

---

저자 (Authors)	김성근
출처 (Source)	<a href="#">한국지능정보시스템학회 학술대회논문집</a> , 1994.4, 293-307(15 pages)
발행처 (Publisher)	<a href="#">한국지능정보시스템학회</a> Korea Intelligent Information Systems Society
URL	<a href="http://www.dbpia.co.kr/journal/articleDetail?nodeId=NODE00196013">http://www.dbpia.co.kr/journal/articleDetail?nodeId=NODE00196013</a>
APA Style	김성근 (1994). 재무비율을 기초로 한 주가예측 신경망시스템. 한국지능정보시스템학회 학술대회논문집, 293-307
이용정보 (Accessed)	송실대학교 203.253.***.153 2020/09/29 18:04 (KST)

---

### 저작권 안내

DBpia에서 제공되는 모든 저작물의 저작권은 원저작자에게 있으며, 누리미디어는 각 저작물의 내용을 보증하거나 책임을 지지 않습니다. 그리고 DBpia에서 제공되는 저작물은 DBpia와 구독계약을 체결한 기관소속 이용자 혹은 해당 저작물의 개별 구매자가 비영리적으로만 이용할 수 있습니다. 그러므로 이에 위반하여 DBpia에서 제공되는 저작물을 복제, 전송 등의 방법으로 무단 이용하는 경우 관련 법령에 따라 민, 형사상의 책임을 질 수 있습니다.

### Copyright Information

Copyright of all literary works provided by DBpia belongs to the copyright holder(s) and Nurimedia does not guarantee contents of the literary work or assume responsibility for the same. In addition, the literary works provided by DBpia may only be used by the users affiliated to the institutions which executed a subscription agreement with DBpia or the individual purchasers of the literary work(s) for non-commercial purposes. Therefore, any person who illegally uses the literary works provided by DBpia by means of reproduction or transmission shall assume civil and criminal responsibility according to applicable laws and regulations.

재무비율을 기초로 한 주가예측 신경망시스템

중앙대학교 경영학과  
김 성 근

1994년 3월 15일

한국전문가시스템학회  
1994년 춘계학술대회 발표논문

## 요 약

자료분류의 목적으로 사용되는 기법은 매우 다양하다. 기존 연구는 통계 기법보다 인공신경망 기법이 효과적임을 보여주고 있다. 주가수익률 예측의 문제에도 인공신경망 기법이 수차례 적용된 적이 있다. 그러나 이들 연구는 주식시장의 기술지표 또는 경제변수를 이용한 단기주가의 예측에 치중하여왔다. 본 연구의 목적은 상장업체의 재무비율과 주요 경제변수를 토대로 장기주가의 예측을 위한 인공신경망의 구축에 있다. 아울러 이 결과의 분석을 통해 보다 효과적인 인공신경망 구조에 대한 아이디어를 제시한다. 특히 입력변수의 성격에 따른 처리요소간의 연계구조, 입력자료의 종류별 개수와 인공신경망의 성능과의 관계에 대한 방향을 제시한다.

## 1. 서론

자료분류(Classification) 방법은 여러 가지의 경영 의사결정에 자주 적용되고 있다. 여기서 자료분류란 주어진 자료집합에서 유사한 성격의 자료집단을 추출해내고, 이를 토대로 새롭게 발생하는 자료를 이미 파악된 자료집단으로 할당하는 과정을 말한다. 이 기법을 주로 활용하는 문제는 기업도산예측, 회계기법선정, 감사의견결정, 신용등급예측, 신용공여결정 등이다.

자료분류를 위한 기법은 매우 다양하다. Han, et al. (Han, et al. 1993)은 자료분류에 사용되는 기법을 크게 세가지로 분류하고 있다. 우선 첫째로 통계기법을 들 수 있다. 경영의사결정에 주로 적용되어 온 통계기법은 다변량판별분석(Multiple Discriminant Analysis: MDA), logit 및 probit 모형 등이다. 둘째 분류는 귀납적 학습(inductive learning)이다. 이는 인공지능의 한 분야로서 주어진 사례들로부터 일정한 의사결정나무규칙(decision tree rules)을 끌어내고, 이를 토대로 주어진 상황에 대해 예측하는 기법이다. 그 대표적인 예는 1979년 Quinlan 에 의해 고안된 ID3 기법(Quinlan, 1979)이다.

마지막 자료분류 기법은 인공신경망(Artificial Neural Network: ANN)이다. 이것 역시 인간의 정보처리 패턴과 유사한 방식으로 시스템을 구축하고자 했던 인공지능 연구의 한 분야이다. 인공신경망은 여러 개의 처리단위(Processing Element: PE)가 網 구조로 짜여져 있다. 이 처리단위는 계층(layer)이라는 조직단위로 다시 조합된다. 그리하여 인공신경망은 입력계층과 출력계층을 갖고 있다. 아울러 인공신경망의 구성방법에 따라 1 개 이상의 은닉계층(hidden layer)을 가질 수도 있다. 인공신경망에서 각 처리단위는 통상 연계된 여러개의 입력신호로부터 보내어진 자료값을 받아, 여기에 일정한 전이함수(transfer function)를 적용하여 출력값을 내보내는 절차를 밟는다.

이러한 여러가지의 자료분류기법의 성능에 대한 연구는 다양하게 수행되었다. 대부분의 연구는 인공신경망이 보다 유용한 기법임을 말해 주고 있다. (Han, et al. 1993; Kim, 1992)

이러한 결과에 힘입어 인공신경망에 대한 연구는 다른 의사결정의 문제로 급격히 확산되고 있다. 최근에는 동적 특성을 갖고 있는 주가예측 문제에도 적용되고 있다. 최초의 연구는 Lapedes and Farber (1987)에 의한 주가지수의 예측을 위한 인공신경망의 적용이다. 그 연구의 목적은 인공신경망을 이용한 자기회귀(auto regressive) 방법과 단순이동평균방법과 비교하는 것이었다. 연구결과는 10주의 과거 자료를 토대로 다음 주의 주가를 예측하는 것으로 인공신경망이 단순이동평균법보다 나은 예측력을 갖고 있음을 보여주고 있다. 인공신경망을 전형적인 기술적 분석에 적용한 연구도 있었다. 일별 주가흐름을 나타내주는 도표의 특성 즉 최고치, 최저치 등을 토대로 다음의 일별주가를 예측한 연구(Ken-ichi & Testsuji, 1990)이다.

국내에서는 노종래와 이종호에 의해 처음 시도되었다. 이들은 경제관련자료 또는 부문별 주가지수 등의 입력자료를 토대로 종합주가지수의 월별 변화정도와 주별변화 정도를 예측하고자 했다. 또한 이재유는 증권시장 상황을 나타내주는 신용잔고, 고객예탁금, 회사채수익률 등의 입력자료를 토대로 일별주가지수를 예측하였다.

이들 선행연구의 특징은 대부분 주식시장의 전체흐름에 대한 인공신경망을 구축하고자 했던 점이다. 그리하여 고려했던 종속변수가 주로 종합주가지수이었다. 이재유의 연구는 별도로 특정주식의 주가 움직임도 연구하였으나, 각 주식별로 개별적인 인공신경망을 구축했다는 단점을 안고 있다. 또한 기존 연구의 대부분은 주로 일별 주가흐름과 같은 단기적 주가에측을 위한 것이었다.

그리하여 본 연구에서는 개별종목의 주가흐름에 관한 인공신경망을 구축하고자 한다. 이는 일별주가와 같은 단기적 예측이 아니라, 주가와 이에 영향을 미치는 관련변수와의 관계를 인공신경망 모형으로 파악한다는 점에서 기존 연구와 다르다. 본 연구에서 고려된 입력변수는 크게 개별회사의 재무비율 자료와 주식시장 전체의 움직임에 관한 자료로 구분된다. 기존의 인공신경망 응용연구에서 재무비율 자료를 사용한 적이 있으나, 모두 채권등급, 기업도산예측과 같은 회계부분에 국한되었다. 아울러 본 연구에서는 성격이 전혀 다른 두 집단의 입력변수를 토대로 인공신경망을 구성할 때 보다 효과적인 모형의 설계방안을 제시하고자 한다.

## 2. 주가 수익률 예측에 관한 연구

주식가치의 평가에는 다양한 방법이 사용되고 있다. 크게 두 가지로 구분된다. 우선 주식의 본질적 가치를 찾아내는 기본적 분석(fundamental analysis)을 들 수 있다. 여기에는 주가에 영향을 미치는 경제적 요인 (이자율, 물가상승률, 통화량, 국민소득 등)을 분석하는 경제분석, 특정 산업에 관한 요인을 다루는 산업분석, 그리고 특정 기업의 수익성, 안정성, 성장성 등을 분석하는 기업분석으로 세분된다. 즉 기본적 분석이란 주가에 영향을 미치는 기본요인을 체계적으로 분석함으로써 주식의 본질적 가치를 결정하는 것이다. 실제로 이러한 기본적 분석은 재무관리 분야에서 최근까지 활발하게 다루어지고 있다(Francis, 1980).

주식가치의 두번째 평가방법은 기술적 분석(technical analysis)이다. 이는 주가의 영향요인을 개별적으로 분석하기 보다는 주가의 전체적 움직임을 분석하는 방법을 취한다. 이러한 분석은 주가는 근원적으로 주식에 대한 수요와 공급에 의해 결정된다는 점을 기본 전제로 하고 있다(Edwards & Magee, 1966). 그리하여 기술적 분석은 이러한 수요와 공급을 측정하게끔 고안되어 있다. 즉 차트에 표시된 과거의 관련자료로부터 의미있는 패턴을 발견하고, 이를 통해 미래의 주가움직임을 예측하는 방법을 취한다.

최근 인공신경망에 대한 관심이 고조되고 있다. 이의 적용분야는 전자, 의학, 군사, 제조, 경영 등 매우 다양하다. 인공신경망을 주식시장에 적용하려는 연구는 해외에서 부터 시작되었다. 시계열자료에 인공신경망을 적용하여 일정한 패턴을 끌어낸 Lapedes & Farber(1987)의 연구와 바차트(barchart)의 최고가와 최저가를 토대로 변화패턴을 인식하는 Ken-ichi & Testsuji (1990)의 연구가 대표적이라고 할 수 있다. 그외에도 인공신경망의 기법을 이용하여 과거의 경제지표와 기술지표 등을 토대로 동경증권시장의 매매예측 시스템을 개발한 Takashi et al.(1990)의 연구도 있다.

국내에서도 인공신경망을 이용한 주가에측 연구는 노종래와 이종호(1991), 이재유(1993)에 의해 행해졌다. 이들 연구는 주로 종합주가

지수의 움직임을 대상으로 하였으며, 입력자료로 경제변수 또는 주식 시장 지표자료를 포함시켰다.

앞선 이들 연구는 인공신경망을 주로 종합주가지수의 단기적 예측에 치중하였다고 볼 수 있다. 또한 주식가치의 두가지 평가방법 중에서 기술적 분석에 더욱 치중하였다. 신용잔고, 고객예탁금과 같은 주식 시장 지표자료를 이용하여 개별주가의 움직임을 파악하려는 연구도 있었으나, 이도 기술적 분석에 해당한다고 하겠다.

인공신경망의 또 다른 적용분야는 주식의 내재가치를 결정해주는 기본적 분석이라고 할 수 있다. 실제 이와 비슷한 시도가 Yoon & Swales (1991) 에 의해 행해졌다. 이들은 고려했던 입력자료는 결산 자료에 대한 주주들의 반응과 같은 질적 자료 등이 대부분이었다. 개별회사의 재무비율 자료는 채권등급, 기업도산예측과 같은 회계부문의 인공신경망 연구에 사용된 적은 있었지만, 이와 주가와와 관계를 파악하려는 인공신경망 연구는 없었다. 그리하여 개별회사의 재무비율 자료와 주가수익률 사이의 일정한 패턴을 인공신경망으로 끌어낼 수 있는가의 여부도 가치있는 연구라고 할 수 있겠다.

### 3. 연구방법

주식시장 연구에 있어 표본선정은 중요한 문제이다. 본 연구에서는 일간지 주식시세란에 수록된 업체중 계층화 표본추출방식으로 표본을 선정하였다. 즉 국내 주요산업의 총업체수를 감안하여 표본업체의 수를 정한 다음, 무작위 방법으로 해당업종의 표본을 선정하였다. 이때 금융과 보험산업의 경우는 제외되었다. 그 이유는 금융기관 및 보험업체의 재무비율의 형태나 재무제표의 양식이 기타 산업과 상이하기 때문이었다. 선정된 표본업체는 <표 1>과 같다.

<표 1> 업종별 표본업체

업종	업체명	업체수	비율
음식료품	동방유랑, 해태제과, 풍진, 우성식품	4개	9%
섬유의복	전방, 유성, 제일합섬, 삼영모방, 패평양 패션, 조광피혁, 백양	7개	15%
화학석유	금호석유, 동방아그로, 한국포리, 영진약품, 한일약품, 삼일제약, 상아제약, 한국티타늄, 동아타이어, 한양화학	10개	22%
1차금속	동국제강, 부산파이프, 삼아알미늄	3개	7%
조립금속기계	삼익공업, 대우중공업, 세진, 동양기전, 내외반도체, 아남산업, 복두, 한국전자, 한일써킷, 고니정밀, 국제전선, 선도전기	12개	26%
자동차운송장비	아시아자동차, 삼성라디에타	2개	4%
종합건설	극동건설, 동부건설, 신화건설,	3개	7%
도매업	삼성물산, 대성산업, 선경, 진웅, 동방개발	5개	11%
총계		46개	100%

본 연구에서는 국내외 기존연구에서 투자수익률과 밀접한 관계가 있는 것으로 분석된 주요 변수를 포함시켰다

{{  
 ) 이와 같은 독립변수의 선정 배경에 대해서는 Chen & Roll(1986) 및 조지  
 호  
 (1989)에 자세하게 언급되어 있음.

}}  
 . 이들 변수는 크게 두 종류로 구분된다. 첫째는 경제 및 주식시장 변수이다.  
 여기에는 1년  
 만기 정기예금이자율, 주가수익배수(PER), 총통화 증가율, 주주수 증  
 가율, 시가총액 증가율, 주식거래량 증가율 등이 여기에 해당된다.  
 둘째 종류의 변수는 해당업체의 관련자료이다. 여기에는 기존 연구에  
 서 주가수익률과 밀접한 관련이 있는 것으로 파악된 재무비율과 베타  
 계수 등이 해당된다.

그리고 종속변수에 해당하는 주가수익률은 연간 주식투자수익률로 취급하  
 였다. 본 연구에서 변수로 고려된 변수의 종류와 설명은 <표 2>와 같다.

<표 2> 연구에 사용된 변수의 설명

변수군	변수	변수의 설명
해당업체 관련자료	자기자본비율	자기자본/총자본*100
	고정장기적합율	(고정자산 + 투자와 기타자산)/(자기자본 + 고정부채) *100
	순이익증가율	{(당기순이익/전기순이익)-1}*100
	자본금경상이익률	경상이익/총자본*100
	총자본회전율	매출액/총자본*100
	매출액순이익률	당기순이익/매출액*100
	산업대비주가수준	산업평균주가 대비 주가의 수준
	주가수익배수	PER = 주가/주당수익
	베타(Beta)	위험계수
경제 및 주식시장 변수	이자율	1년만기 정기예금 이자율
	주가수익배수	주식시장 전체 PER=주가/주당수익
	총통화증가율	전년도 통화량에 대한 증감율
	종합지수수익률	전년도 종합지수에 대한 증감율

주주수 증가율	전년도 총주주수에 대한 증감율
시가총액 증가율	전년도 시가총액에 대한 증감율
주식거래량증가율	전년도 주식거래량에 대한 증감율

본 연구의 자료는 년간으로 발생하는 자료로서 증권관련 문헌으로 부터 연구자가 직접 채록한 것이다. 즉 결산시점에서 발표된 재무비율 자료를 상장회사총감으로 부터 추출했다. 또한 주가수익배수, 베타계수, 기타 경제변수 등도 당해년도 자료를 기초로 계산된 자료이다. 이들 자료는 당해년도 연말시점의 자료이거나 또는 연말시점에서 계산된 자료이다.

그러나 주가수익률 자료는 당해년도 2월말 주가와 익년도 2월말 주가와와의 차이를 통해 구했다. 기준시점을 12월말이 아니라 익년 2월말로 선택한 것은 다음과 같은 이유 때문이다. 12월 결산법인의 재무비율 정보는 보통 익년 1월 또는 2월에 공개되므로, 이러한 정보효과는 2월 주가에 반영이 되리라고 가정했기 때문이다. 또한 독립변수와 종속변수(주가)를 동일 시점에 추출한 연구결과는 시점 t의 주가를 알기위해 시점 t의 종속변수값을 알고 있어야 한다는 조건 때문에 예측의 목적으로 사용하기는 어려움이 따를 수 있다. 본 연구에서와 같이 주가수익률의 기준시점을 익년 2월말로 선정함으로써, 12월말까지의 정보를 통해 익년 2월의 주가수익률을 예측할 수 있으리라고 본다.

인공신경망의 구축에서는 자료표현의 형태가 중요한 문제이다. 우선 자료의 이진화 여부이다. 이는 입출력 자료를 이진화 즉 0 또는 1로만 표현하는 방법을 취할 것인지, 아니면 실수값 그대로 취할 것인지를 결정해야 한다. 인공신경망의 교육성과를 높이기 위해 또는 입력자료가 비계량적인 질적 자료의 경우 보통 이진자료를 취한다. 또 다른 문제는 입력자료의 정규화 여부이다.

본 연구에서는 산업대비 주가수준이라는 입력변수의 경우 자료의 속성상 이진자료로 취급했다. 나머지 입력변수는 모두 실수로 취급했다. 이는 입력값을 이진수로 취급하면 입력 처리단위가 많아져서 신경망의 크기로 인한 제약이 따르기 때문이다. 다만, 재무비율과 주가수익배수의 경우, 자료값의 범위가 너무 커, 정규화할 필요가 있었다. 예를 들면 매출액순이익률의 경우 전체의 85%가 -2.70에서 5.64에 몰려 있으나, 몇개의 자료값은 54.00 또는 -11.33과 같은 의외치(outlier)가 존재하고 있다. 이러한 문제를 극복하기 위해 다음과 같은 산술식으로 순화하였다.

$$N(X_i) = (X_i - P_x) / (Q_x - P_x)$$

$X_i$  = 변수 X의 i 번째값

$P_x$  = 변수 X의 10% 퍼센타일(percentile)의 값

$Q_x$  = 변수 X의 90% 퍼센타일(percentile)의 값

출력변수의 경우 연간 주가수익률의 증감을 2-out-of-5-code를 사용하여 매우증가, 증가, 조금증가, 조금감소, 감소, 매우감소의 6구간으로 표현한다. <표 3>은 이 표현체계로의 전환기준을 보여주고 있다.



<표 3>에서 나타난 바와 같이 주가수익률이 증가일 경우 첫번째 비트(bit)는 항상 1이며, 감소의 경우 마지막 비트가 항상 1이 되게끔 되어 있다.

<표 3> 출력값의 전환기준

출력코드	의 미	주가수익률 증감
0 0 0 1 1	매우 감소	-0.26 이하
0 0 1 0 1	감 소	-0.09 ~ -0.25
0 1 0 0 1	조금 감소	-0.08 ~ 0.00
1 0 0 1 0	조금 증가	0.00 ~ 0.09
1 0 1 0 0	증 가	0.10 ~ 0.32
1 1 0 0 0	매우 증가	0.33 이상

이 코드체계는 출력값이 어떠한든 5개의 출력 처리단위(PE) 중 반드시 2개는 1을 나타내도록 해준다. 그리하여 학습된 인공신경망을 재현할 경우 신뢰할 수 있는 출력값을 얻은 것인지의 여부를 확인할 수 있게 해준다.

#### 4. 인공신경망 구축과 실행결과

인공신경망의 구축 방안은 매우 방대하다. 문제와 입출력자료의 특성에 따라 전혀 다른 종류의 인공신경망을 구축할 수 있다. 최근의 인공신경망 연구에서 가장 많이 활용되는 방안은 역전파 인공신경망이다. 그 이유는 가장 단순하며, 이해하기가 용이하기 때문이다. 본 연구에서도 역전파 신경망을 이용하기로 한다.

인공신경망의 종류를 선정한 다음, 필요한 것은 계층의 수와 각 계층의 처리단위의 갯수를 선정하는 것이다. ADALINE, MADALINE, 또는 Hopfield 신경망 등은 오직 한 개 또는 두 개의 계층으로 구성된다. 그러나 역전파(back-propagation) 인공신경망에서는 여러 개의 은닉 계층이 허용된다. 인공신경망에서 은닉계층은 입력자료로부터의 특성을 추출하는 역할을 한다고 볼 수 있다. 은닉계층의 증가는 인공신경망의 처리능력을 증가시킨다. 그러나 반대로 신경망의 학습을 복잡하게 한다. 즉 학습시 발생하는 오류의 추적을 어렵게 한다는 단점을 갖고 있다.

본 연구에서는 가장 단순하면서도 많이 활용되고 있는 1개의 은닉계층과 입출력계층으로 구성된 신경망을 선정했다. 입력계층은 16개의 처리단위로 구성된다. 여기서 9개는 특정사에 국한된 입력자료를 처리하기 위함이고, 나머지 7개는 특정주식에 관계없이 주식시장 전체에 관련된 입력자료를 처리하기 위한 것이다. 은닉계층은 8개의 처리단위로 구성되어 있다. 출력계층은 앞에서 지정한 바와 같이 5개의 처리단위로 구성되어 있다.

각 계층의 처리단위가 정해지면, 계층별 처리단위간의 연결에 관한 결정을 해야한다. 이 연결방안에도 매우 다양하다. 보통의 연구에서는 각 계층간을 모두 연결하는 방식을 취한다. 그러나 본 인공신경망

입력계층의 처리단위가 전혀 다른 성격의 두 집단으로 구성되어 있으므로, 입력계층과 은닉계층간을 모두 연결할 경우 은닉계층의 본래 기능을 제대로 발휘할 수 없을 지도 모른다. 즉 성격이 다른 두 집단의 혼합으로 인해 입력자료의 특성 추출이 오히려 불가능하게 될 가능성도 있다는 것이다.

그리하여 본 연구의 인공신경망은 은닉계층의 4개 처리단위는 9개의 입력처리단위 (즉 특정사의 재무비율자료 등)와 연결짓고, 또한 나머지 4개의 처리단위는 7개의 경제 및 주식시장 자료의 입력단위와 연결지우는 방식을 취한다. 그리고 은닉계층과 출력계층간의 처리단위는 모두 연결되게끔 설계되었다. <그림 1>은 본 연구에서 설계된 인공신경망의 구조를 보여주고 있다. <그림 1>의 좌측하단의 사각형이 수렴을 촉진하기 위한 Bias 처리단위이다.

구성된 인공신경망의 계층별 처리단위의 기능을 살펴보면 <표 4>와 같다.

<표 4> 처리단위별 학습 함수 및 규칙

계층	입력계층	은닉계층	출력계층
학습함수	선형	Sigmoid	Sigmoid
합계방식	Sum	Sum	Sum
학습규칙	-	Delta Rule	Delta Rule

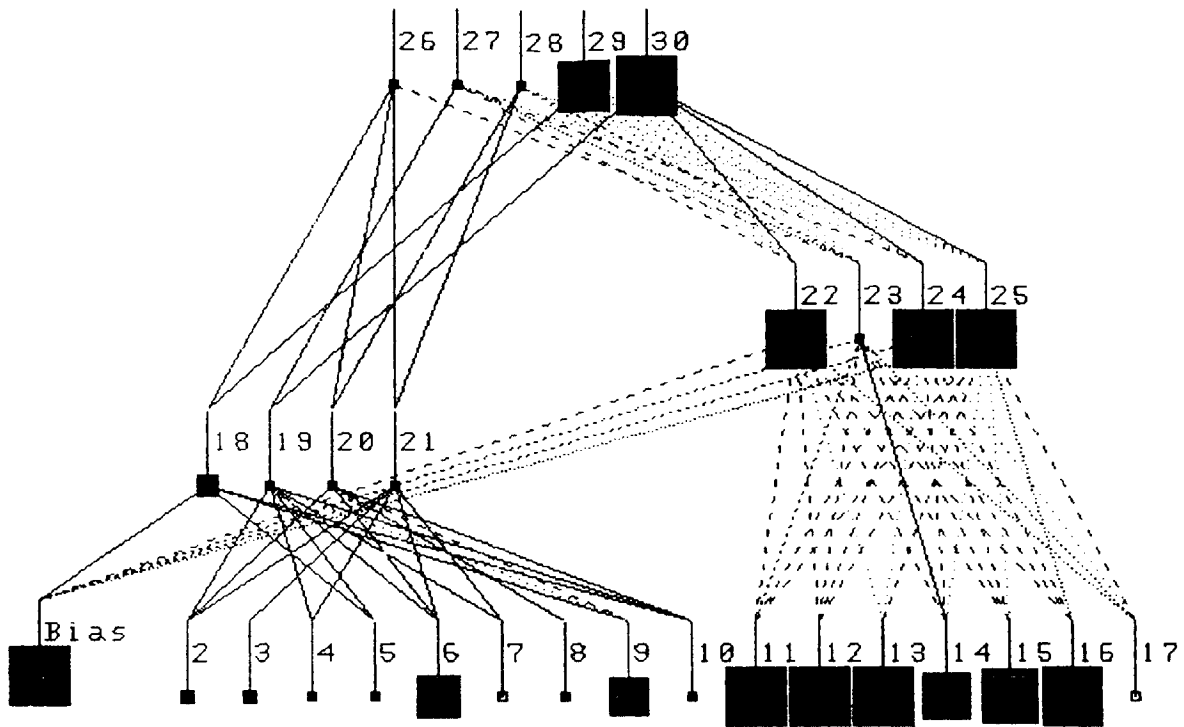
교육에 사용된 자료는 표본업체의 1989년과 1990년의 자료가 사용되었다. 모두 78

```
{{
  ) 한 업체당 2개의 교육자료가 포함되어야 하나, 일부 업체는 상장된 시점에 따라 오직 한 개 이하의 교육자료가 포함되어 있을 수도 있다. 한 예로 "진웅"의 경우 89년 5월에 상장되었기 때문에 1990년 자료만 교육에 포함되었다.
}}
```

개의 교육자료를 35만번 반복하였다. 교육의 반복횟수는 5만번 단위로 증가시켰으며, 35만번일때 가장 수렴효과가 높았다. 35만번 이상을 반복시킨 결과는 오히려 수렴효과가 떨어졌다.

교육을 시킨 결과는 <표 5>와 같다. 실제의 수익률을 정확하게 수렴한 경우가 전체의 72%가 되고, 수익률이 증가 또는 감소할 것인지의 여부를 맞게 수렴한 경우는 78%에 해당된다. 전체의 13%가 교육을 시켜도 수렴하지 못했다. 나머지 6%는 이미 약정되어 있는 출력코드 체계와 다른 형태로 수렴된 경우이다.

〈그림 1〉 인공신경망의 구조



NeuralWorks Professional II (tm)

Copyright (c) 1987-1989 by NeuralWare, Inc. All Rights Reserved.

〈표 5〉 교육결과의 내용

내용	자료수
일치 (실제의 수익률과 일치)	56
유사1(증감의 정도가 한 단위 틀림)	3
유사2(증감의 판단만 맞음)	4
상위 (증감의 판단 자체가 틀림)	10
불능 (출력형태에 대한 해석이 안됨)	5
합 계	78

여기서 교육수험이 잘 안된 이유를 찾아볼 필요가 있다. 일반적으로 교육자료가 골고루 분포되어 있을 경우에 교육수험 효과가 높다고 알려져 있다. 본 인공신경망의 교육자료의 경우별 분포와 수험결과는 〈표 6〉에 주어져 있다.

〈표 6〉 교육자료의 경우별 분포와 교육수험 결과

교육자료의 분류	일치	유사1	유사2	상위	불능	총계
00011	12	0	4	0	2	18
00101	22	2	0	1	0	25
01001	8	0	0	0	1	9
10010	2	1	0	4	1	8
10100	12	0	0	4	0	16
11000	0	0	0	1	1	2

첫 세개의 분류는 교육수험효과가 매우 높은 것으로 나타났다. 그러나 나머지 세 분류의 경우 수험효과가 떨어짐을 알 수 있다. 이는 상대적으로 교육자료의 수가 부족함이 기인하는 것 같다. 이 점은 교육자료가 골고루 분포되어 있을 경우에 교육수험효과가 높아진다는 사실을 다시 입증하고 있다.

이렇게 구축된 인공신경망을 Hold-out 표본에 적용하여 예측효과를 살펴보았다. 예측에 사용된 자료는 총 26개 자료가 사용되었다. 이들은 1991년 자료 중에서 각 자료 종류별로 5개씩 무작위 선정되었다. 그러나 자료종류 10010 과 10100 은 1991년 자료 집합중에서 각 2 개와 4개 밖에 존재하지 않았기 때문에 이들만 예측자료에 포함되었다. 〈표 7〉은 자료 종류별 예측자료에 포함된 개수를 나타내주고 있다.

〈표 7〉 자료종류별 예측자료의 수

자료 종류	자료의 수
-------	-------

00011	5
00101	5
01001	5
10010	2
10100	4
11000	5
계	26

〈표 8〉은 예측결과를 요약된 형태로 보여주고 있다. 수익률을 정확하게 예측한 경우는 전체의 50%이고, 증감의 정도가 한 단위만 틀린 경우가 19%이다. 그리하여 수익률의 증가 또는 감소의 방향을 제대로 예측한 경우는 전체의 69%에 해당하는 셈이다. 그러나 전체의 23%가 증감의 판단 자체가 틀린 경우이며, 또한 8%는 출력형태에 대한 해석이 불가능한 경우이다.

〈표 8〉 예측결과의 요약

내용	자료수	백분율
일치 (실제의 수익률과 일치)	13	50%
유사1(증감의 정도가 한 단위 틀림)	5	19
유사2(증감의 판단만 맞음)	0	0
상위 (증감의 판단 자체가 틀림)	6	23
불능 (출력형태에 대한 해석이 안됨)	2	8
합 계	26	100

〈표 9〉은 예측에 사용된 자료의 분포와 예측결과를 보여주고 있다. 교육수련의 결과에서 보듯이 마지막 세 분류의 자료에서 예측성고가 낮음을 알 수 있다. 그리하여 본 인공지능망의 연구에 있어 교육자료의 편중이 교육 및 예측효과에 지장을 초래한다고 볼 수 있다. 이는 오직 2개년간의 교육자료만을 대상으로 했기 때문이다.

〈표 9〉 예측자료의 분포와 예측결과

예측자료의 분류	일치	유사1	유사2	상위	불능	총계
00011	5	0	0	0	0	5
00101	5	0	0	0	0	5
01001	2	3	0	0	0	5
10010	0	1	0	1	0	2
10100	1	0	0	2	1	4
11000	0	1	0	3	1	5
총계	13	5	0	6	2	26

## 5. 판별분석과의 비교

특정 예측기법의 효용성은 다른 기법을 이용한 예측결과와의 비교를 통해 검증될 수 있다. 기존 연구는 인공신경망의 결과를 판별분석(Discriminant Analysis), logit 분석, 회귀분석, 귀납적 추론 등의 결과에 비추어 검증되었다 [Han, et al. 1993, Kim, 1992].

구체적으로 어떠한 분석결과와 비교될 것인가의 여부는 문제의 성격과 자료의 속성에 따라 다르다. 종속변수가 범주형 자료값을 취하는 경우에는 판별분석의 결과와 가장 많이 비교된다고 말할 수 있다.

판별분석의 구체적 실시 방안은 매우 다양하다. 본 연구에서는 단계별(stepwise) 변수 선정방법이 사용되었다. 인공신경망에서와 같은 동일한 교육자료로 대상으로 분석한 결과 16개의 독립변수중 판별분석에 포함된 변수는 모두 6개이다. <표 10>은 판별분석 모형에 포함된 변수를 해당 Wilk's Lambda 값과 함께 포함된 순서대로 나타내주고 있다.

<표 10> 판별분석 모형에 포함된 변수

변 수	포함된 순서	Wilk's Lambda	유의수준
추가수익비율	1	.49412	.0000
총자본경상이익률	2	.40076	.0000
순이익증가율	3	.33514	.0000
산업대비주가수준	4	.27125	.0000
자기자본비율	5	.23529	.0000
베타	6	.21801	.0000

도출된 판별분석 모형을 예측자료에 적용하여 얻은 예측결과는 <표 11>과 같다. 여기에 사용된 예측자료는 인공신경망 분석에 사용된 동일한 자료이다.

<표 11> 예측자료의 분포와 예측결과

예측자료의 분류	일치	유사1	유사2	상위	불능	총계
00011	5	0	0	0	0	5
00101	0	5	0	0	0	5
01001	0	1	4	0	0	5
10010	0	0	0	2	0	2
10100	0	0	0	4	0	4
11000	0	0	0	5	0	5
총계	5	6	4	11	0	26

예측기법과의 비교평가를 위해 비모수통계 기법이 적용되었다. 예측기법의 예측력에는 차이가 없다는 점을 귀무가설로 하여 <표 12>와 같은 contingency table을 작성하였다. 예측 결과는 크게 3가지로 구분하였다. 즉 정확하게 예측한 경

우, 정확하지는 않으나 증감의 방향만 맞은 경우 (유사한 예측), 그리고 예측불능이거나 정확하지 못한 예측 (상이한 예측) 으로 구분하였다. 이를 토대로  $\chi^2$  검정을 실시한 결과, 유의수준 10% 수준에서 귀무가설을 기각할 수 있다.

〈표 12〉 두 예측기법 실행결과의 비교

	정확한 예측	유사한 예측	상이한 예측	합계
인공신경망	13	5	8	26
판별분석	5	6	15	26
합계	18	11	23	

$$\chi^2 = 5.77690 \quad p=0.0557$$

## 6. 연구의 요약 및 추후 연구방향

본 연구의 목적은 주가수익률 예측의 문제에 인공신경망을 적용하고, 그에 따른 결과를 분석하고, 나아갈 방향을 제시하는 것이다. 1989년과 1990년의 자료를 대상으로 교육시킨 신경망을 1991년의 자료를 대상으로 그 예측력을 분석해 보았다.

그 결과 수익률을 정확하게 예측한 경우는 전체의 50%, 수익률의 증가 또는 감소의 방향을 제대로 예측한 경우는 전체의 69%에 해당한다. 그러나 전체의 31%가 증감의 판단 자체가 틀리거나, 예측결과를 해독할 수 없는 경우이다. 이러한 결과는 앞으로 개선될 여지가 많음을 말해주고 있다.

개선의 방향으로서는 우선 입력자료의 분포에 유의해야 한다는 점을 지적할 수 있다. 입력자료가 골고루 분포되어 있지 않을 경우 이에 따른 교육효과도 낮을 뿐만 아니라 예측효과도 지장을 받을 가능성이 있음을 본 연구는 실증적으로 보여주고 있다.

또한 본 연구는 입력자료의 성격이 전혀 다른 두 개의 집단으로 구성되어 있을 경우 은닉계층과 입력계층 처리단위간의 연결을 분리시킴으로써 효과적인 인공신경망을 구축할 수 있다는 점을 제시했다.

본 연구의 한계점으로는 우선 입력자료를 2년으로 국한함으로써 교육 효과에 영향을 미쳤다는 점을 지적할 수 있다. 이는 앞으로 보다 풍부한 교육자료를 대상으로 재연구함으로써 입력자료에 따른 신경망 효과를 보다 상세하게 설명할 수 있으리라고 본다.

또한 본 연구에서는 은닉계층이 하나인 역전파신경망 한가지 만을 조사했다. 앞으로는 다양한 종류의 구조를 시험해 봄으로써 주식수익률 예측문제에 적합한 구조를 선정할 수 있으리라고 본다.

## 참고문헌

- 노종래, 이종호, "신경회로망을 이용한 종합주가지수의 변화율 예측", 제1회 인공지능, 신경회로망 및 퍼지시스템 종합학술대회 논문집, 1991.
- 이재유, "인공신경망을 이용한 주가예측", 중앙대학교 대학원 석사학위논문, 1993.
- 조지호, "주식투자 수익률의 예측에 관한 연구", 증권학회지, 제11집, 1989.
- 한국상장회사협의회, "상장회사총감", 1993.
- Chen, N. and R. Roll, "Economic Forces and the Stock Market", Journal of Business, V. 59, 1986.
- Edwards, R. D. and John Magee, Jr., "Technical Analysis of Stock Trends", 5th Edition, Stock Trends Service, 1966.
- Francis, Jack "Investments: Analysis and Management", 3rd Edition, McGraw Hill, 1980.
- Han, Ingoo, Youngsig Kwon, and Hongkyu Cho, "Performance of Artificial Intelligence Models in Business Classification", 한국전문가시스템학술대회 논문집, 1993.
- Ken-ichi Kamijo and Testsuji Tanigawa, "Stock Price Pattern Recognition -- A Recurrent Neural Network Approach", INNS, Vol. 1, 1990.
- Kim, J., "A Comparative Analysis of Rule-Based, Neural Networks, and Statistical Classification Systems for the Bond Rating Problem", Unpublished Doctoral Dissertation, Virginia Commonwealth University, 1992.
- Nelson, Marilyn and W.T. Illingworth, "A Practical Guide to Neural Nets", Addison Wesley, 1991.
- Lapedes, A. & Farber, R. "Non-Linear Signal Processing using Neural Networks: Prediction and System Modeling", Los Alamos National Laboratory Report, LA-UR-87-2662, 1987.
- Quinlan, J., "Discovering rules by Induction from Large Collection of Examples", in D. Michie (ed), Expert Systems in the Micro Electronic Age, Edinburgh University Press, 1979.
- Takashi Kimoto, Kazuo Asakawa, Morio Yoda, and Masakazu Takeoka, "Stock Price Prediction System with Modular Neural Networks", International Joint Conference on Networks, 1990.
- Yoon, Youngoho & George Swales, "Predicting Stock Price Performance: A Neural Network Approach", Proceedings of the Twenty-Fourth Annual Hawaii International Conference on System Science. Vol.1, 1991.