



## 신경 회로망과 통계적 기법을 이용한 종합주가지수 예측 모형의 개발

Development of the KOSPI (Korea Composite Stock Price Index) forecast model using neural network and statistical methods

---

저자 (Authors)	이은진, 민철홍, 김태선 Eun Jin Lee, Chul Hong Min, Tae Seon Kim
출처 (Source)	<a href="#">전자공학회논문지-CI 45(5)</a> , 2008.9, 95-101(7 pages) <a href="#">The Institute of Electronics Engineers of Korea - Computer and Information 45(5)</a> , 2008.9, 95-101(7 pages)
발행처 (Publisher)	<a href="#">대한전자공학회</a> The Institute of Electronics and Information Engineers
URL	<a href="http://www.dbpia.co.kr/journal/articleDetail?nodeId=NODE01059474">http://www.dbpia.co.kr/journal/articleDetail?nodeId=NODE01059474</a>
APA Style	이은진, 민철홍, 김태선 (2008). 신경 회로망과 통계적 기법을 이용한 종합주가지수 예측 모형의 개발. 전자공학회논문지-CI, 45(5), 95-101
이용정보 (Accessed)	승실대학교 203.253.***.153 2020/09/29 18:06 (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.

논문 2008-45CI-5-12

# 신경 회로망과 통계적 기법을 이용한 종합주가지수 예측 모형의 개발

(Development of the KOSPI (Korea Composite Stock Price Index)  
forecast model using neural network and statistical methods )

이 은 진\*, 민 철 홍\*\*, 김 태 선\*\*\*

(Eun Jin Lee, Chul Hong Min, and Tae Seon Kim)

## 요 약

주가지수는 경제 및 정치적 상황을 포함한 다양한 주변 환경에 영향을 받는 관계로 정확한 주가지수 예측모형의 개발은 매우 어려운 문제로 여겨지고 있다. 본 논문에서는, 신경회로망과 통계적인 방법을 이용하여 종합주가지수(KOSPI)를 예측하는 에이전트 시스템 기법을 제안한다. 예측오차의 평균 및 편차를 최소화하기 위해서, 에이전트시스템은 특징추출, 변수선정, 예측엔진선정 및 분석을 위한 부(sub)에이전트 모듈들을 포함하고 있다. KOSPI(Korea Composite Stock Price Index) 예측을 위한 에이전트시스템 구현의 첫 번째 단계로서, 주성분분석을 이용하여 22개의 표준기본경제지표에서 12개의 경제지표를 추출하였다. 열두 개의 추출된 경제지표들은 예측하고자하는 예측일에 따라 최량부분적합법을 이용하여 다시 한 번 입력 변수들을 선정하게 된다. 성능평가를 위해 주가지수의 변동폭이 다른 두 종류의 실험데이터를 대상으로 예측을 진행한 결과 30일의 연속적인 종합주가지수예측에 있어 11.92포인트의 평균오차를 보였다. 또한, 예측시점에 따라 관련이 높은 기본지표의 종류 및 개수가 다르게 나타나므로 제안한 주가예측 에이전트시스템 구조가 유용함을 보였다.

## Abstract

Modeling of stock prices forecast has been considered as one of the most difficult problem to develop accurately since stock prices are highly correlated with various environmental conditions including economics and political situation. In this paper, we propose a agent system approach to predict Korea Composite Stock Price Index (KOSPI) using neural network and statistical methods. To minimize mean of prediction error and variation of prediction error, agent system includes sub-agent modules for feature extraction, variables selection, forecast engine selection, and forecasting results analysis. As a first step to develop agent system for KOSPI forecasting, twelve economic indices are selected from twenty two basic standard economic indices using principal component analysis. From selected twelve economic indices, prediction model input variables are chosen again using best-subsets regression method. Two different types data are tested for KOSPI forecasting and the prediction results showed 11.92 points of root mean squared error for consecutive thirty days of prediction. Also, it is shown that proposed agent system approach for KOSPI forecast is effective since required types and numbers of prediction variables are time-varying, so adaptable selection of modeling inputs and prediction engine are essential for reliable and accurate forecast model.

**Keywords :** 신경 회로망, 회귀분석, 종합주가지수, 에이전트시스템

\* 학생회원, \*\*\* 정회원-교신저자, 가톨릭대학교 정보통신전자공학과

(Dept. of Information, Communications and Electronics Eng., The Catholic University of Korea)

\*\* 학생회원, 가톨릭대학교 컴퓨터공학과

(Dept. of Computer Science and Eng., The Catholic University of Korea)

※ 본 연구는 2008년도 가톨릭대학교 교비연구비의 지원에 의해 수행된 과제입니다.

접수일자: 2008년8월20일, 수정완료일: 2008년9월5일

## I. 서 론

오늘날 종합주가지수는 경제의 고도발달과 선물이나 옵션, 펀드 등 다양한 금융상품이 개발되면서 많은 사람들이 관심을 가지게 되었다. 그러나 현재 주식시장은 다양한 주가 변동 요인에 의해서 불규칙한 변화를 가지며, 주가 지수와 이를 예측하는 요소들 간의 시간이 지남에 따라 구조적 관계도 불규칙하게 변하는 것이 일반적이다<sup>[1]</sup>. 이러한 주식시장에서의 주가 예측은 다양한 방법으로 연구되고 있다. 최근 시간이 지남에 따라 요소들 간의 구조적 불안 관계와 비선형적인 문제를 해결하기 위해 다양한 분석 방법이 제시되고 있다. 실례로 회귀분석과 VAR 분석을 이용하여 순 매각 비율로 주가지수 수익률 관계를 월별, 분기별, 일별로 분석하여 비슷한 예측률을 보인 연구<sup>[2]</sup>, 주가변동경향에 대한 확률적인 예측시스템을 구현한 연구<sup>[3]</sup>와 선형 회귀모형과 비선형 인공지능망 모형을 이용하는 방법을 비교분석하여 선형 회귀모형에 비해 신경망 모형의 예측오차분산이 감소됨을 보인 연구<sup>[4]</sup>도 진행되었다. 그러나 주가변동과 관련 높은 변수들의 선별을 위한 과정 없이 재무제표의 정보만을 이용하여 예측률이 만족스럽지는 못했다<sup>[4]</sup>. 또한, 주식의 상한가치 잔량과 일일 거래량의 관계를 증명하기 위해 회귀분석을 통해 주가의 단기예측을 잔량비와 수익률 간의 유의적 상관관계가 있음을 증명한 연구<sup>[5]</sup>도 있었고, 최근에는 유전자 알고리즘(GA)과 신경망을 결합하여 상장사의 주가를 예측하는 연구<sup>[6]</sup>도 진행되었다. 이 연구에 따르면 유전자알고리즘을 통해 경제 지표변수와 회사의 재무제표 중의 일부를 선출하고 신경회로망을 이용하여 주가 예측모형을 구현하

였다<sup>[6]</sup>. 또한 산업별 주가지수를 이용하여 변수를 선정하고 그 변수를 기반으로 퍼지 시스템을 이용한 예측모델 연구<sup>[7~8]</sup>도 진행되고 있다. 이처럼 적합한 변수선택과 주가의 예측률을 높이기 위한 연구가 다양하게 진행되고 있지만 주가지수에 영향을 주는 많은 요소와 불규칙한 주가지수의 예측모형을 구현하기란 쉽지 않다. 또한 적합한 변수선택이 되었다 하더라도 경제규모와 경제상황이 변함에 따라 주가지수에 영향을 주는 기본경제지표가 변화하므로 주가지수 예측모델은 경제상황과 날짜, 시간, 주가지수의 낙폭 등에 따라 새로운 기본 경제지표 변수가 선택되어야 한다. 본 논문에서는 그림 1.에 에이전트 시스템과 같이 경제 상황, 예측날짜, 주가지수의 낙폭 등에 따른 기본 지표 변수 선택과 선택된 기본 지표 변수로부터 추론되는 예측결과를 선택적으로 활용할 수 있는 주가예측 에이전트 시스템 구조를 제안한다. 제안하는 에이전트시스템은 예측오차의 평균 및 산포의 최소화를 위하여 예측시점의 상황 및 요구에 따라 적절한 처리가 가능하도록 특징추출, 변수선정, 예측엔진선정 및 분석을 위한 부(sub)에이전트 모듈들을 포함하고 있다. 본 논문에서는 예측방법에 있어 상하패턴의 예측이 아닌 예측일의 종합주가지수를 예측하였다. 또한 종합주가지수에 대해 영향을 미치는 변수를 추출하기 위한 분석으로서 다변량 분석 중 하나인 주성분분석과 최량 부분 적합을 식별하는 회귀분석 방법을 통하여 직접 가치 있는 변수들을 추정해 내는 방법을 이용하였다. 이렇게 종합주가지수에 미치는 영향이 변수들을 선별하여 신경 회로망과 다중 회귀분석을 이용하여 종합주가지수를 예측하고 비교분석하였다.

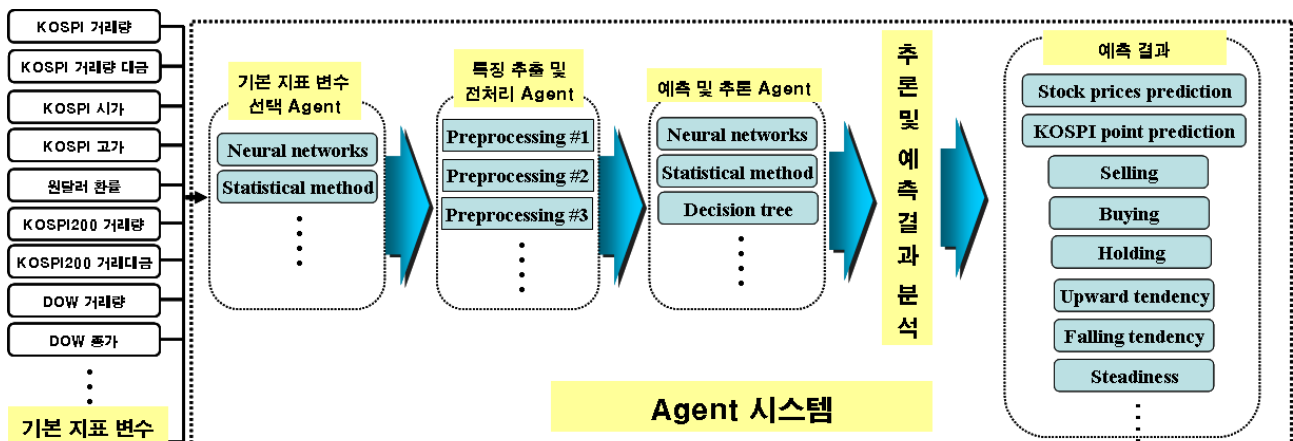


그림 1. 주가예측 모델의 에이전트 시스템 구조

Fig. 1. Agent system structure for stock index prediction model.

## II. 본 론

### 1. 데이터 구성

본 논문에서는 선행연구들을 통하여 일반적으로 사용되는 종합주가지수에 영향을 끼치는 기본 지표 변수 22가지를 선택하여 표 1.에 표시하였다<sup>[1,9]</sup>. 예측을 위한 데이터 기간은 약 12개월로 일별 종합주가지수를 이용하였다. 선택한 전체 자료 중 종합주가지수의 변동이 큰 기간 167일(2006년 10월 19일~2007년 6월 10일)과 변동이 작은 기간 116일(2006년 10월 19일~2007년 4월 5일)로 각각 나누어 구성하였으며, 전체 자료의 70%를 학습을 위한 Training data로 사용하였고 나머지 30%를 Test data로 구성하였다. 종합주가지수의 예측날짜는 현재 시점  $t$ 일로 부터  $t+1$ 일과  $t+3$ 일로 정하였다. 표 1.은 다양한 선행연구들로부터 선별된 주가에 영향을 미치는 기본 지표 변수들이다.

표 1. 입력 기본 지표 변수

Table 1. Standard basic Input variables for KOSPI forecast.

번호	지표 변수	번호	지표 변수
1	KOSPI 거래량	12	KOSPI200 거래량
2	KOSPI 거래대금	13	KOSPI200 거래대금
3	KOSPI 시가	14	KOSPI200 시가
4	KOSPI 고가	15	KOSPI200 고가
5	KOSPI 저가	16	KOSPI200 저가
6	DOW 거래량	17	KOSPI200 종가
7	DOW 시가	18	NASDAQ 거래량
8	DOW 고가	19	NASDAQ 시가
9	DOW 저가	20	NASDAQ 고가
10	DOW 종가	21	NASDAQ 저가
11	원달러 환율	22	NASDAQ 종가

### 2. 변수선택

#### 가. 주성분 분석을 통한 변수 선택

주성분 분석은 여러 개의 변수로부터, 전체 변동을 잘 설명할 수 있는 중요한 몇 개의 성분을 얻는데 목적이 있어 적은 수의 변수를 만드는데 유용하다. 따라서 표 1.과 같이 선별된 22가지 기본 지표 변수들 중 실제 종합주가지수에 가장 큰 영향을 주는 기본 지표 변수를 선별하기 위해 다양한 에이전트 시스템 중에서 주성분 분석이 선택되었을 경우를 가정하여 실험하였다. 주성분 분석으로 추출된 기본 경제 지표 변수는 총 12가지

표 2. 주성분 분석을 이용한 추출 변수

Table 2. Selected variables from principal component analysis.

번호	주성분 분석으로 추출된 변수
1	KOSPI 저가
2	KOSPI 고가
3	KOSPI 시가
4	KOSPI200 저가
5	KOSPI200 시가
6	KOSPI200 고가
7	KOSPI200 종가
8	NASDAQ 고가
9	DOW 종가
10	DOW 저가
11	DOW 시가
12	환율

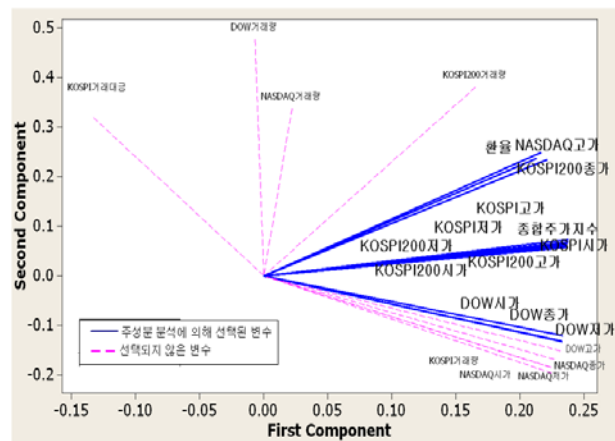


그림 2. 주성분 분석의 적재 그래프

Fig. 2. Loading plot results from principal components analysis.

로서 표 2.와 같으며, 추출된 기본 지표 변수를 이용하여 종합주가지수의 예측모델 자료로 사용하였다. 그림 2.은 주성분 분석 결과로 나타난 적재 그래프이다. 주성분 분석에서 나타난 그림에서는 서로 적은 각도차이를 가지면 비슷한 변동을 하고 있는 것이다.

#### 나. 회귀분석(최량 부분 적합)을 통한 변수 선택

본 논문에서는 주성분 분석으로 추출된 12가지 변수 중 예측하고자하는 예측일과 종합주가지수의 낙폭에 따라 선택되어지는 변수와 변수개수를 확인하고 선택된 변수에 따른 주가 예측률을 확인하기 위해 표 3.과 같이 최량 부분 적합을 통하여 변수들을 선정하였다. 회귀 분석 방법 중 최량 부분 적합은 예측 변수에서 최적의 모형을 식별하는 방법으로 전체 변동 중 회귀 식에

표 3. 최량 부분 적합을 이용한 추출 변수  
Table 3. Chosen variables from best-subsets regression.

순서	총 DATA 일수	Training 일수	Test 일수	예측일	최량 부분 적합으로 선택된 변수
1	167일	116일	51일	1일 후	1, 4, 7, 8, 9, 11 1, 4, 7, 9, 11, 12 1, 4, 7, 9, 10, 11, 12
2	167일			3일 후	3, 5, 7, 8, 10, 11, 12 3, 4, 5, 7, 10, 11, 12 3, 5, 7, 8, 10, 12
3	116일	81일	35일	1일 후	2, 3, 5, 6, 7, 10, 11 2, 5, 7, 10, 11 2, 3, 5, 6, 7, 10, 11, 12
4	116일			3일 후	3, 5, 7, 10, 11 3, 5, 7, 10, 11, 12 3, 4, 5, 7, 10, 11, 12

의해 설명되는 변동이 높은 R-제곱 값과 적합치의 평균 제곱 오차와 관련이 있는 작은 Mallows C-p통계치를 이용하여 추출하였다.

데이터의 구성은 종합주가지수의 변동 폭이 큰 기간(167일)과 변동 폭이 작은 기간(116일)의 두 그룹으로 구분하여 실험하였으며, 현재 시점 기준 t일로부터 다음 날인 t+1일과 3일 후인 t+3일을 구분하였다. 표 3.은 최량 부분 적합을 이용한 추출 변수이며 최량 부분 적합으로 선택된 변수의 숫자는 표 2.에서 주성분 분석으로 추출된 12가지 변수 숫자와 동일하다.

### III. 실험

#### 1. 신경 회로망을 이용한 주가예측 모델

신경회로망은 인간의 정보처리 패턴과 유사한 방식의 시스템으로 비선형적 문제를 해결 할 수 있을 뿐만 아니라 많은 데이터양의 자료를 처리할 수 있으며, 규칙성을 가지지 않는 특이한 경우들도 학습할 수 있게 함으로써 보다 일반적인 보편화된 예측모형으로 연구되고 있다. 신경회로망은 응용범위와 목적에 따라 다양한 종류를 가지고 있으나, 본 논문에서는 오류 역전파 신경회로망(feedforward error backpropagation neural network)을 이용하여 주가예측 모델을 구현하였다. 출력된 신호는 전달함수에 따라 다양한 형태로 신호를 출력하게 되는데 본 실험에서는 다양한 결과 값을 출력하기 위해서 sigmoid 전달함수를 사용하였으며, 신경 회로망의 학습에서는 입력 데이터와 목표 데이터를 페구간 [0.1, 0.9]사이의 실수로 정규화 하였다. 입력 데이터에는 주성분분석과 최량 부분 적합을 이용하여 종합주

가지수에 크게 영향을 미친다고 판단되는 변수그룹들을 선출하였다. 예측 방법은 종합주가지수의 변동이 큰 기간과 변동이 작은 기간 각각의 t+1일과 t+3일, 총 4가지 경우에 대하여 각각 3가지씩의 변수그룹을 선출하여 총 12가지의 예측 모델을 구현하였다[표 3참조]. 그리고 예측 모델의 성능 평가를 위해서 본 논문에서는 식 (1)과 같이 RMSE(Root Mean Squared Error)를 이용하였다. 변수  $N$ ,  $X_k^t$ ,  $X_k^p$ 는 각각 데이터의 수, 목표 값, 측정값을 나타낸다.

$$RMSE = \sqrt{\frac{\sum_{k=s} (X_k^t - X_k^p)^2}{N}} \quad (1)$$

#### 2. 회귀 분석을 이용한 주가예측 모델

회귀 분석은 한 변수 혹은 여러 변수가 다른 변수에 미치는 영향력의 크기를 수학적 관계식으로 추정하고 분석하는 방법으로 그 접근 방법으로 단순회귀, 다중회귀 등의 방법이 있으며, 종속변수에 따라 로지스틱 회귀분석 등을 사용할 수 있다. 그 중 다중 회귀분석은 2개 이상의 예측변수와 1개의 반응변수와의 관계를 수학적으로 나타내는 것으로, 다중 회귀분석 예측 모델을 구현하기 위해서 입력 데이터와 출력 데이터는 신경 회로망을 이용한 예측 모델에서와 같이 전처리과정을 거치지 않고 실제 기본 지표 변수 값을 사용하였다. 데이터 구성은 신경 회로망을 이용한 주가예측 모델과 동일한 방법으로 변수그룹을 구성하여 총 12가지의 예측모델을 구현하였으며, 회귀 분석의 성능평가를 위해 RMSE 방법을 이용하였다.

#### 3. 예측결과 및 분석

본 논문에서는 신경 회로망과 회귀 분석을 같은 조건에서 예측 모델을 구현하였다. 표 4.는 신경 회로망을 이용하여 예측한 결과이고, 표 5.는 회귀 분석을 이용한 예측 결과이다. 표 4.의 결과를 보면 종합주가지수의 변동 폭이 큰 167일 데이터기간에서의 t+1일 예측 결과 중 최량 부분 적합으로 선택된 첫 번째 6개의 지표 변수가(1, 4, 7, 8, 9, 11) 다른 두 번째, 세 번째로 선택된 지표 변수들을 이용하여 예측한 결과들 보다 작은 Test RMSE 값을 보이는 것을 알 수 있었으며, 같은 데이터 기간 중 t+3일 예측 결과에서는 세 번째 선택된 지표 변수 6개(3, 5, 7, 8, 10, 12)에서 38.76[point]로 첫 번째,

표 4. 신경 회로망을 이용한 예측결과

Table 4. KOSPI forecasting results using neural network.

순서	총 DATA 일수	Training 일수	Test 일수	예측일	NN [point]	
					Training error	Test error
1	167일	116일	51일	1일 후 (t+1)	<b>14.61</b>	<b>25.74</b>
					16.16	47.59
					17.52	67.62
2	167일	116일	51일	3일 후 (t+3)	17.57	73.68
					17.82	49.70
					<b>17.24</b>	<b>38.76</b>
3	116일	81일	35일	1일 후 (t+1)	5.46	23.39
					6.24	20.85
					<b>5.20</b>	<b>17.41</b>
4	116일	81일	35일	3일 후 (t+3)	32.11	23.65
					32.06	29.42
					<b>31.87</b>	<b>23.45</b>

표 5. 회귀 분석을 이용한 예측결과

Table 5. KOSPI forecasting results using regression.

순서	총 DATA 일수	Training 일수	Test 일수	예측일	regression [point]	
					Training error	Test error
1	167일	116일	51일	1일 후 (t+1)	<b>13.81</b>	<b>12.84</b>
					12.08	21.34
					11.20	16.20
2	167일	116일	51일	3일 후 (t+3)	23.84	37.31
					19.57	26.18
					<b>19.95</b>	<b>25.61</b>
3	116일	81일	35일	1일 후 (t+1)	9.36	16.81
					9.89	13.54
					<b>10.82</b>	<b>11.92</b>
4	116일	81일	35일	3일 후 (t+3)	15.57	22.27
					16.89	31.92
					<b>18.25</b>	<b>19.39</b>

두 번째 선택된 지표 변수들을 이용한 예측한 결과들보다 작은 Test RMSE 값을 보였다. 반면, 종합주가지수의 변동 폭이 작은 기간 116일 중 t+1일의 주가를 예측한 결과를 보면 세 번째로 선택된 8개의 변수(2, 3, 5, 6, 7, 10, 11, 12)에서 17.41[point]로 가장 작은 Test RMSE 값을 보였으며, t+3일 예측 결과에서도 세 번째로 선택된 7가지 변수(3, 4, 5, 7, 10, 11, 12)에서 가장 작은 Test RMSE 값을 보였다.

표 5.의 회귀 분석 방법을 이용한 종합주가지수 예측결과를 보면 지수 낙폭이 큰 167일에서의 t+1일, t+3일 모두 신경회로망을 이용한 예측결과와 동일한 변수에서 각각 12.84, 25.64[point]로 가장 작은 Test RMSE 값을 보였으며, 지수 낙폭이 작은 116일에서도 신경회로망을 이용한 예측결과와 동일한 변수에서 각각 11.92, 19.38[point]로 가장 작은 Test RMSE 값을 보였다.

상기 실험에서는 종합주가지수의 변동 폭이 다른 기

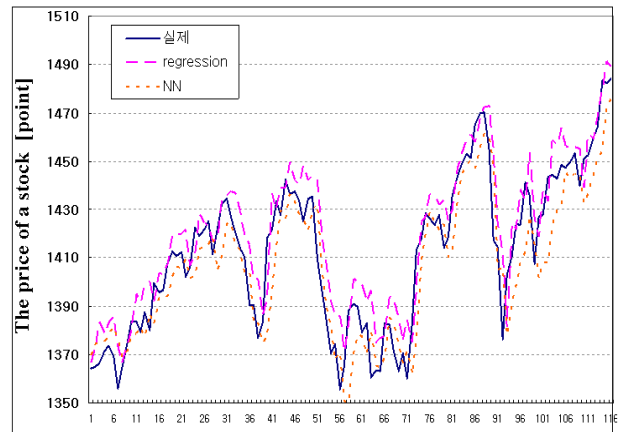


그림 3. 167일 데이터 중 t+1예측을 위한 Training 결과  
Fig. 3. Training results of 167 days data for t+1 day prediction.

간과 예측하고자하는 일수에 따라 기본 지표변수의 개수와 종류가 변하고 그에 따른 예측 결과도 다르게 나타나는 것을 알 수 있었다. 표 4.와 표 5.의 예측 결과에서 볼 수 있듯이 회귀 분석 방법이 신경회로망을 이용한 예측 결과보다 오차의 평균인 RMSE가 낮은 결과를 보인 것을 알 수 있다. 일반적으로 신경회로망은 비선형적 요소들의 학습 패턴에서 탁월한 학습능력을 발휘하여 다양한 예측 모델에 활용되고 있으나 본 실험에서는 통계적 방법으로 추출된 기본 지표 변수가 예측구간 내에서 종합주가지수의 변화와 선형적 관계가 있었음을 추정할 수 있다. 이러한 원인은 예측대상으로 삼은 기간의 주가지수가 단기간의 관점에서는 지속적인 등락이 있었지만 전체 추세로는 기형적으로 급상승하는 형태를 취한 것에 기인한 것으로 본다. 예측을 위한 전체기간에 걸쳐 회귀분석이 우수한 결과를 나타내었는지를 판단하기 위해서는 예측오차의 산포분석이 추가적으로 필요하므로 RMSE값의 크기만으로 예측모델의 우수성을 언급하기는 무리가 있다.

그림 3.는 종합주가지수의 변화폭이 큰 167일 중 t+1일을 예측한 신경회로망 예측모델과 회귀분석을 이용한 예측모델의 Training결과 그래프이다. 두 예측 모델의 Training RMSE 값은 각각 14.61, 13.81[point]로서 실제 종합주가지수 그래프와 동일한 경향을 갖으며 Training된 것을 볼 수 있다. 또한 그림 4.은 종합주가지수의 변화 폭이 큰 167일 중 t+1일의 신경회로망과 회귀 분석 방법을 이용하여 Test를 통하여 예측한 결과를 그래프로 표시한 것으로 각각의 Test RMSE 값은 25.74, 12.84[point] 이다.

그림 5.은 종합주가지수의 낙폭이 컸던 기간(167일)



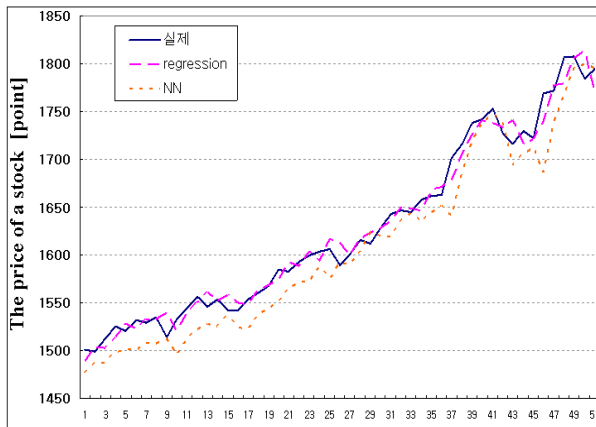


그림 4. 167일 데이터 중 t+1예측을 위한 Test 데이터  
Fig. 4. Test results of 167 days data for t+1 day prediction

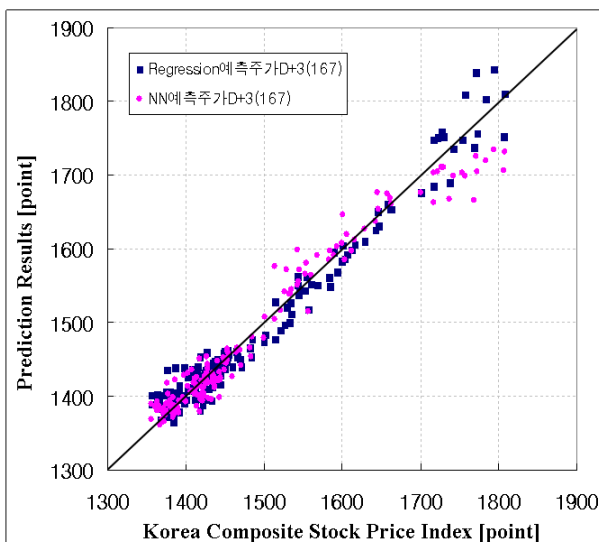


그림 5. 167일 데이터의 t+3 예측결과와의 데이터 산점도  
Fig. 5. t+3 day prediction scatter plot for 167 days data

에 대한 예측결과를 산점도로 표시한 그래프이다. 앞에서 언급한 바와 같이 예측모형의 지속적인 신뢰성 확보를 위한 성능분석을 위해서는 예측오차의 산포분석이 필요하다. 종합주가지수의 변동폭이 대체로 학습데이터와 유사한 부분에서의 분산은 회귀 분석과 신경회로망을 이용한 예측값에 대한 편차에 유의성이 없는 것으로 나타났다. 그러나 종합주가지수가 학습데이터보다 크게 커진 부분에서는 (1700~1900[point])에서는 회귀분석을 이용한 예측결과는 대체로 과대추정(over-estimate)되었고, 신경회로망을 이용한 경우에는 과소추정(under-estimate)된 것으로 보인다. 이는 종합주가지수가 1700[point]이하인 부분에서의 경향과는 차이가 있다. 따라서 비선형이 학습가능하다는 특징으로 기존의 통계적인방식에 비해 단기적인 주가지수의 변동예측에

유리했던 신경회로망도 상승 혹은 하락의 폭이 학습데이터들의 범위를 크게 벗어나는 경우에는 그 성능에 한계를 가질 수밖에 없고, 또한 예측시점에 따라 지속적으로 주가와 상관도가 높은 변수들을 선정하기 위해서는, 고정된 개수 및 형태의 입력변수 사용보다는 적응적으로 입력변수를 선택할 수 있는 시스템이 필요하다고 본다.

#### IV. 결 론

본 논문에서는, 신경회로망과 통계적인 방법을 이용하여 종합주가지수를 예측하는 에이전트 시스템 기법을 제안한다. 제안한 시스템은 예측오차의 평균 및 편차를 최소화하기 위해서, 22가지의 표준기본경제지표로부터 12개의 경제지표를 추출하였고, 열두 개의 추출된 경제지표들은 에이전트 시스템의 변수선정모듈에서 예측하고자하는 예측일의 특성에 따라 최량 부분 적합법을 이용하여 입력 변수들의 개수 및 형태를 재선정하게 된다. 성능평가를 위해 주가지수의 변동폭이 다른 두 종류의 실험데이터를 대상으로 예측을 진행한 결과 30일의 연속적인 KOSPI 예측에 있어 11.92[point]의 평균오차를 보였다. 예측을 위한 엔진으로 회귀분석과 신경회로망을 적용한 결과 주가예측을 위한 시점 및 기간에 따라 그 성능이 달라짐을 알 수 있었다. 본 연구결과는 경제 상황, 예측날짜, 주가지수의 낙폭 등에 따른 기본지표 변수 선택과 선택된 기본 지표 변수로부터 추론되는 예측결과를 선택적으로 활용하여 향후 보다 실질적인 주가예측 에이전트 시스템을 구성을 위한 기초가 될 것으로 기대한다.

#### 참 고 문 헌

- [1] 이희춘, "KODEX200 주가 수익 예측에 대한 연구", 한국경영교육학회, 제37집, 345-358쪽, 2005.
- [2] 고광수, "주식형 펀드의 현금흐름과 주식시장", 한국증권학회학술지, 제31집, 71-107쪽, 2002.
- [3] 강병우, 한동수, "한국종합주가지수 변동 경향에 대한 확률적 예측 시스템", 한국정보과학회 가을 학술발표논문집, Vol. 33, No. 2(A), 500-504쪽, 2006.
- [4] 최성섭, 구형건, 김영건, "장부가치와 주당이익을 이용한 선형회귀 모형과 신경망 모형의 주가예측", 재무관리연구, 제17권, 제1호, 161-180쪽, 2000.
- [5] 오성민, 김성집, "주식의 상한가치 잔량과 일일거래량의 관계를 통한 주가의 단기예측에 관하여",

대한산업공학회 춘계공동학술대회 논문집, 457-460 쪽, 2000.

- [6] 김성환, “유전자 알고리즘과 신경망을 이용한 주가 패턴 예측 시스템”, 중앙대학교 대학원 전기공학과 석사학위논문, 2003.
- [7] 김해균, “지식정보에 기초한 퍼지시스템을 이용한 주가지수예측”, 부산대학교 대학원, 전기공학과 석사학위논문, 2001.
- [8] WEI YANG, “stock price prediction vased on fuzzy logic”, Proxceedings of the Sixth International Conference on Machine Learning and Cybernetics, Vol. 3, pp19-22, 2007. 8.
- [9] 강대훈, “판별분석, 인공신경망, SVM을 이용한 수익률 성능 비교 연구(KOSPI 지수)”, 연세대학교 대학원 학위논문, 2004. 2.

---

#### 저 자 소 개

---



**이 은 진**(학생회원)  
 2008년 가톨릭대학교 반도체  
 시스템공학과 학사 졸업.  
 2008년 현재 가톨릭대학교 대학원  
 정보통신전자공학과  
 석사과정.  
 <주관심분야 : 컴퓨터, 지능형시

스템, 반도체>



**민 철 흥**(학생회원)  
 2005년 가톨릭대학교 반도체  
 시스템공학과 학사졸업.  
 2007년 가톨릭대학교  
 컴퓨터공학과 석사졸업.  
 2008년 현재 가톨릭대학교 컴퓨터  
 공학과 박사과정.  
 <주관심분야 : 지능형 시스템, 반도체공정>



**김 태 선**(정회원)-교신저자  
 1991년 인하대학교 전기공학과  
 학사 졸업.  
 1993년 인하대학교 전기공학과  
 석사졸업.  
 1998년 Georgia Institute of  
 Technology, 전기 및  
 컴퓨터공학과 박사 졸업.

1999년~2001년 삼성전자 반도체 책임연구원.  
 2001년~현재 가톨릭대학교 정보통신전자공학부  
 부교수.

<주관심분야 : 지능형시스템, BIT융합기술, 지능  
 형 반도체 제조기술>