

비대칭 퍼지 학습률을 이용한 퍼지 신경회로망 모델

Fuzzy Neural Network Model Using Asymmetric Fuzzy Learning Rates

김용수

대전대학교 컴퓨터공학부

Yong-Soo Kim

Division of Computer Engineering

Daejeon University

E-mail : kystj@dju.ac.kr

요 약

본 논문에서는 LVQ(Learning Vector Quantization)을 퍼지화한 새로운 퍼지 학습 법칙을 제안하였다. 퍼지 LVQ 학습 법칙 3은 기존의 학습률 대신에 퍼지 학습률을 사용하였는데, 기존의 LVQ와는 달리 비대칭인 학습률을 사용하였다. 기존의 LVQ에서는 분류가 맞거나 틀렸을 때 같은 학습률을 사용하고 부호만 달랐으나, 새로운 퍼지 학습 법칙에서는 분류가 맞거나 틀렸을 때 부호가 다를 뿐만 아니라 학습률도 다르다. 이 새로운 퍼지 학습 법칙을 무감독 신경회로망인 improved IAFC(Integrated Adaptive Fuzzy Clustering) 신경회로망에 적용하여 감독 신경회로망으로 변형하였다. Improved IAFC 신경회로망은 유연성이 있으면서도 안정성이 있다. 제안한 supervised IAFC 신경회로망 3의 성능과 오류 역전파 신경회로망의 성능을 비교하기 위하여 iris 데이터를 사용하였는데 Supervised IAFC 신경회로망 3가 오류 역전파 신경회로망보다 성능이 우수하였다.

1. 서론

신경회로망은 뉴런들이 상호 연결된 네트워크이고, 이 뉴런들은 연결강도에 의해 상호 연결되어 있다. 이 연결강도들은 신경회로망의 성능을 개선하기 위해서 조절이 되는데, 연결강도를 조절하는 학습법칙은 신경회로망의 성능을 좌우하는 중요한 요소들 중의 하나이다. 학습법칙에는 무감독 학습법칙과 감독 학습 법칙과 강화 학습 법칙이 있다. LVQ는 감독 학습 법칙의 하나이다. LVQ는 분류가 맞으면 승자의 연결강도를 입력 벡터 쪽으로 조정하고, 분류가 틀리면 연결강도를 입력 벡터 쪽으로 부터 멀리 보낸다 [1, 2]. 이러한 LVQ를 퍼지화하는 연구가 진행되어 왔는데, Chung과 Lee는 퍼지 소속도를 학습률과 같이 사용하여 LVQ를 퍼지화 하였다 [3]. 그들은 목표 소속도와 실제의 소속도의 차이에 학습률을 곱한 것을 사용하였으나 현실적으로 목표

소속도를 얻는 것이 어려운 문제점을 안고 있다. Kusumoputro 등은 또한 LVQ를 퍼지화 하였다 [4]. 그들은 퍼지 소속도와 학습률을 사용하여 LVQ를 퍼지화 하였다.

본 논문에서는 LVQ를 퍼지화한 새로운 퍼지 학습 법칙을 제안한다. 이 새로운 학습 법칙은 LVQ와 같이 분류가 맞으면 승자의 연결강도를 입력 벡터 쪽으로 조정하고 분류가 틀리면 연결강도를 입력 벡터 쪽으로 부터 멀리 보내는 것은 같으나, 분류가 맞느냐 틀리느냐에 따라 조정하는 학습률이 다르다. 분류가 틀리면 퍼지 학습률인 반복횟수의 함수와 Π 소속 함수와 퍼지 소속도, μ_i 의 조합이 사용된다. 반면에, 분류가 맞으면 퍼지 학습률인 반복횟수의 함수와 Π 소속 함수와 1과 퍼지 소속도의 차이, $1-\mu_i$,의 조합이 사용된다. 분류가 틀리면 입력벡터가 선정된 클러스터에 소속되어 있는 정도를 반영하는 퍼지 소

속도, μ_i , 를 사용하여 선정된 클래스의 연결강도를 입력 벡터로부터 멀리 보낸다. 반면에, 분류가 맞으면 1과 퍼지 소속도의 차이, $1 - \mu_i$, 를 사용하는데, 이는 결정 경계선의 근처에 존재하는 입력 벡터들이 최적의 결정 경계선에 대한 정보를 더 가지고 있다는 것에 기반을 둔 것이다. 결정 경계선에서 멀리 떨어져 있으면서 한 쪽에 치우쳐 있는 입력 벡터는 그 클래스에 대한 퍼지 소속도가 커서 $1 - \mu_i$ 인 가중치가 작다. 반면에 결정 경계선 근처에 있는 입력 벡터는 선정된 클래스에 대한 퍼지 소속도가 작아서 $1 - \mu_i$ 인 가중치가 크다. 이러한 가중치를 사용하여 최적의 결정 경계선에 대한 정보를 더 가지고 있는 입력 벡터에 대한 반영도를 크게 하는 것이다.

제안한 LVQ를 퍼지화한 퍼지 학습 법칙을 무감독 신경 회로망인 improved IAFIC 신경회로망에 적용하여 감독 신경 회로망인 Supervised IAFIC 신경회로망 3를 제안한다 [5, 6]. Improved IAFIC 신경회로망은 ART-1 신경회로망 [7] 과 Kohonen의 Self-Organizing Feature Map [8]의 장점을 취합한 것이다. Improved IAFIC 신경회로망은 ART-1 신경회로망과 유사한 제어 구조를 가지고 있기 때문에 ART-1 신경회로망과 같이 안정성과 유연성을 가지고 있다. 이 신경회로망은 과거에 배운 내용을 보전하는 안정성이 있으면서도 새로운 입력 벡터를 반영하는 유연성이 있다. ART-1 신경회로망은 잡음에 민감한데 비하여 improved IAFIC 신경회로망은 ART-1 신경회로망에 비하여 덜 민감하다. 이는 학습 법칙이 다른 데에서 기인한다. 한편 Kohonen의 Self-Organizing Feature Map은 임의의 작은 수로 연결강도를 초기화 해야 하기 때문에 Underutilization 문제가 있다. Improved IAFIC 신경회로망은 ART-1 신경회로망과 유사한 제어 구조를 사용하기 때문에 연결강도를 임의의 작은 수로 초기화 할 필요가 없어서 Underutilization 문제가 없다. Improved IAFIC 신경회로망은 vigilance parameter 를 사용하여 클래스의 크기와 개수를 조절한다. 그러나 improved IAFIC 신경회로망의 vigilance parameter 의 의미는 ART-1 신경회로망의 vigilance parameter 와는 다르다. 이는 vigilance parameter 의 유클리디안 영역에서의 재해석에 기인한다. Improved IAFIC 신경회로망에서 vigilance parameter 는 거리 임계값 또는 클러스터 반지름과 관계가 있다 [9].

제안한 퍼지 신경회로망의 성능과 오류 역전파

신경회로망의 성능을 비교하기 위하여 iris 데이터를 사용하였다. 이 iris 데이터는 클러스터링 알고리즘의 성능을 비교하는데 많이 사용되어 왔다.

2. Supervised IAFIC 신경회로망 3

Supervised IAFIC 신경회로망 3는 ART-1 과 유사한 제어구조를 가지고 있다. 따라서, Supervised IAFIC 신경회로망 3는 안정하면서도 유연하다. 또한 임의의 작은 수로 연결강도를 초기화할 필요가 없기 때문에 underutilization 문제를 해결한다. 이 퍼지 신경회로망 들은 클래스들의 크기와 개수를 조절하기 위하여 vigilance parameter를 사용하는데, 이 vigilance parameter 의 의미는 ART-1 신경회로망의 vigilance parameter와는 다르다. Supervised IAFIC 신경회로망 3은 유클리디안 거리를 비유사 척도 (dissimilarity measure)로 사용한다. 따라서, vigilance 테스트에서 이 퍼지 신경회로망은 비유사도가 vigilance parameter보다 작으냐를 테스트한다.

입력 벡터가 Supervised IAFIC 신경회로망 3에 가해진 후, 출력 뉴런들 간의 경쟁은 winner-take-all 식으로 일어난다. 입력 벡터와의 유클리디안 거리가 제일 작은 연결강도를 가진 출력 뉴런이 경쟁에서 승리한다. 즉 I번째 출력 뉴런,

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

이 경쟁에서 승리한다. \mathbf{x} 는 입력 벡터이고, $\mathbf{v}_i(t)$ 는 i번째 출력 뉴런의 연결강도이다.

출력 뉴런들 중에 승자를 결정한 후에, supervised IAFIC 신경회로망 3은 다음의 vigilance 기준에 의해 vigilance test 를 한다:

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

여기서 T는 vigilance parameter 이다.

퍼지 소속도 μ_I 는 다음과 같이 정의 된다:

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

여기서 n은 committed 출력 뉴런의 수이고

$m^{[1]}$ 은 weight exponent이며 실험적으로 2로 놓는다. 위의 새로운 vigilance 기준은 퍼지 소속도를 이용하여 입력 벡터의 존재하는 클래스의 연결강도들과의 상대적 거리를 고려한 것이다. 결정 경계선에 가까운 곳에 입력 벡터가 위치할 때 클래스들 간의 경쟁이 치열하기 때문에 입력 벡터의 위치가 조금만 변해도 승자를 결정하는 것에 영향을 미칠 수 있다. 반면에 입력 벡터가 경쟁이 약한 곳에 위치할 때는 입력 벡터의 위치가 변하여도 승자를 결정하는 것에 영향을 미치지 않는다. 따라서, 입력 벡터가 경쟁이 치열한 곳에 위치할 때는 입력 벡터와 승자의 연결강도 간의 유클리디안 거리에 큰 가중치를 주고, 경쟁이 약한 곳에 위치할 때는 작은 가중치를 준다. 이는 Mahalanobis 거리가 유클리디안 거리에 통계적 성질을 고려하여 가중치를 다르게 주는 것과 비교할 수 있다. Mahalanobis 거리는 공분산(covariance)을 이용하여 공분산이 크면 공분산 행렬의 역을 곱하므로 가중치가 작고, 공분산이 작으면 가중치가 크다 [10]. 그러나, 존재하는 클래스의 개수가 1 이면 $\|x - v_i(t)\| \leq T$ 를 vigilance 기준으로 사용한다.

만약 승자가 vigilance 테스트를 만족하면은 supervised IAFC 신경회로망 3은 승자의 연결강도를 다음의 식에 의하여 조정 한다:

$$v_i(t+1) = v_i(t) + f(t) \frac{\|x - v_i(t)\|}{T} (1 - u_i) (x - v_i(t))$$

if x is classified correctly,

$$v_i(t+1) = v_i(t) - f(t) \frac{\|x - v_i(t)\|}{T} u_i (x - v_i(t))$$

if x is classified incorrectly,

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

위의 식에서 $f(t)$ 는 $1/(1+k(t-1))$ 이다.

$\Pi(x, v_i(t), T)$ 는 아래의 식으로 정의 된다:

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

위의 퍼지 학습 법칙에서 $1 - \mu_i$ 를 사용하는 이유는 결정 경계선 근처에 있는 입력 벡터가 최적의 결정 경계선에 대한 정보를 더 많이 가지고 있기 때문에 연결 강도를 조정할 때 결정 경계선 근처에 있는 입력 벡터를 더 반영하기 위한 것이

다. 결정 경계선으로부터 멀리 있는 특정의 클래스에 치우쳐 있는 입력 벡터는 퍼지 소속도가 크기 때문에 $1 - \mu_i$ 를 사용하여 반영도를 낮춘 것이다. 이에비해 결정 경계선 근처에 있는 입력 벡터의 선정된 클래스에 대한 퍼지 소속도가 작아 μ_i 를 사용하여 반영도를 높인것이다. 분류가 틀렸을 때는 퍼지 소속도를 사용하는데 이는 입력 벡터가 잘못 선정된 클래스에 소속되어 있는 정도를 반영하여 연결강도를 입력 벡터로부터 멀리 보내는 것이다.

알고리즘을 요약하면 아래와 같다.

- [1] 파라미터들(t, k)과 입력 뉴런들과 출력 뉴런들을 초기화 한다.
- [2] 입력 벡터 x 를 가한다.
- [3] 입력 벡터가 첫 번째 입력 벡터이면 (4)로 가고; 그렇지 않으면 (5)로 간다.
- [4] 입력 벡터를 첫 번째 출력 뉴런의 연결강도로 하고 (2)로 간다.
- [5] 승자인 출력 뉴런을 찾는다.
- [6] 승자인 출력 뉴런이 vigilance 테스트를 만족하면 (7)로 가고; 그렇지 않으면 (8)로 간다.
- [7] 승자인 출력 뉴런의 연결강도를 조정하고 (2)로 간다.
- [8] 첫 번째 uncommitted 출력 뉴런을 활성화하고 입력벡터를 첫 번째 uncommitted 출력 뉴런의 연결강도로 한다. (2)로 간다.

3. 실험 및 결과

제안한 퍼지 신경회로망의 성능과 오류 역전과 신경회로망의 성능을 비교하기 위하여 클러스터링 알고리즘의 성능을 비교하는데 많이 쓰이는 iris 데이터를 사용하였다. 오류 역전과 신경회로망은 감독 학습 법칙을 사용한 신경회로망 중에서 잘 알려진 신경회로망 모델이다. 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 subspecies 로부터 25개의 데이터를 사용하였다. 나머지 75개의 데이터는 테스트 데이터로 사용하였다.

반복 학습을 하는 중에 현재의 연결강도와 이전의 연결강도의 차이인 $\sqrt{\|V(t+1) - V(t)\|}$

이 0.01보다 작으면 연결강도 들이 실험적으로 수렴한 것으로 보고 supervised IAFC 신경회로망 3의 훈련을 끝내도록 하였다. 훈련을 끝낸 후 supervised IAFC 신경회로망 3을 테스트 하였다. 그림 1은 supervised IAFC 신경회로망 3의 성능과 오류 역전과 신경회로망의 성능을 비교하였다.

		Desired		
		1	2	3
Actual	1	23		
	2	2	25	
	3			25

(1) Supervised IAFC 신경회로망 3

		Desired		
		1	2	3
Actual	1	20		
	2	5	25	
	3			25

(2) 오류 역전과 신경회로망

그림 1 iris 데이터를 사용한 두 개의 신경회로망들의 성능 비교

Supervised IAFC 신경회로망 3를 훈련시키기 위해 7번의 반복 학습을 하였으며, T가 1.7이고 k가 0.5일 때 2개의 오류를 발생하였다. 오류 역전과 신경회로망은 5개의 오류를 발생시켰다. 그림 2는 Supervised IAFC 신경회로망 3를 훈련시킬 때 반복 횟수에 따른 오류의 개수를 보여준다.

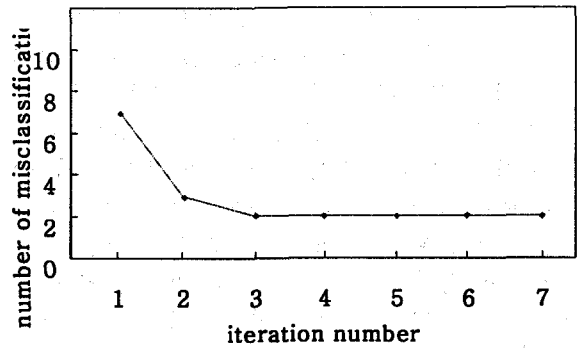


그림 2 Supervised IAFC 신경회로망 3를 훈련시킬 때 반복 횟수에 따른 오류의 개수

4. 결론

본 논문에서 LVQ를 퍼지화한 새로운 퍼지 학습 법칙을 제안하였고, 이 새로운 퍼지 학습 법칙들을 사용하는 Supervised IAFC 신경회로망 3를 제안하였다. Supervised IAFC 신경회로망 3의 성능과 오류 역전과 신경회로망의 성능을 비교하기 위하여 iris 데이터를 사용하였다. Supervised IAFC 신경회로망 3가 오류 역전과에 비해 적은 오류를 발생시켰으며, 훈련에 필요한 반복횟수가 적음을 보여주었다.

4. 참고문헌

- [1] Kohonen, T., "Self-Organizing Map," Proceeding of the IEEE, Vol. 78, No.9, pp.1464-1480, 1990.
- [2] Lin, C-T, and Lee, C. S. G., Neural Fuzzy Systems: A Neuro-Fuzzy Synergism to Intelligent Systems, Upper Saddle River, NJ : Prentice-Hall, 1996.
- [3] Chung, F.-L., and Lee, T., "A Fuzzy Learning Model for Membership Function Estimation and Pattern Classification," Proceeding of the third IEEE Conference on Fuzzy Systems, Vol. 1, pp. 426-431, 1994.
- [4] Kusumoputro, B., Widyanto, M. R., Fanany, M. I., and Budiarto, H., "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.
- [5] Kim, Y. S., and Mitra, S., "An adaptive integrated fuzzy clustering model for

- pattern recognition," Fuzzy Sets and Systems, Vol. 65, pp. 297 - 310, 1994.
- [6] Kim, Y. S., "An Unsupervised neural network Using a Fuzzy Learning Rule," Proceedings of 1999 IEEE International Fuzzy Systems Conference, Vol. 1, pp. 349-353, 1999.
- [7] Carpenter, G. A., and Grossberg, S., "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] Kohonen, T., Self-Organization and Associative Memory, 3rd ed., NewYork : Spring-Verlag, 1989.
- [9] Moore, B., "Art-1 and Pattern Clustering," Proceedings of the 1988 Connectionist Models Summer School, pp. 174-185, 1989.
- [10] Tou, J. T., and Gonzalez, R. C., Pattern Recognition Principles, MA : Addison Wesley, 1974.