

# 실험계획법을 이용한 다층 퍼셉트론 인공 신경망의 구조 설계

이 석 호\*, 강 대 천\*, 이 찬\*, 강 무 진\*\*

\* 성균관대학교 대학원

\*\* 성균관대학교

## 1. 서 론

경험과 학습을 바탕으로 새로운 상황에 대처하는 인간의 신경계에서의 신경세포들의 상호작용을 규명하는 일은 많은 과학자들을 매료시켜 왔다. 이와 함께, 생물학적인 신경계를 닮은 인공적인 신경망을 구축하여 감지하고, 인식하고, 구별하고, 판단하는 일에 이용하고자 하는 노력도 끊임없이 진행되어 왔다. 인공신경망은 생물학적 신경계에서 처럼 뉴우런이라는 정보처리 요소들로 이루어지고, 각 뉴우런은 활성화함수를 이용하여 입력신호를 출력신호로 변환한다. 뉴우런 간의 신호전달은 연결 링크를 통하여 일어나고, 링크마다의 가중치에 의해 전달강도가 결정된다. 즉, 인공신경망은 뉴우런의 개수와 연결상태로 표현되는 신경망 구조(Architecture)와 뉴우런의 활성화함수(Activation Function), 그리고 연결링크의 가중치를 구하는 방법(학습 알고리즘)에 의해 결정된다고 할 수 있다.

공학적인 견지에서 인공신경망은 분류기(Classifier)의 한 방법으로서, 경험적인 실험을 통해 아무리 복잡한 분류 모델이라도 구현시킬 수 있는 것이 장점이다. 오랫동안 실용화되지 못하던 인공신경망이 근래에 들어 공학적, 산업적으로 많이 쓰이기 시작하게 된 것은, 매우 단순하고 명료한 학습(Learning 또는 Training)방법이 개발된 덕분이다. 다른 많은 인공지능의 기법과 비교할 때, 인공신경망은 복잡한 추론 법칙과 많은 양의 Knowledge Base가 없어도 단순한 학습법에 의해 경험을 축적할 수 있고, 다량의 정보를 고도의 병렬, 분산형으로 처리할 수 있으며, 입력자료에 의해 생긴 잡음, 변형, 파손 등에 강하다는 등의 장점을 갖는다. 반면, 시스템이 커짐에 따라 학습시간이 기하급수적으로 길어지고, 완성된 인공신경망을 분류기로서 사용할 때의 예측오차를 추정하기가 어려우며, 최적 신경망 구성을 위한 정형화된

방법이 없고, 학습의 연속성이 없는 것 등이 한계점으로 되어 있다. 이외에도 학습결과 도달한 해(Solution)가 최소 오차를 갖는 최적(Global Optimum)인 지 확인할 방법이 없다는 것도 문제점이 될 수 있으나, 설사 부분적 해(Local Optimum)를 얻었다 하더라도 일정한 오차 범위 내에서 분류기능을 수행할 수 있다면 실제적인 이용상의 문제는 없다고 볼 수 있다.

인공신경망의 구성은 전적으로 많은 경우의 수에 대한 테스트에 의존하게 되는 데, 이 경우 시간도 대단히 많이 소요될 뿐더러 체계적으로 가장 좋아 보이는 신경망 모델에 도달하였는 지도 확실치 않다. 따라서, 본 논문에서는 실험계획법을 다층 퍼셉트론 신경망의 구조설계에 적용하여 적은 실험횟수로 적합한 신경망 모델을 구성할 수 있음을 검증하고, 실제 사례에 적용하여 그 유용성을 제시하고자 한다.

## 2. 인공 신경망의 공학적 이용

전술한 바와 같이, 인공신경망은 분류기로서의 적용상 단순성과 범용성 때문에 의료진단이나 항공기와 선박의 항로 계획, 음성인식 등 일반적인 응용 외에, 공학적으로도 그 응용분야가 날로 증가하고 있다. 신호처리나 영상인식을 이용한 제조라인에서의 부품 인식, PCB(Printed Circuit Board)나 기계제품의 조립성 검사 등은 전형적인 생산공학적 이용사례이며, 동작기계에서의 최적 공구경로 결정, 최적 질삭 조건 선정 등의 적용에도 알려져 있다. 또한, 로봇의 제어나 AGV(Automated Guided Vehicle)의 위치 추정 및 경로 계획, 기계시스템의 결함 진단, 그리고 자동차 설비의 적응제어등 진단 및 제어 분야에의 이용이 급증하고 있다. 한편, 영업에서의 수요 예측이나 생산계획, 품질관리, 설비의 작업 감시, 그리고

공수예측과 공정관리 등 생산공학적인 활용도 점차 확산되고 있다.

인공신경망은 그 구조와 학습이론에 따라 수없이 많은 모델이 개발되었지만, Backpropagation을 이용하는 다층 퍼셉트론 신경망과 뉴우런의 횡적 연결 구조를 갖는 단층 Hopfield신경망, 그리고 Simulated Annealing으로 Global Minimum을 보장하는 Boltzmann Machine과 새로운 학습시 과거의 학습내용이 지워지지 않고 연속성이 유지되는 ART(Adaptive Resonance Theory)신경망 등이 실용적으로 많이 쓰이는 편이다. 이들은 각각 원리상의 강점을 가지고 있지만, 너무 복잡하거나 기억용량 또는 계산시간, 실시간성, 그리고 일반성 결여(예: Binary신호만 처리) 등의 문제로 아직까지는 Backpropagating Multilayered Perceptron Model이 가장 널리 쓰이고 있다.

### 3. 다층 Perceptron 인공신경망 구성에 관한 고찰

다층 퍼셉트론 신경망은 입력층(Input layer)과 하나 이상의 은닉층(Hidden layer), 그리고 출력층(Output layer)으로 구성된다. 입력층과 출력층은 입력 유니트와 출력 유니트를 가지고 있는데, 이들은 각각 분류를 위한 파라메타와 알고자 하는 출력값들을 대표하고 있고, 은닉층이라 불리는 중간층은 각 유니트의 입출력 특성을 비선형으로 함으로써 신경망의 능력을 향상시키기 위한 것이다. Fig.1은 한 개의 은닉층을 갖는 3층 퍼셉트론 신경망의 구조를 보여 준다. 다층 퍼셉트론의

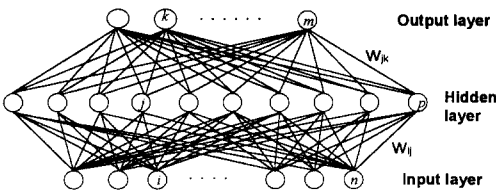


Fig.1 The typical architecture of the back-propagating perceptron

학습은 흔히 오류 역전파 알고리즘(Error Backpropagation Algorithm)을 이용하는데, 입력 유니트에 입력 신호가 들어오면 이것은 중간층의 각 뉴우런으로 전달되고, 이 신호는 다시 출력 뉴우런으로 전달된다(Feed

Forward). 출력 뉴우런의 출력값과 목표값(Target)을 비교하여 그 차(Error)를 감소시키는 방향으로 연결강도( $w_{ik}$ ,  $w_{ij}$ )를 조정해 나아가는 것(Backpropagation)이 학습의 과정이다.<sup>[1]</sup> 오류 역전파 학습 알고리즘은 여러 문헌에 자세히 기술되어 있으므로 추가적인 설명을 생략하고, 여기에서는 Backpropagating Multilayered Perceptron Network 구성에 관한 고려사항에 관하여 고찰하기로 한다.

### 은닉층과 은닉 뉴우런의 수

현재까지 연구된 바에 의하면, 최적 은닉층 수와 은닉뉴우런 수의 결정에 관한 이론이 아직까지 완전하게 정립되어 있지 않고, 경험에 의해 또는 실험적으로 결정하고 있는 실정이다. 일반적으로 은닉 뉴우런이 부족하면 학습이 잘 되지 않아 잘못된 결과를 보여줄 수 있으며, 반대로 너무 많은 은닉 뉴우런은 Overfitting이나 학습시간의 장기화 등 좋지 않은 결과를 가져온다. 은닉층의 수에 관하여는, 은닉 뉴우런의 숫자가 일정할 때 두 층으로 뉴우런을 분산시키는 것이 단층 일때보다 학습이 빠르다고 알려져 있지만, 두 개의 은닉 층이 단일 은닉층을 갖는 신경망보다 분류 기능이 우수하다는 보장은 없다.<sup>[2]</sup> 또한, 하나의 은닉 층을 갖는 퍼셉트론 구조로 모든 종류의 분류 문제를 모델링 할 수 있고, 대부분의 경우에 충분히 좋은 결과를 줄 수 있다고 알려져 있다. 이 경우, 가장 좋은 결과를 내는 은닉 뉴우런의 수가 존재하는데, Kolmogorov의 정리에 의하면 그 값은 입력 유니트의 수가  $n$ 일 때  $(2n+1)$ 이내에 존재한다.<sup>[3]</sup>

### 학습과 수렴 문제

인공신경망의 학습은 출력 오차를 최소로 하는 연결가중치를 구하는 과정이다. 그러나, 수학적으로 완벽하게 오차가 0(zero)이 될 때까지 학습을 계속할 수는 없으므로, Error값이 설정치  $E_{conv}$ 보다 작아지면 수렴했다고 보고 학습을 마치게 된다. 그러나, Fig.2에서 볼 수

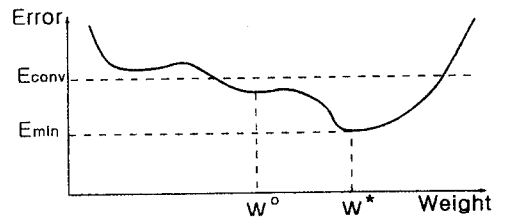


Fig.2 Error surface diagram

있는 바와 같이, 수렴조건을 만족하여도 진정한 최적해  $W^*$ 가 아닌 유사 최적해(Local Minimum)  $W^o$ 에 도달할 가능성도 있다. 또한, 경우에 따라서는 도달 가능한  $E_{\min}$  값이  $E_{\text{conv}}$ 보다 커서 결코 수렴될 수 없는 상황이 발생할 수도 있다. 1회의 학습마다 가중치를 조정하는 정도를 결정하는 변수로 학습율이 있는 데, 너무 작게 잡으면 학습시간이 너무 길어지고, 너무 크면 오차의 최소값에 이르지 못하고 진동하여 수렴하지 않을 위험이 있다. 이 위험을 줄이면서 학습속도를 빠르게 하고 Local Minimum에 빠질 수 있는 경향도 줄이는 방법으로 Momentum계수를 사용할 수 있다. 이 때의 학습효과는 다음과 같이 표시된다.

$$W_{ij}(t+1) = W_{ij}(t) + \Delta W_{ij}(t) + \alpha \Delta W_{ij}(t-1)$$

( $\alpha$  는 Momentum Parameter)

인공신경망의 학습이 수렴되지 않을 때에는 은닉층 구성을 바꾸어 보거나 가중치 초기값을 바꾸어 보는 것이 좋은 방법이다. 일반적으로 다른 초기값들로 시작한 학습은 각기 다른 해(Solution)에 도달하는 데, 어느 경우라도 오차의 범위 내에 들어오기만 하면 좋은 것은, 설사 해가 Global Minimum 아닌 Local Minimum이라도 관계없는 것과 마찬가지이다.<sup>[4]</sup>

#### 학습 데이터와 예측오차

신경망의 학습을 위하여는 Target Pattern을 갖는 학습 데이터가 필요한데, 이 데이터 샘플은 신경망으로 분류하고자 하는 모든 패턴을 포괄할 수 있어야 한다. 만족할 만한 오차로 학습이 수렴되었다 해도, 신경망이 분류기로서 이용될 때 어느 정도의 오차의 범위에서 기능을 발휘할 지 그 예측오차를 추정하는 것은 매우 어렵고, 어쩌면 학습오차를 줄이거나 수렴의 문제보다 더 중요한 일일 것이다. 예측 오차 추정의 한 방법으로 Leaving-One-Out법이 있는데, Sample Data를  $n$ 등분하여 그 중 하나의 Set을 테스트 데이터로 하고 나머지  $(n-1)$ 개의 Set을 학습데이터로 이용하는 것이다. 이때 테스트 데이터 Set을 바꾸어 가며 학습과 실험을 하면  $n$ 회의 학습오차 및 예측오차를 구하게 된다. 이 값들의 평균값을 예측오차로 삼으면, 실험 횟수가 많아지는 부담은 있지만 학습완료된 인공신경망의 예측성능에 관한 좋은 척도가 될 수 있다.<sup>[5]</sup>

#### 학습횟수 (Epoch)

전체 학습 데이터를 1회 학습시키는 것을 Epoch이라 하는데, 일반적으로 학습이 진행되며 Epoch수가 증가하

면 학습 오차는 줄어들게 된다. 그러나, 예측오차는 Epoch이 증가한다고 하여 계속 작아지지는 않고, 어느 단계를 넘으면 다시 커지는 과학습(Overtraining)현상이 발생하게 된다(Fig.3). 따라서, 학습 오차를 작게 함과 동시에 예측오차가 최소가 되는 Epoch에서 학습을 종료하는 것이 중요하다.

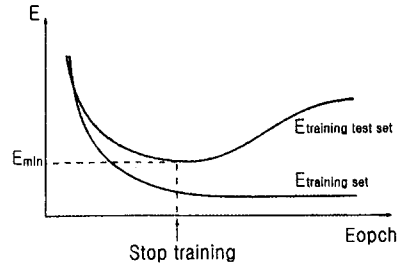


Fig.3 Trends of training and test error

#### 4. 실험계획법에 의한 신경망 설계

앞에서 고찰한 바와 같이, 인공신경망을 구성하고 학습시키는 데 있어서, 은닉층 수와 은닉 뉴우런 수, 학습율, 수렴조건, 연결 강도의 초기값, 학습횟수 등 많은 변수를 고려해야 하고, 이들의 조합으로 발생하는 모든 경우의 수를 테스트해보고 최적 신경망을 얻는 것은 실제적으로 불가능하다 할 수 있다. 그러나, 실험의 횟수를 최소로 하여 최대의 정보를 얻음으로써 최적의 Configuration을 찾는 방법으로 실험계획법(DoE: Design of Experiments)이 알려져 있는데 여기에서는 직교배열 표에 의한 실험계획법을 이용하기로 한다.

##### 직교배열표에 의한 실험계획법

실험결과에 영향을 주는 요인들이 직교성(Orthogonal)을 갖도록 실험을 계획하여 Data를 구하면, 주어진 실험횟수로 제어 인자(Control Factor)의 신뢰성을 잘 검토할 수 있다. 이러한 원리를 이용하여 실험계획을 짜도록 고안된 표를 직교배열표(Orthogonal Array)라 한다. 즉, Factor들 간의 직교성을 이용하여 만들어 놓은 표를 말한다. 직교배열표에는 2, 3, 4, 5수준 및 혼합수준계등이 있으나, 제어인자가 계량치여서 현재의 조건보다 큰 쪽이 좋은가 작은 쪽이 좋은가가 명확하지 않을 때에는 3수준계를 사용하면 충분하다. 3수준계는  $L_3 = (3^{(3^m-1)/2})$ 로 나타내는데, 여기서  $m$ 은 2이상의 정수로서  $3^m$ 은 실험의 크기,  $(3^m - 1)/2$ 는 직교배열표의 열의 수를 의미

한다.  $m=2$ 인 경우,  $L_9(3^4)$  직교배열표는 Table 1과 같고, 표 중의 숫자는 수준을 표시한다.<sup>[6]</sup>

Expt.No	Column			
	1	2	3	4
1	1	1	1	1
2	1	2	2	2
3	1	3	3	3
4	2	1	2	3
5	2	2	3	1
6	2	3	1	2
7	3	1	3	2
8	3	2	1	3
9	3	3	2	1

Table 1  $L_9(3^4)$  Orthogonal array

본 연구에서는, 하나의 은닉층을 갖는 다층 퍼셉트론 신경망 구성에 있어서의 주요 제어 인자인 은닉뉴런 수, 학습율, 학습횟수를 직교배열표에 의한 실험계획법으로 찾음으로써, 인공신경망의 최적 설계를 시도하였다.

### 5. 사례연구 및 결과

선박의 건조에 있어서, 설계가 완료되어 상세 사양이 결정되기 전에는 작업에 소요되는 공수를 계산 할 수 없다. 그러나, 과거에 건조해 본 경험치를 바탕으로 신경망을 이용하여 분류기를 구성하면 몇 몇 주요 입력 변수만으로도 작업공수를 예측할 수 있다. 본 연구에서는 선박건조 작업 중, 배의 조립에 해당하는 선각공수를 추정하는 일에 실험계획법을 이용한 인공 신경망 설계기법을 적용해 보았다. 입력 유니트는 선종, 강재중량, 선주계수이고, 출력 유니트는 가공, 소조립, 패널 조립, 대조립, D/H(Dock House) 내업 조립, 본선 조립, D/H 외업 조립, 선각계등 여덟 개이며, 그 구성은 Fig.4와 같다. 최적의 될 관측치인 특성치를 다음과 같이 정의하고,

$$Error_{char} = \frac{|총실투입공수 - 총예측공수|}{총실투입공수} \times 100$$

제어인자로써는 중간층의 은닉 뉴런 수, 학습율, 학습횟수로 선정하고, 제어인자가 계량치 이므로 3수준계의 직교배열표를 사용하였다. Sampling 방법에서  $n$ 을 5로 하여 전체 Sample Data Set인 26Pattern중 5개를 Test Set으로 하고 나머지 21개의 Pattern을 Training Set으로 하여 반복실험을 하였다.

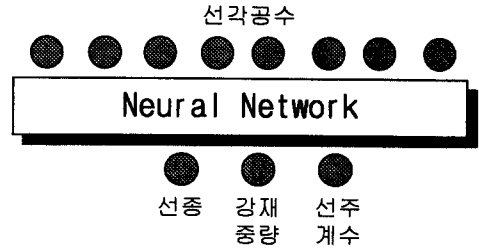


Fig.4 Composition of the neural network

### 결과

제어인자 및 수준표를 Table2와 같이 구성하고, 실험 Matrix를 3수준 직교배열표( $L_9(3^4)$ )을 이용하여 Table3과 같이 작성하였다. 9가지 실험조합(Experiment1은  $[A_1, B_1, C_1]$ , Experiment2는  $[A_1, B_2, C_2]$ 등)에 따라 실험을 실행하여 특성치  $Error_{char}$ 를 구하였다. 각 실험조합의 특성치는 Leaving-one-out에 따라 5번 수행한 특성치의 평균을 구하여 대표값으로 하였다.

FACTOR \ LEVEL	LEVELS		
	1	2	3
A) Hidden Neuron	3	5	7
B) Learning Rate	0.01	0.02	0.1
C) Epoch	10000	20000	30000

Table 2 Control factor for the neural network ( Experiment 1 )

Expt.No	Column				특성치
	1(A)	2(B)	3(e)	4(C)	
1	1	1	1	1	$E_1$
2	1	2	2	2	$E_2$
3	1	3	3	3	$E_3$
4	2	1	2	3	$E_4$
5	2	2	3	1	$E_5$
6	2	3	1	2	$E_6$
7	3	1	3	2	$E_7$
8	3	2	1	3	$E_8$
9	3	3	2	1	$E_9$

Table 3 Orthogonal array for experiment

각 인자들의 효과정도를 표시하기 위해 동일 수준들의 특성치를 합하여 그림으로 표시하면 Fig.5와 같이 되어

제어인자들의 영향추이를 알 수 있었다. 관측하고자 하는 특성치는 오차이므로 Fig.5에서 최소값이 되는 수준을 각 제어인자의 최적값으로 볼 수 있다. 즉, 첫 번째 실험에서는 은닉 뉴우런이 3개, 학습율이 0.02, 학습횟수가 10000근처에서 제어인자의 최적값이 있음을 알 수 있었다. 더 정확한 최적제어인자 값을 구하기 위하여, 제어인자의 수준을 미세하게 분할하여 2차 실험을 하였고, Table4는 이때의 수준치를 보여준다. 2차 실험에 대한 직교배열표는 Table3과 같고, 실험 결과는 Fig.6에 나타내었다.

FACTOR \ LEVEL	LEVELS		
	1	2	3
A) Hidden Neuron	4	6	12
B) Learning Rate	0.015	0.03	0.05
C) Epoch	8000	13000	22000

Table4 Control factor for the neural network ( Experiment 2 )

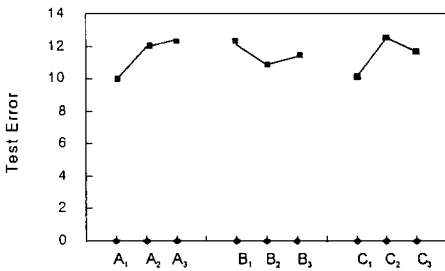


Fig.5 Plots of control factor effects ( Experiment 1 )

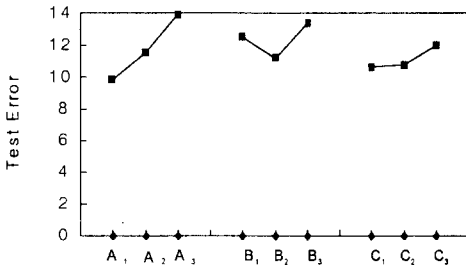


Fig.6 Plots of control factor effects ( Experiment 2 )

이 결과로부터 가장 좋은 Network성능은 은닉 뉴우런이 4개, 학습율이 0.02, 학습횟수가 10000일 때임을 알 수 있었다. 이 신경망을 이용한 선각공수의 예측 결과는 Table5와 같다.

출력 유닛	학습 오차(%)	예측오차(%)
가공	1.40	2.90
소조립	2.50	2.50
패널 조립	3.17	4.10
대조립	1.80	3.50
D/H 내업 조립	11.70	25.9
본선 조립	2.40	4.00
D/H 외업 조립	10.0	19.4
선각계	0.80	5.00

Table5 RMS training and test error rate for the final neural network

### 결론

다층 퍼셉트론 인공신경망 구성시 은닉층과 은닉 뉴우런의 수, 학습율, 학습횟수, 연결가중치의 초기값 등 많은 변수를 고려해야 하는데, 이들의 조합으로 발생하는 경우의 수는 너무 많아 계수적(Enumerative)으로 실험을 하여 최적 모델을 찾는 것은 현실적으로 불가능한 형편이다. 본 연구에서는 거의 모든 종류의 분류문제를 모델링할 수 있는 3층 퍼셉트론 신경망 구성의 최적화에 은닉 뉴우런 수, 학습율, 학습횟수를 제어 인자로 하는 실험계획법을 적용하여 짧은 시간에 좋은 결과를 주는 신경망을 얻을 수 있었다. 개발된 신경망을 선각 건조의 조립 공수 예측에 적용하여 매우 좋은 적응율을 성취하였다. 두층 이상의 은닉층을 갖는 퍼셉트론 모델이나 다른 구조를 갖는 신경망에도 실험계획법을 적용하면 좋은 결과를 얻을 수 있을 것으로 전망된다.

### [참고문헌]

1. Jacek M. Zurada, "Artificial Neural System" West Publishing Company Press, 1992
2. S. Y. Kung & Yu Henn Hu, " A Frobenius Approximation Reduction Method(FARM) for Determining Optimal Number of Hidden Units", International Joint Conference on Neural Networks, VOL. II, 1991
3. Laurene Fausett, " Fundamentals of Neural Networks" Prentice Hall Press, 1994
4. Robert Hecht Nielsen, "Neurocomputing," Addison-Wesley Publishing Company Press, 1990
5. Sholom M. Weiss & Casimir A. Kulikowski, " Computer Systems That Learn", Morgan Kaufmann Publishers, Inc
6. 박성현, "현대실험계획법", 민영사