

# 일상어휘를 기반으로 한 선물 가격 예측모형의 개발

김광용\*, 이승용\*\*

## 요 약

본 논문은 인공지능경망과 귀납적 학습방법 등의 인공지능 방법과 선물가격결정에 대한 기존 재무이론을 사용하여 일상어휘로 표현되는 파생상품 가격예측 모형을 개발하는데 있다. 모형의 개발은 1단계로 인공지능경망이나 기존의 선물가격결정이론(평균보유비용모형이나 일반균형모형)을 이용하여 선물 가격을 예측한 후, 서로 비교분석하여 인공지능경망 모형의 우수성을 확인하였다. 귀납적 학습방법중 CART 알고리즘을 사용하여 If-Then 규칙을 생성하였다. 특히 실용적 측면에서 선물가격의 일상어휘화를 통한 모형개발을 여러 가지 방법으로 시도하였다. 이러한 선물가격 예측모형의 유용성은 일단 If-Then 규칙으로 표현되어 전문가의 판단에 확실한 이론적인 근거를 제시할 수 있는 장점이 있으며, 특히 의사결정지원시스템으로 활용될 경우 매우 유용한 근거자료로 활용될 수 있다. 이러한 선물가격 예측모형의 정확성은 분석표본과 검증표본으로 나누어 검증표본에서 세가지 기본모형(평균보유비용모형, 일반균형모형, 인공지능경망 모형)과 각 모형의 귀납적 학습방법 모형의 다른 3가지 어휘표현방법 3가지를 모형별로 비교 분석하였다. 분석결과 인공지능경망모형은 상당한 예측력을 갖고 있는 것으로 판명되었으며, 특히 CART를 기반으로 한 일상어휘기반의 선물가격예측 모형은 예측력이 높은 것으로 나타났다.

## 1. 문제제기

1996년 5월 3일 주가지수(KOSPI 200) 선물이 거래된 이후 선물시장가격은 보유비용모형을 이용한 이론가격에 비해 괴리율이 매우 높아 논란의 대상이 되고 있다. 주가지수선물시장이 개설될 때 많은 사람들은 선물가격이 현물지수보다 높게 형성될 것으로 예상하였다. 왜냐하면 보유비용모형에 의하면 순보유비용은 이자율에서 배당수익률을 뺀 값으로 결정되는데, 우리나라의 경

우 배당수익률이 이자율보다 낮은 관계로 선물가격이 현물지수보다는 높게 형성되어야 하기 때문이다. 그러나, 선물가격은 96년 6월말의 만기 이후 저평가 및 백워드이션(backwardation) 현상을 보이는 경우가 많았다.

주가지수선물의 가격괴리현상을 분석한 기존의 연구를 살펴보면, 주식과 선물거래의 과세차이[Cornell & French(1983)], 배당금의 불확실성[Peters(1985)], 새로운 시장에서의 일시적 현상[Figlewski(1984)], 차입금리와 대여금리의 차이[Gould(1988)], 선물가격과 선도가격의 차이[MacKinlay & Ramaswamy (1988)], 포지션 제한[Brennan & Schwartz(1990)] 등에 의해 가격괴리를 설명하고 있다. 그러나, Modest(1984), Cornell(1985), Yadav & Pope(1990) 등은 상기의 요인들이 선물가격에 그리 큰 영향을 미치지 않는다는 연구결과를 발표하였다. 더구나, Finnerty & Park(1988), Merrick(1989), 그리고 Yadav & Pope(1990)은 상기의 요인들이 서로 상쇄작용을 하여 차익거래불가영역을 오히려 좁히는 역할도 한다고 주장한다.

Modest & Sundaresan(1983)은 미국 S&P 500 지수선물시장에서 선물의 저평가현상은 공매도 대금의 100% 이용이 불가능하기 때문이라는 주장을 펴고 있다. 그들은 배당금과 거래비용을 고려하여 이론가격과 차익거래불가영역을 추정하고 결과 투자자가 공매도 대금을 사용할 수 없는 경우 차익거래 기회가 존재하지 않았으나, 공매도, 대금을 100% 사용할 수 있는 투자자에게는 차익거래 기회가 존재하였음을 보이고 있다. 이와 유사한 연구로 Puttonen & Martikainen(1991)은 공매도 대금의 이용제한 때문에 선물의 저평가현상을 검증할 때는 고평가현상을 검증할 때보다 추가적인 거래비용(이자비용)을 고려해야 한다는 주장을 펴고 있다. 또한, Brenner et. al.(1990)은 일본 주가지수선물시장에서 선물의 저평가현상의 상당부분은 거래비용으로 설명될 수 있음을 보여주고 있다. 한편, McKinlay & Ramaswamy(1988)는 미국 S&P500 지수선물시장에서 차익거래불가영역을 초과하는 가격괴리가 존재하였으며, 특히 선물계약의 만

\* 숭실대학교 경영학부 교수, Tel:820-0597, gygim@saint.soongsil.ac.kr

\*\* 숭실대학교 경영학과 대학원

기가 길수록 가격괴리가 증가함을 보여주고 있다.

이상에서 논의한 바와 같이 시장불완전성, 거래제도 및 제약 등이 주가지수 선물 가격 형성에 미치는 영향에 관한 논의는 미국과 일본의 경우를 중심으로 상당한 연구가 이루어져 왔다. 우리나라의 경우 이와 관련된 현안을 연구해 보는 것은 다음과 같은 이유에서 그 시사하는 바가 매우 크다고 할 수 있다. 첫째, 우리나라의 경우 기관 투자자는 주식의 공매도를 할 수가 없다. 1996년 9월 2일부터 '유가증권 대차거래제도'가 시행되었으나, 이 제도가 주식대여자 뿐만 아니라 차입자에게 별로 매력적이지 못한 측면이 있다. 첫째, 주식대여자는 매도물량으로 자신들이 보유하고 있는 주식의 미실현손실이 증가할 것을 우려하는 반면, 현행 유가증권대차거래제도상의 회계처리기준은 막대한 주식평가손을 안고 있는 기관 투자자들이 증권에탁권을 통하여 매도차익 거래를 하는 것을 꺼리게 하고 있다. 둘째, 증권회사나 은행, 투신, 보험 등 기관 투자자들은 보유주식을 매도함으로써 준차익거래를 실행할 수 있으나, 원가주의에 근거한 유가증권 회계처리제도로 인해 시장하락기에 보유주식의 매도(상품평가손의 실현)를 꺼리기 때문이다. 이와 같이 우리나라의 경우 공매도의 제약과 관련된 매도차익거래의 부재는 주가지수선물의 저평가현상을 지속시키는 주된 요인이 될 수 있음을 시사하고 있다.

주가지수선물의 가격결정모형 중 가장 잘 알려진 모형은 보유비용모형(cost of carry model)이다. 이 모형에 의하면 선물 포지션은 주식과 채권으로 복제될 수 있으며 차익거래불가논리에 의해 선물가격이 주가지수, 이자율 그리고 배당수익률에 의해 결정된다. 이와 같이 보유비용모형은 논리적 단순성으로 인해 주가지수선물뿐만 아니라 거의 모든 선물계약의 가격결정에 널리 이용되고 있다. 우리나라의 경우에도 주가지수선물의 가격결정을 논의할 때 보유비용모형의 채택을 당연시하고 있다.

KOSPI 200 선물의 시장가격이 보유비용모형의 이론가격과 상당한 괴리를 보일 때 우리는 가격결정모형에 오류가 있다고 생각

해 볼 수 있다. 예를 들어, 보유비용모형은 현물가격(KOSPI 200)이 외생적으로 주어진다고 가정하고 차익거래불가논리를 이용하는 부분균형모형이다. 따라서 현물시장과 선물시장의 동태적인 상호작용을 고려하지 못하는 단점을 지니고 있다. 또한 보유비용모형에서는 이자율과 현물가격의 변동성이 일정하다는 비현실적인 가정이 전제되어 있다.

시장가격이 보유비용모형이 제시하는 이론가격과 상당한 괴리를 보인다면 우리나라 주가지수 선물시장에 적합한 대체적인 모형은 무엇인지를 살펴보고자 하는 것이 본 논문의 내용이며, 특히 본 연구에서는 보유비용모형외에 일반균형모형과 인공신경망모형을 이용하여 우리나라 주가지수선물의 가격결정을 실증분석하고자 한다.

## 2. 보유비용모형

주가지수선물의 이론가격은 보유비용의 개념과 차익거래불가논리에 의해 결정할 수 있다. 거래비용 및 공매도에 대한 제약이 없고 차입금리와 대출금리가 동일한 완전시장의 가정 하에서 주가지수선물의 이론가격은 현물지수에 순보유비용(이자비용 - 배당수익)을 더하여 계산한다. 즉, 배당금을 배당탁일에 지급받는 것으로 의제한 전기 현금배당금에 근거하여 각 배당금의 만기시 가치를 구하고 연속복리를 가정하는 경우의 주가지수선물 이론가격( $F_t$ )은 다음과 같이 결정된다.

$$F_t = S_t \cdot e^{r(T-t)} - D(t, T) \quad (1)$$

이때  $S_t$ 는 주가지수,  $D(t, T)$ 는 시간  $t$ 부터 만기( $T$ )까지의 기간동안 받게 되는 배당금의 만기 가치, 즉,

$$D(t, T) = \sum_{s=t}^T e^{r \cdot (T-s)} D_s, \quad \text{이때}$$

$D_s$ 는  $s \in [t, T]$ 일에 지급되는 배당금의 액수,  $r$ 은 무위험수익률을 나타낸다.

### 3. 일반균형모형

Hemler & Longstaff (1991)은 이자율과 시장변동성이 확률적이라는 가정 하에 주가지수선물의 일반균형 가격결정모형을 제시하고 있다. 그들의 모형에서 주가지수선물가격은 주식시장, 선물시장, 그리고 자금시장간의 상호작용에 의해 결정된다. 특히, 주가지수선물가격, 주가지수, 이자율 등이 내생적으로 결정되며, 시장변동성도 선물가격에 영향을 미치게 된다. 즉,

$$F(S, r, \sigma, D(t, T), \tau) = [S - D(t, T)] A(\tau) \cdot \exp[B(\tau)r + C(\tau)\sigma] \quad (2)$$

이때  $F$ 는 주가지수 선물가격,  $S$ 는 주가지수,  $r$ 은 이자율,  $\sigma$ 는 주가지수 수익률의 변동성,  $D(t, T)$ 는 시간  $t$ 부터 만기( $T$ )까지의 기간동안 받게 되는 배당금의 만기가치, 즉,

$$D(t, T) = \sum_{s=t}^T e^{r \cdot (T-s)} D_s,$$

이때  $D_s$ 는  $s \in [t, T]$ 일에 지급되는 배당금의 액수, 그리고  $\tau$ 는 현재시점( $t$ )부터 선물 만기일( $T$ )까지의 시간을 나타내며,  $A(\tau)$ ,  $B(\tau)$ ,  $C(\tau)$ 는 모두  $\tau$ 의 함수이다.

Hemler & Longstaff (1991) 모형의 실증적 시사점중 가장 중요한 것은 “배당금조정 후 주가지수선물가격과 현물지수의 비율에 자연대수를 취한 값이 이자율 및 시장변동성과 선형회귀식으로 표현될 수 있다”는 것이다. 즉,

$$L_t = \beta_0 + \beta_1 r_t + \beta_2 \sigma_t + \varepsilon_t \quad (3)$$

이때  $L_t = \ln\left(\frac{F_t}{S_t - D(t, T)}\right)$ ,  $S_t$ 는 주가지수,  $D(t, T)$ 는 시간  $t$ 부터 만기( $T$ )까지의 기간동안 받게 되는 배당금의 만기가치, 즉,

$$D(t, T) = \sum_{s=t}^T e^{r \cdot (T-s)} D_s, \quad \text{이때 } D_s \text{는}$$

$s \in [t, T]$ 일에 지급되는 배당금의 액수,  $r_t$ 는 이자율, 그리고  $\sigma_t$ 는 시장변동성을 각각 나타낸다.

일반균형모형에 의하면 식 (3)의 추정계수가 다음과 같아야 함을 의미한다.  $\beta_0 \neq 0$ ,

$\beta_1 > 0$ ,  $\beta_2 \neq 0$ . 한편, 보유비용모형에 의하면 배당금조정 후 주가지수선물가격과 주가지수의 비율에 자연대수를 취한 값이 이자율에만 의존한다. 따라서, KOSPI 200 선물가격이 보유비용모형을 따른다면 식 (3)의 추정계수가 다음과 같은 제약조건을 만족시켜야 한다.  $\beta_0 = 0$ ,

$\beta_1 > 0$ ,  $\beta_2 = 0$ . 즉, 이자율은 선물가격의 형성에 영향을 미치더라도 시장의 변동성은 선물 가격에 어떠한 영향도 미치지 않아야 한다. 따라서, 보유비용모형은 일반균형모형의 특수한 형태임을 알 수 있다.

일반균형모형을 이용하여 이론가격을 얻기 위해서는 다음과 같은 단계를 거쳐야 한다.

(1) 1997년 7월 7일부터 1998년 7월 21일까지의 KOSPI 200 선물시장가격, KOSPI 200, 배당금, 이자율, KOSPI 200의 변동성을 이용하여 식 (2)의  $\beta_0$ ,  $\beta_1$ ,  $\beta_2$ 를 추정한다.

(2)  $\beta_0$ ,  $\beta_1$ ,  $\beta_2$ 의 추정계수  $\hat{\beta}_0$ ,  $\hat{\beta}_1$ ,  $\hat{\beta}_2$ 와 1998년 7월 22일의 KOSPI 200, 배당금, 이자율, 변동성 자료를 이용하여 다음과 같이 1998년 7월 22일자 KOSPI 200 선물의 이론가격을 산출한다.

$$F_t = [S_t - D(t, T)] \cdot \exp[\hat{\beta}_0 + \hat{\beta}_1 r_t + \hat{\beta}_2 \sigma_t] \quad (4)$$

(3) 날짜가 이동함에 따라 추가적인 자료를 이용하여 (1)과 (2)의 과정을 반복적으로 수행하여 이론가격 시리즈를 산출한다.

#### 4. 인공신경망 모형

인공신경망은 인식이나 의사결정에 있어 컴퓨터보다 우수한 인간두뇌에 착안하여 인간의 두뇌를 모방하는 연구에서 비롯되었다. 인공신경망 접근방법은 매우 뛰어난 학습능력과 인지능력이 있어 기업도산과 같이 '분류문제(classification problem)에 있어서 여타 통계적 방법이나 귀납적 학습방법보다는 우수한 인식능력이 있음이 알려져 있다.

인공신경망의 장점은 비교적 쉽게 비선형적 특성을 갖는 금융문제들을 모형화하여 예측할 수 있다는 것이다. 또한 자료가 불확실하고, 불완전하여도 아주 효과적으로 모델을 생성하고 또한 학습능력 및 지식발견에서도 아주 뛰어난 능력을 보여 주고 있다. 이미 기존의 많은 논문들은 인공신경망 모형이 기업부도 예측에서 통계학적 모형보다 월등히 뛰어나다는 연구결과를 많이 보여주었으며 그외 많은 금융문제해결에도 우수한 결과를 보여주었다. 인공신경망의 가장 큰 단점은 설명력이 없다는 것이다. 다시 말해서 사용자에게 왜 그런 결정을 했는지에 대한 설명을 하지 못한다는 것이다. 이러한 문제를 극복한 인공지능을 이용한 다른 접근은 디시전트리(decision tree)의 지식표현을 사용한 귀납적학습방법(rule induction)이나 여러 인공지능기법을 통합한 통합모형등으로 현재 많은 연구가 이루어지고 있다.

인공신경망모형에서 사용된 모형은 일반적으로 많이 사용되는 역전파모형(Backpropagation Architecture)을 이용하

였다. 역전파 모형중에서도 다음의 【그림 4-1】과 같은 3계층모형을 이용하였다. 인공신경망모형의 입력변수는 현물가격, 잔존일수, GARCH(1,1)모형을 이용한 변동성, 이자율, 배당락지수, 어제의 현물가격, 이틀전의 현물가격, 어제의 선물가격, 이틀전의 선물가격 등 9개로 선정하고 출력변수는 선물가격으로 선정되었다.

#### 【기호설명】

$S_t$  :  $t$  시점의 현물가격

$S_{t-1}$  :  $t-1$  시점의 현물가격

$\tau$  : 만기일까지의 잔존일수

$\sigma$  : 변동성(Garch(1,1))

$r$  : 이자율(CD 금리)

$D(t, T)$  : 배당락 지수

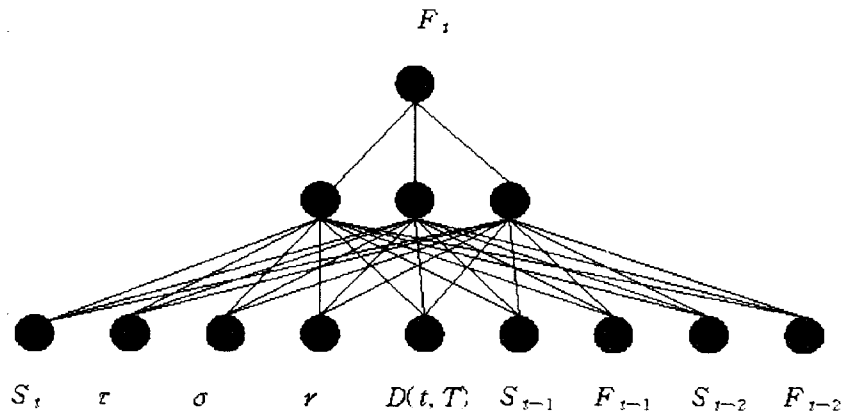
$F_{t-1}$  :  $t-1$  시점의 선물가격

$S_{t-2}$  :  $t-2$  시점의 현물가격

$F_{t-2}$  :  $t-2$  시점의 선물가격

인공신경망모형에서 모형의 설계시 가장 중요한 것은 히든 노드의 수를 결정하는 것이다. 본 실험에서는 파일럿실험에서 히든노드의 수를 결정하였다. 자료의 훈련방식을 자료를 무작위적으로 선택하여 훈련받는 무작위방식(Random Selection)을 이용하였다. 인공신경망모형의 학습율과 모멘텀은 각각 0.05와 0.5로 설정하였으며, 실험을 위해 사용된 프로그램은 Ward System사에서 개발된 NeuralShell이고, 사용된 PC는 Pentium 200MHz를 이용하였다.

【그림 4-1】 인공신경망 모형



## 5. 실증분석

### 1) 분석에 사용된 자료 및 기간

본 연구에서 분석기간은 1997년 7월 7일부터 1998년 9월 31일까지 약 1년여이다. 본 연구에 이용된 자료는 KOSPI 200, KOSPI 200 선물 시장가격(최근월물과 차근월물), 배당락지수(최근월물과 차근월물), 양도성 예금증서(CD) 91일물의 연수익률 등이며, 모든 자료는 증권거래소로부터 구입하였다. 시장변동성은 KOSPI 200 지수 수익률의 GARCH(1,1) 변동성을 이용하였다. 선물시장이 주식시장보다 15분 늦게 종료함에 따라 현물과 선물의 종가(closing prices)를 이용하여 가격과리울을 측정하는 경우 단순히 시간차에 의한 오차가 지속적으로 발생하므로 본 연구에서는 오전 11시 30분의 가격자료를 이용하였다.

· 기간 : 1997년 7월 7일 ~ 1998년 9월 30일(365개 자료)

· 자료종류 : Nearest Futures Price

· 이자율 : CD 금리(비율)

· 변동성 : GARCH(1,1)의 변동성(비율)

· 잔존기간 : 잔존일수/365

· Training set : 초기 333개(2개의 자료는 2-lagged 자료 때문에 없어짐)

· Testing set : 초기 30개.(1998년 7월 22일부터)

· 입력자료 : ①현물가격(Kospi200) ②잔존일수(Time to expiration), ③배당락지수, ④변동성(Volatility), ⑤이자율(Interest rate), ⑥어제종가(Lag closing price), ⑦어제현물가격(Lag Kospi200), ⑧이틀전 종가(2-lagged closing price), ⑨이틀전 현물가격(2-lagged Kospi200)

· 출력자료 : ① 11시 30분 종가(Closing price)

### 2) 예측성과 측정치

본 실험에서 사용한 각 모형의 성과에 대한 측정치는 과리울의 평균, 과리울<sup>2</sup>의 평균, MAD(Mean Absolute Deviation), MSE(Mean Square Error)를 사용하였다.

과리울의 평균은 식(1)에서 보는 바와 같이 실제가격과 모형가격의 차이를 모형가격으로 나누어 준 결과들의 평균이다.

과리울의 평균 =

$$\sum_{i=1}^N \frac{(X_{actual, i} - X_{model, i})}{X_{model, i}} \quad (1)$$

MAD는 식(2)서 보는 바와 같이 실제가격과 모형가격의 차이를 절대값을 취한 다음 나온 결과들의 평균이다.

$$MAD = \frac{\sum_{i=1}^N |X_{actual, i} - X_{model, i}|}{N} \quad (2)$$

MSE는 식(3)와 같이 실제가격과 모형가격의 차이를 제곱한 결과들의 평균이다. MAD와 MSE는 본 실험에서 에러를 측정하는데 의미 있는 측정도구로 사용되었다.

$$MSE = \frac{\sum_{i=1}^N (X_{actual, i} - X_{model, i})^2}{N} \quad (3)$$

### 3) 실험결과

파일럿실험을 바탕으로 3개의 히든노드와 4개의 히든노드 인공신경망모형 모두를 실험자료 30개에 실험하고 일반균형모형과 보유비용모형을 비교하였다. 또한 투자전문프로그램인 NeuroShell Trader라는 프로그램을 이용하여 본 실험에서 설정된 인공신경망모형과의 성과를 비교하였다.

#### (1) 보유비용모형과 일반균형모형의 비교

1998년 7월 22일부터 1998년 9월 30일까지 1개월동안 KOSPI 200 선물가격의 과리울을 보여 주고 있다. 최근월물의 경우 보유비용모형이 주는 이론가격으로부터의 과리울은 -7.73% ~ -0.41%의 분포를 보이고 있으며, 평균적으로 -3.85%의 가격과리울을 보였다. 한편, 일반균형모형이 주는 이론가격으로부터의 과리울은 -4.46% ~ +3.59%의 분포를 보이고 있으며, 평균적으로 -0.05%의 가격과리울을 보였다. 평균 가격과리울을 단순 비교하더라도 보유비용모형보다 일반균형모형이 시장가격에 가까운 이론

가격을 제시하고 있다는 사실을 확인할 수 있다. 이 보다 더 중요한 사실은 보유비용모형이 시장 가격에 비해 고평가하는 경향을 보여 가격과리율이 음(-)의 값에 편향되어 있는 반면, 일반균형모형은 가격과리율이 0을 중심으로 상하로 고르게 분포되어 있다는 점을 들 수 있다.

차근월물의 경우 보유비용모형이 주는 이론가격으로부터의 과리율은 -9.60% ~ -2.28%의 분포를 보이고 있으며, 평균적으로 -5.85%의 가격과리율을 보였다. 한편, 일반균형모형이 주는 이론가격으로부터의 과리율은 -5.04% ~ +1.52%의 분포를 보이고 있으며, 평균적으로 -1.23%의 가격과리율을 보였다. 최근월물의 경우와 마찬가지로 평균 가격과리율을 단순 비교할 때 보유비용모형보다 일반균형모형이 시장가격에 가까운 이론가격을 제시하고 있다. 이상의 분석결과를 요약하면, 시장가격에 가까운 이론가격을 제시한다는 측면에서 일반균형모형이 보유비용모형보다 우월함을 알 수 있다.

## (2) 인공신경망 모형

### ① 인공신경망모형의 훈련결과

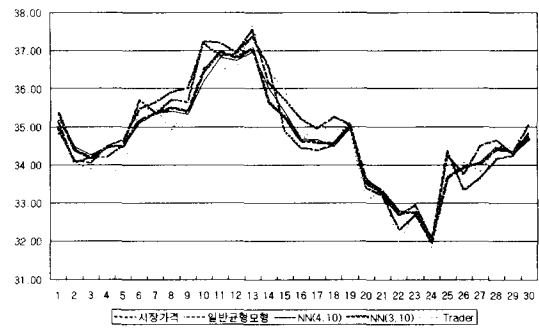
본 실험에서의 인공신경망모형은 3개와 4개의 히든노드를 가진 인공신경망모형을 훈련시켜 모형을 실험한 것이다. 훈련모형의 결과는 부록의 A에 첨부하였다. 훈련시간을 10분으로 설정하였다. 평균 훈련수는 3개의 인공신경망모형이 21155번, 4개의 인공신경망모형이 19755번을 기록하였다. 3개와 4개의 히든노드 인공신경망모형의  $R^2$ 의 평균은 각각 0.9975, 0.9975를 기록하여 매우 높게 나타났다.

### ② 모형간의 가격비교

실제시장가격과 각 모형들의 가격 비교는 【 그림 5-1 】에서 제시하였다. 【 그림 5-1 】의 X축의 15와 17사이를 보면 일반균형모형이 실제가격과 상당한 차이가 나는 것을 볼 수 있다. 즉 급격한 감소의 패턴을 기존 재무이론이 잘 잡지 못하는 것을 알 수 있다. 또한 3개와 4개의 히든노드 인공신경망모형의 가격은 실험자료 30개 모두를 실험해도 가격예측의 차이는 있지만 시장가

격에 대한 같은 패턴을 보이고 있는 것을 알 수 있다. NeuroShell Trader는 초반에 차이가 큰 반면 후반에 갈수록 실제가격과의 차이가 적어짐을 볼 수 있다.

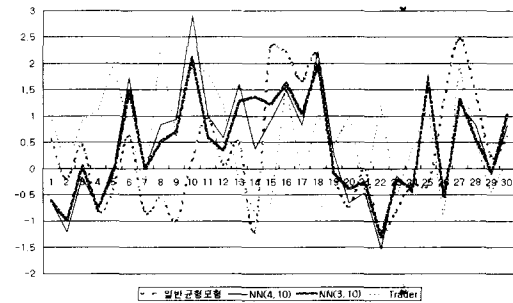
【 그림 5-1 】 전체모형가격과 시장가격과의 비교



### ③ 모형간의 성과비교

모형가격의 성과를 비교할 때에는 측정도구를 이용하여 측정도구의 크기를 가지고 분석하게 된다. 과리율의 평균이나 과리율<sup>2</sup>의 평균, MAD, MSE 등의 값이 적을수록 실제가격과 가까운 값을 가지게 된다. 다음 【 그림 5-2 】은 모형가격별로 과리율을 표시한 것이다.

【 그림 5-2 】 모형가격들의 과리율



다음의 【 표 5-1 】은 기존의 선물가격결정모형과 인공신경망모형을 측정도구를 이용하여 어느 모형이 실제가격에 더 가까운지를 보여주고 있다. 과리율의 평균과 MAD에서는 일반균형모형가격이 다른 모형에 비해 실제가격과 거의 차이가 없음을 보여주고 있다. 하지만 과리율<sup>2</sup>의 평균과 MSE에서는 3개의 히든노드 인공신경망모형이 다른

모형에 비해 실제가격과 더 가까운 모형일  
을 알 수 있다.

【 표 5-1 】 기존의 선물가격결정모형과  
인공신경망모형의 비교

	일반균형 모형가격	보유비 용가격	4-히든 인공신경 망가격	3-히든 인공신경 망가격	Nshell trader 가격
파리올 의 평균	0.3045	-3.3525	0.4575	0.4446	0.6085
파리올 <sup>2</sup> 의 평균	1.2307	13.5406	1.2638	1.0352	1.2020
MAJ	0.1052	-1.2020	0.1345	0.1592	0.2103
MSE	0.1463	1.7361	0.1552	0.1266	0.1456

따라서 전반적으로 주가지수 선물가격의 예  
측에서는 일반균형모형과 3개의 히든노드를  
가진 인공신경망 모형이 우수함을 알 수 있  
었다. 특히 일종의 최적화 기법을 사용하고  
있는 투자전문 프로그램인 Neuroshell  
Trader의 경우 미국 선물자료를 이용하여  
자체적으로 인공신경망의 최적화기법을 만  
들어놓은 것인데 그 예측 성능은 매우 높은  
것으로 밝혀져 향후 인공신경망 기법의 일  
상적 활용화에 많은 기여를 할 것으로 판단  
된다.

## 6. 귀납적 학습방법을 이용한 모형실험

귀납적 학습방법을 이용한 실험에서는  
CART(Classification and Regression Tree)  
알고리즘을 이용하여 규칙을 생성하였다.  
CART 알고리즘은 의사결정트리를 형성하는  
데 CHAID, QUEST 등과 함께 널리 사용되는  
의사결정트리 알고리즘 중의 하나이며, 일  
반적으로 시간의 변화에 따른 종속변수의  
효과를 알아보는 데에 사용된다.

이 실험에서는 선물가격의 향방을 예측하  
는 규칙을 찾아내어 선물가격모형의 퍼지화  
에 도움을 주고자 한다.

### 1) 자료구성

이 실험을 위하여 입력변수와 출력변수를  
재구성하였다. 입력변수에는 배당타지수  
(ND), 잔존일수를 반영한 이자율( $MCD[=CD \times$   
 $MAT]$ ), 잔존일수를 반영한 변동성( $MGARCH[=$   
 $GARCH \times MAT]$ ), 이틀전과 전날종가와의

Kospi200 비율( $KP_1[=KP_{t-1}/KP_t]$ ), 삼일전과  
이틀전과의 Kospi200 비율( $KP_2[=KP_{t-2}/KP_t]$ ),  
이틀전과 전날의 선물가격의 비율( $NF_1[=NF_{t-2}$   
 $/NF_{t-1}]$ ), 삼일전과 이틀전과의 선물가격의  
비율( $NF_2[=NF_{t-3}/NF_{t-1}]$ )로 선정하였고, 출력  
변수로는  $NF_t$ 와  $NF_{t-1}$ 를 비교하여 2그룹, 4  
그룹, 6그룹으로 나누어 자료를 구성하였  
다. 2그룹의 자료는 올라가는 방향(UP)이면  
1을, 내려가는 방향(DOWN)이면 0로 구분하  
였으며, 4그룹은 2그룹보다 더 세분화하여  
시장가격이  $\pm 2$  Point를 기준으로 하여 -2  
Point이하 감소는 1, -2 Point이상 0 Point  
미만 감소는 2, 0 Point이상 2 Point이하  
증가는 3, 2 Point이상 증가는 4로 설정하  
였다. 6 그룹은  $\pm 1$  Point와  $\pm 2$  Point를  
기준으로 자료를 구성하였다.

실험에 사용된 자료는 출력변수에서 언급  
한 2그룹, 4그룹, 6그룹으로 나누고 각각의  
그룹을 검증표본은 약 30개로 정하고, 특히  
일정시점에 영향을 받지않는 모형의 검증을  
위하여 분석표본을 기간별로 5개 그룹으로  
나누어 분석표본수를 기간별로 늘려 나가면  
서 한달 단위의 검증표본을 사용하였다.

### 2) 실험결과

실험에서 사용된 CART 알고리즘은 SPSS의  
AnswerTree를 이용하여 실험하였다. 실험에  
서 사용된 예측치에 대한 계산은 (예측된  
결과갯수)/(전체갯수)로 계산되어졌다. 다  
음은 그룹에 따른 예측력이 제시되었다.

#### ① 가격향방의 2그룹결과

가격의 향방을 Up과 Down으로 나누어 출  
력변수를 구성하였을 때 결과가 다음의 【  
표 6-1】에 제시되었다. 표를 살펴보면 2  
그룹에서의 예측력은 약 91.1%이상을 기록  
하였다.

【 표 6-1 】 자료구성에 따른 2그룹의  
예측결과

	자료의 구성		예측력 2그룹
	분석	검증	
1	211	29	93.1%
2	241	29	93.1%
3	271	29	89.6%
4	301	29	89.6%
5	331	31	90.3%

분석표본의 증가에 따른 예측력은 증가하는 방향은 아니지만 상당히 안정적인 결과를 보이고 있다.

2그룹의 Rule은 자료 5번의 Rule만 부록에 제시하였다.

② 가격향방의 4그룹결과

가격의 향방을 Up과 Down으로 나누고 이에 대한 Up과 Down을 세분화시켜 Highly Up과 Highly Down으로 자료를 더 나누어 출력변수를 구성하였을 때 결과가 다음의 【 표 6-2 】에 제시되었다. 표를 살펴보면 4그룹의 예측력은 약 80%를 기록하였고, 4그룹을 2그룹으로 변환시켜 측정한 예측력은 약 92.5%이상을 기록하였다.

분석표본의 증가에 따른 예측력은 2그룹과 비슷한 경향으로 상당히 안정적인 결과를 보이고 있다.

4그룹의 Rule은 자료 5번의 Rule만 부록에 제시하였다.

【 표 6-2 】 자료구성에 따른 4그룹의 예측결과

	자료의 구성		예측력	
	분석	검증	4그룹	2그룹
1	211	29	68.9%	86.6%
2	241	29	86.2%	96.5%
3	271	29	79.3%	93.1%
4	301	29	79.3%	93.1%
5	331	31	83.9%	93.5%

③ 가격향방의 6그룹결과

가격의 향방을 Up과 Down으로 나누고 이에 대한 Up과 Down을 더 세분화시켜 Deeply High Up, Highly Up, Highly Down, 그리고 Deeply High Down으로 자료를 더 나누어 출력변수를 구성하였을 때 결과가 다음의 【 표 6-3 】에 제시되었다. 표를 살펴보면 6그룹의 예측력은 59.4%를 기록하였고, 6그룹에서 4그룹으로 변환시킨 평균 예측력은 약 74.2%를 기록하였고, 6그룹을 2그룹으로 변환시켜 측정한 평균 예측력은 약 89.9% 이상을 기록하였다.

분석표본의 증가에 따른 예측력은 2그룹과 비슷한 경향으로 상당히 안정적인 결과를

보이고 있다.

6그룹의 Rule은 자료 5번의 Rule만 부록에 제시하였다.

【 표 6-3 】 자료구성에 따른 6그룹의 예측결과

	자료의 구성		예측력		
	분석	검증	6그룹	4그룹	2그룹
1	211	29	55.2%	68.9%	89.7%
2	241	29	62.1%	75.9%	93.1%
3	271	29	51.7%	72.4%	86.2%
4	301	29	65.5%	75.9%	89.7%
5	330	32	62.5%	78.1%	90.8%

종합적으로 살펴보면 2그룹, 4그룹, 6그룹으로 나누어 2그룹으로 변환시킨 예측력은 각각 평균 91.1%, 92.5%, 89.9%를 기록하고 있으며, 대부분의 자료에서 아주 높은 예측력을 보이고 있다. Up과 Down을 표현한 그룹 중 4개의 그룹으로 나눈 4그룹의 예측력이 92.5%로 높은 결과를 보이고 있다.

참고문헌

Brooks, C., "Predicting Stock Index Volatility: Can Market Volume Help?", *Journal of Forecasting*, Vol.17. (1998), 59-80

Cheng, W., McClain, B., and Kelly, Christopher, "Artificial neural networks make their mark as a powerful tool for investors," *Review of Business*, 18(Summer, 1997), 4-9

Cornell, B. and French, K., "Taxes and the Pricing of Stock Index Futures," *Journal of Finance*, 38(June, 1983), 675-694

Donaldson, R. Glen, "Forecast Combining with Neural Networks," *Journal of Forecasting*, Vol.15. (1996), 49-61

Helmer, M., "The Quality Delivery Option in Treasury Bond Futures Contracts.", Ph.D. Disstertation, Univ. of Chicago, (1988)

Helmer, M and Longstaff, F., "General Equilibrium Stock Index Futures Price:



Theory and Empirical Evidence," 26(Sep. 1991), 287-308

Hutchinson, J., Lo, A., and Poggio, T., "A Nonparametric Approach to Pricing and Hedging Derivative Securities Via Learning Networks," Journal of Finance, 49(July, 1994), 851-889.

Kamar, A., "Market Trading Structure and Asset Pricing: Evidence from the Treasury-Bill Markets," Review of Financial Studies, 1(winter 1988), 357-376

Mackinley, A. and Ramaswamy, K., "Index-Futures Arbitrage and the Behavior of Stock Index Future Prices," Review of Financial Studies, 1(summer 1988), 373-413

#### 【부록】

##### 1. 2그룹의 Rule

```
if kpt1 <= 1.00141938 and kpt1 <= 0.9766600485
then
  node=2
  prediction=up
  probability=1.000

if kpt1 <= 1.00141938 and kpt1 > 0.9766600485
and nft1 <= 0.951756492
then
  node=4
  prediction=down
  probability=0.700

if kpt1 <= 1.00141938 and kpt1 > 0.9766600485
and nft1 > 0.951756492 and kpt1 <= 1.000803974
and mgarch <= 0.1359926715
then
  node=7
  prediction=up
  probability=0.860

if kpt1 <= 1.00141938 and kpt1 > 0.9766600485
and nft1 > 0.951756492 and kpt1 <= 1.000803974
and mgarch > 0.1359926715
then
  node=8
  prediction=down
  probability=0.667

if kpt1 <= 1.00141938 and kpt1 > 0.9766600485
and nft1 > 0.951756492 and kpt1 > 1.000803974
then
  node=9
  prediction=down
  probability=0.571

if kpt1 > 1.00141938 and kpt1 <= 1.0089048195
and mcd <= 0.0284691785 and mcd <= 0.009868493
then
  node=13
  prediction=down
  probability=0.563

if kpt1 > 1.00141938 and kpt1 <= 1.0089048195
and mcd <= 0.0284691785 and mcd > 0.009868493
then
  node=14
  prediction=down
  probability=1.000

if kpt1 > 1.00141938 and kpt1 <= 1.0089048195
```

```
and mcd > 0.0284691785
then
  node=15
  prediction=up
  probability=0.750
if kpt1 > 1.00141938 and kpt1 > 1.0089048195
and mgarch <= 0.157606356 and kpt2 <=
0.9976909215
then
  node=18
  prediction=down
  probability=0.833

if kpt1 > 1.00141938 and kpt1 > 1.0089048195
and mgarch <= 0.157606356 and kpt2 >
0.9976909215 and nft2 <= 1.075994006
then
  node=20
  prediction=down
  probability=0.990

if kpt1 > 1.00141938 and kpt1 > 1.0089048195
and mgarch <= 0.157606356 and kpt2 >
0.9976909215 and nft2 > 1.075994006
then
  node=21
  prediction=down
  probability=0.900

if kpt1 > 1.00141938 and kpt1 > 1.0089048195
and mgarch > 0.157606356
then
  node=22
  prediction=down
  probability=0.714
```

##### 2. 4그룹의 Rule

```
if kpt1 <= 1.00141938 and kpt1 <= 0.960525525
and nd <= 1.23
then
  node=3
  prediction=highly Up
  probability=0.960

if kpt1 <= 1.00141938 and kpt1 <= 0.960525525
and nd > 1.23
then
  node=4
  prediction=Up
  probability=1.000

if kpt1 <= 1.00141938 and kpt1 > 0.960525525
and kpt2 <= 0.938567928
then
  node=6
  prediction=highly Up
  probability=0.412

if kpt1 <= 1.00141938 and kpt1 > 0.960525525
and kpt2 > 0.938567928 and mcd <= 0.0489726025
and kpt1 <= 1.000803974
then
  node=9
  prediction=Up
  probability=0.833

if kpt1 <= 1.00141938 and kpt1 > 0.960525525
and kpt2 > 0.938567928 and mcd <= 0.0489726025
and kpt1 > 1.000803974
then
  node=10
  prediction=Down
  probability=0.500

if kpt1 <= 1.00141938 and kpt1 > 0.960525525
and kpt2 > 0.938567928 and mcd > 0.0489726025
then
  node=11
  prediction=highly Up
  probability=0.600

if kpt1 > 1.00141938 and kpt1 <= 1.0313713085
and nft2 <= 0.9540879495
then
  node=14
  prediction=Down
  probability=0.400
```

```

if kpt1 > 1.00141938 and kpt1 <= 1.0313713085
and nft2 > 0.9540879495 and nft2 <=
1.0580663235 and mgarch <= 0.151772178
then
  node=17
  prediction=Down
  probability=0.849

if kpt1 > 1.00141938 and kpt1 <= 1.0313713085
and nft2 > 0.9540879495 and nft2 <=
1.0580663235 and mgarch > 0.151772178
then
  node=18
  prediction=Up
  probability=0.500

if kpt1 > 1.00141938 and kpt1 <= 1.0313713085
and nft2 > 0.9540879495 and nft2 >
1.0580663235
then
  node=19
  prediction=Down
  probability=0.500

if kpt1 > 1.00141938 and kpt1 > 1.0313713085
and kpt1 <= 1.0466785465
then
  node=21
  prediction=highly Down
  probability=0.500

if kpt1 > 1.00141938 and kpt1 > 1.0313713085
and kpt1 > 1.0466785465
then
  node=22
  prediction=highly Down
  probability=0.950

```

### 3. 6그룹의 Rule

```

if kpt1 <= 1.00141938 and kpt1 <= 0.965574107
and kpt2 <= 0.9812891695
then
  node=3
  prediction=Deeply High Up
  probability=0.958

if kpt1 <= 1.00141938 and kpt1 <= 0.965574107
and kpt2 > 0.9812891695
then
  node=4
  prediction=High Up
  probability=0.667

if kpt1 <= 1.00141938 and kpt1 > 0.965574107
and kpt2 <= 0.9821983535 and nft1 <=
0.9637181425 and nft2 <= 1.0012903225
then
  node=8
  prediction=Up
  probability=0.360

if kpt1 <= 1.00141938 and kpt1 > 0.965574107
and kpt2 <= 0.9821983535 and nft1 <=
0.9637181425 and nft2 > 1.0012903225
then
  node=9
  prediction=Deeply High Up
  probability=1.000

if kpt1 <= 1.00141938 and kpt1 > 0.965574107
and kpt2 <= 0.9821983535 and nft1 >
0.9637181425
then
  node=10
  prediction=High Up
  probability=0.625

if kpt1 <= 1.00141938 and kpt1 > 0.965574107
and kpt2 > 0.9821983535 and nft1 <=
1.034743868 and kpt1 <= 0.97337471
then
  node=13
  prediction=High Up
  probability=0.667

if kpt1 <= 1.00141938 and kpt1 > 0.965574107
and kpt2 > 0.9821983535 and nft1 <=
1.034743868 and kpt1 > 0.97337471
then
  node=14

```

```

prediction=Up
probability=0.781

if kpt1 <= 1.00141938 and kpt1 > 0.965574107
and kpt2 > 0.9821983535 and nft1 > 1.034743868
then
  node=15
  prediction=Up
  probability=0.294

if kpt1 > 1.00141938 and kpt1 <= 1.0299667705
and kpt1 <= 1.015914167 and nft2 <=
0.9540879495
then
  node=19
  prediction=Up
  probability=0.444

if kpt1 > 1.00141938 and kpt1 <= 1.0299667705
and kpt1 <= 1.015914167 and nft2 >
0.9540879495 and nft2 <= 1.0580663235
then
  node=21
  prediction=Down
  probability=0.750

if kpt1 > 1.00141938 and kpt1 <= 1.0299667705
and kpt1 <= 1.015914167 and nft2 >
0.9540879495 and nft2 > 1.0580663235
then
  node=22
  prediction=Up
  probability=0.455

if kpt1 > 1.00141938 and kpt1 <= 1.0299667705
and kpt1 > 1.015914167 and mgarch <=
0.0781052605 and nft1 <= 1.0123469835
then
  node=25
  prediction=High Down
  probability=0.682

if kpt1 > 1.00141938 and kpt1 <= 1.0299667705
and kpt1 > 1.015914167 and mgarch <=
0.0781052605 and nft1 > 1.0123469835
then
  node=26
  prediction=Deeply High Down
  probability=0.444

if kpt1 > 1.00141938 and kpt1 <=
1.0299667705 and kpt1 > 1.015914167 and mgarch
> 0.0781052605
then
  node=27
  prediction=Down
  probability=0.583

if kpt1 > 1.00141938 and kpt1 > 1.0299667705
and kpt1 <= 1.0466785465 and mgarch <=
0.0475216025
then
  node=30
  prediction=High Down
  probability=0.667

if kpt1 > 1.00141938 and kpt1 > 1.0299667705
and kpt1 <= 1.0466785465 and mgarch >
0.0475216025
then
  node=31
  prediction=Deeply High Down
  probability=0.733

if kpt1 > 1.00141938 and kpt1 > 1.0299667705
and kpt1 > 1.0466785465
then
  node=32
  prediction=Deeply High Down
  probability=0.950

```