In the IEEE Conference on *Computational Intelligence in Financial Engineering*, New York, 1996.

Building Long/Short Portfolios Using Rule Induction

George H. John
Computer Science Department
Stanford University
Stanford, CA 94305
`gjohn@cs.stanford.edu`
`http://robotics.stanford.edu/~gjohn/`

Peter Miller
Lockheed Martin AI Center
3251 Hanover Street
Palo Alto, CA 94304
`pmiller@aic.lockheed.com`
`http://hitchhiker.space.lockheed.com/~recon/`

Abstract

We approach stock selection for long/short portfolios from the perspective of *knowledge discovery in databases* and *rule induction*: given a database of historical information on some universe of stocks, discover rules from the data that will allow one to predict which stocks are likely to have exceptionally high or low returns in the future. Long/short portfolios allow a fund manager to independently address value-added stock selection and factor exposure, and are a popular tool in financial engineering. For stock selection we employed the Recon[1] system, which is able to induce a set of rules to model the data it is given. We evaluate Recon's stock selection performance by using it to build equitized long/short portfolios over eighteen quarters of historical data from October 1988 to March 1993, repeatedly using the previous four quarters of data to build a model which is then used to rank stocks in the current quarter. When trading costs were taken into account, Recon's equitized long/short portfolio had a total return of 277%, significantly outperforming the benchmark (S&P500), which returned 92.5% over the same period[2]. We conclude that rule induction is a valuable tool for stock selection.

1 Introduction

In recent years, long/short investing has become a popular method in financial engineering. The technique allows fund managers to add value by identifying not only those stocks likely to outperform their benchmark, but also those stocks likely to underperform. By buying the strong performers and selling the underperformers short, a manager hopes to achieve positive returns while potentially decreasing risks (Jacobs & Levy 1993). Since many of the long positions might have offsetting short positions with similar factor exposures, the major source of return is uncorrelated with broad market moves. This decrease in risk can either be a serendipitous effect, or it can be engineered into the portfolio via constraints given to a portfolio optimizer.

Primarily because of the complicated investment objectives involved in a long/short portfolio, quantitative methods are used by many long/short managers. We present a quantitative modeling system, Recon, with roots in the research communities of Artificial Intelligence, Machine Learning, Neural Networks, and Statistics. The system has previously been shown to create long portfolios with excess return of 3.5% per quarter over four and a half years (John, Miller & Kerber 1996). Because of the relative unpopularity of short-selling and thus less efficient pricing, we hope to find even better results using Recon to build long/short portfolios.

Recon addresses the supervised learning problem: given a database, build a model which allows the prediction of one target field given the rest. In the stock selection problem, the field of interest is the return of a stock, and the database would include such other fields as current price-book ratio, earnings estimates, and measures of trends in the stock price. Recon is a rule induction algorithm, which means that the model it produces is a set of simple if-then rules. This allows an analyst to understand and perhaps modify the model before applying it to new data. Many other problems fit into the supervised learning framework: asset pricing, database marketing, bond rating, and credit scoring are just a few examples (Hutchinson, Lo & Poggio 1994, Anthes 1995, Utans & Moody 1991).

The rest of this paper discusses long/short strategies, the Recon system and how it processes and models data, and our experiments using Recon for building long/short portfolios. We then describe related work and present our conclusions.

[1] Recon is a trademark and service mark of Lockheed Martin Missiles and Space, a subsidiary of the Lockheed-Martin Corporation.

[2] As with human fund managers, past performance is no guarantee of future returns.

2 Long/Short Strategies

As of early 1995, estimated assets under long/short management were \$12–\$15 billion, up from \$3–5 billion two years earlier Bensman (1995). The popularity of these portfolio management strategies is due to many attractive practical and theoretical properties. Theoretically, Michaud (1993) shows that for any given level of risk, the return to a long/short portfolio is expected to be higher than a long-only portfolio. Hennessee (1994) shows that this phenomenon is even stronger when looking at "downside risk." Practically, because of low short-selling volume (only 3% of NYSE and NASDAQ volume as of late 1993), underperforming stocks are likely to be less efficiently priced, leaving more room for profit by astute short pickers (Price 1989, Dravo 1993).

The ability to short-sell gives a manager more freedom. For example, a long-only manager might strongly believe that a company, which constitutes 2% of her benchmark, will go bankrupt. The only way she can express this belief is to not hold the stock at all. In a long/short portfolio, she could bet against the company as strongly as she could normally bet for a company in a long portfolio (Jacobs & Levy 1995).

Actual long/short fund managers have experienced good results. Hedge funds using long/short equity positions had an average Sharpe ratio (Sharpe 1994) of 0.6 during the period 1990-1994, versus 0.3 for the S&P 500 (Van 1995). Management costs are typically high compared with long-only portfolios, due to the added complexity in portfolio maintenance.

2.1 Mechanics and Costs

The following small example should provide a concrete depiction of long/short mechanics for readers new to the idea. It also gives experienced readers a detailed model of our assumptions about trading costs and margin requirements.

Assume the current price of IBM stock is \$100/share, Apple is at \$50, and a six-month S&P 500 index futures contract price is \$625. We have \$10 million and expect IBM's price to rise and Apple's to fall.

We invest \$7M in IBM. Assume .5% one-way transaction costs, so our investment in IBM is actually 69,650 shares after costs, worth \$6,965,000. We sell short \$7M of Apple. Our broker borrows shares from someone else, sells them, and puts the proceeds (\$6.965M after transaction costs, or 139,300 shares' worth) into our account. The broker demands that we put up an additional 50% (\$3,482,500) as a margin to cover possible losses. We present our IBM stock, which the broker values at 50 cents on the dollar and accepts as sufficient initial margin.

We want our portfolio to behave as if we had also invested \$10M in the market. Assume that our IBM and Apple holdings together are uncorrelated with the market, so that our current effective investment in the market is \$0. We buy 32 S&P 500 index futures contracts (the value of an S&P index futures contract is 500× the quoted price) which approximately satisfies our goal. This costs us \$3200 (\$100 per contract), and our broker requires us to put up 10% of the value as initial margin. We put \$1M in a margin account with the broker.

We have now spent \$8M and are holding almost \$2M in cash. During the quarter, whenever Apple pays a dividend we must pay the dividend to the broker. Whenever IBM pays a dividend, we receive it as cash. When the S&P index rises or falls, the difference between the new price and the old prices (times 32 contracts times 500) is added to or subtracted from our margin account. Our Apple margin account also varies. During the quarter we get interest on our margin accounts equal to the Treasury bill rate minus .5%, and interest at the Tbill rate on our cash account. At times we may be required to deposit more money in our margin accounts if they fall below the maintenance margin.

At the end of the quarter, say IBM stock is worth \$105, Apple is worth \$40, and an S&P 500 index futures contract is \$650. We sell our IBM stock, getting \$7,276,684 (after another .5% transaction charge) for a return of 4.0% plus dividends received. We close our short position in Apple, the broker uses \$5,628,000 to buy back the shares it previously borrowed, for a return of 19.1%, minus dividends, plus the interest on the margin. We close our long position in S&P futures; nothing happens except that we receive all funds in our margin account, which now contains our original \$1M plus interest, plus \$.8M in profits due to the change in the index.

2.2 Usefulness of Quantitative Models

In multi-factor models such as the BARRA United States Equity Model or the multi-factor CAPM model described by Sharpe (1982), the return to a security is expressed as a function of the returns to various factors to which the security is exposed. For example, IBM stock is exposed to the market as a whole, the large company factor, the computer industry factor, and so forth. With a long/short technique, a fund manager can control the net exposure of the portfolio to factors by defining a set of constraints on net factor exposure. Without expressing constraints on factor exposures, a long/short portfolio may easily be more risky than a simple long portfolio. For example, going long in airline stocks (which are highly negatively correlated with oil prices) and short on oil stocks effectively places a double bet that oil prices will drop (Bensman 1995).

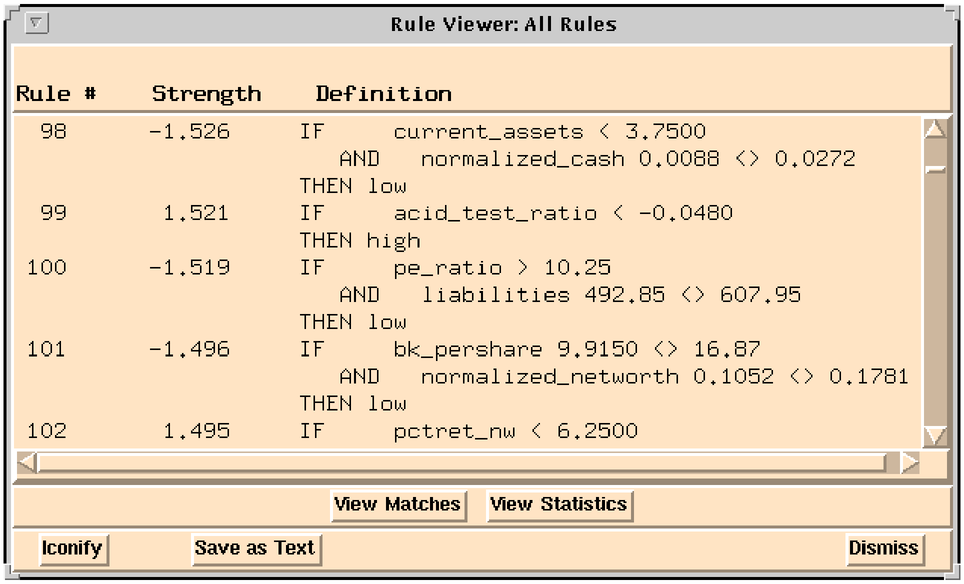

Figure 1: Example of rules learned by the Recon rule induction system.

Quantitative models are essential here—a manager needs to supply predicted returns on a large number of stocks so that she can actually find positive excess return portfolios meeting strict exposure constraints. The stricter the factor exposure limits, the fewer portfolios will meet the constraints, and the more stocks need to be included in the opportunity set in order to still find a good (high-return) portfolio satisfying the constraints. As the size of the selection universe grows, either the analysis staff must grow commensurately, or quantitative methods must be employed.

3 The Recon System

The Recon system (Simoudis, Livezey & Kerber 1994, Simoudis, Livezey & Kerber 1995) is an integrated system for the exploration and analysis of large databases (John et al. 1996). The exploratory tools allow a user to graphically define concepts and new features, and to visualize data. The principal analytical tool is a rule induction system developed at Lockheed-Martin.

The rule induction algorithm in Recon is a descendant of the systems described in Kerber (1991) and Kerber (1992). The algorithm searches through the space of rules to find those that are likely to be useful in classifying items in a future database. Some examples of rules that were found by Recon are shown in Figure 1. The first rule reads "If `current assets` is less than 3.75 and `normalized cash` is between 0.0088 and 0.0272 then predict `unexceptional` (low) return, with strength 1.526." (These rules were learned from ValueLine data. We are not permitted to show rules learned from the database used in this paper, which was provided to us by a mutual fund management company.) Note that Recon is not an expert system, even though it does learn rules. Expert systems are collections of rules painstakingly written by a human knowledge engineer, while Recon's rules are learned automatically from a database.

Each rule has an associated strength. Since rules may overlap, the class (`exceptional` or `unexceptional`) that Recon finally predicts for a stock is a combination of the predictions made by each matching rule. A rule's prediction is weighted by the amount of evidence supporting the rule, allowing Recon to not only make a prediction but also to assign some measure of certainty or strength to the prediction.

When used to learn a set of rules from a database, Recon's rule induction algorithm begins by first discretizing all numeric features. For example, the price-to-book ratio might be segmented into three intervals which could be interpreted as an indicator that a stock is an aggressive-growth, growth, or value stock. All features are discretized using the ChiMerge algorithm described in Kerber (1992). ChiMerge uses the training data, including the class label, to build intervals such that the distribution of class values in neighboring intervals is statistically dissimilar. The rule induction algorithm then begins exploring the space of all possible rules. It begins with all single-antecedent rules and then adds conditions using a combination of several heuristics to concentrate on interesting rules, ultimately yielding a set of rules with high predictive accuracy. Each rule is characterized by a set of statistics that describe the (training-sample estimates of the) conditional distribution of the class given whether or not the antecedent of the rule is true; thus, the search through the space of rules is essentially just a search through the space of the multivariate statistics of the data that are pertinent to the classification problem.

The Recon system allows an analyst to explore the discovered rules by viewing the training cases that led to a particular rule's creation, or viewing statistics about a rule. One can also determine which features are important by seeing which ones tend to participate in most of the rules. Similarly, features which are never tested did not provide information useful to the classification problem, according to Recon. Unrepresentative training data can yield rules that disagree with an analyst's knowledge of a problem; in such cases the analyst can edit the set of rules, perhaps removing the offending rules. For each test case, Recon shows the relative strength of its membership in each class, and can show the set of rules participating in the prediction.

Recon performs a search through a database to find statistically salient rules, and presents them to an analyst in a comprehensible format. If the resulting rules can result in superior performance in stock selection (or credit scoring or other problems), a strong case exists for the use of a tool like Recon because of the understandability of the rules it discovers.

4 Experiments

To evaluate Recon as a stock-selection tool for long/short strategies, we used its predictions to form portfolios and analyzed their performance using standard financial methods. Below we describe how a model is learned from training data, and how Recon's model was used to construct a portfolio over an out-of-sample period in our simulation. We then discuss the results, examining the risk and return of the created portfolios.

In the simulation, we repeatedly pretend that the beginning of a quarter in our database is the current date. We use Recon to build a return model for the previous four quarters, and apply that model to the current quarter. For each of the previous four quarters we have a table with about 1300 records (rows, one per stock) and 100 fields (columns, such as price to earnings ratio and relative strength). All of the fields were measured at the beginning of each quarter, except for the future three month return, which is measured at the end of each quarter.

Since Recon uses a binary classification algorithm, it needs a true/false value to predict. We added two fields to the data in each quarter—the `outperformer` field and the `underperformer` field.

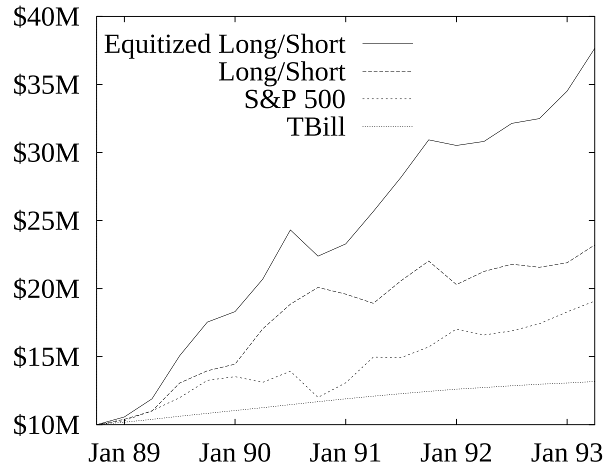

Figure 2: Out-of-sample cumulative returns for the period October 1988 – March 1993.

`Outperformer` was set to `true` in the top 20% of stocks in each quarter, and `underperformer` was set to `true` for the bottom 20%. We collect all four quarters together into a large training set of about 5200 records, and use Recon to build a model (which looks similar to Figure 1) to predict the `outperformer` field, and another model to predict the `underperformer` field.

We then apply these models to the data available "today" in the simulation. For each stock, Recon assigns a strength that the stock will be an outperformer and an underperformer. We use these strengths to produce two ranked lists on the selection universe: the long list is ranked by the strength of the `outperformer` field, and the short list is ranked by the `underperformer` field. Not having a portfolio optimizer available, and lacking some important fields in the historical data (such as volatility), we decided to simply take the top fifty stocks from each list as our long and short stocks to form two market value-weighted portfolios to be used as components of a larger long/short portfolio.

The Recon equitized long/short portfolio is formed by investing 70% of our current assets into both the long and short portfolios, with a 100% investment in S&P index futures as described in Section 2.1. The Recon long/short portfolio is the same but has no investment in futures. To estimate the performance of these portfolios, we use the costs and margin requirements from Section 2.1, and assume that at the end of each quarter we completely liquidate our positions, so that trading costs are assessed quarterly on our entire invested capital. (The actual turnover for the long and short portfolios was 80%.) Figure 2 shows the returns to both portfolios, along with the returns to the S&P 500 portfolio and Treasury bills.

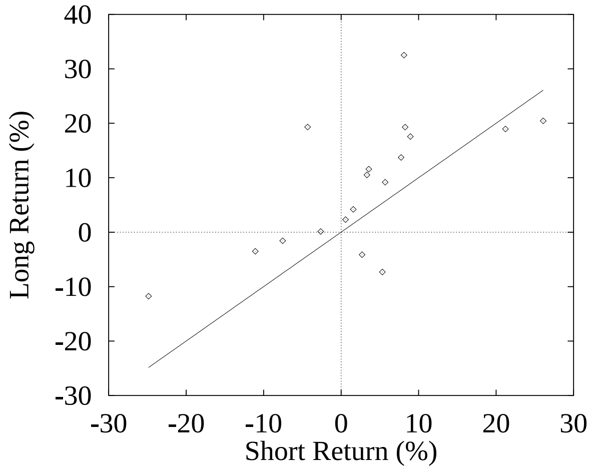

Figure 3: Scatterplot showing the quarterly Recon long portfolio returns versus the Recon short portfolio returns. Positive net return for points above the $y = x$ line.

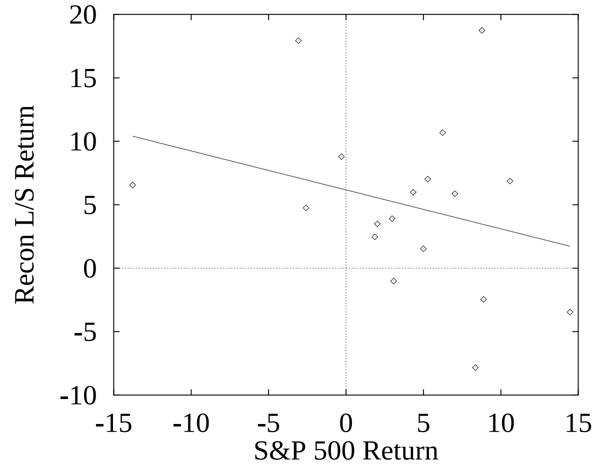

Figure 4: Scatterplot showing the quarterly Recon Long/Short portfolio returns versus the S&P 500 index returns, and the line with slope Beta and intercept Alpha.

Results of trading should always be considered relative to an appropriate benchmark. Jacobs & Levy (1993) suggest the S&P 500 as a benchmark for the equitized long/short portfolio and Treasury bills for the long/short portfolio. Indeed, they manage a long/short fund and are actually compensated based on their performance relative to Tbills (White 1991). However, the Recon long/short portfolio had much higher volatility than Tbills (6.71% vs. 0.49%), so we instead use combinations of S&P futures and Tbills to construct custom benchmarks with the same risk as the two Recon portfolios. The Recon equitized long/short portfolio had higher risk than the S&P 500 (7.87 % vs. 6.29%), but the S&P/Tbill combination portfolio with equivalent risk (which had 125% futures investment) generated only 5.2% quarterly return, as opposed to 7.91% for the long/short equitized portfolio. The customized benchmark for the Recon long/short portfolio actually had a 107% position in S&P futures and generated 4.66% average quarterly return, versus 4.99% for Recon's portfolio. The Sharpe ratio was .81 for the Recon equitized long/short portfolio, .51 for the Recon long/short portfolio, .37 for the S&P 500, and .46 for the custom benchmark for the equitized Recon portfolio.

Within the Recon long/short portfolios, the source of excess return is the difference between the returns to the long and short portfolios. Figure 3 shows a scatterplot of the long and short portfolio returns in each quarter. Since the net return is the long portfolio return minus the short portfolio return (since we sell these stocks short), we profit whenever points are above the $y = x$ line. Figure 4 shows the correlation between these net returns and the S&P 500 return. The correlation is negative with a slope, or Beta, of -.31, and an intercept, or Alpha, of .06. The R^2 value for the linear regression indicates that only about 8.3% of the variation in Recon's returns is explained by the S&P 500 returns.

Although we were not able to control exposure to other factors besides the market, we did investigate the net positions of the long and short portfolios. We looked at two factors: growth/value and size. Stocks were classified as growth or value stocks using price to book ratios as in Capual, Rowley & Sharpe (1993). The long/short portfolio had an average net positive 20% exposure to growth stocks, with 69% standard deviation. Stocks were classified as small, medium, or large by sorting all stocks by market capitalization and putting an equal number into each bin. The net exposure to size was -16.3% to small, -8.3% to medium, and +24.6% to large, with standard deviations of 26.5%, 27.6% and 40.4% respectively. The long and short portfolios did not exactly balance each other—the long/short portfolio actually doubled the long portfolio's bet on growth stocks, but at the same time it hedged the long portfolio's bet on small stocks.

Figure 5 shows the quarterly average return and standard deviation of a portfolio with 70% investment in the long and short portfolios, with varying amounts invested in S&P futures. With a 60% futures overlay, the risk is still roughly 6.7%, the same as with no overlay, but the quarterly return is 1.75% greater. Figure 6 shows the quarterly average return and standard deviation of a portfolio with 100% investment in S&P futures, with varying amounts invested in a long and short portfolio. Note again that by combining assets of low correlation we

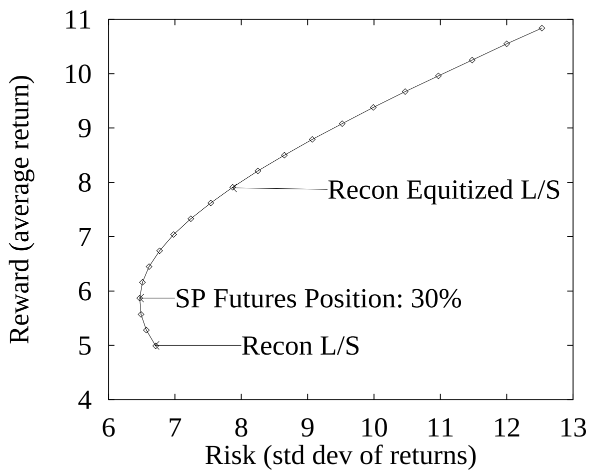

Figure 5: Risk-reward curve for increasing investment in S&P 500 index futures. The lowest point represents a 0% position, the top a long 200% position, and points are in 10% increments.

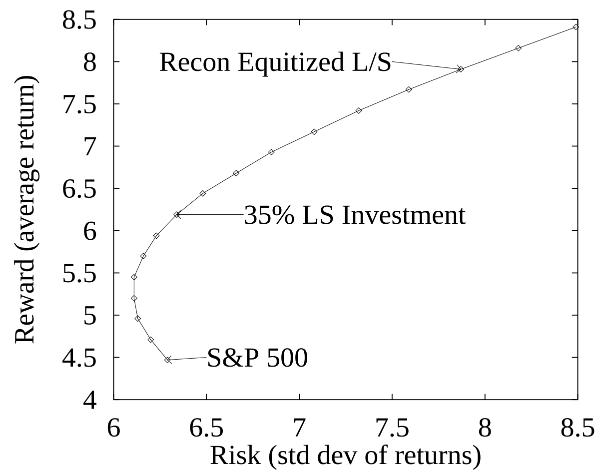

Figure 6: Risk-reward curve for increasing investment in a Long/Short portfolio. The bottom point has zero investment, the top has 80% investment, and points are in 5% increments.

can get extra returns with no penalty in risk: a 35% investment in long and short portfolios gives 1.72% higher return than the plain S&P portfolio but with the same standard deviation.

Caution is called for when interpreting results of simulations such as ours. Freeman (1992) gives a good consumer's guide to understanding simulated returns, delineating many possible errors that might lead to optimistically biased results. Markowitz & Xu (1994) and Black (1995) mention several ways in which out-of-sample data can be overused, leading to optimistic results. We did use the first fourteen quarters of our out of sample period several times to set a few experimental parameters. The last four quarters constitute a strict out-of-sample set, which was only used once. Because the Recon portfolios still strongly outperform their benchmarks during the strict out-of-sample year, we do not feel that the previous fourteen quarters of results are optimistic. Additionally, by using higher margins and costs than necessary we have attempted to avoid any unintentional optimism in our results. We have not discussed potential problems with margin calls, but since the worst single-quarter returns for either Recon portfolio was only -8% and since we kept a 20% cash buffer, we do not believe this would have been a problem.

5 Related Work

Related work in finance deals with other methods for adjusting the exposure of a portfolio. Long/short strategies are not the only way to get a near-zero beta portfolio—a manager can also short S&P 500 futures to hedge (Hull 1993, McGee 1995).

Levin (1996) used a neural network on monthly data to form equitized long/short portfolios, generating an excess return of 175.4% during the period August 1990 – February 1995. Levin reports positive results in real-time trading, with live trading expected soon. Apte & Hong (1995) used a similar rule-induction algorithm to form long portfolios on IBM retirement fund data, achieving an excess return of 160% over 5 years of monthly trading. Both studies assumed 1% round-trip transaction costs and used the S&P 500 as an appropriate benchmark. In previous work with Recon we observed an excess return of 144% during the period October 1988 – March 1993, also using 1% round-trip costs. The appropriate benchmark was custom-designed but very close to S&P 500 (John et al. 1996).

6 Conclusion

For many significant reasons, long/short equity management is growing in popularity. When subject to a number of risk controls, such strategies typically require quantitative models for estimating future excess returns.

Any modeling tool used for money management should produce a comprehensible model. With millions of real dollars at stake, the ability to understand and interact with the model is a necessity. Neural networks are very popular with quantitative managers, but their opacity makes them a less than ideal tool (Barr & Bhagat 1994). We believe that Recon's ability to discover patterns and present them as easily understood rules, explain why it included certain rules, interact with an analyst to modify the rules, and explain its predictions, makes it an excellent tool for financial modeling and stock selection.

Of course, to be a useful tool, the learned rules must be good predictors of future performance. Our experiments show that the long/short portfolio selected by Recon significantly outperformed the benchmark over four and a half years of trading.

Acknowledgments

The authors would like to thank Craig Scholl at Hewlett-Packard and Professors Peter DeMarzo and William Sharpe at Stanford for their help in understanding the mechanics and costs of long/short strategies. Ron Kohavi, Karl Pfleger, and Mehran Sahami provided comments on an earlier draft. The first author's work was partially supported by an NSF Graduate Research Fellowship.

References

Anthes, G. H. (1995), "Customer 'data mining' pays off", *ComputerWorld* **29**(20), May 15, pp. 1,28.

Apte, C. & Hong, S. (1995), Predicting equity returns from securities data with minimal rule generation, in U. Fayyad, G. Piatetsky-Shapiro & P. Smyth, eds, *Advances in Knowledge Discovery and Data Mining*, AAAI Press/MIT Press, chapter 22, pp. 537–556.

Barr, D. S. & Bhagat, S. A. (1994), "Computers as investment strategists", *Pension Management*, November, pp. 12–15.

Bensman, M. (1995), "Psst, we're market-neutral", *Institutional Investor* **29**(1), January, pp. 37–44.

Black, F. (1995), "Estimating expected return", *Financial Analysts Journal* **51**(1), January-February, pp. 168–171.

Capual, C., Rowley, I. & Sharpe, W. F. (1993), "International value and growth stock returns", *Financial Analysts Journal* **49**(1), January-February, pp. 27–36.

Dorfman, J. R. (1994), "Garzarelli hedges with mutual fund", *Wall Street Journal*, May 16, p. C1.

Dravo, E. (1993), "The long (and short) of short selling", *Financial World*, September 14, pp. 58–60.

Freeman, J. D. (1992), "Behind the smoke and mirrors: gauging the integrity of investment simulations", *Financial Analysts Journal*, November-December, pp. 26–31.

Hennessee, E. L. (1994), "Hedge funds can smooth out bumpy returns", *Pensions & Investments*, September 5, pp. 39–40.

Hull, J. C. (1993), *Options, Futures, and other Derivative Securities*, second edition, Prentice Hall, Englewood Cliffs, New Jersey.

Hutchinson, J. M., Lo, A. W. & Poggio, T. (1994), "A nonparametric approach to pricing and hedging derivative securities via learning networks", *Journal of Finance* **49**(3), July, pp. 851–889.

Jacobs, B. I. & Levy, K. N. (1993), "Long/short equity investing", *Journal of Portfolio Management* **20**(1), pp. 52–63.

Jacobs, B. I. & Levy, K. N. (1995), "Market-neutral strategy limits risk", *Pension Management*, July, pp. 39–43.

John, G. H., Miller, P. & Kerber, R. (1996), Stock selection using Recon, in Y. Abu-Mostafa, J. Moody, P. Refenes & A. Weigend, eds, *Neural Networks in Financial Engineering*, World Scientific.

Kerber, R. (1991), Learning classification rules from examples, in *Proceedings of the AAAI-91 Workshop on Knowledge Discovery in Databases*, AAAI.

Kerber, R. (1992), Chi Merge: Discretization of numeric attributes, in *AAAI-92: Proceedings of the Tenth National Conference on Artificial Intelligence*, AAAI Press/MIT Press, pp. 123–128.

Levin, A. U. (1996), Stock selection via nonlinear multifactor models, in *Advances in Neural Information Processing Systems*, Morgan Kaufmann Publishers, Denver, Colorado.

Markowitz, H. M. & Xu, G. L. (1994), "Data mining corrections", *Journal of Portfolio Management*, Fall, pp. 60–69.

McGee, S. (1995), "Fund managers get a jump with stock-index futures", *Wall Street Journal*, February 21, p. C1.

Michaud, R. O. (1993), "Are long-short equity strategies superior?", *Financial Analysts Journal*, November/December, pp. 44–49.

Price, K. (1989), "Long on safety: don't sell short selling short", *Barron's*, November 6, pp. 16,38.

Sharpe, W. F. (1982), "Factors in New York Stock Exchange security returns, 1931-1979", *The Journal of Portfolio Management*, Summer, pp. 5–19.

Sharpe, W. F. (1994), "The Sharpe ratio", *Journal of Portfolio Management* **21**(1), Fall, pp. 49–58.

Simoudis, E., Livezey, B. & Kerber, R. (1994), Integrating inductive and deductive reasoning for database mining, in *Proceedings of the AAAI Knowledge Discovery in Databases Workshop*, Seattle, WA.

Simoudis, E., Livezey, B. & Kerber, R. (1995), Using Recon for data cleaning, in U. M. Fayyad & R. Uthurusamy, eds, *First International Conference on Knowledge Discovery and Data Mining (KDD-95)*, pp. 282–287.

Utans, J. & Moody, J. (1991), Selecting neural network architectures via the prediction risk: Application to corporate bond rating prediction, in *The First International Conference on Artificial Intelligence Applications on Wall Street*, IEEE Computer Society Press.

Van, G. P. (1995), "Taming hedge funds", *Pension Management*, August, pp. 20–24.

White, J. A. (1991), "How Jacobs and Levy crunch stocks for buying – and selling", *The Wall Street Journal*, March 20, p. C1.