}^{old} + \Delta w_{ji} \quad (3)$$

$$\Delta w_{ji} = \begin{cases} \lambda & if(v^{new} - u^{new} \le s) \\ d \bullet (T - \dfrac{v^{new} - u^{new}}{v^{old} - u^{old}}) & elseif(\dfrac{v^{new} - u^{new}}{v^{old} - u^{old}} \le T) \\ MAX(w_{ji}^{old} \bullet (\dfrac{v^{old} - u^{old}}{v^{new} - u^{new}} - 1.0), \dfrac{-w_{ji}^{old}}{2}) & otherwise \end{cases} \quad (4)$$

식 (4)는 가중치 변화분을 나타낸 식으로 세가지로 구분하여 조정된다. 우선 새롭게 구성되는 특징범위가 일정범위 이내, 즉 가장 이상적인 작은 범위 이내이면 일정 상수값을 증가시킨다. 반면 일정범위 보다 크지만 새롭게 생성된 범위가 이전보다 작거나, 일정비율 이내에서 증가한 경우라면 증가분에 반비례하여 결정한다. 그러므로 매개변수 T 는 1보다 큰 1근처의 값이다. 세 번째로 생성된 범위가 지나치게 커진 경우이다. 이는 식에 보인 바와 같이 커진 비율에 비례하여 가중치 값을 감소시킨다. 그러므로 식에서 s 는 하이퍼박스의 특징영역에 대한 이상적인 표준크기이고, λ와 d는 학습률 상수인데 특징영역이 확장함에 따라 해당 특징에 대한 가중치가 얼마나 급격히 증가시킬 것인가를 결정하는 매개변수이다.

각 가중치 값은 하이퍼박스가 생성되는 시점에 초기값 1.0으로 설정된다. 식 (2)에서 가중치 값이 항상 1.0이라면 이는 기존 FMM의 특성과 유사해 진다. 즉 기존의 모델은 모든 특징에 대하여 균일한 가중치를 고려하는 개념으로 볼 수 있다. 그림 1은 이러한 가중치 개념의 비교를 그림으로 보인 것이다.

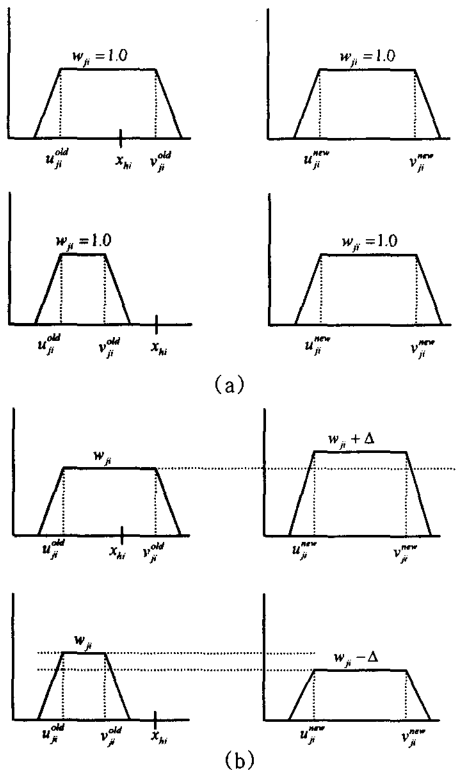

그림 1. 하이퍼박스 확장에 따른 가중치값 조정 비교: (a)FMM 신경망 모델, (b)제안된 모델.

본 논문에서는 제안한 모델에서 학습데이터에 포함될 수 있는 비정상적인 패턴이나 노이즈 패턴의 영향를 최소화하는 방법이 고려되었다. 즉 학습과정 중 비정상적인 데이터에 의하여 어떤 특징값이 범위가 큰 범위로 증가하는 경우 한꺼번에 확장하지 않고 점진적으로 범위를 증가시키도록 하는 방법을 채택하였다. 식 (5),(6) 및 (7)은 이러한 개념으로 수정된 확장방법을 보인다. 식에서 θ는 하이퍼박스의 확장 또는 생성여부를 결정하는 파라미터이다.

$$if \ n\theta \ge \sum_{i=1}^{n} (\max(v_{ji}, x_{hi}) - \min(u_{ji}, x_{hi})) \quad (5)$$

$$then \ if(x_{hi} < u^{old})$$

$$u_{ji}^{new} = u^{old} - \frac{1}{w_{ji}}(u^{old} - x_{hi}) \quad (6)$$

$$if(x_{hi} > v^{old})$$

$$v_{ji}^{new} = v^{old} + \frac{1}{w_{ji}}(x_{hi} - v^{old}) \quad (7)$$

식에서 보인 바와 같이 새롭게 정해지는 하이퍼박스의 특징범위는 새롭게 계산된 범위와 이전 범위의 비율을 고려해서 점진적으로 증가시키게 된다. 이는 기존의 여

러 번의 학습과정을 통해서 조정된 결과가 한 개의 잘못된 비정상적 데이터로 인하여 크게 왜곡되는 것을 방지할 수 있게 한다.

4. 실험결과 및 고찰

제안된 이론의 유용성을 평가하기 위하여 실제 패턴 분류 실험으로 Iris데이터와 Cleveland 의료 데이터[4]를 사용하였다. 또한 FMM모델과 제안한 모델을 비교하기 위해 평가 기준을 분류의 오류 개수와 식 (8)로 표현되는 오류율을 적용하였다.

$$E = \frac{1}{pm}\sum_{i=1}^{p}\sum_{k=1}^{m}\left|c_{ik} - d_{ik}\right| \qquad (8)$$

식(8)에서 p는 테스트 패턴의 수, m은 클래스 수, c_{ik}와 d_{ik}는 i번째 입력패턴에 대한 k번째 클래스의 각각 실제 출력값(actual output)과 기대 출력값(desired output)이다. 정확한 분류여부뿐만 아니라 오분류에 대해서도 얼마나 바람직한 값의 변화를 보였는가를 평가하였다.

표 1은 Iris 데이터에 대한 실험 결과이다. 잘 알려진 바와 같이 Iris 데이터는 총 3개 클래스에 150개의 패턴으로 이루어져 있다. 표에 보인 바와 같이 이 실험에서 오류갯수는 기존의 FMM 과 유사하였으나 오류율에서 보다 개선된 결과를 보였다. 실험에 사용된 각 파라미터의 값은, θ= 0.3, γ= 0.5, T=1.2, d=0.1 및 s=0.05 로 설정하였다.

표 1. Iris 데이터에 대한 실험결과

학습패턴 수	FMM 신경망		제안된 모델	
	오류패턴 수	에러율	오류패턴 수	에러율
30	4	0.62277	4	0.57643
60	3	0.63080	3	0.59236
90	1	0.63094	1	0.59241
120	0	0.63222	0	0.59498
149	0	0.63720	0	0.60504

두 번째 실험은 Cleveland 병원의 의료 자료[4]를 사용하여 학습데이터의 양에 따른 오분류 개수를 비교하였다. 이 데이터는 5개의 클래스에 총 297개의 패턴으로 이루어지고 각 패턴은 13개의 특징으로 구성되는 데이터이다.

표 2는 이를 이용한 분류 실험 결과이다. 표에서 보인 바와 같이 학습데이터의 수가 증가할수록 분류성능이 좋아지며, 기존의 FMM 신경망에 비하여 제안된 모델이 보다 개선된 성능을 보임을 알 수 있다.

5. 결론

제안한 모델의 핵심은 기존의 FMM 신경망에서 가중치 개념을 사용한 확장이다. 본 연구에서는 학습과정에서 각 특징에 대한 가중치를 평가하기 위해 새로운 소속 함수와 하이퍼박스 확장방법을 정의하였다. 이는 기존의 FMM 신경망이 갖는 일부 약점을 보완한다. 우선 학습패턴상의 특징값에 대한 발생빈도를 고려한 학습이 된다는 특성은 노이즈나 왜곡된 패턴에 의한 바람직하지 못한 영향을 최소화 한다. 실제 데이터를 사용한 실험에서 제안된 모델은 기존의 FMM 신경망에 비해 분류성능 개선 등의 의미 있는 결과를 보였다.

또한 특징과 하이퍼박스간의 상대적 기여도가 서로 다른 값으로 평가됨으로써 주어진 문제에서 특징과 클래스와의 관계 및 효과적인 특징선정기법 등에 활용될 수 있다. 이는 FMM신경망의 갖는 퍼지집합 특성을 장점을 그대로 수용할 뿐만 아니라 추가로 규칙형태의 지식생성이나 전문가 지식을 반영하는 학습기법으로의 확장을 위한 특성으로서도 의미가 있다. 향후 연구로는, 제안된 모델이 갖는 가중치 특성으로부터 각 특징에 대한 상대적 중요도를 평가하고 이를 특징 선정 및 규칙생성 과정에 활용하는 연구를 고려하고 있다.

표 2. Cleveland 의료 데이터에 대한 실험 결과

학습패턴수	오류 패턴 수	
	FMM 신경망	제안된 모델
10	35	32
20	29	27
30	21	19
40	14	12
50	7	5

* 이 연구는 과학기술부 뇌과학 연구개발사업으로 수행되었음.

6. 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] 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] C.L. Blake, and C. J. Merz, " UCI Repository of machine learning databases [http://www.ics.uci.edu/~mlearn/MLRepository.html]," University of California Irvine, Department of Information and Computer Science,1998.

[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.