

## 강화 학습을 이용한 퍼지 최소-최대 신경망의 학습 방법

곽병동, 박광현, 변증남  
한국과학기술원

# the Novel Learning Method of Fuzzy Min-Max Neural Network by Using the Reinforcement Learning

Byoung-Dong Kwak, Kwang-Hyun Park, Z. Zenn Bien

Dept. of EECS, KAIST

E-mail : goodkwak@ctrsys.kaist.ac.kr, akaii@robotian.net, zbien@ee.kaist.ac.kr

### Abstract

퍼지 최소 최대 신경망(Fuzzy Min Max Neural Network)은 많은 장점을 가진 분류기로 널리 사용 되고 있다. 그러나 최초의 퍼지 최소 최대 신경망은 몇 가지 단점을 가지고 있으며 그 중에 학습 결과가 학습 파라 미터에 민감한 점을 들 수 있다. 본 논문에서는 퍼지 최소 최대 신경망의 학습에 영향을 주는 학습 파라 미터를 사용하지 않고 강화 학습을 이용하여 신경망을 학습하는 방법을 제안한다. 이 방법에서는 학습 파라 미터 없이 하이퍼 박스의 수와 잘못 분류된 결과에 따라 보답(reward)을 주는 강화 학습을 이용하여 퍼지 최소 최대 신경망을 학습시킨다. 결과로는 학습 데이터에 대해 오분류가 없고 최초의 학습 방법의 결과 보다 작은 하이퍼 박스 수를 갖는 퍼지 최소 최대 신경망이 얻어졌다. 이는 학습 파라미터를 이용한 학습 방법으로 생긴 많은 수의 하이퍼 박스로 인한 일반화 능력의 감소를 막고 하드웨어 구현 시 많은 하이퍼 박스로 인한 어려움을 덜어 줄 수 있다.

### I. 서론

Simpson (1992)이 제안한 새로운 신경망 회로 모델인 퍼지 최소 최대 신경망은 퍼지 집합을 패턴 클래스로 이용하였고 이 퍼지 집합은 패턴 공간에서 최소 최대값으로 정의되는 박스의 형태로 하이퍼 박스

(Hyperbox)로 불린다.[1] 하이퍼 박스는 입력 패턴의 클래스에 대한 소속 정도를 나타내주어 분류가 이루어지고, 신경망의 학습은 이 하이퍼 박스의 영역을 조절하여 이루어 진다. 이러한 특성으로 퍼지 최소 최대 신경망은 간단하고 빠른 학습이 가능하게 되고 하드웨어 구현 역시 간단하게 이루어질 수 있다. 그러나 Simpson의 알고리즘의 큰 두 가지 문제는 학습 데이터의 입력 순서에 따라 학습 결과가 달라질 수 있는 점과 학습 파라미터인 하이퍼 박스의 최대 크기의 결정에 학습 결과가 매우 민감하다는 점이다. 특히 하이퍼 박스의 최대 크기의 제한은 입력 데이터의 분포를 오분류 없이 학습할 수 있는 최소의 크기에 의해 결정된다. 그리고 전체 특성 공간(feature space)에 동일하게 가해 지기 때문에 일반화 능력에 큰 영향을 준다. 이 문제를 해결하기 위한 여러 제안으로는 M. Meneganti의 출력 클래스 수만큼의 하이퍼 박스를 미리 만들고 시작하여 하이퍼 박스의 크기는 크게 하고 수는 적게하는 방법[2]과 Chen Xi의 하이퍼 박스의 크기를 큰 값에서 시작해서 에러가 없는 부분은 유지하고 에러가 생긴 부분에 대해서 하이퍼 박스의 크기 제한을 줄이면서 학습을 하는 방법[3] 등이 있다.

본 논문에서 제안하는 새로운 퍼지 최소 최대 신경망의 학습법은 기존의 하이퍼 박스의 크기의 제한으로 하이퍼 박스의 확장(Expansion)의 여부를 결정하는 방법의 한계를 극복하고자 강화학습(Reinforcement

Learning)을 이용하였다. 새로운 방법을 제안한 동기로는 학습 데이터의 입력이 모두 끝나고 난 후에 학습의 결과를 알 수 있다는 점과 각 학습 데이터의 입력된 순간에 하이퍼 박스를 확장(Expansion) 혹은 생성(Creation)하는 행동(Action)을 한다는 점을 착안해서, 지연된 보답(Delayed Reward)과 상태(State) 그리고 행동(Action)을 이용한 학습 방법인 강화 학습(Reinforcement Learning)을 이용하게 되었다.

본 논문의 구성은 퍼지 최소 최대 신경망의 간단한 소개, 강화 학습의 적용을 위한 문제 형식화, 제안된 방법, 실험 결과 그리고 결론이다.

## II. 퍼지 최소-최대 신경망

FMMNN 은 최소점과 최대점으로 정의되는 하이퍼 박스(hyperbox)로 이루어져 있다. 하이퍼 박스는 정의된 영역 내부에서는 소속 정도를 1로 하이퍼 박스에서 멀어질수록 소속 정도가 줄어드는 소속 함수를 가진 퍼지 집합이다. FMMNN 은 하이퍼 박스를 학습 데이터를 이용하여 각 클래스 별로 만들고 특성 공간에 배치하여 패턴을 분류하게 된다.

학습은 Expansion, overlap test, 그리고 contraction 으로 이루어져 있다. 학습 방법은 새로운 학습 데이터가 입력되면 가장 가까운 같은 클래스의 하이퍼 박스를 찾고, 같은 클래스의 하이퍼 박스가 없거나 일정 거리 이상에 존재하게 되면 새로운 하이퍼 박스를 생성하고 그렇지 않다면 찾아진 하이퍼 박스를 새로운 학습 데이터를 포함하기 위해 Expansion 을 한다. 이때 다른 클래스의 하이퍼 박스와 Overlap test 를 하게 되고 Overlap 이 발생하면 기존의 박스의 크기의 변화를 최소화하는 방향으로 Contraction 을 한다.

Expansion-Contraction 의 학습 과정에서 하이퍼 박스의 크기를 제안하는 학습 파라 미터는 매우 큰 영향을 미치게 된다. 큰 학습 파라 미터는 하이퍼 박스의 크기를 크게 하여 복잡한 분포를 가진 데이터를 학습하기가 힘든 반면 에러를 학습하지 않을 수 있다. 작은 학습 파라 미터는 하이퍼 박스의 크기를 작게 하여 복잡한 분포를 가진 데이터는 학습을 할 수 있으나 필요 이상의 하이퍼 박스를 생성시켜 일반화 능력을 감소시킬 수도 있다.

## III. 강화 학습 적용을 위한 문제 형식화

강화 학습은 상태(State)와 행동(Action) 그리고 보답(Reward)를 이용해서 각 상태 혹은 상태 행동 쌍(State Action Pair)에 보답을 이용한 가치(Value)를 부여하여 최선의 정책(Optimal Policy)을 찾아내는 방법이다.[4] 퍼지 최소 최대 신경망의 학습을 위한 상태(State)와 행동(Action)의 정의는 다음과 같다.

$$S = \{s_{t,i} | s_{t,i} = (B_{t,1}, B_{t,2}, \dots, B_{t,k_{i-1}}, x_i), t_i \in \{1, \dots, 2^{i-2}\}, i \in \{2, 3, \dots, N\}\}$$

여기서 N 은 학습 데이터의 수이고 i 는 입력된 데이터가 몇 번째인가를 나타내며  $t_i$  는 i 번째 학습 데이터가 입력되었을 때 있을 수 있는 상태 중에서 몇 번째 것인가를 나타낸다. 즉 상태는 학습 데이터 1 개가 입력될 때의 그 이전에 존재하는 하이퍼 박스와 입력된 데이터로 결정 된다.

$$A(s_{t,i}) = \{Expand, Create\}, \forall s_{t,i}$$

행동은 위의 식과 같이 확장 혹은 생성 두 가지가 된다. 위의 정의에 따른 상태와 Immediate Reward 는 다음 그림 1 과 같다.

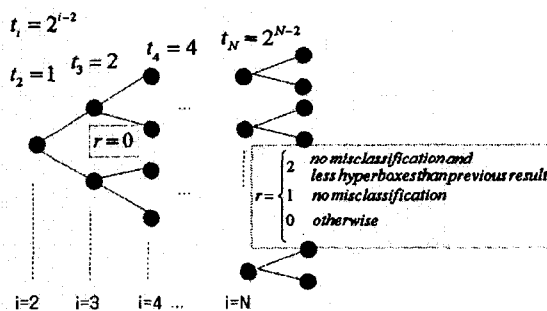


그림 1. 상태 행동 쌍과 Immediate Reward

누적 보답(Accumulated Reward)는 다음과 같다.

$$R_{t,i} = \sum_{k=0}^{\infty} \eta^k r_{t_{i+k}, i+k+1}$$

그림에서 소개된 Immediate Reward 는 학습 데이터의 입력이 끝나고 학습 데이터에 대한 결과가 나오기 전에는 모두 ' 0 ' 이고 학습 결과가 나왔을 때는 학습 데이터에 대해 오분류가 없고 이전 결과의 하이퍼 박스 수보다 적은 박스 수를 가지면 ' 2 ', 오분류만 없으면 ' 1 ' 그 외에는 ' 0 ' 을 주었다. 이는 학습 데이터에

대한 오분류가 없고 하이퍼 박스 수가 적은 학습 방법에 큰 보답을 준다는 것을 의미한다.[4]

형식화된 문제를 풀기 위한 방법을 제안하기 앞서 문제의 특성을 고찰해보면, 확장이 많이 될수록 적은 수의 하이퍼 박스를 갖는 결과가 나올 가능성이 높다는 점과 학습 데이터의 입력이 모두 끝이 나고 평가가 되어 보답을 받는 것까지를 하나의 에피소드로 하고 Monte Carlo 방법으로 최적의 학습 방법을 얻기 위해서는 엄청난 량의 계산 량이 필요하다는 것을 알 수 있다. 그러므로 최적의 답과 계산 시간 사이의 절충이 필요하다.

#### IV. 학습 방법

제안된 방법은 초기에 사용자가 정의한 수만큼 확장에 치우친 정책( $\pi(s, a)$ )으로 에피소드를 만들어 상태 행동 값( $Q(s, a)$ )을 갱신한다. 그 다음에는  $\epsilon$ -greedy 정책으로 계속 에피소드를 만들면서 상태 행동 값을 갱신한다. 결과의 수렴을 위해서는 엄청난 시간이 걸리기 때문에 절충안으로  $\epsilon$  을 줄이면서  $\epsilon$ -greedy 정책으로 에피소드를 만들어 내면서 나온 가장 좋은 즉 하이퍼 박스 수가 가장 적고 오분류가 없는 신경망 학습 결과를 저장하고 일정 횟수가 반복되면 프로그램은 끝난다. 제안된 방법을 정리하면 아래와 같다.

1. 초기화
  - 1)  $\pi(s, a): \pi(s, \text{Expand}) = 0.8, \pi(s, \text{Create}) = 0.2$
  - 2) 모든  $Q(s, a) = 0$
2. 사용자 정의 횟수만큼 반복
  - 1) 초기 조건에 따라 에피소드 생성
  - 2) 에피소드의 모든 상태 행동 쌍에 대해:
    - i) R(누적 보답)을 계산한다.
    - ii) 상태 행동 쌍의 가치를 갱신한다.
 
$$Q(s, a) \leftarrow Q(s, a) + \alpha [R(s, a) - Q(s, a)]$$
3. 종료 조건이 만족할 때까지 반복
  - 1)  $\epsilon$ -greedy 정책을 이용하여 에피소드 생성
    - i)  $\epsilon \leftarrow \epsilon / p$  단  $p=1$  그리고 사용자 정의 횟수 반복 후 1씩 증가
  - 2) 2.2)와 동일
  - 3) 최소의 하이퍼 박스 수를 가지며 오분류가 없는 FMMNN의 하이퍼 박스 수와 FMMNN 자체를 저장

#### V. 실험

##### 5.1 설정

초기 반복 횟수 = 10000 번  
 $\alpha = 0.1$   
 $\epsilon = 0.2$   
 종료 조건: 77500 번 반복

##### 5.2 데이터 셋

데이터 셋은 2 가지를 사용하였다. 첫 번째는 제안된 학습 방법의 유용성을 쉽게 확인 할 수 있도록 2 차원의 데이터를 사용하였고 학습 데이터만으로 구성되어 있다. 두 번째는 실제 분류기의 성능 평가에 많이 사용되는 IRIS 를 사용했다.[5] 성능 평가를 위해 학습 데이터와 테스트 데이터로 나누어 사용되었다.

>> 데이터 셋 1

1. 2 차원 데이터
2. 학습 데이터: 3 클래스, 클래스 당 15 데이터
3. 분포 : 2 개의 클래스는 근접해서 섞여있고 나머지 클래스는 떨어져 있다.

>> 데이터 셋 2: IRIS

1. 4 차원 데이터
2. 학습 데이터: 3 클래스, 클래스 당 15 개의 데이터
3. 테스트 데이터: 3 클래스, 클래스 당 35 개의 데이터
4. 분포 : 데이터 셋 1 과 유사

두 데이터 모두 퍼지 최소 최대 신경망에서 사용하기 위해 0~1 사이 값으로 정규화 하였다. 데이터 셋 2 에서 학습과 테스트 데이터는 임의로 선택하였다.

##### 5.3 실험 방법

데이터 셋 1, 2 를 이용하여 FMMNN 을 4 절에서 제안된 방법으로 학습시키고 테스트한다.

##### 5.4 결과

데이터 셋 1 을 이용한 기존의 학습 방법과 강화 학습을 이용한 학습 결과는 각각 그림 2 와 그림 3 과 같다. 그림에서 다른 색은 다른 클래스를 의미하며, 점은 학습 데이터의 위치를 박스는 하이퍼 박스를 나타낸다.

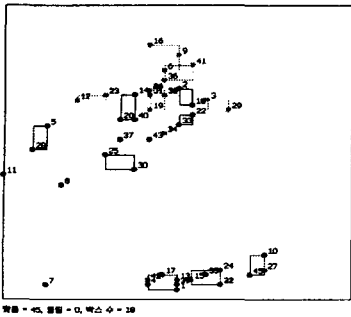


그림 2. 기존 학습 방법의 결과,  $\theta = 0.089$ , 박스 수 = 18

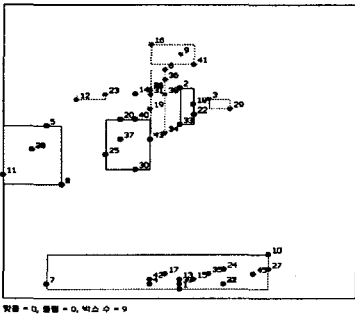


그림 3. 강화 학습을 이용한 학습 방법의 결과, 박스 수 = 9

그림 2, 3 에서 보는 바와 같이 강화 학습을 사용한 결과에서 하이퍼 박스의 수가 18 개에서 9 개로 줄어 들었고 붉은 색으로 표시된 클래스의 경우 하나의 하이퍼 박스 만이 생성되었다.

학습 결과뿐만 아니라 일반화 능력을 검증하기 위해서 IRIS 데이터를 학습 데이터 45 개 테스트 데이터 105 개로 나누어 실험한 결과는 표 1 과 같다.

performance	No. of hyperboxes	Misclassification (total No. of test data)
Method		
FMMNN With $\theta = 0.305$	9	7 (105)
FMMNN with RL	5	3 (105)

표 1. IRIS 데이터에서 두가지 방법의 퍼지 최소 최대 신경망의 학습, 테스트 결과

## 6. 결론

박스의 크기를 제한하여 학습을 하는 기존의 방법

의 한계를 극복하기 위해 강화 학습을 이용한 퍼지 최소 최대 신경망의 학습 방법을 제안하였다. 강화 학습의 적용을 위하여 퍼지 최소 최대 신경망의 학습에서 상태(State)와 행동(Action) 그리고 보답(Reward)을 정의하고 최적의 정책을 찾는 것과 계산 시간의 감소 사이에서 절충안으로 변형된 Monte Carlo 방법을 이용하여 두 가지 데이터 셋에 대해 신경망을 학습시켰다. 데이터 셋 1 에 대한 학습에서 제안된 방법은 기존의 방법의 결과인 18 개의 절반인 9 개의 하이퍼 박스를 갖는 신경망을 얻었다. 제안된 학습 방법의 일반화 성능을 알아보기 위한 IRIS 데이터 셋을 이용한 실험의 결과로는 박스 수 5 개와 오분류 3 개로 기존 방법의 박스 수 9 개와 오분류 7 개 보다 우수한 성능을 보였다. 이는 박스 수의 증가로 인한 일반화 능력(Generalization Ability)의 감소를 막은 결과로 보인다. 제안된 방법으로 복잡한 분포를 갖는 데이터를 분류하기 위한 퍼지 최소 최대 신경망의 학습이 매우 효율적으로 이루어 질 수 있었다.

제안된 방법의 문제점으로는 상태의 수가 학습 데이터가 많아 질수록 매우 증가하게 되고 계산 시간도 매우 길어진다는 점이 있다. 이를 해결하기 위해서 퍼지 최소 최대 신경망의 학습에서 상태의 새로운 정의가 요구된다.

## 7. 참고문헌

- [1] Patrick K. Simpson "Fuzzy Min-Max Neural Networks—Part 1: Classification," IEEE Transaction on Neural Networks, Vol. 3, NO. 5, September 1992.
- [2] Meneganti M., Saviello F.S., Tagliaferri R. "Fuzzy neural networks for classification and detection of anomalies," IEEE Transactions on Neural Networks, Vol. 9 Issue: 5, Page(s): 848 -861 Sep 1998
- [3] Chen Xi, Jin Dongming, Li Zhijian, "Recursive training for multi-resolution fuzzy min-max neural network classifier," Proceedings of 6th International Conference on Solid-State and Integrated-Circuit Technology, Vol. 1 , pp. 131-134, 2001
- [4] Richard S. Sutton, Andrew G. Barto "Reinforce Learning An Introduction," the MIT Press, 1998.
- [5] C. Blake, E. Keogh, C. J. Merz. UCI repository of machine learning databases, 1998  
<http://www.ics.uci.edu/~mllearn/MLRepository.html>