

FMM 신경망에서 가중치 요소와 하이퍼박스 중첩효과 분석

박현정^o 김호준

한동대학교 전산전자공학부

hjpark79@empal.com^o, hjkim@handong.edu

Analysis of Weight Factor and Hyperbox Overlapping Effects in FMM Neural Networks

Hyun Jung Park^o, Ho Joon Kim

School of Computer Science and Electronic Engineering, Handong Global University

요약

본 연구에서는 FMM 신경망의 학습 알고리즘에서 하이퍼박스 확장과정에 수반되는 중첩현상을 분석하고, 이에 대한 축소 과정의 특성과 이를 보완하기 위한 새로운 활성화 함수에 관하여 고찰한다. 하이퍼박스 중첩 영역에 속하는 패턴 데이터는 그 분류 결과가 왜곡될 수 있다. 왜냐하면 학습과정에서 하이퍼박스의 특징범위는 특징값의 빈도요소를 고려하지 않음으로 인하여 극소수의 비정상적 데이터에 관해서도 동일 수준으로 민감하게 확장되기 때문이다. 본 논문에서는 특징집합에서 가중치와 빈도요소를 반영하는 모델로서 이러한 중첩현상의 영향을 개선하는 방법론을 소개한다. 제안된 이론은 단순화된 패턴집합에 대하여 그 유용성을 이론적으로 고찰하며, 실제 패턴분류 문제에 적용하여 실험적으로 평가한다.

1. 서론

1992년 Simpson등에 의해 제안된 퍼지 최대최소 신경망(FMMNN : Fuzzy Min-Max Neural Network)은 매우 간결하면서도 강력한 학습방법을 갖는 패턴 분류 모델이다[1]. 이후 FMM 신경망 분류 모델은 Gabrys 등에 의해 내부연산 및 데이터의 표현형태에 대한 일반화된 모델로 제안된바 있으며 이 외 다양한 신경망 기반 패턴 분류기법에 관한 연구가 활발하게 이루어 지고 있다[2-5].

FMM 신경망은 특징의 분포나 발생 빈도를 고려하지 않아 노이즈 패턴에 의해 성능이 왜곡되고 이에 의한 분류 성능의 저하가 단점으로 지적되었다[3]. 이러한 단점을 보완하기 위하여 특징의 발생 빈도와 범위를 고려한 가중치 개념을 학습 알고리즘과 활성화 특성에 적용한 수정된 퍼지 최대최소 신경망(WFMMNN : Weighted Fuzzy Min-Max Neural Network) 모델이 제안된 바 있다[3].

본 연구에서는 FMM 신경망의 학습 과정 중 하이퍼박스 확장 후 서로 다른 클래스에 속한 하이퍼박스 간의 중첩현상을 분석하였다. 기존의 FMM 신경망에서는 서로 다른 클래스에 속해있는 하이퍼박스의 중첩 영역이 발생하면 중첩 영역에 대한 패턴의 분류가 모호해 지므로 이를 해결하기 위하여 중첩 영역을 축소(Contraction)한다. 그러나 이러한 학습 알고리즘은 노이즈에 의해 하이퍼박스가 확장되고 중첩된 경우 이미 학습된 하이퍼박스 정보를 손실시킬 수 있다. 이에 본 연구에서는 WFMM 신경망에서 특징의 범위와 빈도를 고려한 활성화 특성을 적용함으로써 하이퍼박스 축소과정을 생략할 수 있게 하였다. 또한 가중치 요소와 중첩 효과를 분석하고 FMM 신경망의

본 논문의 2장에서는 FMM 신경망과 가중치 요소의 학습 알고리즘의 특성에 대하여 기술하였으며, 3장에서는 학습 효과를 분석하고, 4장에서는 아이리스 데이터 분류 문제를 통하여 제안된 이론의 타당성을 고찰한다.

2. FMM 신경망과 가중치 요소

FMM신경망은 퍼지 집합을 활용하는 교차 학습 신경망으로 하이퍼박스 기반의 패턴 분류 모형이다. 하이퍼박스는 소속함수에 상응하는 n 차원 최소, 최대값으로 정의된다[1]. 주어진 패턴 A_h 에 대하여 임의의 하이퍼박스 j 에 대한 소속함수 $b_j(A_h)$ 는 식 (1)과 같이 정의된다.

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

식에서 $A_h = (a_{h1}, a_{h2}, \dots, a_{hn})$ 은 입력패턴으로 총 n 개의 특징값으로 이루어진다. 또한 각 특징에 대한 최소점과 최대점은 $U_j = (u_{j1}, u_{j2}, \dots, u_{jn})$, $V_j = (v_{j1}, v_{j2}, \dots, v_{jn})$ 로 표현된다. γ 는 특징범위의 가장자리에서 퍼지 소속함수의 기울기를 결정하는 매개변수이다.

FMM 신경망은 확장 및 생성, 중첩 테스트, 축소의 3단계 학습과정으로 구성되어 있다. 첫번째 단계는 입력된 패턴을 가장 크게 반응하는 하이퍼박스에 확장 하거나 혹은 새로운 하이퍼박스로 생성한다. 확장 여부는 식(2)에 의해 결정한다.

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

하이퍼박스 축소과정과 비교하였다.

식에서 x_{hi} 는 입력 패턴이고 θ 는 하이퍼박스의 최대 크기를 결정하는 매개변수이다. 두번째 단계는 확장된 하이퍼박스를 다른 클래스에 속하는 하이퍼박스와 중첩되었는지 테스트하고 중첩이 된 경우 마지막 단계로 중첩 영역을 제거한다.

학습 과정에 특징의 빈도와 발생 범위를 고려한 WFMM 신경망의 활성화함수는 식(3)과 같다.

$$b_j(A_i) = \frac{1}{\sum_{i=1}^n w_{ji}} \cdot \sum_{i=1}^n w_{ji} [\max(0, 1 - \max(0, \gamma \min(1, a_{hi} - v_{ji}))) + \max(0, 1 - \max(0, \gamma \min(1, u_{ji} - a_{hi}))) - 1.0] \quad (3)$$

식 (3)에서 w_{ji} 는 j 번째 하이퍼박스와 i 번째 특징 간의 연결 가중치를 의미하며, 나머지 표기는 식 (1)과 동일하다. 식에서 보인 바와 같이 WFMM 신경망에서는 하이퍼박스와 특징과의 관계가 w_{ji} 로 각 특징에 가중치로 적용되었다. 학습과정에서 빈도변수값이 최종적으로 유지되면 이로부터 w_{ji} 값이 식 (4)에 의해 산출된다.

$$w_{ji} = \frac{\alpha f_{ji}}{R} \quad (4)$$

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

가중치는 식에서 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.

WFMM 신경망의 학습과정은 하이퍼박스 확장 및 생성의 단일 과정이며 FMM 신경망의 하이퍼박스의 최소값과 최대값 정보 이외에 발생빈도 정보를 함께 학습한다.

3. 학습 효과 분석

본 절에서는 중첩 현상에 대한 해결 방안으로, FMM 신경망의 하이퍼박스 축소과정과 WFMM 신경망의 가중치를 이용한 패턴 분류의 영향을 분석하였다. 모델의 유용성을 이론적으로 보이기 위하여 다음과 같이 단순화된 패턴집합을 고려한다. 다시 말해서 총 2종류의 패턴클래스, $C = \{C_1, C_2\}$ 와 2종류의 특징집합, $X = (x_1, x_2)$ 으로 이루어지는 9개의 학습패턴, $P = \{p_1, p_2, \dots, p_9\}$, $p_i = (c_i, (x_{1i}, x_{2i}))$ 이 다음과 같이 주어졌다고 가정한다.

$$p_1 = (1, (0.4, 0.4)) \quad p_2 = (2, (0.7, 0.7)) \quad p_3 = (1, (0.5, 0.5))$$

$$p_4 = (2, (0.6, 0.6)) \quad p_5 = (1, (0.5, 0.45)) \quad p_6 = (2, (0.7, 0.6))$$

$$p_7 = (1, (0.45, 0.45)) \quad p_8 = (2, (0.65, 0.7)) \quad p_9 = (2, (0.45, 0.45))$$

그림 1.의 (a), (b)는 학습과정에서 p_1 에서 p_9 에 의한 하이퍼박스의 확장형태를 보인 것으로 p_9 가 주어지면 그림의 (a)에서 보는 바와 같이 서로 다른 클래스의 하이퍼박스 중첩 영역이 발생하게 된다. 이를 해결하기 위해 FMM 신경망에서는 (b)와 같이 하이퍼박스를 축소하게 되고 제안된 WFMM 신경망은 (a)와 같이 축소되지 않은 상태에서 가중치를 학습함으로써 해결한다.

학습이 완료된 상태에서 FMM 신경망의 하이퍼박스 특

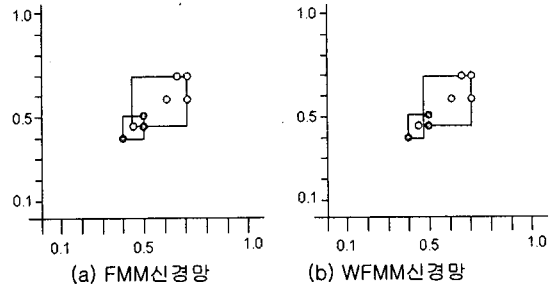


그림 1. FMM과 WFMM 신경망 학습에 대한 중첩효과 분석

, 0.7}으로 나타난다. 이 정보를 이용하여 $\gamma = 0.1$ 로 하고 패턴 p_3 와 패턴 p_5 의 소속값을 구해보면 $H_1(p_3) = 0.99875$, $H_2(p_3) = 1.0$, $H_1(p_5) = 0.99875$, $H_2(p_5) = 1.0$ 의 값으로 모두 하이퍼박스2(H_2)에 가장 큰 소속값을 갖고 오 분류된다. 그러나 WFMM 신경망에서는 $H_1 = \{(0.4, 0.5), (0.4, 0.5), 3\}$, $H_2 = \{(0.45, 0.7), (0.45, 0.7), 3\}$ 의 값을 가지며 이의 정보를 가중치 값으로 활용하면 $H_1(p_3) = 5/7$, $H_2(p_3) = 2/7$, $H_1(p_5) = 5/7$, $H_2(p_5) = 2/7$ 의 소속값을 갖고 바르게 분류됨을 알 수 있다.

또한 제안된 모델은 만일 다른 클래스에 속하는 두개의 패턴이 거의 동일한 값을 갖는 경우, 각 패턴은 특징의 발생 범위와 빈도에 의한 가중치로 판단되기 때문에 노이즈 패턴에 의한 영향을 최소화 할 수 있다.

4. 실험 결과 및 고찰

서로 다른 클래스에 속한 하이퍼박스 중첩현상을 해결하는 방안으로 하이퍼박스 축소를 사용하는 기존의 FMM 신경망 모델과, 학습된 가중치를 이용한 제안된 모델의 분류 성능을 비교하였다. 실험결과와 분석은 오류 패턴의 개수와 오류누적합을 비교, 분석하였으며 오류누적합은 식(5)으로 계산된다.

$$E = \frac{1}{pm} \sum_{i=1}^p \sum_{k=1}^m |c_{ik} - d_{ik}| \quad (5)$$

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.

실험은 패턴분류 문제로 잘 알려진 Iris 데이터를 적용, 분류실험을 수행하였다. 아이리스 데이터는 총 3개의 클래스에 각 50개 씩 총 150개의 데이터 패턴으로 구성되어 있다. 실험에서는 중첩현상 이 분류에 미치는 영향을 분석하기 위해 Iris의 versicolor와 virginica의 패턴만을 이용하였다. Iris의 setosa 클래스는 그림 2에 보인 바와 같이 다른 클래스와의 구별이 명확하기 때문에 제외하였다.

실험에 사용된 파라미터 값들은 여러번의 튜닝을 통해, $\theta = 0.25$, $\gamma = 0.1$, $s = 0.0001$ 의 값으로 설정하였다.

성은 $H_1 = \{(0.4, 0.475), (0.4, 0.5)\}$, $H_2 = \{(0.475, 0.7), (0.45$

표 1은 각 특징의 조합별 오류패턴과 오류누적합을 나타낸 것이다. (F_1, F_2)의 특징 조합을 사용하는 경우는 그림 3에서 보는 바와 같이 중첩된 영역에 의한 손실이 커서 가중치를 고려한 신경망이 개선된 분류 성능을 보였다. 그러나 (F_3, F_4)의 특징 조합을 사용한 경우 그림 4에 보인바와 같이 중첩 영역에 의한 정보손실이 작아 유사한 수준의 성능을 보였다.

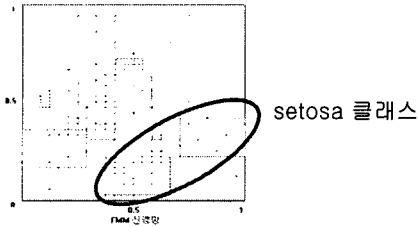


그림 2. FMM 신경망의 2차원 특징 공간상의 하이퍼박스 분포

표 1. Iris 데이터에 대한 실험결과

Feature 조합	FMM 신경망		제안된 모델	
	오류패턴	오류누적합	오류패턴	오류누적합
(F1, F2)	32(68)	0.01494	23(77)	0.01051
(F1, F3)	5(95)	0.00539	2(98)	0.00188
(F1, F4)	6(94)	0.00412	6(94)	0.00411
(F2, F3)	9(91)	0.00426	6(94)	0.00294
(F2, F4)	5(95)	0.00283	4(96)	0.00294
(F3, F4)	2(98)	0.00233	2(98)	0.00224

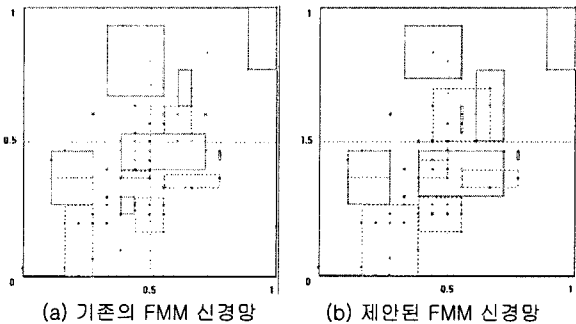


그림 3. Iris 데이터의 1번째 및 2번째 특징 학습 결과1

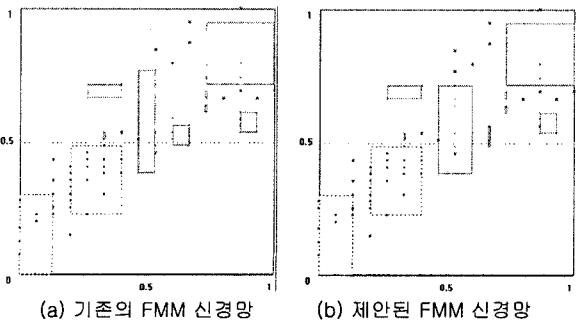


그림 4. Iris 데이터의 3번째 및 4번째 특징 학습 결과2

5. 결론

본 연구에서는 FMM 신경망의 학습 과정에서 발생하는 하이퍼박스 중첩현상을 해결하기 위한 두가지 방법을 비교 분석하였다. 서로 다른 클래스에 속한 하이퍼박스 간의 중첩 현상은 패턴의 분류를 모호하게 하기 때문에 이를 해결하기 위한 방법론이 필요하다. 기존의 FMM 신경망에서는 하이퍼박스가 중첩되지 않도록 각 하이퍼박스를 축소하여 이를 해결하였지만 특징의 분포를 고려하지 않았기 때문에 노이즈 패턴에 의한 영향을 크게 받고 하이퍼박스 정보를 심하게 왜곡시킬 수 있었다. 또한 하이퍼박스 축소 과정으로 인하여 특징 공간에서 학습된 하이퍼박스 정보의 일부가 사라져 패턴 간의 하이퍼박스 중첩이 많을수록 패턴 분류 성능이 심하게 왜곡될 가능성이 있다. WFMM 신경망은 특징의 발생 빈도와 그 발생 범위를 고려하여 학습된 하이퍼박스 정보를 이러한 정보를 가중치로 활용할 수 있다. 각 하이퍼박스의 특징별 고유 가중치를 가지고 중첩영역의 축소 과정을 생략하여 축소로 인한 오분류율을 줄이고 중첩 영역에 속한 패턴을 비교적 바르게 분류할 수 있는 기준을 제시하였다.

* 이 연구는 산업자원부 뇌과학 연구개발사업으로 수행되었음

6. 참고 문헌

- [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, and Sudhir Gupta, " A Weighted Fuzzy Min-Max Neural Network for Pattern Classification and Feature Extraction," Proceeding of International Conference on Computational Science and Its Application, Part.4, pp.791-198, 2004
- [4] 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.
- [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.