

이 재 원 교수지도

석사학위 청구논문

신경망 기반 주가 예측 시스템의  
성능평가 요소 분석

2005

성신여자대학교 교육대학원

교육학과 전자계산교육전공

김 현 진

신경망 기반 주가 예측 시스템의  
성능평가 요소 분석

이 재 원 교수지도

이 논문을 석사학위 논문으로 제출함.

2005년 5월

성신여자대학교 교육대학원

교육학과 전자계산교육전공

김 현 진

# 인 준 서

김 현 진 의 석사학위 논문을 인준함.

심사위원 (인)

심사위원 (인)

심사위원 (인)

성신여자대학교 교육대학원

# 논문 개요

주식 시장을 예측하려는 많은 연구들이 있었으나, 주식 시장의 가격 변동이나 기대 수익을 예측하는 것은 매우 어려운 작업이다. 한때 ‘효율적 시장가설’로 인해, 학계에서는 시장 분석 방법론들에 대해 회의적 이었으나, 최근 인공지능 분야, 특히 기계 학습 분야의 연구자들은 인공지능 기법과 기술 분석 기법의 결합에 의해 시장 평균을 초과하는 수익의 달성이 가능하다는 연구 결과들을 제시하고 있다.

본 논문에서는 ‘NNshell’이라는 신경망 기반 주가 예측 시스템의 성능 평가 및 시스템 최적화를 위한 통합 시뮬레이션 도구의 분석을 통해, 보다 체계화된 성능 평가 및 거래 시스템의 거래 성능 요소를 알아보려고 한다.

# 목 차

논문 개요

I. 서론 .....	1
II. NNShell : 통합 시뮬레이션 도구 .....	4
III. 체계화된 성능평가 .....	9
1. PPT 기준 소팅하기 .....	12
1.1 PPT .....	12
1.2 PPT 기준 소팅하기 .....	13
2. 일별 PPT 기준 소팅하기 .....	15
3. 적중률 기준 소팅하기 .....	17
3.1 적중률.....	17
3.2 적중률 기준 소팅하기 .....	19
4. 수익금 기준 소팅하기 .....	22

IV. 거래 시스템의 요소 분석 .....	23
1. 거래 대금의 규모 .....	23
2. 예측치 베이스 .....	25
2.1 한 엔진에서의 예측치 베이스의 변화 .....	25
2.2 2개의 다른 엔진에서의 예측치 베이스의 변화 .....	27
2.3 적절한 예측치 베이스 값 .....	30
V. 결론 .....	32

참고문헌

ABSTRACT

## 표 목차

<표 1> 거래 하한 금액의 변화에 따른 엔진별 투자 결과 .....	23
<표 2> A 엔진에서의 예측치 베이스의 변화에 따른 수익금과 거래 횟수의 변화 .....	27
<표 3> B 엔진에서의 예측치 베이스의 변화에 따른 수익금과 거래 횟수의 변화 .....	28
<표 4> 예측치 변화에 따른 수익금의 변화 .....	30

## 그림 목차

<그림 1> NNshell 창 -1 .....	4
<그림 2> 예측치 파일 - 종목번호, 날짜, 예측치 값 .....	5
<그림 3> 매수 할인율 .....	6
<그림 4> 이익 실현율 구간 설정 .....	6
<그림 5> 손절율 구간 설정 .....	7
<그림 6> NNshell 창 -2 .....	8
<그림 7> NNshell 실행 화면 .....	9
<그림 8> Policy 검색하기 .....	10
<그림 9> 특정 거래정책을 이용한 데이터 검색하기 .....	10
<그림 10> 특정 파라미터의 데이터 검색 결과 .....	10
<그림 11> PPT 기준으로 소팅했을 경우 .....	12
<그림 12> 구체적인 거래내역이 나타나 있는 결과 파일 .....	14
<그림 13> 일별 PPT 기준으로 소팅했을 경우 .....	16
<그림 14> 결과 파일의 예시 .....	17
<그림 15> 이익 실현율이 높은 결과 파일 A .....	19

<그림 16> 수익금이 많은 결과 파일 B .....	20
<그림 17> 적중률 기준으로 소팅했을 경우 .....	21
<그림 18> 수익금 기준으로 소팅했을 경우 .....	22
<그림 19> 거래 하한 금액에 따른 엔진별 투자 결과 그래프 .....	24
<그림 20> 예측치에 따른 예측치 데이터의 분포 .....	25
<그림 21> 예측치의 변화에 따른 거래빈도와 수익률간의 관계 .....	26
<그림 22> A 엔진에서의 수익금의 변화 .....	27
<그림 23> A 엔진에서의 거래횟수의 변화 .....	27
<그림 24> B 엔진에서의 수익금의 변화 .....	28
<그림 25> B 엔진에서의 거래횟수의 변화 .....	28
<그림 26> 예측치 변화에 따른 수익금의 변화 그래프 .....	31

# I. 서론

주식 투자는 가장 대중적인 투자 방법 중 하나이며 따라서 주식 시장을 예측하려는 많은 연구들이 있었다[5, 6, 8, 9]. 그러나 많은 잡음(noise), 비정상성(non-stationarity), 그리고 비선형성으로 인해 금융 시계열의 가격 변동이나 기대 수익을 예측하는 것은 매우 어려운 작업이다. 비록 많은 투자 전문가들은 시장 예측을 위해 기본 분석(fundamental analysis)이나 기술 분석(technical analysis) 기법을 사용하고 있지만, 특정 시점에 알려진 모든 정보가 그 시점의 가격에 이미 반영되어 있으므로, 미래 가격의 변동을 예측하는 것은 불가능하다고 주장하는 ‘효율적 시장 가설(EMH: Efficient Market Hypothesis)’[10]로 인해, 한때 학계에서는 시장 분석 방법론들에 대해 회의적인 입장이 지배적이었다. 여전히 효율적 시장 가설을 완전히 기각할 수 있는 연구 결과가 보고되고 있지는 않지만, 최근 인공지능 분야, 특히 기계 학습 분야의 연구자들은 시장이 처음 생각했던 것만큼 효율적이지는 않아서 인공지능 기법과 기술 분석 기법의 결합에 의해 시장 평균을 초과하는 수익의 달성이 가능하다는 연구 결과들을 제시하고 있다. 이러한 연구들에서는 주로 신경망, 의사결정트리, SVM(Support Vector Machine)과 같이 본질적으로 예측 문제에 적합한 감독 학습(supervised learning)을 사용하였다[3, 4, 7].

연구 [9]에서는 각 주식에 대해 분리된 네트워크를 구축하는 방법과, 같은 네트워크를 공유하는 방법 모두를 실험해 보고 각각의 성능을 비교하였다. 이 연구에서는 다양한 경제상황에서의 성능을 평가하였으며 국내총생산, 물가, 실업률, 금리, 환율 등의 거시 경제 변수와 개인의 경제행위, 기업의 경제활동 등의 미시 경제 변수를 고려하였다. 위험분석 요인을 성능 평가 방법에 고려했다는 점은 유의미하다. 그러나 단지 연간 수익률만을 가지고 시스템의

성능을 평가하는 단순한 평가척도를 사용하고 있다. 연구 [8]에서는 주가를 예측하는데 웹 게시판의 자료를 활용하였다. 웹 게시판의 정보를 활용하여 주가를 예측하는 것은 의미 있는 시도이다. 이 연구에서는 성능 평가 방법으로 개별 주식에 대한 주가의 상승률과 하락률을 사용하였다. 그러나 테스트 기간이 200일에 불과하고, 22개의 한정된 주식종목에 대해 실험하였기 때문에 통계적 신뢰도가 미약한 것으로 판단된다.

이러한 기존의 연구들에서는 많은 경우 검증기간이 충분히 장기적이지 못하거나 거래 횟수 등 검증 데이터 수의 부족으로 인해 발생할 수 있는 성능 평가의 통계적 신뢰도 문제에 대한 언급이 부족하다. 또한 다량의 데이터로 검증한 경우라 할지라도 평균수익률이나 거래 성공률 등 비교적 단순한 척도에 의존하는 경우가 많았다. 실제로 장기간에 걸쳐 투자를 함에 있어서는, 투자 기간 동안의 투자자의 심리 등과 같이 기존의 단순한 성능평가 척도들만으로는 표현될 수 없는 문제들을 반영할 수 있는 다양한 성능 평가 수단이 추가로 고려될 필요가 있다.

본 논문에서는 ‘NNShell’이라는, 신경망 기반 주가 예측 시스템의 성능 평가 및 시스템 최적화를 위한 통합 시뮬레이션 도구의 분석을 통해, 보다 체계화된 성능 평가 및 거래 시스템의 거래 성능 최적화 사례를 보인다. 먼저, 보다 현실적인 고려 사항들을 반영할 수 있는 몇 가지 성능 평가 척도들을 제시하고 각각의 필요성과 의미에 대하여 사례를 제시한다. 또한 기존의 연구와 달리 이러한 평가 수단들에 의한 평가가 평가 그 자체로 그치는 것이 아니라 예측 시스템의 최적화를 위한 수단으로 활용될 수 있음을 보인다. 이러한 시스템 요소 분석에서는 일차적으로 예측 시스템의 예측 결과와 거래 정책 파라미터들이 통합적으로 고려되어 각 조합별 성능 평가가 이루어지고, 다음 단계에서 사용자의 선호도에 따라 시스템의 거래 성향이 조정된다.

본 논문은 총 5장으로 구성되어있다. 제 II장에서는 통합 시뮬레이션 도구인 NNshell을 소개하면서 여러 거래 정책들에 대하여 알아보았다. 제 III장

에서는 NNshell의 수행 결과를 바탕으로 여러 가지 성능평가 방법의 특징과 장·단점에 대하여 설명했다. 제 IV장에서는 주식 거래에 있어서 거래 대금의 규모, 예측치 베이스의 최적화에 대하여 알아보았고 마지막으로 제 V장에서는 본 논문의 결과를 요약하고 향후 과제로 끝맺는다.

## II. NNshell

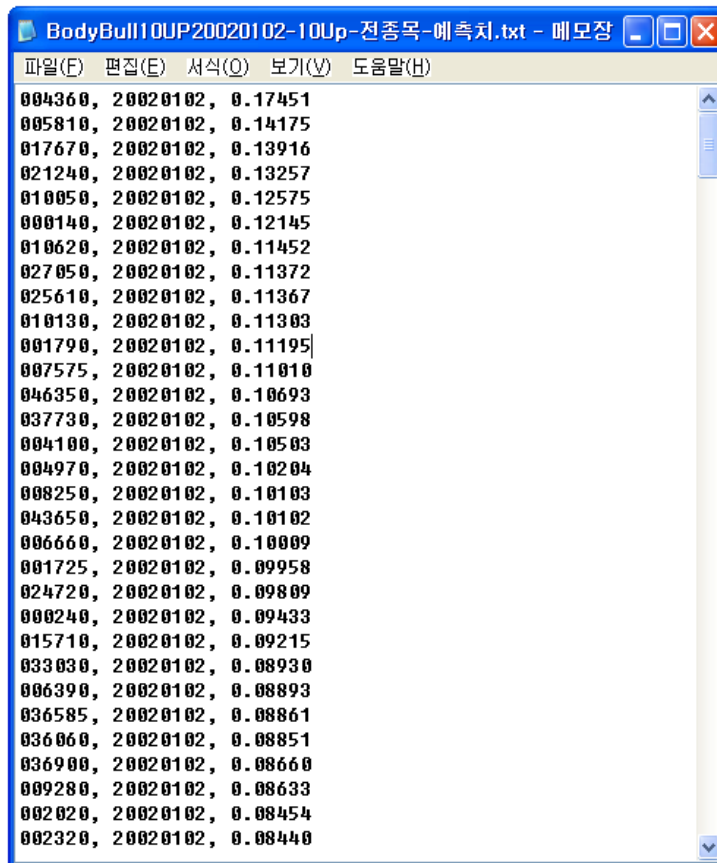
### : 통합 시뮬레이션 도구

본 절에서는 [1, 2]의 연구에서 개발/활용된 통합시뮬레이션 도구인 NNshell을 소개한다. NNshell은 신경망 기반 주가 예측 시스템의 성능 평가와 거래 시스템의 거래 성능 최적화를 통합적으로 수행할 수 있는 기능을 가지고 있다.



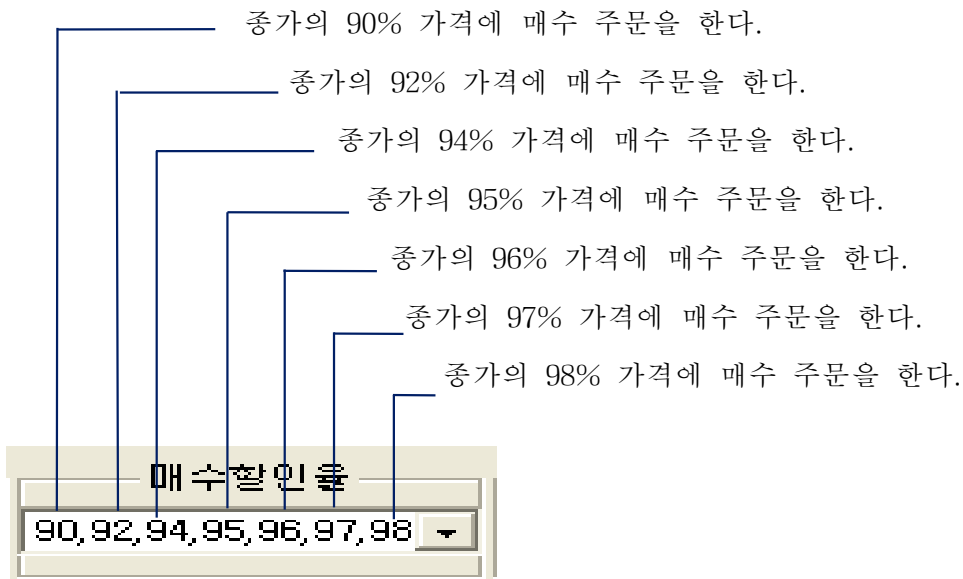
<그림 1> NNshell 창 -1

- ① 예측치 베이스 : 투자 대상으로 삼을 종목에 대한 예측치의 최저 임계치.  
 즉, 얼마 이상의 예측치를 가진 종목에 투자할 것인지를 설정한다. 여기서 예측치란, 예측 엔진의 결과 값으로, -0.9 ~ +0.9 사이의 값으로 표현되며 숫자가 0.9 에 가까울수록 해당엔진이 그 종목의 주가가 오를 가능성이 높다고 판단하였다는 의미이다. 엔진은 모든 거래일에 대해 모든 투자 가능 종목에 대한 예측치를 담고 있는 예측치 파일을 생성한다.
- ② 예측치 상한 : 투자 대상으로 삼을 종목에 대한 예측치의 최고 임계치.



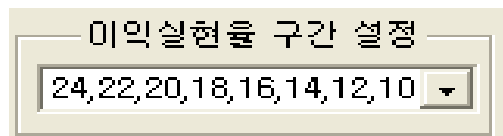
<그림 2> 예측치 파일 - 종목번호, 날짜, 예측치 값

- ③ 매수 할인율 : 당일 종가의 몇 % 가격에 다음날 매수 주문을 할 것인지를 설정한다. 투자하려는 종목의 주가가 상승하리라는 기대가 큰 경우에는 매수 할인율을 높게 설정해서라도, 해당 종목을 구입하는 것이 유리하다. 그러나 주가의 상승이 불확실할 경우에는 매수 할인율을 최대한 낮게 설정하는 것이 좋다.



<그림 3> 매수할인율

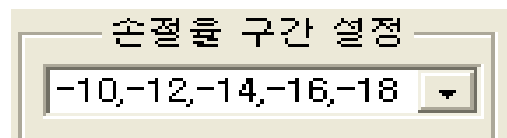
- ④ 이익실현율(target profit rate) 구간 설정 : 매수 가격 대비 몇 % 이상 상승하면 매도할 것인지 설정한다. 그림 4의 앞에서부터 차례로 24%, 22%, 20%, 18%, 16%, 14%, 12%, 10% 의 순



<그림 4> 이익실현율 구간 설정

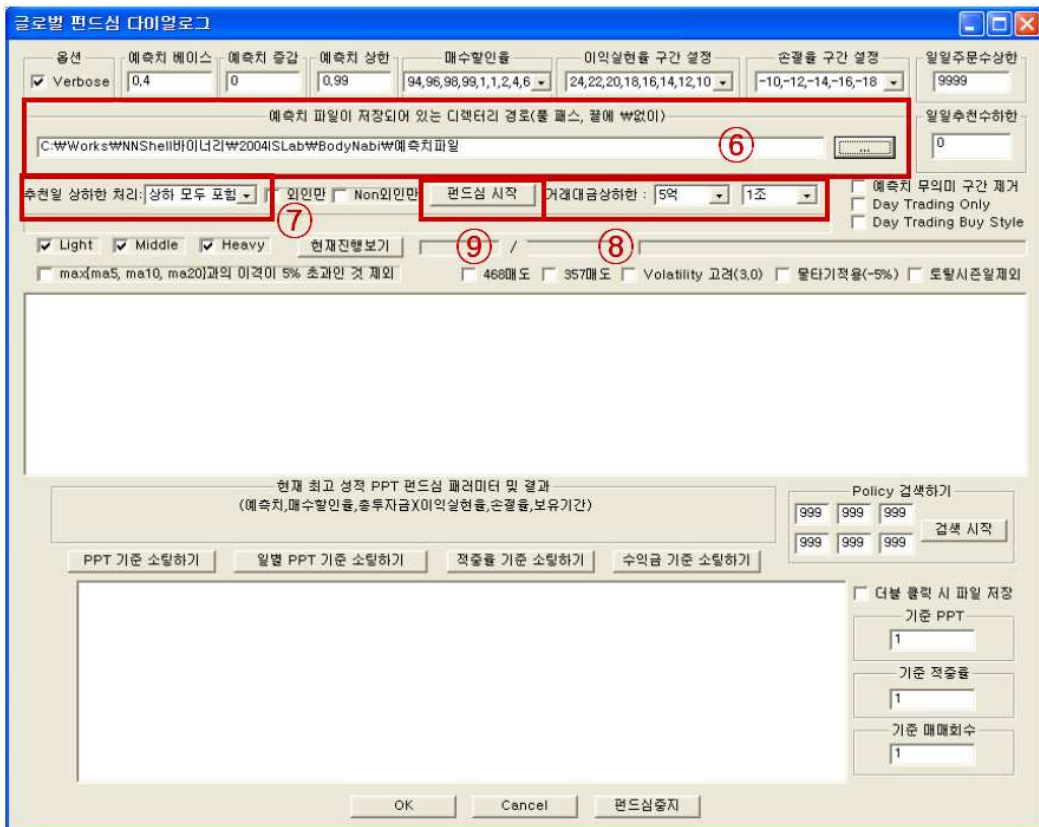
서로, 해당 % 만큼 상승할 경우 매도한다. 매수 할인율과 마찬가지로, 투자하려는 종목의 기대치가 큰 경우, 이익 실현율을 높게 설정해서 최대한 높은 값에 매도하는 것이 유리하다. 그러나 기대치가 낮을수록 이익 실현율을 낮게 설정하는 것이 좋다.

- ⑤ 손절율(stop-loss rate) 구간 설정 : 매수 가격 대비 몇 % 이상 하락하면 매도할 것인지 설정한다. 그림 5의 앞에서부터 차례로 -10%, -12%, -14%, -16%, -18% 의 순서로, 해당 % 만큼 하락할 경우 매도한다. 앞의 매수 할인율, 이익 실현율과 마찬가지로, 투자 종목의 기대치가 큰 경우에는 손절율 구간을 크게 설정해서 주가가 반등하기를 기다리는 것이 유리하다. 그러나 기대치가 낮은 종목의 경우, 손절율 구간을 작게 설정해서 낮은 %의 하락일 지라도 주식을 매도하는 것이 좋다.



<그림 5> 손절율 구간 설정

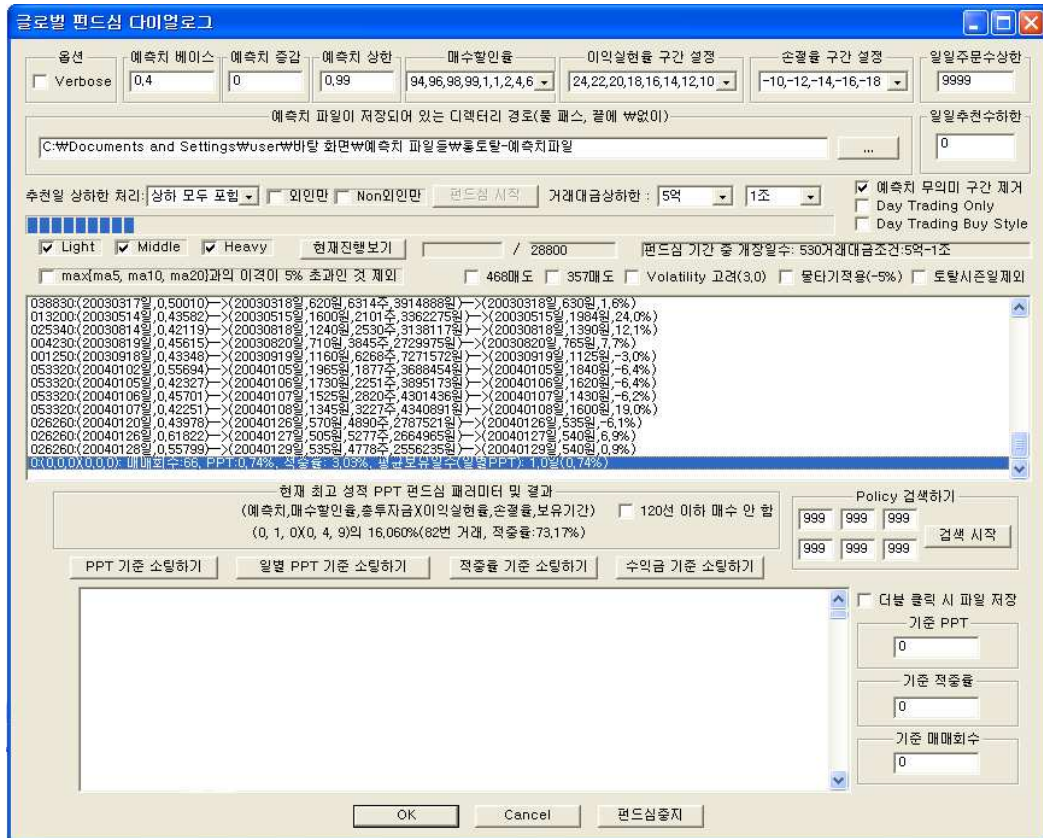
- ⑥ 예측치 파일이 저장되어 있는 디렉토리 : 예측치 파일이 저장되어 있는 디렉토리를 설정한다.
- ⑦ 추천일 상하한 처리 : 상하 제한 폭(±15%)까지 급등하거나, 급락한 종목의 포함 여부를 설정한다.
- ⑧ 거래대금 상하한 : 하루에 한 종목에서 거래되는 금액의 상·하한 규모를 설정한다.
- ⑨ 펀드십 시작 : 설정을 모두 끝내고 프로그램을 실행시킨다.



<그림 6> NNshell 창 -2

# III. 체계화된 성능평가 - 평가 척도별 의미

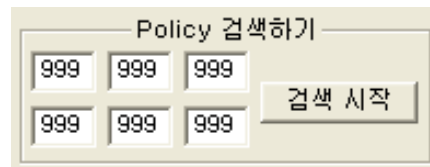
본 절에서는 2절에서 설명한 NNshell의 수행 결과를 바탕으로, PPT 기준, 일별 PPT 기준, 적중률 기준, 수익금 기준으로 정렬(sorting)하고, 각 정렬의 특징과 장단점을 알아보려고 한다. NNshell의 모든 파라미터의 설정을 마치고 ‘펀드십 시작’을 클릭하면 다음과 같은 화면이 나타난다.



<그림 7> NNshell 실행 화면

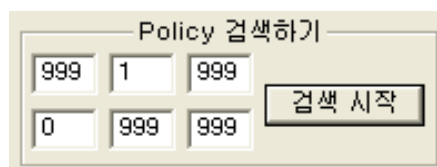
결과 값들은 (예측치, 매수할인율, 예비 인자, 이익실현율, 손절율, 보유기간)의 순으로 표시되며, 현재 창에 표시된 값은 (0, 0, 0, 0, 0, 0) 조합의 경우, 결과 값이 나타나 있다. 현재 창을 보면 거래가 이루어진 종목의 번호, 매매회수, PPT, 적중률, 평균보유일수를 알 수 있다.

이 외에도 각 값들의 다양한 조합을 통해 수백~수만 가지 경우의 거래 결과가 표시되는데 특정한 거래정책을 바탕으로 이루어진 거래 결과를 검색하고 싶을 경우, 창의 오른 쪽에 위치한 ‘Policy 검색하기’를 이용한다. ‘Policy 검색하기’의 배열 순서는 결과 값들의 배열 순서와 같다.



<그림 8> Policy 검색하기

예를 들어, 표시된 거래 결과 중, 종가의 92% 가격에 매수 주문을 하고 24% 상승하면 매도하는 결과를 찾고 싶을 경우 Policy 검색하기 창에 그림 9와 같이 입력하고 ‘검색시작’을 클릭하면 해당 결과만을 보여준다.



<그림 9> 특정 거래정책을 이용한 데이터 검색하기

401	(0, 1, 0X0, 0, 1)	수주	18652000원	전	12,50%	PPT: 7,77%	24번매매	보유	2,00	일	PPT: 3,89%
402	(0, 1, 0X0, 0, 2)	수주	25250962원	전	25,00%	PPT: 10,52%	24번매매	보유	2,79	일	PPT: 3,77%
403	(0, 1, 0X0, 0, 3)	수주	26949039원	전	41,67%	PPT: 11,23%	24번매매	보유	3,46	일	PPT: 3,25%
404	(0, 1, 0X0, 0, 4)	수주	29456739원	전	50,00%	PPT: 12,27%	24번매매	보유	3,96	일	PPT: 3,10%
405	(0, 1, 0X0, 0, 5)	수주	32284626원	전	58,33%	PPT: 13,45%	24번매매	보유	4,38	일	PPT: 3,07%
406	(0, 1, 0X0, 0, 6)	수주	36036057원	전	62,50%	PPT: 15,02%	24번매매	보유	4,71	일	PPT: 3,19%
407	(0, 1, 0X0, 0, 7)	수주	38320509원	전	70,83%	PPT: 15,97%	24번매매	보유	5,00	일	PPT: 3,19%
408	(0, 1, 0X0, 0, 8)	수주	38637476원	전	70,83%	PPT: 16,10%	24번매매	보유	5,17	일	PPT: 3,12%
409	(0, 1, 0X0, 0, 9)	수주	42338049원	전	79,17%	PPT: 17,64%	24번매매	보유	5,33	일	PPT: 3,31%
411	(0, 1, 0X0, 1, 1)	수주	18252000원	전	12,50%	PPT: 7,61%	24번매매	보유	2,00	일	PPT: 3,80%
412	(0, 1, 0X0, 1, 2)	수주	24850962원	전	25,00%	PPT: 10,35%	24번매매	보유	2,79	일	PPT: 3,71%
413	(0, 1, 0X0, 1, 3)	수주	26549039원	전	41,67%	PPT: 11,06%	24번매매	보유	3,46	일	PPT: 3,20%
414	(0, 1, 0X0, 1, 4)	수주	29056739원	전	50,00%	PPT: 12,11%	24번매매	보유	3,96	일	PPT: 3,06%
415	(0, 1, 0X0, 1, 5)	수주	31884626원	전	58,33%	PPT: 13,29%	24번매매	보유	4,38	일	PPT: 3,04%
416	(0, 1, 0X0, 1, 6)	수주	35636057원	전	62,50%	PPT: 14,85%	24번매매	보유	4,71	일	PPT: 3,15%

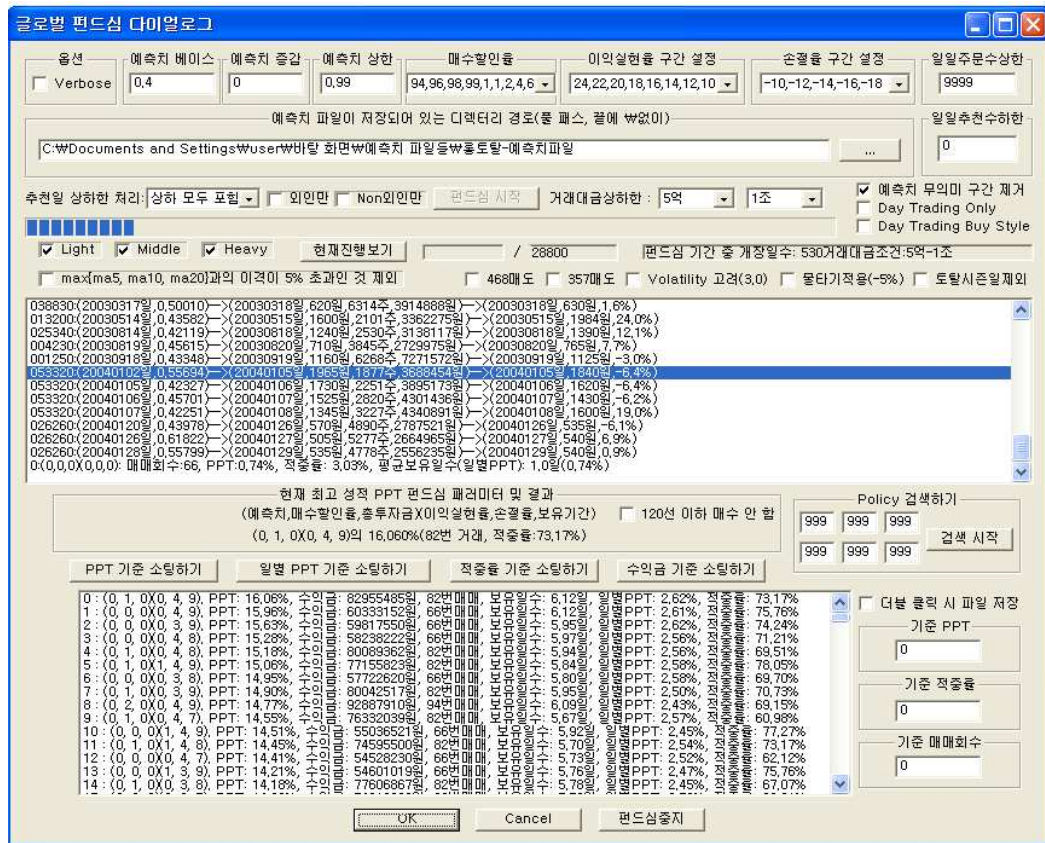
<그림 10> 특정 파라미터의 데이터 검색 결과

찾고자 했던 ‘종가의 92% 가격에 매수 주문을 하고 24% 상승하면 매도 하는 거래 정책’을 이용하는 모든 결과 데이터들이 그림 10과 같이 나타난다. 단, 해당 결과가 없을 경우에 화면에는 아무런 내용도 표시되지 않는다.

# 1. PPT 기준 소팅하기

## 1.1 PPT

PPT는 profit per trade의 약자로서 거래 당 순수수익, 즉, 거래 수수료와 거래세금을 제외한 순수한 수익을 말한다. 여기서, 거래수수료는 증권사에 납부하는 금액으로 매수·매도 시 각각 부과되며 거래세금은 주식을 매수할 경우에는 부과되지 않고 매도할 경우에만 부과된다. 거래 수수료와



<그림 11> PPT 기준으로 소팅 했을 경우

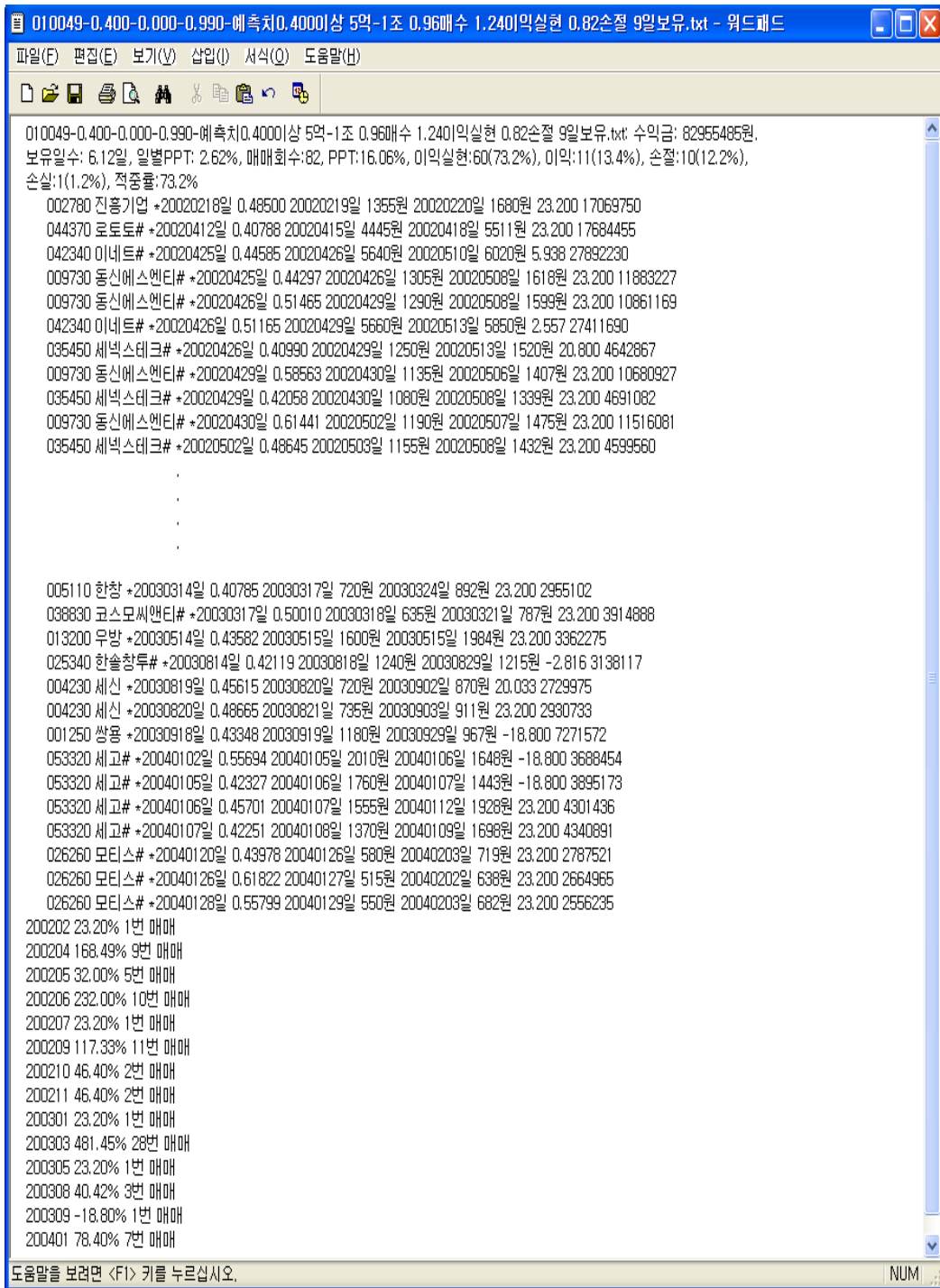
거래 세금은 거래대금의 0.4%가 부과되기 때문에, 빈번히 거래하는 day-trading의 경우 거래 수수료와 거래세금도 고려해야 한다. PPT를 중요한 정렬 기준으로 보는 이유도, 빈번한 거래일 경우 거래 수수료와 거래세금도 적지 않은 금액이기 때문이다. 단기 투자인 day-trading에 집중되었을 경우, PPT를 기준으로 하여 거래하면 유리하다.

$$\text{PPT} = \frac{\text{전체이익 (\%)}}{\text{거래횟수}} - 0.4$$

## 1.2 PPT 기준 소팅하기

‘PPT 기준 소팅하기’를 선택했을 경우, 그림 11과 같은 결과가 나타나며 PPT가 높은 순에서 낮은 순으로 내림차순 정렬된다. 현재 결과에서는 (0, 1, 0, 0, 4, 9)가 가장 높은 PPT(16.06%)를 보이는 조합이며, 82번의 매매를 통해 총 82,955,485원의 수익을 얻었다. 하지만 위의 결과만으로 실제 거래에 적용하기에는 문제가 있다. 매수·매도 가격, 어떤 종목을 거래했는지, 며칠을 보유했는지 등의 세부적 사항이 나타나 있지 않기 때문이다. NNshell에서 ‘더블 클릭시 파일 저장’을 선택한 후, 결과창 하단에 나타나 있는 0번째 결과를 더블클릭하고 파일을 저장할 폴더를 선택하면, 다음 그림 12와 같은 파일이 생성된다.

‘종목번호, 종목이름, 추천날짜, 예측치, 매수일자, 매수가, 매도일자, 매도가, 이익(%), 투자금액’의 순으로 결과 값이 표시되며 마지막에는 거래가 이루어진 시점과 이익, 매매 횟수가 표시된다. 현 정책으로 거래했을 경우 2003년 9월에 발생한 1번의 매매를 제외하고는 모두 이익을 얻었음을 알 수 있다.



<그림 12> 구체적인 거래내역이 나타나 있는 결과 파일

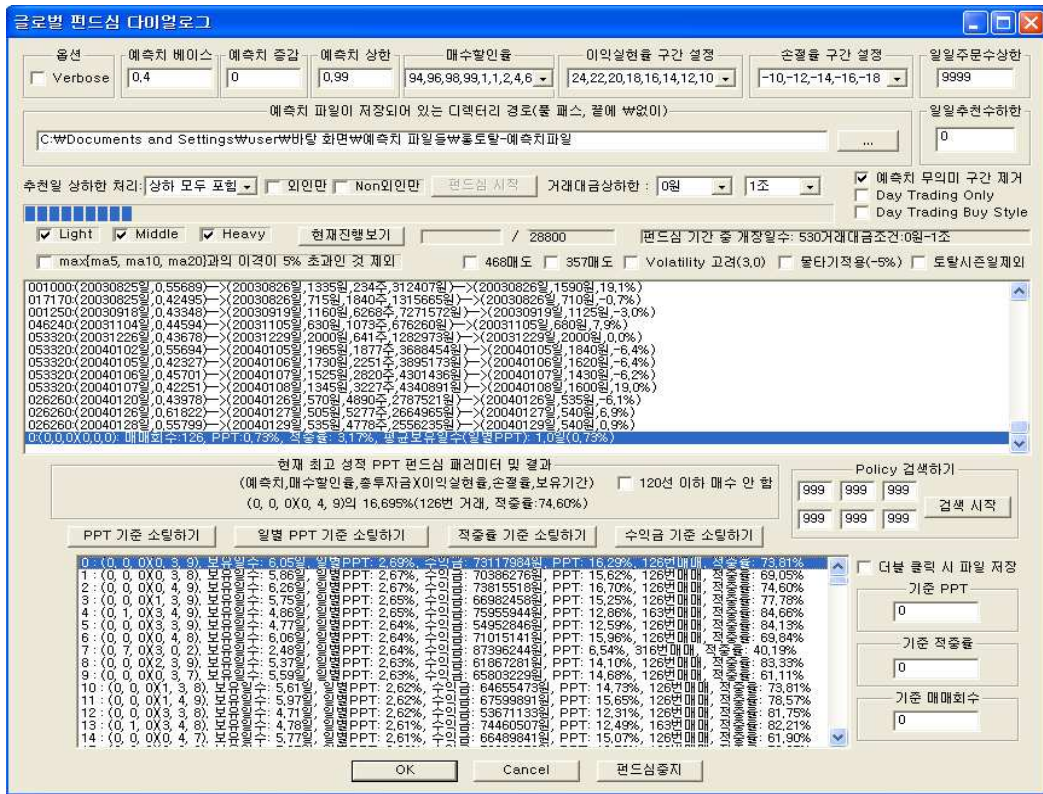
## 2. 일별 PPT 기준 소팅하기

일별 PPT는 날짜별 수익을 기준으로 정렬한다. PPT는 수익을 얻기까지의 소유된 기간에 대한 고려가 없기 때문에, PPT에 대한 시간 효율을 고려한 것이 일별 PPT 이다. 일별 PPT는 기존 논문에서 말하고 있는 PPT 보다 실제 투자에서의 시간적 효율을 추가적으로 고려할 수 있는 평가 척도이다.

$$\text{일별 PPT} = \frac{\text{PPT}}{\text{보유일수}}$$

예를 들어, 투자 결과 같은 20%의 이익달성을 했다 하더라도, 4일 만에 20%의 PPT를 달성한 투자자 A와 10일 만에 20%의 PPT를 달성한 투자자 B를 비교해 보면, 투자자 A가 자금 회전력이 높다는 것을 알 수 있다. 즉, B가 10일 동안 투자한 자금을 이용할 수 없는데 반하여, 4일 만에 20%의 PPT를 달성한 투자자 A의 경우, 나머지 6일간 다른 주식에 대한 투자를 통해 추가적인 수익을 올릴 수 있는 것이다.

투자자의 거래 형태가 단기 투자인 day-trading에 집중되었고 보유자금의 유동률을 높이고 싶을 경우, 일별 PPT를 기준으로 하여 거래하면 유리하다.

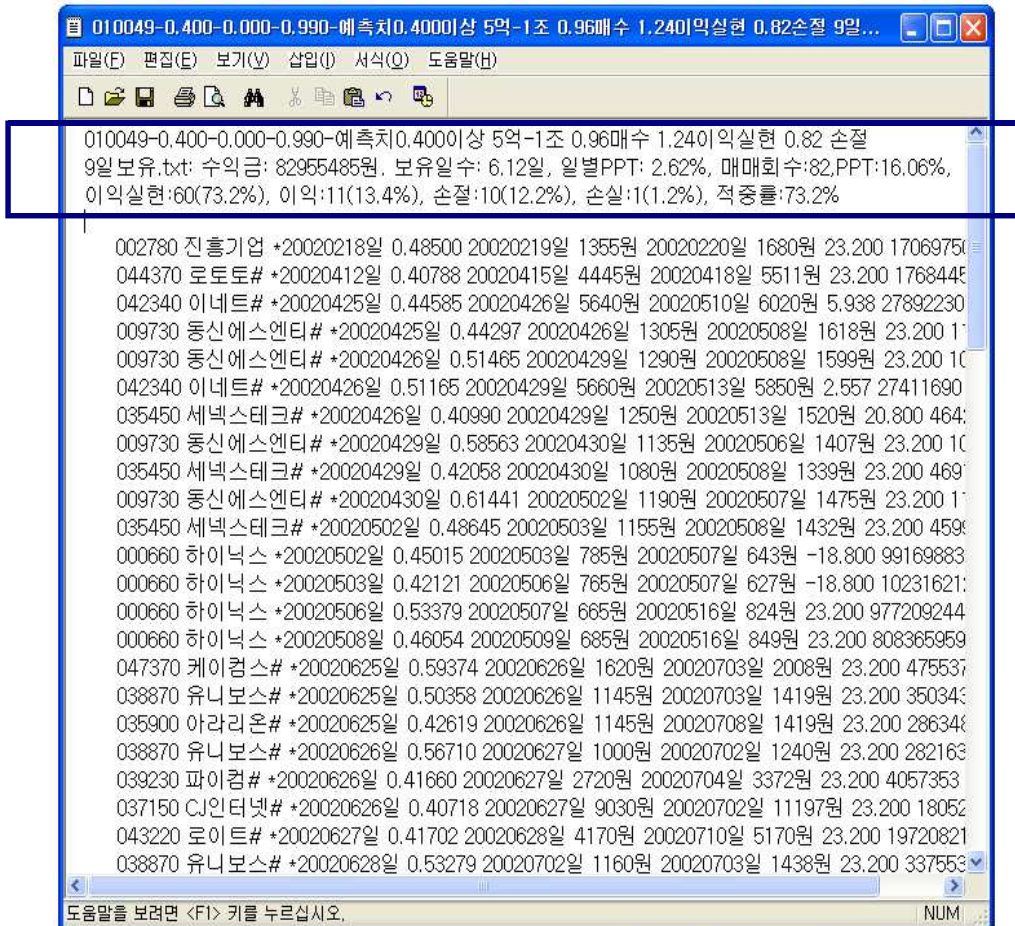


<그림 13> 일별 PPT 기준으로 소팅 했을 경우

### 3. 적중률 기준 소팅하기

#### 3.1 적중률

적중률이란 엔진이 예측한 결과를 실제 주식시장에 적용시켰을 때의 정확도이다. 적중률의 수치는 이익실현을 달성한 주식의 %로 측정하며 그림 14의 경우 73.2%의 이익실현율을 달성했다.

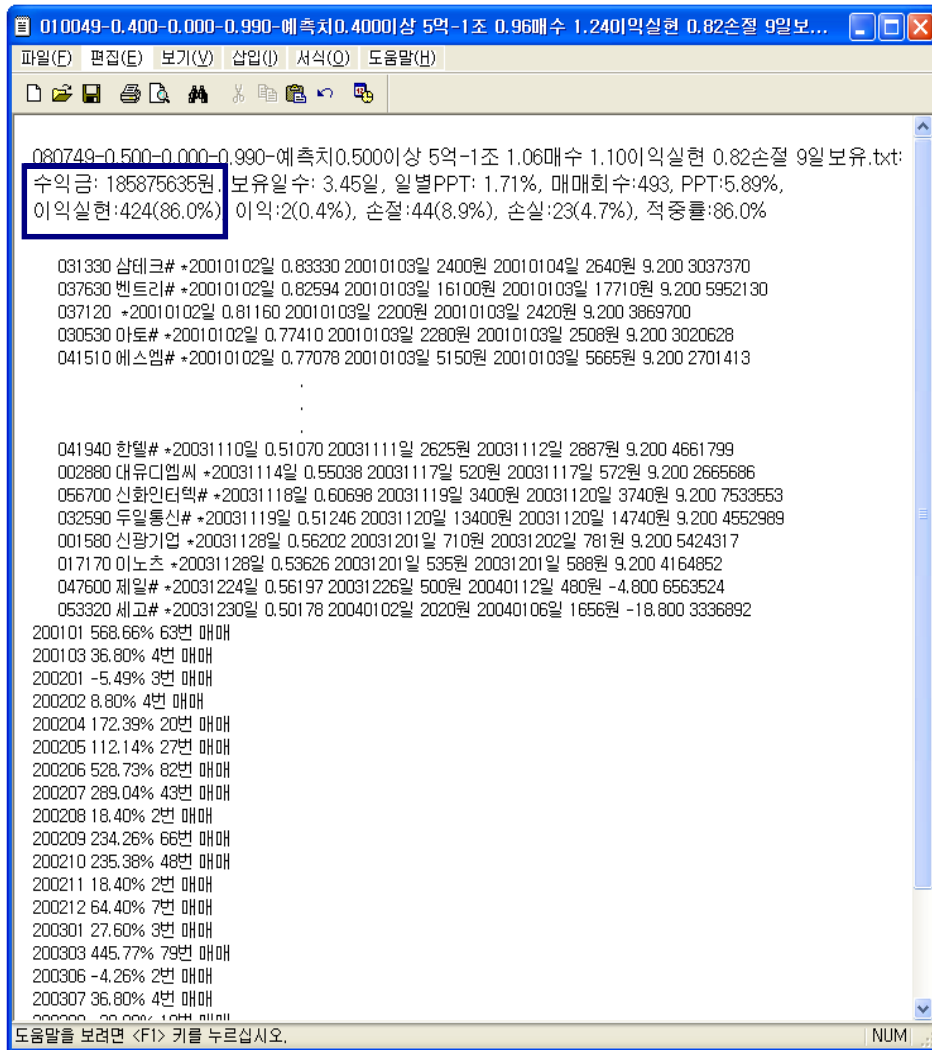


<그림 14> 결과파일의 예시

- ▶ 이익실현 - NNshell 프로그램의 실행 시 설정한 거래 정책인, 이익 실현율을 달성한 주식의 수이다. 그림 14의 경우 이익 실현율은 1.24로 설정되어 있으므로 24%의 이익실현율을 달성한 주식의 수가 전체 거래의 73.2% 이다.
- ▶ 이익 - 비록 목표치인 이익실현율을 달성하지는 못했지만, 수익을 얻은 주식의 수이다. 그림 14에서는 11개의 주식의 이익을 얻었으며 13.4%의 이익률을 보였다.
- ▶ 손절 - NNshell 프로그램 실행 시 설정한 손절을 만큼 하락하여, 매도한 주식의 수이다. 그림 14의 경우 손절율은 0.82로 설정되어 있으므로 -18%의 손해를 입어 매도한 주식의 수가 전체 거래의 12.2% 이다.
- ▶ 손실 - 손절을 만큼 하락하지는 않았으나, 손해를 입은 주식의 수이다. 그림 14에서는 1개의 주식의 손해를 입었으며 1.2%의 손실율을 보였다.

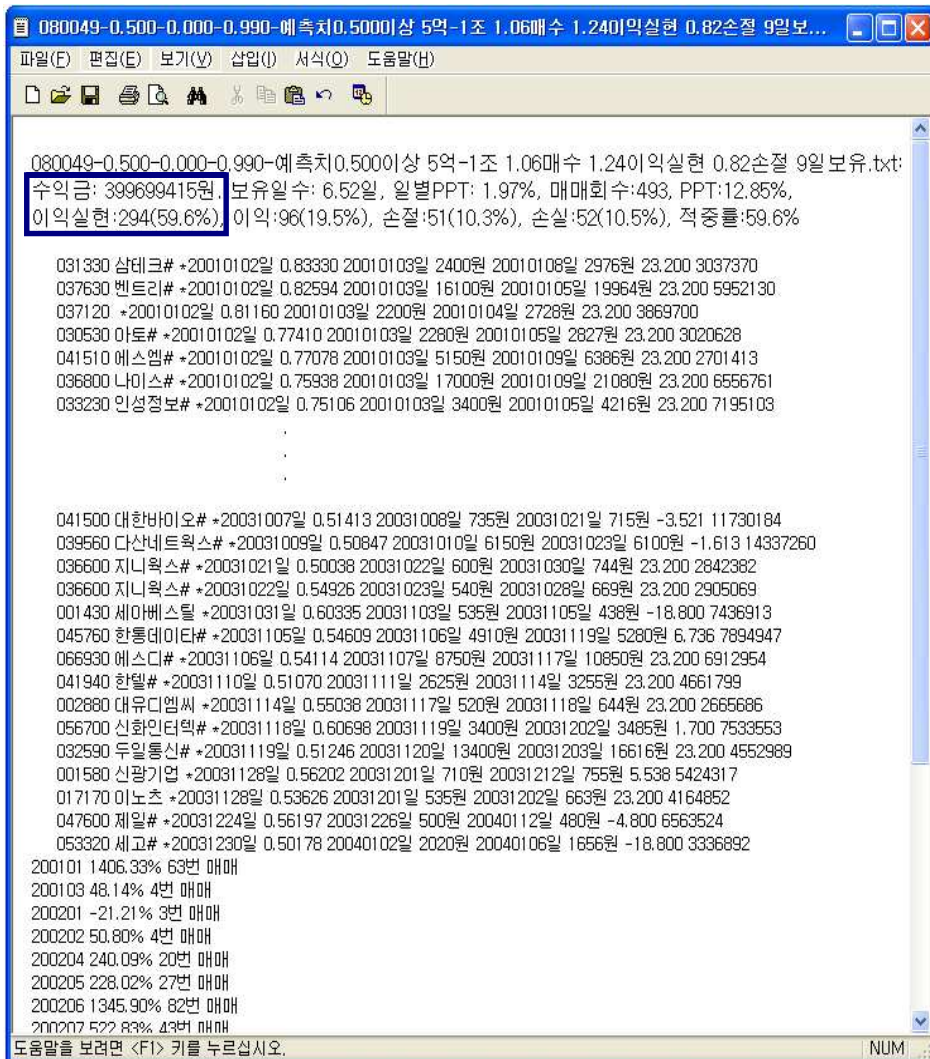
### 3.2 적중률 기준 소팅하기

적중률이 높은 결과 파일을 기준으로 하여 투자할 경우, 다른 요인들을 기준으로 한 결과 파일에 투자할 경우에 비하여 조금 더 안전한 투자를 할 수 있다.



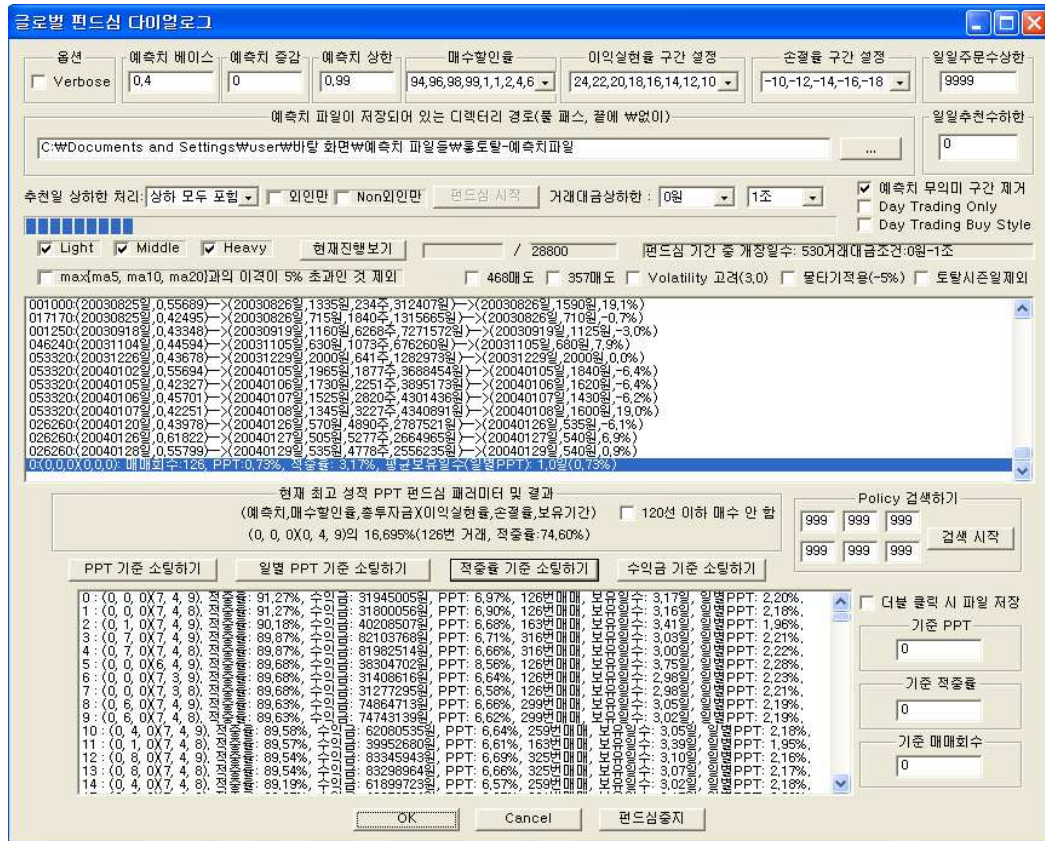
<그림 15> 이익실현율이 높은 결과파일 A

안전한 거래라는 것은 거래(매수, 매도)를 통해 수익을 얻는 경우이든, 손실을 보는 경우이든 결과 파일이 예측한 흐름대로 진행될 확률이 높다는 의미이다. 다음 그림 15, 16은 NNshell의 결과 파일들이다. 수익금을 비교하면 399,699,415원인 B가 185,875,635원인 A 보다 많다. 그러나 이익실현율, 다시 말해 적중률을 비교하면 A가 86.0%로 B의 적중률인 59.6%보다 높다.



<그림 16> 수익금이 많은 결과파일 B

결국, 투자자는 거래를 함에 있어, 불안감을 덜 느끼며 거래를 할 수 있는 것이다. 그러나 적중률이 높다고 해서 가장 많은 수익을 얻을 수 있는 것은 아니다.



<그림 17> 적중률 기준으로 소팅 했을 경우



# IV. 거래 시스템의 요소 분석

## 1. 거래 대금의 규모

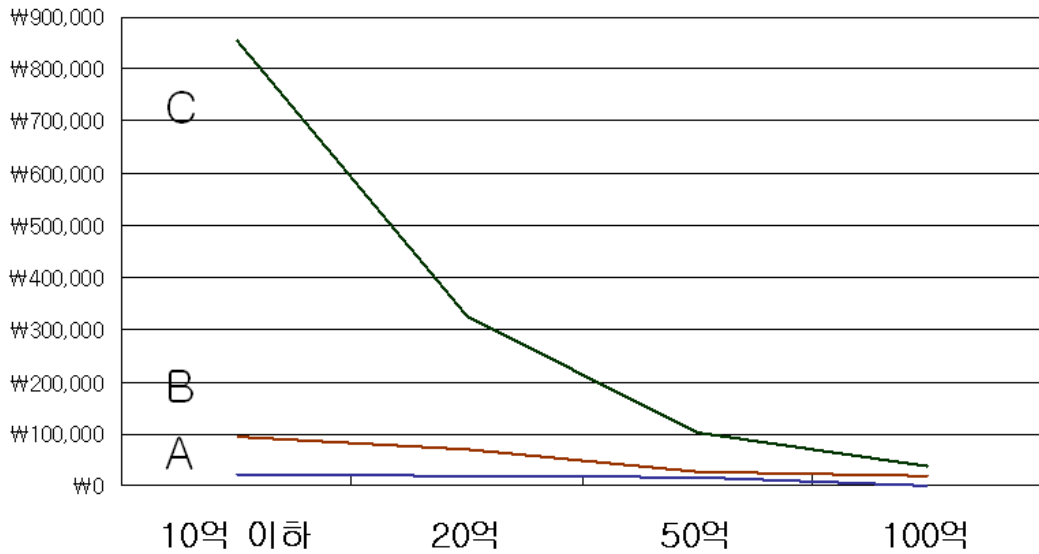
주식 시장에는 무수히 많은 종류의 주식들이 있으며 각 주식들마다 거래되는 금액에도 많은 차이가 있다. 투자자가 거래를 시도할 때, 거래 대금의 규모가 얼마나 되는 종목을 선택하는 것이 유리한지 알아본다. 거래대금의 상한보다는 하한이 더 많은 영향을 미치기 때문에 하한금액의 변화를 가지고 몇 가지 실험을 해 보았다.

A, B, C라는 각각의 엔진에 대하여 예측치 값은 0.4로 고정시킨 후, 거래 하한 금액만을 각기 다르게 설정한 후, NNshell 프로그램을 실행시켜 보았다. 그 결과는 다음 표 1과 같다(거래 하한 금액이 0원~10억 까지는 수익금의 변화가 없었다).

엔진명 거래하한금액	A	B	C
10억 이하	₩22,721,151	₩853,119,900	₩95,246,800
20억	₩18,960,478	₩325,050,697	₩68,557,626
50억	₩14,913,298	₩102,819,213	₩27,439,469
100억	₩10,600,000	₩38,739,042	₩18,794,286

<표 1> 거래 하한 금액의 변화에 따른 엔진별 투자 결과

표 1에서 보면, 거래대금의 하한값이 낮을수록 총 수익은 증가하며, 하한값이 높을수록 총수입은 감소한다. 거래대금의 하한값이 높다는 것은 거래되는 금액이 많다는 것, 즉, 대형주라는 것을 의미한다. 대형주의 특성상 PPT의 변화가 적기 때문에 같은 금액을 투자한다 해도 수익금이 적을 수밖에 없다. 하지만 주식이 상장폐지 되거나 급락할 가능성은 적기 때문에 위험성이 적다는 장점이 있다.



<그림 19> 거래 하한 금액의 변화에 따른 엔진별 투자 결과 그래프  
(수익금의 단위 : 천원)

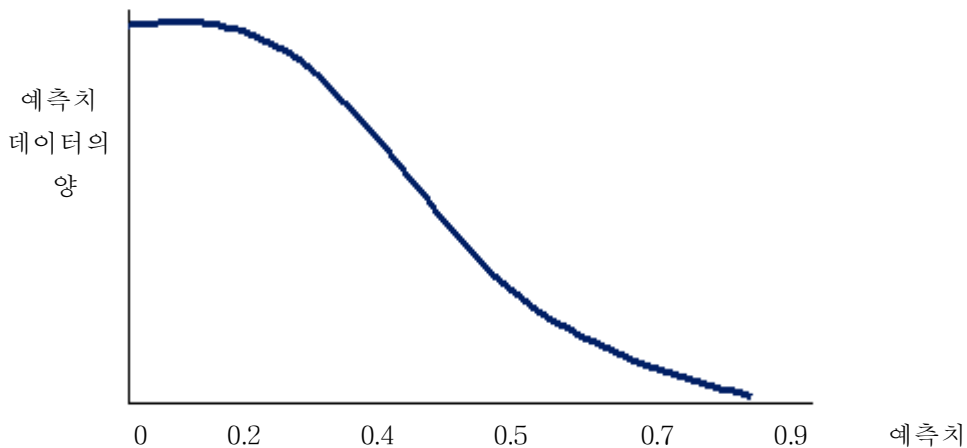
또한, 투자자의 투자 자금이 많을수록, 하한 값이 낮은 업종에 투자하면 불리하다. 그 이유는 하한 값이 낮은 업종, 즉, 거래대금이 적은 업종에 투자자의 많은 투자자금은 해당 업종의 전체적인 주가의 흐름에 영향을 미치기 때문이다.

## 2. 예측치 베이스

### 2.1 한 엔진에서의 예측치 베이스의 변화

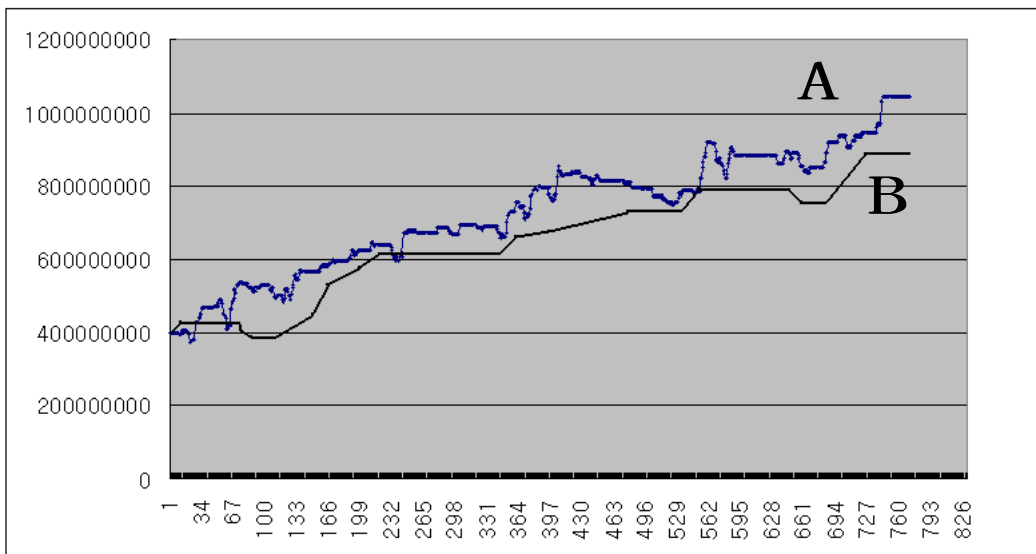
NNshell 프로그램을 실행시킬 때, 예측치 베이스 값을 결정하는 것은 까다로운 문제이다. 예측치 베이스를 낮게 설정했다는 것은, 예측치가 낮은 종목도 투자를 고려한다는 의미이므로 빈번한 매수·매도가 발생한다. 수익금만을 놓고 본다면, 예측치 베이스가 낮을수록 유리하다. 투자하는 주식의 종류가 많아지기 때문에 플러스 수익을 내는 종목과 마이너스 수익을 내는 종목 모두가 많아져서 전체적 수익은 증가하기 때문이다. 하지만 예측치 베이스가 낮다는 의미는 다시 말하면, 예측의 정확도가 떨어진다는 의미이다. 예측의 정확도가 떨어지기 때문에 투자를 통해 마이너스 수익을 얻을 수 있는 위험성도 증가한다.

그림 20은 예측치의 데이터 분포를 나타내고 있다.



<그림 20> 예측치에 따른 예측치 데이터의 분포

예측치의 변화에 따른 거래빈도와 수익률간의 관계를 알아보기 위해 한 가지 실험을 해 보았다. 하나의 엔진에서 예측치 값을 낮게 설정한 경우를 A, 예측치 값을 높게 설정한 경우를 B라 하고, 다른 속성들(예를 들면, 예측치 상한, 거래 대금 등)은 모두 같게 설정한 후 NNshell을 실행시켜 보았다. A의 경우, 프로그램이 추천한 종목 중, 예측치가 다소 낮은 종목이라도 거래를 시도한다. 그리하여 빈번한 매수·매도가 발생하며 수익률은 높다. 즉, 위험을 감수하더라도 수익률이 높은 결과를 추구하는 거래형태이다. 반면, B의 경우는 예측치가 높은 종목에만 집중적으로 거래를 시도하여, A에 비해 그리 많은 매수·매도가 발생하지는 않는다. 수익률은 낮지만, 안전을 우선시하는 거래 형태이다.



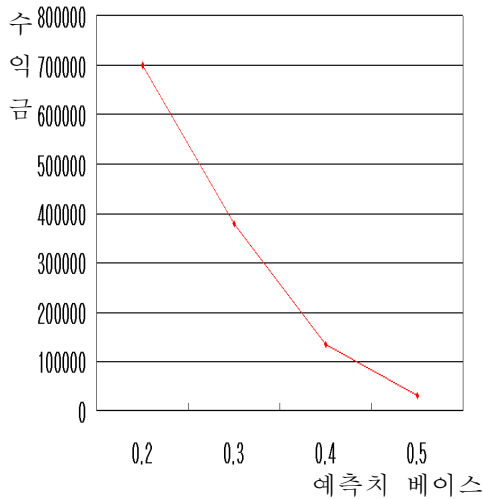
<그림 21> 예측치의 변화에 따른 거래빈도와 수익률간의 관계

## 2.2 2개의 다른 엔진에서의 예측치 베이스의 변화

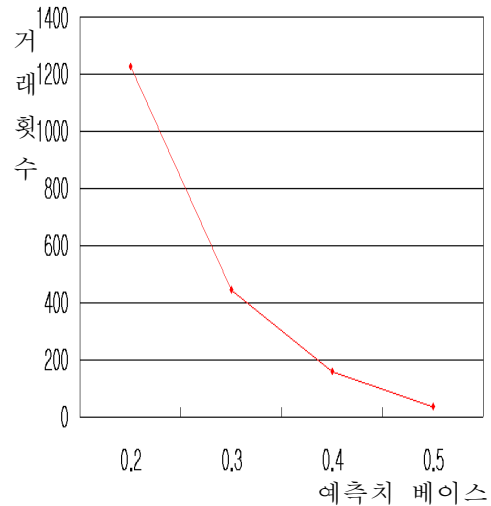
이번에는 각기 다른 2개의 엔진에서 예측치 베이스 값을 변화시켜 주어 실험해 보았다. 다음 표 2, 표 3은 A, B 라는 엔진에서 예측치 베이스 값을 변화시켜 주어 실험한 결과이다.

예측치 베이스	수익금	거래 횟수
0.5	₩30,537,067	33번
0.4	₩135,498,456	156번
0.3	₩377,819,126	442번
0.2	₩699,875,391	1226번

<표 2> A 엔진에서의 예측치 베이스의 변화에 따른 수익금과 거래 횟수의 변화



<그림 22> A 엔진에서의 수익금의 변화  
(수익금의 단위 : 천원)

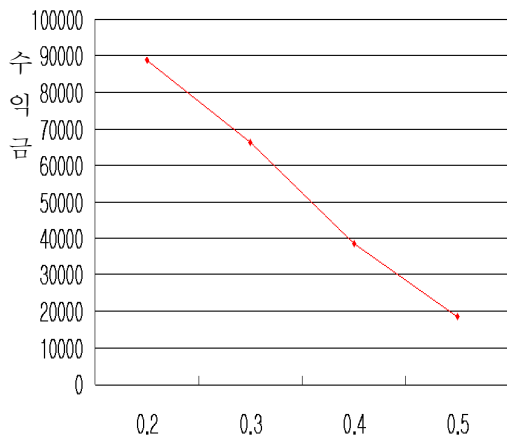


<그림 23> A 엔진에서의 거래횟수의 변화

A 엔진에서, 예측치 베이스의 변화에 따른 수익금의 변화를 그래프로 나타낸 것이 그림 22이고, 예측치 베이스 변화에 따른 거래횟수의 변화를 그래프로 나타낸 것이 그림 23이다. 예측치 베이스가 높아짐에 따라 수익금과 거래횟수가 모두 감소함을 알 수 있다.

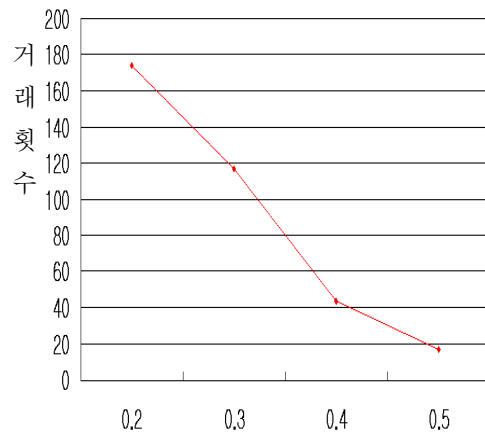
예측치 베이스	수익금	거래 횟수
0.5	₩18,470,574	17번
0.4	₩38,739,042	44번
0.3	₩66,138,584	117번
0.2	₩88,886,216	174번

<표 3> B 엔진에서의 예측치 베이스의 변화에 따른 수익금과 거래 횟수의 변화



예측치 베이스

<그림 24> B 엔진에서의 수익금의 변화  
(수익금의 단위 : 천원)



예측치 베이스

<그림 25> B 엔진에서의 거래횟수의 변화

표 3은 B 엔진에서의 예측치 베이스의 변화에 따른 수익금과 거래 횟수의 변화를 나타낸 것이다. 그림 24, 그림 25는 표 3의 수익금과 거래횟수를 나타낸 그래프이다.

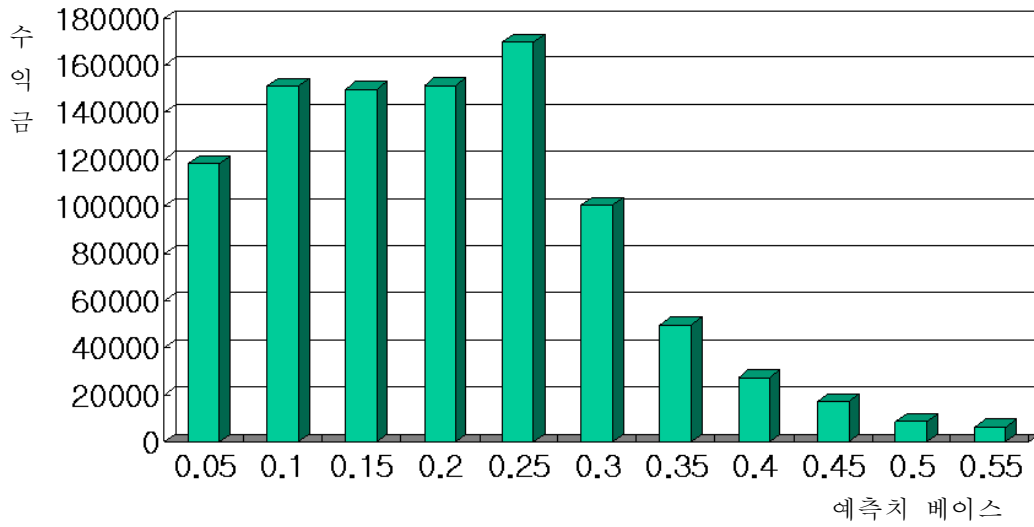
위의 결과와 같이, 한 엔진의 내부에서 예측치 값을 바꾸는 경우나 각기 다른 엔진에서 예측치 값을 바꾸는 경우 모두, 같은 형태의 그래프가 나타난다. 즉, 어떤 엔진이든 예측치 베이스 값이 낮으면 수익금은 많아지고, 거래는 빈번하게 이루어지며, 예측치 베이스 값이 높으면 수익금은 작아지고 거래의 횟수는 줄어든다.

## 2.3 적절한 예측치 베이스 값

위의 실험에 의하면 예측치 베이스 값이 낮을수록 수익금은 증가했다. 그렇다면 예측치 베이스 값을 계속 낮추면 수익금도 계속 증가하는 것인지, 수익금이 최대가 되는 예측치 베이스 값은 얼마인지 알아보기 위해 실험을 해 보았다. 표 1에서 언급한 A, B, C 엔진에 대하여 베이스 값을 계속 낮추어 가며 수익금의 변화를 관찰했다. 실험 결과 세 엔진 모두 수익금 변화의 흐름이 동일했기 때문에 이 논문에서는 엔진 C의 실험 결과만 언급했다.

예측치 베이스	수익금
0.55	₩ 6,200,000
0.50	₩ 8,578,789
0.45	₩ 17,259,200
0.40	₩ 27,439,469
0.35	₩ 49,884,048
0.30	₩ 100,504,458
0.25	₩ 169,857,132
0.20	₩ 151,488,038
0.15	₩ 149,918,357
0.10	₩ 151,142,065
0.05	₩ 118,083,061

<표 4> 예측치 변화에 따른 수익금의 변화



<그림 26> 예측치 변화에 따른 수익금의 변화 그래프

그림 26에서 보면 예측치가 낮을수록 수익금이 커지다가 예측치 베이스 0.25를 기준으로 수익금이 다시 낮아짐을 확인할 수 있었다. 예측치 베이스가 낮아진다고 해서 수익금이 끝없이 높아지는 것이 아니며, 어느 시점이 되면 수익금도 낮아지는 지점이 있다. 이 지점을 본 논문에서는 최소 예측 임계치라 정의한다. 최소 예측 임계치는 예측력 저하에 대한 현실적 기준이 될 수 있다. 이 최소예측 임계치는 엔진마다 다소 차이가 있으나, 이 엔진에서는 0.25이다.

## V. 결론

주식 시장을 예측하려는 기존의 많은 연구들에 더하여, 투자 기간 동안의 투자자의 심리 등과 같은 기존의 단순한 성능평가 척도들만으로는 표현될 수 없는 문제들을 반영할 수 있는 다양한 성능 평가 수단이 추가로 고려될 필요가 있다.

본 논문에서는 ‘NNshell’이라는 신경망 기반 주가 예측 시스템의 성능 평가와 시스템 최적화를 위한 통합 시뮬레이션 도구의 분석을 통해, 보다 체계화된 성능 평가 및 거래 시스템의 거래 성능 요소 분석을 시도하였다.

통합 시뮬레이션 거래 시스템의 성능평가에서는 PPT 기준 소팅, 일별 PPT 기준 소팅, 적중률 기준 소팅, 수익금 기준 소팅의 특징과 각 평가수단이 유용한 경우를 논하였다.

거래 시스템의 최적화에서는 거래 대금의 규모와 예측치 베이스에 변화를 주어 실험해 보았다. 그 결과, 거래 대금의 하한값이 낮을수록 투자자의 총 수익은 증가하며, 하한값이 높을수록 총 수익은 감소한다. 거래대금의 하한값이 높은 대형주의 특성상 PPT의 변화가 적기 때문에 같은 금액을 투자한다 해도 수익금이 적기 때문이다. 하지만 주식이 상장폐지 되거나 급락할 가능성이 거의 없으므로 위험성이 적다는 장점이 있다.

한 엔진의 내부에서 예측치 값만을 바꾸는 경우나 각기 다른 엔진에서 예측치 값만을 바꾸는 경우 모두, 같은 형태의 그래프 흐름이 나타난다. 즉, 어떤 엔진이든 예측치 베이스 값이 낮아지면 수익금은 많아지고, 거래는 빈번하게 이루어지며, 예측치 베이스 값이 높아지면 수익금은 작아지고 거래의 횟수는 줄어든다. 그러나 예측치 베이스가 낮아진다고 해서 수익금이 끝없

이 높아지는 것이 아니며, 어느 시점이 되면 수익금도 낮아지는 지점이 있으며, 이 지점을 본 논문에서는 최소 예측 임계치라 정의한다. 이 임계치 이하의 예측치를 가진 투자 종목의 경우 기대 수익률이 현저히 저하되는 것으로 판단할 수 있다.

이 논문의 실험결과는 소규모 투자에 대한 경우가 대부분이다. 대규모 자금 투자의 경우에는 투자자의 자금이 해당 주가의 전체적인 흐름에 영향을 미칠 수 있기 때문에, 이 논문에 제시된 거래 성능(PPT)을 달성할 수 없을 것으로 판단된다.

향후 연구 과제로는, 'NNshell' 시스템에서 추천한 종목에 투자할 경우, 투자자의 효율적인 자산 분배 정책에 대한 연구가 추가적으로 이루어져야 할 것으로 보인다.

## 참고문헌

- [1] J. O, J. W. Lee, J. Lee, and B. chae, "Dynamic Asset Allocation Exploiting Predictors in Reinforcement Learning Framework", European Conference on Machine Learning, pp. 298~309, 2004
- [2] J. W. Lee, S. Kim, J. Lee, and J. Chae, "An Intelligent stock Trading System based on Reinforcement Learning", IEICE Trans. on Information and Systems, 2003.
- [3] E. W. Saad, D. V. Prokhorov, and D. C. Wunsch, Comparative Study of Stock Trend Prediction Using Time Delay, Recurrent and Probabilistic Neural Networks, IEEE Trans. on Neural Networks, Vol. 9, No. 6, 1998.
- [4] S. T. Chou, H. Hsu, C. Yang, and F. Lai, "A Stock Selection DSS Combining AI and Technical Analysis", Annals of Operations Research, Vol. 75, pp. 335-353, 1997.
- [5] T. Chenoweth, Z. Obradovic, and S. Lee, "Embedding Technical Analysis into Neural Networks Based Trading Systems," Applied Artificial Intelligence Journal, 1996.
- [6] T Hellstrom "A Random walk through the stock market" ph D. Thesis. Department of Computing Science. Umea University. 1998.
- [7] Tim chenoweth, Zoran obradovic, Sauchi Stephen Lee, "Neural Network Trading Systems", 1996

- [8] James D Thomas, Katia Sycara “Integrating Genetic Algorithms and Text Learning for Financial Prediction” , 2000.
- [9] Journana Ghosn, Yoshua Bengio, “Multi Task Learning for Stock Selection” , pp 947-952, 1998
- [10] B. G. Malkiel, A Random Walk Down Wall Street, Norton, New York, 1996.
- [11] 김규영, 김영빈, “한국 주식시장에서 수익률의 예측에 관한 연구 ”, 한국 산업경제학회, Vol. 17, No. 4, pp.1255-1371, 2004
- [12] 최순재, 오정택, “주가장부가치비율 등의 변수를 이용한 주가수익률 예측 모형 및 투자성과에 대한 분석”, 서강경영논총, Vol. 13, No 1, pp. 129-158, 2002

# ABSTRACT

## A Study on Performance Evaluation Parameters of Neural Network based Stock Price Forecasting System

Kim, Hyun Jin  
Major in Computer Science Education  
Graduate School of Education  
Sungshin Women's University

Although there were a lot of researches to forecast the stock market, it is difficult to forecast the fluctuation of price and the expected return in the stock market, Owing to the 'Efficient Market Hypothesis', market analysis methodology was considered skeptical. Nowadays the research workers in the field of machine learning are showing us the result of studies which prove that it is possible to make excess profits through the combining artificial intelligence technique with technical analysis technique. In this thesis, I will demonstrate more systematized performance evaluation and the optimization cases of transaction performance of a stock trading system, through the analysis of an integrated simulation tool for neural network-based stock price forecasting system, called NNshell.