

클래스간의 거리를 고려한 학습법칙을 사용한 퍼지 신경회로망 모델

Fuzzy Neural Network Model Using A Learning Rule Considering the Distances Between Classes

김용수*, 백용선**, 이세열***

Yong-Soo Kim, Yong-Sun Back, and Se-Yul Lee

* 대전대학교 컴퓨터공학과

** 대덕대학 컴퓨터웹정보과

*** 청운대학교 컴퓨터학과

요 약

본 논문은 입력 벡터와 클래스들의 대표값들간의 유클리디안 거리들을 사용한 새로운 퍼지 학습법칙을 제안한다. 이 새로운 퍼지 학습을 supervised IAFC(Integrated Adaptive Fuzzy Clustering) 신경회로망 4에 적용하였다. 이 신경회로망은 안정성을 유지하면서도 유연성을 가지고 있다. iris 데이터를 사용하여 테스트한 결과 supervised IAFC 신경회로망 4는 오류 역전과 신경회로망과 LVQ 알고리즘보다 성능이 우수하였다.

Abstract

This paper presents a new fuzzy learning rule which considers the Euclidean distances between the input vector and the prototypes of classes. The new fuzzy learning rule is integrated into the supervised IAFC neural network 4. This neural network is stable and plastic. We used iris data to compare the performance of the supervised IAFC neural network 4 with the performances of back propagation neural network and LVQ algorithm.

Key words : Learning rule, Fuzzy vector quantization, Supervised IAFC neural network 4, Euclidean distance, Decision boundary.

1. 서 론

신경회로망에 있어서 연결강도를 조정하는 학습법칙은 신경회로망의 성능을 좌우하는 중요한 역할을 한다. LVQ(Learning Vector Quantization)은 많이 사용되어온 감독학습 법칙의 하나인데, LVQ는 분류가 맞으면 승자의 연결강도를 입력 벡터쪽으로 조정하고, 분류가 틀리면 승자의 연결강도를 입력 벡터로부터 멀리 보낸다 [1, 2]. LVQ는 승자의 연결강도를 조정할 때 학습률을 사용하는데, 이 학습률은 승자의 연결강도의 조정량을 조절하면서 연결강도를 수렴하도록 해준다. 그러나, 이 학습법칙은 연결강도가 적절한 값으로 수렴하기 위해서는 연결강도의 초기값을 적절한 값으로 초기화하는 것이 필요하다. 이러한 문제점을 개선하고자 하는 연구가 계속되어 왔다 [3-8]. Bezdek등은 Fuzzy c-Means 알고리즘을 Kohonen의 Self-Organizing Feature Map의 학습법칙에 적용한 FLVQ(Fuzzy Learning Vector Quantization)를 제안하였다 [3]. Chung과 Lee는 퍼지 소속도를 LVQ와 조합한 FLVQ를 제안하였다 [4]. Chung과 Lee의 FLVQ는 퍼지 목적함수를 최적화 하는 식에 의하여 유도되었다. Chung과 Lee의 FLVQ는 원하는 소속도를 알아야 하는데, 현실에서는 원하는 소속도를 아는 것이 쉽지 않은

문제점이 있다. Karayiannis는 입력 벡터들과 클래스들의 대표값들 사이의 거리의 average generalized mean을 최소화 하는 FLVQ를 제안하였다 [5]. 이 FLVQ는 batch learning을 사용하므로 메모리를 많이 사용하는 단점이 있다. Kusumoputro등은 학습률과 퍼지 소속도를 사용한 FLVQ를 제안하였다 [6]. Kim and Bien은 반복횟수의 함수와 Π 함수와 퍼지 소속도를 사용한 FLVQ를 제안하였다 [7]. Kim은 LVQ는 분류가 맞거나 틀리거나에 관계없이 똑같은 학습률을 사용하는데 반하여 분류가 맞느냐 틀리느냐에 따라 다른 학습률을 사용하는 FLVQ를 제안하였다 [8].

본 논문에서는 퍼지소속도와 함께 입력 벡터와 클래스들의 대표값들간의 유클리디안 거리들을 사용한 새로운 FLVQ를 제안한다. 이 새로운 FLVQ는 선정된 클래스의 대표값과 입력 벡터사이의 유클리디안 거리와 다른 클래스들의 대표값들과 입력 벡터사이의 유클리디안 거리들의 상대적인 크기를 학습법칙에 반영한 것이다. 분류가 맞을 때에는 적절한 결정경계선을 얻기 위해서 적절한 결정경계선에 대한 정보를 더 많이 가지고 있는 입력 벡터의 반영도를 높이기 위해 $(1 - u_j)^2$ 을 사용한다. u_j 는 입력 벡터가 선정된 클래스에 소속되어 있는 정도를 나타내는 퍼지 소속도이다. 결정경계선 근처에 있는 입력 벡터는 퍼지소속도 u_j 가 클래스들 간의 경쟁이 치열하기 때문에 선정된 클래스에 치우쳐 있는 입력 벡터의 퍼지소속도 u_j 보다 작다. 이 성질을 이용하여

접수일자 : 2006년 4월 13일

완료일자 : 2006년 6월 5일

$(1 - u_j)^2$ 을 사용하면 결정경계선 근처에 있는 입력 벡터의 반영도를 특정클래스에 치우쳐 있는 입력 벡터의 반영도보다 크게 할 수 있다. 분류가 틀릴 때에는 입력 벡터의 틀린 클래스에 소속되어 있는 정도를 나타내는 퍼지 소속도 u_j 를 사용한다. 입력 벡터의 틀린 클래스에 소속되어 있는 정도를 나타내는 퍼지 소속도가 크다는 것은 그 만큼 틀린 클래스의 대표값을 입력 벡터로부터 멀리 보내야 한다는 것을 의미하므로, 틀린 클래스의 대표값을 조정할 때 그 조정량을 입력 벡터의 틀린 클래스에의 퍼지 소속도에 비례하도록 하였다.

이 새로운 FLVQ를 supervised IAFC 신경회로망 4에 적용하였다. Supervised IAFC 신경회로망 4는 ART-1 신경회로망 [9]와 유사한 제어 구조를 가지고 있기 때문에, ART-1 신경회로망과 같이 안정성과 유연성을 가지고 있다. 제안한 신경회로망은 과거에 배운 내용을 보전하는 안정성이 있으면서도 새로운 입력 벡터가 발생했을 때 이를 반영하는 유연성이 있다. ART-1 신경회로망은 잡음에 민감한데 비하여 제안한 신경회로망은 ART-1 신경회로망에 비하여 잡음에 덜 민감하다. 또한 LVQ와 같이 일부 출력 뉴런이 선택되지 않아 발생하는 underutilization 문제를 제안한 신경회로망은 해결하였다. 제안한 신경회로망은 클래스의 크기와 개수를 vigilance parameter로 조절한다. 이 vigilance parameter는 거리 임계값 또는 클래스의 반지름과 관계가 있다 [10].

제안한 신경회로망의 성능과 LVQ 알고리즘과 오류역전파 신경회로망의 성능을 비교하기 위하여 iris 데이터를 사용하였다. 이 iris 데이터는 클러스터링 알고리즘의 성능을 비교하는데 많이 사용되어 왔다.

2. Supervised IAFC 신경회로망 4

입력 벡터가 supervised IAFC 신경회로망 4에 가해진 후, 출력 뉴런들 간의 경쟁은 winner-take-all 식으로 일어난다. 입력 벡터와의 유클리디안 거리가 제일 작은 연결강도를 가진 출력뉴런이 경쟁에서 승리한다. 즉 I번째 출력뉴런,

$$I = \min_i \| \mathbf{x} - \mathbf{v}_i(t) \|,$$

이 경쟁에서 승리한다. 여기서 \mathbf{x} 는 입력 벡터이고, $\mathbf{v}_i(t)$ 는 i번째 출력뉴런의 연결강도이다. 출력뉴런들 중에서 승자를 결정한 후에 supervised IAFC 신경회로망 4는 다음의 vigilance 기준에 의해 vigilance test를 한다 :

$$e^{-u_i} \| \mathbf{x} - \mathbf{v}_i(t) \| \leq T,$$

여기서 T는 vigilance parameter이다. 퍼지 소속도 u_i 는 다음과 같이 정의 된다:

$$u_i = \frac{\left[\frac{1}{\| \mathbf{x} - \mathbf{v}_i(t) \|^2} \right]^{\frac{1}{m-1}}}{\sum_{j=1}^n \left[\frac{1}{\| \mathbf{x} - \mathbf{v}_j(t) \|^2} \right]^{\frac{1}{m-1}}},$$

여기서 n은 존재하는 committed 출력뉴런의 개수이고, $m \in [1, \infty]$ 은 weight exponent이며 실험적으로 2로 놓는다.

만약 승자가 vigilance 테스트를 만족하면은 supervised IAFC 신경회로망 4는 승자의 연결강도를 다음의 새로운 퍼지 학습법칙에 의하여 조정한다 :

$$\mathbf{v}_I(t+1) = \mathbf{v}_I(t) + f(t) \cdot (1 - u_I)^2 \cdot g(d) \cdot (\mathbf{x} - \mathbf{v}_I(t))$$

if x is classified correctly,

$$\mathbf{v}_I(t+1) = \mathbf{v}_I(t) - f(t) \cdot u_I \cdot h(d) \cdot (\mathbf{x} - \mathbf{v}_I(t))$$

if x is classified incorrectly,

$$v_i(t+1) = v_i(t) \quad \text{for } i \neq I.$$

위의 식에서 f(t)는 반복 횟수의 함수로 $\frac{1}{1+k(t-1)}$ 이다. g(d)는 분류가 맞을 때 승자의 연결강도를 조정하는 양을 조절하며 다음과 같이 정의된다.

$$g(d) = \frac{d_i}{\sum_{i=1}^n d_i - d_i},$$

여기서 d_i 는 입력 벡터와 i번째 클래스의 대표값간의 유클리디안 거리이고, d_I 는 입력 벡터와 승자인 I번째 클래스의 대표값간의 유클리디안 거리이다. n은 존재하는 committed 출력뉴런의 개수이다.

그림1은 클래스간의 거리를 고려한 학습법칙을 설명하기 위한 것이다. 입력벡터 A는 클래스1에 소속되어 있는 정도는 크지만 적절한 결정경계선에 대한 정보는 많지 않다. 이에 비해 입력벡터 B는 클래스1에 소속되어 있는 정도는 작지만 적절한 결정경계선에 대한 정보는 많다. 따라서, 분류가 맞았을 때, 새로운 학습법칙에서 클래스간의 거리를 고려하는 것이 g(d)이다. 입력벡터 A와 클래스1의 대표값 p_1 간의 유클리디안 거리는 $g_{A1}(d)$ 이고, 입력벡터 A와 클래스2의 대표값 p_2 간의 유클리디안 거리는 $g_{A2}(d)$ 이다. 또한 입력벡터 B와 클래스1의 대표값 p_1 간의 유클리디안 거리는 $g_{B1}(d)$ 이고, 입력벡터 B와 클래스2의 대표값 p_2 간의 유클리디안 거리는 $g_{B2}(d)$ 이다. $g_{A1}(d)$ 와 $g_{B1}(d)$ 가 비슷할 때

$$g_{A2}(d) \text{가 } g_{B2}(d) \text{보다 크기 때문에, } g_A(d) = \frac{g_{A1}(d)}{g_{A2}(d)}$$

는 $g_B(d) = \frac{g_{B1}(d)}{g_{B2}(d)}$ 보다 작다. 따라서 입력벡터 A에 대한 클래스1의 대표값 p_1 을 조정하는 양을 조절하는 가중치는 입력벡터 B에 대한 클래스 1의 대표값 p_1 을 조정하는 양을 조절하는 가중치보다 작다. 입력벡터 B와 입력벡터 C는 둘 다 결정경계선 근처에 존재하지만 어디에 위치하느냐에 따라서 클래스1의 대표값 p_1 을 조정하는 양을 조절하는 가중치가 다르다. 입력벡터 C와 클래스1의 대표값 p_1 간의 유클리디안 거리는 $g_{C1}(d)$ 이고, 입력벡터 C와 클래스2의 대표값 p_2 간의 유클리디안 거리는 $g_{C2}(d)$ 이다. $g_{C1}(d)$ 는 $g_{B1}(d)$ 보다 작고, $g_{C2}(d)$ 는 $g_{B2}(d)$ 보다 크다. 따라서, $g_C(d) = \frac{g_{C1}(d)}{g_{C2}(d)}$ 는 $g_B(d) = \frac{g_{B1}(d)}{g_{B2}(d)}$ 보다 작다. 이는 같은 결정경계선 쪽에 입력벡터가 존재하더라도 입력벡터 B가 입력벡터 C보다 적절한 결정경계선에 대한 정보를 더 가지고 있기 때문에 입력벡터 B에 대한 클래스1의 대표값 p_1 을 조정하는 양을 조절하는 가중치가 입력 벡터 C에 대한 클래스 1의 대표값 p_1 을 조정하는 양을 조절하는 가중치보다 크다. $g(d)$ 는 새로운 학습법칙에서 분류가 맞았을 때 입력벡터와 클래스들의 대표값들간의 유클리디안 거리들을 고려하여 선정된 클래스의 대표값을 조정하는 양을 조절하는 가중치를 결정한다.

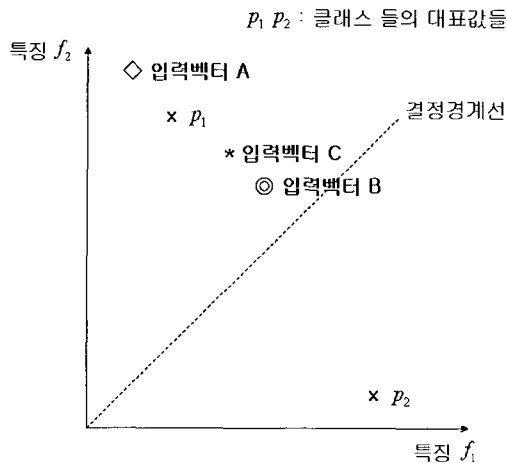


그림 1. 클래스간의 거리를 고려한 학습법칙
Fig 1. Learning rule considering the distances between classes

$h(d)$ 는 분류가 틀릴 때 승자의 연결강도를 조정하는 양을 조절하며 다음과 같이 정의된다 :

$$h(d) = \frac{d_i}{\sum_{i=1}^n d_i}$$

3. 실험 및 결과

제안한 신경회로망의 성능과 LVQ 알고리즘과 오류 역전

파 신경회로망의 성능을 비교하기 위해서 클러스터링 알고리즘의 성능을 비교하는데 많이 쓰이는 iris 데이터를 사용하였다. iris 데이터는 150개의 4차원 데이터로 구성되어 있다. 이 데이터는 3개의 subspecies로부터 추출한 데이터이고, 각 subspecies는 50개의 데이터를 가지고 있다. 150개의 데이터 중 임의로 75개의 데이터를 선정해 훈련 데이터로 사용했는데, 각 subspecies로부터 25개의 데이터를 사용하였다. 나머지 75개의 데이터는 테스트 데이터로 사용하였다.

반복 학습을 하는 중에 현재의 연결강도와 이전의 연결강도의 차이인 $\|v(t) - v(t-1)\|$ 이 0.01보다 작으면 연결강도들이 실험적으로 수렴한 것으로 보고 supervised IAFC 신경회로망 4의 훈련을 끝내도록 하였다. 훈련을 끝낸 후 supervised IAFC 신경회로망 4를 테스트 하였다. 그림 2는 supervised IAFC 신경회로망 4의 성능과 LVQ 알고리즘과 오류 역전파 신경회로망의 성능을 비교한 것이다.

		Output		
		1	2	3
Desired	1	18	7	
	2		25	
	3			25

(a) LVQ 알고리즘

		Output		
		1	2	3
Desired	1	20	5	
	2		25	
	3			25

(b) 오류 역전파 신경회로망

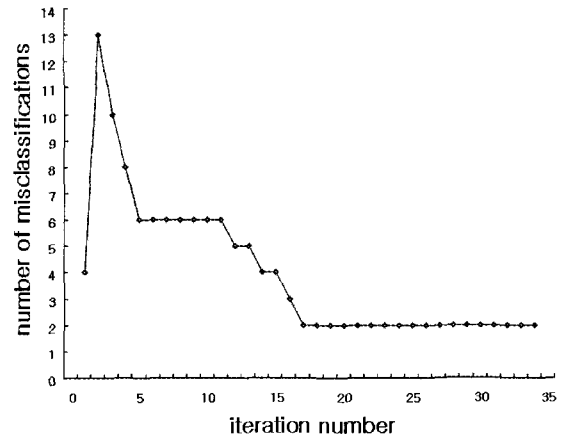
		Output		
		1	2	3
Desired	1	24	1	
	2	2	23	
	3			25

(c) Supervised IAFC 4

그림 2. iris 데이터를 사용한 결과 비교
Fig 2. Comparison of results using iris data

Supervised IAFC 신경회로망 4를 훈련시키기 위해 30번의 반복학습을 시켰으며, T가 1.3이고 k가 0.5일 때 테스트 데이터에서 3개의 오류가 발생하였다. 오류 역전과 신경회로망은 5개의 오류를 발생시켰으며, LVQ 알고리즘은 7개의 오류를 발생시켰다.

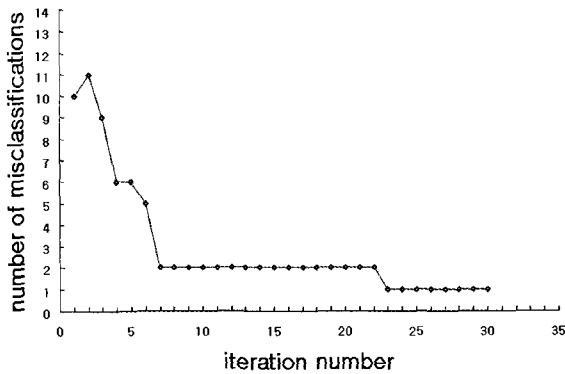
그림 3은 k가 0.5일때 T에 따른 supervised IAFC 신경회로망 4의 성능비교를 보여준다. T가 1.3일 때는 학습을 위하여 30회 반복학습을 하였으며 1개의 오류가 발생하였고, 테스트 데이터에 대해서는 3개의 오류가 발생하였다. T가 1.7일 때는 학습을 위하여 23회 반복학습을 하였으며 1개의 오류가 발생하였고, 테스트 데이터에 대해서는 4개의 오류가 발생하였다. T가 2.2일때는 학습을 위하여 34회 반복학습을 하였으며 2개의 오류가 발생하였고, 테스트 데이터에 대해서는 4개의 오류가 발생하였다. 위의 결과는 T에 따라서 supervised IAFC 신경회로망 4의 성능이 변화함을 보여준다.



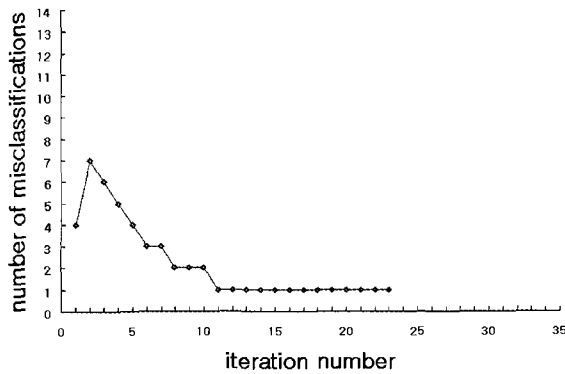
(c) T가 2.2일 때

그림 3. k가 0.5일 때 T에 따른 supervised IAFC 신경회로망 4의 성능비교.

Fig 3. Comparison of performances of the supervised IAFC neural network 4 depending on T when k is 0.5

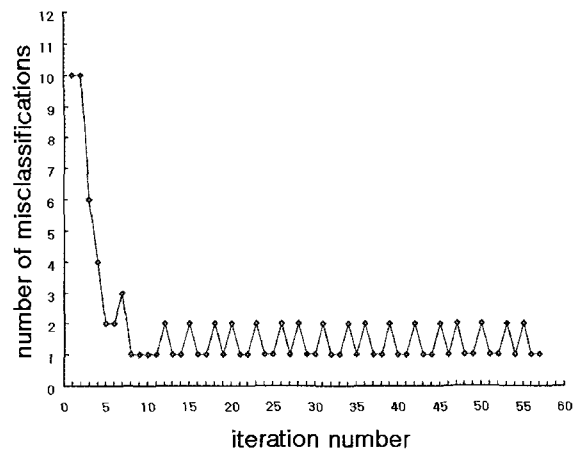


(a) T가 1.3일 때

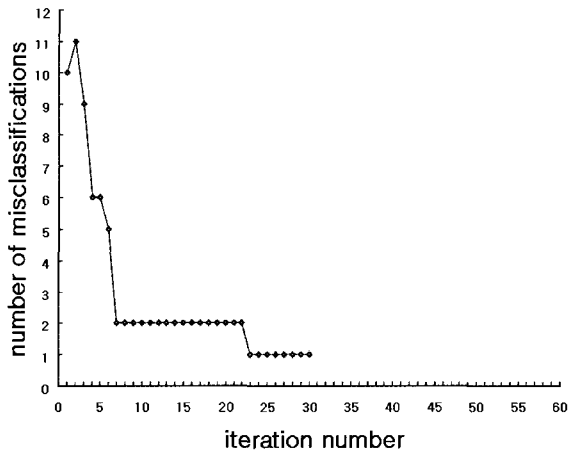


(b) T가 1.7일 때

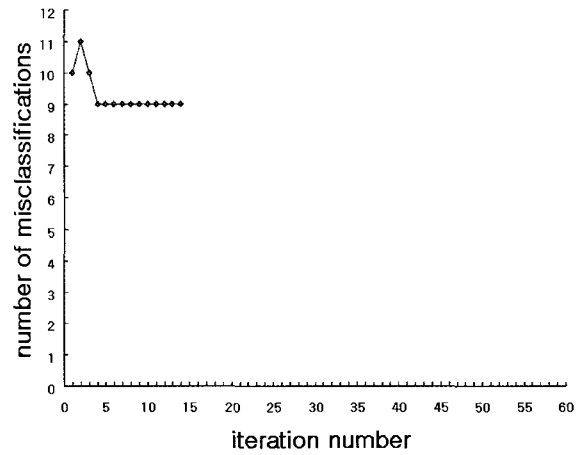
그림 4는 T가 1.3일때 k에 따른 supervised IAFC 신경회로망 4의 성능비교를 보여준다. k가 0.05일때 학습을 위하여 57회 반복학습을 하였으며 1개의 오류가 발생하였고, 테스트 데이터에 대해서는 3개의 오류가 발생하였다. k가 0.5일 때는 학습을 위하여 30회 반복학습을 하였으며 1개의 오류가 발생하였고, 테스트 데이터에 대해서는 3개의 오류가 발생하였다. k가 1.0일 때는 학습을 위하여 27회 반복학습을 하였으며 2개의 오류가 발생하였고, 테스트 데이터에 대해서는 4개의 오류가 발생하였다. k가 1.5일 때는 학습을 위하여 27회 반복학습을 하였으며 4개의 오류가 발생하였고, 테스트 데이터에 대해서는 5개의 오류가 발생하였다. k가 5.0일 때는 학습을 위하여 14회 반복학습을 하였고 9개의 오류가 발생하였고, 테스트 데이터에 대해서는 9개의 오류가 발생하였다. 위의 결과는 k에 따라서 supervised IAFC 신경회로망 4의 성능이 변화함을 보여준다. 대체적으로 k가 커지면 실험적으로 수렴하기 위해 학습에 필요한 반복학습의 횟수가 줄어드나 오류의 개수가 증가 하였다.



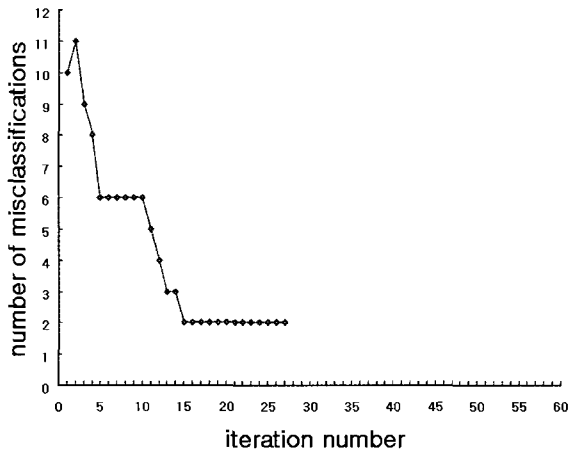
(a) k가 0.05일 때



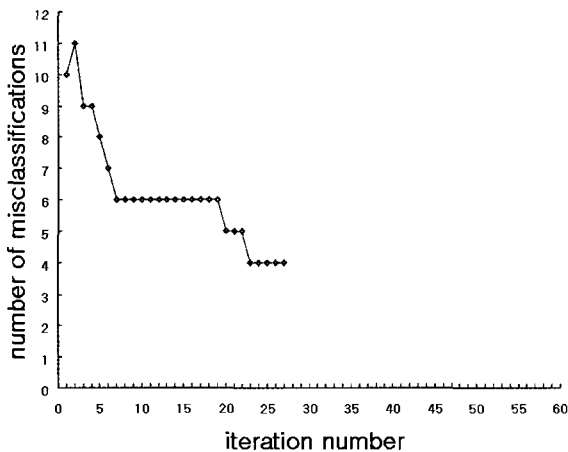
(b) k가 0.5일 때



(e) k가 5.0일 때



(c) k가 1.0일 때



(d) k가 1.5일 때

그림 4. T가 1.3일 때 k에 따른 supervised IAFC 신경회로망 4의 성능비교.

Fig 4. comparison of performances of the supervised IAFC neural network 4 depending on k when T is 1.3

4. 결 론

본 논문에서 클래스간의 거리를 고려한 학습법칙을 제안하였고, 이 새로운 퍼지 학습법칙을 supervised IAFC 신경회로망 4에 적용하였다. 이 supervised IAFC 신경회로망 4는 실험적으로 수렴하기 위하여 필요한 반복횟수가 적었고, supervised IAFC 신경회로망 4는 iris 데이터를 사용하여 테스트한 결과 오류 역전과 신경회로망과 LVQ 알고리즘에 비하여 적은 오류를 발생시켰다

참 고 문 헌

- [1] T. Kohonen, "Self-Organizing Map," Proceedings of the IEEE, vol. 78, No. 9, pp. 1464-1480, 1990.
- [2] C.-T. Lin and C. S .G. Lee, Neural Fuzzy Systems : A Neuro-Fuzzy Synergism to Intelligent Systems, Prentice-Hall, 1996.
- [3] J. C. Bezdek, E. C. Tsao, and N. R. Pal, "Fuzzy Clustering Networks," Proceedings of the First IEEE Conference on Fuzzy Systems, San Diego, pp. 1035-1043, 1992.
- [4] F.-L. Chung and T. Lee, "Fuzzy Learning Vector Quantization," Proceedings of 1993 International Joint Conference on Neural Networks, Nagoya, Vol. 3, pp. 2739-2743, 1993.
- [5] N. B. Karayiannis and J. C. Bezdek, "An Integrated Approach to Fuzzy Learning Vector Quantization and Fuzzy c-Means Clustering," IEEE Transactions on Fuzzy Systems, Vol. 2, pp. 622-628, 1997.
- [6] B. Kusumoputro, M. R. Widyanto, M. I. Fanany, and H. Budiarto, "Improvement of Artificial Odor

Discrimination System Using Fuzzy-LVQ Neural Network," Proceedings of the third International Conference on Computational Intelligence and Multimedia Applications, pp. 474-478, 1999.

- [7] Y. Kim and Z. Bien, "Integrated Adaptive Fuzzy Clustering(IAFC) Neural Networks Using Fuzzy Learning Rules," Iranian Journal of Fuzzy System, Vol. 2, No. 2, pp. 1-13, 2005.
- [8] 김용수, "비대칭 퍼지 학습률을 이용한 퍼지 신경회로망 모델," 퍼지 및 지능 시스템 학회 논문지, 제 15권 제 7호, pp. 800-804, 2005.
- [9] G. A. Carpenter and S. Grossberg, "A Massively Parallel Architecture for Self-Organizing Neural Pattern Recognition Machine," Compute Vision, Graphics, and Image Processing, Vol. 37, pp. 54-115, 1987.
- [10] B. Moore, "Art-1 and Pattern Clustering," Proceedings of the 1988 Connectionist Models Summer School, pp. 174-185, 1989.



백용선(Yong-Sun Baek)

1998년 : 우송대학교 컴퓨터학과 공학사
 1999년 : 대전대학교 대학원 컴퓨터공학과 공학석사
 2003년~현재 : 대덕대학 컴퓨터웹정보과 초빙교수
 2005년 : 대전대학교 대학원 컴퓨터공학과 박사수료

관심분야 : 신경회로망, 퍼지 논리 등
 Phone : +82-42-866-0394
 Fax : +82-42-866-0399
 E-mail : dodo029@ddc.ac.kr



이세열(Se-Yul Lee)

1996년 : 대전대학교 전자물리학과 이학사
 1999년 : 동 대학원 정보통신공학과 공학석사
 2000년 : (주)인소팩 부설기술연구소 연구원
 2003년 : 동 대학원 컴퓨터공학과 공학박사
 2004년~현재 : 청운대학교 컴퓨터학과 전임강사

관심분야 : 침입탐지, 정보보호, 네트워크 보안, 퍼지 논리 등
 Phone : +82-41-630-3225
 Fax : +82-41-634-8700
 E-mail : pirate@cwu.ac.kr

저 자 소 개



김용수(Yong-Soo Kim)

1981년 : 연세대학교 전기공학과 공학사
 1983년 : KAIST 전기 및 전자공학과 공학석사
 1986년 : 삼성전자 종합연구소 주임연구원
 1993년 : Texas Tech Univ. 공학박사
 1995년~현재 : 대전대학교 컴퓨터공학부 부교수

관심분야 : 신경회로망, 퍼지 논리, 패턴인식, 영상처리, 침입 탐지 등
 Phone : +82-42-280-2547
 Fax : +82-42-284-0109
 E-mail : kvstj@diu.ac.kr