

# 조건 확률을 퍼지화한 학습 법칙을 사용하는 퍼지 신경회로망 모델

## The Fuzzy Neural Network Utilizing A Fuzzy Learning Rule

김 용수, 함 창현, 백 용선

Yong Soo Kim, Chang Hyun Ham, Yong Sun Baek

대전대학교 컴퓨터공학과

Dept. of Computer Engineering, Taejon University

(Tel) 042-280-2547

(Fax) 042-284-0109

### ABSTRACT

학습법칙은 신경회로망의 성능을 좌우하는 중요한 요소의 하나이다. Kohonen의 학습법칙등이 개발되어 사용되어 왔으나 Underutilization 문제가 있어 실제 사용상에 문제가 있어 왔다. 본 논문에서 제시하는 학습법칙은 이를 부분적으로 해결하였다. 또한 이 학습법칙을 ART(Adaptive Resonance Theory)-1과 Kohonen의 자기 구조 특징 지도의 장점을 조합한 개선된 IAFC(Integrated Adaptive Fuzzy Clustering)신경회로망에 적용하였고, 성능을 평가하기 위해 가우시안 분포의 데이터와 IRIS 데이터를 각각 사용하여 실험하였다.

### I. 서론

Kohonen의 학습법칙은 현재의 데이터와 승자의 연결강도의 차이에 학습률을 곱하여 조정하는데 이 학습률은 연결강도를 수렴시키기 위하여 0과 1사이의 임의의 작은 수를 취하여 시간이 지남에 따라 0으로 감소하도록 한다 [1]. 이에 비해 Huntsberger와 Ajjimarangsee는 위의 학습률 대신에 퍼지소속도를 사용하였는데

연결강도가 수렴하는 지에 대한 증명이 없었다 [2]. Kim과 Mitra는 학습률 대신에 반복횟수의 함수와 퍼지소속도와 Within-cluster 소속도를 조합하여 사용하였다 [3]. Chung과 Lee는 경쟁에서 승리한 승자와 패자 모두의 연결강도를 조정하였고, 이 연결강도를 조정하기 위해 학습률과 퍼지 소속도를 조합하여 사용하였는데

Kim과 Mitra의 학습법칙과 크게 다른 점은 모든 출력 뉴런의 연결강도를 조정한다는 것이다 [4]. 본 논문에서 제시하는 학습법칙은 Kim과 Mitra의 학습법칙과 Chung 과 Lee의 학습법칙을 조합한 것으로, 이는 누설 경쟁 학습(leaky competitive learning)을 퍼지화한 것과 조건 확률의 이론을 퍼지화 한 것에 기반을 둔 것이다. 누설 경쟁 학습은 승자뿐만 아니라 패자의 연결강도를 데이터와 현재의 연결강도 차이에 학습률을 곱하여 조정되 승자의 연결강도를 조정하기 위한 학습률이 패자의 연결강도를 조정하기 위한 학습률보다 크도록 하는 것이다. 이러한 학습률의 크기를 퍼지 소속도로 조절하는 것이 퍼지 누설 경쟁 학습이다. 즉, 승자는 퍼지 소속도가 크고, 패자는 퍼지 소속도가 작은 것을 이용하는 것이다. 조건 확률은 여러 데이터 클래스 중에서 하나의 클래스를 선정할 확률과 그 클래스내에서 데이터가 존재할 확률을 곱한 것인데 이러한 확률들을 Bezdek이 개발한 퍼지 소속함수 [5]를 변형한 것과 Pal등이 개발한  $\Pi$ 함수로 대체 한 것이다 [6]. 위에 언급한 학습법칙을 ART(Adaptive Resonance Theory)-1[7]과 Kohonen의 자기구조 특징 지도(Self-Organizing Feature Map)을 조합한 구조에 적용하였고, vigilance test에 사용하는 유사척도는 퍼지 소속도를 활용한 상대적 거리를 사용하였다. 이 퍼지 신경회로망의 성능을 Kohonen의 자기 구조 특징 지도의 성능과 비교하기 위하여 가우시안 분포의 데이터와 IRIS 데이터를 사용하였다.

## II. 본론

학습법칙은 신경회로망의 성능을 좌우하는 중요한 요소의 하나이다. 이 학습법칙 중에서 Kohonen의 학습법칙이 가장 많이 사용되어왔고 그 식은 아래와 같다.

$$V_i(t+1) = V_i(t) + \lambda(X - V_i(t))$$

여기서  $V_i$ 는 승자의 연결강도이고,  $X$ 는 데이터이며,  $\lambda$ 는 학습률로 0과 1사이의 값을 가지며 수렴시키기 위하여 시간이 무한대로 갈 때 0으로 간다. 이 학습 법칙은 초기에 연결강도의 적절한 값을 모르기 때문에 임의의 작은 수로 초기화하는데 Underutilization문제가 발생할 수 있다.

Huntsberger와 Ajjimarangsee는 학습률 대신에 퍼지 소속도를 사용했는데 그 식은 아래와 같다.

$$V_i(t+1) = V_i(t) + \mu_i \cdot N_i(t)(X - V_i(t))$$

여기서  $\mu_i$ 는 데이터  $X$ 의  $i$ 번째 클러스터에 대한 퍼지 소속도이고,  $N_i(t)$ 는 시간  $t$ 에서의 주변 함수이다. 그러나, 이 학습법칙은 퍼지 소속도를 최초로 학습법칙에 사용하였으나, 연결강도가 수렴하는지에 대한 증명이 없었다.

Kim과 Mitra는 최초로 조건확률(Conditional probability)를 퍼지화 함과 동시에 반복횟수의 함수와 조합하여 사용하였고 그 식은 아래와 같다.

$$V_i(t+1) = V_i(t) + f(t) \cdot \mu_i^2 \cdot$$

$$\Pi(X, V_i(t), T)(X - V_i(t))$$

여기서  $f(t)$ 는 반복횟수의 함수로

$$f(t) = \frac{1}{1+k(t-1)} \quad (k \text{는 상수}) \text{이다. } \mu_i [5] \text{는}$$

$$\mu_i = \frac{\left[ \frac{1}{\|X - V_i\|^2} \right]^{\frac{1}{m-1}}}{\sum_{j=1}^n \left[ \frac{1}{\|X - V_j\|^2} \right]^{\frac{1}{m-1}}}$$

이고  $m$ 은 weight exponent로 실험적으로 2로 놓았다. Bezdek의  $\mu_i$ 와 다른점은  $n$ 이 클러스터링 결과에 따라 변한다는 것이다.

$\Pi(X, V_i(t), T)$ [6]는

$$\Pi(x, v_i^{(old)}, T) = \begin{cases} 1 - 2 \left( \frac{|x - v_i^{(old)}|}{T} \right)^2, & \text{when } 0 \leq |x - v_i^{(old)}| \leq \frac{T}{2} \\ 2 \left( 1 - \frac{|x - v_i^{(old)}|}{T} \right)^2, & \text{when } \frac{T}{2} \leq |x - v_i^{(old)}| \leq T \\ 0, & \text{when } |x - v_i^{(old)}| \geq T \end{cases}$$

이다.  $\mu_i$ 는 데이터 X의 i번째 클러스터에 대한 퍼지 소속도이다. Chung과 Lee는 승자와 패자 모두의 연결강도를 조정하고 이 연결강도를 조정하기 위해 학습률과 퍼지 소속도를 같이 사용하였다.

본 논문에서 제시하는 학습법칙은 Kim과 Mitra의 학습법칙과 Chung과 Lee의 학습법칙을 조합한 것으로, 이는 누설 경쟁학습을 퍼지화 한 것과 조건확률의 이론을 퍼지화한 것에 기반을 둔 것으로 그 식은 아래와 같다.

$$V_i(t+1) = V_i(t) + f(t) \cdot \mu_i^2 \cdot$$

$\Pi(X, V_i(t), T)(X - V_i(t))$  for all i 이다. 위의 식이 Kim과 Mitra의 학습법칙과 다른 점은 경쟁에서 승리했든 패배했든에 관계없이 모든 클러스터의 연결강도를 조절하는 것이다.

위의 학습법칙을 ART-1과 Kohonen의 자기 구조 특징지도의 장점을 조합한 개선된 IAFC 신경회로망에 적용하였고 [3], 성능을 Kohonen의 자기 구조 특징 지도와 비교하기 위하여 가우시안 분포의 데이터와 IRIS 데이터를 사용하여 실험하였다. IRIS 데이터는 3개의 클러스터로 이루어져 있고 150개의 4차원 데이터이다. 각 클러스터는 50개의 데이터로 이루어져 있다. 사용한 가우시안 분포의 데이터는 2개의 클래스로 이루어져 있으며 200개의 2차원 데이터이다. 첫 번째 클래스는 평균값이 (2,5)이고 표준 편차가 1이며 100개의 데이터로 이루어져 있고 두 번째 클래스는 평균값이 (5,2)이고 표준 편차는 1이며 100개의 데이터로 이루어져 있다. 개선된 IAFC 신경회로망에 있어서 IRIS 데이터의 경우는 K가 0.5일 때는 vigilance parameter T값이 1.6에서 2.0사이에서 제대로 클러스터링하여 3개의 클러스터를 형성하였고 k가 1.0일 때는 vigilance parameter T값이 1.7에서 2.0 사이에서 제대로 클러스터를 형성하였다 [그림 1].

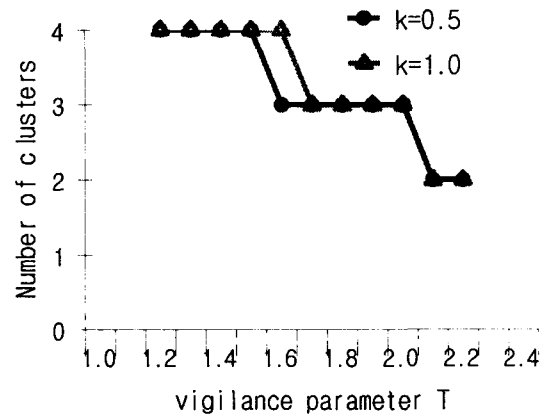


그림 1. vigilance parameter와 k값과 cluster 갯수

가우시안 분포의 데이터의 경우는 k가 0.5일 때 vigilance parameter T값이 4.0과 5.5사이에서 제대로 클러스터링하여 2개의 클러스터를 형성하였다. 반복수행에 있어서 각각의 특징의 조정된 양을 자승한 것을 모든 특징에 대하여 더한 것이 0.00001보다 적을 때 수렴한 것으로 보고 중단 시켰다. IRIS 데이터의 경우는 k가 0.5이고 T가 1.7일 때 error가 11개 발생하였고 32회 반복 수행하였다. 가우시안 분포의 데이터의 경우는 k가 0.5이고 T가 5.5일 때 error가 9개 발생하였고 50회 반복 수행하였다.

Kohonen의 자기 구조 특징 지도의 경우에는 IRIS 데이터에 대해서 연결강도를 0과 1사이의 가우시안 분포를 갖는 임의의 수로 초기화 했을 때 1개의 출력 뉴런만 선정되었고, 0과 7사이의 가우시안 분포를 갖는 임의의 수로 초기화 했을 때 2개의 출력 뉴런만 선정되었다. 이러한 결과는 임의의 수로 초기화 하는 것이 쉽지 않음을 보여준다. 따라서, 데이터 중 임의로 선정하여 연결강도의 초기값으로 사용하는 방법을 사용하였는데 선정된 값과 학습률에 따라 error가 16개 또는 14개 발생하였고 필요한 반복횟수는 83회였다. 가우시안 분포의 데이터

에 대해서는 임의의 작은 값으로 연결강도를 초기화했을때는 190개 가까이 error가 나올 때도 있고 데이터 중에서 임의로 선정하여 초기화 했을때는 10개의 error가 발생하나 700회 이상의 반복횟수가 필요하였다.

### III. 결론

Kohonen의 자기 구조 특징지도 신경회로망을 가우시안 분포 데이터와 IRIS 데이터에 대하여 여러 가지 값을 연결강도의 초기값으로 사용하여 테스트해 본 결과 처리하고자 하는 데이터 중에서 임의로 선정한 데이터 값을 연결강도의 초기값으로 사용했을 때 가장 좋은 결과를 얻었다. 또한 같은 데이터들을 사용하여 개선된 IAFC 신경회로망의 성능을 테스트 해본 결과는 개선된 IAFC 신경회로망의 구조의 우수성을 입증하였다. Kohonen의 자기 구조 특징 지도 신경회로망의 연결강도의 초기값으로 데이터중 임의로 선정한 경우에 비해서도 개선된 IAFC 신경회로망이 실험적으로 수렴하는데 필요한 반복횟수와 error면에서 우수함이 입증되었다.

### IV. 참고문헌

1. T. Kohonen, Self-Organization and Associative Memory, 3rd ed., Springer-Verlag, Berlin, 1984.
2. T. L. Huntsberger and P. Ajjimarangsee, "Parallel Self-Organizing Feature Maps for Unsupervised Pattern Recognition," Int'l. J. General Systems, Vol. 16, NO. 14, pp. 357-372, 1990.
3. Y. S. Kim and S. Mitra, "An Adaptive Integrated Fuzzy Clustering Model for Pattern Recognition," Fuzzy Sets and Systems, Vol. 65, pp. 297-310. 1994.
4. F. L. Chung and T. Lee, "Fuzzy Competitive Learning," Neural Networks,

Vol. 7, No. 3, pp. 538-551, 1994.

5. J. C. Bezdek, Pattern Recognition with Fuzzy Objective Function Algorithms, Plenum Press, New York, 1981.

6. S. K. Pal and S. Mitra, "Fuzzy Dynamic Clustering Algorithm," Pattern Recognition Letters, Vol. 11, 1990.

7. G. A. Carpenter and S. Grossberg, "A Massively Parallel Architecture for A Self-organizing Neural Pattern Recognition Machine," Computer Vision, Graphics, and image Processing, Vol. 37, pp. 54-115, 1987.

8. Yong S. Kim, "An Unsupervised Neural Network Using a Fuzzy Learning Rule," 1999 IEEE International Fuzzy Systems Conference Proceedings, Vol. I, pp. 349-353, Seoul, Korea, 1999 August.