「정보시스템연구」 제9권 제1호
한국정보시스템학회
2000년 6월, pp . 27-44

인공 신경망의 학습에 있어 가중치 변화방법과 은닉층의 노드수가 예측정확성에 미치는 영향

김 진 백* · 김 유 일**

<목 차>

Ⅰ. 서론
Ⅱ. 인공 신경망의 실험모형
2.1 인공 신경망의 지식획득과정
2.2 인공 신경망 실험모형의 구성
Ⅲ. 가중치 변화방법 및 노드수에 따른 인공 신경망의 예측률 변화
3.1 실험설계
3.2 실험결과
3.2.1 epoch size에 따른 예측률변화
3.2.2 노드수에 따른 예측률 변화
3.2.3 다변량 판별모형과의 비교
Ⅳ. 결론 및 향후 연구
참고문헌
Abstract

Ⅰ. 서 론

인공 신경망(artificial neural network)에 관한 이론적 발전과 더불어 H/W의 성능 향상으로 인해 여러 경영문제들을 해결하기 위해 인공 신경망의 구축에 대한 관심이 고조되고 있다. 특히 패턴매칭, 분류, 예측 등을 필요로 하는 재무분야는 인공 신경망의 효과적 응용분야로 인식되고 있다. 재무분야를 대상으로 한 인공 신경망의 기존 연구들은 주로 기업의 부실예측이나 채권의 등급결정 등과 같은 분류문제에 대한 응용들이 주류를 이루고 있다. 1980년대 초까지 이러한 재무분야의 문제를 다루기 위해 주로 통계적 기법들을 사용하였다. 특히 부도예측을 위해서는 Altman의 선행연구(1968)이후 판별분석(discriminant analysis)을 많이 이용하였다(Deakin, 1972; Pantalone & Platt, 1987).

* 동명정보대학교 유통경영학과
** 부산대학교 경영학부

최근들어 인공 신경망을 이용한 부도예측에 관한 연구가 많은 관심을 보이고 있다. 인공 신경망을 이용한 부도예측대상들은 크게 금융기관과 일반기업으로 나눌 수 있다. Tam & Kiang(1992)과 Salchenberger et al.(1992) 등은 금융기관의 부도예측을 위해 인공 신경망을 이용하였으며, Coats & Fant(1991-1992), Fletcher & Goss(1993), Wilson & Sharda(1994), Jo et al.(1997) 등은 일반기업의 부도예측을 위해 인공 신경망을 이용하였다. 그 외에도 채권등급결정을 위해 인공 신경망을 이용한 연구에는 Kim et al.(1993)의 연구가 있다.

그러나 이러한 선행연구들은 인공 신경망과 전통적인 통계적 기법간의 비교를 통해 인공 신경망의 우수성만을 검증하였다. 따라서 대부분의 선행연구들은 인공 신경망의 성과에 많은 영향을 미치고 있는 가중치 변화방법에 대한 문제는 다루지 않았다. 가중치는 인공 신경망의 학습에 있어 중요한 부분으로서 이의 변화방법에 따라 예측결과에 많은 영향을 줄 수 있다. 가중치의 변화방법은 크게 2가지로 나누어 볼 수 있다. 첫째는 각 입력 데이터 별로 오차를 수정하기 위해 가중치를 변화시키는 방법이며, 둘째는 일정한 데이터를 입력한 후 누적된 오차에 대해 이를 수정하기 위해 가중치를 변화시키는 방법이다. 첫번째 방법은 매번 가중치를 변화시키므로 인공 신경망의 훈련시간은 많이 소요되나 신속히 오차를 반영하면서 학습을 진행시킬 수 있다. 두번째 방법은 첫번째 방법과 반대로 인공 신경망의 훈련시간을 단축시키면서 누적된 오차를 반영하게 된다.

인공 신경망 모형의 성과에 영향을 미치는 다른 요인으로는 은닉층(hidden layer)의 노드수가 있다. 은닉층의 노드수 결정문제 역시 현재까지 체계적으로 이루어지지 않고 있다. 그러나 은닉층의 노드수도 가중치 변화방법과 같이 인공 신경망의 성과에 중요한 영향요소라 할 수 있다. 따라서 본 논문에서는 가중치 변화방법과 은닉층의 노드수가 예측정확성에 어떠한 영향을 미치는가를 살펴봄으로써 인공 신경망의 효과적인 이용문제를 검토하고자 한다.

II. 인공 신경망 실험모형

2.1 인공 신경망의 지식획득과정

인공 신경망은 인간의 뇌를 모방한 모형으로서 노드간의 연결강도를 조정하므로서 지식을 표현한다. 인공 신경망에서의 연결강도는 노드간의 가중치를 의미하며, 노드간의 적절한 가중치를 결정하는 문제는 매우 어려운 문제이며, 특히 모집단분포에 대한 정보가 없을 경우에는 더욱 그러하다. 가중치를 결정하기 위한 일반적인 방법중 하나는 예(examples)를 통해 인공 신경망이 학습을 하도록 하는 것이다(Hinton, 1989). 그러므로 가중치의 변화는 반복적인 학습과정을 통해 이루어진다. 따라서 인공 신경망의 지식획득과정은 주어진 예에 대해 가장 적합한 가중치를 찾는 과정이다.

인공 신경망의 지식획득과정인 학습방법에는 입력내용에 대한 바람직한 출력결과를 알려주는 감독학습(supervised learning) 방법과 그렇지 않은 무감독학습(unsupervised learning) 방법이 있다. 현재 인공 신경망의 구축에는 구축의 용이성과 훈련시간상의 문제로 인해 감독학습 방법을 많이 이용하고 있다. 따라서 인공 신경망에 있어서 지식획득과정은 입력결과에 대한 인공 신경망의 실제 출력결과와 입력된 기대 출력(expected output)간의 오차를 계산해서 이 오차가 원하는 크기 이하가 될 때까지 반복실험을 통해 가중치를 변화시키는 과정이며, 학습이 완료된 시점에서의 노드간의 연결상태가 지식의 표현형태가 된다.

인공 신경망의 지식획득과정으로 많이 이용되는 감독학습방법으로는 Rumelhart et al.(1986)에 의해 개발된 오류 역전파(error backpropagation)규칙이 있다. 오류 역전파 규칙은 일반화된 델타규칙(generalized delta rule)이라고도 불리는 방법으로 어떠 노드의 활성이 다른 노드의 출력에 영향을 주었다면 두 노드간의 연결가중치를 조정하는 방법으로 훈련자료에 존재하는 규칙적 현상과 예외적 현상을 처리할 수 있는 능력이 있다. 오류 역전파규칙은 전향전파(forward propagation)와 후향전파(backward propagation)로 구성된다. 표본 자료의 수가 s개, 각 자료의 입력값이 $X_i = (x_{i1}, x_{i2}, \cdots, x_{im})$, $(1 \le i \le s)$ 그리고 기대 출력값이 $D_i = (d_{i1}, d_{i2}, \cdots, d_{in})$ 이라고 가정하면 전방향 전파에서는 X_i를 입력층에 제공하므로서 초기 가중치 집합 W를 통해 인공 신경망의 실제 출력값 $Y_i = (y_{i1}, y_{i2}, \cdots, y_{in})$을 계산하게 된다. 그후 인공 신경망의 실제 출력값 Y_i 와 기대 출력값 D_i 를 비교해서 양자간의 차이를 제곱한 $(y_{ij} - d_{ij})^2$, $(1 \le j \le n)$을 각 출력노드에 대해 계산한다. 그리고 이들 차이를 다음과 같이 합한 E를 계산한다.

$$E = \sum_{j=1}^{n} \frac{(y_{ij} - d_{ij})^2}{2}$$

인공 신경망은 시스템의 실제 출력값인 Y_i와 기대 출력값인 D_i의 차이를 최소화하도록 가중치 집합 W를 조정하는 것이 목적이므로 인공 신경망의 지식획득과정은 E의 최소화문제로 볼 수 있다. 따라서 전향전파의 단계에서 구한 E를 출력층에서 입력층으로 층별로 역전파하게 되며, 지식획득을 위한 인공 신경망의 이러한 과정을 후방향 전파단계라고 한다.

후방향 전파단계에서는 기울기법(descent method)에 의해 최적의 가중치 집합을 찾는 단계이다. 이때 각 가중치의 변화 방향과 크기는 다음의 Δw_{ij} 공식을 이용한다

$$\Delta w_{ij} = -\frac{\partial E}{w_{ij}} \varepsilon, \quad 0 < \varepsilon < 1$$

위 수식에서 ε는 최적해에 접근하는 속도를 결정하기 위해 Δw_{ij}의 크기를 조절하기 위한 매개변수이다. 각 노드간의 새로운 연결가중치(W^{new})는 결국 다음과 같이 계산된다.

$$W^{new} = W^{old} + \Delta w_{ij} \qquad \text{where } W^{old} : \text{기존 연결가중치}$$

이때 가중치를 갱신시키는 시점은 2가지 방법으로 이루어질 수 있다. 즉, 가중치 W를 갱신할 때 각 자료(X_i, D_i)별로 실행결과를 반영할 수도 있고, Δw_{ij}를 일정 자료군 단위로 누적시켜서 할 수도 있다. 만약 가중치 변화시점을 반영해서 새로운 연결가중치를 나타내면 다음과 같이 표현할 수 있다.

$$W^{new} = W^{old} + \frac{\sum_{j=1}^{p} \Delta w_{ij}}{p}, \quad 1 \le p \le s$$

만약 $p=1$이면, 이는 각 자료별로 즉각 오차를 반영하기 위해 가중치를 변화시키는 경우가 되며, $p=s$이면, 이는 모든 자료를 인공 신경망에 입력한 후 오차평균을 가중치 변화에 이용하는 경우가 된다. 따라서 p의 크기인 epoch size에 따라 반영되는 가중치 크기가 다르게 된다. 그러므로 인공 신경망기법은 최적화기법이 아니므로 인공 신경망의 학습과정이 끝난 후, 최적의 가중치 W는 epoch size에 따라 다를 수 있다. 이는 곧 epoch size에 따라 인공 신경망의 성과가 변할 수 있다는 것을 의미한다. 기존의 연구에서는 이러한 문제점을 언급하지 않고 있다. 인공 신경망의 학습과정에서 epoch size의 결정문제는 최적해 탐색과 학습시간이라는 2가지 관점에서 중요한 의미를 지닌다. 최적해 탐색관점에서 epoch size가 중요한 이유로 가중치 변화의 크기가 다르므로 근사해 기법인 인공 신경망의 해가 다를 수 있다는 것이다. 그리고 학습시간의 관점에서 epoch size가 중요한 이유는 epoch size에 따라 가중치 변화횟수가 결정되므로 학습시간이 변하기 때문이다. 따라서 본 논문에서는 이러한 점을 고려해서 인공 신경망의 성과에 영향을 주지 않으면서 학습시간을 최소화하기 위해 epoch size를 변동시킬 수 있는지 여부를 알아보고자 한다.

2.2 인공 신경망 실험모형의 구성

대부분의 인공 신경망 모형은 단일의 입출력층(input/output layer)과 소수의 은닉층(hidden layers)으로 구성된다. 입력층은 학습에 필요한 입력 요소 노드의 집합이다. 이들 노드들은 은닉층의 각 노드와 완전 연결(fully connected) 형태로 연결되나 동일 계층내의 노드들은 상호 연결되지 않는다. 각 층별로 최적의 노드수를 결정하는 명확한 규칙이 없기 때문에 노드의 수는 휴리스틱적인 방법이나 시행착오적인 방법으로 결

정된다.

은닉층은 없거나 1개 이상의 계층으로 구성되며, 은닉층의 수도 노드수의 결정과 같이 명확한 결정규칙이 없다. Tam & Kiang(1992)의 연구에 따르면 은닉층이 없는 경우와 1개가 있는 경우를 비교한 결과, 1개가 있는 인공 신경망이 우수한 결과를 나타냈다. 그리고 Collins et al.(1988), Dutta & Shekhar(1988), 그리고 Salchenberger et al.(1992) 등의 연구에 의하면 인공 신경망의 은닉층 수를 2개 이상으로 증가시켜도 단일 은닉층을 갖는 인공 신경망보다 예측률이 나아지지 않는 것으로 나타났다. 따라서 대부분의 선행연구에서 인공 신경망의 성과를 높이고 학습시간을 줄이기 위해 단일 은닉층을 갖는 인공 신경망을 이용하고 있다.

부실예측용 인공 신경망 모형의 구축을 위해 일반법인중 각각 796개의 건전기업과 부실기업에 대해 총 35개의 재무변수 관련 자료를 수집해서, 각각 398개씩 자료의 수를 동일하게 학습 및 테스트 자료군에 할당하였다. 자료군의 크기를 동일하게 한 이유는 Jain & Chandrasekaran(1982)과 Wilson & Sharda(1994)의 연구에서 동일한 비율로 구성된 자료군을 이용할 경우 인공 신경망의 예측 정확률이 증가하는 것으로 나타났기 때문이다. 또한 외생변수의 통제를 위해 모든 표본 자료는 외부감사를 받지 않는 중소규모의 일반법인을 대상으로 1995년부터 1997년 사이의 재무제표를 중심으로 자료를 수집하였다.

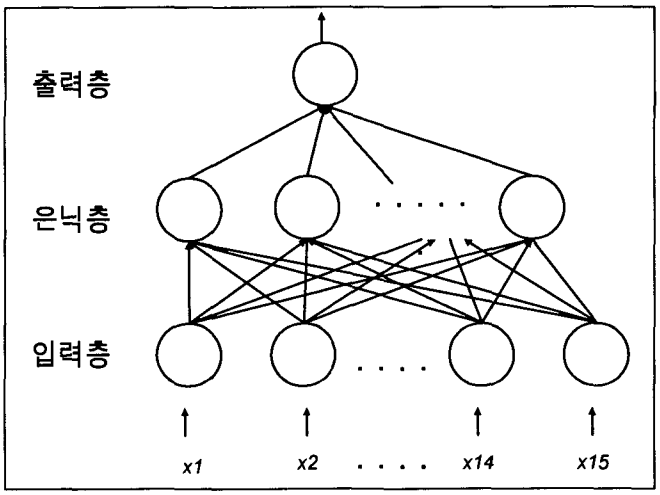
<표 1> 인공 신경망의 입력변수

지 표 그 룹	변 수 명
규모(2/3)	업력, 매출액,
생산성(1/2)	부가가치율
성장성(2/3)	매출액증가율, 총자산증가율
수익성(5/10)	매출액순이익율, 매출원가율, 손익분기점율, 자기자본순이익율, 총자본경상이익율
안정성(3/11)	매출채권대매입채무, 유보액대총자산비율, 총차입금대매출액
추세(1/1)	금융비용부담율 증가분
활동성(1/5)	재고자산회전율

* 괄호속의 숫자는 해당 그룹의 전체 변수중 단계적 방법에 의해 선정된 변수의 수

일반적으로 실험에 사용할 입력변수의 수가 너무 많으면 변수들간의 다중공선성(multicolliiarity) 문제가 발생할 수 있다(Jo et al.(1997)). 또한 인공 신경망 모형의 패턴 발견을 어렵게 할 수도 있다(Adriaans & Zantinge(1990); Altman(1993); Berry & Linoff(1997)). 이러한 문제의 해결을 위해서 Jo et al.(1997), Salchenberger et

al.(1992), 그리고 Sung et al.(1999) 등의 연구에서는 단계적 방법(stepwise method)을 이용해서 입력변수의 수를 조절하였다. 특히 Jo et al.(1997)의 연구에서는 t-test와 단계적 방법을 이용해서 각각 입력변수를 선택해서 다변량 판별모형과 인공 신경망모형의 예측률을 비교한 결과, 단계적 방법에 의해 선택된 변수를 이용하였을 때 예측률이 보다 우수한 것으로 나타났다. 따라서 본 논문에서도 이러한 문제를 해결하기 위해 통계 패키지를 이용해서 단계적 방법에 따라 7개 지표그룹으로 되어 있는 총 35개의 변수중 15개를 입력변수로 선정하였다(<표1> 참조). 또한 은닉층의 노드수는 Kolmogrov 정리에 따라 1 ~ 2n+1(n: 입력 노드의 수)개 사이로 설정하였으며, 출력층의 노드수는 1개로 해서 <그림 1>과 같은 인공 신경망의 실험모형을 구축하였다.



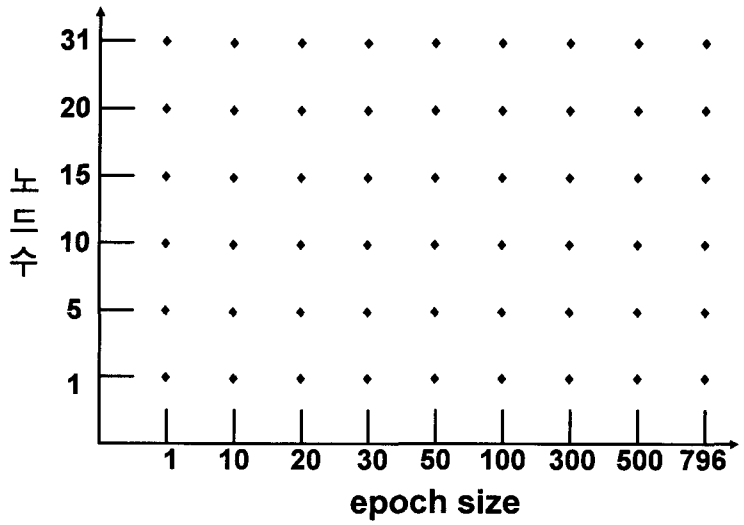
<그림 1> 인공 신경망의 실험모형

III. 가중치 변화방법 및 노드수에 따른 인공 신경망의 예측률 변화

3.1 실험설계

가중치 변화시점을 결정하는 epoch size와 은닉층의 노드수 변화에 따른 인공 신경망의 예측률을 알아보기 위해 <그림 2>와 같은 실험구조를 설정하였다. 은닉층의 노드수 범위는 Kolmogrov 정리에 따라 구한 최소값(1개)과 최대값(31개)을 중심으로 실험대상점을 선정하였다. 그리고 epoch size 요인은 가중치 변화가 가능한 최소(1개) 및 최대(796개) 범위를 고려해서 실험대상점을 선정하였다. 따라서 실험대상점은 총 54개가 선정되었다.

만약 epoch size가 1이면 이는 각 자료별로 기대 출력값과 인공 신경망의 실제 출력값을 비교해서 오차부분을 즉각 가중치 변화에 반영하는 것을 의미하므로 새로운 가중치는 $W^{old} + \Delta w_{ij}$가 된다. 그리고 eopch size가 796이면 각 자료별로 발생하는 오차를 자료군 전체에 대해 누적시킨 후 모든 자료군이 인공 신경망에 제공된 후 오차평균을 가중치변화에 반영하므로 새로운 가중치는 $W^{old} + \sum_{j=1}^{796} \Delta w_{ij} / 796$가 된다.



<그림 2> 인공 신경망의 실험구조

추가적으로 인공 신경망 모형의 예측력이 어떠한가를 알아보기 위해 전통적으로 재무분야에서 많이 이용하는 다변량 판별모형(multivariate discriminant model)을 단계적 방법으로 추출한 재무변수를 이용해서 구축한 후 양자간의 예측력도 비교할 것이다.

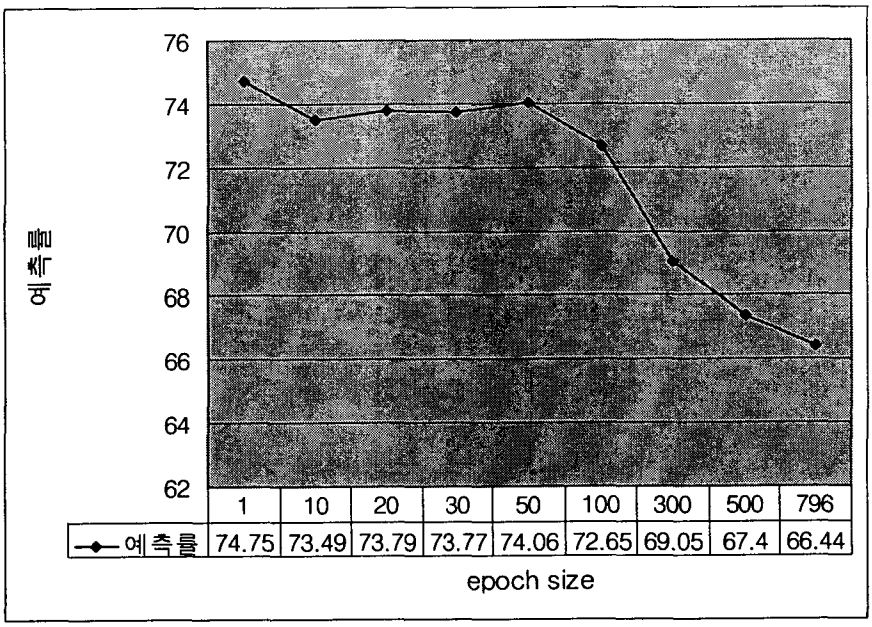
3.2 실험결과

가중치의 변화시점과 은닉층의 노드수를 변화시키면서 실시한 인공 신경망의 전체적인 실험결과는 <표 2>와 같다. 실험결과에 따르면 예측률이 가장 높은 경우는 은닉층의 노드수가 1개이면서 epoch size가 20과 50인 경우로서 예측률이 76.01%였다. 그리고 예측률이 가장 낮은 경우는 은닉층의 노드수가 15개이면서 epoch size가 796일 경우로서 예측률이 50.25%였다. 따라서 인공 신경망 모형은 적용방법에 따라 예측률 차이가 많은 것으로 나타났다. 실험결과에 대한 검토의 편의를 위해 epoch size에 따른 예측률 변화, 은닉층의 노드수에 따른 예측률 변화, 다변량 판별모형과의 비교로 나누어 살펴보기로 한다.

<표 2> 인공 신경망의 실험결과(단위 : %)

노드수 / epoch size	1	5	10	15	20	31	평 균
1	74.12	74.88	74.50	74.62	75.38	75.00	74.75
10	73.49	73.12	73.25	73.50	73.62	74.00	73.49
20	76.01	72.99	73.37	72.62	73.62	74.12	73.79
30	73.87	73.12	74.00	74.00	73.75	73.87	73.77
50	76.01	73.49	74.25	72.74	73.75	74.12	74.06
100	74.75	73.75	73.49	63.82	75.00	75.13	72.65
300	74.37	71.73	68.97	53.64	71.74	73.87	69.05
500	72.62	70.36	67.21	51.38	69.85	72.99	67.40
796	70.23	69.98	63.70	50.25	71.11	73.37	66.44
평 균	73.94	72.60	71.41	65.17	73.09	74.05	71.71

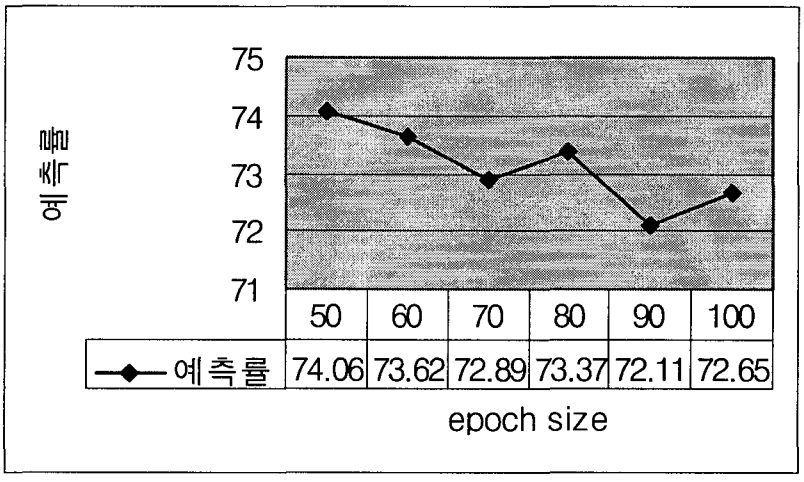
3.2.1 epoch size에 따른 예측률 변화



<그림 3> epoch size별 예측률 변화

가중치 변화시점이 예측 정확률에 미치는 영향을 알아보기 위해 epoch size별 평균 예측률을 살펴보면 <그림 3>과 같다. epoch size 변화에 따른 예측률 변화추이는 일정한 크기의 epoch size까지는 예측률의 변동이 크지 않다가 특정 크기(epoch size=50) 이상으로 epoch size가 증가함에 따라 급격히 예측률이 하락하는 것으로 나타났다. 특히 epoch size가 1일 때 예측률이 74.75%로서 가장 높았으며, epoch size가 796일 때 예측률이 66.44%로서 가장 낮게 나타났다. 이러한 연구결과는 지원철(1995)의 연구에서도 동일하게 나타났다.

epoch size가 50에서 100이 됨에 따라 예측률 차이가 급격히 하락하므로 보다 구체적인 예측률 변화시점을 알아보기 위해 이 구간에 대한 epoch size를 10씩 변화시키면서 예측률 변화를 분석해 본 결과 <그림 4>와 같은 결과를 나타내었다. <그림 4>에서 보는 바와 같이 epoch size가 50 이상이 됨에 따라 전체적인 예측률이 하락함을 알 수 있다.



<그림 4> epoch size변화에 따른 예측률 변화 추가분석

epoch size가 1, 50, 796인 3가지 경우를 대상으로 예측률 차이에 따른 통계적 유의성 검정을 위해 쌍표본 t 검정(paired-sampled t-test)을 실시해본 결과는 <표 3>과 같았다. 먼저 epoch size가 1인 경우와 50인 경우간의 예측률 차이가 유의한가를 알아본 결과 유의수준 p=0.1을 기준으로 양자사이에는 유의한 차이가 없는 것으로 나타났다. 그러나 epoch size가 1인 경우는 유의수준 p=0.1을 기준으로 796인 경우와 비교해서 예측률 차이가 유의한 것으로 나타났다. 그리고 epoch size가 50인 경우도 유의수준 p=0.1을 기준으로 796인 경우와 비교해서 예측률 차이가 유의한 것으로 나타났다.

<표 3> epoch size별 예측률 차이분석(* : p < 0.1)

epoch size	예측률	*t* 값	유의수준
1	74.75	1.215	0.279
50	74.06		
1	74.75	2.412	0.061*
796	66.44		
50	74.06	2.326	0.068*
796	66.44		

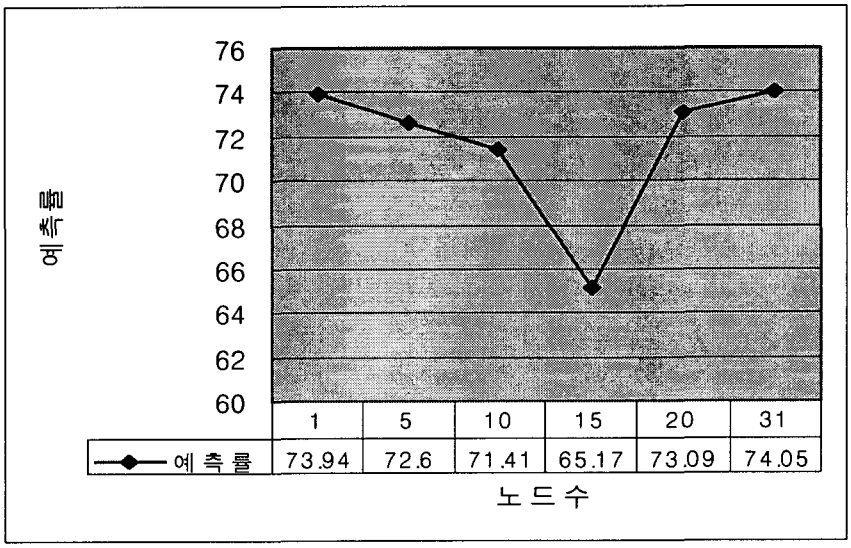
따라서 본 부도예측 실험에서는 오류 역전파규칙에 따른 학습방식의 경우 가중치 변화시점이 예측률에 많은 영향을 미침을 알 수 있으며, 특히 각 학습자료를 인공 신경망 모형에 제시할 때마다 발생하는 오류를 즉각적으로 반영하는 것이 보다 우수한 학습방식임을 알 수 있었다. 그러나 이렇게 가중치를 변화시킬 경우에는 학습시간이 많이 소요된다는 문제점이 발생한다. 따라서 epoch size를 일정한 크기까지 증대시켜도 예측률에 큰 차이가 없었으므로 높은 성과를 거두면서 훈련시간을 단축하기 위해서는 허용 가능한 범위까지 epoch size를 증가시키는 것이 바람직한 설계방법임을 알 수 있다.

그러나 epoch size를 얼마나 크게 하는 것이 예측률과 훈련시간과의 최적 상호상쇄(trade-off)인가 하는 점은 해결해야 할 문제의 중요성과 응용분야에 달려있다. 즉, 해결해야 할 문제의 예측결과와 실제결과의 차이가 미치는 영향도가 크다고 한다면 가능하면 작은 epoch size를 선택하는 것이 좋을 것이다. 그리고 Chung & Silver(1992)의 연구에 따르면 예를 통한 학습(learning by examples) 방법들의 성과는 응용분야에 따라 차이가 있기 때문에 응용분야에 따라 eopch size 변화에 따른 예측률 변화정도에 차이가 있을 수 있다. 그러므로 적용영역별로 epoch size의 변화에 따른 예측률 변화를 검토한 후 epoch size를 결정해야 할 것이다.

3.2.2 노드수에 따른 예측률 변화

은닉층의 노드수를 변화시킬 경우 예측률 변화는 <그림 5>와 같았다. 노드수를 증대시킴에 따라 예측률은 하락하다가 다시 일정지점(15개)이 지나면 상승하는 것으로 나타났다. 특히 은닉층의 노드수가 최대치인 31개일 때 최대 예측정확률인 74.05%를 나타냈으며, 은닉층의 노드수가 15개일 때 최소 예측정확률인 65.17%를 나타냈다. 이러한 결과는 이재식 & 한재홍(1995)의 연구에서도 유사하게 나타났다. 이들은 입력층

과 출력층의 노드수를 각각 9개와 1개로 고정킨 후, 은닉층의 노드수를 2, 5, 10, 15, 20개로 변화를 주면서 실험한 결과 2개일때가 가장 예측률이 높았으며, 다음으로는 20개일 경우가 높은 것으로 나타났다. 이들의 연구에서는 은닉층의 노드수 변화에 따른 구체적인 예측률은 제시하지 않아지만, 예측률이 V자형 포물선을 따르는 형태가 본 논문의 실험결과와 유사하였다.



<그림 5> 은닉층의 노드수에 따른 예측률 변화

<그림 5>에 나타난 예측률의 차이를 통계적으로 비교하기 위해 쌍표본 t 검정을 실시한 결과는 <표 4>와 같다. 즉, 노드수의 변화에 따른 최대 예측률과 최소 예측률 간의 차이는 유의수준 p=0.1을 기준으로 유의한 것으로 판정되므로 은닉층의 노드수가 31개일 경우가 15개일 경우보다 우수한 예측률을 갖는 것으로 나타났다.

<표 4> 은닉층의 노드수에 따른 예측률 차이분석(* : p < 0.1)

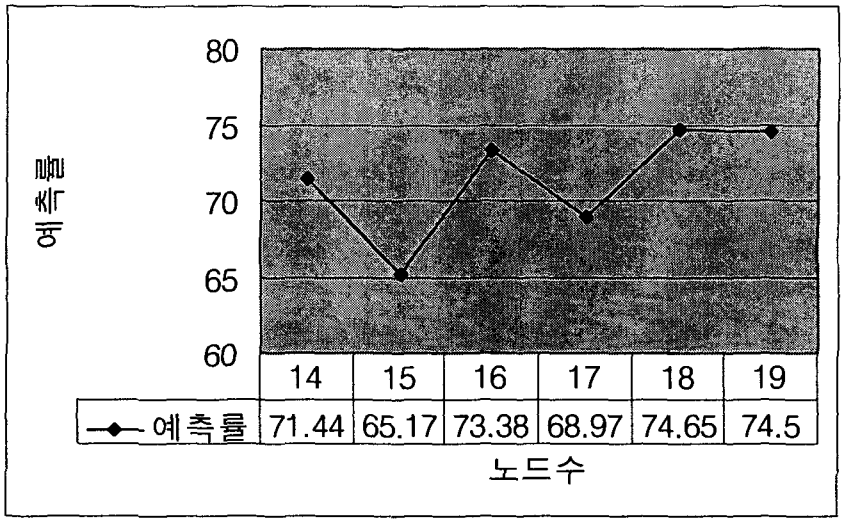
노드 수	예측률	t 값	유의수준
31	74.05	-2.609	0.031*
15	65.17		

은닉층의 노드수가 15인 지점이 가장 낮은 예측률을 나타내는가를 알아보기 위해 추가적인 분석을 위해 은닉층의 노드수를 14부터 20사이까지 1씩 증가시켜 가면서 추가적인 실험을 해본 결과는 <그림 6>과 같았다. <그림 6>에 나타난 것과 같이 은닉층의 노드수를 증감시킨 결과 예측률이 하락하지 않고 모두 은닉층의 노드수가 15인

경우보다 높은 것으로 나타났다. 특히 은닉층의 노드수가 18과 19인 경우는 본 실험에서 가장 높은 예측률을 나타낸 은닉층의 노드수가 31인 경우보다 높은 예측률을 나타내었다. 따라서 은닉층의 노드수가 15를 기준으로 예측률이 상승함을 알 수 있다.

이상의 실험을 통해서 볼 때 오류 역전파규칙의 경우 은닉층의 노드수도 예측률에 많은 영향을 미침을 알 수 있다. 일반적으로 은닉층의 노드수가 너무 많으면 인공 신경망 모형이 입력 자료의 패턴을 기억하기 때문에 일반화 능력이 부족해지므로 학습과정에서는 결과가 좋으나, 테스트과정에서는 결과가 나빠진다. 이러한 사항은 Salchenberger et al.(1992)의 연구에서도 입증되었다. Lodewyck & Deng(1993)의 연구에서는 은닉층의 노드수 변화에 따라 예측성과의 변동은 있으나 일정한 추세는 없는 것으로 나타났다. 그러나 여기서는 은닉층의 노드수가 일정 수준(15개)까지 늘어남에 따라 예측률이 하락하다가 그 이상의 수준에서는 다시 예측률이 증가하는 상이한 결과가 나타났다.

이러한 현상이 초기에 부여한 가중치에 의한 영향인가를 알아보기 위해 초기 가중치 부여를 위한 난수 seed를 변화시켜 은닉층의 노드수를 15로 해서 반복실험을 실시해 보았다. 그 결과 난수 seed를 바꾸어도 유사한 예측률(66.51%)을 나타내었다. 따라서 은닉층의 노드수가 일정한 크기까지 증가함에 따라 인공 신경망의 성과는 하락하다가 이를 기준으로 다시 상승하는 것으로 나타났다. 따라서 인공 신경망의 설계시 은닉층의 노드수 결정문제에 많은 관심이 집중되어야 함을 알 수 있다.



<그림 6> 은닉층의 노두수 변화에 따른 예측률 변화 추가분석

3.2.3 다변량 판별모형과의 비교

인공 신경망의 성과를 평가하기 위해 재무자료를 이용한 분류문제에 많이 이용되는 통계적 기법인 다변량 판별모형과 인공 신경망 모형의 예측률을 비교한 결과는 <표 5>와 같았다. <표 5>에 나타난 바와 같이 다변량 판별모형은 건전기업의 분류에서 인공 신경망 모형보다 높은 예측률을 나타냈으며, 부도기업의 분류에서는 인공 신경망 모형이 다변량 판별모형보다 높은 예측률을 나타냈다. 평균적으로 보면, 다변량 판별모형이 인공 신경망 모형보다 높은 예측률을 나타냈다. 그러나 건전기업의 분류 정확성보다 부도기업의 분류 정확성이 비용의 측면에서 보다 중요하다(Barth et al.(1989)). 따라서 인공 신경망 모형이 다변량 판별모형보다 전체적으로는 예측률이 낮았지만, 부도기업의 예측률이 상대적으로 높기 때문에 다변량 판별모형이 인공 신경망 모형보다 유용한 모형이라고 보기는 어렵다.

<표 5> 다변량 판별모형과 인공 신경망 모형의 예측률 비교

	다변량 판별모형	인공 신경망 모형
건전 기업	83.92	70.24
부도 기업	64.32	73.18
평 균	74.12	71.71

대부분의 선행연구에서는 <표 6>에 나타난 바와 같이 인공 신경망 모형이 다변량 판별모형보다 높은 예측률을 나타내었다. 그러나 본 논문에서는 상이한 결과를 나타내었다. 이러한 현상의 이유중 하나는 기업의 규모가 비교적 작은 일반법인을 대상으로 자료를 수집하였기 때문이라 할 수 있다. Bansal et al.(1993)의 실험에 의하면 자료의 부정확성을 증대시킬수록 인공 신경망 모형의 예측률이 하락하는 것으로 나타났다. 즉, 자료가 부정확할수록 다변량 판별모형의 예측률이 인공 신경망 모형의 예측률보다 높을 가능성이 증대하는 것이다. 따라서 일반법인은 규모가 작아 경영활동비율들이 대기업보다 불균형을 이룰 가능성이 높다고 볼 때 이 실험에서와 같이 다변량 판별모형이 인공 신경망 모형보다 예측률이 높을 수 있다. 이러한 예측을 가능하게 하는 또다른 이유는 이재식 & 한재홍(1995)의 연구에서도 찾아 볼 수 있다. 이들의 연구 역시 중소기업을 대상으로 인공 신경망 모형을 이용해서 부도예측을 한 결과, 재무정보만을 이용하였을 경우 64.4%가 가장 높은 예측률을 나타내는 것으로 나타났다. 그러나 이들의 연구에서는 자료수의 불충분 문제와 훈련 및 테스트 자료의 불균형 문제를 해결한다면 다소의 예측정확률은 상승할 수 있을 것이나 기존의 연구에서 처럼 인공 신경망 모형이 우수한 예측정확률을 나타낼지는 의문이다. 그렇다면 본 논문에서의 예측정확률은 매우 높다고 볼 수 있을 것이다.

<표 6> 선행연구와의 예측률 비교

선 행 연 구	다변량 판별모형	인공 신경망 모형
Coats & Fant	85.50	93.50
Tam & Kiang	83.30	86.95
Jo et al.	82.22	83.79

Ⅳ. 결론 및 향후 연구

IMF 이후 기업의 부도사태가 속출함에 따라 기업의 건전성 여부에 대한 투자자들의 관심이 고조되고 있다. 따라서 여러 금융기관에서 인공지능기법을 이용한 기업평가 시스템의 개발에 많은 관심을 보이고 있으며, 인공 신경망은 이를 위한 대안중 하나로 언급되고 있다. 인공 신경망이 기업평가시스템의 개발방법으로 주목받는 이유는 기존의 통계적 기법보다 성과가 우수하다는 선행연구들이 있었기 때문이다. 하지만 인공 신경망은 본 논문의 실험결과에서 나타난 바와 같이 모형의 구성 및 가중치 변화방법에 따라 성과에 많은 차이를 보일 수 있다.

본 논문에서 밝혀진 바에 따르면 인공 신경망을 이용한 모형구축시 경험적 방법에 따라 은닉층의 노드수를 결정할 경우 개발자가 원하는 수준의 성과를 얻기 어렵다는 것을 알 수 있다. 또한 가중치의 변화방법도 적절한 성과를 거두면서 신속한 모형개발을 위해서는 epoch size를 적정 수준으로 증대시키는 것이 바람직함을 알 수 있다. 그러나 epoch size가 변동할 경우 예측률의 추이는 응용분야에 따라 다를 수 있으며, 또한 예측률이 실제 결과에 미치는 영향이 클 경우에는 모형개발의 신속성과 은닉층의 노드수 및 epoch size의 크기간에 상호상쇄 관계를 고려해서 모형을 구축해야 할 것이다.

본 논문과 관련해서 추가적으로 수행되어야 할 연구분야로는 인공 신경망 모형의 업종별 예측률을 비교하는 것이다. 이는 인공 신경망 모형이 모든 업종의 분류예측문제에 적합한 모형이 아닐 수 있다는 것이다. 따라서 인공 신경망 모형의 업종별 성과를 비교해서 적합한 적용분야를 찾는 것도 향후 연구가 될 수 있을 것이다.

또한 가중치 집합은 다차원 공간에 존재하는 반응표면상의 한 점이라 할 수 있다. 이러한 측면에서 보면 결국 가중치 결정을 위해 반응표면 분석기법의 도입이 가능할 수 있다. 따라서 최적의 가중치 집합을 결정하기 위해 인공 신경망의 학습방법과 반응표면 분석기법과의 결합도 향후 연구가 될 수 있을 것이다.

참 고 문 헌

[1] 이재식 & 한재홍, "인공신경망을 이용한 중소기업도산예측에 있어서의 비재무정보의 유용성 검증," *한국전문가시스템학회지,* 창간호, 1995, pp.123-134.

[2] 지원철, "신경망을 이용한 시계열 분석 : M1-Competition Data에 대한 예측성과 분석," *한국전문가시스템학회지,* 창간호, 1995, pp.135-148.

[3] Adriaans, P., & D, Zantinge, *Data Mining,* Addison-Wesley, 1996.

[4] Altman, E. I., "Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy," *Journal of Finance,* September, 1968, pp.589-609.

[5] Altman, E. I., *Corporate Financial Distress and Bankruptcy,* John Wiley and Sons, 1993.

[6] Bansal, A., R. J. Kauffman, & R. R. Weitz, "Comparing the Modeling Performance of Regression and Neural Networks as Data Quality Varies : A Business Value Approach," *Journal of Management Information Systems,* Summer, 1993, Vol.10, No.1, pp.11-32.

[7] Barth, J. R., R. D. Brumbaugh, D. Sauerhaft, & G. H. Wang, "Thrift Instiurion Failures : Estimating the Regulator's Closure Rule," In Kaufman, G. G.(Eds.), *Research in Financial Services,* Vol.1, JAI Press, 1989.

[8] Berry, M., & G. Linoff, *Data Mining Techniques : For Marketing, Sales, and Customer Support,* John Wiley and Sons, 1997.

[9] Chung, H., & M. S. Silver, "Rule-based Expert Systems and Linear Models : an Empirical Comparison of Learning-by-examples Methods," *Decision Sciences.* Vol. 23, 1992, pp.687-707.

[10] Coats, P. K., & L. F. Fant, "A Neural Network Approach to Forecasting Financial Distress," *The Journal of Business Forecasting,* Winter, 1991-1992, pp.9-12.

[11] Collins, E., S. Ghosh, C. Scofield, "An Application of a Multiple Neural-network

Learning System to Emulation of Mortage Underwriting Judgements," *Proceedings of the IEEE International Conference on Neural Networks,* 1988, Vol.2, pp.459-466.

[12] Deakin, E. B., "A Discriminant Analysis of Predictors of Business Failure," *Journal of Accounting Research,* 1972, pp.167-179.

[13] Dutta, S., & S. Shekhar, "Bond-rating: A Non-conservative Application of neural Networks, *Proceedings of the IEEE International Conference on Neural Networks,* 1988, Vol.2, pp.443-450.

[14] Fletcher, D., & E. Goss, "Forecasting with Neural Networks - An Application Using Bankruptcy Data," *Information & Management,* Vol.24, 1993, pp.159-167.
Hinton, G. E., "Connectionist Learning Procedures," *Artificial Intelligence,* 1989, Vol.40, pp.185-234.

[15] Jain, A., & B. Chandrasekaran, "Dimensionality and Sample Size Considerations in Pattern Recognition Practice," In Krishnaiah, P., & L. Kanal(Eds.), *Handbook of Statistics,* Vol.2, North-Holland, 1982, pp.835-855.

[16] Jo, H., I. Han, & H. Lee, "Bankruptcy Prediction Using Case-based Reasoning, Neural Networks, and Discriminant Analysis," *Expert System With Applications,* Vol.13, No.2, 1997, pp.97-108.

[17] Kim, J. W., H. R. Weistroffer, & R. T. Redmond, "Expert System for Bond Rating : a Comparative analysis of Statistical, Rule-based and Neural Network Systems," *Expert Systems,* Vol.10, No.3, August, 1993, pp.167-171.

[18] Lodewyck, R. W., & P. Deng, "Experimentation with a Back-propagation Neural Network - an Application to Planning End User System Development," *Information and Management,* Vol.24, 1993, pp.1-8.

[19] Pantalone, C., & M. Platt, "Predicting Failure of Savings and Loan Associations," *AREUEA Journal,* 1987, Vol.15, No.2, pp.46-64.

[20] Rumelhart, D. E., G. Hinton, & R. Williams, "Learning Internal Representation by Error Propagation," In Rumelhart, D. E., & McClelland(Eds.), *Parallel Distributed Processing,* MIT Press, 1986, pp.318-362.

[21] Salchenberger, L. M., E. M. Cinar, & N. A. Lash, "Neural Networks : A New Tool for Predicting Thrift Failures," *Decision Sciences,* Vol.23, 1992, pp.899-916.

[22] Sung, T., N. Chang, & G. Lee, "Dynamics of Modeling in Data Mining : Interpretive Approach to Bankruptcy Prediction," *Journal of Management Information Systems,* Vol.16, No. 1, Summer, 1999, pp.63-85.

[23] Tam, K. Y., & M. Y. Kiang, "Managerial Applications of Neural Networks : The Case of Bank Failure Predictions," *Management Science,* Vol.38, No.7, 1992, pp.926-947.
Wilson, R. L., & R. Sharda, "Bankruptcy Prediction Using Neural Networks," *Decision Support Systems,* Vol.11, 1994, pp.545-557.

<Abstract>

The Influence of Weight Adjusting Method and the Number of Hidden Layer's Node on Neural Network's Performance

Jin-Baek Kim · Yuil Kim

The structure of neural networks is represented by a weighted directed graph with nodes representing units and links representing connections. Each link is assigned a numerical value representing the weight of the connection. In learning process, the values of weights are adjusted by errors. Following experiment results, the interval of adjusting weights, that is, epoch size influenced neural networks' performance. As epoch size is larger than a certain size, neural networks' performance decreased drastically. And the number of hidden layer's node also influenced neural networks' performance. The networks' performance decreased as hidden layers have more nodes and then increased at some number of hidden layer's node. So, in implementing of neural networks the epoch size and the number of hidden layer's node should be decided by systematic methods, not empirical or heuristic methods.