Title	Neural Network Approach to Predicting Stock Exchange Movements using External Factors
Author(s)	Madden, Michael G.; O'Connor, Niall
Publication Date	2005
Publication Information	"Neural Network Approach to Predicting Stock Exchange Movements using External Factors", Niall O'Connor and Michael G. Madden. Proceedings of AI-2005, 25th International Conference on Innovative Techniques and Applications of Artificial Intelligence, Cambridge, Dec 2005.
Item record	http://hdl.handle.net/10379/193

A Neural Network Approach to Predicting Stock Exchange Movements using External Factors

Niall O'Connor and Michael G. Madden
National University of Ireland, Galway
Galway, Ireland.
niallaoconnor@gmail.com, michael.madden@nuigalway.ie
www.it.nuigalway.ie/m_madden

Abstract

The aim of this study is to evaluate the effectiveness of using external indicators, such as commodity prices and currency exchange rates, in predicting movements in the Dow Jones Industrial Average index. The performance of each technique is evaluated using different domain specific metrics. A comprehensive evaluation procedure is described, involving the use of trading simulations to assess the practical value of predictive models, and comparison with simple benchmarks that respond to underlying market growth. In the experiments presented here, basing trading decisions on a neural network trained on a range of external indicators resulted in a return on investment of 23.5% per annum, during a period when the DJIA index grew by 13.03% per annum. A substantial dataset has been compiled and is available to other researchers interested in analysing financial time series.

1. Introduction

The Dow Jones Industrial Average (DJIA) index was launched in 1896 with 12 stocks, and is now the world's most often-quoted stock exchange index, based on a price-weighted average of 30 significant companies traded on the New York Stock exchange (NYSE) and Nasdaq. The index is used as a general indication of how the market reacts to different information. Financial institutions offer mutual funds based on this index, enabling investors to capitalise on market growth. Several researchers in the past have applied machine learning techniques such as neural networks in attempts to model predict movements in the DJIA and other stock exchange indices. A common approach involves the use of technical indicators which are derived from the DJIA time series itself, such as moving averages and relative strength indices. This relies on past events in the time series repeating themselves to produce reliable predictions. Although machine learning studies using technical indicators, such as those of Yao & Tan [1] and Rodrígues et al. [4], have claimed successful returns, the key limitation of these approaches is that such models do not capture the *cause* of the movements in the market.

The earnings of companies are affected by both internal influences such as product development and external influences such as the cost of energy and the currency exchange rates with foreign markets. The external factors tend to affect the majority of companies in the same way; for example, a rise in energy costs results

in a decline in profitability and thus an adverse effect on the value and share price of companies other than energy suppliers. Thus, we hypothesise that such factors will have an observable effect overall on the Dow Jones index, and consideration of them should improve one's ability to predict movements in the index.

Accordingly, this work seeks to identify some prominent external indicators of the Dow Jones index, and use neural networks to model the effect these indicators have on the index. If an effective model is created it will be possible to predict, to some extent, future movements in the index based on current and past data, thus capitalising on the prior knowledge.

Clearly, there is an enormous range of other factors that would not be accounted for by this approach, but an effective model should perform better than random and better than the baseline growth in the index over the testing period. For that reason, this work places emphasis on the evaluation of model performance. Model predictions are used to drive a trading strategy so that the profitability of models may be assessed. Models are also evaluated relative to simple benchmarks.

The contributions of this work are:

- As described in Section 3, a substantial dataset has been compiled with daily values of the DJIA, derived technical indicators, and external indicators. It is available by email from the second author for use by the research community.

- A profit-based evaluation procedure for financial prediction systems is proposed, based on simulations with a simple trading strategy. As discussed in Section 4, this is more meaningful than evaluating systems based on the error between predicted and actual values, as is sometimes done.

- A neural network approach is shown to be successful in predicting movements of the DJIA in Section 5, provided that external factors are considered. These results may be used as a baseline against which to compare other prediction techniques in the future.

2.   Related Research

Many of the papers published in this domain are based entirely on analysing the stock market index time series itself, along with derived quantities, without reference to external indicators. For example, Rodrígues et al. [4] develop a rather simple model based entirely on the time series of the Madrid Stock Market General Index. The previous nine days of the index were used as inputs and a buy/sell signal was the output. The research concluded that the neural network trading model was superior to a strategy of simply buying and holding stocks for a bear market (period of decline) and stable market (period of neither growth nor decline), but for a bull market (period of growth) the model performed poorly when compared with the buy-hold strategy for that period. It is difficult to determine if the model created was effective; it stands to reason that a buy-hold strategy cannot be profitable for a period of market decline or stagnation, which implies that a different benchmark against which to compare against might have been more appropriate.

Likewise, in the work of Yao & Tan [1], the effectiveness of a time series model based on the FOREX (foreign currency exchange market) with no external input parameters was evaluated. The paper discusses the concept of market efficiency, which is the time taken for asset prices to react to new information in the market. It is claimed by the Efficient Market Hypothesis that in an efficient market, prices react essentially instantaneously, so that traders cannot capitalise on new information and asset prices reflect all information available to the market. The neural network model created by Yao & Tan performed well for most foreign exchange markets except the Japanese Yen/US Dollar exchange. Better results were seen when moving average technical indicators were incorporated into the model, except again in the Yen/Dollar market. Yao & Tan suggest that technical analysis is not suitable for this market as it is highly efficient and the use of technical indicators would be widely adopted by traders in this market.

Other related research includes neural network approaches to forecasting trends in the Kuala Lumpur Stock Exchange [13], and the Taiwan stock index [9], an approach using genetic programming [2], and a combined genetic algorithm and neural network system for trading of individual stocks [10, 11]. In each of these, forecasting is based solely on the past movements of the time series of interest.

This paper considers the Dow Jones index, which is an average of significant companies in the New York Stock Exchange, which in turn is the largest stock exchange in the world handling volumes of over 1 million trades per day, so it should be highly efficient. The efficiency of the Dow Jones makes it unlikely to be a profitable candidate for technical analysis, resulting in a need for our approach considering external factors.

Hellström and Holmström [12] provide a good introductory tutorial on stock market prediction using both technical and external fundamental indicators. In predicting movements of the USD/GBP exchange rate, Anastasakis and Mort [14] use other exchange rates as fundamental indicators of the USD/GBP, and show how the addition of external information to the model results in a marked improvement in the root mean squared error (RMSE). However, as will be noted in Section 4, low RMSE is not necessarily correlated with correct predictions of the direction of market movement. Lendasse et al. [15] describe an approach to forecasting movements in the Belgian Bel-20 stock market index, with inputs including external factors such as security prices, exchange rates and interest rates. Using a Radial Basis Function neural network, they achieve a directional success of 57.2% on the test data.

Using fundamental and technical analyses is not the only approach to this problem. An interesting study was conducted by Lavrenko et al. [3], predicting trends in stock prices of companies based on news articles relating to these companies. Starting with a time series of the company's stock value, trends are extracted using piecewise linear regression, to re-describe the time series as a sequence of trends. Each trend is assigned a label according to its slope. The time-stamped news articles are aligned with the trends, with a news article being considered to be associated with a trend if its time-stamp occurs within a certain timeframe before the trend occurs. It is possible that a single news article can affect more than one trend. A language model is built relating the typical language used in a news with the trend it is associated with. For instance, a news story corresponding to an upward trend might have words such as "merger" or "higher earnings" contained

within it. A Bayesian classification model is created to classify future trends from new news articles. At first glance the results from this approach are quite impressive: over a period of 40 days trading on Yahoo.com stock, a net profit of $19,000 was generated from a principal of $10,000. However, 570 trades were made over the 40 days, which by any standard is a large number of trades; this amounts to an average of $50 profit per transaction or a 0.5% return on investment per transaction. The market simulation did not take into account transaction cost, which for such a large number of trades may have had a significant negative impact on profits. While the Yahoo.com stock represents one significant positive result, when tested over other time periods the mean return was -$9,300 (loss).

3. Data Set Compilation

For the purposes of this work, a data set has been compiled containing daily opening and closing values of the DJIA index, and corresponding values for a range of external indicators. (Note that one day's closing value of the index can be slightly different from the next day's opening value, due to the recent introduction of after hours trading between institutions' private exchanges.)

In choosing external indicators, an important consideration is of course whether an indicator is likely to have a significant influence on the movement of the index, so that indicators are selected that tend to have an impact on the earnings of the companies in the DJIA. As will be described in Section 5, the relevance of the chosen external indicators was determined experimentally by adding them as inputs to the neural network models and assessing whether they improved performance.

Two other criteria had to be satisfied: high frequency of observations and high availability of historic data. These criteria preclude the use of some external factors even though they may be significant. For instance, the Federal Reserve interest rate is announced quarterly, while the network is trained on daily data, so there is difficulty in representing such an occurrence. One cannot simply interpolating the value between announcements, as in a real-time system one will not know the future announcement. Conversely, a step-wise representation, where the value is constant at all other times except when a change is announced, would be problematic as announcements tend to cause short-term changes in the index. Furthermore, the rate change itself does not hugely affect the earnings of the companies, but rather than how traders view their investments. This factor was therefore not considered.

Even though different companies are dependant on different resources and markets, the earnings of all will be affected to varying degrees by external factors such as the value of oil or foreign currency exchange rates.

To represent oil prices, the daily spot values of WTI Cushing Crude Oil were included in the data set because it is a common oil type internationally traded and a large volume of historic data is available for it. The choice of currency exchange rates was made by selecting the largest US trading partners with the largest volume of historic data available: US Dollar/Canadian Dollar, US Dollar/Japanese Yen and US Dollar/Pound Sterling.

The data set was formed from figures taken from three sources:

1. Yahoo.com Finance Section [5] for the daily spot values of DJIA

2. US Energy Information Administration [6] for data on the daily price of WTI Cushing Crude Oil

3. OANDA.com[7] for currency exchange rates

Since the data was taken from multiple sources, the representation of non-trading days differed. The data from Yahoo.com removed weekends and public holidays from their data store, while data from EIA did not and simply used the previous trading day's closing value for the subsequent non-trading days' entries. It was necessary to use a uniform representation across all data sources, therefore the other the Yahoo.com data streams was used as the standard form and all other sources were adjusted to conform to this standard. In some cases public holidays were on different days, as in the case of foreign currency markets being different to US public holidays. This gave rise to some "missing" values in the data, which were substituted with the previous trading days' values.

Since global currency markets close at different times there is the potential for "future" data being supplied unintentionally to the model. To guard against this, the only current-day element of the input vector is the opening value of the Dow Jones index; all other values are taken from the previous five days. This buffers models from the variations in the closing times of global markets.

The data set also includes technical indicators derived from the DJIA spot values, specifically the daily gradient of the DJIA, calculated as (Closing–Opening)/Opening, and 10-day and 30-day moving averages of opening values.

The working data set begins on 2 Jan 1986 and ends on 4 Feb 2005. It comprises 4818 data points and is available by email from the second author in normalised and raw representations.

4. Model Construction and Evaluation

4.1. Evaluation Methodology

To determine the effectiveness of training and thus determine if the resulting model is effective for the desired application, four simple benchmark functions were evaluated over the test set. They estimate the current day's closing value as:

1. Average of previous 5 days' opening values

2. Average of previous 10 days' opening values

3. Average of previous 30 days' opening values

4. One day lag i.e. today's closing value is the same as yesterday's.

Each benchmark function was used to predict the current day's closing value and its effectiveness was measured by the error between the functions output and the actual closing value. The performance was computed in three ways:

1. Root Mean Squared Error (RMSE)

2. Error in Dow Jones Points (i.e. RMSE re-scaled to DJIA units)

3. Directional Success (i.e. how often a rise/fall was correctly predicted)

The following table lists the performance of each of the benchmarks described above when applied to the 500 days of test data (18 Dec 2002 to 13 Dec 2004):

	RMSE	DJ Points	Dir. Success
Benchmark 1 – **5 day moving Ave**	0.08143	48.04	51.9%
Benchmark 2 – **10 day moving Ave**	0.00913	53.88	50.3%
Benchmark 3 – **30 day moving Ave**	0.01620	95.59	49.8%
Benchmark 4 – **One day Lag**	0.00443	26.47	40.2%

Table 1: Performance of Benchmark Approaches to DJIA Prediction

RMSE is often optimised in neural network applications, although from an end-user's point of view it can be quite abstract. By using the Error in DJ Points, the progress of training can be viewed in the context of the application. However, it is hazardous to infer model application performance from either of these error measures. Directional success is an important metric as buy/sell decisions will be based on predictions that the index will rise/fall. Examining Benchmark 4, it is clearly has lower error than the other benchmarks yet it has poor directional success, indicating a lack of ability to generalise. There is an intuitive rationale for this: we would expect that a day's closing value would be close to the previous day's closing value, so that Benchmark 4 should have low RMSE, but the change could be a rise or fall with almost equal probability, so the previous day's closing value is not a good basis for trading decisions on the current day.

4.2. Trading Strategy

While the accuracy of a model may be measured using RMSE, Error in DJ Points or Directional Success, what is ultimately needed is a measure of the effectiveness of the model in relation to its use in driving decisions to buy/sell shares. A fourth application-specific measure of model performance is therefore introduced: Return on Investment (ROI). This is computed by basing trading decisions on the output of the model. A simple trading strategy is proposed here: *No Threshold All In/Out*. If the market is predicted to rise by any amount (no threshold) this signals a *buy*, while if the market is predicted to fall this signals a *sell*.

We assume the existence of an idealised Index Tracking Fund that exactly mirrors the movements of the DJIA. (As discussed in the Conclusions, real-world tracking funds are not so precise.) We further assume that fractional amounts of the idealised fund may be traded, rather than just whole units. We start with an initial capital amount C_0 = \$10,000. Then, the first day that a *buy* signal is received from the model, an investment is made in the idealised fund using the full capital amount (all in). *Buy* signals on subsequent days are ignored (they are treated as *hold* instructions) until a *sell* signal is received, when all investment units currently held are sold, yielding a capital amount C_1. Because of the assumptions, this is computed simply as $C_1 = C_0 D_S/D_B$ where D_B is the closing value of the DJIA index on the day of the *buy* and D_S is the corresponding value on the day of the *sell*.

Subsequent days' *sell* signals are again treated as holds, until another *buy* is received, and the process is repeated. The overall ROI over the investment period is the percentage gain in investment capital. While the test period spans 500 working days, for clarity ROI is expressed as an average per annum figure.

We may also account for transaction charges. While such charges vary between brokerage institutions, we assume a flat-rate charge of $8 per trade, which is typical of some reputable online trading services[1]. We do not deduct charges directly from the investment capital but assume they are accounted for separately. However, transaction charges are deducted at the end when computing ROI. Results for ROI are reported with and without transaction charges because of the potential variability of transaction charges, and because if a larger starting capital sum were used, flat-fee charges would be proportionately less significant.

The profits from the benchmark approaches using this strategy are shown below.

	ROI Per Annum	ROI Per Annum with Tr. Charges
Benchmark 1	6.90%	1.78%
Benchmark 2	3.49%	1.53%
Benchmark 3	5.9%	2.78%
Benchmark 4	-14.20%	-24.10%

Table 2: Performance of Benchmark Approaches in Terms of Return on Investment

These figures must be taken in the context of the market environment during the test period. The Dow Jones for this period saw growth with a daily market direction rise 52.8% of the time. Furthermore, if one was to invest on the first day of the test set and hold until the last day, one's investment would have grown 13.03% per annum. Using the trading strategy described, an oracle with perfect knowledge of the future would attain a maximum ROI of 234.81% per annum over this period, or 223.86% per annum when accounting for transaction costs.

Comparing the metrics of the benchmarks as presented in Tables 1 and 2, it clear that several metrics are needed to accurately determine the effectiveness of a model during training and ultimately for the desired application.

4.3. Neural Network Architecture & Parameters

As will be described in Section 5, a range of experiments have been conducted. In all cases, the aim was to predict the current day's closing DJIA index value, given that day's opening value and other inputs such as some previous days' opening values of the DJIA and moving averages over several previous days. Multi-layer feed-forward neural networks were used in all experiments, trained using the back-propagation with momentum algorithm. The training parameters lay within the following ranges:

1. Learning Rate, η (0.001 – 0.2)

---

[1] http://personal.fidelity.com/accounts/services/content/brokerage_commission_index.shtml

2. Momentum, μ (0.003 – 0.3)

3. Flat Spot elimination, c (0.1)

In all cases, a range of different parameter settings and configurations of hidden nodes were evaluated, and the most successful ones are documented in Section 5.

For all experiments, networks were trained on 4000 days' data (11 Feb 1987 – 17 Dec 2002) and tested on 500 days' data (18 Dec 2002 – 13 Dec 2004). Remaining data points were unused.

5. Experiments & Results

A range of experiments were carried out, in each case using feed-forward neural networks to predict the current day's DJIA index closing value. In the first experiment, predictions were based on the current and previous 5 days' opening values, and in successive experiments additional inputs were added.

As stated earlier, all networks were trained on 4000 days' data, from 4000 days' data from 11 Feb 1987 to 17 Dec 2002 and tested on 500 days' data from 18 Dec 2002 to 13 Dec 2004. For each experiment, a range of different training parameters and numbers of hidden nodes were tried, and the best results are listed here.

For each experiment, error in predicting training and test set outputs are reported, in terms of both RMSE and DJ Points, along with the directional success and annual return on investment on the test set, with and without transaction charges.

5.1. Details of Experiments

Experiment 1 – Simple Time Series Experiment

| Input Data: | Current day's Dow Jones Opening Value |
| | Previous 5 days' Dow Jones Opening Values |

Network Architecture: 6-10-5-1 (115 Weights)

Training Parameters: $\eta = .01 \rightarrow .001$, $\eta = .03 \rightarrow .003$, c=0.1, 4000 epochs

Prediction Error:	RMSE	DJ Points
Training Set	0.04144	224.51
Test Set	0.0150	88.3

Directional Success on Test Set	53.30%
Annual ROI on Test Set	9.96%
Annual ROI inc. Transaction Charges	8.03%

Experiment 2 – Technical Indicators

| Input Data: | Current day's Dow Jones Opening Value |
| | Previous 5 days' Dow Jones Opening Values |

10-day Moving Average of Opening Values
30-day Moving Average of Opening Values

Network Architecture: 8-10-5-1 (135 weights)

Training Parameters: $\eta = .01 \rightarrow .001$, $\mu = .03 \rightarrow .003$, c=0.1, 4000 epochs

Prediction Error:	RMSE	DJ Points
Training Set	0.0456	269.43
Test Set	0.0146	86.26

Directional Success on Test Set	53.74%
Annual ROI on Test Set	8.56%
Annual ROI inc. Transaction Charges	7.65%

Experiment 3- Transformation of Time Series

Input Data: Current day's Dow Jones Opening Value
Previous 5 days' Dow Jones Opening Values
Previous 5 days' Daily Gradients of Dow Jones

Network Architecture: 11-9-7-1 (169 weights)

Training Parameters: $\eta = .01 \rightarrow .001$, $\mu = .03 \rightarrow .003$, c=0.1, 4000 epochs

Prediction Error:	RMSE	DJ Points
Training Set	0.0395	233.57
Test Set	0.0147	86.21

Directional Success on Test Set	52.70%
Annual ROI on Test Set	12.08%
Annual ROI inc. Transaction Charges	9.25%

Experiment 4 – Addition of Crude Oil Data

Input Data: Current day's Dow Jones Opening Value
Previous 5 days' Dow Jones Opening Values
Previous 5 days' WTI Cushing Crude Oil Price
(Price per Barrel)

Network Architecture: 11-9-7-1 (169 weights)

Training Parameters: $\eta = .01 \rightarrow .001$, $\mu = .03 \rightarrow .003$, c=0.1, 4000 epochs

Prediction Error:	RMSE	DJ Points
Training Set	0.04086	241.14

Test Set	0.0150	88.486

Directional Success on Test Set	53.31%
Annual ROI on Test Set	18.44%
Annual ROI inc. Transaction Charges	12.53%

Experiment 5 - WTI Cushing Crude Oil and Currency Data

Input Data:
Current day's Dow Jones Opening Value
Previous 5 days' Dow Jones Opening Values
Previous 5 days' WTI Cushing Crude Oil Price
Previous 5 days of the USD/YEN exchange rate
Previous 5 days of the USD/GBP exchange rate
Previous 5 days of the USD/CAN exchange rate

Network Architecture: 26-39-20-1 (1814 Weights)

Training Parameters: $\eta = .01 \rightarrow .001$, $\mu = .03 \rightarrow .003$, c=0.1, 4000 epochs

Prediction Error:	RMSE	DJ Points
Training Set	0.03416	211.0
Test Set	0.017087	88.8

Directional Success on Test Set	54.3%
Annual ROI on Test Set	20.52%
Annual ROI inc. Transaction Charges	18.28%

Experiment 6 – Currency Data, Crude Oil and Gradient of Dow Jones

Input Data:
Current day's Dow Jones Opening Value
Previous 5 days' Dow Jones Opening Values
Previous 5 days' Daily Dow Jones Gradients
Previous 5 days' WTI Cushing Crude Oil Price
Previous 5 days of the USD/YEN exchange rate
Previous 5 days of the USD/GBP exchange rate
Previous 5 days of the USD/CAN exchange rate

Network Architecture: 31-37-20-1 (weights 1907)

Training Parameters: $\eta = .01 \rightarrow .001$, $\mu = .03 \rightarrow .003$, c=0.1, 6000 epochs

Prediction Error:	RMSE	DJ Points
Training Set	0.03656	215.7
Test Set	0.0145	85.7

Directional Success on Test Set	55.1%
Annual ROI on Test Set	23.42%
Annual ROI inc. Transaction Charges	21.10%

5.2. Discussion

The purpose of Experiment 1 is to determine the performance of a neural network trained solely on the Dow Jones time series. While this shows profitability of 9.93% per annum, it is not greater the market growth (13.03% per annum over the test period, as stated in Section 4.3). This is to be expected, because of the efficient nature of the Dow Jones, and is comparable with the results of Yao & Tao [1], where similar experiments were performed on the Japanese Yen foreign exchange markets.

In Experiment 2, two derived indicators (10-day and 30-day moving averages of the DJ spot values) are added to the input data. In this case the profitability of the model falls significantly even though the directional success rises slightly. This highlights the importance of the profitability metric: while we would expect the profitability to rise with increased generalization ability as indicated by the directional success, this model appears to have got its predictions right on less profitable days than other models.

In Experiment 3, an alternative derived indicator, the gradient of the index in the past several days, is used. The results indicate that this is useful; yet the profit is not greater than market growth and the directional success is poorer than the previous two experiments.

In Experiment 4, crude oil prices are added to the input vector and it is clear that this external variable has had a positive effect on the generalization capabilities of the model. The profitability of the model increases relative to the earlier experiments.

In Experiment 5, the effect of the external variables can be seen clearly. The crude oil data and the currency exchange rates are all used. Greater generalization is seen with a return on investment of 20.52% per annum, which is significantly better than market growth.

In Experiment 6, the gradient of the Dow Jones is added to the input data, which has an interesting effect. In Experiment 3, when this was combined with the Dow Jones time series data only, significantly poorer results were seen. In contrast with this, in Experiment 6 it is added to the external variables of Experiment 5 with marked improvements in generalisation ability and profitability.

It is clear that by adding external indicators to the input vector, the overall performance in terms of profitability and directional success of the model has improved significantly. On the other hand, the accuracy in terms of Dow Jones Points of the best neural network model is poor, when compared to the benchmarks. The least accurate benchmark is Benchmark 3, with an average error of 95 Dow Jones points, while our model only achieves an accuracy of 99.7 Dow Jones points. While the generalization capabilities of the model allows profitable

trades since the signal to buy and sell is still valid, the poor accuracy makes it difficult to refine the trading strategy to maximize profits. As well as pointing to the weakness of the trading strategy used, this also indicates a disparity between RMSE and profitability, which in turn indicates that a neural network architecture that directly optimised profitability (or a strongly related quantity) might be better for this application.

5.3. Comparison with Profitability of Benchmarks

Of the simple benchmarks that were presented in Section 4, the one that performed the best was Benchmark 3, with an annual return on investment of 5.92%. As was noted in Section 4, this is not as good as simply buying and holding for the duration of the test period.

Figure 1 is a graph of profits accumulated over the test period when using Benchmark 3, Buy/Hold and the network of Experiment 6. As it shows, using a neural network trained with external variables as well as variables derived from the DJIA index itself is superior to the other approaches considered in this paper.
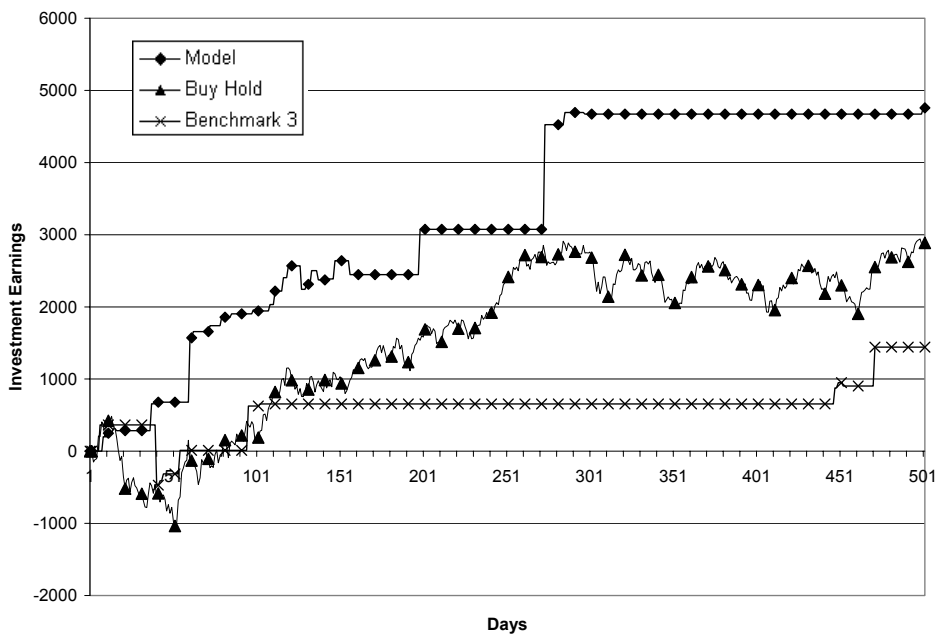
Figure 1: Comparison of Growth in Investment Earnings when Using Experiment 6 Model, Buy/Hold Strategy and Benchmark 3 Strategy

5. Conclusions

Even though, as observed in the Introduction, there is an enormous range of factors that are not accounted for by this approach, the analyses presented here demonstrate the benefit of considering some external factors when predicting stock exchange movements. While it is uncertain whether the addition of further external

factors will result in further gains in profitability, it is clear from the analysis presented here that their use is warranted. In Experiment 6, they led to a return on investment of 23.4% per annum (excluding transaction charges) during a period when the DJIA index grew by 13.03% per annum. In Experiments 4 and 5, there was less benefit arising from including fewer external factors.

Conversely, Experiments 1 to 3 have demonstrated little benefit in the commonly-used strategy of predicting future movements based solely on past trends. Also, the analysis of the performance of neural networks and benchmark strategies show a disconnect between being able to predict the closing value of the index with low error and being able to translate this into a profitable trading strategy. This implies that analysts should avoid reporting results in terms of RMSE but should compute a return on investment using a method like the one described in this paper, if results are to be convincing. It also implies that a machine learning approaches to this problem would be more successful if they sought to optimise profitability, or a strongly related quantity, directly.

As this work demonstrates, the use of domain-specific metrics and domain knowledge must be accommodated at many stages in machine learning application development, from data collection, to training, to performance evaluation and finally in the presentation of results. Greater attention to this can produce performance gains and results that are more relevant to domain experts.

There are many possible extensions to this work. Most obviously, there are many other machine learning techniques that could be substituted for the neural networks used here; regression trees, support vector machines and nearest neighbour algorithms are among the possibilities. In addition, there is likely potential for improving the profitability of the system by using more sophisticated trading mechanisms. One simple refinement would be to hold rather than buying/selling if the predicted gain in the index would not cover the transaction costs. Another refinement would be to invest an amount proportionate to the expected benefit, rather than the all in/out approach used here.

It should be noted that, in the experiments presented here, the neural network was not retrained during the testing period, although it is likely that it the performance of the system could be improved if it was retrained periodically (weekly or even daily) so as to include the most recent data.

Another practical consideration that could be addressed in the future is the assumption in this work of the existence of an idealised Index Tracking Fund. There are several such trackers in existence, such as the iShare [7] Dow Jones U.S. Industrial Sector Index Fund. We have conducted some initial experiments using the predictions output by Experiment 6 model to trade in iShares. However, it was found that this fund does not track the Dow Jones perfectly and a significant time lag was observed. This affected the timing of the model, as the end-of-day closing price predicted by the model was not realised in the iShares fund until the following day. The solution is to train the system on iShares data. However, at present there is not sufficient historical data available on the fund to substitute it for the DJIA data used. As more fund data becomes available, this could be explored.

Finally, there are other forms of analysis that could be performed with the dataset that has been compiled as part of this work. It includes daily opening and closing values of the DJIA, some derived technical indicators, and external indicators including daily oil prices and currency exchange rates, all covering a period from 1986 to 2005. This dataset is available by email from the second author and may be of use to other researchers engaged in other types of time series data mining.

References

1. Yao, J. & Tan, C.J.: "A Case Study on Using Neural Networks to Perform Technical Forecasting of FOREX", Neurocomputing, 2000.
2. Li, J. & Tsang, E.P.J.: "Improving Technical Analysis Predictions: An Application of Genetic Programming", Florida Artificial Intelligence Research Symposium, USA, 1999.
3. Lavrenko, V., Schmill, M., Lawrie, D. & Ogilvie, P.: "Mining of Concurrent Text and Time Series", Proc. 6th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Workshop on Text Mining, 2000.
4. Fernández-Rodríguez, F., González-Martel, C., Sosviall-Rivero, S.: "On the Profitability of Technical Trading Rules based on Artificial Neural Networks: Evidence from the Madrid Stock Market", Economics Letters, 2000.
5. Historic Dow Jones Data, Yahoo Finance, http://finance.yahoo.com.
6. WTI Cushing Crude Oil data, US Dept. of Energy Information Administration, http://www.eia.doe.gov.
7. Foreign Currency Data, Oanda, The Currency Site, http://www.oanda.com/.
8. iShare Dow Jones ETF: http://www.ishare.com.
9. Chen, A.S., Leung, M.T. and Daouk, H.: "Application of Neural Networks to an Emerging Financial Market: Forecasting and Trading the Taiwan Stock Index", Computers & Operations Research, Vol. 30, No. 6, 2003.
10. Thawornwong, S., D. Enke, and C. Dagli, "Genetic Algorithms and Neural Networks for Stock Trading Prediction and Technical Signal Optimization", 33rd Annual Meeting of the Decision Sciences Institute in San Diego, 2002.
11. Thawornwong, S., D. Enke, and C.H. Dagli, "Using Neural Networks and Technical Analysis Indicators for Predicting Stock Trends", Intelligent Engineering Systems through Artificial Neural Networks, Vol. 11, 2001.
12. Hellström, T. & Holmström, K.: "Predicting the Stock Market", Technical Report IMa-TOM-1997-07, Mälardalen University, Sweden, 1997.
13. Yao, J., Tan, C.J. & Poh, H.L.: "Neural Networks for Technical Analysis: A Study on KLCI". International Journal of Theoretical and Applied Finance, Vol.2, No.2, 1999.
14. Anastasakis, L. and Mort, N.: "Neural Network-Based Prediction of the USD/GBP Exchange Rate: The Utilisation of Data Compression Techniques for Input Dimension Reduction", Technical Report, University of Sheffield, 2000.
15. Lendasse, A., de Bodt, E., Wertz, V. and Verleysen, M.: "Non-Linear Financial Time Series Forecasting – Application to the Bel 20 Stock Market Index", European Journal of Economic and Social Systems, Vol. 14, No. 1, 2000.