Optimizing the Sharpe Ratio for a Rank Based Trading System

Thomas Hellström

Department of Computing Science
Umeå University, 901 87 Umeå, Sweden
thomash@cs.umu.se
http://www.cs.umu.se/~thomash

Abstract. Most models for prediction of the stock market focus on individual securities. In this paper we introduce a rank measure that takes into account a large number of securities and grades them according to the relative returns. It turns out that this rank measure, besides being more related to a real trading situation, is more predictable than the individual returns. The ranks are predicted with perceptrons with a step function for generation of trading signals. A learning decision support system for stock picking based on the rank predictions is constructed. An algorithm that maximizes the Sharpe ratio for a simulated trader computes the optimal decision parameters for the trader. The trading simulation is executed in a general purpose trading simulator ASTA. The trading results from the Swedish stock market show significantly higher returns and also Sharpe ratios, relative the benchmark.

1 Introduction

The returns of individual securities are the primary targets in most research that deal with the predictability of financial markets. In this paper we focus on the observation that a real trading situation involves not only attempts to predict the individual returns for a set of interesting securities, but also a comparison and selection among the produced predictions. What an investor really wants to have is not a large number of predictions for individual returns, but rather a grading of the securities in question. Even if this can be achieved by grading the individual predictions of returns, it is not obvious that it will yield an optimal decision based on a limited amount of noisy data. In Section 2 we introduce a rank measure that takes into account a large number of securities and grades them according to the relative returns. The rank concept has in a previous study [5] shown a potential for good predictability. In Section 4, perceptron models for prediction of the rank are defined and historical data is used to estimate the parameters in the models. Results from time series predictions are presented. The predictions are used as a basis for a learning decision support system for stock picking described in Section 5. The surprisingly successful results are discussed. Section 6 contains a summary of the results together with ideas for future research.

2 Defining a Rank Measure

The k-day return $R_k(t)$ for a stock m with close prices $y^m(1), ..., y^m(t_1)$ is defined for $t \in [k+1, ..., t_1]$ as

$$R_k^m(t) = \frac{y^m(t) - y^m(t-k)}{y^m(t-k)}. \tag{1}$$

We introduce a rank concept A_k^m, based on the k-day return R_k as follows: The k-day rank A_k^m for a stock s_m in the set $\{s_1, ..., s_N\}$ is computed by ranking the N stocks in the order of the k-day returns R_k. The ranking orders are then normalized so the stock with the lowest R_k is ranked -0.5 and the stock with the highest R_k is ranked 0.5. The definition of the k-day rank A_k^m for a stock m belonging to a set of stocks $\{s_1, ..., s_N\}$, can thus be written as

$$A_k^m(t) = \frac{\#\{R_k^i(t) | R_k^m(t) \geq R_k^i(t), 1 \leq i \leq N\} - 1}{N-1} - 0.5 \tag{2}$$

where the # function returns the number of elements in the argument set. This is as integer between 1 and N. R_k^m is the k-day returns computed for stock m. The scaling between -0.5 and $+0.5$ assigns the stock with the median value on R_k the rank 0. A positive rank A_k^m means that stock m performs better than this median stock, and a negative rank means that it performs worse. This new measure gives an indication of how each individual stock has developed relatively to the other stocks, viewed on a time scale set by the value of k.

The scaling around *zero* is convenient when defining a prediction task for the rank. It is clear that an ability to identify, at time t, a stock m, for which $A_h^m(t+h) > 0, h > 0$ means an opportunity to make profit in the same way as identifying a stock, for which $R_h(t+h) > 0$. A method that can identify stocks m and times t with a mean value of $A_h^m(t+h) > 0, h > 0$, can be used as a trading strategy that can do better than the average stock. The hit rate for the predictions can be defined as the fraction of times, for which the sign of the predicted rank $A_h^m(t+h)$ is correct. A value greater than 50% means that true predictions have been achieved. The following advantages compared to predicting returns $R_h(t+h)$ can be noticed:

1. The benchmark for predictions of ranks $A_h^m(t+h)$ performance becomes clearly defined:
   - A hit rate $> 50\%$, for the predictions of the sign of $A_h^m(t+h)$ means that we are doing better than chance. When predicting returns $R_h(t+h)$, the general positive drift in the market causes more than 50% of the returns to be > 0, which means that it is hard to define a good benchmark.
   - A positive mean value for predicted positive ranks $A_h(t+h)$ (and a negative mean value for predicted negative ranks) means that we are doing better than chance. When predicting returns $R_h(t+h)$, the general positive drift in the market causes the returns to have a mean value > 0. Therefore, a mere positive mean return for predicted positive returns does not imply any useful predicting ability.

2. The rank values $A_k^1(t), ..., A_k^N(t)$, for time t and a set of stocks $1, ..., N$ are uniformly distributed between -0.5 and 0.5 provided no return values are equal. Returns R_k^m, on the other hand, are distributed with sparsely populated tails for the extreme low and high values. This makes the statistical analysis of rank predictions safer and easier than predictions of returns.
3. The effect of global events gets automatically incorporated into the predictor variables. The analysis becomes totally focused on identifying deviations from the average stock, instead of trying to model the global economic situation.

3 Serial Correlation in the Ranks

We start by looking at the serial correlation for the rank variables as defined in (2). In Table 1 mean ranks $A_1^m(t+1)$ are tabulated as a function of $A_k^m(t)$ for 207 stocks from the Swedish stock market 1987-1997. Table 2 shows the "*Up fraction*", i.e. the number of positive ranks $A_1^m(t+1)$ divided by the number of non-zero ranks. Table 3 finally shows the number of observations of $A_1^m(t+1)$ in each table entry. Each row in the tables represents one particular value on k, covering the values $1, 2, 3, 4, 5, 10, 20, 30, 50, 100$. The label for each column is the mid-value of a symmetrical interval. For example, the column labeled 0.05 includes points with k-day rank $A_k^m(t)$ in the interval [0.00, ..., 0.10 [. The intervals for the outermost columns are open-ended on one side. Note that the stock price time series normally have 5 samples per week, i.e. $k = 5$ represents one week of data and $k = 20$ represents approximately one month. Example: There are 30548 observations where $-0.40 \leq A_2^m(t) < -0.30$ in the investigated data. In these observations, the 1-day ranks on the following day, $A_1^m(t+1)$, have an average value of 0.017, and an "*Up fraction*" = 52.8%.

Table 1. Mean 1-step ranks for 207 stocks

	k-day rank									
k	**-0.45**	**-0.35**	**-0.25**	**-0.15**	**-0.05**	**0.05**	**0.15**	**0.25**	**0.35**	**0.45**
1	0.067	0.017	-0.005	-0.011	-0.011	-0.004	-0.005	-0.010	-0.014	-0.033
2	0.060	0.017	0.002	-0.004	-0.010	-0.003	-0.007	-0.015	-0.017	-0.032
3	0.057	0.016	0.003	-0.005	-0.003	-0.008	-0.011	-0.011	-0.015	-0.034
4	0.054	0.018	0.003	-0.003	-0.005	-0.008	-0.011	-0.013	-0.012	-0.032
5	0.051	0.015	0.004	-0.002	-0.004	-0.009	-0.010	-0.009	-0.016	-0.032
10	0.040	0.013	0.005	-0.001	-0.003	-0.006	-0.007	-0.009	-0.012	-0.030
20	0.028	0.008	0.003	-0.003	-0.002	-0.002	-0.006	-0.011	-0.009	-0.019
30	0.021	0.007	0.002	0.004	-0.003	-0.003	-0.006	-0.006	-0.011	-0.015
50	0.014	0.005	0.000	-0.000	-0.001	-0.002	-0.005	-0.004	-0.006	-0.010
100	0.007	0.003	0.001	-0.002	-0.003	-0.004	-0.004	-0.004	-0.003	-0.008

Table 2. Fraction up/(up+down) moves (%)

	k-day rank									
k	**-0.45**	**-0.35**	**-0.25**	**-0.15**	**-0.05**	**0.05**	**0.15**	**0.25**	**0.35**	**0.45**
1	59.4	52.9	49.1	47.3	48.0	49.6	49.5	48.2	47.8	46.4
2	58.4	52.8	49.7	48.9	48.4	49.7	48.9	47.4	47.6	46.2
3	58.1	52.4	50.3	48.9	49.1	49.0	48.1	48.1	47.8	46.1
4	57.5	52.5	50.4	49.2	49.0	48.7	48.0	47.9	48.6	46.3
5	57.1	52.0	50.4	49.4	49.1	48.5	48.2	48.6	47.7	46.3
10	55.6	51.7	50.4	49.8	49.3	48.8	48.7	48.5	48.2	46.3
20	53.8	51.1	50.2	49.6	49.4	49.5	49.0	48.3	48.7	47.8
30	52.7	50.9	50.3	50.8	49.1	49.2	48.8	48.9	48.5	48.4
50	52.0	50.7	49.6	49.9	49.6	49.6	49.0	49.3	49.1	48.9
100	51.4	50.4	49.9	49.5	49.2	49.2	49.0	49.2	49.6	49.1

The only clear patterns that can be seen in the table are a slight negative serial correlation: negative ranks are followed by more positive ranks and vice versa. To investigate whether this observation reflects a fundamental property of the process generating the data, and not only idiosyncrasies in the data, the relation between current and future ranks is also presented in graphs, in which one curve represents one year. Figure 1 shows $A_1^m(t+1)$ versus $A_1^m(t)$. I.e.: 1-day ranks on the following day versus 1-day ranks on the current day. The same relation for 100 simulated random-walk stocks is shown in Figure 2 for comparison.

From Figure 1 we can conclude that the rank measure exhibits a mean reverting behavior, where a strong negative rank in mean is followed by a positive rank. Furthermore, a positive rank on average is followed by a negative rank on the following day. Looking at the "*Up fraction*" in Table 2, the uncertainty in these relations is still very high. A stock m with a rank $A_1^m(t) < -0.4$ has a positive rank $A_1^m(t+1)$ the next day in no more than 59.4% of all cases. However, the general advantages described in the previous section, coupled with the shown correlation between present and future values, do make the rank variables very interesting for further investigations. In [3] the observed mean reverting behavior is exploited in a simple trading system. The rank measure in the next section is used both as input and output in a model for prediction of future ranks.

4 Predicting the Ranks with Perceptrons

For a stock m, we attempt to predict the h-day-rank h days ahead by fitting a function g_m so that

$$\hat{A}_h^m(t+h) = g_m(I_t) \tag{3}$$

where I_t is the information available at time t. I_t may, for example, include stock returns $R_k^m(t)$, ranks $A_k^m(t)$, traded volume etc. The prediction problem 3

Table 3. Number of points

	k-day rank									
k	**-0.45**	**-0.35**	**-0.25**	**-0.15**	**-0.05**	**0.05**	**0.15**	**0.25**	**0.35**	**0.45**
1	30878	30866	31685	30837	30434	31009	31258	30539	30951	31550
2	30926	30548	31427	30481	30442	31116	31263	30435	30841	31675
3	30922	30440	31202	30404	30350	31146	31061	30449	30814	31697
4	30887	30315	31052	30320	30371	31097	31097	30328	30777	31776
5	30857	30293	30951	30275	30191	31049	31144	30254	30701	31816
10	30755	30004	30648	29958	30004	30875	30889	30155	30571	31775
20	30521	29635	30306	29591	29679	30560	30580	29836	30377	31692
30	30388	29371	30083	29388	29567	30349	30437	29652	30190	31503
50	30117	29006	29728	28979	29306	29876	30109	29236	29927	31159
100	29166	28050	28790	28011	28238	29015	29049	28254	29012	30460

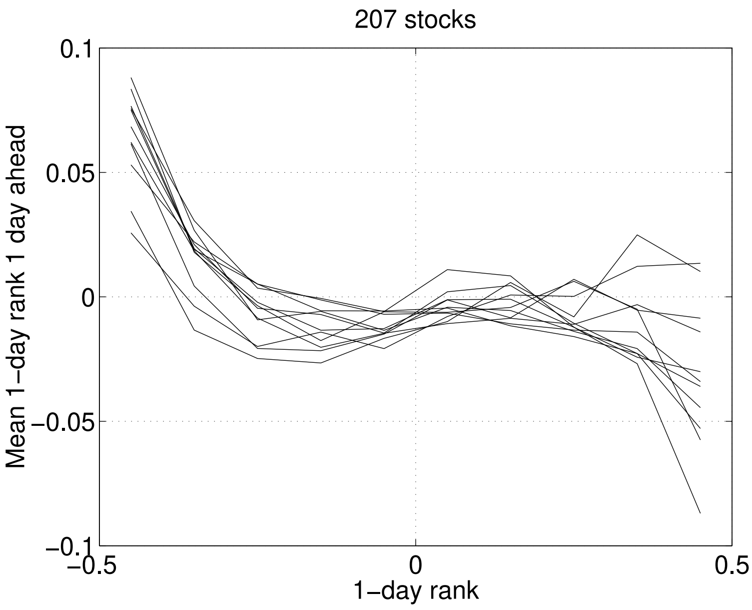

Fig. 1. 1-day ranks $A_1^m(t+1)$ versus $A_1^m(t)$. Each curve represents one year between 1987 and 1997.

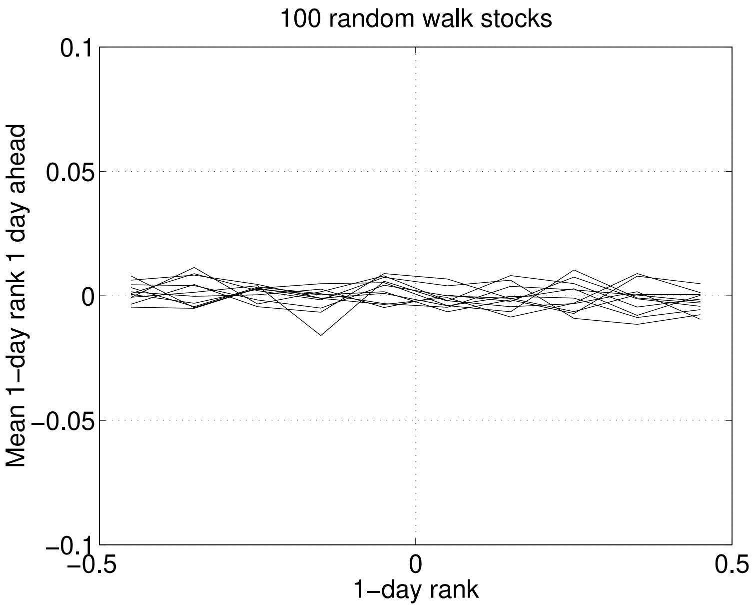

Fig. 2. 1-day ranks $A_1^m(t+1)$ versus $A_1^m(t)$. Each curve represents one year with 100 simulated random-walk stocks.

is as general as the corresponding problem for stock returns, and can of course be attacked in a variety of ways. Our choice in this first formulation of the problem assumes a dependence between the future rank $A_h^m(t+h)$ and current ranks $A_k^m(t)$ for different values on k. I.e.: a stock's tendency to be a *winner* in the future depends on its *winner* property in the past, computed for different time horizons. This assumption is inspired by the autocorrelation analysis in Hellström [5], and also by previous work by De Bondt, Thaler [1] and Hellström [3] showing how these dependencies can be exploited for prediction and trading. Confining our analysis to $1, 2, 5$ and 20 days horizons, the prediction model 3 is refined to

$$\hat{A}_h^m(t+h) = g_m(A_1^m(t), A_2^m(t), A_5^m(t), A_{20}^m(t)). \tag{4}$$

The choice of function g_m could be a complex neural network or a simpler function. Our first attempt is a perceptron, i.e. the model is

$$\begin{aligned}&\hat{A}_h^m(t+h) = \\ &f(p_0^m + p_1^m A_1^m(t) + p_2^m A_2^m(t) + p_3^m A_5^m(t) + p_4^m A_{20}^m(t))\end{aligned} \tag{5}$$

where the activation function f for the time being is set to a linear function. The parameter vector $p^m = (p_0^m, p_1^m, p_2^m, p_3^m, p_4^m)$ is determined by regression on historical data. For a market with N stocks, N separate perceptrons are built, each one denoted by the index m. The h-day rank A_h^m for time $t+h$ is predicted from the 1-day, 2-day, 5-day and 20-day ranks, computed at time t. To facilitate further comparison of the m produced predictions, they are ranked in a similar way as in the definition of the ranks themselves:

$$\begin{aligned}&\hat{A}_h^m(t+h) \leftarrow -0.5+ \\ &(\#\{\hat{A}_h^i(t) | \hat{A}_h^m(t) \geq \hat{A}_h^i(t), 1 \leq i \leq N\} - 1)\tfrac{1}{N}.\end{aligned} \tag{6}$$

In this way the N predictions $\hat{A}_h^m(t+h), m = 1, ..., N$, get values uniformly distributed between -0.5 and 0.5 with the lowest prediction having the value -0.5 and the highest prediction having the value 0.5.

4.1 Data and Experimental Set-Up

The data that has been used in the study comes from 80 stocks on the Swedish stock market from January 1, 1989 till December 31, 1997. We have used a sliding window technique, where 1000 points are used for training and the following 100 are used for prediction. The window is then moved 100 days ahead and the procedure is repeated until end of data. The sliding window technique is a better alternative than cross validation, since data at time t and at time $t+k, k>0$ is often correlated (consider for example the returns $R_5^m(t)$ and $R_5^m(t+1)$). In such a case, predicting a function value $A_1^m(t_1+1)$ using a model trained with data from time $t > t_1$ is cheating and should obviously be avoided. The sliding window approach means that a prediction $\hat{A}_h^m(t+h)$ is based on close prices $y^m(t-k), ..., y^m(t)$. Since 1000 points are needed for the modeling, the predictions are produced for the years 1993-1997.

4.2 Evaluation of the Rank Predictions

The computed models $g_m, m = 1, ..., N$ at each time step t produce N predictions of the future ranks $A_h^m(t+h)$ for the N stocks. The N predictions $\hat{A}_h^m, m = 1, ..., N$, are evenly distributed by transformation 6 in $[-0.5, ..., 0.5]$. As we shall see in the following section, we can construct a successful trading system utilizing only a few of the N predictions. Furthermore, even viewed as N separate predictions, we have the freedom of rejecting predictions if they are not viewed as reliable or profitable[1]. By introducing a cut-off value γ, a selection of predictions can be made. For example, $\gamma = 0.49$ means that we are only considering predictions $\hat{A}_h^m(t+h)$ such that $|\hat{A}_h^m(t+h)| > 0.49$.

The results for 1-day predictions of 1-day ranks $\hat{A}_1^m(t+1)$ for a $\gamma = 0.0$ and 0.49 are presented in Tables 4 and 5. Each column in the tables represents performance for one trading year with the rightmost column showing the mean values for the entire time period. The rows in the table contain the following performance measures:

1. $Hitrate_+$. The fraction of predictions $\hat{A}_h^m(t+h) > \gamma$, with correct sign. A value significantly higher than 50% means that we are able to identify higher-than-average performing stocks better than chance.
2. $Hitrate_-$. The fraction of predictions $\hat{A}_h^m(t+h) < -\gamma$, with correct sign. A value significantly higher than 50% means that we are able to identify lower-than-average performing stocks better than chance.
3. $Return_+$. 100·Mean value of the h-day returns $R_h^m(t+h)$ for predictions $\hat{A}_h^m(t+h) > \gamma$.
4. $Return_-$. 100·Mean value of the h-day returns $R_h^m(t+h)$ for predictions $\hat{A}_h^m(t+h) < -\gamma$.
5. $\#Pred_+$. Number of predictions $\hat{A}_h^m(t+h) > \gamma$.
6. $\#Pred_-$. Number of predictions $\hat{A}_h^m(t+h) < -\gamma$.
7. $\#Pred$. Total number of predictions $\hat{A}_h^m(t+h)$.

All presented values are average values over time t and over all involved stocks m. The performance for the one-day predictions are shown in the Tables 4 and 5. In Table 4 with $\gamma = 0.00$, the hit rates $Hitrate_+$ and $Hitrate_-$ are not significantly different from 50% and indicate low predictability. However, the difference between the mean returns ($Return_+$ and $Return_-$) for positive and negative rank predictions shows that the sign of the rank prediction really separates the returns significantly. By increasing the value for the cut-off value γ to $\gamma = 0.49$, the hit rate goes up to 64.2.0% for predicted positive ranks (Table 5). Furthermore, the difference between the mean returns for positive and negative rank predictions ($Return_+$ and $Return_-$) is substantial. Positive predictions of ranks are in average followed by a return of 0.895% while a negative rank prediction in average is followed by a return of 0.085%. The rows $\#Pred_+$ and $\#Pred_-$ show

[1] As opposed to many other applications, where the performance has to be calculated as the average over the entire data set.

the number of selected predictions, i.e. the ones greater than γ and the ones less than $-\gamma$ respectively. For $\gamma = 0.49$ these numbers add to about 2.7% of the total number of predictions. This is normally considered insufficient when single securities are predicted, both on statistical grounds and for practical reasons (we want decision support more often than a few times per year). But since the ranking approach produces a uniformly distributed set of predictions each day (in the example 80 predictions) there is always at least one selected prediction for each day, provided $\gamma < 0.5$. Therefore, we can claim that we have a method by which, every day we can pick a stock that goes up more than the average stock the following day with probability 64%. This is by itself a very strong result compared to most published single-security predictions of stock returns (see for example Burgess and Refenes [2], Steurer [8] or Tsibouris and Zeidenberg [9]).

5 Decision Support

The rank predictions are used as basis for a decision support system for stock picking. The layout of the decision support system is shown in Figure 3. The 1-day predictions $\hat{A}_1^m(t+1)$ are fed into a decision maker that generates buy and sell signals that are executed by the ASTA trading simulator. The decision maker is controlled by a parameter vector x, comprising the threshold values of the step functions that generate buy and sell signals from the rank predictions. Two other parameters control the amount of money the trader is allowed to invest in one single trade, and how large a fraction of the total wealth should be kept in cash at all times. The learning element comprises an optimizing algorithm, which finds the parameter vector x that maximizes the mean annualized Sharpe ratio for the simulated trader. The learning period is 2 years and the found optimal x is then used to control the trading in the following year. The procedure is repeated yearly, using the last 2 years as learning period. The ASTA system is a general-purpose tool for development of trading and prediction algorithms. A technical overview of the system can be found in Hellström [4] and examples of usage in Hellström [3] and Hellström, Holmström [6]. More information can also be found at http://www.cs.umu.se/~thomash. The rank measure and also the prediction algorithm described in Section 4 is implemented in ASTA and therefore the test procedure is very straightforward. A transaction cost of 0.15% (minimum 90 Swedish crowns $\sim$ 10 USD) is assumed for every buy or sell order.

5.1 Trading Results

The annual trading profit is presented in Table 6. As can be seen, the performance is very good. The trading strategy outperforms the benchmark (the Swedish Generalindex) consistently and significantly every year and the mean annual profit made by the trading is 129.4%. The mean annual index increase during the same period is 27.4%. The Sharpe ratio which gives an estimate of a risk adjusted return shows the same pattern. The average Sharpe ratio for the trading strategy is 3.0 while trading the stock index Generalindex gives 1.6.

By studying the annual performance we can conclude that these differences in performance is consistent for every year 1993-1997. Further more, the number of trades every year is consistently high (seven buy and seven sell per week), which increases the statistical credibility of the results. The trading results are also displayed in Figure 4. The upper diagram shows the equity curves for the trading strategy and for the benchmark index. The lower diagram shows the annual profits.

Let us look at possible reasons and mechanisms that may lie behind the good results. In [7], Lo and MacKinley report on positive *cross-autocovariances* across securities. These cross effects are most often positive in sign and are characterized by a *lead-lag* structure where returns for large-capitalization stocks tend to lead those of smaller stocks. Initial analysis of the trades that the rank strategy generates, expose a similar pattern, where most trading signals are generated for companies with relatively low traded volume. A positive *cross-autocovariances* can therefor provide part of an explanation to the successful trading results.

6 Conclusions

We have successfully implemented a model for prediction of a new rank measure for a set of stocks. The shown result is clearly a refutation of the Random Walk Hypothesis (RWH). Statistics for the 1-day predictions of ranks show that we are able to predict the sign of the threshold-selected rank consistently over the investigated 5-year-period of daily predictions. Furthermore, the mean returns that accompany the ranks show a consistent difference for positive and negative predicted ranks which, besides refuting the RWH, indicates that the rank concept could be useful for portfolio selection. The shown experiment with an optimizing trading system shows that this is indeed the case. The mean annual profit is 129.4% compared to 27.4% for the benchmark portfolio, over the investigated 5-year-period. The risk adjusted return, as measured by the Sharpe ratio, exhibits the same relation. The trading system gives a Sharpe ratio of 3.0 while trading the benchmark portfolio gives only 1.6.

Of course, the general idea of predicting ranks instead of returns can be implemented in many other ways than the one presented in this paper. Replacing the perceptrons with multi layer neural networks and also adding other kind of input variables to the prediction model (4) are exciting topics for future research.

References

1. W. D. Bondt and R. Taler. Does the stock market overreact? *Journal of Finance*, 40(3):793–805, 1985.
2. A. N. Burgess and A. N. Refenes. The use of error feedback terms in neural network modelling of financial time series. In C. Dunis, editor, *Forecasting Financial Markets*, Financial Economics and Quantitative Analysis, pages 261–274. John Wiley & Sons, Chichester, England, 1996.
3. T. Hellström. ASTA - a Tool for Development of Stock Prediction Algorithms. *Theory of Stochastic Processes*, 5(21)(1-2):22–32, 1999.

4. T. Hellström. ASTA - User's Reference Guide. Technical Report UMINF-00.16 ISSN-0348-0542, Department of Computing Science Umeå University, Umeå Sweden, 2000.
5. T. Hellström. Predicting a Rank Measure for Portfolio Selection. *Theory of Stochastic Processes*, 6(22)(3-4):64–83, 2000.
6. T. Hellström and K. Holmström. Parameter Tuning in Trading Algorithms using ASTA. In Y. S. Abu-Mostafa, B. LeBaron, A. W. Lo, and A. S. Weigend, editors, *Computational Finance 1999*, pages 343–357, Cambridge, MA, 1999. MIT Press.
7. A. W. Lo and A. C. MacKinley. *A Non-Random Walk Down Wall Street*. Princeton University Press, Princeton, 1999.
8. E. Steurer. Nonlinear modelling of the DEM/USD exchange rate. In A.-P. Refenes, editor, *Neural Networks in the Capital Markets*, chapter 13, pages 199–212. John Wiley & Sons, Chichester, England, 1995.
9. G. Tsibouris and M. Zeidenberg. Testing the efficent markets hypothesis with gradient descent algorithms. In A.-P. Refenes, editor, *Neural Networks in the Capital Markets*, chapter 8, pages 127–136. John Wiley & Sons, Chichester, England, 1995.

Table 4. 1-day predictions of 1-day ranks $|\hat{A}_1(t+1)| > 0.00$

$Year:$	**93**	**94**	**95**	**96**	**97**	**93-97**
$Hitrate_+$	51.1	53.4	53.3	53.0	52.5	52.7
$Hitrate_-$	51.8	53.6	53.4	53.2	52.6	52.9
$Return_+$	0.389	0.101	0.155	0.238	0.172	0.212
$Return_-$	0.253	-0.176	-0.094	0.057	0.008	0.010
$\#Pred_+$	7719	8321	8313	8923	8160	41510
$\#Pred_-$	7786	8343	8342	8943	8172	41664
$\#Pred$	15505	16664	16655	17866	16332	83174

Table 5. 1-day predictions of 1-day ranks $|\hat{A}_1(t+1)| > 0.49$

$Year:$	**93**	**94**	**95**	**96**	**97**	**93-97**
$Hitrate_+$	59.7	65.1	67.9	66.7	61.2	64.2
$Hitrate_-$	52.7	53.2	56.4	59.4	56.7	55.7
$Return_+$	1.468	0.583	0.888	0.770	0.745	0.895
$Return_-$	1.138	-0.236	-0.402	-0.040	-0.055	0.085
$\#Pred_+$	211	215	218	228	214	1088
$\#Pred_-$	222	220	220	234	217	1115
$\#Pred$	15505	16664	16655	17866	16332	83174

Table 6. Trading results for the trading system shown in Figure 3.

Year :	**93**	**94**	**95**	**96**	**97**	Mean	Total
Profit	120.8	116.7	142.9	165.8	100.8	129.4	6102
Index profit	52.1	4.6	18.3	38.2	23.8	27.4	222
Diff.	68.7	112.2	124.6	127.7	76.9	102.0	5880
#.trades	640	706	700	770	671	697	3487
Sharpe	1.7	2.4	3.2	4.5	3.1	3.0	
Index sharpe	2.8	0.1	1.4	2.4	1.4	1.6	

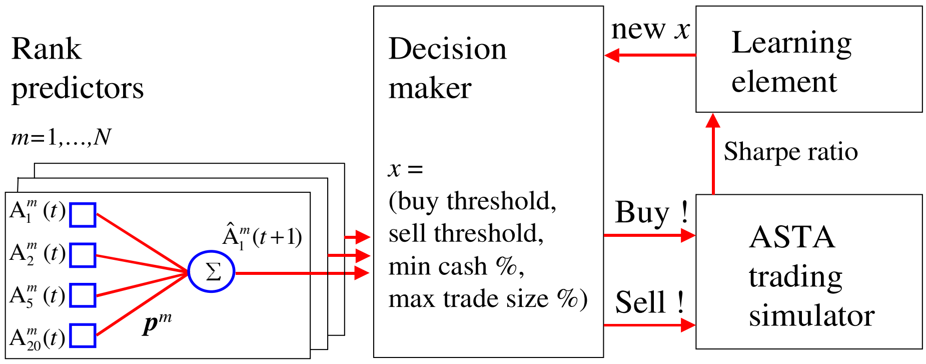

Fig. 3. The trading system is based on 1-day rank predictions for N stocks, and a decision maker parameterized by the x-vector. The learning element comprises a numerical optimizer, which finds the x that maximizes the Sharpe ratio for the historical data.

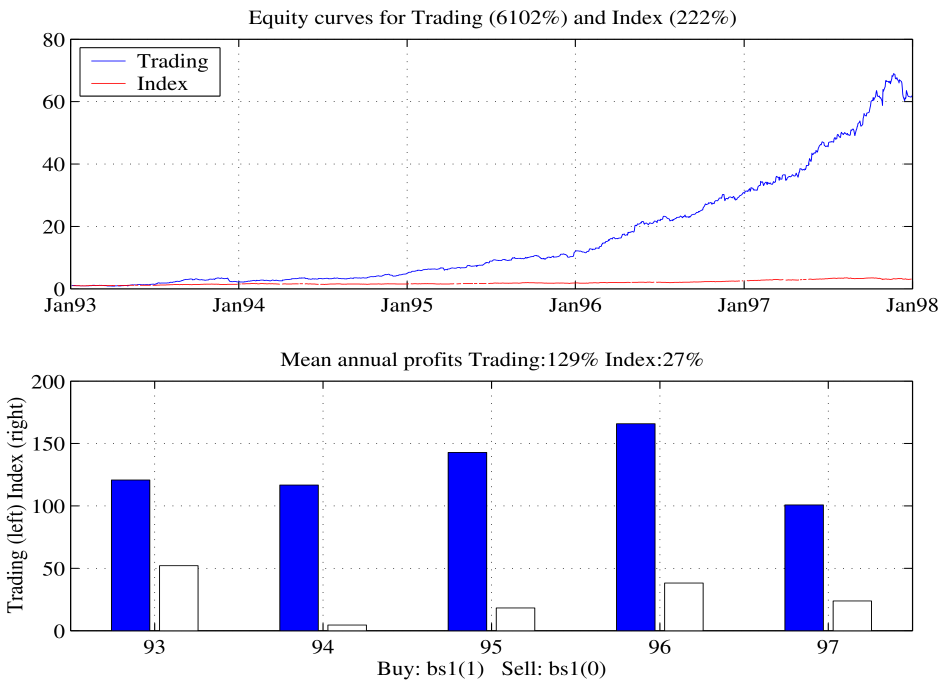

Fig. 4. Performance for the simulated trading with stock picking based on 1-day rank predictions as shown in Figure 3. The top diagram shows the equity curves while the lower diagram displays the annual profits. The trading outperforms the benchmark index consistently every year.