패턴 인식을 위한 감독학습을 사용한 IAFC(Integrated Adaptive Fuzzy Clustering)모델

IAFC(Integrated Adaptive Fuzzy Clustering)Model Using Supervised Learning Rule for Pattern Recognition

김용수, 김남진
대전대학교 컴퓨터공학부
Yongsoo Kim, Namjin Kim
Division of Computer Engineering
Daejeon University
E-mail:kystj@dju.ac.kr

이재연,지수영,조영조
지능형로봇연구단
한국전자통신연구원
Jaeyeon Lee, Suyoung Chi, Youngjo Cho
Intelligent Robot Research Division
ETRI

이세열
청운대학교 컴퓨터과학과
Seyul Lee
Dept. of Computer Science
Chungwoon University

요 약

본 논문은 패턴인식을 위해 사용할 수 있는 감독학습을 이용한 supervised IAFC neural network 1과 supervised IAFC neural network 2를 제안하였다. Supervised IAFC neural network 1과 supervised IAFC neural network 2는 LVQ(Learning Vector Quantization)를 퍼지화 한 새로운 퍼지 학습법칙을 사용하고 있다. 이 새로운 퍼지 학습 법칙은 기존의 학습률 대신에 퍼지화 된 학습률을 사용하고 있는데, 이 퍼지화된 학습률은 조건 확률을 퍼지화 한 것에 근간을 두고 있다. Supervised IAFC neural network 1과 supervised IAFC neural network 2의 성능과 오류역전파 신경회로망의 성능을 비교하기 위하여 iris 데이터를 사용하였는데, 실험결과 supervised IAFC neural network 2 의 성능이 오류역전파 신경회로망의 성능보다 우수함이 입증되었다.

1. 서론

신경회로망은 뉴런들이 상호 연결된 네트워크이고, 이 뉴런들은 연결강도들에 의하여 상호 연결되어 있다. 연결강도들은 신경회로망의 성능을 개선하기 위하여 조정이 된다. 따라서, 연결강도들의 조정을 조절하는 학습법칙은 신경회로망의 성능을 좌우하는 중요한 요소들 중의 하나이다. 학습법칙들은 크게 2 그룹으로 나눌 수 있는데, 하나는 무감독 학습법칙이고, 다른 하나는 감독 학습법칙이다. 무감독 학습법칙은 신경회로망에 입력 벡터만을 가하고 감독 학습법칙은 입력 벡터와 원하는 값을 신경회로망에 가한다.

Kohonen 학습 법칙은 무감독 학습법칙들 중에서 널리 사용되고 있다. Kohonen 학습 법칙은 선정된 클래스의 연결 강도를 입력 벡터와 선정된 클래스의 연결강도의 차이에 비례하여 조정한다. 이 차이는 연결 강도들을 수렴시키기 위해 사용되는 학습률에 의해 곱해 진다[1,2]. Huntsberger와 Ajjimaranges는 학습률 대신에 퍼지 소속도를 사용하여 Kohonen 학습 법칙을 퍼지화하였다[3]. 그러나, 그들의 학습 법칙은 연결강도들이 수렴하는 것을 보장하지 않는다. Kim과 Mitra는 Kohonen 학습 법칙을 퍼지화했는데[4], 학습률 대신에 반복 횟수의 함수와 Π 소속함수[5]와 퍼지소속함수를 사용하였다. Chung과 Lee 또한 학습률 대신에 학습률과 퍼지소속도의 조합을 사용함으로써 Kohonen 학습 법칙을 퍼지화하였다[6]. Bezdek은 Fuzzy C-Means 모델을 Kohonen 학습 법칙과 통합함으로써 Kohonen 학습

법칙을 퍼지화하였다[7].

LVQ는 감독학습 법칙의 하나이다. LVQ는 분류가 맞으면 승자의 연결강도를 입력 벡터쪽으로 조정하고, 분류가 틀리면 연결강도를 입력 벡터쪽으로부터 멀리 보낸다[2,7]. Chung과 Lee는 퍼지 소속도와 LVQ를 통합한 퍼지 LVQ를 제안하였다[8]. 그들은 학습률과 목표 소속도와 실제의 소속도의 차이의 조합을 사용하였다. 본 논문은 LVQ를 퍼지화하는 2개의 새로운 퍼지 학습 법칙을 제안하였다. 이 퍼지 학습 법칙들은 반복횟수의 함수와 Π소속 함수와 퍼지 소속함수를 LVQ의 학습률 대신에 사용하였다.

제안한 퍼지 학습 법칙들을 improved IAFC 신경회로망에 적용하였다[9,10]. Improved IAFC 신경회로망은 ART-1과 유사한 제어 구조를 가지고 있기 때문에 ART-1과 같이 안정성(Stability)과 유연성(Plasticity)을 동시에 가지고 있다[11]. 이 신경회로망은 과거에 배운 내용을 보전하는 안정성이 있고 새로운 입력 벡터를 반영하는 유연성이 있다. Improved IAFC 신경회로망은 vigilance parameter를 사용하여 클러스터의 크기와 개수를 조절한다. 그러나 Improved IAFC 신경회로망의 vigilance parameter의 의미는 ART-1의 vigilance parameter와는 다르다. 이는 vigilance parameter의 유클리디안 영역에서의 재해석에 기인한다. Improved IAFC 신경회로망에서 vigilance parameter는 거리 임계값(distance threshold) 또는 클러스터 반지름(cluster diameter)과 관계가 있다[12]. ART계열의 신경회로망들을 제외한 많은 신경회로망 모델들은 연결강도들을 임의의 작은 수들로 초기화를 시킬 필요가 있다. 이러한 초기화는 underutilization문제를 일으킬 수 있다. Improved IAFC 신경회로망은 이러한 문제를 해결했다.

제안한 신경회로망들의 성능과 오류역전파 신경회로망의 성능을 비교하기 위하여 iris 데이터를 사용하였다. 이 iris 데이터는 클러스터링 알고리듬들의 성능을 비교하기 위하여 많이 사용되어 왔다.

2. Supervised IAFC 신경회로망

Supervised IAFC 신경회로망들은 ART-1과 유사한 제어구조를 가지고 있다. 따라서, 이들은 안정하면서도 유연하다. 또한 임의의 작은 수로 연결강도를 초기화할 필요가 없기 때문에 underutilization 문제를 해결한다. 이 신경회로망들은 클러스터들의 크기와 개수를 조절하기 위하여 vigilance parameter를 사용하는데, vigilance parameter의 의미는 ART-1의 vigilance parameter와는 다르다. Supervised IAFC 신경회로망들은 유클리디안 거리를 비유사척도(dissimilarity measure)로 사용한다. 따라서, vigilance 테스트에서 이 신경회로망들은 비유사도가 vigilance parameter보다 작으냐를 테스트한다.

입력 벡터가 supervised IAFC 신경회로망들에 가해진 후, 출력 뉴런들간의 경쟁은 winner-take-all 식으로 일어난다. 입력 벡터와의 유클리디안 거리가 제일 작은 연결 강도를 가진 출력 뉴런이 경쟁에서 승리한다. 즉 I번째 출력 뉴런,

$$I = \min_i \| x - v_i \|,$$

이 경쟁에서 승리한다. x는 입력 벡터이고, v_i는 I번째 출력 뉴런의 연결강도이다.

출력 뉴런들 중에 승자를 결정한 수에, supervised IAFC 신경회로망들은 다음의 vigilance 기준에 의해 vigilance test를 한다:

$$e^{-ru_I} \| x - v_i \| \leq T,$$

여기서 T는 vigilance parameter이고, r은 상수로서 일반적으로 1로 놓는다.
퍼지 소속도 u_i는 다음과 같이 정의 된다:

$$u_I = \frac{\left[\frac{1}{\| x - v_I \|}\right]^{\frac{1}{m-1}}}{\sum_{j=0}^{n}\left[\frac{1}{\| x - v_j \|^2}\right]^{\frac{1}{m-1}}},$$

여기서 n은 committed 출력 뉴런의 수이고, $m \in [1,\infty]$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 클래스간의 경쟁이 치열하기 때문에 입력 벡터의 위치가 조그만 변해도 승자를 결정하는 것에 영향을 미칠 수 있다. 반면에 입력 벡터가 경쟁이 약한 곳에 위치할 때는 입력벡터의 위치가 변하여도 승자를 결정하는 것에 영향을 미치지

않는다. 따라서, 입력 벡터가 경쟁이 치열한 곳에 위치할 때는 입력 벡터와 승자의 연결강도간의 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 Mahalanobis 거리는 작고, 공분산이 작으면 Mahalanobis 거리가 크다[13]. 그러나, 존재하는 클러스터의 개수가 1이면 $\| x-v_i \| \le T$를 vigilance 기준으로 사용한다.

만약 승자가 vigilance 테스트를 만족하면은 supervised IAFC neural network 1은 승자의 연결강도를 다음의 식에 의하여 조정 한다:

$$v_I(t+1) = v_I(t) + f(t)\cdot\prod(x, v_I(t), T)\cdot u_I\cdot(x - v_I(t))$$
if x is classified correctly,
$$v_I(t+1) = v_I(t) - f(t)\cdot\prod(x, v_I(t), T)\cdot u_I\cdot(x - v_I(t))$$
if x is classified incorrectly,
$$v_i(t+1) = v_i(t)$$
for $i \neq I$,

위의 식에서 f(t)는 1/1+k(t-1)이고 $\Pi(x, v_i(t), T)$는 아래의 식으로 정의 된다:

$$\prod(x, v_i(t), T) = 1 - 2\left(\frac{\| x - v_i(t)\|}{T}\right)^2 \quad when\ 0 \le \| x - v_i(t)\| \le \frac{T}{2},$$
$$2\left(\frac{\| x - v_i(t)\|}{T}\right)^2 \quad when\ \frac{T}{2} \le \| x - v_i(t)\| \le T,$$
$$0 \quad when\ \| x - v_i(t)\| \ge T.$$

위의 식에서 사용된 퍼지학습률은 조건 확률의 퍼지화에 근거를 두고 있다. 조건 확률은 존재하는 클래스들 중에서 하나의 클래스를 선정할 확률에 입력 벡터의 선정된 클래스에 존재하는 확률을 곱한 것이다. 퍼지소속도는 존재하는 클래스들 중에서 하나의 클래스를 선정할 확률을 퍼지화한것이다. Π함수는 입력 벡터가 선정된 클래스에 존재할 확률을 퍼지화한것이다.

다른 한편으로 supervised IAFC neural network 2는 승자의 연결강도를 아래의 식에 의해서 조정 한다:

$$v_I(t+1) = v_I(t) + f(t)\cdot\prod(x, v_I(t), T)\cdot(1 - u_I)\cdot(x - v_I(t))$$
if x is classified correctly,
$$v_I(t+1) = v_I(t) - f(t)\cdot\prod(x, v_I(t), T)\cdot(1 - u_I)\cdot(x - v_I(t))$$
if x is classified incorrectly,
$$v_i(t+1) = v_i(t)$$
for $i \neq I$,

알고리듬을 요약하면 아래와 같다.

(1) 파라미터들(t, k)과 입력 뉴런들과 출력 뉴런들을 초기화 한다.
(2) 입력 벡터 x를 가한다.
(3) 입력 벡터가 첫 번째 입력 벡터이면 (4)로 가고; 그렇지 않으면 (5)로 간다.
(4) 입력 벡터를 첫 번째 출력 뉴런의 연결 강도로 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 신경회로망의 성능과 오류역전파 신경회로망의 성능을 비교하기 위하여 클러스터링 알고리듬의 성능을 비교하는데 많이 쓰이는 iris 데이터를 사용하였다. 오류역전파 신경회로망은 감독학습을 사용한 신경회로망 중에서 잘 알려진 신경회로망 모델이다. iris 데이터는 150개의 4차원 데이터로 구성되어 있다[14].

이 데이터는 3개의 subspecies로부터 추출한 데이터이고, 각 subspecies는 50개의 데이터를 가지고 있다. 150개의 데이터 중 임의로 75개의 데이터를 선정해 훈련 데이터 집합으로 사용했는데, 각 subspecies로부터 25개의 데이터를 사용하였다. 나머지 75개의 데이터는 테스트 데이터 집합으로 사용하였다.

반복학습을 하는 중에 현재의 연결강도와 이전의 연결강도의 차이인 $\sqrt{\| V(t+1) - V(t)\|^2}$이 0.01보다 작으면 연결강도들이 실험적으로 수렴한 것으로 하여 supervised IAFC 신경회로망들의 훈련을 중단시키도록 하였다. 훈련이 중단된 이후에 supervised IAFC 신경회로망들을 테스트하였다. 다음 표는 supervised IAFC 신경회로망들의 성능

과 오류역전파 신경회로망의 성능을 비교한 것이다.

Actual \ Desired	1	2	3
1	19		
2	6	25	
3			25

(a) Supervised IAFC neural network 1

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

(b) 오류역전파 신경회로망

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

(c) Supervised IAFC neural network 2

그림 1. iris 데이터를 사용한 3개의 신경회로망들의 성능 비교

Supervised IAFC neural network 1은 훈련시키기 위해 11번 반복 학습을 하였으며, T가 1.55이고 k가 0.5일 때 6개의 오류를 발생시켰다. 그림 2는 Supervised IAFC neural netwroks 1을 훈련시킬 때 반복횟수에 따른 오류의 개수를 보여준다. 오류역전파 신경회로망은 5개의 오류를 발생시켰다. Supervised IAFC neural netwrok 2는 학습시키기 위해 9번의 반복학습을 시켰으며, 테스트할 때 2개의 오류를 발생시켰다. 그림 3은 Supervised IAFC neural network 2를 훈련시킬 때 반복학습에 따른 오류의 개수를 보여준다.

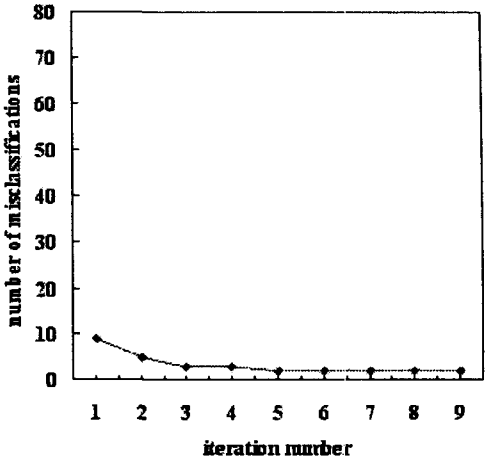

그림 2. Supervised IAFC neural network 1을 훈련시킬 때 반복횟수에 따른 오류의 개수

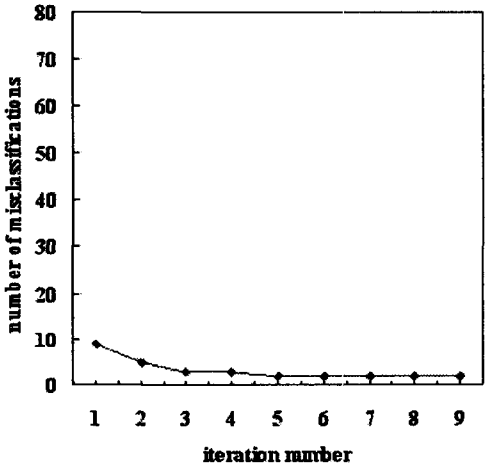

그림 3. Supervised IAFC neural network 2를 훈련시킬 때 반복횟수에 따른 오류의 개수

4. 결론 및 향후 과제

본 논문은 새로운 퍼지 학습 법칙들을 사용하는 supervised IAFC neural network 1과 supervised IAFC neural network 2를 제안하였다. 이 새로운 퍼지 학습법칙들은 LVQ의 퍼지화에 근거를 둔 것이다.

Supervised IAFC neural network 1 과 오류역전파 신경회로망과 supervised IAFC neural network 2의 성능을 비교하기 위하여 iris 데이터를 사용하였다. Supervised IAFC 신경회로망들은 훈련을 위한 반복횟수가 다른 신경회로망 모

델들에 비해 아주 작고, supervised IAFC neural network 2는 오류역전파 신경회로망보다 성능이 우수함을 보여주었다.

감사의 글 : 본 논문은 정보통신부에서 지원하는 " URC 과제" 에 의하여 지원되었으며 이에 감사드립니다.

5. 참고 문헌

[1] T. K. Kohonen, Self-Organization and Associative Memory, 3rd ed., New York : Springer-Verlag, 1989.

[2] C-T Lin and C. S. G. Lee, Neural Fuzzy Systems A Neuro-Fuzzy Synergism to Intelligent Systems, New Jersey : Prentice-Hall, 1996.

[3] T. L. Huntsberger and P. Ajjimarangsee, "Parallel Self-Organizing feature Maps for Unsupervised Pattern Recognition," Int. J. General Systems, Vol. 16, No. 4, pp. 357-372, 1990.

[4] Y. S. Kim and S. Mitra, "Integrated Adaptive Fuggy Clustering(IAFC) Algorithm," Proceedings of the Second IEEE International Conference on Fuzzy Systems, Vol. 2, pp. 1264-1268, San Francisco, 1993.

[5] S. K. Pal and S. Mitra, "Fuzzy dynamic clustering algorithm, " Pattern Recognition Letters, Vol. 11, p. 525-535, 1990.

[6] F. L. Chung and T. Lee, "Fuzzy Competitive Learning, " Neural Networks, Vol.7, No.3, pp. 539-551, 1992.

[7] J. C. Bezdek, E. C. Tsao, and N. R. Pal, "Fuzzy Kohonen Clustering Networks," Proceedings of the First IEEE conference on Fuzzy Systems, pp. 1035-1043, San Diego, 1992.

[8] F-L Chung and T. Lee, "A fuzzy Learning Model for Membership Function Estimation and Pattern Classification," Proceedings of the third IEEE Conference on Fuzzy Systems, Vol.1, pp. 426-431, 1994.

[9] Y. S. Kim and S. Mitra, "An adaptive integrated fuzzy clustering nodel for pattern recognition," Fuzzy Sets and Systems, Vol. 65, pp. 297-310, 1994.

[10] Y. S. Kim, "An Unsupervised neural network Using a Fuzzy Learning Rule," Proceedings of 1999 IEEE International Fuzzy Systems Conference, Vol. 1, pp. 349-353, Seoul, 1999.

[11] 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.

[12] B. Moore, "ART1 and Pattern Clustering," Proceedings of the 1988 Connectionist Models Summer School, pp.174-185, San Mateo, 1989.

[13] J. T. Tou and R. C. Gonzalez, Pattern Recognition Principles, MA : Addison Wesley, 1974.

[14] E. Anderson, "The IRISes of the Gaspe Penninsula," Bulletin American IRIS Society, Vol. 59, pp. 2-5, 1935.