

개념 학습에 의한 신경 회로망 컴퓨터

이 기한^o, 황 회음, 서울대학교 컴퓨터 공학과

김 춘석, 숭실대학교 전기공학과

A Neural Network for Concept Learning : Recognitron

Lee Ki Han^o, Whang Hee Yoong, Seoul Nat'l Univ. Dept. of Comp. Eng.

Kim Choon Suk, Soong Sil Univ. Dept. of Elec. Eng.

< Abstract >

Concept is the set of selected neurons in a stable state of a neural network. The Recognitron uses a parallel feedback structure to support concept learning. A number of clusters can exist in response to a given input, each of which make up a selective neuron.

There are supervised and unsupervised learning methods in concept learning. In this paper, we have chosen unsupervised learning. Also, a new concept called relaxational learning has been introduced to stop runaway weights

1. 서론

신경회로망 컴퓨터를 통한 패턴 분류나 인식에 관한 방법이 최근에 많이 대두되었다. 패턴인식이나 분류는 주어진 입력패턴을 원하는 패턴으로 인식하거나 분류(clustering)하는 것이다. 이때 원하는 결과를 도출해내기 위해서 신경회로망 컴퓨터 외부에서 미리 정보를 지정하면, 이를 관리자 학습(supervised learning)에 의한 학습방식이라고 하고, 신경회로망 컴퓨터 자체에 의해서, 즉 외부정보가 없이 원하는 결과를 얻으면 관리자 없는 학습(unsupervised learning)에 의한 학습방식이라 한다.

개념에 의한 학습이란, 시간 t에서의 input을

$r_1(t), \dots, r_n(t)$ 라고 하고 이를 $r(t) = [r_1(t), \dots, r_n(t)]$ 로 정의하자. 그리고, t+1 시간에서의 $r_1(t+1), \dots, r_n(t+1)$ 은 아래와 같이 정의된다.

$$r(t+1) = [w * r(t) -] - (1)$$

이때, w는 weight를 나타내며, 는 threshold이다.

$$r(t+1) = T * r(t) \text{ --- (2)}$$

라고 다시 쓸 수가 있다. 여기서, T는 형태변환 함수라고 한다.

이 식 2가 안정된 상태, 즉 $r = T * r$ 이 될 때까지 학습을 시키면 입력 패턴에 대한 적당한 패턴의 안정된 형태 변환에 의한 학습이 이루어진다. 이 때 선택 뉴uron (selective neuron)은

입력패턴들에 대한 적당한 규합(cluster)의 결과를 보여주는 것이다. 이러한 학습방식을 개념학습이라고 부른다. 현재까지 이러한 개념화(conceptualization), 추상화(abstraction), 그리고 범주화(catagorization)에 의한 학습방식은 관리자 학습방식이 대부분이었다. (1)

1972년 Kohonen은 개념을 연관(association)이라고 보고, correlation matrix memory로 구현을 했다(1). 그 외에도 Ameri(2), Grossberg등이 구현을 했다(3, 4, 5).

본 논문에서 제안한 Recognitron은 관리자가 없는 학습방식(unsupervised learning), 즉 외부로부터의 정보가 없이 자신이 개념학습을 하는 신경회로망 컴퓨터이다.

2. Recognitron의 구조

Recognitron은 관리자가 없는 학습방식에 의한 개념학습을 하므로 입력 시그날에 의해서 자신이 규합과 분류를 한다. Recognitron의 구조를 살펴보면 그림 1과 같다.

여기서 $r(1), \dots, r(n)$ 은 외부에서 주는 입력시그날로서 0과 1 값을 갖는다. 이를 $r = [r(1), r(2), \dots, r(n)]$ 이라고 정의하자. $f(1), \dots, f(n)$ 은 결과시그날로서 이도 역시 0과 1 값을 갖는다. 또한 $f = [f(1), f(2), \dots, f(n)]$ 이라고 정의하자.

Feed forward signal w_{ij} 는 입력 r 과 출력 f 를 완전히 연결(fully connected)하고, f 를 활성화(active) 또는 비활성화(inactive) 시키는 역할을 한다. 이때 feed forward weight 값은 0에서 1 사이의 무작위 추출에 의한 값이다.

Feed backward 시그날인 fb_{ji} 는 출력 f 와 입력 r 을 완전히 연결하고, 학습된 후에 feed forward weight 값을 진정시키는 역할을 하고, 그 값은 0 과 1 이다. 자세한 것은 3장 Recognitron의 학습방식에서 설명하겠다. 많은 입력패턴들이 Recognitron에 외부로부터 들어오면, 이들을 적당한 규합(cluster)으로 분리할 해낸다. 그림 2는 이를 보이고 있다.

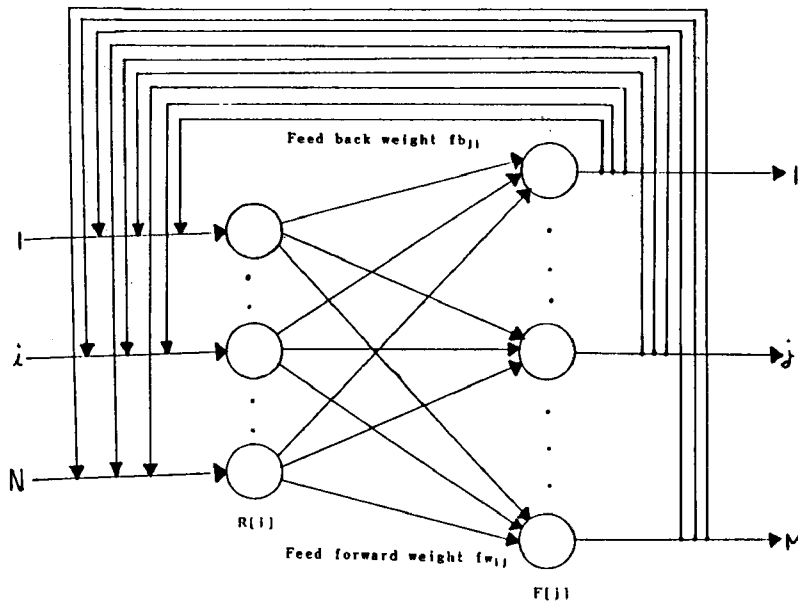


그림 1. Recognitron의 구조

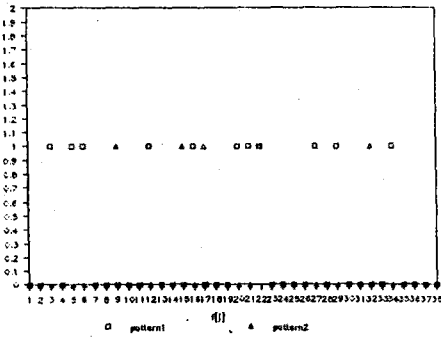


그림 2 입력에 의한 규합의 예

3. Recognitron의 학습방식

Recognitron의 학습방식은 두 개로 나누어 지는데 첫째는, 입력패턴들을 적당히 규합하는 개념 학습이고, 둘째는, 규합이 끝난후에 feed forward weight들을 조정시키는 weight 진정학습으로 나누어진다.

3.1 개념학습 방식

Recognitron은 관리자가 없는 개념학습방식을 취한다고 위에서 언급했다. 0과 1의 조합으로 구성된 입력과 무작위로 추출된 0과 1 사이의 feed forward weight fw_{ij} 를 가지고 아래와 같은 학습 함수로써 학습을 시킨다.

$$fw_{ij}' = fw_{ij} + l_f * fw_{ij} * r_i \quad (3)$$

여기서 fw_{ij}' 는 변화된 weight값이고 l_f 는 학습비로서 학습시 fw_{ij} 값을 어느정도 비율로서 증가시킬 것인가를 외부에서 지정해준다. 이 l_f (학습비)는 얼마나 빨리 개념화 하느냐를 결정하는데 아주 중요한 역할을 한다. Recognitron의 각 뉴런(neuron)들이 firing하기 위한 변환함수는 nonlinear 함수의 하나인 sigmoid형태를 취하고 있다. 그림 3은 sigmoid 변환함수를 보인

것이다. 개념화 학습을 끝내는 조건은 적어도 하나의 $f(i)$ 가 나오는 것이다. 따라서 식 3에 의해서 결과 f 가 나오면 개념화 학습은 끝난것이다.

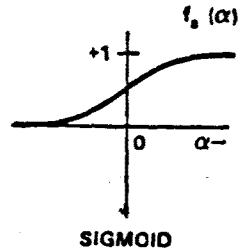


그림 3. Sigmoid 변환 함수

3.2 Feed forward weight 진정학습

Feed forward weight 학습에 의하면 $r(i) = 1$ 인것의 feed forward weight fw_{ij} 은 모두 증가하고, $r(i) = 0$ 인 것의 fw_{ij} 는 변화가 없게된다. 따라서, $r(i) = 1$ 인 것에 연결된 $f(j)$ 가 활성화(active)된다면 fw_{ij} 는 개념학습이 이루어진 것인데 $f(j) = 0$ 인 것에 연결된 fw_{ij} 는 불필요한 변화다. 따라서 fb_{ji} 이라는 feed backward weight를 두어서 이들 fw_{ij} 의 weight를 진정시켜야 하겠다. fb_{ji} 는 아래와 같이 정의된다.

$$f(j) = 0 \text{ and } r[i] = 1 \text{ 면 } fb_{ji} = 1$$

$$f(j) = 1 \text{ 면 } fb_{ji} = 0 \quad (4)$$

식 4가 의미하는 것은 $f(j)$ 의 결과가 0이고, $r(i)$ 가 1인 노드 사이의 fw_{ij} 는 진정을 시켜야 할 필요가 있고, $f(j)$ 가 1 이면 $r(i) = 1$ 이건 0 이건 상관없이 그대로 둔다. 즉, 개념학습이된 fw_{ij} 값인 것이다. 식 5는 이를 나타낸다.

$$fw_{ij} = fw_{ij} + l_f * fw_{ij} * r_i * fb_{ji} \quad (5)$$

이러한 학습을 weight 진정학습이라고 부른다.

여기서 l_f 는 개념화 학습시 사용했던 값을 그대로 쓴다.

4. 실험 및 결과

Recognitron의 개념학습은 식 (3)과 (5)에서 보듯이 학습비 l_f 와 학습횟수에 상관관계가 있다.

4.1 고정된 학습횟수에 의한 개념학습

학습횟수를 500으로 고정시키고 개념학습을 시작했다. 그림 4는 패턴 1과 패턴 2의 규합을 보인 것이다.

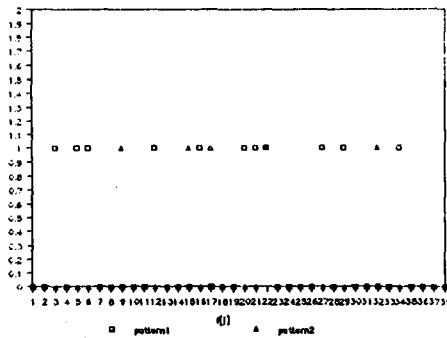


그림 4. 패턴 1과 패턴 2의 규합

그림 5는 학습비의 증가에 의한 규합의 갯수의 변화를 보인 것이다. 패턴 1은 학습비가 0.5인 경우 규합이 29개인데, 0.005이하에서는 11개의 규합으로 수렴하고, 패턴 2는 학습비가 0.5인 경우 규합이 26개이고, 0.005이하에서는 5개의 규합으로 수렴함을 볼 수 있다.

표 1은 패턴 1과 2에 관한 학습비 증가에 따른 규합을 보인 것이다. 이를 살펴보면 패턴 1과 2의 경우 학습비가 0.005가 적당하다고 본다.

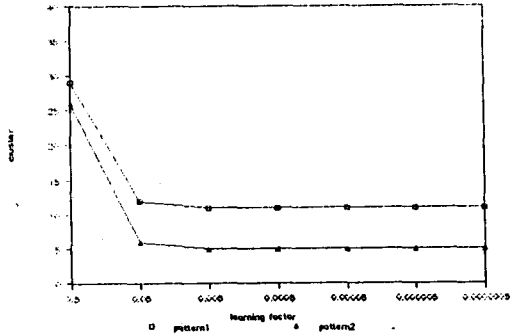


그림 5. 고정된 학습횟수하에서 학습비 변화에 의한 개념학습

표 1. 고정된 학습횟수하에서 학습비 변화에 의한 개념학습

rate of learning	pattern1	pattern2
0.5	29	26
0.05	12	6
0.005	11	5
0.0005	11	5
0.00005	11	5
0.000005	11	5
0.0000005	11	5

4.2 고정된 학습비에 의한 개념학습

4.1에서 살펴본 바에 의하면 학습비는 0.005가 적당하다. 그림 6은 학습횟수의 증가에 의한 규합을 보인 것이다.

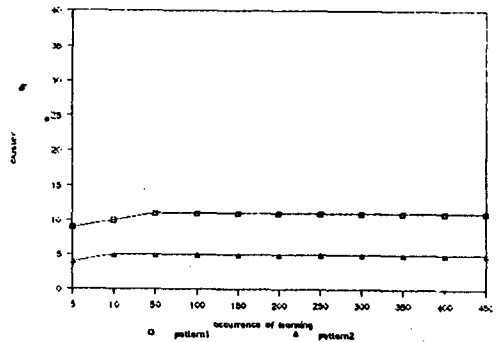


그림 6. 고정된 학습비하에서 학습횟수 변화에 의한 개념학습

여기서 살펴보면 패턴 1의 경우 학습횟수가 50회 이상이면 패턴의 규합이 수렴되는 것을 볼 수 있고, 패턴 2의 경우는 학습 횟수가 10회 이상이면 패턴의 규합이 수렴하는 것을 볼 수 있다. 표 2는 이 관계를 보인 것이다.

표 2. 고정된 학습비하에서 학습횟수 변화에 의한 개념학습

rate of learning	pattern1	pattern2
5	9	4
10	10	5
50	11	5
100	11	5
150	11	5
200	11	5
250	11	5
300	11	5
350	11	5
400	11	5
450	11	5

따라서 학습횟수는 패턴 1의 경우 50회, 패턴 2의 경우 10회 이면 개념학습이 이루어진다고 본다.

5. 결론

본 논문에서 제안한 Recontron은 판리자가 없는 학습을 하는 신경회로망 컴퓨터이다. 그리고 특이한 것은 개념학습을 하고나서 필요이상으로 흥분된 weight값을 진정시키는 진정학습이 따로 존재한다는 점이고, Reconitron은 parallel feedback 시그날을 사용한다는 점이다. Reconitron이 개념을 학습하는데에는 학습비를 얼마로 하느냐와 학습횟수를 얼마로 하느냐가 서로 연관된 문제이고, 이들의 값을 선택하는 것이 중요한 결정 요소이다.

앞으로는 이들을 multi-layer로 연결해서 개념학습을 하는 보다 좋은 Reconitron을 만들 예정이다. 보다 좋은 Reconitron이 갖추어야 할 조건은 위치, 크기, 모양에 무관한 패턴을 분류 또는 규합해야 한다. 이들이 필요한 이유

는 개념학습이 완전히 되기위한 필요조건이기 때문이다.

6. 참고문헌

1. T. Kohonen, "Correlation matrix memories," IEEE Trans. on Comp. 1972. 1-21, 353- 359

2. S. I. Amani, "Neural Theory of Association and Concept Translation," Bio. cyb 1977, 26, 175 - 185

3. Douglas L. Reilly, L. N. Cooper, Charles Elbaum, "A Neural Model for Catagory Learning," Bio. cyb. 1982.

4. J. A. Anderson, et al, "Distinctive features, catagorical perception, and probability learning some application of a neural model," Psych. Review, 1977, 84, 413 - 451

5. J. A. Feldman and G. E. Hinton, Chapter 8 : Catagorization and Selection Neurons by Parallel Models of Associative Memory.