Investigating the size and value effect in determining performance of Australian listed companies: a neural network approach

Justin Luu **Paul J. Kennedy**

Faculty of IT,
University of Technology, Sydney,
PO Box 123, Broadway, NSW 2007,
AUSTRALIA,
Email: paulk@it.uts.edu.au

Abstract

This paper explores the size and value effect in influencing performance of individual companies using backpropagation neural networks. According to existing theory, companies with small market capitalization and high book to market ratios have a tendency to perform better in the future. Data from over 300 Australian Stock Exchange listed companies between 2000–2004 is examined and a neural network is trained to predict company performance based on market capitalization, book to market ratio, beta and standard deviation. Evidence for the value effect was found over longer time periods but there was less for the size effect. Poor company performance was also observed to be correlated with high risk.
Keywords: multilayer perceptron, size and value effect, company performance prediction.

1 Introduction

The stock exchange is an exceedingly fluid, dynamic and engaging entity. It facilitates thousands of transactions which occur simultaneously from traders striving to outbid and outsell each other. From the moment it opens there is unceasing activity until the second it closes. Decisions to buy, sell or hedge are based on analysis of sophisticated theoretical models or the instinct of a speculator. New information about company developments and stock recommendations are continuously made available while papers are released on new and different ways in which the market can be exploited. But can the market really be exploited?

Eugene F. Fama (1965) described how an active market filled with well informed and "intelligent participants" leads to a situation where the stock price reflects its actual value. This is due to the situation in which investors compete for new available information about the stock for profit. The stock will then promptly reflect the new price that the information retains. This is known today as the Efficient Market Hypothesis (EMH).

The EMH is a controversial idea, even today, as many investors and active fund managers truly believe that there is value in exploiting the timing of market. However, the great irony of the EMH is the market's ability to promptly correct itself when presented with news regarding a new inefficiency or mis–pricing: news which many investors attempt to exploit. This implies that it is not possible to make above-average returns. Once new information becomes available it "triggers a rapid process of adjustments, and re-prices the stock to its "correct" level" (Kingdon 1997).

However, there also exist anomalies in the market which contradict the EMH such as the size and value effect as described in the work by Fama and French in (1993). The size effect states that stocks with smaller portfolios of companies will perform better in the future while the value effect suggests that firms with a high book ratio in relation to its market price will also outperform.

The aim of this work is to investigate the EMH by testing existence of the size and value effects using a backpropagation multilayer perceptron (Reed & Marks II 1999) (Bishop 1995). In the process we examine the attributes from the three factor model developed by Fama and French that describe the size and value effects.

There has also been other work in the prediction of stock performance including studies by Gaunt (2004) and Albanis and Batchelor (2000).

Evidence of the three factor model as an effective pricing model in an Australian context can be seen with Gaunt (2004) which updates the study Halliwell, Heaney and Sawicki (1999) by examining Australian companies from the period of 1991–2000. The analysis shows that the three factor model has more explanatory power in predicting the future return on assets than a simpler one factor model, CAPM (see Section 2.2). Unlike our study, Gaunt (2004) divided the dataset into 25 portfolios ranking them into varying amounts of book to market value and market capitalization. Gaunt's results were consistent with Halliwell (1999) in that the three factor model provided better explanatory power than the traditional CAPM model for performance of portfolios. He found evidence of both the size and value effects and observed that less risky stocks offered better raw return. Our work is similar to Gaunt in that we investigate Australian companies. However, our study extends to more recent data and we also apply a neural network to predict individual stock performance.

Albanis and Batchelor (2000) describe different models of analysis for identifying high performing shares. They used analysis techniques including probabilistic neural networks, vector quantization, recursive partitioning and rule induction to investigate stock performance on the London stock exchange from 1991 to 1997. They observed that nonlinear approaches gave better classification performance than linear methods. Our work differs from this study in terms of recency of the data and the Australian context.

In section 2 we describe the Efficient Market Hypothesis and the Three Factor Model of return. Next, in section 3 we describe the data attributes used as proxies to components of the Three Factor Model and the source and preprocessing applied to the data. In section 4 we describe the multilayer perceptron used to learn the relationship between data attributes and the performance of the company. Next in section 5 we detail the experiments run and their results. That section includes an account of initial trials predicting the performance class of companies and investigations into relationships between input data attributes using linear regression and scatterplots. Also we examine the effect on the learning task of using longer periods of data as well as brief details about other investigations. In section 6 we provide a summary of the results of the work and, finally, in section 7 we outline what we think are interesting further directions for study.

2 The Problem Definition

The aim of this paper is to implement a neural model to test Fama and French's theory of size and value effect in influencing stock performance. The theory suggests that firms with low market capitalization (i.e. small firms) and firms with high book to market ratios tend to perform better in the future. This investigation also tests the validity of the EMH where excess returns due to the size and value effect can only be gained through taking on extra risk.

2.1 Efficient Market Hypothesis

The efficient market hypothesis states that at any given time, stock prices fully reflect all available information of the asset's value. All past stock information can be reflected in the current stock price where this price changes only with the availability or release of new information. The present value of the stock is determined by discounting the expected future cash–flows (or dividends) of the stock by using all information investors have available to them (Kingdon 1997).

Using the present value model, the value of a stock can computed as

$$\mathrm{PresentValue} = \mathrm{CF}_1/(1{+}r){+}\ldots{+}\mathrm{CF}_i/(1{+}r)^i{+}\ldots{+}\infty \tag{1}$$

where CF represents the cash flow for the period, and r represents the required rate of return for the period. The term $\mathrm{CF}_i/(1+r)^i$ represents the return for year i. The present value is the total sum of all the future values.

This means that the main determinant of a stock price lies in r. Generally speaking, the riskier the company, the higher the rate of return demanded by investors and the higher r (Brailsford & Heaney 1998). For instance, if a firm is risky, investors may only want to buy the stock if a return of 23% is guaranteed. Hence if the expected annual dividend for the rest of the company's life is \$0.50, then the expected value of the share is (\$0.50/1.23) + (0.5/1.232) + (0.5/1.233) + ... + ∞ = \$2.17. So if new information about a firm's future cash flow indicates that the current market price is lower than the stock expected value (e.g. \$2), then according to the EMH, investors will see that the share is undervalued resulting in purchase of more quantities of the stock, thus raising it to its fair value (i.e. back to \$2.17).

2.2 Three Factor Model

Using the three factor model, Fama and French (1993) argue that the rate of return r for a portfolio of stocks is determined by three attributes: (i) its return in relation to the market; (ii) its size; and (iii) its book to market ratio. These are known as the Capital Asset Pricing Model (CAPM), the size effect and the value effect respectively. The size effect indicates that portfolios of firms with low market capitalization (smaller firms) will perform better than the average market return in the long run, while the value effect suggests that portfolios of firms with high book to market ratios will also perform higher than average. According to Fama and French's three factor model, portfolios of firms with higher book to market ratios and low market capitalization tend to perform well (better than the market) as they tend to be more risky. To account for the extra risk, they will require a higher rate of return r from the stock in the future. This is modeled by the equation

$$r = \mathrm{CAPM} + b_s \times \mathrm{SMB} + b_\sigma \times \mathrm{HML} + \alpha. \tag{2}$$

where CAPM refers to a model used by investors to determine the rate of return for valuing a portfolio of stocks. CAPM is defined as

$$\mathrm{CAPM} = r_f + \beta(E[r_m] - r_f) \tag{3}$$

where r_f is the rate of the risk free asset (generally the government bond rate), $E[r_m]$ is the expected return of the market and β is the sensitivity of the stock to the market. In this study, r_f and $E[r_m]$ are constant for all companies at particular time periods, allowing us to use β to be an effective proxy for CAPM.

The Three Factor Model extends CAPM by adding two other factors namely SMB (small minus big) and HML (high minus low). Both these factors reflect the excess return that stocks of smaller companies and stocks with high book to market ratios are capable of delivering. The coefficients b_s and b_σ show the relative scale of the factors in relation to the portfolio, with $b_s = 1$ representing a portfolio having small capitalization and $b_s = 0$ representing a portfolio with large capitalization. Similarly, b_σ shows how high the firm's book to market ratio is compared to the market.

In summary, the Three Factor Model predicts a firm's future expected return on the basis of its return in relation to the market, its size and its book to market ratio.

3 Data

We examine evidence of the following in affecting the performance of companies:

1. CAPM model of returns;
2. size effect; and
3. value effect.

However, as the Three Factor Model is used to examine *portfolios* of stocks rather than individual stocks and due to the difficulty of obtaining information on portfolios of stocks with the inherent inconsistencies of the available data, we will proxy data items for the factors in equation (2). The proxy data items for individual companies (Beta, Market Capitalization and Book to Market Ratio respectively) are used instead of the actual CAPM, size effect and value effect. Hence the neural network will examine the effect that Beta, Market Capitalization, and Book to Market Ratio have on the future return of each individual firm. Additionally, another data item "Standard Deviation" will be used as a proxy and control factor for volatility.

Financial data for a group of companies listed on the Australian Stock Exchange (ASX) for the years 2000–2004 were obtained from Aspect Huntley [1] and

[1] See `http://www.aspectfinancial.com.au/af/aerhome?xtm-licensee=aer` and `http://www.aspectfinancial.com.au/af/finhome?xtm-licensee=finanalysis` for details.

from quarterly reports from the Australian Graduate School of Management (AGSM). To ensure the consistency of returns over the time period, only companies which reported returns in June by Aspect Huntley and companies with no missing data were used. From the 1315 companies reported by the AGSM, 346 were used in this study. Despite the filtering, the final dataset still represents a broad cross section of the ASX.

Performance of these companies is divided into three categories: high performing, medium performing and low performing. High performing companies are those with market return over the calculated period falling into the top third of the group, with the medium comprising of the next third and low performing companies in the bottom third.

Data for Beta, Market Capitalization and Standard Deviation were obtained from the AGSM while Book to Market Ratio and the category rankings came from Aspect Huntley.

Next we describe the individual input factors.

Market Capitalization of a firm is the value of the total amount of stock it has outstanding. It can be calculated by multiplying its current share price by the number of stocks it has on issue. The Market Capitalization is a useful indicator of the size of a company, and has been used to evaluate the effectiveness of Fama and French's size effect in determining returns (Fama & French 1995).

The Book to Market Ratio is the historical (or accounting value of a firm) divided by its Market Capitalization. It will be used to determine whether a stock is over or undervalued. The book value represents the net assets of a firm calculated by subtracting its current assets from its current liabilities.

Standard Deviation measures volatility or risk of an investment by showing the average amount by which it deviates from the mean. Generally speaking, the higher the standard deviation of a stock the higher its risk. It will be used as the control factor for risk to test the size effect and value effect. It is understood from Fama and French's work that the smaller companies and value stocks encounter greater risk, and hence it is necessary to determine how much excess return is attributed to higher risk and how much is attributed to the size factor.

4 Neural Network Model

A fully connected feed forward multilayer perceptron (MLP) was trained using backpropagation and momentum. Multilayer perceptrons were chosen for this study due to their known ability to act as universal approximators (Haykin 1999) and hence are suitable for this type of non-linear problem.

Preliminary investigations compared performance of the MLP with Support Vector Machines (SVM) and Naive Bayes classifiers on the datasets. Performance by the SVM was not markedly better than with the MLP so we chose to continue with the MLP in the investigations.

All networks in this study had four inputs, three output neurons and one hidden layer. The four inputs correspond to the input factors: Beta, Market Capitalization, Book to Market Ratio and Standard Deviation respectively. The three output neurons correspond to the category ranking classes: low, medium and high performing respectively. Our networks contained between three and eight neurons in the hidden layer. All neurons in the hidden and output layers used the sigmoid transfer function.

All data was scaled into $[0, 1]$ before presentation to the network. Additionally, Market Capitalization was first transformed with the logarithmic function before scaling so as to compress the range of possible values due to the enormous variation in the values.

The output class for a particular input pattern was the one associated with the output neuron having the highest value.

5 Experiments

5.1 Initial Results

Initially we compared two methods of training the MLP: (i) stopping training using a holdout set; and (ii) training for a fixed number of epochs with estimation of test error using 10–fold cross validation (Witten & Frank 2005).

The dataset of companies was divided randomly into three datasets for training with the holdout set method. Seventy percent of the data was presented to the MLP for training. The next 15% was used to determine when to stop training the MLP. When the error on this set started to increase training was stopped. The final 15% was used for quoting the accuracy of the MLP in the classification task. The other method of training the MLP used a 10–fold cross validation scheme with all available data.

Networks were trained on data from 2000 to 2001. Results are shown in Table 1. The results show that the best accuracy arose using the holdout method with 7 neurons in the hidden layer. This accuracy is much greater than random choice of the company's class of performance. While this is an improvement on the random model of 33%, it is not a high number. These initial results over the 2000–2001 period do not strongly support the value and size effects due to the limited predictability using the input variables.

We next investigate potential reasons for the relatively poor classification accuracy.

5.2 Investigating relationships between the input variables and the output

Linear regression analysis is used to investigate the degree to which the input variables affect the output results. In particular, we are interested in which input variables are the strongest predictors of the output class. Table 2 shows a regression of inputs to the real–valued output (i.e. the actual performance rather than the class). There is only a very weak relation between the output and the input attributes as evidenced by an R–squared value of 0.058. The only factor which has a strong linear relation with the output variable is the Standard Deviation, as indicated by its $\rho < .05$ within a 95% confidence interval. The negative coefficient for the Standard Deviation demonstrates that the higher the volatility (an indicator of risk) of the company, the higher the likelihood of a poor return in the future. Beta, Market Capitalization and Book to Market Ratio do not show any statistically significant linear relationship with performance.

Table 2: Results from linear regression of inputs to performance. 2000–2001 data.

Attribute	Coefficients	P–value
Intercept	30.91	0.13
Market Capitalisation	11.10	0.67
Beta	-37.88	0.25
Standard deviation	-61.49	0.02
Book to market ratio	58.70	0.08

Figures 1 and 2 show scatter plots of the input attributes for the 346 companies for the 2000 to 2001 period. All values on the axis have been scaled between 0 and 1. The plots show that there is a high

Table 1: Results of training neural networks, 2000–2001 data.

Hidden Neurons	Parameters	Accuracy	Mean Absolute Error
	Training using the holdout method		
3	Learning rate 0.1, momentum 0.5	47.83%	0.38604
4	Learning rate 0.2, momentum 0.5	45.45%	0.36732
5	Learning rate 0.2, momentum 0.4	50.00%	0.38785
6	Learning rate 0.1, momentum 0.7	45.65%	0.37204
7	**Learning rate 0.1, momentum 0.7**	**58.70%**	**0.34210**
8	Learning rate 0.1, momentum 0.7	47.83%	0.39517
	Training with fixed 500 epochs and 10–fold crossvalidation		
3	Learning rate: 0.3, momentum 0.2	50.00%	-

degree of overlap between the classes and that they cannot be easily differentiated. This is one reason why the MLP has difficulty with this problem. Figure 1 plots Beta against Market Capitalisation and shows that the dataset is unevenly skewed to smaller companies. This is expected due to market's value dominated by a small number of very large companies. Although most of data for the different categories is overlapping, a slight pattern emerges for firms with a Market Capitalisation greater than 0.5. There are very few low performing firms in this region and more high performing than mid performing suggesting that larger firms perform better than smaller firms and contradicting the size effect.

5.3 Longer periods

The data used up to this point has been from the period June 2000 to June 2001 which coincides with the period directly after the market crash caused by technology stocks. This period was a time of considerable restructure in the stock market globally. The size and value effect are known to be more apparent after a longer period of time (Arnott 2005). It is hence necessary to investigate longer periods of time.

Table 3 summarizes test accuracies for MLPs trained on different periods. All networks were trained for 500 epochs (learning rate: 2, momentum: 3) and had a topology of three hidden neurons. The table shows, as expected, that MLPs trained with data over a longer period were generally more accurate. This suggests that the size and value effect is evident in the dataset, however require the dataset to be analyzed from longer periods of time in order for it to emerge more clearly.

Figure 3 plots Standard Deviation against the Book to Market Ratio for the period 2000–2003. Data points for each class of company cluster reasonably clearly into three distinct areas. This suggests that the higher the Standard Deviation, the lower the future returns for the firm implying that firms which are more volatile tend to perform poorly in the long term, and reinforcing our findings from the linear regression. The value effect can also be observed in Fig. 3 with the low performing firms clustered towards the lower end of the Y–axis (book to market ratio) while high performing firms are spread towards the higher areas of the graph. This suggest that underpriced firms (with high Book to Market Ratios) will revert back to their true value over the three year period and generally exceed the market's performance.

Figure 4 plots Beta against Market Capitalisation. It is interesting to compare this graph with Fig. 1 which plots the same values over the shorter period. In Fig. 1 large firms (with Market Capitalization greater than 0.5) tend to be the best performers, however after three years the majority of firms with large Market Capitalisation actually becomes mid performing firms. This shows that the large companies which may have performed well in 2000–2001 but after another two years their performance fell to a medium level. This anomaly may have been caused by investors favoring well established large companies over the more risky technology stocks after the crash of 2000, pushing their stock price up, but then abandoning them later when investors realized they were overpriced. This observation does not entirely reinforce the size effect as there is a mix of high and low performing smaller firms, suggesting that the size of a firm does not have a direct bearing on its future performance. Even though only the data from 2000-2003 has been presented here, analysis of the other periods also yielded similar findings.

5.4 Other Investigations

Whilst investigating the relationship between Market Capitalisation and risk, a MLP was trained to determine the current Market Capitalisation from the Beta and Standard Deviation. Using this analysis a correlation was discovered between the size of a company and it's risk. Specifically, that larger companies had a lower Standard Deviation and had a positive correlation with Beta. This reinforces Fama and French's suggestions indicated earlier that smaller firms were underpriced due to their inherent riskiness.

Investigations of classifying companies into five performance classes rather three resulted in a 10–fold cross validation scheme yielded accuracy of 34.50% compared to random selection of 20% over the 2002–2004 period. The confusion matrix (Table 4) shows that the neural network has difficulty differentiating between the high performers and the low performers. It was observed that high performing stocks tended to be either very large or very small firms.

Table 4: Confusion matrix of classification into five performance classes rather than three.

Very High	High	Medium	Low	Very Low	classified as ←
48	5	1	3	23	Very High
31	2	7	8	23	High
9	10	9	11	13	Medium
15	2	7	14	28	Low
18	3	0	7	45	Very Low

Other investigations which were conducted included the removal of Technology firms from the dataset (as classified by the Global Industry Classification Standard), in order to control for the influence that the technology crash of 2000 had on the results. It was observed that the network found it more difficult to predict the performance of the remaining companies after the removal. Elimination of outliers from the data set also did not yield any improvement in the results.

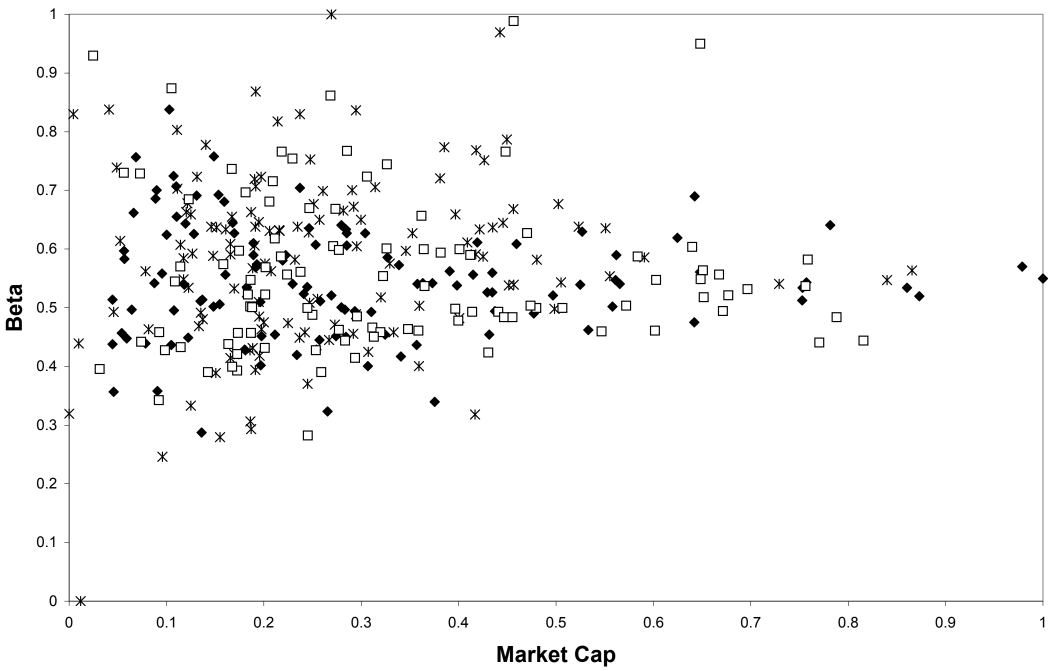

Figure 1: Scatter plots of attributes for 2000–2001 data: Beta vs Market Capitalisation. × = low performing companies, ♦ = medium performing companies, □ = high performing companies.

Table 3: Accuracy of MLP trained on different periods of data. Accuracy is the 10–fold cross validation value.

Period from	2000				2001			2002	
Period to	2001	2002	2003	2004	2002	2003	2004	2003	2004
Accuracy (%)	50.00	51.16	58.55	56.52	45.09	52.75	54.49	59.36	57.02

6 Conclusion

This study examined evidence for the size and value effect by building MLP models to predict the class of performance of ASX–listed companies over the period 2000–2004.

We were able to classify 58.70% of the companies correctly on 2000–2001 data with an MLP built using the holdout method. This network had 7 neurons in the hidden layer. By contrast, the best network we were able construct with 3 neurons in the hidden layer training for 500 epochs had an accuracy estimated by 10–fold cross validation of 50.00%.

Investigation of the 2000–2001 dataset with linear regression suggested that only the Standard Deviation attribute showed any significant relationship with the performance of the firms. Standard Deviation was inversely proportional to the future return of the company suggesting that investment into volatile companies led to poor investment return. Scatter plots of the attributes did not show distinct areas for different performance classes.

However, analysis of longer time periods, specifically 2000–2003, showed evidence of the value effect with the accuracies of 3–neuron hidden layer MLPs considerably higher than that of the 2000–2001 time periods. There is only weak evidence for the size effect, with small firms tending to be both high performers as well as poor performers. It was, however, discovered that the high performing large firms from 2000 had fallen to medium performance over a longer time period.

In terms of the three factor model, the only factor which was clearly observed was the value effect as proxied by the Book to Market Ratio. There was a relationship between Book to Market Ratio and the future return on a stock over the long term, implying that underpriced firms will eventually revert to their fair value. The findings show evidence against the EMH as existence of the value effect represents an anomoly in the expected behavior of stocks. The outperformance of high book to market firms also cannot be simply attributed to higher risk, with our findings indicating that higher Standard Deviation leading to lower returns in the future. Although this study did not *directly* examine the three factor model as individual stocks were examined rather than portfolios and also proxied for values, it does offer new insights into the size and value effect on the Australian stock market.

Our results support Gaunt's findings in (2004) that there was a value effect, however, we were unable to find strong evidence for the size effect. Our findings also support the study by Albanis and Batchelor (2000) in that nonlinear methods resulted in more accurate models than linear models.

7 Future Work

Further work with other financial measures to predict future investment return, such as the price to earnings ratio or the debt to equity ratio, would yield interesting insights into the EMH.

Also, given the superior performance of the seven hidden neuron MLPs over those with three hidden neurons, it would also be interesting to extend the additional modeling of longer time periods and finer granularity performance classes to these networks.

Moreover, extending Gaunt's study to more recent data with portfolios of stocks rather than individual companies may yield additional insights into the size and value effects.

Also, it would be interesting to compare models built with the MLP with an SVM. As mentioned

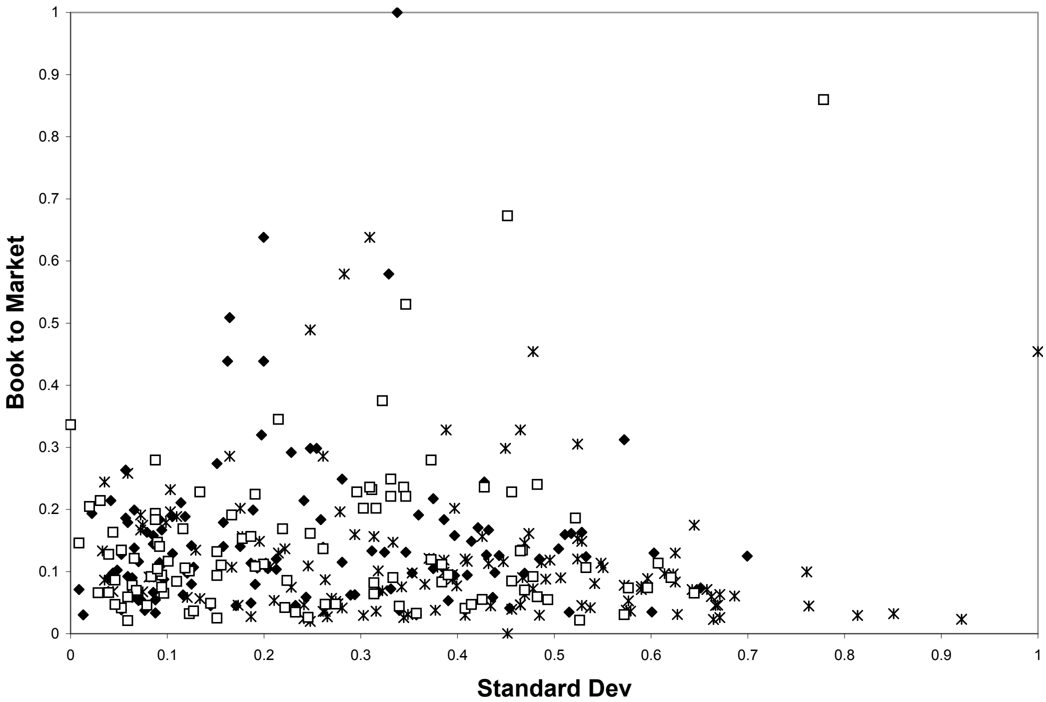

Figure 2: Scatter plots of attributes for 2000–2001 data: Book to market ratio vs standard deviation. × = low performing companies, ♦ = medium performing companies, □ = high performing companies.

above, we conducted initial experiments comparing performance of an MLP and an SVM in this task and found that there was not significant differences in the results. However, more careful modelling, particularly using lessons learnt in this work, may allow more accurate SVMs to be constructed.

References

Albanis, G. T. & Batchelor, R. A. (2000), Five classification algorithms to predict high performance stocks, *in* C. Dunis, ed., 'Advances in quantitative asset management', Kluwer Academic Publishers, pp. 295–318.

Arnott, R. D. (2005), 'Disentangling size and value', *Financial analysts journal* pp. 12–15. CFA Institute.

Bishop, C. M. (1995), *Neural Networks for Pattern Recognition*, Oxford University Press, Oxford.

Brailsford, T. & Heaney, R. (1998), *Investment: Concepts and applications in Australia*, Harcourt, London.

Fama, E. E. & French, K. (1993), 'Common risk factors in the returns on stocks and bonds', *Journal of Financial Economics* (33), 3–56.

Fama, E. E. & French, K. (1995), 'Size and book to market factors in earnings and returns', *Journal of Finance* (50), 131–155.

Fama, E. F. (1965), 'Random walks in stock market prices', Paper No. 16 in the series of Selected Papers from the Graduate School of Business, University of Chicago, 1965. Reprinted in the Financial Analysts Journal (Sept–Oct 1965), The Analyst Journal, London (1966), The Institutional Investor, 1968.

Gaunt, C. (2004), 'Size and book to market effects and the Fama French three factor asset pricing model: evidence from the Australian stockmarket', *Accounting and Finance* **44**, 27–44.

Halliwell, J., Heaney, J. & Sawicki, J. (1999), 'Size and book to market effects in Australian share markets: a time series analysis', *Accounting Research Journal* **12**, 122–137.

Haykin, S. (1999), *Neural networks: a comprehensive foundation*, 2nd edition edn, Prentice–Hall.

Kingdon, J. (1997), *Intelligent Systems and Financial Forecasting*, Springer–Verlag Telos.

Reed, R. D. & Marks II, R. J. (1999), *Neural Smithing*, MIT Press, Cambridge, Massachusetts.

Witten, I. H. & Frank, E. (2005), *Data mining: Practical machine learning tools and techniques*, second edn, Elsevier.

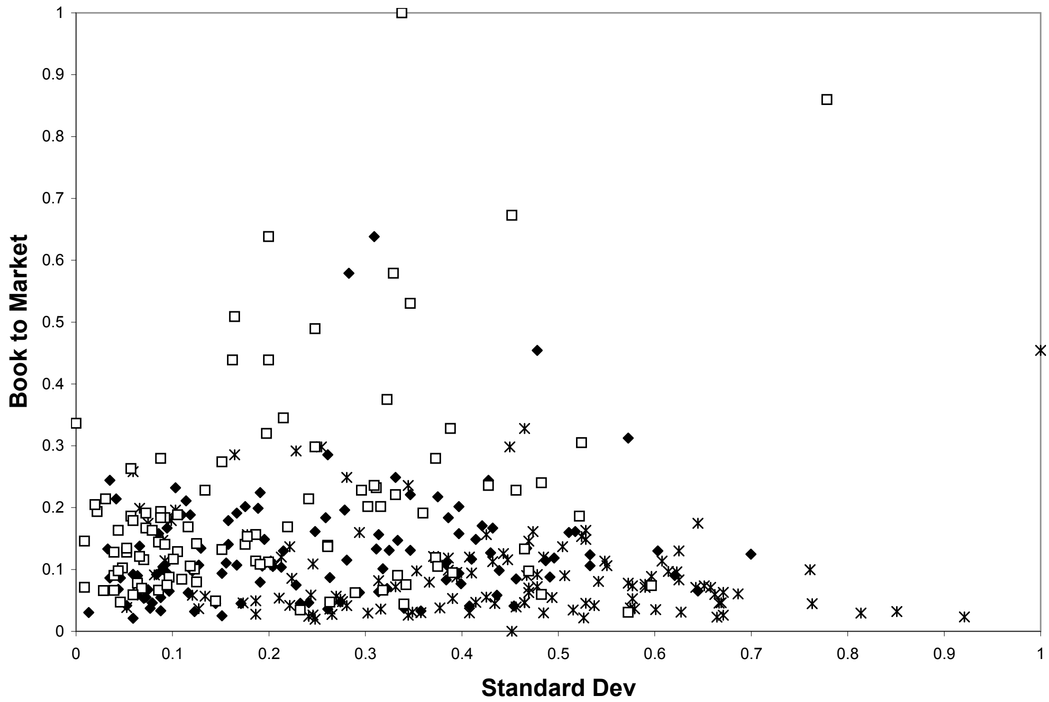

Figure 3: Scatter plots of attributes for 2000–2003 data: Standard deviation vs book to market ratio. × = low performing companies, ⧫ = medium performing companies, □ = high performing companies.

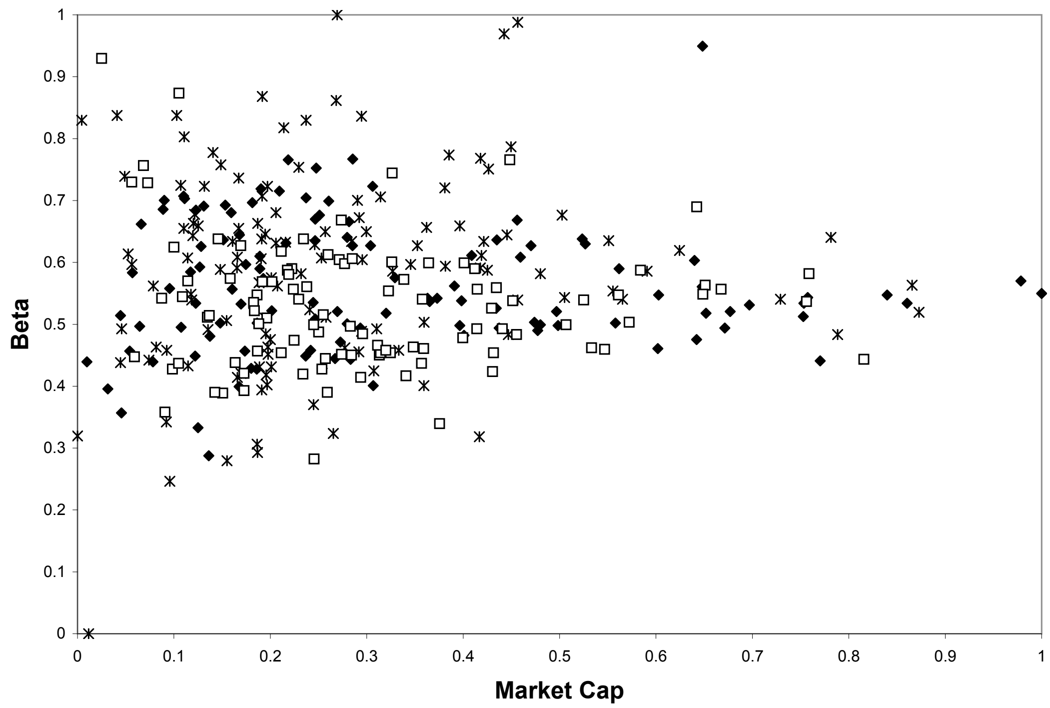

Figure 4: Scatter plots of attributes for 2000–2003 data: Market capitalisation vs beta. × = low performing companies, ⧫ = medium performing companies, □ = high performing companies.