

뉴럴네트워크의 개요와 적용



임 춘 성*

1. 서 론

1980년대 자연과학 분야의 관심과 초점이 집중된 학문적 진보 중에서 뉴럴네트워크(neural network)를 빼 수는 없다. 사실 뉴럴네트워크는 1950년대 이미 현재 인공지능(artificial intelligence)이라 불리우는 학문 분야와 ‘인간의 두뇌활동의 이해를 위한 전산모형화’라는 같은 목적을 위한 서로 다른 접근방법으로 출발하였고 상당한 이론적인 업적을 성취하였다. 그러나, 전형적인 인공지능이 그 접근방법·지식과 정보처리과정의 심볼(symbol)화의 이해와 이론전개의 용이성으로 인해 순항을 거듭한 반면에, 뉴럴네트워크는 인간신경망의 병렬적 패턴인식과 학습의 수학적 모델화라는 난제에 임하여 더딘 행보를 취하게 되었다. 1980년대에 뉴럴네트워크가 다시 부각이 되었던 이유로 크게 3 가지를 들수 있다. 첫째, 흔히 심볼리즘(symbolism)이라 부르는 인공지능적 접근방법의 한계가 그 이론과 응용성 모든 측면에서 강렬히 부각되어 콘NECTIONISM(connectionism)이

라 부르는 뉴럴네트워크 대안을 찾게 되었다. 둘째, 고도화된 전산기기가 범용화되고 뉴럴네트워크 전용회로가 개발, 사용되면서 신속하고 대용량의 처리가 요구되는 뉴럴네트워크의 적용성이 현실화되었다. 세번째이자 특히 중요한 이유로서, 풀리지 않았던 병렬적 패턴인식과 학습의 수학적 모델화에 대한 몇가지 난제가 백프로파게이션(Backpropagation)과 같은 알고리듬의 개발로 해결되었다는 것이다.

1980년대가 이러한 뉴럴네트워크의 이론적 재도약으로 특성화할 수 있다면 1990년대는 여러 분야의 제반문제에 뉴럴네트워크가 적용되어 기존의 다른 문제해결방법을 대체하거나 보완하는 적용확산 기간이라 말할 수 있다. 이는 ‘인간의 두뇌활동의 이해’라는 본래적인 뉴럴네트워크의 심리/생리학적 목적과는 관계 없이 제반 문제의 해결에 뉴럴네트워크의 방법론적 강점을 활용하는 공학적 접근방안(engineering approach)의 확산을 의미하며, 본 논문의 내용 또한 이러한 측면에서 기술되고 있다. 다음은 뉴럴네트워크의 정의와 특

* 연세대학교 산업시스템공학과, 교수

성을 약술하고 계속하여 뉴럴네트워크의 종류와 뉴럴네트워크 적용시의 지침을 정리해 본다.

2. 뉴럴네트워크의 정의와 특성

2.1 뉴럴네트워크의 정의

뉴럴네트워크를 간략하게 정의하면, ‘다량의 처리요소(processing element 또는 node)들이 복잡하게 상호 연결되어 있는 병렬적, 분산적 정보처리 구조’라 할 수 있다(Hecht-Nielsen, 1990). 이는 다분히 뉴럴네트워크의 구조적 정의이며 부연 설명하면 다음과 같다. (그림 1 참조)

- (i) 대부분의 뉴럴네트워크는 다수의 처리요소로 구성된 몇 개의 층(layer)으로 구성되어 있으며,
- (ii) 각 처리요소는 외부 또는 다른 처리요소로부터 다수의 입력치를 받으며,
- (iii) 처리요소간의 연결선(connection)에는 할당된 가중치(weight)가 있어,
- (iv) 각 처리요소는 입력치와 입력이 들어오는 연결선의 가중치를 특정한 수학적 함수를 통하여 변환하여 자신의 출력치로 산정하게 되며,
- (v) 이를 다른 처리요소의 입력치 또는 전체구조의 출력으로 사용하게 된다.
- (vi) 연결선에 할당된 가중치들은 일정한 수학적 알고리듬(학습식)을 사용하여 점진적으로 변하게 된다.

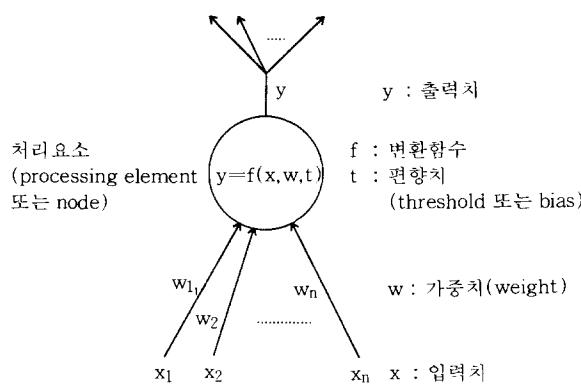


그림 1 뉴럴네트워크의 처리요소

이러한 구조를 통하여 뉴럴네트워크가 수행하는 기능은, 바람직한 입력치와 출력치의 관계를 설정하는 근사적인 함수 또는 맵핑(approximated implementation of a bounded mapping or function)이다. 뉴럴네트워크에서의 이러한 함수의 추정은 입력과 출력에 대한 사전 지식이나 통계분포적인 가정에 근거하지 않고 단순히 주어진 관계 예(example)들, 다시 말해서 주어진 입력치와 상응하는 출력치(또는 입력치 만)로 구성된 예들을 써 뉴럴네트워크를 학습시키는 것을 의미한다. 따라서, 통계학적인 관점으로 뉴럴네트워크는 회귀분석(regression)과 유사한 기능을 수행하고 있으나, 회귀분석이 관계함수의 기본구조를 가정하고 파라메터(parameter)를 추정하는 데에 반하여, 뉴럴네트워크는 기존의 전제가 전혀 없는 추정 방법이다. 요컨대, 뉴럴네트워크는 준비된 관계 예들을 이용하여 함수학습식이라 불리우는 알고리듬을 통하여 입력과 출력의 관계를 학습추정하는 기능을 지니고 있다.

2.2 뉴럴네트워크의 특성

뉴럴네트워크의 적용에 대한 특성은 다음의 4가지로 요약될 수 있다.

1) 학습(learning)

앞서 언급한 바와 같이 뉴럴네트워크는 관계함수 설정을 위하여 확률분포와 함수기본구조(선형, 2차식 등)와 같은 정량적인 값이나 논리적인 지식에 대한 가정이 필요하지 않다. 단순히 관계의 예가 확보되어 있을 경우에는 학습식을 이용한 학습과정을 통하여 원하는 관계함수를 근사할 수 있다.

2) 추상화(abstraction)

관계함수에 대한 추정이 학습을 통하여 뉴럴네트워크의 병렬구조(parallel structure)에 분산 추상화(distributed abstraction)되어 학습된 지식이 경마한 손상이나 오류에도 크게 침해받지 않고 보존된다(graceful degradation).

3) 일반화(generalization)

학습을 통하여 추상화된 관계함수의 지식은 처리요소간의 연결선에 할당된 가중치값의 형태로

저장되는데, 이는 관계 예의 핵심적인 내용을 간추린 것으로 관계 예에 포함되지 않은 새로운 입력에 대하여 저장된 관계함수의 맵핑을 통하여 올바른 출력을 산출하게 된다. 특히, 뉴럴네트워크의 입력치를 구성하는 입력벡터(input vector)의 일부분의 값이 부정확하더라도 올바른 결과를 도출하는 것이 기대되어진다(fault-tolerance).

4) 속도(speed)

뉴럴네트워크에서는 주어진 입력에 대한 출력의 산출시간이 저장된 관계 함수의 복잡도에 대하여 사실상 독립적이라 볼 수 있다. 이는 뉴럴네트워크의 병렬구조에 기인하는 것으로 최근 뉴로칩(neuro chip)의 개발과 확산 보급으로 실시간 응용(real-time application)분야에 더욱 활용되고 있다.

3. 뉴럴네트워크의 종류

뉴럴네트워크의 종류는 주로 뉴럴네트워크의 학습학습식에 의하여 구분된다. 학습학습식에 따라 그 대략의 구조와 기능이 결정되므로 흔히 학습식의 이름으로 해당 뉴럴네트워크를 명명하게 되는데, 특정 뉴럴네트워크들을 포괄하는 분류로 크게 관리형(supervised), 비관리형(unsupervised), 그리고 혼합형(hybrid) 뉴럴네트워크의 3가지가 있다. 관리형의 경우는 하나의 관계 예가 입력치와 이에 상응하는 출력치의 쌍으로 구성되어 있고, 비관리형은 입력치만으로 구성된다. 혼합형은 복합적인 형태로 학습되는 것을 지칭한다. 본

논문에서는 여러 뉴럴네트워크중에서도 특히 그 적용빈도가 높은 관리형의 일종인 백프로퍼게이션(Backpropagation)과 비관리형의 기본 형태가 되는 경쟁학습(Competitive Learning)에 대하여 간략하게 그 알고리듬과 기능에 대하여 설명하기로 한다. 그외 뉴럴네트워크는 일반적인 뉴럴네트워크의 문헌(예, Hertz et al., 1991)을 참조하기 바란다. 표 1은 두 개의 뉴럴네트워크를 개요적으로 비교하고 있다.

3.1 백프로퍼게이션(Backpropagation)

백프로퍼게이션(Rumelhart et al, 1986) 알고리듬의 개발은 뉴럴네트워크의 진흥에 촉진 역할을 했다고 해도 과언이 아니다. 뉴럴네트워크 분야에서 오랜 기간 동안 풀지 못했던 비선형 분류(nonlinear classification)의 문제를 다층(multi layer) 네트워크구조에 비선형계획법(nonlinear programming)의 일종인 그데리언트 방법(gradi-ent method)을 적용하여 해결하였다. 백프로퍼게이션은 관리형 뉴럴네트워크의 일종으로 네트워크에 입력치가 입력되면 입력치가 가중치와 처리요소 별로 규정된 변환 함수 그리고 편향치와 연산되어 계속 상위층으로 그 과정이 진행된다. 이리하여 최상위층 처리요소의 출력치와 미리 준비된 ‘정답’ 출력치와 비교하게 되고, 그 차이를 최소화하기 위하여 국부적인 그레디언트 방법을 적용하여 다시 역방향 층별로 연결선의 가중치를 변환하게 된다. 그림2와 표2에는 백프로퍼게이션의

표 1 관리형 백프로퍼게이션(Backpropagation)과 비관리형 경쟁학습(Competitive Learning)의 비교(Leem, 1996a)

	Backpropagation	Competitive Learning
classification	(the most popular) supervised learning	(the basic principle of) unsupervised learning
a training sample	input and correct output	input only
network structure	multiple layers	single layer in general
function	error minimization through gradient method	clustering for vector quantization
transformation in nodes	nonlinear sigmoid	linear sum
learning speed	slow	fast

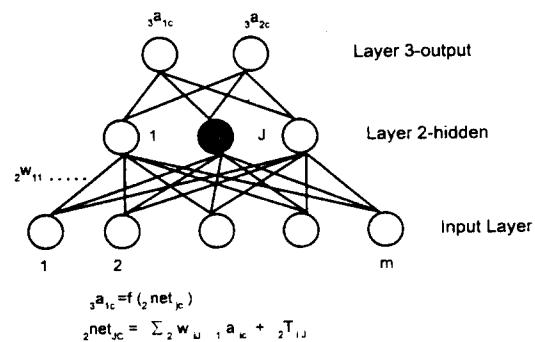


그림 2 백프로퍼게이션이 수행되는 3층구조의 뉴럴네트워크

표 2 백프로퍼게이션 알고리듬

- Initialize weights and thresholds to small random values [usually, weights are uniformly distributed between (-0.3, 0.3) and threshold between(0.015, 0.015)].
- Present input /output pairs for the learning set(C pairs). Input vector I may be continuos-valued.
- Specify desired outputs($d(I), \dots, d(n)$) for each case.
- Caculate actual outputs for each case c:
 $a_{kc} = f_k(\text{net}_{kc}) = 1[1 + \exp - (\text{net}_{kc})]$
where
 $a_k = \text{activation}(\text{output}) \text{ of node } k \text{ at layer } k \text{ for case } c,$
 $\text{net}_{kc} = \text{net input to node } j \text{ at layer } k \text{ for case } c$
 $= \sum_k (W_{ij})(k-a_{kc}) + kT_j$
- W_{ij} = weights on connection from unit i at layer k-1 to unit j at k
and
 T_j = negative threshold of unit i at layer k.
At output layer, let $Y^c(j) = a_{kc}$
- Adapt weights after all cases c=1,..,C are computed, starting at output layer and using
 $\Delta_k W_{ij} = \mu \sum_k (D_{kc})(k-a_{kc})$
and
 $\Delta_k T_j = \mu \sum_k (D_{kc})$
where μ is a learning rate. If node j is an output node. Then
 $\Delta D_{jc} = -\partial E / \partial (k \text{net}_{kc}) = Y^c(j)(1 - Y^c(j))(d^c(j) - Y^c(j))$
Otherwise,
 $\Delta_{k-1} D_{ic} = (k-1-a_{kc})(1-k-a_{kc}) \sum_k (k D_{kc})(k W_{ji})$
where i is all nodes on layer k.

알고리듬이 정리되어 있다.

백프로퍼게이션은 대부분의 경우 안정된 결과를 주고 다중 구조의 특성상 임의의 복잡한 분류 경계(classification 또는 decision boundary)를 생성(그림 3 참조)한다는 강점을 지니고 있어 그 활용성이 매우 높다. 반면에, 학습 속도가 느린 편이며 국부 최적화(local optima) 결과를 산출할 위험이 있는 것이 단점으로 지적되고 있다.

3.2 경쟁 학습(Competitive Learning)

경쟁 학습(Rumelhart and Zipser, 1985)은 비관리 뉴럴네트워크의 기본 형태이며 주어진 입력치만으로 학습을 하게 되는데 이는 입력치들의 분포에 대한 유목화(categorization)을 의미하며 통계학에서의 클러스터링(clustering)을 수행하는 것이다. 경쟁 학습을 포함한 비관리형 뉴럴네트워크는 대부분 입력층과 출력층으로 구성된 2개층구조를 지니고 있는데 출력층의 처리요소별로 정의된 가중치 벡터(weight vector)들이 입력분포의 집단들의 중앙으로 점진 이동하여 그 집단들을 대표하여 유목화를 학습하게 된다(그림 4 참조)

경쟁 학습과 다른 비관리 뉴럴네트워크는 학습

STRUCTURE	TYPES OF DECISION REGIONS	EXCLUSIVE OR PROBLEM	CLASSES WITH MIXED REGIONS	MOST GENERAL REGION SHAPES
SINGLE-LAYER	HALF PLANE BOUNDED BY HYPERPLANE			
TWO-LAYER	CONVEX OPEN OR CLOSED REGIONS			
THREE-LAYER	ARBITRARY CONCAVE LISHED BY NUMBER OF HIDDEN			

그림 3 백프로퍼게이션의 기능(Lippman, 1987)

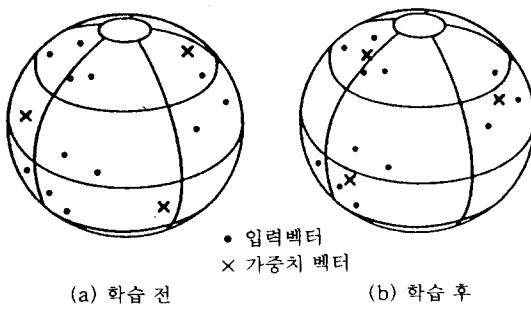


그림 4 경쟁학습의 기능

의 형태가 명시적이며, 관계 예의 입력별로 출력 정답을 일일이 구할 필요가 없다는 강점을 지니고 있다. 또한 학습속도가 빠르며 입력치의 구성 요인(입력벡터의 사이즈)이 상당한 수에 이르는 경우에 관리형 뉴럴네트워크에 비해 상대적인 적용상의 강점을 갖고 있는 것을 평가받고 있다. 그럼 5와 표 3은 경쟁학습의 알고리즘을 정리한 것이다.

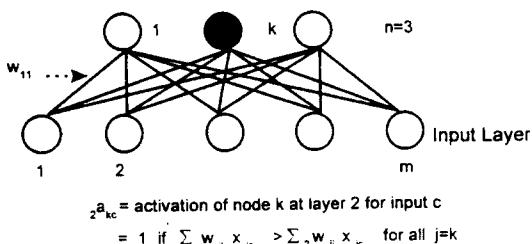


그림 5 경쟁학습이 수행되는 뉴럴네트워크

표 3 경쟁학습 알고리듬

1. Initialize weight vectors for the k output nodes to either small random values or small uniform values.
 2. Present input vector, x . Input x is multiplied by the weights on the connections from input nodes to each output node. Thus, the neural network finds
 $x^T w(j)$
 for each output node, J , which is its input. A competition ensures to determine the output node with the largest net input. Denote the winning node, J^*
 3. Update the weights vector for J^* by the following learning rule. Where α is a learning rate and $0 < \alpha < 1$:
 $w'(J^*) = w(J^*) + \alpha(x - w(J^*))$
 4. Present next input vector

4. 뉴럴네트워크 적용 지침

문제의 해결방법으로 뉴럴네트워크 적용을 고려할 때 몇 가지 대답해야 하는 질문이 있다. ‘언제 또 왜 뉴럴네트워크를 적용해야 하는가?’ ‘어떤 종류의 뉴럴네트워크를 적용해야 하는가?’ ‘입력치와 출력치는 어떠한 방식과 형태로 준비되어야 하는가?’ ‘의미있는 결과를 위한 관계 예의 숫자는?’ ‘어떻게 결과를 평가할 것인가?’와 같은 질문을 신중히 대답해 볼 필요가 있다. 뉴럴네트워크가 이미 확산 적용기에 있어 그 장점과 적용시 주의해야 할 점에 대한 구체적 보고들이 적지 않으므로, 다른 대안으로서의 방법론과의 신중한 비교가 선행되지 않은 무작정 적용은 의미가 없다고 할 수 있다. 본 논문에서는 위의 질문사항의 일부에 해당하는 뉴럴네트워크의 적용이 우선적으로 고려되는 문제의 특성을 나열해 본다(더욱 자세한 사항은(Leem, 1996b) 참조).

- i) 입력과 출력의 상관관계는 예견되나 추정에 도움이 되는 확률분포, 식의 기본구조, 법칙 등이 존재하지 않아 그 관계식 설정이 어려운 경우
 - ii) 입출력관계가 비선형이 예견되는 경우
 - iii) 입력치가 불완전하거나 부정확한 경우
 - iv) 입력치로 사용되는 벡터가 상당한 수의 요인으로 구성된 경우
 - v) 입출력 또는 입력만으로 구성된 관계 예가 상당수 존재하거나 용이하게 도출할 수 있는 경우
 - vi) 실시간 문제 해결이 선호되는 경우
 - vii) 문제에 대한 가용한 정보가 정성적보다는 정량적일 경우
 - viii) 전통적인 문제해결방법론(예; 통계분석, 패턴인식, 최적화기법, 인공지능 등)이 성공적이지 못한 경우

여타 다른 방법론도 그려하듯이 뉴럴네트워크 또한 그것의 의미있는 적용과 결과를 위해서는 뉴럴네트워크의 장점과 단점 그리고 해당 문제의 성격을 고려한 적용시의 주의할 점을 제대로 이해하는 것이 중요하다. 뉴럴네트워크를 ‘만병통치약’

으로 과대 해석하는 것도 피해야 하겠지만, 올바른 기능의 이해와 효과적인 적용 방안에 대한 고려가 선행되지 않은 채로 시도를 포기하는 것은 더욱 배제해야 하겠다.

참고문현

1. Hecht-Nelson, R.(1990) *Neurocomputing*, Addison-Wesley.
2. Hertz, J., Krogh, A. and Palmer, R.G.(1991), *Introduction to the Theory of Neural Computation*, Addison-Wesley.
3. Leem, C.S.(1996a) "An Automated Fault Diagnosis for Manufacturing Process Monitoring and Control", *International Journal of Modelling*

- and *Simulation*, 16, 4, in press.
4. Leem, C.S.(1996b) "Operational Guide on the Application of Neural Networks to Intelligent Manufacturing Systems", *Proceedings of the '96 Computers and Industrial Engineering Conference*, in press.
5. Lippman, R.P.(1987) "An Introduction to Computing with Neural Nets", *IEEE ASSP Magazine*, 4, 2, 4-22.
6. Rumelhart, D. and Zipser, D.(1985), "Feature Discovery by Competitive Learning", *Cognitive Science*, 9, 75-112.
7. Rumelhart, D., Hinton, G. and Williams, R. (1986), "Learning Representations by Backpropagating Errors", *Nature*, 323, 533-536. 

● 한국전산구조공학회 광고게재 안내 ●

한국전산구조공학회에서는 학회지 '전산구조공학'을 정기적으로 연 4회 발행하고 있습니다. 저희 학회지에 광고게재를 하시기 원하시는 분을 위하여 다음과 같이 안내합니다.

1. 학회지는 매년 3, 6, 9, 12월에 발행하며 학회지에는 전산구조 공학 분야에 관련되는 건설분야, 전산기기, 계측기기, 소프트웨어, 전산해석 및 설계용역, 전산자료 및 서적 등의 광고를 게재해 드리고 있습니다.
2. 광고게재료(1면 1회당)
 - 1) 전산구조공학회지 : 표 2(앞표지 내면) 1,000,000원
표 3(뒤표지 내면) 1,000,000원
표 4(뒤표지) 1,500,000원
간지 700,000원
내지(본문 여백 : 행사안내 및 도서광고) 200,000원
 - 2) 연 4회 이상 게재할 경우 우대내용(특별회원사가 아닌 경우)
연 4회 이상 광고를 게재할 경우는 상기 광고게재료의 20% 할인혜택을 드립니다.
 - 3) 특별회원사 우대내용
본 학회의 특별회원사가 게재하는 광고는 상기 광고게재료의 10%를, 연간 게재 광고는 30%를 할인해 드립니다.
 - 4) 기타사항
본 학회 사무국(전화 : (02)565-0035~6)으로 연락을 주시면 자세한 안내말씀을 드립니다.