AETLA를 이용한 이진 신경회로망의 최적 합성방법

Optimal Method for Binary Neural Network using AETLA

성상규 · 정종원 · 이준탁
동아대학교 전기공학과

Sang-Kyu Sung, Jong-Won Jung, Joon-Tark Lee
School of Electrical Engineering, Dong-A University
sksungkor@hanmail.net, jtlee@mail.donga.ac.kr

ABSTRACT

In this paper, the learning algorithm called advanced expanded and truncate algorithm(AETLA) is proposed to training multilayer binary neural network to approximate binary to bunary mapping. AETLA used merit of ETL and MTGA learning algorithm. We proposed to new learning algorithm to decrease number of hidden layer. Therefore, learning speed of the proposed AETLA learning algorithm is much faster than other learning algorithm.

Keywords : AETLA, LS, MTGA, ETL, SITV

Ⅰ. 서 론

초창기의 신경회로망은 1950년대 미국의 프랭크 로젠블럿(Frank Rosenblatt)에 의해 처음 소개된 Perceptron은 비교적 정확히 기술된 계산에 의한 최초의 신경회로망 모델이었다. 그에 의해 개발된 단층 퍼셉트론은 이 모델은 패턴인식 연구의 바탕이 되었고, 이 장치들의 기본 요소는 임계논리 유니트(TLU: Threshold Logic Unit)이고, 특정한 생물 기관에 국한되지 않는 일반적인 지능시스템의 기본적인 성질들을 규명하는데 이용되었다. 그러나, 퍼셉트론의 주된 기능적인 제한점은 출력 유니트가 선형 분리 가능한 패턴들만을 분류할 수 있는 한계를 가지고 있다. 1980년대 전방향(feedforward) 3층 네트워크가 제안되었다. 일반적으로, 학습은 다층일 경우 백 프로퍼게이션(Back Propagation)알고리즘을 이용한다. 백 프로퍼게이션(Back Propagation)알고리즘은 다층 퍼셉트론에 관련된 가중치 및 임계값들에 관한 해를 반복적으로 구하는 일반적인 방법으로 신경회로망을 만들 때 몇 개의 뉴런을 사용하여 어떤 연결 구조를 만들 것인지를 결정하고, 목표치를 구하여, 계속해서 반복을 수행하여, 이미 결정된 목표값에 근사시키기 위해 많은 반복(iteration)을 수행한다. 따라서, 많은 반복(iteration)을 수행함으로써, 많은 학습시간을 필요로 하는 단점을 가지고 있다. 최근에는 기하학적인 방법을 이용한 ETL(Expanded and Truncate Learning)알고리즘이 제안되었다. 주어진 훈련입력의 기하학적인 분석에 기반을 둔 ETL은 요구된 separating hyperplane의 함수를 찾아내고 뉴런의 모든 수의 임계값과 모든 수의 가중치를 결정하는 것이다. 이들 hyperplane들은 다른 hyperplane으로부터 원하는 같은 출력을 가진 입력을 분할하는데 있다. 따라서, 두 이웃하는hyperplane들 사이에 위치한 훈련입력은 원하는 같은 출력을 가지는 구조로 구성되어있고, 이진 공간의 기하학적인 구조를 이용하여, 전체 패턴으로부터 한 개씩 학습을 하는 알고리즘으로써, 처음 학습이 끝나고 나서, 합성된 회로망이 변하는 한계를 가지고 있다. 이보다 진보된 학습알고리즘인 MTGA(MSP Term Grouping Algorithm)은 디지털 논리 합성 방법중의 하나인 MSP(Minimal Sum of Product) 형태와 MTGA에서 제안된 필요조건인 Unate의 특징을 이용하여 은닉층의 뉴런의 개수를 최소화하여, 학습시간을 단축시킨다. 그러나, MTGA 학습알고리즘의 필요조건인 Unate의 특성을 만족하지 못하는 패턴은 합성되지 못하고, 하나

의 은닉층을 구성하게 된다. 따라서, 본 논문에서는 ETL과 MTGA 알고리즘의 장점을 이용하여, 새로운 학습알고리즘인 AETLA(Advanced ETL Algorithm)을 제안한다. AETLA은 이진 논리회로의 최소화된 논리회로의 표현방법과 드모르강 법칙을 이용하여, 최적의 합성된 형태를 구성하여, 최적의 학습 알고리즘을 설계하고자 한다.

II. 본 론

2.1 이진 신경회로망의 구조

n-비트 훈련입력벡터의 집합이 주어져 있다고 가정하고, 원하는 출력의 2진 집합이 n-비트 입력벡터는 n-차원의 hypercube의 vertex들처럼 생각한다. 또한, 훈련입력벡터의 두 class들은 net 함수로 표현할 수 있는 (n-1) 차원의 hyperplane에 의해 분리할 수 있다고 가정한다. net (X, T) =

$$w_1x_1 + w_2x_2 + \ldots w_nx_n - T = 0 \quad (1)$$

여기서, w_i's와 T는 상수이다. 이 경우, 훈련입력의 집합은 LS(Linearly Separable) 선형적으로 분리가능하다고 말하고 (n-1) 차원의 hyperplane은 separating hyperplane들 아래의 hard-limiter 활성함수를 가진 n-입력 뉴런에 의해 성립될 수 있다

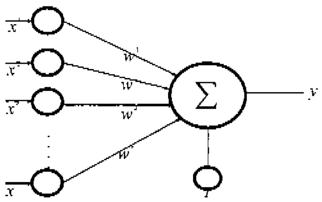

그림 1. 이진 신경회로망의 구조

$$y=1 \ : \ \sum_{i=1}^{n} w_i x_i - T \geq 0$$
$$y=0 \quad \text{otherwise} \quad (2)$$

여기서, y: 뉴런의 출력

w_i: i번째 입력과 뉴런사이의 연결 가중치

x_i: i번째 입력

T: 뉴런의 임계치

2.2 가중치와 임계치의 결정

최소화된 그룹의 각 단위 뉴런에 대해서 가중치와 임계치를 결정하기 위해서 RHS(Reference HyperSphere)방법을 이용한다. RHS의 식은

$$(x_1-\frac{1}{2})^2+(x_2-\frac{1}{2})^2+\ldots + (x_n-\frac{1}{2})^2=\frac{n}{4} \quad (3)$$

훈련 vertex들로부터 선형적으로 분리가능하다면, 그곳에는 hyperplane이 존재해야만 한다. Hypersphere를 찾기 위해서 hypersphere의 중심이 SITV(Set Included True Vertex)내 성분들이 중심이 위치하고 있다는 것을 고려한다. 만약, 이 hypersphere가 분리된다면, 이것은 최소의 반지름이 되고, 이와 반대로 SITV내에 모든 성분을 포함시키기 위해서 긴 반경을 가지고 있다면, hypersphere이 선택될 것이고, 아래와 같이 separating hypersphere의 식으로 표현할 수 있다.

$$(x_1-\frac{c_1}{c_0})^2+(x_2-\frac{c_2}{c_0})^2+\ldots + (x_n-\frac{c_n}{c_0})^2=r^2 \quad (4)$$

이 separating hypersphere가 RHS와 교차할 때 (n-1)차원의 hyperplane을 찾을 수 있다. (3)로부터 (4)를 빼고, c_0를 곱하면, separating hyperplane은

$$(2c_1-c_0)x_1+(2c_2-c_0)x_2+\ldots + (2c_n-c_0)x_n-T=0 \quad (5)$$

즉, separating hyperplane가 존재한다면, SITV내의 true vertex v_t에 대해

$$\sum_{i=1}^{n}(2c_i-c_0)v_t^i - T \geq 0 \quad (6)$$

그리고, 나머지 vertex들로부터 다른 training vertex v_r에 대해

$$\sum_{i=1}^{n}(2c_i-c_0)v_r^i - T < 0 \quad (7)$$

따라서, SITV내의 true vertex v_t와 training vertex v_r 은

$$\sum_{i=1}^{n}(2c_i-c_0)v_t^i > \sum_{i=1}^{n}(2c_i-c_0)v_r^i \quad (8)$$

여기서,

t_{min} : SITV내의 모든 vertex들 사이에

$\sum_{i=1}^{n}(2c_i - c_0)v_i^i$ 의 최소값.

f_{max} : 나머지 vertex 사이에 $\sum_{i=1}^{n}(2c_i - c_0)v_r^i$ 의 최대값.

만약, $t_{min} > f_{max}$ 라면, separating hyperplane은

$$(2c_1 - c_0)x_1 + (2c_2 - c_0)x_2 + \dots + (2c_n - c_0)x_n - T = 0 \quad (9)$$

이 되고, separating hyperplane은 존재한다.

여기서, $T = [\frac{t_{min} + f_{max}}{2}]$ 와 $[x]$는 x와 크거나 같은 수이고, 만약, $t_{min} \le f_{max}$라면, separating hyperplane은 존재하지 않는다. 예를 들어, 입력변수의 함수 $f(x_1, x_2, x_3)$인 6-훈련입력을 고려한다 고려한다. 만약, 입력이 {000, 010, 011, 111}일 때 $f(x_1, x_2, x_3)$는 1이고, 입력들이 {001, 100}일 때 $f(x_1, x_2, x_3)$는 0이고, 입력 vertex들이 {101, 110}일 때, $f(x_1, x_2, x_3)$가 "don't care"인 경우를 가정한다. SITV는 {000, 010, 011}에 대해 t_{min}=Minimum $[-3x_1 + x_2 - x_3]$이고, 따라서 t_{min}=0, f_{max}=-1이다. $t_{min} > f_{max}$와 T=0이기 때문에 hyperplane $-3x_1 + x_2 - x_3 = 0$는 나머지 vertex로부터 SITV {000, 010, 011}내에 vertex가 분리된다. 더 많은 true vertex를 포함하기 위해서 다른 true vertex가 처음처럼 같이 사용함으로서 선택되고, 만약 새로운 trial vertex가 SITV에 부가할 수 있는지 테스트하고, 이 과정은 더 이상의 true vertex와 SITV내에 부가되지 않을 때까지 계속 반복한다. 따라서, 알고리즘은 두 separating hyperplane으로 바뀌게 되고, 즉, 은닉층내의 요구된 두 뉴런을 말한다. 두 번째 요구된 hyperplane는

$$(2C_1 - C_0)x_1 + (2C_2 - C_0)x_2 + \dots + (2C_n - C_0)x_n = T \quad (10)$$

여기서,

$C_0 = 5$, $C_1 = 1$, $C_2 = 2$, $C_3 = 2$로 계산되어지고, $(2C_1 - C_0)$가 가중치로 결정된다. 즉, $w = (2C_1 - C_0)$가 된다.

따라서, $-3x_1 - x_2 - x_3 - T = 0$이 되고, $t_{min} = -3$과 $f_{max} = -5$.이고, $t_{min} > f_{max}$이고, T=-4이므로, 요구된 hyperplane은 $-3x_1 - x_2 - x_3 + 4 = 0$로 결정된다. 예제에 대한 3층 신경회로망의 구조와 은닉층내의 가중치와 임계치를 그림.2와 표.1에 나타내었다.

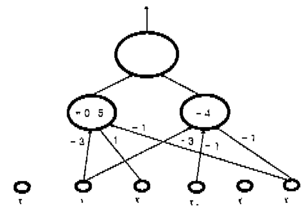

그림. 2. 예제에 대한 3층 신경회로망의 구조

표 1. 은닉층내의 가중치와 임계치

Neuron	Input			
	W1	W2	W3	T
1	-3	1	-1	-0.5
2	-3	-1	-1	-4

2.3 EPS(Expanded Product of Sum)와 이진 신경 회로망의 합성

적은 개수의 그룹으로 전체함수를 표현하기 위해서는 한 개의 그룹에 최대한 많은 수의 항을 포함시켜야 한다. 본 논문에서는 이진 신경회로망을 합성하기 위해서 입력된 학습패턴으로부터, EPS(Expanded Product of Sum)의 형태를 이용하였다. EPS는 boolean 표현방식으로써, 예를 들면, $x + y$의 출력을 얻기 위해서 $(x+y)y$의 함수에 $(x+y)\cdot 1$을 확장함으로써, 최소화된 그룹을 구성할 수 있다.

2.4 실험결과

여기에서는 본 논문에서 제시한 학습알고리즘과 기존의 학습알고리즘을 비교하기 위해 실제적인 실험을 통하여 검증하고 있다.

2.4.1 양자화된 원의 분할

6bit으로 양자화된 2차원 원 문제에 대하여 2차원의 각 축을 3 bit를 사용하여, 각각을 양자화하면 전체 공간이 64개의 사각형으로 구분되며, 가운데 중앙의 12개에 대해서는 '1'의 기대값을 할당하고, 나머지 52개에 대해서는 '0'의 기대값을 할당했을 때, 이를 구분하는 이진 신경회로망을 합성하는 문제에 적용하였다. 기존의 학습알고리즘인 ETL 학습알고리즘은 그림. 3과 같이 1개의 출력층과 5개의 은닉

층이 필요하였다.

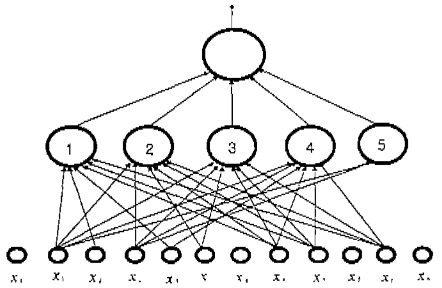

그림. 3. ETL 학습알고리즘의 3층 신경회로망의 구조

표 2. 은닉층내의 가중치와 임계치

Neuron	Input						
	W1	W2	W3	W4	W5	W6	T
1	-3	3	1	3	3	1	7
2	-29	-3	-1	-3	3	1	-5
3	-29	-3	-1	-3	3	1	-26
4	-19	-13	1	-3	3	1	-20
5	-16	-16	0	0	0	0	-24

MTGA 학습알고리즘은 그림. 5와 같이 1개의 출력층과 4개의 은닉층이 필요하였다.

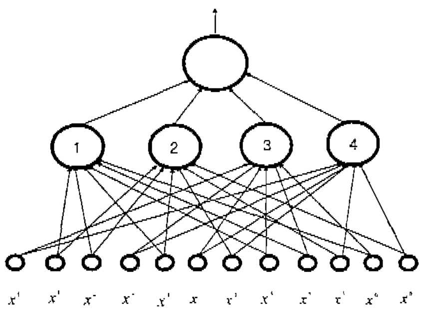

그림. 4. MTGA 학습알고리즘의 3층 신경회로망의 구조

표 3. 은닉층내의 가중치와 임계치

Neuron	Input						
	W1	W2	W3	W4	W5	W6	T
1	-3	3	1	-3	3	-1	6.5
2	-3	3	1	3	-3	1	5.5
3	3	-3	1	-3	3	1	5.5
4	3	-3	-1	3	-3	1	4.5

본 논문에서 제안한 AETLA 학습알고리즘은 입력패턴으로부터, 함수를 확장시키고, 그림. 6과 같이 1개의 출력층과 4개의 은닉층으로 구성된 회로망이 합성되었다.

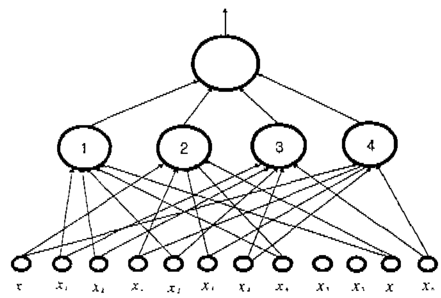

그림. 5. AETLA 학습알고리즘의 3층 신경회로망의 구조

표 4. 은닉층내의 가중치와 임계치

Neuron	Input					
	W1	W2	W3	W4	W5	T
1	-3	3	1	-3	-1	4.5
2	3	-3	-1	-3	1	3.5
3	-3	3	1	3	1	5.5
4	3	-3	-1	3	1	4.5

III. 결 론

본 논문에서 제시한 AETLA 학습알고리즘은 ETL 학습알고리즘의 확장의 원리를 이용하였다. 비교적 우수한 MTGA 학습알고리즘에 비해 MSP의 형태와 Unate의 특성을 만족하지않은 형태의 함수일지라도, 확장의 원리를 이용함으로써, 최소의 은닉층을 구성할 수 있었고, 각 은닉층에 연결되는 노드의 수도 감소함으로써, 학습시간을 단축시킬 수 있었다. 앞으로, 제안한 학습알고리즘을 이용하여, 실제 플랜트에 적용시키고자 한다.

IV. 참고문헌

[1] P. L. Bartlett and T. Downs, *Using random weight to train multilayer network of hard-limiting units*, IEEE Trans. Neural Networks, pp202-210, 1992.

[2] M. L. Brady, R. Rayhavan, and J. Slawny, *Back propagation fails to separate where perceptrons succeed*, IEEE Trans. Circuits Systems., pp.665-674, May 1989.

[3] S. Park, J. H. Kim, and H. Chang, *A learning algorithm for discrete multilayer perceptron*, in Proc. Int. Sysp. Cricuit Systems, Singapore, Jun.1991.

[4] J. H. Kim, S. K. Park, Han, H. Oh, and M. S. Han, *The geometrical learning of binary neural network*, IEEE Trans. Neural Networks, vol. 6. no. 1. pp.237-247, Jan. 1995.