School of Information Technology

Bond University

Trading in the Australian Stockmarket using Artificial Neural Networks

by

Bruce James Vanstone

Submitted to Bond University in fulfillment of the requirements for the degree
Doctor of Philosophy

November 2005

Abstract

This thesis focuses on training and testing neural networks for use within stockmarket trading systems. It creates and follows a well defined methodology for developing and benchmarking trading systems which contain neural networks.

Four neural networks and consequently four trading systems are presented within this thesis. The neural networks are trained using all fundamental or all technical variables, and are trained on different segments of the Australian stockmarket, namely all ordinary shares, and the S&P/ASX200 constituents.

Three of the four trading systems containing neural networks significantly outperform the respective buy-and-hold returns for their segments of the market, demonstrating that neural networks are suitable for inclusion in stockmarket trading systems.

The fourth trading system performs poorly, and a number of reasons are proposed to explain the poor performance. It is significant, however, that the trading system development methodology defined in this thesis clearly exposes the potential failure when testing in-sample, long before the neural network would be used in real trading.

Overall, this thesis concludes that neural networks are suitable for use within trading systems, and that trading systems developed using neural networks can be used to provide economically significant profits.

Statement of original authorship

This thesis represents my own work and contains no material which has been previously submitted for a degree or diploma at this University or any other institution, except where due acknowledgement is made.

Signature

Witness

Date

Additional Publications

The following is a list of publications by the candidate on matters relating to this thesis.

Conference

Vanstone, B. and C. N. W. Tan (2003). A Survey of the Application of Soft Computing to Investment and Financial Trading. Proceedings of the 8th Australian & New Zealand Intelligent Information Systems Conference (ANZIIS 2003), Sydney.

Vanstone, B., G. Finnie, et al. (2004). Applying Fundamental Analysis and Neural Networks in the Australian Stockmarket. Proceedings of the International Conference on Artificial Intelligence in Science and Technology (AISAT 2004), Hobart, Tasmania.

Vanstone, B., G. Finnie, et al. (2004). Enhancing Security Selection in the Australian Stockmarket using Fundamental Analysis and Neural Networks. Proceedings of the 8th IASTED International Conference on Artificial Intelligence and Soft Computing (ASC 2004), Marbella, Spain.

Vanstone, B., G. Finnie, et al. (2005). Evaluating the Application of Neural Networks and Fundamental Analysis in the Australian Stockmarket. Proceedings of the IASTED International Conference on Computational Intelligence (CI 2005), Calgary, AB, Canada, ACTA Press.

Book Chapter

Vanstone, B. and C. N. W. Tan (2005). Artificial Neural Networks in Financial Trading. Encyclopedia of Information Science and Technology. M. Khosrow-Pour, Idea Group. 5**:** 163-167.

Acknowledgments

This thesis was jointly supervised by Professor Gavin Finnie and Adjunct Professor Clarence Tan, both from Bond University. I would like to take this opportunity to formally thank them both for the guiding influence and constructive comments they have made throughout this research. Further, I would also like to thank Professor Gavin Finnie for his contributions to my development in areas apart from research, such as in teaching, and for providing an academic role model for me to aspire to.

I would also like to thank Associate Professor Kuldeep Kumar for the many times he offered his statistical expertise, which I was most grateful to receive.

My thanks also go to Bond University for accepting me as a doctoral candidate, and for the provision of excellent research support and facilities. Without the level of research support I received at this university, from both academic and non-academic staff alike, the journey undertaken in this thesis would have been considerably more difficult.

The contribution of numerous other people is also gratefully acknowledged. These include:

- Graham Johnson and Brian Boniwell, both from Bond University library, for being ever willing to help with locating all manner of research publications,

- Jim McElroy and Travis Johnston, both from IT Support, for being ever willing to assist with all manner of support issues that arose during the lifetime of this thesis

Dedication

This thesis is dedicated to my family.

To my wife, Sue, and my children Daisy and Serena, for their continued support and understanding over the years that this journey has taken.

To my parents, Patricia and James, for instilling in me the values of dedication and hard-work, without which this thesis may never have been completed.

Table of Abbreviations

A variety of specialized abbreviations are used within this thesis. The more obscure of these terms are listed below.

ADX:	Average Directional Index
AMEX:	American Stock and Options Exchange
AORD:	Australian All Ordinaries Index
ASX:	Australian Stock Exchange
ATR:	Average True Range
DAX:	German Stock Exchange Index
DJIA:	Dow Jones Industrial Average
FTSE:	London Stock Exchange Index
IBEX:	Spanish Stock Exchange Index
KOSPI:	Korean Stock Exchange Index
MACD:	Moving Average Convergence/Divergence
MOM:	Momentum Indicator
NIKKEI:	Tokyo Stock Exchange Index
NYSE:	New York Stock Exchange
RSI:	Relative Strength Index
SESALL:	Singapore All Equities Index
STOCHK:	Stochastic (Momentum) Indicator
TOPIX:	Tokyo Stock Exchange Price Index
TUNINDEX:	Tunisian Stock Exchange Index

Table of contents

List of tables

List of figures

List of Equations

Chapter 1 Introduction

1.1 Background to the research

Almost every individual is influenced in some way by movements in the stockmarket. These could be direct influences, for example, the changing account balances of a day trader, or they could be much longer-term, for example, the balance of an individual's pension fund. Despite significant advances in certain areas of financial understanding, there is still no formal model that describes the underlying mechanics of the stockmarket.

Indeed, the issue of 'market efficiency', or what is known in the popular press as the 'random walk theory', is one of the most hotly debated and thoroughly examined areas within the field of finance. Yet, market efficiency still remains a hypothesis, and volumes of well respected publications gather increasing amounts of evidence against it.

An interesting early review of some of this evidence is provided by Lehmann (1991), who then claimed it is now 'open season on the efficient market hypothesis'. A comprehensive review of deficiencies in the efficient market hypotheses is provided by Haugen (1999). There is also speculation amongst academics as to whether efficiency is in some way related to the maturity of the market itself. Los (2000) adds some credibility to this suggestion, by performing nonparametric testing on all six major asian stockmarkets and finding them all to be inefficient, and that none exhibit random walk behaviour.

Since the early work of Fama (1970), it is common to discuss three forms of efficiency when discussing the Efficient Market Hypothesis (EMH). A market is 'weak form efficient' if it is not possible to consistently earn excess returns using past prices and returns. A market is 'semistrong efficient' if it is not possible to consistently earn excess returns using any public information. A market is 'strongly efficient' if it is not possible to consistently earn excess returns using any information, including private information.

From a trading point of view, market efficiency in general means that it is not possible to consistently earn excess returns using any available information. In essence, then, the only thing that causes security prices to change is new information. By its definition, the arrival and timing of new information is unpredictable. Therefore, in an efficient market, security prices should appear to be generated randomly.

This thesis does not wish to enter the fray which dominates acceptance or otherwise of the efficient market hypothesis. After all, even the efficient market hypothesis accepts the notion of anomalies and bubbles in pricing to some extent. Instead, this thesis concentrates on the development of financially rewarding trading methodologies using ANNs, and leaves the findings of this thesis to stack up alongside the great many other published works which document ways to exploit share price movements in a market which may or may not be technically 'efficient'.

As demonstrated within the literature review, a great deal of well established literature exists which demonstrates persistent, abnormal security price deviations from those that would be implied by a strictly efficient market. There is also an interesting obvious paradox here, as described by Lorie et al (1985), in which the market could only ever be efficient if some investors believed it to be inefficient. In other words, if all participants believed the market to be efficient, there would be no incentive to seek out new information, and therefore, no ability for this information to be assimilated into share prices.

The challenge of stockmarket trading is to find ways to signal when security prices deviate (or are expected to deviate) from the expected spectrum, and to take advantage of these opportunities. Trading is conducted in an environment in which trading decisions incur costs, and in which the data used to drive the information process is 'noisy'. A ‘noisy’ environment is one characterized by a low signal-to-noise ratio, one in which determinant signals may be swamped amidst other, non-relevant information. It is in noisy environments such as this that traditional computing typically gives way to soft computing, as the rigid conditions applied by traditional computing cannot be met. There is also acceptance within the academic community that the relationship between security prices (and returns), and the variables that constitute that price (return), changes over time as described by Refenes et al. (1993) and also by Thawornwong and Enke (2004). In other words, the structural mechanics of the market may change over time, and their effects on prices are also changing. Further, there is a general acceptance in the academic community that many of the relationships concerning security prices (and returns) are most likely non-linear. As suggested by Ferreira et al. (2004), one of the most relevant characteristics of ANNs is their ability to represent complex non-linear relationships without making prior assumptions about the data distribution.

This thesis concerns itself with the development of mechanical stockmarket trading strategies using artificial neural networks (ANNs) as signal generators. It focuses on developing and benchmarking ANNs, and the subsequent issues of testing these ANNs within the context of valid stockmarket trading systems. As such, it aims to contribute knowledge that will one day lead to a model of the underlying mechanics of the stockmarket pricing process.

Further, it aims to present a well-defined methodology that can be used to create and benchmark trading systems. By presenting a well-defined methodology, it is hoped that this thesis can also address many of the deficiencies of published research in this area. Absence of a well-defined methodology means that very little published work is directly comparable. This has lead to a plethora of ANNs being developed, some of which demonstrate superior predictability in their own domains. However, as these ANNs are very rarely sited within trading systems, the actual effect of using them as trading signal generators cannot be assessed. When these ANNs are sited within trading systems, it is often unclear as to why various parameter choices which drive the development of the host trading system have been made.

When developing and testing ANNs, there appears growing confusion over how to effectively benchmark the ANNs to be used for trading. For example, Azoff (1994) and Thawornwong and Enke (2004) both question what to measure and suggests traditional measures of forecasting performance may not be strongly related to profits from trading. For example, as shown in the literature review, some researchers attempt to predict actual values of their chosen target, others try to predict the direction of change of the target from its last recorded observation. Thawornwong and Enke state that approaches such as predicting direction (or sign) may lead to higher observed levels of accuracy, but this does not necessarily lead to higher profitability. Azoff claims that the effectiveness of an ANN developed for trading cannot be assessed by its prediction accuracy. For example, an ANN with low prediction accuracy may be much more profitable than an ANN with higher prediction accuracy, if the ANN with the lower prediction accuracy is better at predicting large price moves.

In essence, there is a need for a formalized trading system development methodology, and also a formalized trading system benchmarking methodology. Both are accommodated by this thesis.

1.2 Research problem and hypotheses

Due to some of the problems outlined above, it is still not possible to answer the question:

"Can ANNs be used to develop economically significant stockmarket trading systems?".

From the variety of research summarized in the literature review, it is clear that a great deal of research in this area has taken place outside of the constraints imposed by real-world trading. This directly threatens its applicability to industry. According to Zirilli

(1997), and confirmed by this literature review, issues such as accounting for transaction costs, and money management are rarely mentioned. Yet these are issues that must be addressed for the question posed above to be answered, as they are practical limitations on our theory.

This thesis will attempt to answer the above question within the constraints and scope of the 10-year sample period (from 1994-2003) using data for Ordinary shares from the Australian stockmarket. Further, it will attempt to answer this question within the practical constraints of transaction costs and money management imposed by real-world trading. Although a formal statement of hypotheses is left until section 3.10, it makes sense to discuss the way in which this thesis will address the above question.

In this thesis, two neural networks will be trained using fundamental data, and two neural networks will be trained using technical data. A trading system development methodology will be defined, and these neural networks will be sited into valid trading models. These valid neural trading models will then be comprehensively tested out of sample, and benchmarked to their buy-and-hold naïve equivalents. In this way, the benefits of incorporating neural networks into trading strategies can be exposed and quantified. Once this process has been undertaken, it will be possible to answer the thesis question.

1.3 Justification for the research

The seminal ideas of Harry Markowitz provided a turning point in modern investment finance. These ideas yielded the 'efficient frontier' for portfolios, and resulted in an algorithmic procedure for choosing portfolio weights so as to minimize variance for any feasible expected return. This was made possible by increasing the weights on subsets of sufficiently anti-correlated stocks. However, as discussed by Borodin et al (2004), known universal portfolio selection algorithms do not seem to provide any substantial benefit over a naïve investment strategy.

Over time, this had lead to a proliferation of portfolio management approaches, and consequently, to a proliferation of trading strategies, with the universal goal of outperforming the market.

The widespread introduction of soft-computing techniques has led to an ever increasing number of researchers attempting to use techniques such as ANNs to select securities, with the same goal of outperforming the market. Studies of this nature often falter due to a lack of understanding of the real-world constraints that afflict traders. These concern such areas as transaction costs, money management, and occasionally, even data quality (see for example, Versace et al(2005) who use particularly poor quality data sourced from Yahoo). In this context, money management concerns the appropriate amount of capital to apply to a trading position, and is also known as position allocation, or capital allocation. Further, some research makes inappropriate judgments about a markets ability to tolerate short selling of securities. There is definitely a need for researchers to gain a greater understanding of how to develop mechanical trading systems to enable us, as academics, to stay relevant in this area.

According to Thawornwong and Enke (2004), ANNs appear to offer the ability to generate higher profit with lower risks than the naïve buy-and-hold approach, conventional linear regression and the random-walk model.

It is tempting to see techniques like ANNs as a panacea for predicting the market, but it is important to realize that ANNs are simply a step in the correct direction to a model which gives deeper understanding of the underlying mechanics of the stockmarket. There is general acceptance amongst the academic community that ANNs are superior at modeling and predicting relationships within the stockmarket domain, due in part to their ability to model non-linear relationships in the noisy environment which the stockmarket represents. There are other researchers, such as Pan (2005), who postulates that as humans seem remarkably poor at predicting changes in the stockmarket, the AI techniques which mirror human intelligence may also not be up to the task of beating the market. Others such as Soros (2003) suggest that developing ANNs for predicting the market may well turn out to be a never-ending evolutionary process due to reflexivity or ever changing cycles in financial markets.

Trading systems may be created by using a variety of methods. They may be based on either fundamental variables or technical variables. They may be simple rule based systems, or pattern based systems, or may rely on a variety of statistical observations for their implementation. The challenge with the majority of trading approaches is they are linear implementations, usually deciding to buy or sell securities based on thresholds in the underlying variables which have been identified as relevant by the trader. As previously mentioned, there is growing acceptance that the underlying characteristics of the market are non-linear. In essence, it is this observation which suggests that ANN based trading systems may yield better results compared to conventional methods.

Finally, there is a significant lack of work carried out in this area in the Australian stockmarket. As such, this thesis draws heavily on results published mainly within the United States, from both academics and (to a lesser extent) from practitioners. One interesting aspect of this point is that it will be interesting to see how much of the published research on stockmarket anomalies is applicable to the Australian market. This is important as some researchers (see, for example, Pan et al (2005)) accept that each market is different, and has its own unique personality.

1.4 Outline of the report

The second chapter, the literature review, details the building blocks for much of the research. As always in a study of this nature, this thesis stands on the body of work of previous researchers, and the literature review describes this body of work, in terms of its content, and contribution. The literature review takes on particular significance when deriving suitable inputs for ANNs. As noted by Thawornwong and Enke (2004), there must be some solid justification as to why a particular set of ANN inputs were selected. Much of the support for the ANN input choices made in this thesis comes directly from the literature review.

The third chapter, the methodology, describes the actual thesis hypotheses in detail, the data, and the mechanics of the study that takes place in this thesis. It also covers such issues as software and hardware used in the study, and the methodology of creating mechanical trading systems.

The fourth chapter, the analysis of data, presents the neural networks and trading systems developed. It describes the choices that needed to be made in creating host trading systems for each network, and justifies these choices in terms of the literature. Finally it subjects each trading system to a barrage of out-of-sample testing, and describes in detail the key metrics that explain the workings of the trading systems themselves.

The fifth chapter, conclusions, restates the thesis hypotheses, and discusses them in terms of the analysis of data. Conclusions are drawn, and the thesis findings are put into perspective. Finally, the next steps for further research are considered.

1.5 Definitions

The term *security* is used within this thesis to describe the Ordinary Shares of the relevant company being discussed.

The term *tradeable* is used to describe a strategy that respects real-world trading constraints, with respect to such issues as timing, costs, and information availability.

1.6 Delimitations of scope

The thesis concerns itself with data for Ordinary shares in the Australian Stockmarket during the period 1994 – 2003.

Chapter 2 Literature Review

2.1 Introduction

The arena of securities trading is dominated by practitioners applying the techniques of one of two main frameworks, namely, Value Investment and Technical Analysis. Both frameworks provide methods by which practitioners can make qualitative and quantitative judgments concerning the future price movements of individual securities. Practitioners attempt to increase their future investment returns by implementing trading strategies based on those judgments. This activity takes place within an environment in which trading decisions incur costs, and in which the data used to drive the information processes is 'noisy'. A noisy environment is one characterized by a low signal-to-noise ratio, or where the determinant signals are buried amidst other non-relevant information. It is in noisy environments such as this that traditional computing typically gives way to soft computing, as the rigid conditions applied by traditional computing cannot be met.

This section of the thesis explores the three parent disciplines of Value Investment, Technical Analysis, and Soft Computing, which together form the conceptual frameworks of the thesis. The parent disciplines of Value Investment and Technical Analysis are reviewed in historical context, sketching out the development of those disciplines, and reviewing their academic credibility, and their application to this thesis. In the case of Soft Computing, the discipline is reviewed with regard to that portion of the literature which deals with applying soft computing to investment trading, and an existing classification model is extended to allow a more detailed analysis of the area than would otherwise have been possible. This sets the scene for the focal area of the thesis, as shown in Figure 1 below, namely the application of Value Investment, Technical Analysis, and Soft Computing to increase trading returns.

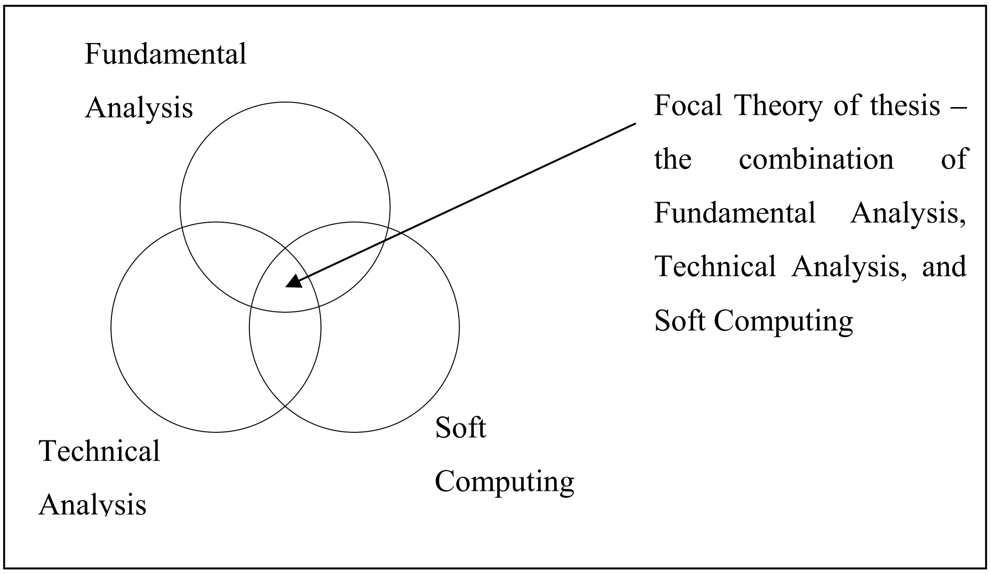

Figure 2-1 Data Theory and Focal Theory of this thesis

2.2 Value Investment

2.2.1 Introduction

As mentioned earlier, the key historical developments in Value Investment will be reviewed initially, followed by a discussion of their academic credibility, and finally their applicability to this thesis.

2.2.2 Historical Evolution and Credibility

The discipline of Value Investment begins with Benjamin Graham, commonly referred to within the field of Finance as the "Father of Value Investment". Graham began teaching his approach to investment at Columbia University in 1928, and published his first book, Security Analysis, in 1934. This book defined the framework of Value Investment, and is now in its 5^{th} Edition.

The principle of Value Investment lies with the discovery of a securities 'intrinsic value', through fundamental analysis techniques. The study of a securities financial and accounting history is a prime driver for determining a securities intrinsic value. Unlike proponents of the Efficient Market Hypothesis (EMH), the value investor believes that the market does not accurately price securities. According to Cottle et al. (1988), the value investor believes the market price of a security is driven by a number of irrational processes, and the market value of a security only occasionally coincides with the securities 'true' (intrinsic) value. Lowe (1996) describes Value Investors as seeking to determine the intrinsic value of securities, and actively seek out those securities whose market value is significantly less than the securities intrinsic value.

Graham's value investment philosophy is well entrenched, in part due to the success of Warren Buffett, who is widely recognized as the world's greatest 20th century investor. Buffett credits his success to Graham, however, according to Bierig (2000), rather than just seeing a balance sheet as a frozen snapshot of a company, Buffett broadens his definition of value, and investigates the dynamics of the company. In this sense, Buffet has become subjective rather than objective.

There are many credible definitions of intrinsic value, and hence, of a value investment. According to Cottle et al. (1988), Graham defined intrinsic value as 'the value which is justified by assets, earnings, dividends, definite prospects, and the factor of management'. Buffett (1996), defines intrinsic value as '…the discounted value of the cash that can be taken out of a business during its remaining life'.

According to Lowe (1996), Graham regularly emphasized three main concepts in teaching Value Investing. These were:

- the 'right' attitude,
- the importance of a margin of safety, and
- Intrinsic Value.

Graham tried to define the 'right' attitude by stating that an investor should seek an investment, as opposed to speculating, and defined an investment as a proposition that offers both safety of capital and a reasonable expectation of a return, and anything else as speculating.

Graham's book, The Intelligent Investor was initially published in 1949, and was last published in its 4th edition. In each edition, Graham defined the manner in which a 'defensive' and an 'enterprising' value investor should build a portfolio. A defensive investor was considered to be an investor without the time and inclination to consider investing a full-time business, an 'enterprising' investor was the opposite. The instructions Graham gave changed slightly in each version of the book, as it was updated to current market conditions. In the last edition of The Intelligent Investor (1973), Graham stated a defensive investor should pay attention to three main factors, namely,

- size of firm,
- capitalization, and
- price-earnings ratio.

A defensive investor should select stocks that:

- had at least $50 million in assets or annual sales,
- be in the upper quarter or third of its industry in size,
- the equity (at book value) should be at least 50% of total capitalization for industrial companies, at least 30% of total capitalization for utilities
- the price should not exceed 25 times average earnings of past seven years, and not to exceed 20 times earnings of the latest 12 month period.

Oppenheimer and Schlarbaum (1981) tested Graham's defensive portfolio strategy to determine its usefulness. They extracted the rules provided to investors in each of the four editions of The Intelligent Investor, and using publicly available stock information, found that positive risk-adjusted rates of return were delivered to defensive investors following Graham's criteria from 1956 to 1975. Rates of return were 3% – 3.5% higher than was achieved using a buy-and-hold strategy (in a frictionless market). When the various market frictions were taken into account, the typical defensive investor achieved rates of return 2% - 2.5% higher than the buy-and-hold strategy. The small firm size effect (see Banz (1981), and Reinganum (1981)) is not considered as a suitable bias for these results, as the defensive strategy called for buying stocks only in the top third or quarter of their industry (in terms of size). The authors state that '…it is reasonable to conclude that our evidence contradicts the semi-strong form of the efficient markets hypothesis'.

According to Lowe (1994), Benjamin Graham also published ten attributes of an undervalued stock, that investors could use to locate undervalued companies. These were:

- earnings-to-price yield double the AAA bond yield,
- P/E four-tenths highest average P/E in most recent 5 years,
- dividend yield two-thirds the AAA bond yield,
- price two-thirds tangible book value per share,
- price two-thirds NCAV (Net Current Asset Value),
- total debt less than tangible book value,
- current ratio greater than or equal to 2,
- total debt less than or equal to net quick liquidation value,
- earnings doubled in most recent 10 years, and
- no more than two declines in earnings of 5 percent or more in the past 10 years.

It was noted that few companies could meet all 10 criteria.

Many value investment textbooks publish similar lists of Grahams ten points. For example, the ten points listed as Grahams by Brandes (1989) are similar, yet not identical. Brandes replaces 'total debt less than or equal to net quick liquidation value' by 'total debt should be less than twice net current assets', and 'earnings doubled in most recent 10 years' by 'earnings growth should have been at least 7 percent per annum compounded over the previous decade'. The remaining eight points are identical to the list published by Lowe. Brandes also stated that Graham required a stock to meet only one of the first 5 tests, and only one of the last 5 tests.

According to Lowe (1996), in 1975, Benjamin Graham published his ideas on a finely tuned portfolio, giving both stock selection rules, and basic portfolio management rules. For portfolio rules, Graham suggested:

- a portfolio size of about 30 stocks, and
- a target of 50% profit obtained within 2 years. Issues that didn't appreciate within the 2 year timeframe should be sold out at their market price.

The 30 stocks were to be selected by applying the following rules:

- earnings-to-price ratio of twice the last 12 months yield on AAA bonds,
- attractive dividend yield,
- price below book value,
- price well below previous 2 year high,
- P/E lower than the 7-to-10 year average P/E.

Following on from Grahams success other researchers tried to identify better and more reliable ways to determine if a security was undervalued, or would yield a return disproportionate to its cost and risk. Much of this work was centered on Benjamin Graham's ideas as described above. It also focused on the detection of stock pricing anomalies, that is, cases where investment returns were exaggerated for some particular value-based (fundamental) reason.

A large number of anomalies were detected and documented, and, in many cases, provide support for some of the stock selection 'rules' initially laid down many years earlier by Graham. The remainder of this section on Value Investment considers some of this research, and presents it according to the specific fundamental factors considered.

Basu (1977) investigated whether stocks with low P/E ratios earned excess returns when compared to stocks with high P/E ratios. It was found that during the study period (April 1957 – March 1971), portfolios built from low P/E stocks earned higher returns than those portfolios built from higher P/E stocks, even after adjusting returns for risk. The study concluded that there is an information content present in publicly available P/E ratios, which could offer opportunities for investors, and that this was inconsistent with the semi-strong from of the EMH. There are some clear parallels with the first two guidelines of Graham's 10-point list here. The first guideline suggested earnings-to-price yield be double the AAA bond yield. The earnings-to-price yield is the inverse of the P/E ratio, and ensuring it is greater than the AAA bond yield effectively capped the P/E ratio. In this manner, it steered investors away from high P/E stocks. The second guideline required P/E be four-tenths highest average P/E in most recent 5 years, again effectively steering the investor away from high P/E stocks.

In 1981, Banz (1981) focused on the 'size effect'. Essentially, the size effect concerns the relationship between the market capitalization of a firm, and its return. Banz reports that during the study period (1936 – 1975), common stock of small firms had higher returns than the common stock of large firms, even after adjusting for risk. Banz also raises the issue that the size effect may just be a proxy for one or more other factors, which are correlated with size, an interpretation he also applies to Basu's findings concerning the P/E effect.

Also in 1981, Reinganum (1981) described a misspecification of the simple one-period CAPM model, namely, that data on firm size can be used to create portfolios that earn abnormal returns. From studying small firms listed on the New York and American Stock Exchanges, during the period from 1963 to 1977, Reinganum discovered average rates of return for small firms to be nearly 20% per year greater than those of large firms.

In 1984, Rosenberg et al. (1984) presented two strategies aimed at exploiting fundamental information to increase returns. The first, the "book/price" strategy buys stocks with a high ratio of book value to market price, and sells stocks with the reverse. The second strategy, "specific return reversal" computes specific returns per stock, and relies on the observation that specific returns tend to reverse in the subsequent month. Thus, this strategy buys stocks with negative specific returns in the preceding month, exploiting this reversal. The study sourced data from Compustat, on 1400 of the largest companies, from 1980 to 1984, and stocks were priced mainly from the NYSE. The study demonstrated statistically significant results of abnormal performance for both strategies, and suggests that prices on the NYSE are inefficient. Here, the first strategy provides support for Graham's fourth guideline, namely that price be two-thirds tangible book value per share, effectively steering the investor toward stocks with a higher book value than price.

DeBondt and Thayler (1987) present evidence that investors tend to overreact when considering recent data. This overreaction led to a reversal effect, with stocks that had been prior 'losers' likely to become future 'winners'. The researchers also investigate seasonality patterns in returns data. They demonstrate that the winner-loser effect is not primarily a size effect, and the earnings of 'winner' firms and 'loser' firms show reversal patterns consistent with overreaction. In terms of seasonal influence, DeBondt and Thayler report that excess returns for 'losers' are negatively related to both long-term and short-term formation performance, particularly in January. For 'winners', they find that January excess returns are negatively related to the excess returns for the prior December.

Detailed research from Fama and French (1992) surveys the above style of anomaly detection, and conclude that if asset-pricing is rational, then size and the ratio of book value of a stock to its market value must be proxies for risk, as opposed to reflecting market inefficiency.

Lakonishok et al (1994) find that a wide range of value strategies (based on sales growth, Book-to-market, Cash flow, earnings, etc) have produced higher returns, and refute Fama and French's claims that these value strategies are fundamentally riskier. Using data from end-April 1963 to end-April 1990, for the NYSE and AMEX, Lakonishok et al find evidence that the market appears to have consistently overestimated future growth rates for glamour stocks relative to value stocks, and that the reward for fundamental risk does not explain the 10% - 11% higher average returns on value stocks. This study lends further support for Grahams fourth guideline, again effectively steering the investor toward stocks with a higher book value than price.

Fama and French (1995) respond to Lakonishok et al by focusing on size and book-to-value, and form portfolios of stocks partitioned by these variables from the NYSE, AMEX and NASDAQ, from 1963 to 1992. Their results demonstrate that both size and BE/ME (book-to-market equity) are related to profitability, but find no evidence that returns respond to the book-to-market factor in earnings. They conclude that size and BE/ME are proxies for sensitivity to risk factors in returns. Their results also suggest that there is a size factor in fundamentals that might lead to a size-related factor in returns.

Later, Fama and French (1998) study returns on market, value and growth portfolios for the US and twelve major EAFE countries (Europe, Australia, and the Far East). They recognize that value stocks tend to have higher returns than growth stocks, finding a difference between low B/M (Book-to-market) stocks and high B/M stocks of 7.68% per year on average. They find similar value premiums when investigating earnings/price, cash flow/price and dividend/price. They find that value stocks outperform growth stocks in twelve of thirteen major markets during 1975 – 1995. They also find a value premium in emerging markets. Fama and French conclude that these results are explained by a one-state-variable ICAPM (or a two-factor APT) that explains returns with the global market return and a risk factor for relative distress.

Frankel and Lee (1998) estimate firms fundamental values (V) using I/B/E/S concensus forecasts and a residual income model. They find that V is highly correlated with stock price, and that the V/P ratio is a good predictor of long-term returns. They state that this effect is not explained by a firm's market beta, B/P ratio, or total market capitalization (size). They also find evidence that errors in consensus analysts forecasts are predictable, and these prediction errors can be exploited by incorporating the error with V/P. They conclude that the evidence suggests that firm's value estimates may well provide a better forecast ability than simply using ratios, and that prices converge to value estimates gradually over greater than 12 month horizons. They also state that the predictability of long-term forecast errors in consensus forecasts is consistent with a long-term mispricing hypothesis. Finally, the authors also acknowledge that the results may demonstrate yet another proxy for cross-sectional risk differences, but state that this is an unlikely conclusion.

Reinganum (1988) studied the fundamental and technical characteristics of top performing US stocks during the period 1970 – 1983. Of the fundamental ratios studied, Reinganum found that a price/book ratio of less than 1 was a common characteristic of the winners, and suggested this was an important aspect of an investment strategy. He also discovered that very low P/E ratios, and prices, were not a necessary ingredient of a successful strategy, and that low market capitalization was not important either. In summary, Reinganum singles out 9 features of stocks set for explosive price change. These are:

- The price/book ratio is less than 1
- The five-year growth rate based on quarterly earnings is positive
- Quarterly earnings are accelerating, that is, there is a positive change in the percentage change in quarterly earnings
- Pre-tax profit margins are positive
- There are fewer than 20 million common shares outstanding
- The relative strength rank of the stock is at least 70
- The relative strength of the stock in the current quarter is greater than the rank in the previous quarter
- The O'Neil Datagraph rating is at least 70
- The stock is selling within 15% of its maximum price during the previous two years

Philips (1999; 2002) attempts to develop simple analytical expressions for the expected return and fair value of an equity market, as a function of (amongst other things) fundamental valuation ratios.

Piotroski (2000) investigates whether fundamental analysis can be used to provide abnormal returns, and right shift the returns spectrum earned by a value investor. In anomaly terms, Piotroski focused on high book-to-market securities, and shows that the mean return earned by a high book-to-market investor can be shifted to the right by at least 7.5% annually, and a simple investment strategy based on high book-to-market securities generates a 23% annual return between 1976 and 1996. The research is stimulated by the observation that portfolios of high book-to-market firms normally contain several strong performing firms (achieving strong returns), and many deteriorating ones (achieving poor returns). Piotroski defines three different classes of financial performance signals, namely:

- Profitability,
- Leverage, Liquidity and source of funds, and,
- Operating Efficiency.

From these three classes of signals, nine simple signals are defined, and an aggregate score of the nine signals is used to rank the constituents. The nine signals involve seven fundamental variables, namely:

- net income before extraordinary items,
- cash flow from operations, (both scaled by the beginning of year total assets),
- leverage,
- liquidity,
- whether external financing has been raised recently,
- current gross margin scaled by total sales, and
- current year asset turnover ratio.

Within the portfolios constructed from the higher aggregates, Piotroski notes that the returns are concentrated in small and medium sized companies, companies with low share turnover, and firms with low analyst following. It is also noted that superior performance is not dependant on initial low share prices. Again, support is found for Graham's fourth guideline in this study. Of further interest is the determination that one-sixth of the annual return difference between the ex-ante strong and weak firms is earned over the four three-day periods surrounding earning announcements. This information is of obvious interest to those advocating market timing approaches.

Kanas (2001) finds a non-linear relation between stock returns and the fundamental variables of dividends and trading volume. Regarding dividends, there has long been speculation that changes in dividends signal future changes in earnings. Early work in this area comes from Litzenberger and Ramaswamy (1979), who discovered a strong positive relationship between dividend yield and expected return for NYSE stocks. This idea is investigated by Benartzi et al. (1997) who find limited support for it. They conclude, however, that in spite of the lack of future earnings growth, firms that increase dividends have significant (though modest) positive excess returns for the next 3 years.

On a final note regarding dividends, there are also a number of artificial relationships maintained in the market. For example, the Dow Dividend Theory is a contrarian investment strategy, which relies on the fact that a reduction in dividend payments will be seen as a negative sign by investors, particularly if the company reducing the dividend payout is in the DOW-30 (the top 30 stocks in the DJIA). For this reason, these firms attempt to maintain dividend stability, as temporary dips in the stock price will raise the dividend yield. The original "Dogs of the Dow" strategy sorted the 30 stocks by dividend yield at the beginning of the year, and then created a portfolio from the 10 stocks with the highest yield. Interestingly enough, this strategy became the basis of "The Motley Fools" popularity. There is much debate regarding the economic superiority of this strategy, and the timeframe in which benefits appear to accrue, and a number of variants have been proposed, such as the Dow-10 investment strategy, by McQueen et al (1997).

Aby et al. (2001) focus on combining fundamental variables to screen stocks for value. This is a reasonably common approach, with some authors reporting outstanding results. Aby et al. developed portfolios based on four fundamental conditions, namely: Single Valued P/E (P/E<10), Market Price < Book Value, established track record of return on Shareholder Equity (ROE > 12%), and dividends paid out less than 25% of earnings. They conclude that when the four criteria are used to screen stocks, quality investments seem to result, again providing support for Graham's guidelines. The authors state that higher yields do not seem to provide good long term returns, possibly due to the use of retained earnings to enhance equity per share. Overall, the main contribution of this work is to establish a relationship between ROE (> 12), and share price performance. The research alludes to the fact that Buffett believes 12 is an appropriate value for ROE in (US) domestic markets. The authors find that the value of 12 for ROE provides a clear line of demarcation between performance and non-performance is share price terms. The authors tested the filter criteria against the Value Line database between August 31, 1989 to August 31, 1999. The filter conditions described cut the database down from 6000 possible stocks to just 14. These 14 yielded an average return of 30.55% per year for the ten years. It is interesting to note that in earlier work (2001), the same authors had focused on shares with simply a low P/E and a market price below book value, and had concluded that this filter method did not produce satisfactory returns.

Olson and Mossman (2002) compared ANN forecasts with forecasts developed by OLS (ordinary least squares) and logistic regression (logit) techniques. They attempted to forecast one year ahead Canadian stock returns, using 61 accounting ratios for 2352 Canadian companies over the period 1976 – 1993. Olson and Mossman conclude that back propagation neural networks outperform the best regression alternatives for both point estimation and classification for high and low returns. Further, they find that the superiority of the ANN models translates into greater trading profitability, and conclude that fundamental analysis adds value to abnormal return trading strategies within the Canadian market.

This portion of the literature review has presented key developments and concepts in the area of Value Investment, from tracing the seminal work by Graham, through to the latest research in Value based investment. It has also tried to link together some of the guidelines produced by Graham, and affectionately known by Value Investors as “the ten points”, with later related supporting research.

Finally, some interesting thoughts regarding the long-term future of Value Investment are provided by Siegel (2000), interesting because the article appeared shortly before the “tech-wreck”. Siegel encourages value investors to continue with value investment, and perhaps consider widening their definitions of value. Within a very short period of time, the “tech-wreck” ensued and value investors were more than compensated for their tenacity.

2.2.3 Applicability

Essentially, the discussions above reveal that there are many credible variations of what may constitute a value investment. Further, as stated by Haugen (1999), a stock is only a value stock at a point in time. That is, it is not the stock itself; it is the conditions that the stock happens to be trading in that cause it to be considered as a value stock. This view is supported by Reinganum (1988), who suggests ‘there may be more than one way to skin the performance cat’, implying that a variety of valid value strategies can be composed to outperform their respective indices.

From the point of view of this thesis, the works reviewed above clearly demonstrate a number of persistent anomolies within the pricing mechanisms of the stockmarket. The anomalies documented and described above concern themselves specifically with fundamental variables. As such, the use of fundamental variables as inputs for ANNs are well within the remit of this study.

2.3 Technical Analysis

2.3.1 Introduction

As mentioned earlier, the key historical developments in Technical Analysis will be initially reviewed, followed by a discussion of their academic credibility, and finally their applicability to this thesis.

2.3.2 Historical Evolution and Credibility

According to Edwards et al. (2001), modern Technical Analysis begins with the work of Charles Dow, who in 1884 drew up an average of the daily closing prices of 11 important stocks. Dow published a series of articles in the Wall Street Journal between 1900 and 1902, documenting stock price movements he had observed in the averages. These articles were the first to describe systematic phenomena in the stock markets.

Although Dow's work represents the beginning of modern technical analysis, it is worthy of note that markets and analysis techniques existed centuries before this, notably in Japan since 1730, where the first futures contracts (in rice) were traded. Tvede (1999) reports that interest in the future prices of the 'futures' ran high, with the Japanese government suspending the forward market in 1869 due to excessive volatility.

Returning to modern technical analysis, Dow saw these averages he had devised as being representative of the general business economy, and envisioned the averages as a way of predicting future business conditions. Dow's reasoning for this was based on the fact that most investors at that time were intimately acquainted with the current industrial situation, and their involvement with the stock market represented their preparedness to wager on their future facts, hopes and fears. In such a system, Dow reasoned that the price fluctuations within the averages represented the combined facts, hopes, and fears of all interest parties, a kind of barometer describing the combined appraisals of all participants.

Dow's work was covered and updated by S. A. Nelson in The ABC of Stock Speculation, published in 1903, and updated further in 1922 when William Peter Hamilton published The Stock Market Barometer. Finally, in 1932 Robert Rhea published the book, Dow Theory, which has become accepted as the definitive version of Dow's original theories.

Richard W Schabacker continued investigating patterns described in Dows work during the 1920s and 1930s, reasoning that the patterns described by Dow in the averages, must to some extent be present in the individual stocks that composed those averages. According to Edwards et al. (2001), Schabacker showed that the signals present in the averages were indeed present in the constituent stocks, and published his work in three books, Stock Market Theory and Practice, Technical Market Analysis, and Stock Market Profits.

Later in life, Schabacker was joined by Robert D Edwards and in 1942 Edwards was joined by John Magee. The first thorough descriptions of the many patterns which Dow described in his averages (and indeed, in individual stock records) was published by Edwards and Magee in 1948, in the book Technical Analysis of Stock Trends. This book is still in print, currently in its eighth edition.

Many people regards Wilders (see 1978) contributions as the most important contemporary work in the field of Technical Analysis, introducing an array of new technical indicators, and a variety of new trading system techniques. Many of the modern technical indicators in use today are directly based on Wilder's work.

In a final note on Dows Theory, Brown et al. (1998) studied William Peter Hamilton's track record, and concluded that Hamilton's timing strategies yielded high Sharpe Ratios and positive alphas for the period 1902 – 1929.

Today, a manual of technical analysis is likely to be composed of techniques relating to one of three primary classifications, namely:

- Charting (mainly pattern matching),
- Indicators, and
- Esoteric approaches.

This paper will focus on the use of technical indicators within technical analysis, and to a smaller extent, on pattern matching approaches. The primary purpose for this is that reading charts and pattern matching are generally practiced inconsistently, even between knowledgeable analysts. The patterns seen in charts are often highly subjective, and without rigorous definition.

Finally, several esoteric approaches are discussed in a variety of technical analysis literature; these techniques are excluded from this paper as having no scientific justification. Warnecke (1987) provides example of some of the criticisms often leveled at esoteric approaches, such as the Elliot Wave Theory (for a description of Elliot Wave Theory, the reader is encouraged to refer to Prechter (1995)). Additional esoteric approaches concern relationships between the length of womens skirts and stock market price movements (named the 'hemline indicator'), and the 'Super-Bowl indicator', which states that if a team from the pre-merger National Football League beats a team from the old American Football League, the stockmarket will advance in the following year. Interestingly, Strong (1988) reports this indicator has been correct in 19 of the 21 years the Super Bowl had been played.

The main principles of Technical Analysis dictate that:
- Prices move in trends,
- Volume goes with the trends,
- A trend, once established tends to persist.

These primary principles are consistent with a variety of authors of Technical Analysis manuals and papers, yet the actual definition and premises of Technical Analysis seems quite variable. Pring (1999) describes the art of technical analysis as '…to try to identify trend changes at an early stage and maintain an investment or trading posture until the weight of evidence shows or proves that the trend has reversed'. Rotella (1992) states 'Technical analysis is the study of past market behaviour to determine the current state or condition of the market'. Other definitions provide coverage of market information in terms of market variables, such as this definition provided by the Australian Securities Institute (2003), 'The study of behaviour of market participants, as reflected in price, volume and open interest for a financial market, in order to identify stages in the development of price trends'.

There is a considerable base of evidence supporting Technical Analysis, and an equally considerable base of evidence opposing Technical Analysis. The remainder of this section reviews key papers which present evidence both for and against Technical Analysis.

Fama (1965) presents an appropriate starting point for this work. This paper presents a considerable amount of evidence which supports the random-walk hypothesis. In essence, this states that successive changes in stock prices are independent, identically distributed random variables. The most important implication of this hypothesis is that this implies a series of price changes has no memory, which further implies that the study of past prices cannot provide a useful contribution to predicting future prices. The natural implication here is that studying chart patterns (a major area in technical analysis) is of no value to the stockmarket investor. Fama's research also provides support for the Mandelbrot hypothesis, which states that empirical distributions of price changes conform better to stable Paretian distributions (with characteristic exponents less than 2) than to the normal distribution.

In direct contrast to the views expressed by Fama, Wilder (1978) describes an array of new trading systems and techniques, all of which are based on historic stock data. Wilder's book is aimed at the practicing Technical Analyst, and does not attempt to persuade non technical analysts to use the strategies, thus, it does not enter the debate on whether technical analysis is a suitable tool for trading stocks.

Kamara (1982) tests the random character of futures prices, and does find some degree of serial correlation. However, whilst some dependencies were detected, these were distinctly related to the sample periods chosen. Further, it was not clear whether sufficient dependencies existed to enable an investor to generate sufficient profits to make the pursuit of these dependencies viable. Also of note, Kamara found that a distribution of futures prices was better approximated by a mixture of two normal distributions, as opposed to a single normal distribution.

LeBaron (1983) documents and describes six market inefficiencies, and considers that these inefficiencies exist due to the way many agents (fund managers, etc) need to act prudently in terms of their investment of clients funds. LeBaron argues that this is the main reason why value effects don't quickly get arbitraged away.

The effectiveness of chartists in technical analysis is considered by Neftci and Policano (1984), who investigate two important charting mechanisms, namely, slopes (essentially trendlines), and moving averages. The authors study used closing prices for various gold and T-Bill futures contracts between January 22, 1975 and March 6, 1980 (for gold), and January 12, 1976 and July 3, 1980 (for T-Bills futures). Using the two techniques, and a set of buy-and-sell rules based on them, Nefti and Policano conduct tests for market efficiency. For the moving average method, they find a significant relationship between moving average signals and futures prices. The results for the slope method using trendlines are mixed, and are probably best described as inconclusive. Of interest is the fact that a significant set of parameters for one commodity were often insignificant for another commodity, perhaps indicating specifics of particular commodities related to hedging versus speculation, or the thinness of that particular market. Others cite this finding as evidence that markets have their own 'personalities'.

Murphy (1986) tests the effectiveness of Technical Analysis by examining the performance of publicly offered futures funds. Data was sourced from all (US) funds which employ only technical trading strategies, listed between May 1980 and April 1985, in the first 'Funds Review' section of Commodities Magazine. The study shows returns from technical funds to be inferior to base stockmarket returns, and also inferior to the T-Bill market, over the sample interval. Specifically, Murphy finds:

No evidence of abnormal returns from technical trading

No evidence that technical funds can outperform a naïve buy-and-hold strategy

Murphy concludes that the findings are consistent with the EMH, and that the futures market is technically efficient.

Laderman (1987) provides mainly anecdotal evidence of the effectiveness of technical analysis, and states that a large degree of interest in technical analysis is due to correct predictions of some key technicians, during periods of market corrections. Laderman also documents an increased reliance and respectability for technical analysis due to the growing relationship between the stock markets, and the futures (and options) market.

Murphy (1988) demonstrates that different sectors of the market move in relationships with other sectors, and discusses the ways in which Technical Analysis can be used to analyse related markets, a field now known as Intermarket Analysis.

A detailed study of the attitudes of Finance PhDs and other investment professionals was conducted by Strong (1988) . The three controversial issues addressed were:

- Informational efficiency of the markets,
- The value of technical analysis, and,
- The importance of risk-adjusted performance measurement

Concerning market efficiency, Strong finds substantial agreement between academics and non-academics for the '*semi*efficient market hypothesis' (as opposed to a single efficient securities universe), and both groups supported the idea of buying low-priced, low P/E stocks. In the area of technical analysis, many Finance PhDs believed that technical analysis should be followed, with over 40% believing in the effectiveness of the advance/decline lines. In the area of performance measurement, a large number of finance PhDs questioned the usefulness of betas, and 'risk-free rates'. Overall, a significant number of investment professionals supported the use of technical analysis.

White (1988) considers the predictions of neural networks for prices of IBM common stock daily returns over an extended period. White does not find any evidence that contradicts the efficient markets hypothesis. White notes that the neural methods of back-propagation effectively reduce training error, whilst the objective in searching for evidence to contradict the EMH would be found by optimizing profit, an opportunity not directly afforded by using neural networks.

Lehmann (1990) considers evidence supporting variation in equity returns, attempting to decide whether the evidence is indicative of predictable changes in expected return, or market inefficiency. Lehmann finds that 'winners' and 'losers' one week often experience reversals of fortune in the following week. The costless portfolio constructed by Lehmann (difference between 'winner' and 'loser' portfolios) showed profit in 90% of weeks. Lehmann concludes that the reversals of fortune are probably reflections of the imbalances in the market for short-term liquidity, and states that 'it is difficult to account for these results within the efficient markets framework'. Lehmann's work is often quoted by practitioners as supporting Technical Analysis, as it supports the idea that price trends occur frequently enough to create profit opportunities for technical traders. Lehmann does not specifically make this statement.

Jegadeesh (1990) examines the predictability of monthly returns on individual securities. Ten portfolios were formed based on the predicted returns using estimates of the regression parameters. The difference between abnormal returns on the extreme decile portfolios was 2.49 percent per month over the period 1934 to 1987. Slightly different values are provided when comparing extreme decile portfolios excluding January results (2.20% per month), and when January was considered separately (4.37% per month).

Jegadeesh rejects the random walk hypothesis, and concludes that returns predictability is due to either market inefficiency, or systematic changes in expected stock returns. This paper is often used to support the principles of technical analysts, as it shows evidence that increases (and decreases) in prices during one month are often reversed out the following month. Patterns of that nature would suggest that investors could profit from technical trading strategies, and would also be a breach of market efficiency.

An investigation of the statistical properties of Technical Analysis is provided by Neftci (1991), in an attempt to determine if there is an objective basis to the popularity of technical analysis. The research attempts to formalize technical analysis rules, on the basis of Markov times, which essentially state that no dependence can be placed on future information when deriving predictions. Neftci examines the relationship of the 150-day moving average to the Dow-Jones index, to evaluate commonly (technically) accepted wisdom regarding Dow-Jones trends. The research finds that the moving average does generate Markov times, and does seem to have some predictive value. The work tested the Dow-Jones industrials for 1911 - 1976. Also, the work reviews several other technical analysis techniques, and finds many do not generate Markov times, hence are effectively relying on 'seeing the future' on a chart, before being able to generate a forecast (for example, according to Neftci, the 'head-and-shoulders' pattern and the trendline crossing do not generate Markov times).

Support for trading rules based on simple moving averages is provided by LeBaron (1997), who uses such rules as specification tests on the process for foreign exchange rates. LeBaron concludes that the exchange rates studied do not follow a random walk, and that the deviations are detected by simple moving average rules.

In the late 80's, full acceptance of Technical Analysis by the academic community was still quite low, so Taylor and Allen (1992) were asked to conduct a survey on behalf of the Bank of England, in November 1988, regarding the acceptance of Technical Analysis by chief foreign exchange dealers in London. Among other findings (reported later in this paper), they found that at least 90% of respondents placed some weight on Technical Analysis, with a skew towards greater acceptance at shorter time horizons. Clearly, Technical Analysis had a much greater acceptance amongst actual practitioners than academics were prepared to accept.

Two popular technical trading rules are tested by Brock et al. (1992), namely Moving Averages, and trading range breaks (Support and Resistance breaks). Using data from the Dow Jones Industrial Average (DJIA) from the first trading day in 1897 to the last trading day in 1986, the authors test combinations of moving averages (and moving average strategies involving fixed and variable length holding periods), as well as the trendline strategy. In all cases, the authors also evaluate the use of a one percent band around the predictions to eliminate some whipsaw action (whipsaw is the tendency of prices to oscillate back and forward across a boundary line). Their findings provide support for the use of technical analysis, in particular to the moving average strategy using a one percent band. The authors also find that buy (sell) signals generate returns that are higher (lower) than ‘normal’ returns, and that the differences are not readily explained by risk. Finally, the authors conclude that their results are consistent with the technical rules having predictive power.

Levich and Thomas (1993) test currency futures contracts in five currencies (British Pound (BP), Canadian Dollar (CD), German Mark (DM), Japanese Yen (JY), and Swiss Franc (SF)) for the period 1976 to 1990 (approximately 3800 daily observations), testing technical trading rules using a bootstrap approach. Their research shows persistent trading profits over the 15 year period using simple filter rules and a variety of commonly researched moving averages. Levich and Thomas conclude ‘the profitability of trend following rules strongly suggest some form of serial dependency in the data, but the nature of that dependency remains unclear’.

Osler and Chang (1995) comprehensively tested the technical analysis pattern known as 'head and shoulders' using daily data from March 1973 to June 1994 for 6 major currencies versus the US dollar. They find that the pattern appears to have predictive power for some currencies, and not others. Nevertheless, they conclude that if 'one had speculated in all six currencies simultaneously, profits would have been statistically and economically significant'.

Neely et al. (1997) use Genetic Algorithms to find effective trading rules, investigating six exchange rates during the period 1981 – 1995. They find that the rules 'discovered' are similar to those in use by technical traders, and the rules produce economically significant out-of-sample returns. The authors find no evidence that the excess returns are attributable to risk, and by the use of bootstrapping procedures, they conclude that the trading rules detect patterns in the data that are not captured by standard statistical models. The authors conclude that they view the results as plausible evidence of market inefficiency.

Inspired by Brock et al. (1992) earlier test of two trading rules in the DJIA, Mills (1997) tests the same two trading rules as Brock, this time in the London Stock Exchange against FT30 index data for the period 1935 – 1994. Mills' results are remarkably similar to Brocks, with Mills coming to the conclusion that trading rules can predict stock prices, and are thus profitable, only in periods when the market is inefficient.

Pruitt and White (1998) demonstrate the effectiveness of the CRISMA trading system, a system based on Cumulative Volume, Relative Strength, and Moving Averages. This system called for traders to buy and sell exchange-listed call options on identified equities, in line with a number of predefined rules. The study consisted of 171 firms over the period from 1976 to 1985, and assumes the purchase of the second shortest maturity option available. The authors demonstrate that even in the presence of maximum 1988 transaction costs, the mean returns to the system were 12.05% per round-trip trade.

A later update on the effectiveness of the CRISMA trading system is provided by Goodacre and Kohn-Speyer (2001), who re-examine the system using US data from the period 1988-1996. They found the performance of the system to be unstable over time, and that the system was of no benefit once market movements, risk, and transaction costs had been taken into account. Finally, the authors conclude the results are consistent with market efficiency.

Lee and Swaminathan (2000) investigate both price momentum and trading volume. They find past trading volume predicts the magnitude and persistence of price momentum. They conclude that past volume helps to reconcile intermediate-horizon underreaction and long-horizon over-reaction effects.

Su and Huang (2003) use combinations of technical indicators (Moving Average, Stochastic Line [KD], Moving average Convergence and Divergence [MACD], Relative Strength Index [RSI] and Moving average of Exchanged Volume [EMA]) to determine trend direction with good results.

Demir et al (2004) study returns to momentum strategies in the Australian equity market. They find that momentum is prevalent in the Australian market, much more so than in overseas markets. The find that momentum strategy returns are robust and prevail over time, and that the profits observed are not explained by size or liquidity.

On a final note regarding the legitimacy of technical analysis, it is occasionally stated that as technical rules become more widely known, the abnormal returns they attempt to identify will be reduced, and the usefulness of the technical rule itself will be destroyed. Silber (1993) finds against this conclusion, instead concluding that 'the continued success of simple technical trading rules is possible as long as there are price smoothing participants in the market'. In this context, Silber's example of price smoothing participants refers to the central banks.

In addition to the above formal sources, a brief literature review was conducted throughout the main practitioners' journal, The Technical Analysis of Stocks and

Commodities. Although some of the articles published in this source were not of academic quality, this search aids in identifying technical variables that are being used by practitioners in the field. These variables, along with those identified in the research above, will later be used to determine the inputs to the technical ANN. It is comforting to notice that many of the ideas emanating from academia are slowly finding their way into the realm of the trader, and vice-versa.

Table 2-1 briefly shows the results of the literature review of the practitioners' journal, by characterizing many of the trading articles by the main technical variable upon which they were based. Not all articles that were reviewed are described below; instead, those articles presented in Table 2-1 are a representative summary of the articles reviewed. In this way, it is possible to determine which technical variables are most in use by practitioners, with the assumption that they are most in use due to the fact that they are useful. A discussion regarding the construction of these technical variables is provided later is section 3.7.1.4.

Technical Variable Classification	Brief Description and Reference
Moving Averages (including simple, and a variety of manipulations based on moving averages)	Reverre (2000) enhances an existing strategy (Bollinger Bands) by using moving averages; Sharp (2000) attempts to remove much of the lag associated with moving averages; Ehlers (2000) also attempts to reduce lag in moving averages, using Hilbert transforms; Pring (2000) describes using simple moving averages to generate signals in a trading strategy; Fries (2001) describes the use and enhancements of elastic moving averages; Ehlers (2001) enhances moving averages by combining maximum entropy spectral analysis and the Hilbert transform; Dormeier (2001) investigates the combination of moving averages and volume; Boomers (2001) discusses trading based on changing the length of moving averages; Schaap (2004) discusses trading using 50 day moving averages; Yoder (2002) describes techniques based on moving averages for trend detection;
Volatility	Levey (2000) investigates the use of volatility in enhancing trading returns; Gustafson (2001) describes the use of volatility based on ATR (Average True Range); Pezzutti (2002) discusses designing trading systems based on breakouts in volatility;
Volume (including derivatives such as On-Balance Volume)	Pring (2000) describes how to use volume to confirm trading signals; Ehrlich (2000) describes On-Balance Volume and its use in combination with a variety of oscillators; Tanksley (2000) describes trading using average price weighted by volume, a technique later re-interpreted

Technical Variable Classification	**Brief Description and Reference**
	by Reyna (2001); Peterson (2001) discusses the role of volume in making assessments of market breadth; Katsanos (2004) gives an overview of trading performance using the flow of volume indicator; Bulkowski (2004) considers the implication of volume on price breakout strategies; Katsanos (2004) discusses trend detection in light of the volume flow indicator; Castleman (2003) describes the use of volatility and volume in adapting trend following systems to changing market conditions; Peterson (2003) describes trading based on consideration of volume change; Gimelfarb (2004) describes using volume to detect changes in trend control; Davies (2004) discusses the use of volume in daytrading systems; Fell (2004) explains the usage of volume weighted moving averages; Ord (2004) explains the significance of volume in price movements
ADX (Average Directional Index, originally created by Welles Wilder)	Boot (2000) describes a simple trading systems based on ADX; Star (2003) explains how ADX is used by a number of leading market technicians; Gujral (2004) describes trading systems based on the ADX indicator
Stochastics	Steckler (2000) and Steckler (2004) describe trading strategies built on the stochastic indicator

Technical Variable Classification	**Brief Description and Reference**
Momentum	Kaeppel (2001) describes trading of sector funds based on momentum; Pring (2001) describes trading strategies based on momentum; Pring (2001) describes further strategies based on trading the divergence between momentum and price; Pring (2001) considers trading the trends that form in charts of momentum; Penn (2002) describes the usage of TRIX, a volume based indicator; Roffey (2002) discusses the relationships between trends and momentum indicators;
RSI (Relative Strength, and associated derivatives)	Pring (2001) discusses strategies based on relative strength; Ehlers (2002) discusses the use of RSI; Siligardos (2003) descirbes using RSI for price projections; Watkins (2003) describes trading strategies based on relative strength;
Variety of other Miscellaneous indicators (MACD (Moving Average Convergence/Divergence), Intermarket, Money Flow, TRIN (TRaders INdex)	O'Brian (2000) demonstrates the use of TRIN in implementing trading systems; Narcouzi (2000) describes the use of the Chaikin Money Flow Indicator for generating and confirming trading signals; Penn (2003) describes the use of intermarket analysis; Blackman (2004) explains the merits of intermarket analysis; Teseo (2003) describes trading based on the analysis of linear regression slopes;

Technical Variable Classification	Brief Description and Reference
Pattern based approaches	Pring (2000) describes the identification of a variety of price patterns, a series of articles continued in Pring (2000); Bulkowski (2000) identifies a number of reliable and significant chart patterns for trading; Bulkowski (2000) also identifies specific patterns of 3-bar length; Pring (2000) extends his previous work on price patterns to those which specifically cover a two unit time period; Chandler (2000) describes the method of trading a specific pattern, the Andrews Pitchfork; Vomund (2000) describes trading with Point-and-Figure charts; Gopalakrishnan (2000) describes trading with pivot points; Hartle (2000) describes trading using triangle formations; Dukas (2000) describes which chart patterns are common during which market phases; Miller (2001) describes enhancements to head-and-shoulders trading; Hill (2001) discusses a specific reliable chart pattern; Narcouzi (2002) describes price patterns that lead changes in trend; Boot (2002) details trading based on analysis of Head & Shoulders patterns;

Table 2-1 Classification of Technical Variables used by practitioners

The majority of pattern based approaches used simple descriptions to describe complex, reasonably poorly defined patterns, so they are excluded from further analysis. Nevertheless, it is interesting to note that the majority of these patterns use increased volume as a confirming attribute of the patterns existence.

Initially, the brief analysis of technical variables being used by publishing practitioners appears to show that almost anything goes. However, a study of the formulae behind many of these indicators shows that in many circumstances, the same basic technical information is simply being put together in a variety of different ways. This issue is further considered in the selection of technical variables, section 3.7.1.4.

2.3.3 Applicability

Essentially, the discussions above reveal that there are mixed feelings in the academic world of finance regarding the suitability of Technical Analysis as an investment technique. However, there are certain facets of Technical Analysis which do appear to have academic acceptance, for example, in addition to those papers cited above, Siegel (2002) 'cautiously' supports the use of Moving Averages. According to Krutsinger (1997), it is also apparent that many expert traders use systems based on channel breakout approaches.

From the point of view of this thesis, the works reviewed above clearly demonstrate that technical analysis is not to be ignored. It has growing academic acceptance, and widespread practitioner acceptance. As such, the use of technical variables as inputs for ANNs are well within the remit of this study.

2.4 Soft Computing

2.4.1 Introduction

Unlike traditional (hard) computing, soft computing is tolerant of imprecision, uncertainty and approximation. The principal constituents of soft computing are Fuzzy Logic, Neural Computing, Evolutionary Computing, Machine Learning and Probabilistic Reasoning. The guiding principle of soft computing is to find ways to exploit the imprecision and uncertainty of the relevant domain, to achieve a robust solution.

As stated earlier, a primary motivation of the investment community is to increase trading returns. Soft-computing offers the distinct possibility of achieving higher returns for two principal reasons. According to Samuel and Malakkal (1990), with the advent of cheaper computing power, many mathematical techniques have come to be in common use, effectively minimizing any advantage they had introduced. Secondly, in order to attempt to address the first issue, many techniques have become more complex. According to Blakey (2002), there is a real risk that the signal-to-noise ratio associated with such techniques may be becoming lower, particularly in the area of pattern recognition. Soft-computing offers the possibility of enhancing predictability through superior ability to determine nonlinear relationships.

It should be noted that implementing soft computing does not necessarily translate into better prediction results, with some researchers finding simple linear time series models better at predicting their desired goals. As an example of this, see Callen et al. (1996), who used neural networks to attempt to predict quarterly accounting earnings for a sample of 296 firms on the NYSE, and found that forecast errors were significantly larger for their neural network compared to certain linear forecasting techniques.

Apart from certain specific examples, however, it is generally accepted amongst the academic community that ANNs offer superior forecasting ability than more conventional methods (see for example, Thawornwong and Enke (2004)).

2.4.2 Historical Evolution and Credibility

2.4.2.1 Soft Computing Classifications

There are a number of approaches within the literatures which deal with applying soft-computing techniques to investment and trading. Although there appears to be no formal segmentation of these different approaches, this paper classifies the literature into the topics proposed by Tan (2001), and augments these classifications with one more category, namely, Hybrid. These categories are:

- Time Series – forecasting future data points using historical data sets. Research reviewed in this area generally attempts to predict the future values of some time series. Possible time series include: base time series data (e.g. Closing Prices), or time series derived from base data, (e.g. Indicators - frequently used in Technical Analysis).

- Pattern Recognition and Classification – attempts to classify observations into categories, generally by learning patterns in the data. Research reviewed in this area involved the detection of patterns, and often, the segregation of base data into 'winner' and 'loser' categories.
- Optimization – involves solving problems where patterns in the data are not known, often non-polynomial (NP)-complete problems. Research reviewed in this area covered the optimal selection of parameters, and determining the optimal point at which to enter transactions.
- Hybrid – this category was used to distinguish research which attempted to exploit the synergy effect by combining more than one of the above styles.

Most soft-computing is data intensive, and relies heavily on a large number of data points being available. The data used for Value Investment is typically drawn from annual statements; hence the volume of data available is low. Perhaps this explains why soft-computing techniques appear to find limited application in the field of Value Investment (and indeed, Fundamental Analysis).

2.4.2.2 Research into Time Series Prediction

The area of time series predictions is normally focused on attempting to predict the future values of a time series in one of two primary ways, either:

- Predicting future values of a series from the past values of that same series, or
- Predicting future values of a series using data from different series

Some researchers focus on attempting to predict future direction of a series (for example increasing from last known value, decreasing from last known value, or no change). Research of this nature is essentially a classification problem, and is discussed in Section 2.4.2.3.

In many ways, these two primary prediction methodologies relate quite closely to technical analysis strategies. For example, the use (and projection) of a moving average over a series of stock prices could be regarded as predicting future values of a series (the moving average) from past values of the same series. Indicators in Technical Analysis are often composed of a number of constituent data items, like price, volume, open-interest, etc. These indicators are commonly used to give indications of future directions of price.

A number of important papers are reviewed below, chosen as they are either representative of current research directions, represent an important change in direction for this style of research, or represent a novel approach to this area of prediction. Typically, current research in this area focuses on predicting returns, or some variable thought to correlate with returns (e.g. earnings). The papers below were chosen as they focus on the use of soft computing techniques. It should be noted that there are cases where standard statistical techniques have been demonstrated to outperform ANNs, as in the case of linear time-series, or non-linear time-series with specific characteristics.

Falas et al. (1994) used Artificial Neural Networks (henceforth, ANNs) to attempt to predict future earnings. The predictions were based on a number of reported accounting variables, broadly following the value investment approach of detecting causal anomalies between fundamental variables, in this case, earnings. Their conclusion showed no significant benefit using an ANN compared to the logit model. Their paper suggested that one reason for this is that the accounting variables chosen were not appropriate earnings predictors. This conclusion represents one of the major problems encountered when working with ANNs, namely, their virtually non-existant explanatory capability. It is not unusual to find conclusions of this type when reviewing ANN research, with non-correlation often being reported as wrongly chosen input variables.

Quah and Srinivasan (2000) demonstrate the use of (mainly) fundamental variables to predict excess returns. Again, this follows the value investment approach concerning anomaly detection. The research uses an ANN approach to determine returns relationships, and then builds portfolios based on the expected high return and low return stocks. The results demonstrate little except that the high return portfolio had bettered the performance of the SESALL index.

As described above, time series prediction may be performed on base data like security prices, or on data derived from this base data. For example, Ruggiero (1997) describes examples of neural networks based on time series prediction of future crossovers of moving averages and other technical indicators. Similar uses are reported by Kaufman (1998), who states that neural networks are ideal for predicting forward-shifted technical indicators, because they tend to have smoother results.

As an example of a novel approach, Chan and Foo (1995) used neural networks to predict the value of future time series data, and then used these predictions to compute various Technical Indicators, used in Technical Analysis. Using the DEM/USD daily data from 01 January 1992 to 30 March 1995, they predicted the high, low and close prices three days into the 'future'. They then computed Donchian's Moving Average (SMA), Lane's Stochastics, and Momentum for these 'future' prices. Their network used the hyperbolic tangent activation function, and momentum learning, with 15 input nodes, 20 hidden nodes (in 1 layer) and 3 output nodes. In testing the model, transaction costs, stop/loss and slippage were taken into account. The model significantly outperformed the trading results of the 'normal' (unpredicted) technical indicators. The authors conclude that the networks ability to predict allows the trader to enter a trade a day or two before it is signalled by regular technical indicators, and that this accounts for the substantially increased profit potential of the network.

Brabazon (2002) uses technical, fundamental and intermarket data to predict 5-day percentage change in the FTSE-100. He finds the neural network to be capable of detecting the structure in the underlying data, but finds that prediction accuracy declines as the time lapse from the model building process increases.

Shao et al. (2003) used stochastic fuzzy neural networks to forecast the earnings per share of 300 public corporations from Shanghai and Shenzhen stock markets, with some success.

Very little work exists within the context of the Australian stockmarket, however, Pan et al. (2005) produced an excellent paper on predicting the AORD using intermarket information. The study covered the time period from Jan 1990 to Aug 2003, and yielded 80% directional prediction correctness.

A variety of other papers within the category of time series prediction have been reviewed, and provide evidence and conclusions consistent with the above. Briefly, Yao and Poh (1995) use Technical Indicators (%K and %D) along with price information to predict future price values. They achieved good returns, and found their models performed better using daily data rather than weekly data. Hobbs and Bourbakis (1995) predict prices of stocks based on the fluctuations in the rest of the market for the same day. They show consistently high rates of return, although the investment is done in a frictionless environment. Paying commissions on the large number of trades instigated would certainly erode much of the benefit from the trading strategy proposed. Austin et al. (1997) develop a neural network that predicts the proper time to move money into and out of the stock market. They used two valuation indicators, two monetary policy indicators, and four technical indicators to predict the four week forward excess return on the dividend adjusted S&P 500 stock index. The results significantly outperformed the buy-and-hold strategy. Kim et al. (2003) use backpropagation ANNs to predict future elements in the price time series in the KOSPI. Mingo et al. (2002) use time delay connections in enhanced neural networks (that is, the addition of time-dependant information in each weight) to forecast IBEX-35 (Spanish Stock Index) index close prices 1 day-ahead. Slim (2004) uses stochastic neural networks for forecast the volatility of index returns in the TUNINDEX (Tunisian Stock Index), and finds that the out-of-sample neural network results are superior to traditional GARCH models. Nenortaite and

Simutis (2004) present a trading approach based on one-step ahead profit estimates created by combining neural networks with particle swarm optimization algorithms. The method is profitable given small commission costs, but does not exceed the S&P500 returns when realistic commissions are introduced. Jaruszewicz and Mandziuk (2004) train ANNs using both technical analysis variables and intermarket data, to predict one day changes in the NIKKEI index. They achieve good results using MACD, Williams, and two averages, along with related market data from the NASDAQ and DAX.

2.4.2.3 Research into Pattern Recognition and Classification

Pattern recognition techniques and classification techniques have been grouped together, as their goal is normally not to predict future values of a time series, but to predict future direction of a time series. For example, the primary goal of chartists (a style of technical analyst) is to attempt to predict trend turning points by studying chart price action, looking for certain patterns. Chartists have noticed that these patterns tend to re-occur, and are reasonably reliable indicators of the future direction of price trends. There is a great deal of these chart patterns, and different analysts attach different weightings to the predictive power of any given pattern. Also, these patterns generally normally need to be confirmed by values from another time series (such as volume) to be considered 'reliable'. For more detail on this area, the reader is encouraged to refer to Pring (1999), or Elder (1993). Non-pattern matching techniques which also attempt to predict future direction of a time series are reviewed in this section, as classification problems. Although the inputs and techniques of each of these areas may be different, their goal is common. Quite often, in addition to predicting future direction of a time series, classification research attempts to classify stocks into two main groups, namely 'winners' and 'losers'.

The papers below were chosen as they focus on the use of soft computing techniques. It should be noted that there are standard statistical techniques such as logistic regression and discriminant analysis which can also be used to predict turning points. Refer to Bhattacharya (2003) for a discussion of these techniques.

An example of chartist pattern recognition is provided by Kamijo and Tanigawa (1990), who investigated pattern recognition ANNs by building recurrent neural networks that were capable of detecting 'triangles' in stock price data. In Technical Analysis, a 'triangle' is a particularly subjective price pattern, often claimed to provide clues to future changes (and directions) of trends. According to Kamijo and Tanigawa, a triangle has non-linear time-elasticity and definite oscillations. Three years of stock data for all names in The First Section of the Tokyo Stock Exchange was analysed by an independent expert, and triangles were manually identified. ANNs were then trained on this data, eventually correctly identifying 15 out of 16 triangles presented. As suggested earlier, whilst it has been shown that ANNs have superior pattern classification ability compared to standard statistical methods, many papers do not demonstrate whether this superior ability can be used to leverage excess profits for the investor. In this particular case, the paper is concerned with the ability to detect triangle patterns, and does not attempt to evaluate the detection in terms of financial returns.

Further work in pattern recognition by Yasunobu et al. (1991) defines an architecture for a Technical Analysis expert system, encompassing knowledge representation, rule definition, inference capacity, and knowledge verification facilities. The system identified recurring patterns in charts, and invoked expert knowledge to determine how best to trade a given pattern. The paper states that 5 main chart types, and 124 trading rules were used with data from the bond futures market. The paper does not say which country the market was made in, nor the time range covered by the data. No statistics are published for the predictive ability of the system, although the paper lists out a number of deficiencies in the system, namely speed, lack of expressive power for certain patterns, and an inability to intuitively recognize a chart pattern after viewing it for a long period of time.

Fu et al. (2001) addresses one of the more complex issues in pattern matching, namely, when looking for a pattern in a time series, how is the length of the sub-series to be matched best determined. The paper matches data from the Hang Seng index against 22 different Technical Analysis patterns, using Genetic Algorithms (henceforth, GA's) to identify the appropriate 'best' time series length. The paper shows that the segmentation method selects segment length dynamically, and these lengths show closer matches to the Technical Analysis patterns than could be achieved using fixed length segments.

While the above papers each demonstrate some ability to predict patterns, none of the papers continue on to draw conclusions relevant to an investor / trader. The papers below, those involving classification, are all focused on trading returns, an area of prime importance to an investor / trader.

A good example of the 'winners' and 'losers' classification is provided by Longo (1996). This research used ANNs to determine non-linear correlations between a set of fundamental and technical variables, and the expected returns from stocks. This approach is similar to Piotroski's (2000), and adds some technical variables to the general search for 'winners' and 'losers'. Longo classified stocks as either 'winner' or 'loser' stocks dependant on their level of returns, and found a high degree of correlation between certain fundamental and technical variables, and their return. Longo also used an ANN to determine market entry and exit points, predicting whether the upcoming return on the S&P 500 was positive or negative. The combination of the best ANNs was used to build a combination portfolio as part of Longo's research, which yielded a compound return of approximately 42.1% over an 18 year period (versus 18.36% for the market). This work is a typical example of using ANNs to classify time series data. The work initially classifies time series projections into 'winner' and 'loser' categories, and then uses classification to determine whether the S&P 500 is likely to yield a positive return ('increasing' and 'decreasing' catagories). Finally, an ensemble approach is implemented by combining the results of the ANNs.

Suh and LaBarre (1995) provides a similar classification system to that of Longo, above, but restrict it purely to the input of fundamental financial variables. The work uses

ANNs to classify stocks as either 'winners' ('losers') based on the ANNs prediction of an increase (decrease) in the change of EPS (Earnings per Share). Portfolios are then built for the two classes of stocks, and it is shown that the return from the 'winner' portfolio exceeds that of the S&P 500 by about 5% over the test period. It should be noted that due to the constraints imposed on selecting the initial stocks for training and testing, the volume of data used is very small.

Skabar and Cloete (2001) produced an ANN for predicting future directions for the DJIA, and used this to determine whether a trader should invest in the index, or a fixed interest security. They found a large discrepancy between the buy and hold strategy (-6.8%), and the predictions of the ANN (43.6%) over a 250 day training window. Over a 40 day test period, the buy and hold strategy had a return of –5.2% compared to –1% for the ANN. The inputs to the ANN were the closing price of the previous day, and a set of moving averages (5, 10, 15, 30, and 60 day moving averages). Skabar and Cloete then used Structured Learning with Forgetting, a method which attempts to strip the ANN parameters down to skeletal form, in an attempt to extract learning rules from the ANN. This produced one rule involving relationships between the 5 day moving average and the 30 day moving average. Skabar and Cloete report that the rule is similar to that commonly used in real world trading of financial markets.

Leus et al. (2001) specifies a GA that develops rules based on a variety of different technical indicators. The approach is to classify stocks as delivering good returns under specific technical analysis situations. The GA developed in the paper also rewards strategies that spend minimum time in the market, effectively attempting to reduce market risk. The paper demonstrates profitable and stable rules for some specific stocks on the Toronto Stock Exchange, and shows that the rules are effective in minimizing the amount of time the system is 'in the market'.

Mizuno et al. (1998) present an ANN using technical analysis to predict buying and selling timing for the TOPIX, based on weekly data from September 1982 to August

1987. This ANN is classed as a classification network, rather than an optimization network, as its output is Buy, Hold or Sell. It does not attempt to predict the actual values and then trade based on those predictions. The paper states that the ANN tends to improve only the prediction accuracy of the most dominant class of inputs, and the paper proposes a method to determine the appropriate amount of sample duplication to avoid this problem. As the most common output (without weightings) is Hold, the paper demonstrates duplicating cases of Buy and Sell inputs, until there is a valid distribution for each result. The research shows that although the ANNs overall performance exceeded that of the technical indicators, all of the mechanisms achieved a lower return than a buy-and-hold strategy. The paper proposes this is due to the fact that the trend of price change was increasing throughout the learning and testing periods.

Another classification expert system is described by Yamaguchi (1989). This system simply classifies stocks as Buy or Don't Buy. It uses moving averages (13 weeks and 26 weeks) to detect dead and gold crosses, and symbolic information to represent the data encoded in Japanese Candlestick charts. The expert system then classifies the stocks, and attempts to determine those which will increase more than 5% throughout the next 2 months. The hit rate for the system was 53.1%.

A variety of other papers within the category of pattern recognition and classification have been reviewed, and provide evidence and conclusions consistent with the above. Briefly, Baba and Handa (1995) used ANNs of 14 input variables to predict increasing (decreasing) trend of a stock price one month into the future. This work is extended further by Baba et al. (2001). Baek and Cho (2000) train an auto-associative ANN to detect "left shoulder" patterns in the Korean composite Stock Price Index, achieving significantly greater performance than the buy-and-hold strategy. Yang and Yang (2003) view trend forecasting as a problem of pattern recognition, with mixed results. Chen et al. (2003) use probabilistic neural networks (PNN) to forecast the direction of the Taiwan Stock Index with good results. They find that the PNN demonstrates stronger predictive power than both the GMM-Kalman filter (Generalized Method of Moments combined with a Kalman filter), and random-walk forecasting models. The inputs to the PNN are economic state input variables, such as interest rates and lagged gross national and domestic products.

2.4.2.4 Research into Optimization

The focus of optimization in this thesis is directed towards research that uses soft-computing specifically to attempt to optimize an otherwise accepted achievable result. Typical of this style of research paper, an already accepted result is discussed, then considered for optimization. The optimization is characteristically proven by excess returns compared to the pre-optimized case.

For example, an Index Arbitrage timing has been proposed by Chen et al. (2001). Their model attempts to optimise the correct entry point timing for index arbitrage positions. Current arbitrage models propose establishing an arbitrage position immediately an opportunity arises; the neural network approach is to attempt to locate the timing when there will be a maximum basis spread for the arbitrage, thereby increasing profit potential. Using data from the Nikkei 225 index, the researchers trained a neural network with three main groups of inputs, namely, a) differences between the Nikkei 225 spot and futures markets, b) Difference ratios between successive periods, and c) technical values of the T-Basis (including rate of change, RSI, MACD, and PSY). The research concludes that the neural model significantly outperforms the traditional approach, with the traditional model yielding 6.35%, and the neural model yielding 16.44%, for the first season of 1996 data.

Tan (1994) evaluates a system built on the principles of moving average oscillators, commonly used in technical analysis. Of primary importance in this type of system is the selection of the time lengths used to define the oscillators. The paper uses GAs to determine the time lengths, and tests the optimized system against the Hang Seng index. The result shows that the returns from the genetically optimized outcome exceed those from the best simple oscillator system by 12.8%.

2.4.2.5 Research into Ensemble Approaches

In this paper, research is classified as an ensemble approach if it combines work from more than one of the areas described above, effectively attempting to leverage the synergy effect by achieving an end result greater than that expected from each of the individual constituents. Amongst soft-computing research, there is a growing trend towards using the ensemble approach to analysis. According to Pan et al (2005), the probability ensemble of neural networks is one of the most promising directions.

Jang et al. (1991) build two neural networks and allow each to learn correlations between price movement trends and a stochastic indicator (Forward K). One of the networks utilizes a 12 day moving window, and the other filters data from the latest trading quarter. The final output is a weighted sum of the outputs from both networks. Further, an adaptive weight adjusting algorithm is used to tune the weights according to the predictive accuracy of the model. The idea is that the ensemble is the weighted output of two different viewpoints, namely, a short-term trend view, and a long-term trend view. The weighted output is shown to produce good market returns against a sample from the Taiwan Stock market. Subsequent to this result, the researchers show that smoothing the prediction of the ensemble using moving averages produces superior trading results.

Other ensemble approaches combine different soft-computing techniques into an integrated predictive system, typically combining ANNs and GAs. For example, Baba et al. (2002) combines ANNs and GAs to predict and trade in TOPIX (Tokyo Stock Exchange Price Indexes). The approach uses an ANN to allow forecasting of the highest and lowest prices of the TOPIX up to 4 weeks into the future. These predictions are then used by a GA to find the most effective way of dealing from this information. The GA is constructed to test parameters for a specific set of trading rules, and optimize these parameters. The research evaluates the effectiveness of the ensemble by comparing it to a system based on similar technical analysis methods, as well as the buy-and-hold strategy. The results demonstrate that the ensemble would be an effective profit making strategy, and that it outperforms the standard technical analysis techniques. The paper also shows the standard technical analysis techniques outperform the buy-and-hold strategy for the same period. It is also notable that the data range testing the strategy is limited to 1 year (Nov. 1996 to Oct. 1997). The paper leaves the issue of extending the data range to future study.

Chou et al. (1997) demonstrate combining a system covering a large number of technical analysis indicators, with fuzzy decision rules forming induction trees. The system uses data from the Taiwan market (January 1990 to April 1995), and compares its output against the buy-and-hold approach, and the four major closed-end mutual funds operating in Taiwan. The research shows the system outperforms the buy-and-hold approach, and outperforms the four major funds in both 1992 and 1993, but curiously, the system performs poorly in 1994. The paper concludes this was due to 'human influence' on the market in that year.

Another example of an ensemble system is provided by Dourra and Siy (2002), which build a fuzzy logic interpretive backend onto a technical analysis system. The system determines the values of various technical analysis indicators, and then uses fuzzy logic to quantify the values of the technical analysis outputs into membership sets for fuzzy logic interpretation. Based on this, the system evaluates expert provided rules in terms of a value being "large", "small" etc, without having to describe what the actual numeric value is. The final system demonstrates superior results, surpassing the S&P500 returns by a considerable margin. The system did account for transaction costs (on a fixed basis per transaction).

Further work on ensemble approaches is to be found from Leigh et al. (2002), which considers combining technical analysis techniques with pattern recognition techniques (to identify bull flags), ANNs to forecast future price (20 days hence), based on previous price and volume values, and GAs to determine the subset of the ANN inputs that improve the correlation observed between the ANN inputs and outputs. The work draws its data from the NYSE, and demonstrates forecasting price changes for the NYSE

Composite Index. These approaches demonstrate the superior decision making results obtained using ensemble approaches.

Expert Systems (henceforth, ES) are also combined into ensembles, with Liu and Lee (1997) using ES based on technical analysis to determine stock recommendations. The system uses Stochastics, RSI, Money Flow, Moving Averages, and Support/Resistance trendline rules to make future predictions based on classical technical analysis rules. The system is capable of explaining to users why certain recommendations have been made. The paper does account for transaction costs (at 1% of trade), and shows promising results. The paper notes that for the time period tested, the returns for the ensemble system (22%) (use of all indicators combined the by expert system), exceeded the returns for all the individual indicators used alone. It further shows that of all the indicators used, the RSI was the only one with a negative return (-1%). Highest individual returns were achieved by Stochastics (11%), and Granvilles's 200-Day Moving Average (13%).

More work on expert systems ensembles is provided by Lam et al. (1996). They combine ES with technical analysis indicators, and fuzzy logic. The system uses moving averages, relative strength index (RSI), Stochastic (K%D), directional movement index (DMI), and stop and reverse parabolic system (SAR) from TA, and combines these value inputs with fuzzy membership functions and expert system rules. The system is evaluated against the stock Cheung Kong (a well known blue chip property stock on the Hong Kong markets), using data from January 1994 to April 1995. Overall the system shows excellent results, and only one case of a false signal. The system demonstrates better than 90% accuracy in issuing buy and sell suggestions.

Another example of ensemble approaches is provided by Lee and Liu (2001), which combines several neural approaches to build a fully integrated stock prediction system. The research covers data from 1990 to 1999 for 33 major Hong Kong Stocks. The stock prediction system uses oscillatory-based recurrent neural networks to forecast long and short term prices. It also uses Gabor filters for feature extraction to enable classification of stock market patterns via a neural oscillatory-based elastic graph matching model. The paper shows that the ensemble model achieves a significantly lower average percentage error than using individual networks, or genetic algorithms, effectively demonstrating a synergy effect from ensemble approaches.

A variety of other papers within the ensemble category have been reviewed, and provide evidence and conclusions consistent with the above. Briefly, Abdullah and Ganapathy (2000) use ensembles of ANNs and establish the predictive ability of the ensemble exceeds that of individual ANNs, again demonstrating the synergy effect. Wong and Lee (1993) demonstrate the use of hybrid systems that are able to discriminate between underperforming and overperforming stocks, as well as address some issues regarding window sizing of data. Chenoweth et al. (1995) demonstrate the use of trading rules with detailed ensembles involving statistical feature selection, simple data filtering neural networks and a symbolic decision rule base. Skabar and Cloete (2002) train neural networks using a genetic algorithm based weight optimization procedure, and test their results on the DJIA. They find their results are significantly better than would have been expected had the price series been random. Thawornwong et al. (2002) find profits from their ensemble of neural and genetic components exceeds the use of ARIMA, MACD and buy-and-hold, and also raise the suspicion that excessive trading is a cause of negative returns, a sentiment raised by other researchers, such as Barber and Odean (2000). Baba (2002), combines ANNs and GA to deal in the TOPIX and NIKKEI-225, and Baba et al. (2003) propose combining GAs, ANNs and technical analysis for enhancing financial prediction. Dong and Zhou (2002) combine saliency analysis with ANNs to study 1960 head-and-shoulders patterns in 753 US stocks. Their results show that pattern length and firm size are the most relevant to future returns. Baba et al (2003) propose an ensemble utilizing traditional technical analysis, neural networks and genetic algorithms, effectively concentrating on enhancing returns from a traditional moving average crossover strategy. Simutis and Masteika (2004) demonstrate the success of a fuzzy-neural-genetic algorithm approach, using technical input variables based on change in price and change in volume, by testing their network over the S&P500 from start 1992 to end 2002, with excellent results. Versace et al. (2005) use genetic algorithms to select architectural features of ensembles of neural networks, and find that this technique offers promise, with the ensemble trained showing a high level of correct up/down (directional) predictions over the testing set. Wang et al. (2002) demonstrate the effectiveness of combining neural networks with rough sets, to capture the generalization capability of neural networks and the rule generation capability of rough sets. Ferreira et al. (2004) demonstrate the principles of combining genetic algorithms with neural networks on five complex time series (Henon, Sunspot, Dow Jones, S&P 500 and Petrobras). Genetic algorithms were used principally to determine the minimum dimensionality necessary to reproduce (to a given accuracy) the phenomenon generator of the time series.

2.4.3 Applicability

Essentially, the literature above demonstrates that there are a vast number of suitable approaches for applying soft computing to investment trading. A summary of the above reviews was presented as a paper by Vanstone and Tan (2003), and concluded that the majority of investment and financial trading research focuses on technical analysis. This is unsurprising, given the data intensive requirements of soft computing. Vanstone and Tans paper also demonstrates that the majority of the work in applying soft computing to investment and financial trading occurs in the area of pattern recognition and classification, and hybrid (ensemble) techniques. This is also to be expected, as this is a primary focus of technical analysts. In terms of soft-computing, Vanstone and Tan's paper demonstrated that the literature appears primarily focused on the use of neural networks, particularly the backpropogation style. This is due in part to the outstanding ability of neural networks to deal with noisy datasets, and due to the simplicity of implementing back-propogation networks. Finally, Vanstone and Tan demonstrated a wide variety of 'success' criteria in common usage, making the comparison of many techniques extremely difficult. Refer to Mitsdorffer (2002) for an overview of the many issues involved when comparing soft-computing techniques.

2.5 Research Contributions

Building on the previously cited work, this study intends to contribute to the current body of knowledge, as described below:

- Contribute a well defined methodology for building and testing trading systems using Artificial Neural Networks
- Provide much needed depth to the study of stockmarket anomalies from an Australian perspective
- Present trading results which are relevant to both institutional and individual investors

It is hoped that by presenting a well defined methodology to build trading systems, other academics who lack this knowledge will be able to join in and contribute to the advancement of trading systems, one day leading to the building of a model that describes the underlying mechanics of the stockmarket pricing process.

2.6 Conclusion

Of the findings presented above, and summarized in Vanstone et al. (2003), it is clear that a large amount of investment trading research takes place using constraints that could not, or should not, be practically implemented. For example, none of the papers reviewed implemented any form of money management, and very few accounted for basic frictions, such as transaction costs. In terms of money management, research by Odean (1998) indicates that individual investors demonstrate a significant preference for selling winner stocks and holding loser stocks, perhaps in the vain hope of mean reversion. Interestingly enough it appears that this belief may have some credibility; although the debate regarding mean reversion has gone on for decades, it has often relied on evidence from relatively small sample sizes. Research from Balvers et al (2000) using stock index data from 18 countries during the period 1969 to 1996 shows strong evidence of mean reversion, with a half-life of 3 to 3.5 years They further find that contrarian investment strategies that exploit mean reversion outperform buy-and-hold, and standard contrarian strategies. A similar finding comes from Poterba and Summers (1998), also using data from 18 countries.

For this reason, the majority of the research is not applicable to the non-academic investment community. As an example of the problems this causes, the strategy detailed in Hobbs and Bourbakis (1995) produces a system which generates a great deal of small trades, of which the majority are profitable. The paper concluded that a profitable approach had been found. However, transaction costs were not accounted for. Taking transaction costs into account (i.e. implementing a real world constraint), the system became extremely unprofitable, as the gain from the trades was less than the cost involved in placing the trades. It is interesting that so few studies take account of transaction costs. Certainly, research exists by Barber and Odean (2000) which demonstrates that poor investment performance is more a result of the transaction costs associated with overtrading, rather than poor portfolio selection. This view is supported by practitioners, such as Burke (1993).

On a final note regarding technical analysis, many of the papers reviewed implemented a simple technical rule, probably due to the fact that it was easy to program. Many of these rules are linear in nature, yet it is precisely the non-linear aspects of technical analysis which appear to make it so rewarding. This point is well made by Jegadeesh (2000), who continues on to suggest that the investigation of pattern approaches may be more rewarding than the simple study of such techniques as moving average crossovers.

The results from this research will be evaluated and benchmarked in terms of trading systems metrics, and 'success' will be evaluated in those terms. There are obvious mismatches between the goals of artificial tools such as neural networks, and academics, and real traders. While neural networks are attempting to minimize forecast errors in training, typically academic research is focusing on trying to increase portfolio profits. While these are both valid goals, a traders' definition of 'success' is different yet again. Typically, a trader views a system as a 'success' if it maximizes profit, subject to minimizing drawdown, minimizing exposure to the market, minimizing position risk, and provides a smooth equity curve. Details of trading goals, and the interpretation of relevant metrics is provided in Section 3.5.

With the approach of benchmarking using trading system metrics, this thesis hopes to start to bridge the gap between academia and industry, and attempt to present systems which are both ‘tradeable’, and useful as models, both for academics and practitioners.

Chapter 3 Methodology

3.1 Introduction

The purpose of this chapter is to discuss the methodology and data employed in this study, and to formally state the hypotheses that will be tested. Other related issues, such as the selection of relevant variables, missing data, data sources, and the software packages utilized in the study are also discussed.

According to Chande (1997), a trading system consists of three major functions, namely:
1. Rules to enter and exit trades,
2. Risk Control, and
3. Money Management

According to Vince (1995), a successful trading system must provide the trader a positive mathematical expectation, that is, an edge. Without a positive expectation, no trading system can provide long-term success, and adjusting various components of the system (such as money management parameters) will only delay (or accelerate!) the inevitable failure. Vince advocates determining rules to enter and exit trades, and then determining whether the rules provide a positive expectation, an approach also advocated by Tharp (1998). If there is a positive expectation, then Vince focuses on selecting the best money management technique for exploiting it. There is no real focus on risk with Vinces approach, except that Vince suggests that if an optimal money management scheme (such as Optimal f) is used, then the trader can expect long-term success, however, there may be serious financial losses along the way to this long-term success. Vince's approach focuses on money management, with the rules to enter and exit trades being reasonably irrelevant, as long as they provide a mathematical edge. Vince's Optimal f technique defines the optimum amount of capital to expose to any given trading opportunity. In practice, very few (if any) traders will trade using the Optimal f technique, as the associated losses along the way to the longer-term success would not be acceptable, and the losses may well include virtually all of a traders capital at some point.

This thesis attempts to determine whether neural networks can be used to create economically feasible trading systems. It does this by developing neural networks to address Chande's first major function; that is, creating neural networks that can signal when to enter and exit trades. Two styles of neural networks are created, one style based on fundamental company data, which is geared toward longer-term investment horizons, and one style based on technical data, which is geared toward short-term trading. Two of each style of neural network are created, one of each style is created for trading the Australian Allshare, and one of each style is created for trading the S&P ASX200.

Chande's next two major functions, risk control and money management are also important components of trading systems. These components are implemented in this thesis so that the trading systems developed are suitable for real-world trading. The details of the implementations of risk control and money management are discussed later in this thesis, in sections 3.4.3 and 3.4.4 respectively.

An outline of the study can be found in Table 3-1.

	Outline of the thesis study
1.	Create the datasets used in this study
2.	Train a neural network for use as a fundamental signal generator (network FNN)
3.	Build and benchmark a trading system based on the signals generated by the FNN
4.	Train a neural network for use as a timing signal generator (network TNN)
5.	Build and benchmark a trading system based on the signals generated by the TNN

Table 3-1 Outline of Thesis Study

As shown in the outline of the thesis study in Table 3-1, after creating the trading datasets, the study attempts to firstly determine whether it is possible to train a neural network to generate buy and sell signals based on the fundamental company specific data. According to the research presented in the literature review, it is indeed possible to filter a set of stocks using only fundamental criteria, to derive a reduced set of higher-than-average return stocks.

A primary motivation for this work is to determine whether neural networks can realistically enhance returns from trading, and capture economically significant profits. As has already been seen from the literature review, creating a neural network which gives high accuracy signals is not sufficient to enable a trader to make economically significant profits. It must also be demonstrated that the signals given enable a trader to capture enough profit to outweigh the various costs of trading, and that the amount of risk that a trader is exposed to can be controlled. For this reason, the neural network trained from the fundamental data must be benchmarked both for accuracy of signals, and for its ability to capture profit from those signals within the constraints of a formalized trading system, which includes risk control and money management.

At this point, the study will proceed to consider the implication of timing in strategies, and this will be done using the framework of technical analysis. This thesis has taken the viewpoint that no variable should be included as an input to the neural networks unless there is significant justification for doing so. For the neural network trained using fundamental data, there is a detailed record of published academic work supporting each variables inclusion. For the neural network to be trained using technical data, there is not as substantial a body of academic published work to rely on. For this reason, the literature review also conducted a brief review of practitioners' journals. To support each technical variable as suitable for inclusion, the opportunity was also taken to create a function profile for each of the technical variables proposed as inputs to the technical neural network. A description of this activity and the supporting function profiles created are included in this thesis in Appendix A.

As outlined above, it must be demonstrated that the signals generated by the technical network enable a trader to capture enough profits to outweigh the costs of trading, and that the amount of risk that a trader is exposed to can be controlled. For this reason, the neural network trained from the technical data must be benchmarked both for accuracy of signals, and for its ability to capture profit from those signals within the constraints of a formalized trading system, which includes risk control and money management.

After completing this study it will be possible to draw conclusions regarding the suitability of using neural networks as signal generators within trading systems.

3.2 Data

3.2.1 Universe of stocks

The data used for this thesis covers the 10-year period from the first day of trading in 1994 to the last day of trading in 2003. Data is sourced solely from the Australian stockmarket, and the only securities considered are Ordinary Shares.

For institutional investors, often trading activities are restricted to a relatively small number of securities, typically the S&P/ASX200, or higher. The S&P/ASX200 is considered an ideal proxy for the Australian market by institutional investors. It covers 90% of the market capitalization for Australia, and its constituents are chosen for their size and liquidity, as described by Standard and Poors (2004). The S&P/ASX200 is considered the 'Investable Benchmark for Australia'.

For many individual investors, the S&P/ASX200 is too restrictive. By its nature, it excludes a vast number of tradeable securities, by virtue of either their size or liquidity. A considerable amount of research documented in the literature review of this thesis demonstrated a clear bias in returns amongst the smaller capitalization end of the market. It is precisely this which attracts many investors into that end of the market, for example, the pursuit of so-called 'penny shares'. This is the belief that a small company can grow much faster than a large company, and hence, its shares can potentially increase in value much more rapidly. There is also concern by many non-traders that often an investor cannot really purchase in the penny-share spectrum of the market without paying a substantial spread premium. This is not a fair generalization, and the ready growing acceptance of CFDs (Contracts for Difference) in trading is directly addressing this concern for many traders.

This thesis will attempt to accommodate both institutional and individual investors, by presenting trading strategy results for both views of the market, namely the S&P/ASX200 component of the market, and the All Share component. It should be noted that the All Share component naturally also includes the S&P/ASX200 constituents.

Further, the data sources used in this study contained data for delisted companies. As such, it is also possible to examine the effects of survivorship bias on trading results in the Australian market. Survivorship bias refers to the tendency for failed companies to be excluded from performance studies, due to the fact that they no longer exist. This is particularly important when examining the application of trading strategies, and very little work has been done in this area. It is worth noting that a large number of related US studies (as discussed in the Literature Review), used the Compustat database as a data source. This does not include some data for delisted companies, which would be essential in any trading study. These studies therefore suffer from survivorship bias. A number of papers address this issue of survivorship and the Compustat database, see for example Fluck et al. (1997).

3.2.2 Sources of Data

This thesis uses both fundamental accounting data, and technical trading data.

Fundamental data is sourced from Aspect Financials, in particular the FinAnalysis research database, produced by AspectHuntley (2004). This source provides detailed fundamental information for all companies listed on the ASX, and also maintains this data for delisted companies. The historic search tools allow for querying back as far as 1989, however, very little data is available until 1992, at which point the database becomes reasonably comprehensive.

Technical Data is sourced from Norgate Investor Services (2004). Full technical data (Date, Open, High, Low, Close and Volume) is provided from the beginning of 1992, and includes historical data for delisted companies. The data is fully adjusted for dilutions, stock splits, and consolidations, using the dilutions factors supplied to data distributors directly by the Australian Stock Exchange. The data is supplied from Norgate in proprietary Metastock format. Programs were written to assess the integrity and consistency of the technical data, specifically; every row for every security was checked to ensure the following conditions were met:

- Open <= High,
- Close <= High,
- Open >= Low,
- Close >= Low,
- Low <= High,
- Open > 0,
- Low > 0,
- High > 0,
- Close > 0,
- Volume > 0

3.2.3 Creating the trading datasets

The trading datasets were created from both the fundamental and technical data sources, using the merging process described in section 3.2.3.2. A trading dataset was created for every security. A brief summary of the key parts of the process is provided in Figure 3-1.

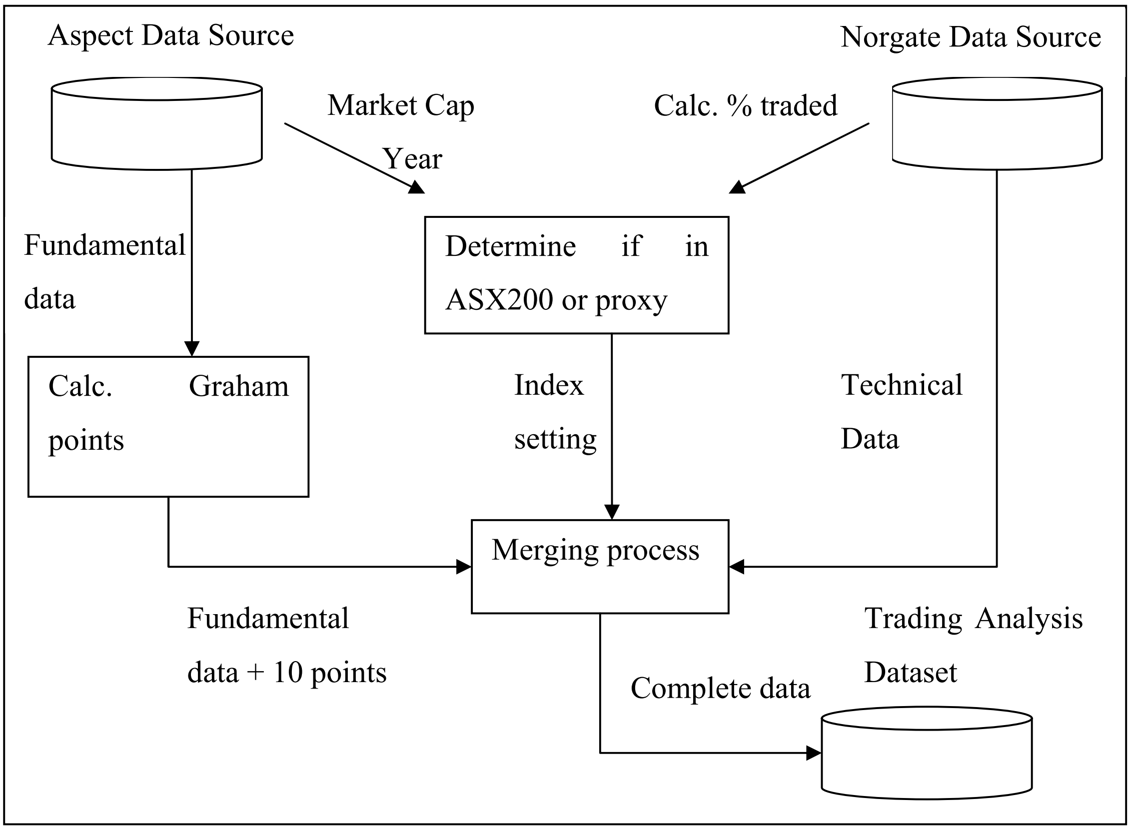

Figure 3-1 Summary of the dataset creation process used for each security

3.2.3.1 Data contents of the datasets

The following tables list the data acquired from each of the two data sources along with a brief description, and the final data contents of each trading dataset produced by the merging process. Where data items in the trading datasets have simply flowed through from the source datasets, they are referred to as 'from <sourcename>'.

3.2.3.1.1 Aspect Fundamental Data

Data values supplied by the Aspect database are normalized to enable comparison over time. Typically, 'normalizing' means that the data has been modified to exclude the effect of pre-tax non-recurring gains or losses. According to the Diversity Investment Group (2004), standardized, and/or normalized data is regularly used by institutional investors and researchers, and is available by subscription at relatively low cost to individual investors.

Aspect source field name	**Description**
ASX Code	The unique identifying code for each security.
Year	The calendar year.
(Annual Per Share Statistics) EPS	Earnings attributable to each ordinary share are measured by net profit after tax, abnormals, preference dividend and outside equity interests, divided by the weighted number of ordinary shares outstanding during the year. The earnings also exclude capital profits/loss and capitalized interest.
(Annual Ratio Analysis) PER	P/E Ratio
(Annual Per Share Statistics) Gross Div. Yield ex. Spec.	Gross Dividend Yield excluding Special dividends
(Annual Ratio Analysis) Price/Book Value	Year end share price divided by Book Value per share
(Annual Ratio Analysis) Year End Share Price	Closing share price on the last day of the companies financial year
(Annual Balance Sheet) Total Curr. Assets	Total Current Assets
(Annual Sundry Analysis) Total Gross Debt	Total Gross Debt
(Annual Per Share Statistics) Weighted Avg. Shares	Weighted average number of shares during the companies financial year
(Annual Ratio Analysis) Current Ratio	Current Assets / Current Liabilities
(Annual Per Share Statistics) Payout Ratio ex. Special	Percentage of dividends paid out to shareholders, excluding Special dividends

Aspect source field name	Description
(Annual Ratio Analysis) ROE	Net profit after Tax before abnormals divided by (shareholders equity – outside equity interests)
(Annual Ratio Analysis) Book value per share	The excess of total assets over total liabilities divided by the number of shares outstanding at the end of the period. This is also sometimes referred to as Carrying Value per share.

Table 3-2 Data sourced from the Aspect database

Note that the current ratio is considered a useful measure of a company's ability to pay its short term debts. A ratio of 2 or more (as required by Graham's strategy) was historically considered desirable, although it is worth noting that many companies in more recent times have reduced this as operating cycles have shortened. Therefore, it is unusual to see a modern company maintaining a current ratio of 2.

3.2.3.1.2 Norgate Technical Data

Data from the Norgate source is provided in proprietary Metastock format. As part of the production of the trading datasets, it was necessary to first translate this data into standard ASCII (specifically .CSV) format. The data also follows the defacto standard for daily trading data, which essentially means that if no trading volume occurred on a given day for a given security, then no pricing information is recorded for that security for that day, even though a market may have been possible.

Norgate source field name	Description
Date	Trading date
Open	Opening price on the trading day
High	Highest price reached on the trading day
Low	Lowest price reached on the trading day

Norgate source field name	Description
Close	Closing price for the trading day
Volume	Number of securities traded on the trading day

Table 3-3 Data sourced from the Norgate files

3.2.3.1.3 Trading Analysis Datasets

The output datasets are produced as a result of the merging process described previously. Essentially, one dataset is produced for each security. The data is created in .CSV format, to allow it to be easily used by any trading or statistical package.

The name of each dataset is its identifying ASX security code. Note that there are a few cases where the Windows operating system prevents this from occurring. Certain 3 character filenames are 'protected' within Windows, and are unable to be used (for example, AUX.CSV, and PRN.CSV amongst others). For these cases, an underscore was appended to each ASX code, and all programs that access these datasets have 'special' code in them to ignore the underscore.

The actual values ascribed to Points 1 to 10 in the table below are described in Section 3.3.1.1. Note that even though the Graham strategy being tested needs only 6 of these 10 points, Graham had originally defined 10 key points. A decision was taken to calculate all 10 points, as it is felt they may be useful in future research.

Output Name	Description
Date	from Norgate
Open	from Norgate
High	from Norgate
Low	from Norgate

Output Name	**Description**
Close	from Norgate
Volume	from Norgate
Point 1	Earnings-to-price yield double the AAA Bond Yield
Point 2	P/E four-tenths highest average P/E in most recent 5 years
Point 3	Dividend yield two-thirds the AAA bond yield
Point 4	Price two-thirds tangible book value per share
Point 5	Price two-thirds Net Current Asset Value
Point 6	Total debt less than tangible book value
Point 7	Current ratio greater than or equal to two
Point 8	Total debt less than or equal to net quick liquidation value
Point 9	Earnings doubled in most recent 10 years
Point 10	No more than two declines in earnings of 5 percent or more in the past 10 years
ASX200	See Section 3.2.3.1.4 for details
PER	from Aspect
DIVYLD	from Aspect
PRICE2BOOK	from Aspect
TOTALCURRENTASSETS	from Aspect
TOTALGROSSDEBT	from Aspect
WAVGSHARES	from Aspect
YEARENDPRICE	from Aspect
CR	from Aspect
EPS	from Aspect
BVPERSHARE	from Aspect
AAARTN	from Aspect

Output Name	Description
PAYOUT	from Aspect
ROE	from Aspect

Table 3-4 Data contents of the target Trading datasets

3.2.3.1.4 Determining the Index attribute

To cater for the needs of the institutional investment community, it was decided to produce results for the S&P ASX200 (the investable benchmark for professionals), as well as producing results for the Allshare for individual investors.

Unfortunately for this study, the S&P/ASX200 was first created in April 2000. For this reason, it was necessary to create a 'tradeable proxy' for the index, for the time period when it was not in existence (January 1994 to March 2000). The main requirements of such a proxy are as previous stated; its constituents must be large companies, with suitable liquidity.

A tradeable proxy was created by selecting the largest 200 companies per year based on market capitalization (as reported by Aspect Financials). These companies were then filtered for liquidity, using the trading data from Norgate Financial Services. Companies chosen on the basis of their market capitalization were excluded if they did not trade on at least 80% of the trading days during their selected period. The value 80% was chosen as it represented the fact that a normal trading week is five days, and thus (on average), a stock was traded four out of the five days. A visual inspection of the data revealed that stocks considered had either high liquidity (traded about 90% of the time, or more), or low liquidity (traded 40% of the time, or less). Thus, the 80% cutoff is considered reasonable.

Five of the companies in the index were delisted and subsequently became worthless. These were FAF (30/05/2000), HIH (04/07/2001), LIB (30/08/2002), ONE (30/08/2001), and PPH (19/12/2003). To represent the fact that these companies had become worthless at the date shown, and therefore to appropriately penalize any trading system that selected them, an additional zero price entry was made to the data files for each of these five stocks. A detailed analysis was made of the delisting outcomes of all other companies used in the index using the 'Delisted' websource (2004), which showed that all other delisted stocks in the index resulted in some financial recompense to the trader.

Finally, even though it was necessary to create the proxy as described above, as noted earlier, the real S&P/ASX200 was first created in April 2000. For this reason, the calculation of the index attribute described above is only relevant for the inclusion of training data.

As the S&P/ASX200 was established before the out-of-sample period, separate, additional programs were written to download the actual constituent list of the S&P/ASX200 for every trading day throughout the entire out-of-sample timeframe. The out-of-sample datasets for the S&P/ASX200 are built directly from this list of daily constituents, as provided by the Standard & Poors website (2003). Therefore, we can be confident that the out-of-sample data used in the thesis matches exactly to the actual S&P/ASX200, and the out of sample results benchmarked from the neural networks created can be compared directly to the S&P/ASX200 returns.

3.2.3.2 Merging process used to create the datasets

Datasets were created that merged together the technical and fundamental data for each security in the study, driven by the technical data. For this reason, it was necessary to ensure that fundamental data was not used for trading decisions before it had been released to the market. Investigations revealed that Aspect Financials had the date of annual report release stored, but not for all companies in the study, and only from September 1998. In other cases, the Australian Stock Exchange Listing Rule 4.6 (1996) could have been applied, which effectively limits the delay the ASX will tolerate for companies to send a copy of the Annual Report to security holders to 17 weeks (or, in the case of a trust, to 3 months). However, a previous Australian study by Halliwell et al. (1999) which worked with fundamental data has allowed a 6 month period for the information to become widely known to the market. The use of an arbitrary time delay is also typical in related US studies (which also use 6 months), as described in the Literature Review. Therefore, this study will do the same to allow for comparability. This decision also suits the concept of tradeability; that is, a system which relies on a trader reacting to fundamental information the day it is released to the market would be extremely difficult to trade in practice.

To allow for trading to take place from the start of the study period, fundamental data was sourced from 2 years before the start of the study period, to allow for the effect of displacing this fundamental data by 6 months. In effect, due to the fact that the fundamental data is always displaced by 6 months, the technical data used covered the period from the start of 1994 to the end of 2003, whilst the corresponding fundamental data was sourced from the start of 1992 to the end of 2002.

As a brief example, consider a company which regular produces its accounts as at the 31st December each year. The 1992 balance sheet data, dated 31st December 1992 can not be acted upon for 6 months. This will be 30th June 1993. The 1993 data, dated 31st December 1993 cannot be acted upon for 6 months. This will be 30th June 1994. In this example, the fundamental data produced at 31 December 1992, is in effect from 30th June 1993 until the 30th June 1994, at which time the accounts published at 31st December 1993 will come into effect. To enable trading to take place in this security from the beginning of the study period of 1994 onwards, it is necessary to acquire fundamental data since 1992, and technical data since the start of 1994. For clarity, examples are diagrammed out in Figure 3-2 and Figure 3-3.

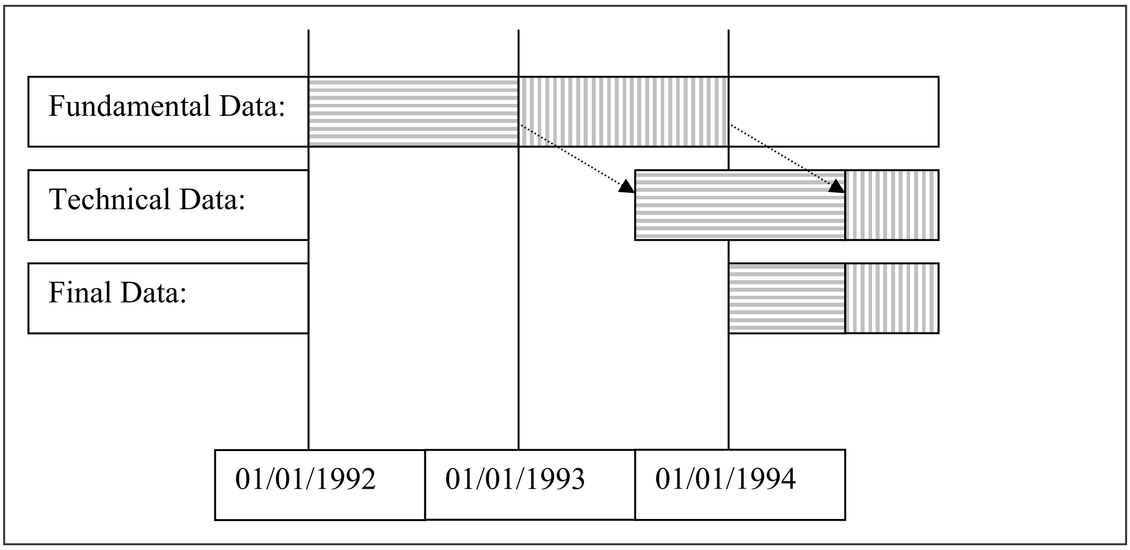

Figure 3-2 Start of data alignment for a security reporting at end of year

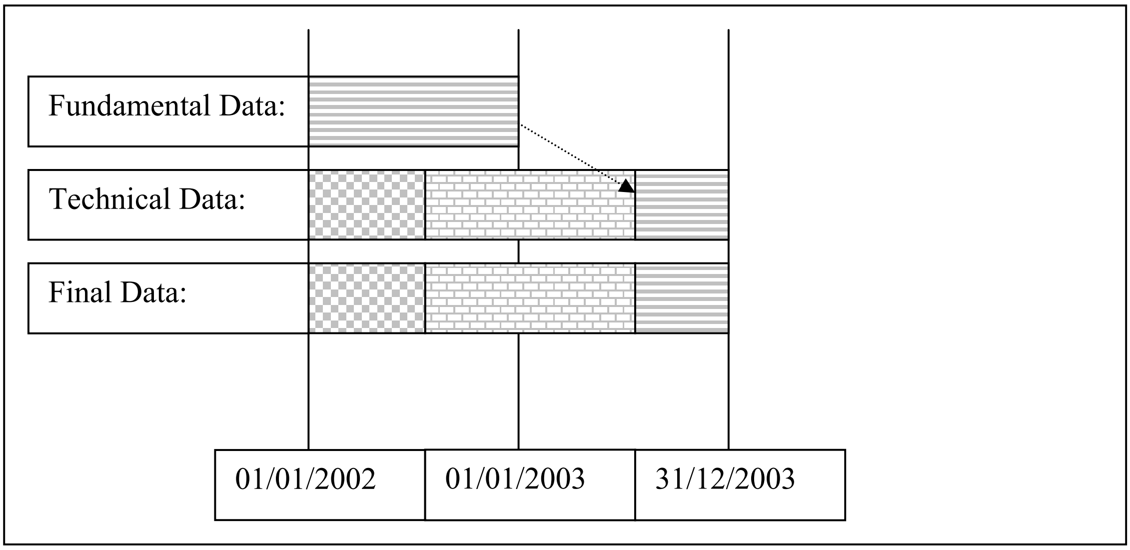

Figure 3-3 End of data alignment for a security reporting at end of year

It should also be noted that the majority of companies produced their accounts as at 30th June each year. No fundamental data was sourced for this study after the end of June 2003. Technical data used must cover the entire study period, however, to allow for the effect of trading out securities acquired based on earlier technical data.

3.2.3.3 Partitioning of the Data

Any study involving optimization, or neural networks must logically separate data which will be used for training, from data that will be used for testing. After creating the trading datasets, this study followed good practice by also physically separated the training data from the testing data.

There is acceptance within the academic community that the relationship between security prices (and returns), and the variables that constitute that price (return), changes over time as described by Refenes et al. (1993) and also by Thawornwong and Enke (2004). In other words, the structural mechanics of the market are changing over time, and their effects on prices are also changing. For this reason, it is necessary to partition data vertically rather than horizontally. A vertical partition of a dataset will divide the dataset into two partitions, one for training, and one for testing. Typically, the training dataset is larger, and covers a significant date range of the overall data, whilst the testing dataset is smaller, and used to provide out-of-sample confidence. These two partitions are typically known as in-sample (training), and out-of-sample (testing) partitions. Using this approach, every security will have its dataset partitioned into training and testing subsets.

The horizontal approach to partitioning splits entire datasets into either a training or a testing block. For example, horizontally partitioning 10 datasets, with 60% in training, and 40% as testing would yield 6 entire datasets used for training, and 4 entire datasets used for testing. This approach is invalid when it is recognized that the structural mechanics change over time, due to the fact that a neural network may well learn correlations that could not have been known in chronological time, and later, exploit these during the testing phase. This may well lead to higher quality predictions, but is clearly unrealistic.

The approach taken, then, is to vertically partition each dataset into a training and a testing set. The actual ratio for the split is generally chosen dependant on how much data is available, and is often arbitrarily chosen. However, some general guidelines can be distilled. For example, Ruggiero (1997) suggests that the data sets used for training should cover a period at least 5 times longer than the desired life of the model to be produced, and suggests using 80% of the data for training, and 20% for testing. Azoff (1994) takes a typical approach, which suggests that the training period should be long enough to cover typical market behaviour, including bull, bear, and sideways moving markets. Kaufman (1998) suggests a 70:30 split between training and testing, Kim and Lee (2004) suggest an 80:20 split, Gately (1996) suggests saving only 10% of the available data for testing, a 90:10 split. From an optimization point of view, Pardo (1992) suggests choosing a period long enough to cover a variety of market activity, and advises choosing a size large enough to generate at least 30 trades for statistical validity. Pardo also notes that the size of the models test window will affect trading shelf life, specifying that the life of the model will be between one-eight and one-quarter of the test window. Pardo further suggests a rule-of-thumb that the walk-forward window (in optimization) should be approximately 10% - 20% of the optimization window.

There are a wide variety of other competing and complementary guidelines available. In essence, the main principle is to capture as much diverse market activity as possible (with a long training window), whilst keeping as long a testing window as possible (to increase shelf life and model confidence). As this study covers a 10-year period, it appears reasonable to split the data 80:20, chronologically. That is, data in the first 8-year period (start 1994 – end 2001) for each security shall be used for training and parameter optimization, and data in the last 2-year period (start 2002 – end 2003) for each security will be used for testing. This split provides a reasonable compromise, and takes the above guidelines into consideration.

The above partitioning will apply to all of the trading datasets. These datasets will be further categorized into four groupings, to facilitate the interpretation of out of sample trading results. As mentioned earlier, this study has access to fundamental and technical data concerning delisted companies. This study also desires to cater to the needs of both institutional and individual investors. For this reason, four categories of datasets are needed:

- Category 1 – contains all trading datasets for constituents of the S&P/ASX200 (and proxy) INCLUDING delisted constituents
- Category 2 - contains all trading datasets for constituents of the S&P/ASX200 (and proxy) EXCLUDING delisted constituents
- Category 3 – contains all trading datasets INCLUDING delisted data
- Category 4 - contains all trading datasets EXCLUDING delisted data

Neural networks must be trained on data which includes delisted securities, to enable the neural network access to data which describes the real world environment. The datasets which exclude delisted data are only used to complement the out of sample benchmarking. This enables the thesis to observe the effect of testing trading strategies on data which contains a survivorship bias, as well as on the data which does not include survivorship bias. As previously mentioned in this thesis, there is published work which presents results which contain survivorship bias, and it would be useful to know to what extent this bias may be influencing published results. For this reason, then, all strategies will provide out of sample benchmarks with survivorship bias, and without survivorship bias.

From a trading point of view, the trader cannot benefit from survivorship bias, so the results for the non-survivorship case will represent the worst case results, and will be directly applicable to the trader.

Finally, as well as being used to provide training material for the neural networks developed, the in-sample data will be used to derive appropriate trading system parameters, such as determining stop-loss levels.

3.3 Strategies

3.3.1 Fundamental Screening Strategies

There are many studies referenced in the literature review which focus on using different fundamental variables to screen stocks. Each of these studies presents a different perspective on this issue. A choice was made to focus on two different studies for this thesis. It was decided to focus on Graham's 6 point screening strategy, and Aby et al.'s strategy. This allows for a broad overview of the literature, as it includes a variety of fundamental variables many of which are re-used in other studies in the literature review, and it also covers a lengthy cross-section in terms of academic research time.

Each of these strategies is described in the appropriate section below, and relevant data issues concerning the strategies are also identified. The fundamental variables necessary to perform the fundamental screening strategies provides the input to the neural networks.

3.3.1.1 Graham's Undervalued Stocks

This study implements the fundamental strategy described by Lowe (1994) in the book, 'Benjamin Graham on Value Investing', with the exception that bonds are not used to balance the portfolio weightings. In this sense, it implements Graham's trading strategy as a stock screening strategy. This strategy was dictated by Graham at age 82, and is widely believed to represent his final thoughts on stock selection.

In essence, the strategy involves buying shares with:

- An earnings yield not less than twice the AAA bond rate for industrials,
- A dividend yield of not less than two-thirds the average AAA bond rate for industrials,
- A price not greater than two-thirds the tangible book value

Additionally, two of the following three criteria must apply:

- A current ratio of at least 2,

- A total debt less than a company's tangible net worth
- A total debt less than twice the companies net current assets

A suitable proxy is used in place of the AAA bond rate for industrials for the Australian market. This study uses the yield on 10-year Commonwealth Treasury Bonds as published by the Reserve Bank of Australia (F02 Capital Market Yields – Government Bonds) (2003).

Each of the individual data items required to determine each of the six 'points' was acquired from Aspect Financial. The values of each of the six points was then computed as either:

- -1: The point could not be computed (insufficient or missing data, division by zero, etc)
- 0: The point was computed, and did not qualify. For example, in point 1, a zero was used to represent the fact that the Earnings Yield was less than twice the AAA bond rate for industrials (actually, the proxy)
- 1: The point was computed, and qualified. For example, in point 1, a 1 was used to represent the fact that the Earnings Yield was NOT less than twice the AAA Bond Rate for industrials (actually, the proxy again)

These six point results were then merged into the trading datasets at the relevant dates as previously described, along with the actual values of each of the variables used to compute each point.

Occasionally, the values for some fundamental variables were missing in the Aspect Database. There are a variety of reasons for some of the data being missing, most often it is due to the fact that companies only disclose what they are required to by law and by generally accepted accounting practice (GAAP). Essentially, the requirements of what must be disclosed change from time to time, and this will impact on any study which uses historical fundamental data. This missing data problem is typical of all studies of this nature, and is not a criticism of the Aspect Financial database. Indeed, this issue occurs regardless of datasource, for example, Halliwell et al. (1999) found a similar problem when using the Australian STATEX files provided by the Australian Stock Exchange.

The effect of this missing data in this study is to render the relevant point setting to -1.

3.3.1.2 Aby & Briscoe Undervalued Stocks

This study also implements the screening process described by Aby et al. (2001), which uses fewer screening variables, specifically:

- P/E < 10,
- MP < BV (market price less than book value),
- ROE > 12,
- DPR < 25 (Dividend payout ratio less than 25%)

Each of the individual data items was acquired from Aspect Financial, and was merged into the datasets at the relevant dates as previously described.

3.3.2 Technical Strategies

As described by the Securities Institute of Australia (2003), one of the basic assumptions underlying modern technical analysis is that ‘prices are not entirely random, and move in trends for significant periods of time’. Traders attempt to exploit this assumption, by developing strategies which allow them to identify and capitalize on trend movements. Indeed, according to Pring (1999), ‘the art of technical analysis is to try to identify trend changes at an early stage, and maintain an investment or trading posture until the weight of evidence shows or proves that the trend has reversed’.

The focus on technical analysis in this thesis concerns the assumption that price movements may not be entirely random. If this is so, then it should be possible to train a neural network to identify situations when there is a greater likelihood of a successful trading outcome. The assumption regarding randomness essentially describes the market concept of ‘efficiency’. As an ‘efficient’ market means that prices (returns) are not predictable from past information, this assumption effectively states that security prices are only impacted by new information. By definition the arrival and timing on new information is unpredictable. Therefore, in an efficient market, security prices appear to be generated randomly. It is this conclusion which leads to the term 'random walk'.

Technical analysis also lends itself to short term trading models, typically it is trying to identify imbalances in the spectrum of supply and demand, and these imbalances are usually arbitraged away with reasonable speed. For this reason, predictions made using technical analysis usually have a short life span, and the trading opportunities they represent must be quickly exploited.

The technical variables deemed to affect returns provide the input to the neural networks. These technical variables are discussed in detail is section 3.7.1.4.

3.4 Trading Considerations

As previously mentioned in section 3.1, there are three major components to a trading system. To implement each of these components, there is a need to make choices and decide on the values for various parameter settings. Clearly, these choices can only be made by reference to the in-sample data.

In this thesis, neural networks have been trained to signal trading opportunities. The output of each of the neural networks is a signal strength value between 0 and 100. There is a need to determine the signal strength threshold that must be exceeded to initiate a trade. In this thesis, function profiles are used to determine signal thresholds. This process is described in detail in section 3.4.2.

The second component, risk control requires a choice of the type of trading stop used, and the selection of an appropriate risk threshold. Again, this can only be chosen by reference to the in-sample data. This process is described in detail in section 3.4.3.

For the third component, there is no need to derive parameters for money management settings, as the money management parameter was fixed at 1% of equity for this thesis. This value could have been determined by implementing a Monte-Carlo simulation against the trades taken from the in-sample testing metric. It was chosen not to do this, however, as the need would then arise to separate the effects of the contribution of the money management approach chosen, from the effects contributed by the neural networks. Issues related to money management are discussed further in section 3.4.4.

Finally, there is a need to determine the overall amount of capital available for running the simulations, and the level of transaction costs to include. The manner in which these values were chosen is described in section 3.4.5.

3.4.1 Constraints

As previously mentioned, the study is confined to the Australian sharemarket. The study is also confined to the 10-year period from start of trading in 1994, to end of trading in 2003. The study is further confined to studying only Ordinary shares. For the study of the S&P/ASX200, it is appropriate that a security should only be acquired if it was a constituent of the index or proxy at the time it was selected for purchase, but it may be sold at any time as required.

The trading systems used in the study are all 'long' systems. That is, the systems implemented only trade long positions, and do not attempt to sell stocks short. The reason for this is that only certain securities can be sold short, and these can only be sold short at certain times. The method the ASX uses to determine which shares can be sold short effectively prevented past replication of the method to determine what would have been eligible at any point in time. In effect, the ASX determine the list of Approved Securities (for short selling), and allow not more than 1/10th of the total quantity in issue of eligible securities to be short sold. The shortsell list is updated every trading day, based on the shortsell list for the previous trading day, as described by the Australian Stock Exchange, in the Short Sales document (2004).

The basic guidelines for determining approved securities for short selling, as stated by the Securities Institute (2003), are:

- The stock must have at least 50 million shares on offer,
- The stock must have a market capitalization of at least $100 million,
- The ASX must be satisfied with the level of liquidity.

As a brief aside, the introduction of CFDs (Contracts for Difference) has alleviated much of the difficulty related to short selling, and as long as a CFD market exists for the required security, then it can be sold short. This offers the ability to open up trading strategies for short selling, and this issue will be further considered as future work.

All trades initiated are day+1 long market orders. This means that after a signal is given, then the trade takes place on the next day the market is open, at market open price. For example, after the market has closed on day t, the trading system is run, and any buy (sell) signals generated are queued for opening positions (closing positions) for the start of the next days trading, day $t+1$. In this way, there is no possibility of acting on information which is not publicly available to all traders.

3.4.2 Determining Buy/Sell rules

In this thesis, we are interested in demonstrating that neural networks can be used to develop trading systems with excess returns. Each neural network developed has fit itself to the characteristics of the market which the training data represents. From inspection of the function profiles for each neural network over the in-sample data, the threshold at which the neural network output signal begins to signal profitable trades can be easily established. From the function profile for each neural network trained (presented in Section 4), again over the in-sample data, it can be clearly seen that the neural networks output signal increases (decreases) as expected return increases (decreases).

Therefore, the buy signal should take account of the individual neural networks threshold, and also take account of whether the signal is increasing in strength, or decreasing in strength from its previous forecast. Naturally, the sell signal should also take account of the threshold, and also take account of whether the signal is increasing in strength, or decreasing in strength from its previous forecast. It is also considered a desirable property of a trading system if the rules for exiting a trade are the contra to the rules for entering it.

Therefore, a general buy and a general sell rule can be explicitly stated, and then applied to each trading system. Where x is the signal strength threshold chosen, then these rules become:

Buy: Buy tomorrow when neural signal output(today) > x, and neural signal output(today) > neural signal output(yesterday)

Sell: Sell tomorrow when neural signal output(today) <= x, and neural signal output(today) < neural signal output(yesterday)

These simple buy and sell rules take account of the threshold and signal strengths, and using the same generic buy and sell rule for each network gives greater confidence of the generalization of the results.

3.4.2.1 Identifying Signal Thresholds

For each neural network, the output is a signal strength rating, scaled between 0 and 100. It is expected that as the numeric value of the signal increases, so should the expected returns to this signal strength. This general principle can be seen by examining a function profile of the signal output of each neural network. The function profile shown for each network graphically illustrates the breakdown of the output values of the neural network (scaled from 0 to 100) versus the average percentage returns for each network output value. The percentage returns are related to the number of days that the security is held, and these are shown as the lines on the function profile graphs (presented in Section 4).

Put simply, the function profiles visualize the returns expected from each output value of the network and shows how these returns per output value vary with respect to the holding period.

3.4.3 Risk Control

3.4.3.1 Investment risk

In the context of stock market investment, the issue of risk has many facets, as an investor may be potentially exposed to a great number of risks. Many of these risks may indeed be unquantifiable, due to the very nature of the investment. According to Spaulding (1997), risk management is an area that is filled with alternative viewpoints, and the investment industry is still trying to agree on the proper way to measure risk. According to Spaulding, standard deviation of portfolio returns is a widely used and accepted measure of dispersion and risk. It may be presented on its own, or more likely it is presented in the context of the Sharpe ratio (see section 3.5.3).

According to Yohannes (1996), the Sharpe ratio is appropriate for use if the portfolio being measured represents the investors only investment. The alternative, the Treynor ratio is appropriate if the investor has substantial other investments in addition to the portfolio being measured. The approach of using the Sharpe ratio is used in this thesis for two main reasons. Firstly, it is considered that the majority of traders' trades would be made through the trading portfolio. This approach is similar to that used by large investment houses, which regularly publish Sharpe ratios for investors to use as a comparison to other investments. Secondly, the calculation of the Treynor ratio involves the calculation of beta, which is, in itself, a controversial variable. This is because there are no uniformly agreed on procedures for beta estimation, and the actual correlation between beta and expected return is still in question. There is also significant disagreement amongst academics regarding the use and interpretation of beta (see for example, Reinganum (1988), and Strong (1988)). For the interested reader, a brief summary of the criticisms of beta are provided in Yohannes (1996), others are mentioned in Thomsett (1989) and Vince (1990).

In summary, the main choice between the Sharpe ratio and the Treynor ratio involves whether the best proxy for risk is the standard deviation of portfolio returns, or whether it is beta. Considering the above discussion, the Sharpe ratio is the preferred measure in this thesis, and it is calculated and presented for all benchmarked portfolios. Finally, as noted by Lin (2005), of all the available risk measures, an investor is most likely to be confronted by Sharpe ratios.

3.4.3.2 Trade risk

In the context of stock market trading (as opposed to investing), a trader is typically concerned with downside risk, which describes how much money is at risk on an individual trade-by-trade basis. This method of approaching risk leads to traders placing orders to sell/buy securities to cover open long/short positions when losses cross pre-determined thresholds. These are known as stop-loss orders.

As investors are typically preoccupied with return, it is also appropriate in this thesis to consider risk to be appropriately controlled by trade risk within the confines of a trading system. After all, this is the entire purpose of a trading system. This method of considering risk is growing in popularity, see for example Kaufman (1998), Longo (1996), and Pocini (2004).

A general framework for considering the issue of risk control is the TOPS COLA approach described by Chande (1997). TOPS COLA is an acronym for "take our profits slowly, cut off losses at once". In effect, it describes the traders approach to risk.

Risk control (as used within this thesis) may therefore be defined as the process for managing open trades using predefined exit orders.

Trend following systems will typically have more losing trades than winning trades. In financial terms, this still leads to a viable system, as long as the value of losing trades is quite low, and/or the value of winning trades is high. Typically, according to Chande, about 5% of the trades made by a trend following system are the 'big ones'. In light of this information, it is easy to see how the TOPS COLA approach can work.

Trading risk is normally contained use stops. These are orders to the market to exit an open trade at a given price. Traders enter an initial stop-loss when they first enter a trade. If the price of the security falls, then a stop-loss is triggered and a loss is taken. Clearly, the closer the stop is to the actual price of the security, the less money will be lost if the price falls, but the more likely it is that a small price fluctuation or random noise will trigger the stop. If the stop price is further away from the security, then there is potentially more money at risk, however, the chance is lower that the stop will be hit by noise. Chande provides evidence to suggest that the use of tight stops may well be degrading long term portfolio performance.

If the price of a security rises, the trader may well adjust the stop loss to a break even stop, or a trailing stop. A break-even stop will cancel out the trade if the price falls, at an amount equal to the trade cost. A trailing stop will increase in value as the security price increase, either proportionately or otherwise. The trailing stop will never move down, it is only ever moved up under the security price, so once the price of the security falls, the stop will be hit.

There are a number of variants of the stops described above, and stops may also be set relative to volatility, or period of time elapsed, or combinations of the above. There are also some specialist stop setting techniques, such as those described by Kase (1996). Also, the use of profit targets is not uncommon. These effectively exit a trade once a predefined amount of profit has been achieved. There are variants of profit exits also, but, in general, they run contra to the principle of "take our profits slowly".

When trading using longer term classical bar charting in the market, it is not uncommon for traders to use very simple stops, typically setting a basic money management stop on the initial purchase, and leaving that stop in for the duration of the trade. The trade entry is determined by the detection of a trend, and the trade exit is controlled by the evidence that a trend has reversed, or ended. This is entirely consistent with the definition provided by Pring (1999), and the use of stops is reserved for exiting out of trades where a trend appeared to begin, but very quickly ended. Finally, Kaufman (1998) shows how the performance of a long-term trending strategy without stops is most consistent, and that the use of fixed value stops may conflict with the strategies objectives.

As can be seen, risk control for a trader concerns protecting open trades where money is at risk through the use of stops. All strategies in this thesis will use stops for risk control. In essence, although there are many styles of stop described above, the use of stops in this thesis is simply to be able to site the neural networks developed within a realistic trading environment. For this purpose, an initial stop-loss will suffice.

The stop-loss threshold is selected by the study of the in-sample MAE as described by Sweeney (1996), and later by Tharp (1998). The MAE studies the Maximum Adverse Excursion (MAE) of a set of trades, in an effort to determine the extent to which favorable (profitable) trades range into unprofitable territory before closing out profitably. This method of risk management allows traders to study the MAE characteristics of a set of trades, to identify preferred stop-loss points.

In this thesis, the MAE technique is used to identify an appropriate stop-loss percentage for the in-sample set of trades. This stop-loss percentage is then used to control trading risk for the out-of-sample trades.

By building a histogram of this data, split according to trades that eventually won (were profitable), and trades that eventually lost (were unprofitable), a visual inspection can be made of a useful stop threshold. This information is very valuable to a trader, as it also gives an indication of how the profit/loss percentages will be affected when the stop is introduced. In this thesis, the stop percentage value determined from the in-sample data will be then used as the stop value in the out-of-sample testing data.

Typically, the exact value chosen is selected by “eyeballing”; what is required is to locate the point at which the number of ‘winning’ trades falls away very sharply, whilst typically the number of ‘losing’ trades does not. According to Chande’s principles, if in doubt, we should err towards selecting a larger value of stop loss than a smaller one. This gives a trade plenty of room to develop a profitable outcome.

3.4.4 Money Management

Money management concerns the actual size of the trade to be initiated, with consideration of account equity and potential for trade risk. It is extremely important for a trader to scale trading positions relative to available capital, indeed, according to Silber (1993) and Gehm (1995), only those who do this can have the staying power necessary to succeed.

Effective money management is related to the risk of ruin, which can help determine the amount of capital needed to trade a given system. Effectively, the risk of ruin increases as the amount per position increases, and decreases as the probability of winning increases, or as the payoff ratio increases.

As every trade carries a potential for loss, there is a need to determine the maximum amount of capital to expose at each trade, given a finite capital base. There are a number of approaches in common usage. For example, the traditional Kelly system, (as described by Balsara (1992)) evaluates the fraction of capital to expose, f, as the difference between the probable expected win and the amount expected to lose. Other developments include Vince's optimal f (1990).

General rules for traders are also provided by market technicians. For example, Elder (1993) and Pring (1999) amongst others, suggests exposing no more than an absolute maximum of 2% of account equity per position.

The issue of money management is a complex one, and it is only relevant here as the goal of the thesis is to ensure that strategies created are tradeable. For this reason, a simple, highly conservative form of money management was selected, namely, the use of 1% of account equity per trade. Not only is this simple to implement, but it also avoids having to determine how much of any effect observed is attributable to the neural networks developed, and how much is attributable to a more obscure form of money management. The reason a value of 1% was chosen rather than the 2% referred to above was firstly, to be conservative, and secondly, as the 2% is suggested as an absolute maximum. Therefore, choosing 2% cannot be considered conservative.

Finally, it should be noted that a number of advanced methods have been used in recent years to determine the optimum money management settings for a given strategy. The most promising of these is Monte-Carlo analysis, where the in-sample trades can be analyzed according to a variety of different money management settings, and a probability function can be established to determine the likelihood of success and failure dependant on the money management settings selected. This is also a fascinating area for my future research, however, it is not pursued in this thesis for the reasons outlined above.

3.4.5 Other Trading Parameters

To enable simulations to take place, there is a need to decide on an appropriate amount of starting capital, and also, an appropriate level of transaction costs. For starting capital, the amount of $100,000 is chosen, as this value represents the value of direct investment of the equal-largest proportion of direct share owners in Australia (20%), as reported by the Australian Stock Exchange (2004).

It is also necessary to select an appropriate transaction cost model. A large number of papers reported in the literature review did not account for transaction costs at all, as discussed by Vanstone et al. (2003). This clearly unrealistic situation greatly exaggerates reported results, particularly for trading systems with large numbers of transactions. For specific examples of the effects of transaction costs on trading systems, see Knez and Ready (1996).

When transaction costs are taken into account, many systems can quickly become unprofitable. Indeed, according to Barber & Odean (2000), it is the cost of trading and the frequency of trading, not portfolio selections, that explain the poor investment performance of most investors. Occasionally, some papers in the literature use a simple model by setting transaction costs equal to 1% of trade size (each way). As this appears to be the main model in usage, this thesis will also adopt the same model. For examples of previous works which use the 1% model, see Chan and Foo (1995), Yao and Poh (1995), Liu and Lee (1997), Dourra and Siy (2002), Chenoweth et al (1995) and Thawornwong et al (2002).

3.5 Evaluation Metrics

3.5.1 Introduction

Trading systems are typically evaluated and compared by traders using traders metrics. These are a set of commonly accepted metrics used by traders, each designed to focus on some particular aspect of trading system behaviour. Further, trading systems can be evaluated and compared using statistical methods. These approaches are both used in this thesis, and the details of the relevant metrics and their significance is described in the following sections.

Whilst these metrics are suitable for evaluating completed systems, and are ideal for evaluating out of sample results, they are not helpful when attempting to determine whether a particular neural net configuration is acceptably trained. For this reason, it is necessary to develop further metrics which can be used in-sample.

The discussion of the purpose and calculation of the in-sample metrics is delayed until after the discussion on training neural networks, and is presented in section 3.7.2.2.

3.5.2 Trading System Metrics

A primary objective of a trading system is to produce profits. However, in itself, this is an unsuitable benchmark for a variety of reasons. The desire to produce a profit must be tempered with considerations such as trading risk, equity curve management, amount of capital required, drawdown, and consistency. These factors combined dictate how 'tradeable' a system would be in practice.

Trading systems are typically assessed according to a variety of metrics. The metrics presented in Table 3-5 are sourced from Babcock Jr (1989), Chande (1997), Ruggiero (1997), Pardo (1992), Kaufman (1998), Tharp (1998), Vince (1990) and Jurik (1999). Table 3-5 presents a common set of these metrics, along with a brief statement describing the metric, and, where relevant, discusses the desired or acceptable range of the metric. Finally, where relevant, the formulas used to calculate the metrics are also presented. Some metrics do not lend themselves to simple formulaic representation, and in these cases, the mechanics of the calculation are explained. Also, as shown in the table below, some metrics are relevant for All Trades, while some are only relevant for all Winning Trades or all Losing Trades. For the purposes of this thesis, a Winning Trade is defined as a trade with a net return greater than zero after trading costs have been taken into account.

It should be remembered that the factors which decide whether a system is acceptable or not are ultimately the choice of the trader. No system should be chosen if it displays undesirable characteristics; however, individual traders would differ on their choice of system, dependant on such issues as their tolerance to risk, their amount of starting capital, and their trading horizon.

Metric	**Brief Description**
Net Profit (Loss)	Total profit (loss) including open positions marked to market with latest closing price for simulation period. Transaction costs are taken into account.
Net Profit (Loss) %	Net Profit (Loss) expressed as a percentage of starting capital.
Annualized Gain %	Net Profit (Loss) divided by the number of annual periods. Also known as Annual Percentage Return (APR).
Exposure %	Market exposure is the actual area of the portfolio equity that was exposed to the market, as calculated on a day by day basis.
All Trades: Number of trades	Total number of round trip trades (open, then close) plus the number of open trades.
All Trades: Average Profit (Loss), Average Profit (Loss) %	Average Profit (Loss) per trade, expressed in dollar terms, and as a percentage.
All Trades: Average Bars Held	Average number of days that a trade is open.
Winning Trades: Number of Trades	Number of Winning Trades.

Metric	Brief Description
Winning Trades: Winning %	Percentage of trades that were winners.
Winning Trades: Average Profit, Average Profit %	Average Profit per winning trade. Includes the effect of trading costs, and does not take open positions into account. Expressed in dollar terms, and as a percentage.
Winning Trades: Average Bars Held	Average number of bars that winning trades are held for.
Winning Trades: Max consecutive wins	Maximum number of winning trades in a row (also know as winning streak).
Losing Trades: Number of Trades	Number of Losing Trades.
Losing Trades: Losing %	Percentage of trades that were losers.
Losing Trades: Average Loss, Average Loss %	Average Loss per Losing Trade. Includes the effect of trading costs, and does not take open positions into account. Expressed in dollar terms, and as a percentage.
Losing Trades: Average Bars Held	Average number of bars that losing trades are held for.
Losing Trades: Max consecutive losses	Maximum number of losing trades in a row (also know as losing streak).

Metric	Brief Description
Max Drawdown, Max Drawdown %, Maximum Drawdown Date	Largest peak to valley decline in the equity curve, on a closing price basis. Reported as both a dollar amount and as a percentage. Date of maximum drawdown is also reported.
Profit Factor (PF)	Used to demonstrate how profitable a system has been historically. A desirable value for this metric is 2 or above, indicating the system has won twice as much as it has lost. $PF = \frac{(W\% \times AW)}{(L\% \times AL)}$ or $\frac{Gross\,\mathrm{Pr}\,ofit}{GrossLoss}$ where W% = Percentage of Winning Trades AW = Average Winning Trade Amount L% = Percentage of Losing Trades ($(1 - W\%)$ LW = Average Losing Trade Amount GrossProfit = Total Amount Won GrossLoss = Total Amount Lost
Recovery Factor (RF)	The Recovery Factor indicates whether a trading system can overcome drawdown effectively. A desirable value for recovery factor is greater than 1. $RF = ABS\left(\frac{NP}{MaxDD}\right)$ where NP = Net Profit MaxDD = Maximum Drawdown

Metric	Brief Description
Payoff Ratio (PR)	The Payoff Ratio is an indicator of how well a system earns profit relative to loss. The higher the payoff ratio the better. Typically, a Payoff Ratio greater than 2.0 is desired. $PR = ABS\left(\frac{AVGW}{AVGL}\right)$ where AVGW = Winning Trades: Average Profit % AVGL = Losing Trades: Average Loss %
Sharpe Ratio (SR)	The Sharpe Ratio measures risk adjusted return, and is also known as the Reward-to-Variability Ratio. Specifically, the ratio indicates the historic average differential return per unit of historic variability of the differential return. Sharpe (1994) provides a detailed discussion of the limitations and uses of the Sharpe Ratio. The traditional formula for calculating the Sharpe ratio is presented below. $SR = \frac{(PR - RFR)}{PSD}$ where PR = Portfolio Return RFR = Risk Free Return PSD = Portfolio Standard Deviation In this thesis, the portfolio return and portfolio standard deviation are annualized by using the average number of days held per trade as a baseline. The annualized average return is then divided by the annualized standard deviation of returns. Note: There is no RFR term used in this thesis. As this thesis studies only the trading return due to price movement, and does not change the cash balance of a portfolio for any other reason, the Sharpe Ratio calculation assumes a zero risk-free rate of return.

Metric	Brief Description
Ulcer Index	Ulcer Index measures overall volatility of a portfolio. It is calculated by the square root of the sum of the squared drawdowns.
Luck Coefficient (LC)	Shows how the largest trade compares to the average trade. Calculated by dividing the percentage profit of the largest winning trade by the average percentage profit of all trades. $LC = \frac{MAXW}{AVGW}$ Where MAXW = Percentage Profit of largest winning trade AVGW = Winning Trades: Average Profit %
Pessimistic Rate of Return (PRR)	An adjustment of the wins to losses ratio for the purpose of estimating the worst expected return based on previous results, as suggested by Vince (1990). According to Vince, a value greater than 2 is good, a value greater than 2.5 is excellent. Pessimistic Rate of Return is calculated by decreasing the number of winning trades by the square root of the total winners, and increasing the number of losing trades by the square root of the number of losers. The result is then computed by multiplying the new number of winners by the average amount won, and dividing this by the new number of losers multiplied by the average amount lost. $PRR = \frac{\left(\left(\frac{\left(W - \left(\sqrt{W}\right)\right)}{T}\right) \times AVGW\%\right)}{\left(\left(\frac{\left(L + \left(\sqrt{L}\right)\right)}{T}\right) \times AVGL\%\right)}$ where W = Number of Winning Trades L = Number of Losing Trades T = Total Number of Trades

Metric	Brief Description
	AVGW = Winning Trades: Average Profit % AVGL = Losing Trades: Average Loss % By comparing this formula to that shown for Profit Factor (PF) it is clear that PRR is a useful extension to calculating Profit Factor alone. Vince's original formula used Average Winning Trade Amount instead of AVGW%, and Average Losing Trade Amount instead of AVGL%. However, using the percentage instead of the amount is now commonplace in these calculations, and as such, it has been used here.

Table 3-5 Trading System Metrics

3.5.3 Statistical Measures

The traders approach to benchmarking trading systems using trading system metrics is an excellent one. It should be noted, however, that it is possible to add extra insight into the expected performance of a trading system by performing the students t-test.

Typically, the approach is to perform a students t-test to determine the likelihood that the observed profitability is due to chance. This is the method recommended by Katz (2000), Katz and McCormick (1997), Knight (1993), Chande (1997), Stakelum (1995), and Kaufman (1998).

The students t-test shows whether there is a systematic bias in the data by showing whether the mean of the net profits is significantly different from zero.

The use of the t-test relies on assumptions of normality and independence. Essentially, these assumptions are constraints upon the usefulness of the t-test in evaluating trading systems.

Typically, the assumption of normality is dealt with by reference to the Central Limit Theorem, which indicates that as the number of cases in the sample increases, the distribution of the sample mean approaches normal. Consequently, as long as the sample size is adequate (generally stated as at least 30 samples), the statistic can be applied with reasonable assurance.

The constraint of independence presents a more difficult issue when comparing trading systems. Essentially, the violation is potentially one of serial dependence, which occurs when cases constituting a sample are not statistically independent of one another. One method of dealing with this issue is to perform a runs test, as described by Vince (1990). The runs test shows whether the sequence of wins and losses in the sample trades contains more or less streaks than would ordinarily be expected in a truly random sequence, which has no dependence between trials. Although a runs test does not prove or disprove dependency, it can be used to determine an acceptable confidence limit in order to accept or reject a hypothesis of dependency. Vince suggests exceeding 95.45% (2 standard deviations) to accept that there is dependency involved. Vince demonstrates the runs test is essentially a matter of obtaining the Z scores for the win and loss streaks of a systems trades, as follows:

$$\text{Z Score} = \left(\frac{(N*(R-0.5))-X}{\sqrt{\frac{(X*(X-N))}{(N-1)}}} \right)$$

where

N = total number of trades,

X = 2 * total number of wins * total number of losses

R = total number of runs in a sequence

Equation 3-1 Computing the Z-Score for a Runs test

3.5.4 Comparing Trading Strategies

The main advantage of calculating trading system metrics is that trading systems can be directly compared on the basis of their trading metrics. Additionally, trading systems can be compared on a statistical basis by performing the independent-measures single factor ANOVA procedure, which effectively quantifies between and within-group variations.

In this thesis, trading systems are compared using the ANOVA procedure, and where a difference is found to be significant, it is explained using the trading system metrics outlined earlier.

The ANOVA procedure is preferred to t-tests as t-tests are limited to situations where there are only two samples to compare. ANOVA can be used to compare two or more samples. ANOVA can also be used to compare samples with unequal sizes, although greatest accuracy is achieved with equal or close sample sizes. However, in testing different trading strategies, it is highly unlikely that sample sizes would be equal, and ANOVA still provides a valid test, especially when the samples are relatively large. ANOVA is also robust to moderate departures from normality. Also, ANOVA is preferred to multiple t-tests, as each t-test contains Type 1 error, and repeatedly performing t-tests will compound the effect of Type 1 error. This deficiency is not present in the ANOVA procedure.

Finally, ANOVA is used in this thesis to compare the results of the ANN based trading systems to the results obtained from buy-and-hold simulations, for each respective ANN based trading strategy developed. These same comparisons could have been made using the Independent Samples t-test.

3.6 Software & Hardware used in this study

Trading algorithms in this study are implemented using WealthLab Developer (Version 3), and its associated partner tool, NeuroLab (for WealthLab Developer 3). Algorithms are programmed directly into the Wealthlab tool using a language called Wealthscript, which is a programming variant based on Pascal and Delphi.

A great many individual programs were also written outside of Wealthlab, to implement the data transformations, calculations and merges used for both fundamental strategies, and also for determining the 'tradeable proxy' index. These were written in Visual Basic using Microsoft Access as the host database.

A stand-alone PC with a Pentium 4 (2.8 GHZ) processor and 2GB RAM were used to perform all neural network training and testing. Using this high-end configuration, each network took approximately 3 days to train.

3.7 Neural Network Considerations

3.7.1 Inputs

An initial consideration of the inputs to the neural network concerns whether to use raw data, or pre-processed data, as the input to the transformation process required by the neural network. A number of approaches in the literature process the input data using natural logs, or ratios of natural logs before supplying it to a neural network. There appears to be a move away from this in some financial-related literature. In early work, Azoff (1994) suggested that the use of direct price values (and raw data input) is preferred to price differences, to prevent destruction of fragile structure inherent only in the original time series. This was confirmed by Longo (1996) who found significantly better results were achieved from the neural networks when using raw data as opposed to transformed data.

The Neurolab tool used in this study allows the inputs to the neural network to be script coded. This script code effectively forms an interface between each input node, and its relevant data field. Following the work of Azoff and Longo, this study will use the raw data values of inputs, and transform them for input into the neural network using a simple linear interpolation between the lowest and highest values in each series to achieve an input set in the range 0 to 1. This method is robust, and quite simple, and is widely used (2000). Outliers were removed to prevent skewing the outcome of the interpolation.

3.7.1.1 Fundamental timeframe

The fundamental timeframe, or the prediction look ahead period is 1 year, chosen primarily as this is the usual timeframe for the production of the fundamental variables. This long timeframe leads to the development of longer-term trading systems. This look ahead period is typical of studies of this nature, see for example, Longo (1996) and Reinganum (1988).

3.7.1.2 Selection of Fundamental Variables

In terms of screening strategies, this study focused on the work of Graham, and Aby et al. The basic screening strategies described by their research are tested in the Australian market. Thus, the selection of fundamental variables which will be inputs to the fundamental neural network are those fundamental variables that were required for the above strategies. The minimum list of fundamental variables required is augmented with:

- the ASX200 indicator, to enable the neural networks which are to be trained on S&P ASX200 data to differentiate between S&P ASX200 data and Allshare data, and
- the Year End Share Price, to enable the neural network to assess fundamental variables like Earnings per share in context

Fourteen variables are required for the fundamental data based ANNs, these are listed below:

- P/E Ratio
- Book Value per share
- ROE
- Payout Ratio

- Dividend Yield
- Price to Book ratio
- Total Current Assets
- Total Gross Debt
- Weighted Average Number of Shares
- Current Ratio
- Earnings per Share
- Year End Share Price
- ASX200 indicator
- AAA return proxy

3.7.1.3 Technical Timeframe

The issue of the required timeframe for the technical variables is a complex one. The general traders approach is to use data from a combination of timeframes, generally two or occasionally three different timeframes. For example, two different timeframes could be used, one timeframe based on 2-3 days, and another based on 2-3 weeks. Indeed, Elder (1993) even suggests three timeframes, with the third timeframe being in months. The main purpose of using more than one timeframe is to try to isolate changes in the shorter term timeframe, and understand whether they are consistent with changes in longer timeframes. It is believed by traders that short term changes that are supported by long term changes are more valid as trading opportunities. This is the reason why many moving average systems use two timeframes, essentially generating buy strength signals when a shorter term moving average crosses from below to above a longer term moving average, and the reverse for exit signals. The expectation is that as the shorter term moving average is leading the longer term moving average, early entry into a potentially long term movement has been signaled.

Specific timeframe lengths could also be determined by searching for cycles, but as already stated earlier, it is generally accepted that the market dynamics are continually changing over time. Thus, there appears little benefit in determining the optimal timeframe at any point in time, as it could not be relied on to hold, thus reducing the expected lifetime of a trading strategy, and adding additional risk (see for example, Balsara et al (1996)). Published work exists which describes cycles already found in the Australian stockmarket, such as a 6-day cycle discovered by Pan et al. (2005), however, this was discovered in the AORD index, and there is no reason to expect it would hold for large numbers of individual stocks, each with its own individual characteristics. For example, it could be expected that certain stocks had radically different cycle lengths, such as bank stocks following the economic cycle, resource stocks following the strength of the industrial cycle in other countries, etc. Other published works use PCA (Principal Component Analysis) (e.g. Raposo et al. (2004)), or Self Organizing Maps (e.g. Chi et al. (2003)).

Essentially, the practical solution, and indeed, the objective, is to select timeframes which are consistent with the traders trading expectations. This idea is consistently presented in literature describing the techniques involved in building trading systems, for example, Ruggiero (1997).

Technical trading strategies are usually short-term, and a number of the articles discussed in Table 2-1 studied 2 and 3 day patterns which occurred within stock prices. Further, it seems logical to relate the choice of the higher order timeframes to the underlying (lower order) timeframes is some obvious way, as was done by Elder, and is usually done in the selection of parameters for moving average systems (for example Baba et al. (2003)).

For experimentation purposes, and because some practical choices must be made, it is proposed to make the following selections:

Selection	Description
Number of timeframes used = 2	There will be two timeframes used for all technical indicators submitted to the technical neural network.
Short Timeframe = 3 days	The period of the short timeframe is 3 days.
Long Timeframe = 3 weeks	The period of the longer timeframe is 3 weeks (actually 15 trading days).

Table 3-6 Selection of Technical Variable timeframes

Thus, all technical variables input to the technical data based ANNs will be computed for both the shorter and the longer timeframes. As the objective using technical analysis is to develop short term trading strategies, the look ahead prediction period will be set to the period of the shortest timeframe, 3 days. According to Jaruszewicz and Mandziuk (2004) this short time frame fits well with the use of technical indicators and oscillators, in particular they stated that 'each oscillator is useful only in a short period of time'.

3.7.1.4 Selection of Technical Variables

The selection of technical variables for input to the technical neural network is influenced by the technical variables discovered during the formal literature review, and is also influenced by those discovered in the studies of the main practitioners journal, as described in Table 2-1.

It is important to note that the ANNs do not have visibility of the actual prices or volume, only ratios and indicators built from the prices and volume. This will prevent the ANNs focusing on whether a stock has a high price (or volume) or a low price (or volume), and allow more for generalization about the ratios and indicators, and their relative relationships. As the goal with neural networks is to encourage generalization, supplying ratios to the neural network is an ideal way to accomplish this. This is because the same ratio can be built from any number of numerators and denominators, and it is unlikely (although not inconceivable) that the exact values of the numerator and denominator are relevant. Technical analysis generally focuses on ideas concerning price and volume behaviour, not the exact price and volume themselves.

As some of these variables used are technical indicators, the relevant formulas used in their construction are provided, as detailed in Bauer (1998), LeBeau and Lucas(1992) and by the Australian Securities Institute (2003). Although it would be possible to train a neural network using only the raw data used in the construction of these indicators, it would take considerably more epochs for an ANN to generate the appropriate values which can be easily computed using known formulas. Hence, it is preferable to supply the network with pre-computed values for these variables. According to Ruggiero (1997), supplying these pre-computed values is simply another form of pre-processing. Ruggiero states that three popular indicators used for neural network development are MACD, Stochastics, and ADX. Later, he also suggests RSI as a candidate. Finally, Zirilli (1997) suggests the same indicators, namely, RSI, Stochastics, MACD, and ADX. This is consistent with the information gained from the reviews of the practitioners journals, and it is clear that all of these indicators are in common usage by technical analysts.

However, one point which is often made clear in trading literature is that whilst one technical variable may appear to work well in one market environment, the same technical variable may perform poorly in another. This is often discussed in trading literature with comments like 'different markets have different personalities', and is a valid criticism of technical variables, and indeed, technical analysis. What is being expressed is that a particular indicator may work well for foreign exchange, or indeed some specific instrument, but it should be checked before being applied to any other market.

To address this concern, it was decided to create complete function profiles for all technical variables considered to establish whether the variable was likely to be useful as a neural network input. The concept of function profiles and the results and interpretation of the function profiles created as part of this thesis are considered in Appendix A.

Table 3-7 lists each classification provided in Table 2-1, and shows which technical variables were used to assess the classification. Typically, the technical variable chosen was the ratio of the shorter timeframe value to the longer timeframe value, unless trading literature appeared to refer to specific values of the variable in question. In this case, the actual technical variable values for both timeframes, as well as the ratio were profiled.

Technical Classification	**Technical Variable Profiled**
Moving Averages (of Price)	A new series created by a simple moving average of closing prices of period 3 divided by a simple moving average of closing prices of period 15. Tested as SMA(close,3) / SMA(close,15).
Volatility	ATR(3) ATR(15) ATR(3) / ATR(15)
Volume	A new series created by a simple moving average of volume of period 3 divided by a simple moving average of volume of period 15. Tested as SMA(volume,3) / SMA(volume,15).
ADX	ADX(3) ADX(15) ADX(3) / ADX(15)
Stochastics	STOCHK(3) STOCHK(15) STOCHK(3) / STOCHK(15)

Technical Classification	Technical Variable Profiled
Momentum	MOM(3) MOM(15) MOM(3) / MOM(15)
RSI	RSI(3) RSI(15) RSI(3) / RSI(15)
MACD	MACD(closing prices)
Other - Although not specifically listed as a technical variable in any of the articles, a number of trading articles concerned whether a stock was trading 'high' relative to its longer term price, or 'low' relative to its longer term price	Lowest price ratio Highest price ratio

Table 3-7 Technical variables chosen for profiling

From the function profiles that were created, the technical variables to be input to the neural network were selected. Table 3-8 provides the actual formula for each of the selected technical variables, and also provides a brief description of its nature and purpose.

Technical Input	Description
Lowest price ratio (LPR)	This new series of ratios is calculated for every day, as soon as there are at least 200 days of data available. It provides the ANN with information relating to the current price performance of this stock relative to its yearly price performance. Note that 200 days is used as an approximate substitute for 1 year, as the exact number of trading days per year varies, due to public holidays, etc. The idea behind creating this technical ratio is to enable the ANN to use the

Technical Input	Description
	current position of the price relative to its yearly low as an influence on other technical variables. $LPR = \left(\frac{lowest(close,200)}{close} \right)$ where close = the closing price lowest(close,200) = the lowest price over the last 200 days According to some traders, stocks represent better buying opportunities when the price is "low", in this case, "low" is measured relative to the previous 200 days. Proponents of this view tend to try to buy stocks at or close to sustained lows, perhaps with the expectation of mean reversion, or a quick "bounce".
Highest Price ratio (HPR)	This new series of ratios is calculated for every day, as soon as there are at least 200 days of data available. It provides the ANN with information relating to the current price performance of this stock relative to its yearly price performance. Note that 200 days is used as an approximate substitute for 1 year, as the exact number of trading days per year varies, due to public holidays, etc. The idea behind creating this technical ratio is to enable the ANN to use the current position of the price relative to its yearly high as an influence on other technical variables. $HPR = \left(\frac{close}{highest(close,200)} \right)$ where close = the closing price

Technical Input	Description
	highest(close,200) = the highest price over the last 200 days In contrast to those traders who think stocks represent better buying opportunities when the price is "low", other traders think that when a stock is approaching a new long-term high, it has better profit potential as the price may well move into uncharted high territory, thus making it difficult for buyers to objectively assess when the price is "high enough". Proponents of this view point to excesses like those experienced before the "dotcom" crash, when virtually all technology stocks were trading at all time highs, and "the sky was the limit". Traders with this belief tend to trade breakout style systems, and like to buy on evidence of price strength.
SMA(price) ratio	This series is calculated for every day, and is the 3 day simple moving average of the daily closing price divided by the 15 day simple moving average of the daily closing price. There is a great deal of support for the use of simple moving averages of price within the trading literature. Typically, traders attempt to establish long positions when a shorter term moving average crosses from below to above a longer term one, or when price moves from below to above a moving average, and establish short positions when the reverse occurs. $SMAPRICERATIO = \left(\frac{SMA(close,3)}{SMA(close,15)} \right)$ where close = daily closing price
SMA(volume) ratio	This series is calculated for every day, and is the 3 day simple moving average of the daily volume divided by the 15 day

Technical Input	Description
	simple moving average of the volume. Many traders rely heavily on volume to confirm a suspected price movement, particularly on sudden increases in volume relative to the normal volume transacted. Indeed, many chart patterns are only "confirmed" when significant volume occurs. $SMAVOLUMERATIO = \left(\frac{SMA(volume,3)}{SMA(volume,15)} \right)$ where volume = daily volume
ATR(3)	Average True Range of the last 3 days. ATR is a measure of volatility, which takes into account price gaps which have occurred in the price movement. It is smoothed with a simple moving average over the time period required. Average True Range gives the ANN an indication of whether the stock is trading with increased volatility, which is generally seen by traders as being related to increasing trading interest. Similarly, a decrease in volatility has opposite connotations. $ATR(3) = WilderMA(TR,3)$ where TR = True Range, which is the actual range, high to low of a bar. It includes any gap between the current bars high or low and the close of the previous bar. The use of WilderMA is described in the section on ADX below.
ATR(15)	Average True Range of the last 15 days, which provides the ANN with information related to volatility for the last 15 days. It is computed as described above.
ATR ratio	This new series is calculated as the ATR(3) divided by the

Technical Input	Description
	ATR(15). This provides the ANN with an indication of the relationship between volatility changes in the short term as compared to volatility changes in the longer term. $ATRRATIO = \frac{ATR(3)}{ATR(15)}$ where ATR(3) and ATR(15) are computed as described above.
ADX(3)	3 day Average Directional Index (ADX). ADX attempts to define an "average" direction for a stock, and determine the extent to which it is trending. It was first introduced by Welles Wilder (1978), and is a popular technical indicator amongst traders. However, its calculation is rather complex, as shown below. $DX = \left(round\left(\frac{(100 \times \lvert DIPlus - DIMinus \rvert)}{\lvert DIPlus + DIMinus \rvert} \right) \right)$ and $ADX = WilderMA(DX,3)$ where $DIPlus = \left(\frac{+DM}{TR} \right) \times 100$ $DIMinus = \left(\frac{-DM}{TR} \right) \times 100$ The DM is the Directional Movement for the day, which is determined by whether the larger part of the current trading range is above or below the last trading day. If it is above, it is termed +DM, and if it is below it is termed –DM.

Technical Input	Description
	The TR is the True Range, which is effectively a refinement of the difference between the low and high price for a period, as it takes gaps in trading prices into account. WilderMA is very similar to a Simple Moving Average (SMA), yet it has a nuance in its calculation method, which tends to introduce a dampening effect. It is similar to an Exponential Moving Average (EMA), yet it reacts slowly to price changes, with a n-period WilderMA giving similar values to a 2n period EMA. $WilderMA(DX,N) = \frac{(PreviousWilderMAValue \times (N-1) + DX)}{N}$ where PreviousWilderMAValue is the value of the last point computed, N = Number of Periods This indicator provides the ANN with information about the strength of trend movements over the last 3 days. This information is important as traders believe there is greater scope for price change when the trend movement is stronger. This in turn gives rise to greater profit making opportunities.
ADX(15)	15 day Average Directional Index, calculated as above except over 15 periods. This provides the ANN with information about the strength of trend movements over the 15 days, which provide depth on how long a trend movement has lasted. Clearly, traders are interested in sustained trend movements, as these offer a greater opportunity to make profit.
Stochastic(3)	The Stochastic Oscillator is a momentum oscillator originally

Technical Input	Description
	developed by George Lane. Technical analysts consider it a useful oscillator when price is contained within a broad trading range, or in a slow moving trend. The version used in this thesis is known to traders as %K, which is the 'fast' stochastic. It is extremely sensitive as %K, and is often slowed down by traders, who smooth it over 3 days using an SMA, whereupon it is known in traders literature as %D. $\%K = \left(\frac{C - L_N}{H_N - L_N} \right) \times 100$ where C = Closing Price L_N = Lowest price for N days H_N = Highest price for N days N = number of days
Stochastic(15)	This is the stochastic oscillator for the last 15 days. It provides the ANN with information about whether the momentum is increasing or decreasing in the longer time frame.
StochK ratio	This new series is calculated as the Stochastic(3) divided by the Stochastic(15). This provides the ANN with an indication of the relationship between momentum price changes in the short term as compared to momentum price changes in the longer term. $STOCHKRATIO = \frac{Stochastic(3)}{Stochastic(15)}$ where Stochastic(3) and Stochastic(15) are computed as described above.
RSI(3)	Relative Strength Index (RSI) is one of the classic technical

Technical Input	Description
	indicators, extremely popular with traders. It is a momentum measuring indicator, first introduced by Welles Wilder (1978). It measures the internal strength of the stock and is provided to allow the ANN to react to changes in price strength of a stock in the short term. $RSI = 100 - \left(\frac{100}{(1 + RS)} \right)$ where $RS = \left(\frac{AVGUP}{AVGDOWN} \right)$ where AVGUP = average of N bar up closes AVGDOWN = average of N bar down closes where N = number of days
RSI(15)	This is the RSI calculation over 15 days. It is provided here to allow the ANN to react to changes to strength in the longer term.
RSI ratio	This new series is the RSI(3) divided by the RSI(15). This provides the ANN with an indication of the relationship between short and long term relative strength. $RSIRATIO = \frac{RSI(3)}{RSI(15)}$ where RSI(3) and RSI(15) are computed as described above.
MACD	MACD is the Moving Average Convergence Divergence indicator developed by Gerard Appel. It is constructed as the

Technical Input	Description
	difference between a 26 day exponential moving average and a 12 day exponential moving average, both of closing prices.

Table 3-8 Description of Technical Variables profiled

After completing the function profile of each variable listed in Table 3-7, for both the S&P/ASX200 and Allshare in-sample datasets, the following 17 technical variables were decided on as suitable inputs for the neural network. The justification for these decisions can be found in Appendix A.

- Lowest Price Ratio
- Highest Price Ratio
- SMA(Price) Ratio
- SMA(Volume) Ratio
- ATR Ratio
- ADX(3)
- ADX(15)
- StochK(3)
- StochK(15)
- StochK Ratio
- MOM(3)
- MOM(15)
- MOM Ratio
- RSI(3)
- RSI(15)
- RSI Ratio
- MACD

The following 3 technical variables were excluded as inputs to the neural network.

- ATR(3)
- ATR(15)
- ADX Ratio

3.7.2 Processing

The neural network tool used in this study is NeuroLab (version 3). It implements the backpropogation model and uses a logistical sigmoid function (range 0 to 1) as the activation function.

3.7.2.1 Hidden Nodes and Layers

There are no standard rules available for determining the appropriate number of hidden layers and hidden neurons per layer, although for greater generalization, the smaller the number of hidden nodes and hidden layers the better. General rules of thumb have been proposed by a number of researchers. For example, Shih (1994) suggests constructing nets to have a pyramidical topology, which can be used to infer approximate numbers of hidden layers and hidden neurons. Azoff (1994) quotes a theorem due to Komolgorov that suggests a network with one hidden layer and 2N + 1 hidden neurons is sufficient for N inputs. Azoff concludes that the optimum number of hidden neurons and hidden layers is highly problem dependant, and is a matter for experimentation. Gately (1996) suggests setting the number of hidden nodes to be equal to the total of the number of inputs and outputs. As another alternative, some researchers, for example Kim at al. (2003) use a brute force approach, and train a great number of ANNs with different configurations, and then select that configuration that performed best. Yet another approach, such as that used by Jaruszewicz and Mandziuk (2004) is to train networks for a fixed number of epochs, or fixing the number of hidden nodes to some arbitrary value, as in Kim and Lee (2004). Zirilli (1997) proposes a formula based on prior knowledge of the number of unique patterns the net is expected to learn, but concedes that if you know your feature space well enough, you can usually determine this number of hidden nodes better yourself. Finally, another reasonably popular method is used by some researchers such as

Kim & Lee (2004) and Versace et al (2005), whereby genetic algorithms are used to select between the combinatorial explosion of possible networks given choices such as network type, architecture, activation functions, input selection and preprocessing.

An alternative approach described by Tan (2001), is to start with a small number of hidden neurons and increase the number of hidden neurons gradually. Tan's procedure begins with 1 hidden layer, containing the square root of N hidden nodes, where N is the number of inputs. Training the network takes place until a pre-determined number of epochs have taken place without achieving a new low in the error functionf For this study, ANNs are trained until no new low had been achieved for 2000 epochs. At this point the network is tested against the in-sample set, and benchmarked. A new neural network is now created with the number of hidden nodes increased by 1, and the training and in-sample testing is repeated. After each test, the metric being used for benchmarking is assessed, to see if the new network configuration is superior. This process continues while the networks being produced are superior, that is, it terminates at the first network produced which shows inferior in-sample results.

This approach to training is an implementation of the early stopping method, which aims to preserve the generalization capabilities of neural networks. It is based on the observation that validation error normally decreases at the beginning of the training process, and begins to increase when the network starts to over-fit. Lack of generalization is caused by over-fitting. In an over-fit (over-training, over-learning) situation, the network has begun to memorize training examples, and is losing the ability to generalize to new situations.

Tan's approach will be used for determining neural network configurations in this thesis. The metrics used for in-sample testing for the neural network trained with fundamental data, and for the neural network trained with technical data are described below.

3.7.2.2 In-Sample testing metrics

As the neural networks developed in this thesis generate trading signals outside of the context of a trading system, the test metric comparing different neural network architectures must be on the basis of their in-sample training. It is not appropriate to test each neural network architecture on the out of sample results and select the best performer.

For this reason, it is necessary to define metrics which can be used to test different neural network architectures, and the metrics presented below are focused on identifying whether the neural network has learnt correctly. For further details on what the in-sample metrics are used for, the reader may wish to review section 3.7.2.1.

3.7.2.2.1 Fundamental in-sample metrics

The neural network trained using fundamental company data will effectively serve as a longer-term screening strategy. It is designed to filter the entire market of stocks, and identify those which have the greatest chance of the highest appreciation within the next one year. The signal generated by the neural network is effectively a prediction of the likely strength of price increase over the next 1 year period, with the output signal value oscillating between zero and one hundred.

To determine how to evaluate a screening strategy, it is necessary to review the purpose of such a strategy. Specifically, a screening strategy is used to reduce (refine) the number of securities that are competing for capital. A traders' requirement is to increase the likelihood of selecting stocks that will significantly increase in value. Thus, a suitable measure of success is to determine the percentage of stocks selected that achieve a pre-specified increase in value. By measuring these values for the entire market, then for the ANN predictions, it can easily be determined whether the neural network is effective. Finally, care must be taken to ensure that a reasonable number of predictions are output by the ANN for this process. Clearly, it would not make sense to select a network with a 100% success rating if there was only 1 trade predicted.

The metric used for screening strategy measurement, termed Filter Selectivity, is defined as:

$$FilterSelectivity = \frac{(ClosedTrades * 100)}{TotalTrades}$$

where

ClosedTrades is the number of trades closed due to meeting the predefined increase in value

TotalTrades is the total number of trades selected by the screening strategy

Equation 3-2 Determining Filter Selectivity

3.7.2.2.2 Technical in-sample metrics

An objective measure of a technical short term trading system is its measure of expectancy. The idea of expectancy in trading was first raised by Tharp (1998), who proposed it as a useful method to compare trading systems. Expectancy is a measure of the expected profit per dollar risked on a fixed position size basis. It is used without money management settings enabled, which is appropriate for the in-sample tests. There are a number of variant formulas for calculating expectancy, this version presented is more conservative than Tharp's; it uses the average loss as the standard of risk (rather than the minimum loss as used by Tharp).

$$EXPECTANCY = \frac{((AW \times PW) + (AL \times PL))}{|AL|}$$

where

AW = Average amount won on profitable trade

PW = Probability of winning

AL = Average amount lost on losing trade (-ve)

PL = Probability of losing

Equation 3-3 Calculating In-sample Expectancy

Secondly, to assess the quality of the ANN architecture chosen, it is also appropriate to consider the 'Average Profit/Loss %', which is a standard trading system metric detailed in Table 3-5.

3.7.2.3 Parameters

The Neurolab tool allows the user to specify parameter settings for Momentum and Training Rate. The tool accepts values between 0.00 and 1.00 in steps of 0.01 for these settings. The implications of these parameters are briefly considered below.

3.7.2.3.1 Momentum

The momentum parameter controls how much of the previous weight adjustment is applied to the current weight adjustment. As larger values are used for momentum, the greater the influence of the current correction term, relative to previous correction terms. Momentum can be used to provide a smoothing effect for weight adjustments. An overview of the possible effects of momentum parameter changes is provided by Tan & Wittig (1993).

3.7.2.3.2 Training Rate

Training rate (also known as Learning Rate) determines the amount of the correction term applied to adjust neuron weights. A small value for training rate tends to increase learning time and decrease the probability of overshooting an optimal solution. A small value also increases the likelihood of being stuck in local minima. Large values for training rate increase the chance of no learning occurring at all. An overview of the possible effects of learning rate parameter changes is provided by Tan & Wittig (1993).

3.7.3 Outputs

Azoff (1994) suggests the network have only one output, to avoid the effect of conflicting outputs pulling the weights in opposing directions during training. In this way the network is effectively focused on one task only.

For the neural network trained with fundamental data, which is attempting to predict the return of 100% or more within a 1 year timeframe, it is appropriate that the output be a nominal scale of the strength of expected returns.

For the neural network trained with technical data, a number of choices exist. While it may seem appropriate to attempt to predict price, much of the earlier research shows that this is a particularly difficult task, possibly due to the fact that price changes do not tend to be smooth. Predicting price direction appears easier, and more likely to be successful, but then the trader has no real way of gauging the strength of the move in that direction. For example, a high degree of directional accuracy may not translate to high returns after costs if the movement in the direction forecast is small, as noted by Azoff (1994) and Thawornwong and Enke (2004). Ruggiero (1996) makes a number of suggestions, such as predicting a smoothed output function, such as a forward shifted technical indicator which is constructed from underlying price data. However, there are a number of inherent advantages and disadvantages in all technical indicators, and whilst they may be smoother than price action, they are typically only suitable for trading whilst market action is favorable to that actual indicator. For example, as Bauer (1998), Pring (1999) and a host of other technical analysts explain, trend based indicators perform well whilst the market is trending, but perform poorly at other times. Oscillators perform well when the market is not trending, but perform poorly otherwise. The temptation to create two technical neural networks, one for each main type of market activity is easily avoided, as then a further methodology would be required to tell which network to use at which point in time. In any event, a number of academics believe that the market actually goes through three phases, trending, non-trending, and chaotic, making the selection of which network to trust at which point in time much more complex. For further information on training neural networks with chaotic constraints, see Slim and Trabelsi (2003). Finally, according to Ruggiero (1997), the decision on what target to predict should be based on three factors, namely, the desired trade frequency, risk / reward criteria, and expertise available.

In consideration of all of the above, it is proposed that a similar output be used for the technical data trained ANN as was used for the fundamental data trained ANN, namely, an indication of the relevant strength of any movement expected over the forecast period. This should give rise to a highly tradeable indicator, which can be expected to perform during both trending and non-trending (and any other!) phases of the market.

It is also worthy of note that in financial trading terms, a low accuracy forecast does not necessarily equate to a low profitability system. For example, a low overall accuracy in forecast is acceptable if there is a tendency to correctly forecast highly profitable moves. Indeed, this is preferable. This is particularly relevant for the neural network trained with fundamental data, which is designed for longer-term trading. According to Chande (1997) only about 5% of the trades made by a trend following system are the 'big ones'.

Ruggiero (1997) suggests post-processing the outputs from neural networks trained for trading purposes. He suggests only relying on the signals if they are some threshold value away from the lower end of the range of possible output predictions, as those predictions closer to the middle of the range are within the error of the model. Similar advice is given by Azoff (1994). Hellstrom and Torgo (2000) take this one step further, by passing signals from a neural network through a classification module, which determines which signals should be taken by further analyzing technical trading aspects of the signals.

In this thesis, function profiles are built from the trained neural networks using the in-sample data, which can then be used to visually determine appropriate thresholds. These thresholds can then be used as the signal threshold values for the trading system when operating against the unseen out-of-sample data.

Outputs in Neurolab are scaled between 0 and 100. The true outputs in Neurolab are in the range of the activation function, which is 0 to 1. However, for ease of use, these output values are reported as integer values between 0 and 100. These reported outputs are simply calculated as the integer part of the true value multiplied by 100.

3.8 Limitations of the study

3.8.1 Ratio Analysis

There are a number of assumptions that are implicit in the use of accounting ratios, particularly when they are used to screen stocks, as in the fundamental data based portion of this thesis. The first assumption covers proportionality. As a ratio is reported as a proportion, it is clear that two equal ratios could have originated from two very different numerators and denominators. Here an assumption is being made that the relationship between numerator and denominator is similar, regardless of size. A detailed examination of this issue is provided by Conroy (1992).

Ratios allow easy comparison across companies, yet there are rarely optimal levels for ratios, and how 'good' a ratio is depends on your point of view. For example, a high DPR (Dividend Payout Ratio) may be favorably considered by a long-term security holder (as a compensation for systemic risk), whilst others may see this as wasteful if the firm could instead have used the money in retained earnings to further grow the company (with the goal of increasing its short term, or intermediate term share price).

The following additional limitations are summarized from a discussion by Hoggett and Edwards (1996). Another limitation of ratio analysis is that negative numbers can distort comparisons. Further, year-end data used for calculation of ratios may not be typical of a firm's position during the year. Historically, there have also been changes to disclosure rules which may be reflected in general-purpose financial reporting. Finally, entities may not be directly comparable, dependant on their use of different accounting methods. For example, the method used to account for inventory can significantly affect a firm's reported financial position.

3.8.2 Neural networks

A number of limitations are inherent in the use of neural networks for a study of this type. Chiefly, these concern the fact that the neural net is a 'black box', and that rule extraction, whilst possible in some limited circumstances, is a particularly difficult and uncertain process (for more detail, as well as a discussion on the use of Support Vector Machines, refer to Mitsdorffer (2002)). Generally, most researchers do not attempt to extract rules from their networks; rather they rely on statistical testing performed on out-of-sample results. This thesis will also rely on statistical testing of out-of-sample results. As an example of the process required to extract rules from a neural network, Skabar and Cloete (2000) describe their detailed analysis of a trained neural network, and their method of determining rules based on intensive observation and relationship to previously known rules. Overall, however, the neural net is not suitable for use as an explanatory tool.

Neural networks also tend to overfit the data if not very carefully controlled during the training process, and can find non-causal patterns in data very easily. There are no rigorous training methodologies that avoid this problem entirely. Determining a good internal structure for the network also tends to be a rather delicate process (refer Tan (1993) for a thorough description), and although a number of useful guidelines exist, there are again no definite steps to success.

Despite these clear limitations with neural networks, they are still considered the tool of choice for investigating non-linear relationships amongst noisy and complex data sets.

3.9 Summary of models

As shown in Table 3-1, this thesis implements two styles of neural network models, one style trained using only company fundamental data, and one style trained using only technical market data. These neural networks are designed to generate trading signals which can enhance returns from trading in the Australian stockmarket. As already discussed, to be effective, these networks must be placed within a valid trading context, incorporating risk control and money management.

Two styles of neural networks are developed, namely:

1. A fundamental (screening) neural network
2. A technical (timing) neural network

Each of these styles of neural networks is trained using data from two different target markets, namely:

1. Australian Allshare
2. S&P ASX200

Each of the four neural networks must be sited within its own valid trading system to be tested effectively. Therefore, there are four trading systems created during this thesis:

1. long-term trading system based on FNN(Allshare), named TS-FNN(Allshare),
2. long-term trading system based on FNN(ASX200), named TS-FNN(ASX200),
3. short-term trading system based on TNN(Allshare), named TS-TNN(Allshare),
4. short-term trading system based on TNN(ASX200), named TS-TNN(ASX200)

Later steps of creating trading systems which could combine both the fundamental neural networks and the technical neural networks for their respective markets is left for future research, however, as noted by Gately (1996), it is an area of interest. Also, there is virtually unlimited scope for creating trading networks using mixed inputs of technical and fundamental variables.

In summary, a brief diagrammatic representation of the architecture of each of the four trading systems follows as Figure 3-4 through Figure 3-7.

Trading System TS-FNN Allshare will host the neural network trained using fundamental data from the Australian Allshare. The architecture of this system is shown in Figure 3-4.

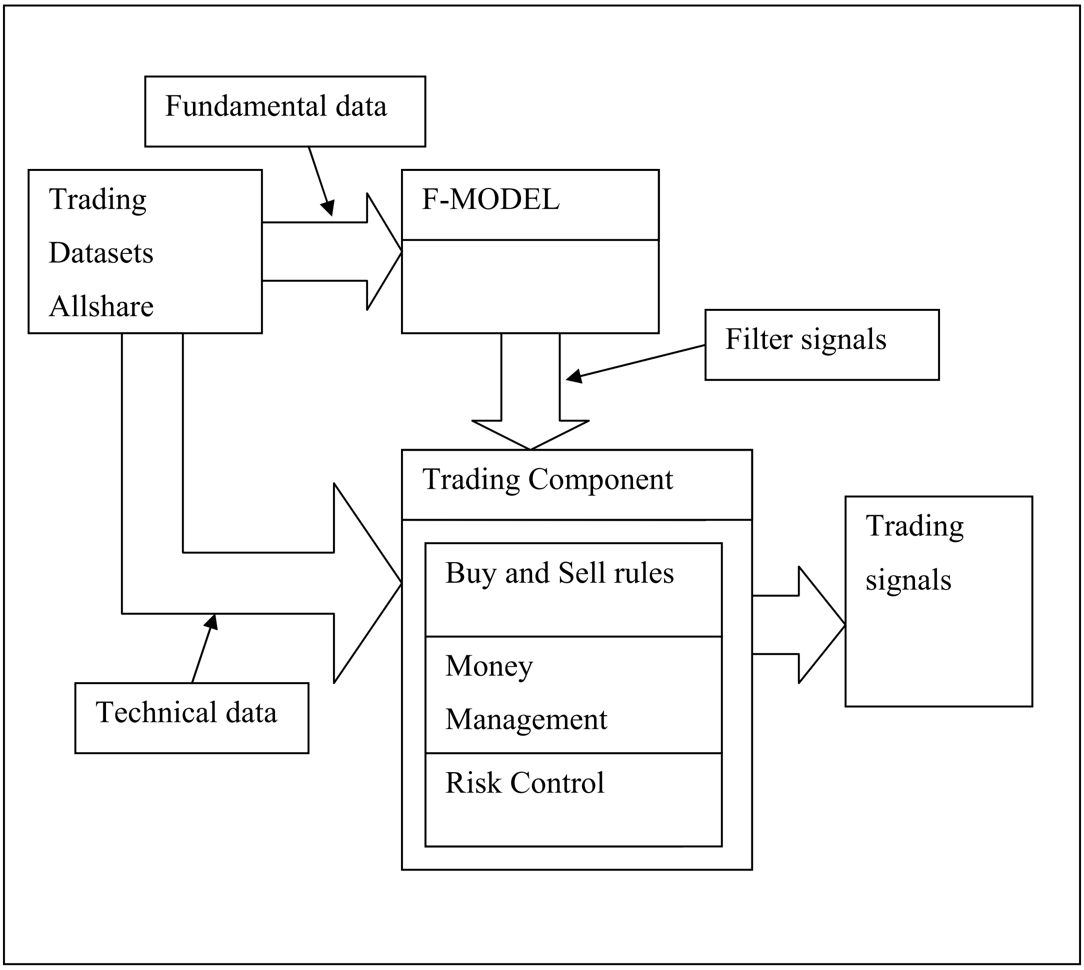

Figure 3-4 Architecture of Trading System TS-FNN Allshare

Trading System TS-FNN ASX200 will host the neural network trained using fundamental data from the Australian S&P/ASX200 constituents. The architecture of this system is shown in Figure 3-5.

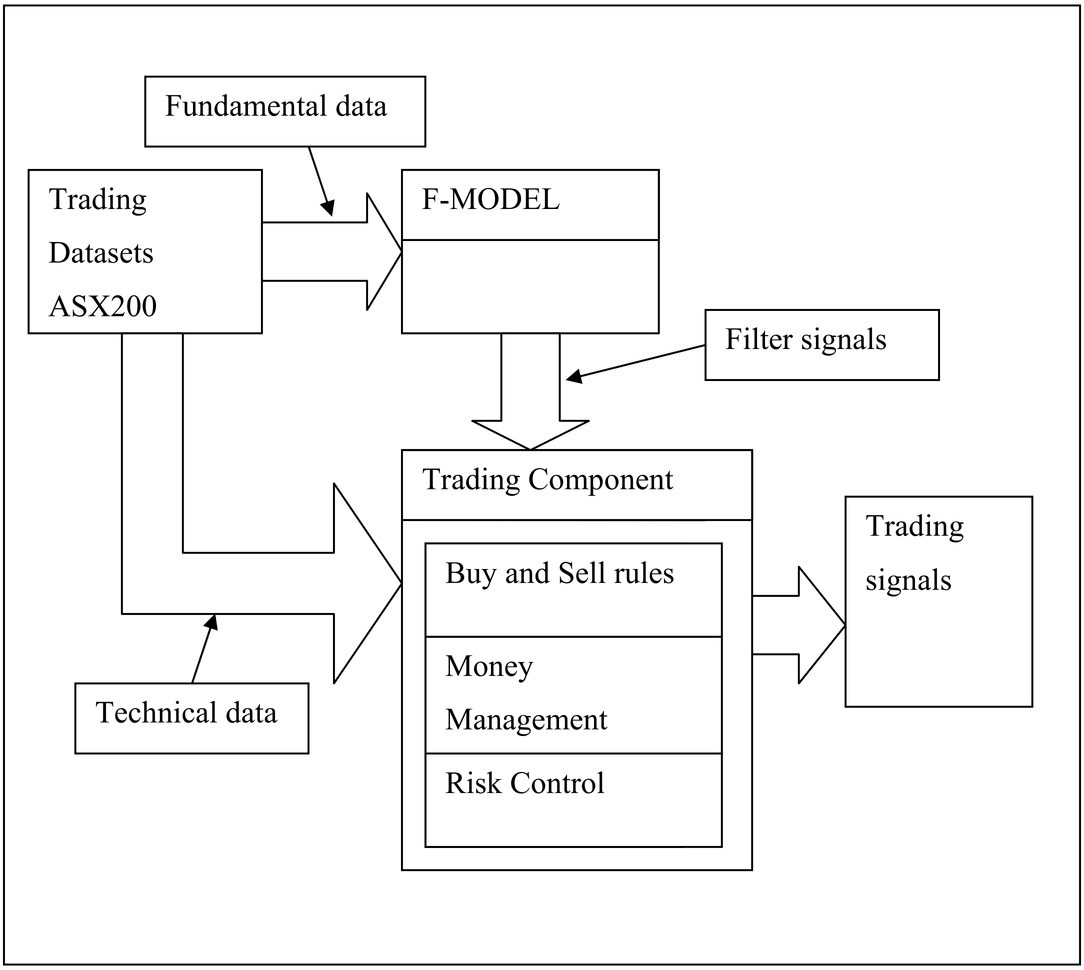

Figure 3-5 Architecture of Trading System TS-FNN ASX200

Trading System TS-TNN Allshare will host the neural network trained using technical data from the Australian Allshare. The architecture of this system is shown in Figure 3-6.

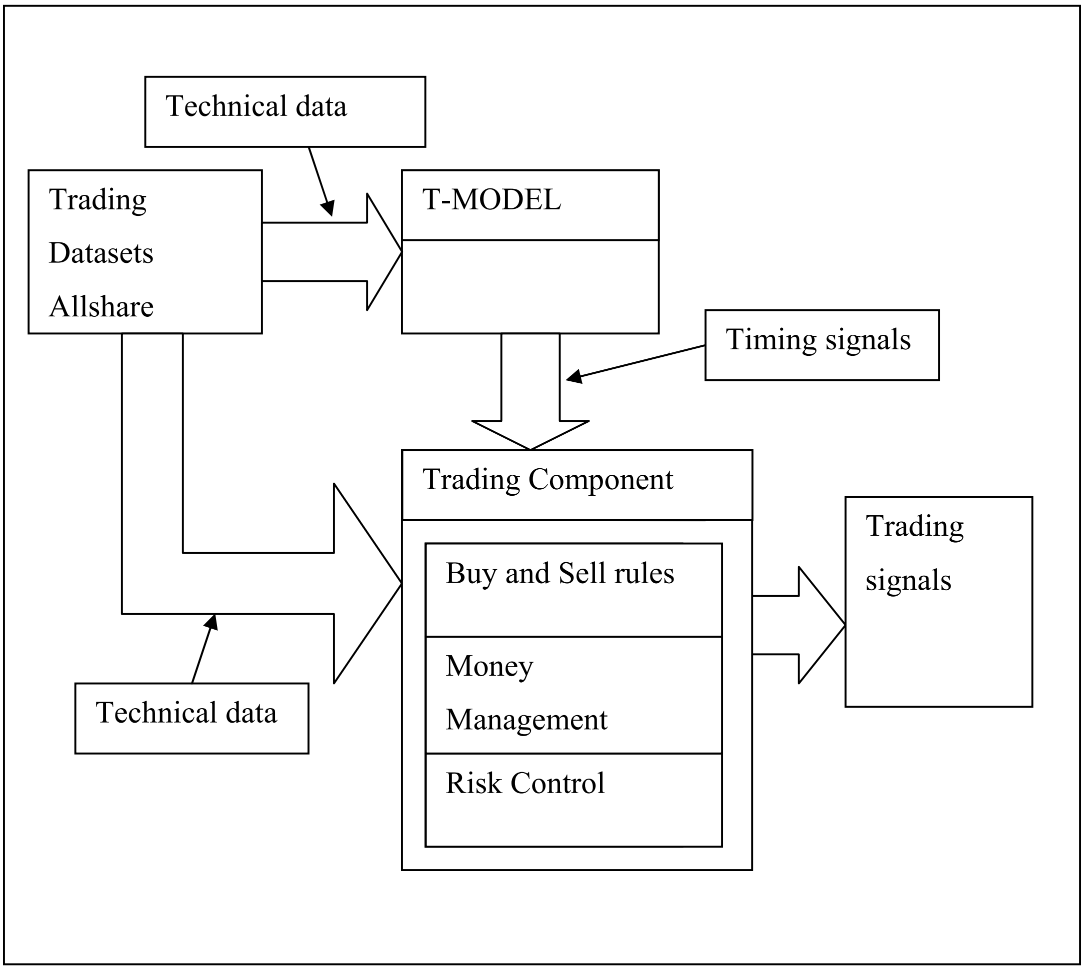

Figure 3-6 Architecture of Trading System TS-TNN Allshare

Trading System TS-TNN ASX200 will host the neural network trained using technical data from the Australian S&P/ASX200 constituents. The architecture of this system is shown in Figure 3-7.

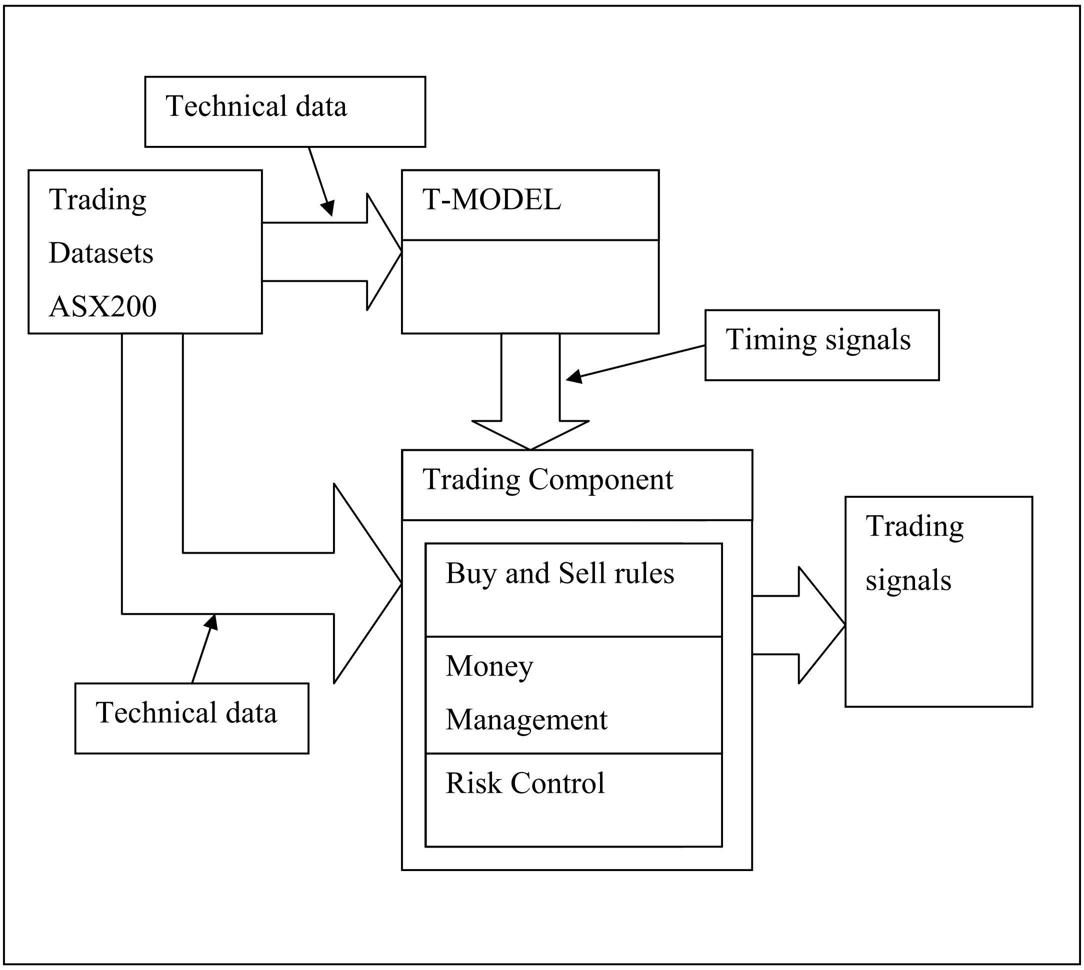

Figure 3-7 Architecture of Trading System TS-FNN ASX200

3.10 Testable Hypotheses

As already discussed, this study provides much needed depth regarding financial trading in the Australian stockmarket. This thesis also demonstrates a valid methodology for building trading systems.

However, the primary objective of this thesis is to demonstrate that neural networks can be used to enhance trading returns, and therefore to build realistic, profitable trading systems. For the hypotheses stated below, 'outperforms' is used to mean that a greater return and consistency has been achieved.

Two previously published fundamental strategies were selected as the basis for the neural networks that were created in this thesis. The justification and detail for the selections are provided in section 3.3.1. The variables used in the selected fundamental strategies were used as the inputs to the neural networks. The primary goal of any trading system is to outperform the market. As such, the testable hypothesis for the neural networks trained with fundamental data is:

Hypothesis 1
Trading strategies created using neural networks trained with fundamental data can outperform the buy-and-hold returns from their respective markets.

Due to the nature of the work published regarding technical analysis, there are no fully disclosed trading strategies that cover the range of technical variables used by the neural networks that were created in this thesis. The primary goal of any trading system is to outperform the market. As such, the testable hypothesis for the neural networks trained with technical data is:

Hypothesis 2
Trading strategies created using neural networks trained with technical data can outperform the buy-and-hold returns from their respective markets.

Taken together, conclusions about these two hypotheses will allow us to answer the initial question posed in this thesis, which was:

“Can ANNs be used to develop economically significant stockmarket trading systems?”.

Chapter 4 Results and Analysis

4.1 Introduction

Four fully trained neural networks were created during this thesis. These are summarized below:

1. neural network trained using ASX Allshare fundamental variables, FNN(Allshare)
2. neural network trained using S&P ASX200 fundamental variables, FNN(ASX200)
3. neural network trained using ASX Allshare technical variables, TNN(Allshare)
4. neural network trained using S&P ASX200 technical variables, TNN(ASX200)

Each of the four neural networks must be sited within its own valid trading system to be tested effectively. For this reason, appropriate signal strength thresholds and stop-loss parameters must be chosen, and these will form the structure of the trading system that will surround each neural network. Therefore, there are four trading systems created during this thesis:

1. long-term trading system based on FNN(Allshare), named TS-FNN(Allshare),
2. long-term trading system based on FNN(ASX200), named TS-FNN(ASX200),
3. short-term trading system based on TNN(Allshare), named TS-TNN(Allshare),
4. short-term trading system based on TNN(ASX200), named TS-TNN(ASX200)

Both the signal strength threshold and the stop-loss parameters for the trading systems must be chosen by examining only the in-sample data. The out-of-sample data may only be accessed once all trading system parameters are chosen, and the trading system structure has been determined. This selection of the required parameters follows ideas described in Lukac and Brorsen (1989).

At this point only, the trading system developed for each neural network is run once only through the appropriate out-of-sample data and the results recorded. It would be invalid to make any changes to the trading system or parameters after the trading systems have been run once on out-of-sample data, as this would represent a violation of what could have been known in chronological time.

The structure of this section of the thesis is the same for each neural trading model.

For each neural network developed:

- the characteristics of the in-sample data are displayed,
- the in-sample metrics are displayed for each neural network iteration developed, and the architecture of the chosen neural network is decided,

For each trading system developed:

- the neural network signal strength threshold parameter is decided,
- the stop-loss parameter is chosen,
- the out-of-sample trading results are displayed,
- a discussion of the results is provided

A detailed discussion of the interpretation of the traders metrics will be provided for the first trading system only. Essentially, the interpretation of traders metrics is independent of the actual trading system, therefore, after a detailed discussion has occurred for the first trading system, a summary discussion will be presented for each of the other three systems developed.

4.2 Trading in the ASX Allshare using neural networks and fundamental data

The neural network FNN(Allshare) has been trained to signal abnormal long-term trading opportunities. It was trained with fundamental data, using a one year (200 days) lookahead period. Results from key stages in the development of the neural network are presented below. To test the neural network properly, it must be sited within a valid trading context. The trading system developed around this neural network is named TS-FNN(Allshare), and this trading system is thoroughly tested out-of-sample, and the results are also presented below.

4.2.1 Training data

Due to the large volume of training data available, and the need to fit the training data within main memory, a simple sampling strategy was implemented. Initially, the minimum, maximum, mean and standard deviation of each fundamental variables time series was recorded, as shown in Table 4-1. Then the sampling strategy was to select every 50th row in each training dataset as a candidate for a neural input row. Every 50th row was then passed to the neural network as long as it did not contain a value that was outside three standard deviations for that particular fundamental variables time series.

Variable	Min	Max	Mean	Stddev
DIVYLD	0.00	22.41	0.03	0.33
PRICE2BOOK	-3,149,989.25	4,666,651.00	304.63	84,304.99
TOTALCURRENTASSETS	-1.00	2,108,700,032.00	76,602,984.00	230,713,424.00
TOTALGROSSDEBT	-1.00	2,121,299,968.00	80,401,776.00	267,895,360.00
WAVGSHARES	-1.00	2,013,090,816.00	140,693,776.00	252,867,152.00
CR	0.00	1,773.06	7.53	29.69
EPS	-587.70	8,607.20	9.62	131.74
PAYOUT	0.00	378.00	0.33	5.61
BVPERSHARE	-5.70	484.11	1.40	12.77
ROE	-90.48	1,953,250.00	318.89	22,842.05
PER	0.00	46,563.88	36.39	748.06

Table 4-1 FNN(Allshare): Characteristics of available data

Using this approach, a total of 22,954 rows were input to the neural network for training purposes. Each of these input rows have been transformed using the process described above. For information purposes only, the characteristics of these input rows are shown in Table 4-2.

Variable	Min	Max	Mean	Stddev
DIVYLD	0.00	1.02	0.02	0.05
PRICE2BOOK	-252,611.34	253,218.59	2.24	6,246.25
TOTALCURRENTASSETS	-1.00	768,743,232.00	63,156,364.00	155,981,616.00
TOTALGROSSDEBT	-1.00	884,087,872.00	64,385,856.00	182,640,016.00
WAVGSHARES	-1.00	899,295,232.00	128,186,928.00	191,159,088.00
CR	0.00	96.60	6.40	15.26
EPS	-385.60	404.84	7.72	33.79
PAYOUT	0.00	17.16	0.25	0.62
BVPERSHARE	-5.70	39.71	1.05	2.54
ROE	-90.48	68,845.04	20.90	1,201.92
PER	0.00	2,280.57	21.30	124.59

Table 4-2 FNN(Allshare): Characteristics of input data

Table 4-3 shows the same summary information for the training target variable. This target is the maximum percentage change in price over the next 200 days, computed for every element *i* in the input series as:

$$\text{Target} = \left(\frac{\left(\max(close_{i+200},\ldots,close_{i+1}) - close_i\right)}{close_i} \right) \times 100$$

Equation 4-1 FNN(Allshare): Computing the target variable

Effectively, this target allows the neural network to focus on the relationship between the input fundamental variables, and the expected yearly forward price change.

Training Output	Min	Max	Mean	Stddev
Output	0.00	100.00	13.71	34.39

Table 4-3 FNN(Allshare): Characteristics of training target

4.2.2 Architecture selected

Each neural network was trained until no new error low had been reached for 2000 epochs. The architecture to be carried forward to out-of-sample testing is selected by reference to the in-sample metric previously detailed. Table 4-4 displays the computed in-sample metric.

Number of Hidden Nodes	Closed Trades	Total Trades	Filter Selectivity
4	4	5	80
5	1	1	100
6	12	17	70.588235
7	22	29	75.862069
8	2	3	66.666667

Table 4-4 FNN(Allshare): In-sample metric values

From the above table, the architecture with 7 hidden nodes is selected. In this research, the neural network with this architecture proceeds on to trading system development.

4.2.3 Derivation of Trading System parameters

4.2.3.1 Signal Strength Threshold

By inspection of the function profile graph below, a signal strength threshold of 20 must be exceeded to initiate trades. The value of 20 is chosen as it is the point at which the neural network begins to signal trades with net positive results. This can be seen in the graph as it represents the signal strength value for which the expected returns are increasing above zero in both the 6 monthly and yearly timeframes.

The trading rules component of the trading system for this neural network consists of the following buy/sell rules:

Buy: Buy tomorrow when neural signal output(today) > 20, and neural signal output(today) > neural signal output(yesterday)

Sell: Sell tomorrow when neural signal output(today) <= 20, and neural signal output(today) < neural signal output(yesterday)

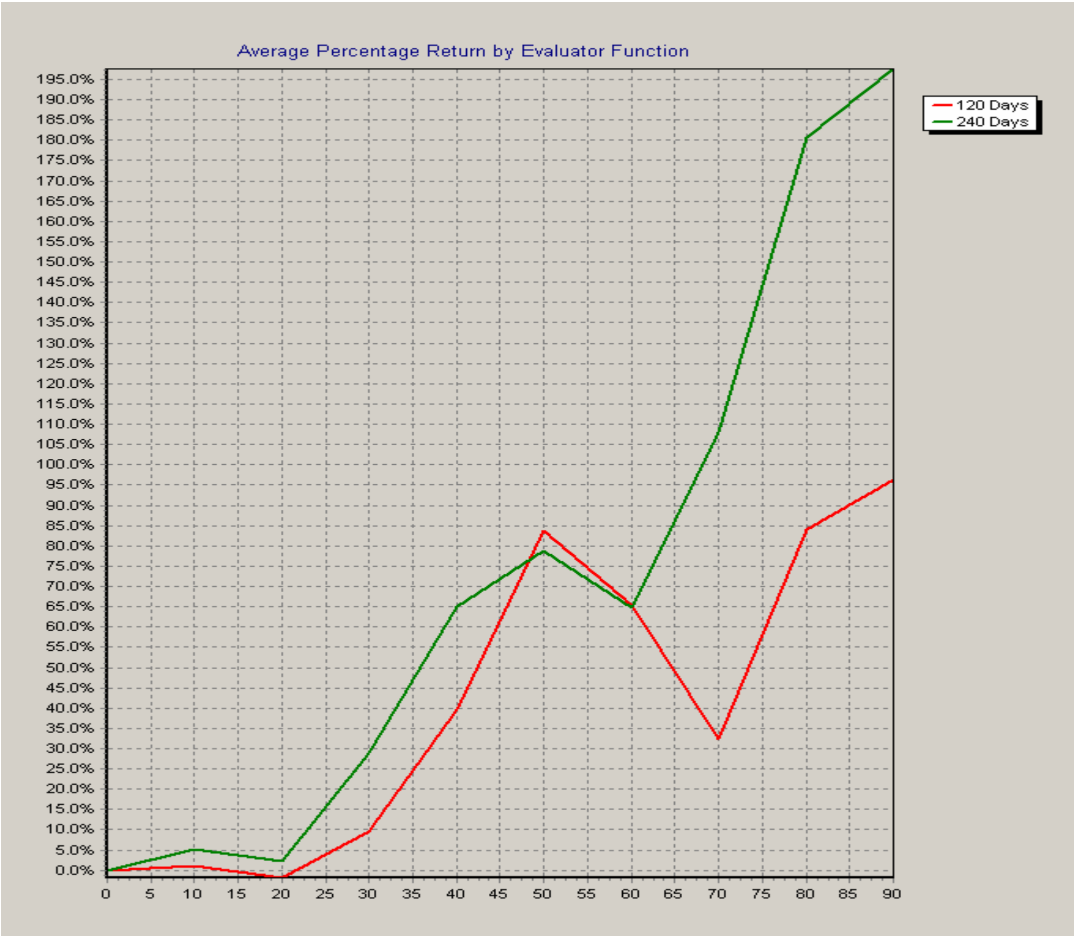

Figure 4-1 FNN(Allshare): In-sample output signal strength graph

4.2.3.2 Stop-Loss threshold

The following figures show the MAE for the in-sample trades.

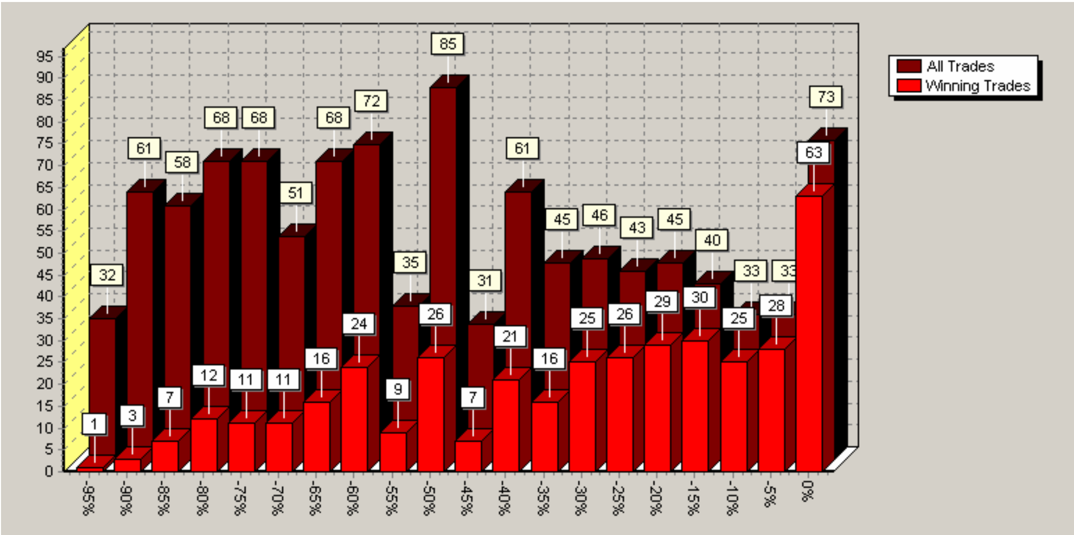

Figure 4-2 TS-FNN(Allshare): In-sample MAE

The MAE shows that few of the in-sample trades which go on to become profitable incur losses of 45% or more during the trades' lifetime (the period between opening and closing the trade). It is at this threshold that a sharp fall in the number of winning trades is observed.

By inspection of the MAE graph, a stop-loss threshold of 45% can be chosen. This value will then be used as the stop-loss threshold for the out-of-sample testing.

Interestingly, a stop value of 35% could also have been justified, although with less commitment. The value of 45 was used as it follows Chande's (1997) and also Tharp's (1998) general guideline of setting wide stops for long-term trading systems. In any event, after the testing was complete and documented, the tests were re-run using the value of 35%, and the final outcomes were extremely similar, with all results still statistically significant.

4.2.4 Out-of-Sample Results

The following tables report the out-of-sample results of the trading system TS-FNN(Allshare), built around the FNN(Allshare) neural network. The out-of-sample period is two years, from the start of trading in January 2002, to the end of trading in December 2003. In this section, results are displayed for the dataset including delisted stocks, as well as for the dataset excluding delisted stocks. As the realistic case is to include delisted stocks, only these results will be included in the discussions. The results are presented for the dataset which excludes delisted stocks for interest only.

Metric	**TS-FNN(Allshare) includes delisted**	**TS-FNN(Allshare) excludes delisted**	**Buy-and-Hold includes delisted**
Net Profit (Loss)	$ 71,833.81	$ 60,281.53	$ 32,379.12
Net Profit (Loss) %	71.83 %	60.28 %	32.38 %
Annualized Gain %	31.21 %	26.70 %	15.11 %
Exposure %	64.85 %	64.50 %	96.71 %
All Trades: Number of trades	145	146	1,367
All Trades: Average Profit (Loss), Average Profit (Loss) %	$ 495.41 54.17 %	$ 412.89 43.60 %	$23.69 32.74 %
All Trades: Average Bars Held	193.52	187.82	472.4
Winning Trades: Number of Trades	64	64	682
Winning Trades: Winning %	44.14 %	43.84 %	49.89 %
Winning Trades: Average Profit, Average Profit %	$ 1,618.14 174.34 %	$ 1,412.12 148.42 %	82.28 113.07 %

Metric	TS-FNN(Allshare) includes delisted	TS-FNN(Allshare) excludes delisted	Buy-and-Hold includes delisted
Winning Trades: Average Bars Held	257.2	248.78	493.13
Winning Trades: Max consecutive wins	7	6	10
Losing Trades: Number of Trades	81	82	685
Losing Trades: Losing %	55.86 %	56.16 %	50.11 %
Losing Trades: Average Loss, Average Loss %	\$ 391.69 40.78 %	\$ 367.00 38.21 %	\$ 34.65 47.24 %
Losing Trades: Average Bars Held	143.2	140.23	451.75
Losing Trades: Max consecutive losses	11	10	11
Max Drawdown, Max Drawdown %, Maximum Drawdown Date	\$ 24,141.49 23.82 % 19/06/2003	\$ 21,816.70 21.51 % 27/03/2003	\$ 22,390.84 20.86 % 13/03/2003
Profit Factor (PF)	3.2641	3.0031	2.3641
Recovery Factor (RF)	2.9755	2.7631	1.4461
Payoff Ratio (PR)	4.2751	3.8847	2.3925
Sharpe Ratio (SR)	1.1161	1.1337	0.7623
Ulcer Index	11.0131	9.8353	11.3687

Metric	TS-FNN(Allshare) includes delisted	TS-FNN(Allshare) excludes delisted	Buy-and-Hold includes delisted
Luck Coefficient (LC)	16.6281	19.5325	9.6076
Pessimistic Rate of Return (PRR)	2.6601	2.3891	2.2075

Table 4-5 TS-FNN(AllShare): Out-of-sample trading metrics

Table 4-6 shows a statistical breakdown of the out-of-sample trades from the TS-FNN(Allshare). The mean of the net profit/loss is tested against the mean of the distribution curve that a random trading strategy would produce, which is generally assumed to be zero under the null hypothesis of no excess returns.

The hypotheses for the t-tests will be:

H_0: $\mu_{profit} = 0$,

H_1: $\mu_{profit} > 0$

Metric	TS-FNN(Allshare) includes delisted
Sample size	145
Sample Mean	495.4056
Sample Standard Deviation (SD)	2821.73841
Standard Error of the Mean	234.33262
t-statistic (P/L > 0)	2.114
Degrees of freedom (df)	144

Metric	TS-FNN(Allshare) includes delisted
t-statistic (5%,df) 1-tailed to exceed	1.655
Lower 95% confidence interval of the Mean	32.2296
Upper 95% confidence level of the Mean	958.5816

Table 4-6 TS-FNN(Allshare): Statistical analysis of mean profit/loss

The t-statistic for the TS-FNN(Allshare) trades allows us to reject the null hypothesis and conclude that the mean profit of the trading system is significantly greater than zero. Specifically, $t(144) = +2.114$, $p = 0.018$ (< 0.05), one tailed.

The results of Vince's runs test for dependency is shown in Table 4-7 below.

Metric	TS-FNN(Allshare) includes delisted
Total Cases	145
Number of Runs	62
Z score	-1.69

Table 4-7 TS-FNN(Allshare): Runs test results

From the Runs Test, we can conclude that there is insufficient evidence to accept dependency amongst the trades.

Finally, the set of trades from the TS-FNN(Allshare) are compared to the set of trades from the buy-and-hold approach, using the ANOVA technique. The results are reported in Table 4-8 below.

	TS-FNN(Allshare) includes delisted	Buy-and-Hold incl delisted
Mean	495.406	23.686
Standard Deviation	2821.7384	95.9704

Table 4-8 TS-FNN(Allshare): Trades compared to Buy-and-Hold trades

The analysis of variance reveals that the trades selected using TS-FNN(Allshare) are significantly different from the trades from the buy-and-hold approach, specifically $F(1,1510) = 38.001$, $p = 0.00$ ($p< 0.05$).

Finally, a brief presentation of the distribution of monthly returns from the strategy is included. This gives additional insight into how the trading system can be expected to perform into the future, and can also be used by traders to establish confidence intervals on monthly returns. This is helpful in identifying whether a model is performing within expectations, and can also aid in detecting when in the future a model has ceased to provide an advantage, and needs to be retrained.

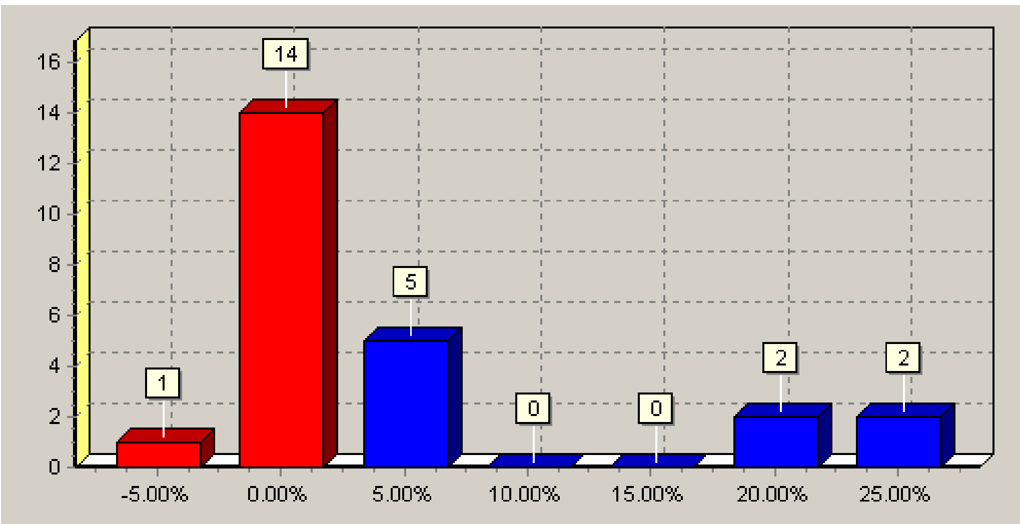

Figure 4-3 TS-FNN(Allshare): Distribution of monthly returns

4.2.5 Detailed discussion of Traders Metrics

Overall, returns from the trading system TS-FNN(Allshare) are excellent. As shown by the results discussed above, it can be concluded that the mean profit of the trades selected by the ANN at the heart of this trading system, FNN(Allshare), are significantly different from the mean profit of the trades which constitute the buy-and-hold strategy. Therefore, we can confidently state that the trading system which embodies this neural network has outperformed the buy-and-hold approach.

4.2.5.1 Net Profit (Loss) / Net Profit (Loss) %

This metric simply describes the increase (decrease) in the final value of the portfolio after all trades have taken place. It shows the final increase (decrease) in dollars, and also that value as a percentage of the initial starting capital. Overall, these are important metrics, however, as mentioned earlier, they must be considered with respect to a number of other factors, such as risk, time in the market, etc. These other factors have their own metrics, which are described and assessed below. From the 'Net Profit (Loss)' and 'Net Profit (Loss) %' metrics point of view, the system being tested, TS-FNN(Allshare), results in more than double the final profit of the buy-and-hold strategy. Therefore, the system being tested is preferable to simply using the buy-and-hold approach.

4.2.5.2 Annualized Gain %

The 'Annualized Gain %' metric provides a valid and simple comparison between any number of trading systems, subject to the factors mentioned above. It is a simple way of comparing systems that have traded over different time lengths, as it effectively annualizes the gross returns of systems. When the time lengths are the same, it provides a simple direct comparison of returns on a yearly basis. From the 'Annualized Gain %'

metrics point of view, the system being tested, TS-FNN(Allshare), is preferable as it has more than doubled the annualized returns achieved by the buy-and-hold approach.

4.2.5.3 Exposure %

The 'Exposure %' metric measures the percentage of time that a trading system was actually in the market. In essence, it describes the percentage of the time the system had trades that were held open overnight, and therefore, were exposed to systemic market risk. By definition, the buy-and-hold strategy will have virtually 100% exposure, as it mandates buying on the first day the security is available, and selling on the last day. As such, the buy-and-hold strategy is heavily exposed to systemic risk, a point made only too clearly to long term investors and pension holders in the aftermath of the terrorist bombings. From a traders point of view, the less exposure the better, therefore the system being tested, TS-FNN(Allshare) is preferable to the buy-and-hold approach. A relevant consideration for traders, albeit one that is outside the scope of this thesis, is the ability to invest unused funds in separate ventures, such as short-term money markets.

4.2.5.4 All Trades: Number of Trades

This metric is simply a count of the number of trades instigated by the system. The neural network at the heart of this system, FNN(Allshare) is designed to signal those securities which can be expected to outperform over the next year. Rather than being a metric that a trader might directly focus on, this metric confirms the operation of the neural network, FNN(Allshare), as selecting a subset of trades with greater than average potential (as previously shown by the ANOVA test). One facet worth commenting on from a trader's perspective is that the transaction costs incurred by any system are a direct function of the number of trades taken. Therefore, a higher return from a smaller number of trades means that a greater amount of capital can be applied to each trade, with the expectation of higher returns. Another factor outside the scope of this thesis, but which is relevant from a practical point of view, is that fewer trades means much simpler operation of a trading system, due to the practical constraints of entering and executing each individual trade. The value for this metric for the system being tested, TS-FNN(Allshare), is much smaller than that for the buy-and-hold approach.

4.2.5.5 All Trades: Average Profit (Loss), Average Profit (Loss) %

These two metrics describe the expected profit (loss) as an average across all trades taken, in essence, they are describing the amount of profit to be expected from the average trade, in both dollar terms, and as a percentage of the trade value undertaken. Subject to the factors mentioned earlier regarding the 'Net Profit(Loss)' metric, the higher this value the better. This value represents the mean of the net trade profit for the system being tested. The ANOVA test has already shown this value to be significantly better than that achieved by the buy-and-hold approach, therefore, from the point of view of this statistic, the system being tested, TS-FNN(Allshare) is significantly better.

4.2.5.6 All Trades: Average Bars Held

The 'All Trades: Average Bars Held' metric details the number of bars that the average trade is held open. In this thesis, one bar corresponds to 1 trading day. In general, traders talk about bars as opposed to exact time periods, with the actual clarification of the time period dependant on the trading system being discussed. For example, a day trading system would also talk about bars, but in that context, a bar could be 1 minute elapsed, or 1 hour elapsed, etc. For the TS-FNN(Allshare), the average holding period is just under 1 year. Rather than being a metric that a trader might directly focus on, this metric confirms the operation of the neural network, FNN(Allshare), as selecting a subset of trades with greater than average potential in the 1 yearly timeframe. From a trader's point of view, there are two implications of this value, both of which are outside the scope of this thesis.

Firstly, there is the issue of dividends. Holders of common stock expect interest payments on their stockholding in the form of dividends. Often these can be re-invested using cost effective dividend reinvestment plans, with the effect of reducing the overall cost base of the stock. Where this option is not available, the dividends can be added to the portfolio to fund further investment. The second implication of the length of the holding period is that a reasonable number of stocks are held over the 1 year holding period (as shown by the traders metrics ‘Winning Trades: Average Bars Held’, and ‘Losing Trades: Average Bars Held’). From a traders point of view, there is beneficial tax treatment for holding a stock investment for over 1 year. As already mentioned, both of the benefits listed are outside the scope of this thesis, however, they represent very real financial benefits to a trader.

4.2.5.7 Winning Trades: Number of Trades, Winning %

These two metrics detail the total number of winning trades, and the percentage of trades that were winners. As above, rather than being a metric that a trader might directly focus on, this metric confirms the operation of the neural network, FNN(Allshare), as selecting a subset of trades with greater than average potential. From a traders point of view, one comment is worth making here, again, the issue it concerns is outside the scope of this thesis. It concerns human ego, and whether an undisciplined user of the trading system would be able to trade a system which lost more trades than it won, even if it won significantly more money than it lost. In essence, this is a question of trader psychology, and perhaps explains why many traders are uncomfortable with trading a system they did not have a hand in developing.

4.2.5.8 Winning Trades: Average Profit, Average Profit %

Given the nature of the system being tested, it is to be expected that the winning trades should show significant gains, both in strict dollar terms, and also as a percentage of capital risked. The values for both of these metrics are far superior to those of the buy-and-hold approach, and the system being tested, TS-FNN(Allshare) is the preferred system on the basis of these metrics.

4.2.5.9 Winning Trades: Average Bars Held

Again, this metric confirms the operating characteristics of the underlying neural network, showing that profitable trades are being correctly signaled, in the correct timeframe. When compared to its counterpart, 'Losing Trades: Average Bars Held', it also confirms the mechanics of the trading system, showing that winning trades are given adequate time to acquire profit, whilst losing trades are cut away much earlier. The trading system being tested, TS-FNN(Allshare) has a much lower value for this metric than that of the buy-and-hold approach, however, it makes no sense to compare these two values. By its very nature, the buy-and-hold approach must buy and hold all stocks, and consequently, its value for this metric will be large.

4.2.5.10 Winning Trades: Max consecutive wins

Although this metric is a regularly published one, it appears to tell very little about the underlying operating characteristics of a system. It describes the number of trades that were winners in a row, also known as the length of the winning streak. It does this without reference to the amount won, or the time taken to win that amount, and as such, appears almost totally useless from a comparative point of view. It is likely that this metric, like the metrics 'Winning Trades: Number of Trades, Winning %' are more closely aligned with a traders individual psychology than anything else. Certainly, there is positive feedback to be gained from a longer winning streak, however, as already mentioned, it seems irrational to focus on the number of consecutive wins without regard to their value, or time taken. In any event, the system being tested has a smaller value for this metric than the buy-and-hold approach.

4.2.5.11 Losing Trades: Number of Trades, Losing %

These two metrics detail the total number of losing trades, and the percentage of trades that were losers. Trades may be realized as losers in this system in two ways. Firstly, the neural network might give a low output signal (below the threshold), after having previously signaled a higher value (above the threshold). If this happens, the system schedules the trade to be closed out on the next trading day, regardless of whether the trade is in profit or not. If the trade is not in profit, this trade will be classified as a loser. The other way for a trade to become a loser in this system is for the price of the stock to decline to such a level that the stop loss is triggered. This trade will automatically be a loser. Rather than being a metric to directly focus on, this metric needs to be considered in combination with the next metrics 'Losing Trades: Average Loss, Average Loss %'. In combination, this group of metrics describes how often the system can expect to lose, and how much it can expect to lose. This point is considered in more detail next.

4.2.5.12 Losing Trades: Average Loss, Average Loss %

These two metrics describe the average amount lost in both dollar terms, and as a percentage of the initiated trade size. As just mentioned, these metrics need to be considered in combination with 'Losing Trades: Number of Trades, Losing %'. It is clear that the average loss which can be taken is related to the frequency of loss. The more often the system loses, the lower the 'Average Loss %' must be. On the other hand, in a system which rarely loses, it is possible to withstand a larger ‘Average Loss %’. When comparing the TS-FNN(Allshare) systems metrics to those of the buy-and-hold approach, it is clear that the TS-FNN(Allshare) loses more often than the buy-and-hold approach (55.86% compared to 50.11%), however, it loses less (40.78% compared to 47.24%). As every system has the potential for loss, perhaps the best way to interpret this set of metrics will again come down to the individual psychology of the trader. The trader must answer the question ‘Is it better to lose less, more often?’, or ‘Is it better to lose more, less often?’. Certainly from a drawdown point of view (drawdown is considered later with its own metrics), it is better to lose less, more often. Large, infrequent losses can severely impact drawdown.

4.2.5.13 Losing Trades: Average Bars Held

This metric describes the number of days a trade is held open, and it is best considered in relation to the earlier metric, ‘Winning Trades: Average Bars Held’. It is desirable to close out losing trades as quickly as possible, as this allows the capital recovered from a losing trade to be cycled back into the next potential trade, allowing that capital to participate in the compounding effect. From this perspective, it is expected that the average bars held for losing trades should be lower than the average bars held for winning trades. Inspection of these two metrics shows that winning trades are held open for 257.2 days on average, whilst losing trades are held open for 143.2 days on average. These figures confirm the trading system to be operating correctly. It makes no sense to compare these figures to those of the buy and hold approach, for reasons already listed under the ‘Winning Trades: Average Bars Held’ metric.

4.2.5.14 Losing Trades: Max Consecutive losses

Much like its earlier counterpart, ‘Winning Trades: Max Consecutive wins’, this metric says very little about the operation of a system. Again, it considers the length of the losing streak, without regard to the amount lost over this period. It is likely the same issues relevant to trader psychology also drive the publication of this metric. To the extent that there is some desire to find a use for this metric, it, like its earlier counterpart could be used to signal when a system begins operating outside of its normal behavioral characteristics, and this could well signal that the neural network needs retraining. In any event, the value for the system being tested, TS-FNN(Allshare) is 11, the same as that of the buy-and-hold approach.

4.2.5.15 Max Drawdown, Max Drawdown %, Max Drawdown Date

Drawdown is the term that describes the portfolio value lost in an equity curve decline from a peak to a valley. The equity curve is simply a graph on a day-to-day basis of the portfolio value. A peak in this graph represents a period where the system accumulated profit, and a valley represents where that value was returned to the market. An equity curve effectively marks the portfolio to market every day, and as such represents the change in aggregate value of the entire set of securities held open on that day. Clearly, an abnormal crisis event affecting the stockmarket will drag down the equity curve of a portfolio, and abnormally 'good' events push it up. Of course, the height of the peaks and valleys are directly affected by which stocks the portfolio is holding, and the maximum drawdown measures this change. It is normal to consider drawdown in terms of the relationship between 'Max Drawdown %', and its related metric, 'Recovery Factor' (described later), which describes how effectively a system can overcome the effects of drawdown. From a traders point of view, the lower the drawdown, the better. The system being tested, TS-FNN(Allshare) has a slightly higher drawdown (23.82%) than the buy-and-hold approach (20.86%). A discussion on the importance of these figures will continue in the discussion for the metric 'Recovery Factor'.

4.2.5.16 Profit Factor

The ‘Profit Factor’ metric is a key metric quoted for trading systems. The higher the value of this metric, the more profitable the system is (or more accurately, ‘has been’). It is desirable for this metric to be greater than 2. Both the system being tested, TS-FNN(Allshare) and the buy-and-hold approach exceed the value of 2, and the TS_FNN(allshare) exceeds 3. On the basis of this metric, the TS-FNN(Allshare) system is more profitable.

4.2.5.17 Recovery Factor

As mentioned earlier (in the discussion of drawdown), this metric describes the ability of a system to overcome drawdown. Drawdown in itself is not a problem, it becomes a problem if a portfolios funds are eroded by drawdown to such an extent that it struggles to become profitable again. This may be due to reduced equity available to take trades, or may even mean that trades cannot be taken due to lack of capital. Therefore, it is important to focus on how effectively a system can overcome a drawdown event. The acceptable value for ‘Recovery Factor’ must be greater than 1. Both the system being tested, TS-FNN(Allshare) and the buy-and-hold approach exceed this figure, and the value for the system being tested is more than double that of the buy-and-hold approach. On the basis of this metric, the system being tested is better at overcoming drawdown.

4.2.5.18 Payoff Ratio

This metric describes how effective a system is at acquiring profit relative to losses. The higher the value of the ‘Payoff Ratio’ the better, and traders typically look for a value greater than 2. Both the system being tested, TS-FNN(Allshare), and the buy-and-hold approach exceed the value of 2, and the system being tested exceeds the value of 4. On the basis of this metric, the system being tested is more effective at acquiring profit.

4.2.5.19 Sharpe Ratio

The Sharpe Ratio is the standard method for comparing a wide range of investments. It is important as it benchmarks the returns achieved with respect to the amount of risk taken to achieve those returns. In portfolio theory, risk is measured in terms of variance, and the Sharpe ratio divides the total return over a portfolio achieved by the standard deviation of that portfolio, after subtracting the risk free rate of return. In this thesis, no attempt is made to remove excess funds from the portfolio, and the cash basis of the portfolio is not adjusted for any reason apart from acquiring and disposing of stocks, therefore, the risk free rate of return is set to zero for both the system being tested, TS-FNN(Allshare), and the buy-and-hold approach (which is always invested). Therefore, the Sharpe Ratio is a direct comparison of the returns achieved relative to risk, and the higher the Sharpe Ratio the better. The system being tested yields a higher Sharpe ratio than the buy-and-hold approach, therefore, on the basis of this metric, the system being tested is preferable.

4.2.5.20 Ulcer Index

This metric measures the overall volatility of a portfolio as a function of the daily drawdowns of that portfolio. As described above, lower volatility is preferable, and the system being tested, TS-FNN(Allshare) has a smaller Ulcer Index than the buy-and-hold approach, therefore, the system being tested is preferable.

4.2.5.21 Luck Coefficient

This metric is a comparison between the largest winning trade and the average trade. It attempts to determine how 'lucky' a system is. For a system which makes a great deal of trades, it is desirable that the system is not 'lucky'. For a system of this nature, the goal is to attempt to predict those stocks with the greatest potential for increase, and acquire them. These are the trades that Chande names 'the big ones', that is, those trades that outperform. From this point of view, it is expected that the system should be deemed lucky by this metric, and the interpretation of this metric is that the system is operating as per requirement.

4.2.5.22 Pessimistic Rate of Return

This metric effectively increases the number of losers in a system by the square root of total losers, and decreases the number of winners in a system by the square root of total winners. Its final calculation is then very similar to 'Profit Factor', using these new adjusted values for winners and losers. According to Vince, if only one metric is used to judge a system, it should be 'Pessimistic Rate of Return'. According to Vince, a value greater than 2 is good; a value greater than 2.5 is excellent. The value of this metric for the system being tested, TS-FNN(Allshare), exceeds that of the buy-and-hold approach, and it also exceeds 2.5. Therefore, on the basis of this metric, the system being tested is excellent.

4.3 Trading in the ASX200 using neural networks and fundamental data

The neural network FNN(ASX200) has been trained to signal abnormal long-term trading opportunities within the S&P ASX200. It was trained with fundamental data, using a one year (200 days) lookahead period. Results from key stages in the development of the neural network are presented below. To test the neural network properly, it must be sited within a valid trading context. The trading system developed around this neural network is named TS-FNN(ASX200), and this trading system is thoroughly tested out-of-sample, and the results are also presented below.

4.3.1 Training data

Due to the large volume of training data available, and the need to fit the training data within main memory, a simple sampling strategy was implemented. Initially, the minimum, maximum, mean and standard deviation of each fundamental variables time series was recorded, as shown in Table 4-9. Then the sampling strategy was to select every 20th row in each training dataset as a candidate for a neural input row. Every 20th row was then passed to the neural network as long as it did not contain a value that was outside three standard deviations for that particular variables time series.

Variable	Min	Max	Mean	Stddev
DIVYLD	0.00	0.24	0.04	0.04
PRICE2BOOK	-13.65	122.29	2.25	4.02
TOTALCURRENTASSETS	-1.00	2,108,700,032.00	273,468,192.00	407,995,392.00
TOTALGROSSDEBT	-1.00	2,121,299,968.00	278,957,792.00	447,763,136.00
WAVGSHARES	-1.00	2,009,524,352.00	331,934,816.00	415,283,104.00
CR	0.00	135.79	2.10	4.84
EPS	-217.30	8,607.20	30.21	256.05
PAYOUT	0.00	19.73	0.55	0.72
BVPERSHARE	-0.20	484.11	3.63	23.40
ROE	-1.63	4.48	0.10	0.22
PER	0.00	15,569.68	56.10	467.17

Table 4-9 FNN(ASX200): Characteristics of available data

Using this approach, a total of 14,577 rows were input to the neural network for training purposes. Each of these input rows have been transformed using the process described above. For information purposes only, the characteristics of these input rows are shown in Table 4-10.

Variable	Min	Max	Mean	Stddev
DIVYLD	0.00	0.16	0.04	0.04
PRICE2BOOK	-9.81	14.31	2.09	2.24
TOTALCURRENTASSETS	-1.00	1,497,454,336.00	264,941,536.00	379,976,256.00
TOTALGROSSDEBT	-1.00	1,622,247,168.00	270,700,544.00	421,340,736.00
WAVGSHARES	-1.00	1,577,784,064.00	324,247,936.00	394,483,648.00
CR	0.00	16.62	1.81	2.38
EPS	-217.30	798.36	23.30	48.64
PAYOUT	0.00	2.71	0.53	0.43
BVPERSHARE	-0.20	73.83	2.58	4.81
ROE	-0.56	0.76	0.10	0.12
PER	0.00	1,457.61	35.63	125.74

Table 4-10 FNN(ASX200): Characteristics of input data

Table 4-11 shows the same summary information for the training target variable. This target is the maximum percentage change in price over the next 200 days, computed for every element *i* in the input series as:

$$\text{Target} = \left(\frac{\left(\max(close_{i+200},...,close_{i+1}) - close_i\right)}{close_i} \right) \times 100$$

Equation 4-2 FNN(ASX200): Computing the target variable

Effectively, this target allows the neural network to focus on the relationship between the input fundamental variables, and the expected yearly forward price change.

Training Target	Min	Max	Mean	Stddev
Output	0.00	100.00	12.38	32.94

Table 4-11 FNN(ASX200): Characteristics of training target

4.3.2 Architecture selected

Each ANN was trained until no new error low had been reached for at least 2000 epochs. The architecture to be carried forward to out-of-sample testing is selected by reference to the in-sample metric previously detailed. Table 4-12 displays the computed in-sample metric.

Number of Hidden Nodes	Closed Trades	Total Trades	Filter Selectivity
4	191	279	68.458781
5	231	398	58.040201

Table 4-12 FNN(ASX200): In-sample metric values

From the above table, the architecture with 4 hidden nodes is selected. In this research, the neural network with this architecture proceeds on to trading system development.

4.3.3 Derivation of Trading System parameters

4.3.3.1 Signal Strength Threshold

By inspection of the function profile graph below, a signal strength threshold of 20 must be exceeded to initiate trades. The value of 20 is chosen as it is the point at which the neural network begins to signal trades with net positive results. This can be seen in the graph as it represents the signal strength value for which the expected returns are greater than zero in both the 6 monthly and yearly timeframes.

The trading rules component of the trading system for this neural network consists of the following buy/sell rules:

Buy: Buy tomorrow when neural signal output(today) > 20, and neural signal output(today) > neural signal output(yesterday)

Sell: Sell tomorrow when neural signal output(today) <= 20, and neural signal output(today) < neural signal output(yesterday)

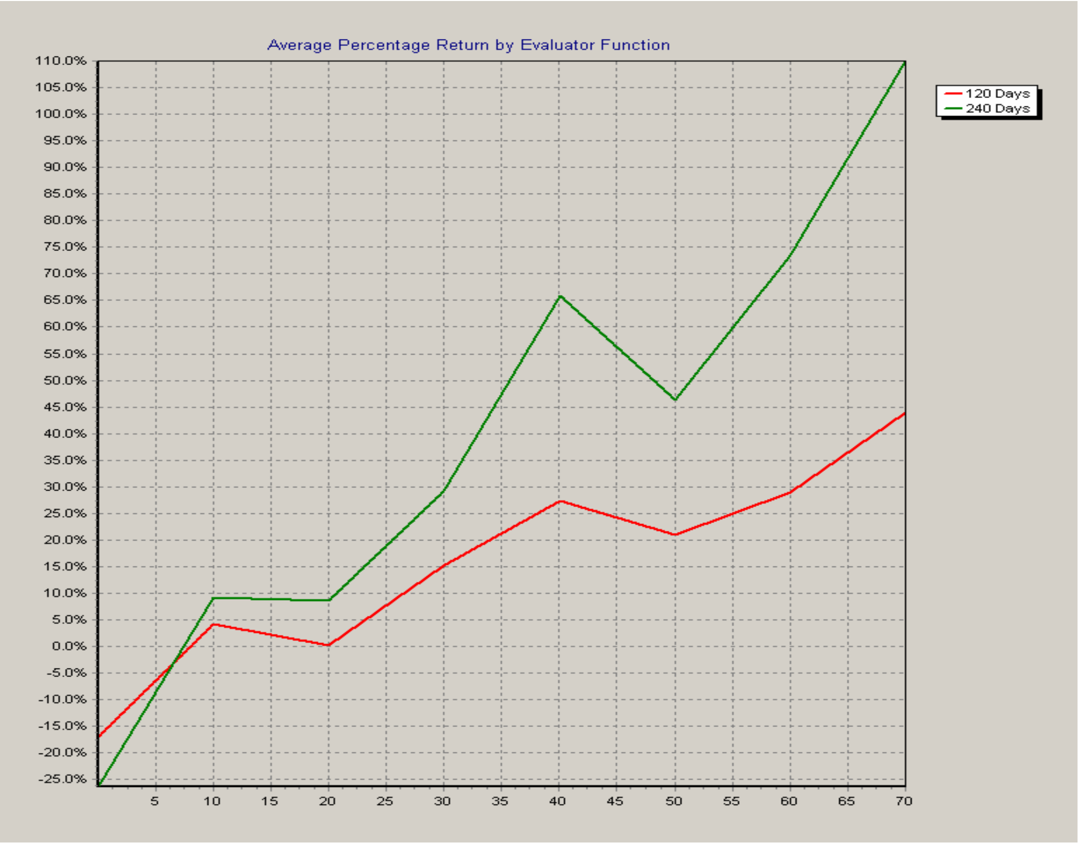

Figure 4-4 FNN(ASX200): In-sample output signal strength graph

4.3.3.2 Stop-Loss threshold

The following figures show the MAE for the in-sample trades.

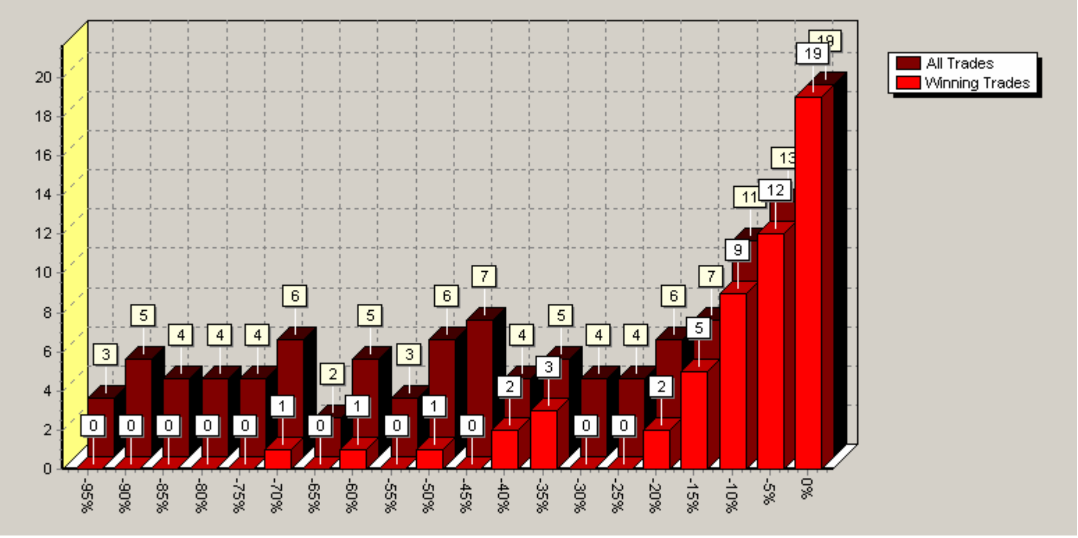

Figure 4-5 TS-FNN(ASX200): In-sample MAE

The MAE shows that very few of the in-sample trades which go on to become profitable incur losses of 45% or more during the trades' lifetime (the period between opening and closing the trade). The stop could have been set at 25% or 30%, however, again Chande's guideline of setting wide stops for longer-term systems is followed.

By inspection of the MAE graph, a stop-loss threshold of 45% can be chosen. This value will then be used as the stop-loss threshold for the out-of-sample testing.

4.3.4 Out-of-Sample Results

The following tables report the out-of-sample results of the trading system TS-FNN(ASX200), built around the FNN(ASX200) neural network. The out-of-sample period is two years, from the start of trading in January 2002, to the end of trading in December 2003. In this section, results are displayed for the dataset including delisted stocks, as well as for the dataset excluding delisted stocks. As the realistic case is to include delisted stocks, only these results will be included in the discussions. The results are presented for the dataset which excludes delisted stocks for interest only.

Metric	**TS-FNN(ASX200) includes delisted**	**TS-FNN(ASX200) excludes delisted**	**Buy-and-Hold includes delisted**
Net Profit (Loss)	$ 3,324.36	$ 2,337.69	$ 1,100.54
Net Profit (Loss) %	3.32 %	2.34 %	1.10 %
Annualized Gain %	1.65 %	1.17 %	0.55 %
Exposure %	7.14 %	6.54 %	100 %
All Trades: Number of trades	17	15	251

Metric	**TS-FNN(ASX200) includes delisted**	**TS-FNN(ASX200) excludes delisted**	**Buy-and-Hold includes delisted**
All Trades: Average Profit (Loss), Average Profit (Loss) %	$ 195.54 19.42 %	$ 155.85 15.47 %	$ 4.38 1.01 %
All Trades: Average Bars Held	197.12	205.4	397.76
Winning Trades: Number of Trades	12	10	125
Winning Trades: Winning %	70.59 %	66.67 %	49.80 %
Winning Trades: Average Profit, Average Profit %	$ 370.31 37.02 %	$ 345.71 34.62 %	$ 147.50 36.71 %
Winning Trades: Average Bars Held	223	240.9	418.66
Winning Trades: Max consecutive wins	4	5	7
Losing Trades: Number of Trades	5	5	126
Losing Trades: Losing %	29.41 %	33.33 %	50.20 %
Losing Trades: Average Loss, Average Loss %	$ 223.88 22.83 %	$ 223.88 22.83 %	$ 137.59 34.41 %
Losing Trades: Average Bars Held	135	134.4	377.03

Metric	TS-FNN(ASX200) includes delisted	TS-FNN(ASX200) excludes delisted	Buy-and-Hold includes delisted
Losing Trades: Max consecutive losses	2	2	7
Max Drawdown, Max Drawdown %, Maximum Drawdown Date	$ 2,373.26 2.35 % 13/03/2003	$ 2,371.16 2.35 % 13/03/2003	$ 17,495.25 17.37 % 13/03/2003
Profit Factor (PF)	3.9697	3.0833	1.0635
Recovery Factor (RF)	1.4007	0.9859	0.0629
Payoff Ratio (PR)	1.6218	1.5166	1.0667
Sharpe Ratio (SR)	0.764	0.577	- 0.0152
Ulcer Index	0.5403	0.6063	8.5491
Luck Coefficient (LC)	2.1692	2.3196	6.0283
Pessimistic Rate of Return (PRR)	1.9131	1.4331	0.8847

Table 4-13 TS-FNN(ASX200): Out-of-sample trading metrics

Table 4-14 shows a statistical breakdown of the out-of-sample trades from the TS-FNN(ASX200). The means of the net profit/loss are tested against the mean of the distribution curve that a random trading strategy would produce, which is assumed to be zero under the null hypothesis of no excess returns.

The hypotheses for the t-tests will be:

H_0: $\mu_{profit} = 0$,

H_1: $\mu_{profit} > 0$

Metric	TS-FNN(ASX200) includes delisted
Sample size	17
Sample Mean	195.5447
Sample Standard Deviation (SD)	383.33717
Standard Error of the Mean	92.9792
t-statistic (P/L > 0)	2.103
Degrees of freedom (df)	16
t-statistic (5%,df) 1-tailed to exceed	1.746
Lower 95% confidence interval of the Mean	-1.5491
Upper 95% confidence level of the Mean	392.6385

Table 4-14 TS-FNN(ASX200): Statistical analysis of mean profit/loss

The t-statistic for the TS-FNN(ASX200) trades allows us to reject the null hypothesis and conclude that the mean profit of the trading system is significantly greater than zero. Specifically, $t(16) = +2.103$, $p = 0.025$ ($p < 0.05$), one tailed.

The results of Vince's runs test for dependency is shown in Table 4-15 below.

Metric	**TS-FNN(ASX200) includes delisted**
Total Cases	17
Number of Runs	9
Z score	0.88

Table 4-15 TS-FNN(ASX200): Runs test results

In the Runs Test, we can conclude that there is insufficient evidence to accept dependency amongst the trades. According to Vince (1990), a negative Z-score implies positive dependency, meaning fewer streaks than the normal probability function would imply, and hence that wins beget wins and losses beget losses.

Finally, the set of trades from the TS-FNN(ASX200) strategy are compared to the set of trades from the buy-and-hold approach, using the ANOVA technique. The results are reported in Table 4-16 below.

	TS-FNN(ASX200) includes delisted	**Buy-and-Hold incl delisted**
Mean	195.54	4.38
Standard Deviation	383.33	206.72

Table 4-16 TS-FNN(ASX200): Trades compared to Buy-and-Hold trades

The analysis of variance reveals that the trades selected using TS-FNN(ASX200) are significantly different from the trades from the buy-and-hold approach, specifically $F(1,266) = 11.873$, $p=0.001$ ($p< 0.05$).

Finally, a brief presentation of the distribution of monthly returns from the strategy is included. This gives additional insight into how the trading system can be expected to perform into the future, and can also be used by traders to establish confidence intervals on monthly returns. This is helpful in identifying whether a model is performing within expectations, and can also aid in detecting when in the future a model has ceased to provide an advantage, and needs to be retrained.

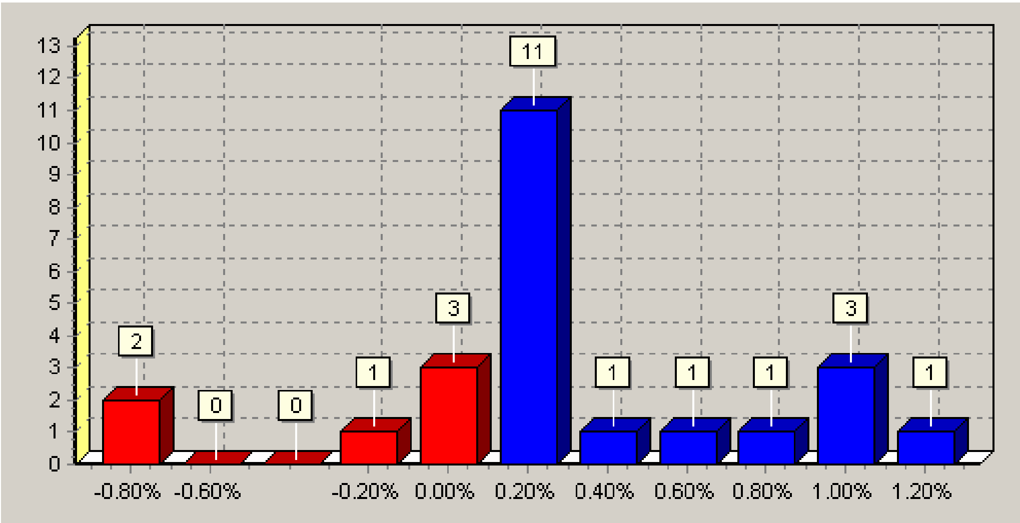

Figure 4-6 TS-FNN(ASX200): Distribution of monthly Returns

4.3.5 Summary

Although returns from this trading system, TS-FNN(ASX200) are not stellar, they are significantly better than the returns achieved by the buy-and-hold benchmark, with much smaller exposure. As shown by the results discussed above, it can also be concluded that the mean profit of the trades selected by the ANN at the heart of this trading system, FNN(ASX200), are significantly different from the mean profit of the trades which constitute the buy-and-hold strategy. Therefore, we can confidently state that the trading system which embodies this neural network has outperformed the buy-and-hold approach.

4.4 Trading in the ASX Allshare using neural networks and technical data

The neural network TNN(Allshare) has been trained to signal abnormal short-term trading opportunities. It was trained with technical data, using a 3 day lookahead period.

Results from key stages in the development of the neural network are presented below. To test the neural network properly, it must be sited within a valid trading context. The trading system developed around this neural network is named TS-TNN(Allshare), and this trading system is thoroughly tested out-of-sample, and the results are also presented below.

4.4.1 Training data

Due to the large volume of training data available, and the need to fit the training data within main memory, a simple sampling strategy was implemented. The sampling strategy was to select every 50^{th} row in each training dataset as a candidate for a neural input row. Every 50^{th} row was then passed to the neural network as long as it did not contain outlier values for any of the technical variables. Rows that contained outlier values were dropped. Outliers were easily identified from the function profiles for each technical variable already created and displayed in Appendix A. Table 4-17 shows the acceptable range, and outlier range as determined for each technical variable. These were selected from the study of the function profiles; all category rows in the function profiles that did not contain at least 1% of the available observations were deemed outliers. In practice, from a study of the rows rejected by this process, it was found that if one technical variable in the row was an outlier, it was likely that several of the other technical variables in the same row were also outliers.

Technical Variable	Acceptable Range	Outliers
SMA(close,3) / SMA(close,15)	> 0.5 to <= 1.5	<= 0.5 or > 1.5
ATR(3) / ATR(15)	> 0 to <= 2.5	<= 0 or > 2.5
SMA(Volume,3) / SMA(Volume,15)	> 0 to <= 3.5	<= 0 or > 3.5
ADX(3)	> 10	<= 10
ADX(15)	> 0 to <= 60	<= 0 or > 60
STOCHK(3)	ANY	none
STOCHK(15)	ANY	none
STOCHK(3) / STOCHK(15)	<= 200	> 200
MOM(3)	> -20 to <= 20	<= -20 or > 20
MOM(15)	> -20 to <= 20	<= -20 or > 20
MOM(3) / MOM(15)	> -50 TO <= 50	<= -50 or > 50
RSI(3)	> 0	<= 0
RSI(15)	> 20 to <= 80	<= 20 or > 80
RSI(3) / RSI(15)	> 0 to <= 2	<= 0 or > 2
MACD	> -5 to <= 5	<= -5 or > 5
LPR	ANY	none
HPR	ANY	none

Table 4-17 TNN(Allshare): Identification of outliers for technical variables

Using this approach, a total of 19458 rows were input to the neural network for training purposes. The characteristics of these input rows are shown in Table 4-18.

Technical Variable	Min	Max	Mean	Stddev
SMA(close,3) / SMA(close,15)	0.51	1.50	1.00	0.08
ATR(3) / ATR(15)	0.00	2.48	0.94	0.38
SMA(Volume,3) / SMA(Volume,15)	0.00	3.50	0.97	0.63
ADX(3)	10.10	100.00	54.03	20.23
ADX(15)	1.31	59.92	23.42	9.86
STOCHK(3)	0.00	100.00	43.14	41.21
STOCHK(15)	0.00	100.00	45.70	33.72
STOCHK(3) / STOCHK(15)	0.00	75.00	1.00	1.41
MOM(3)	-12.00	5.98	0.00	0.26
MOM(15)	-17.99	14.94	-0.01	0.51
MOM(3) / MOM(15)	-45.00	45.00	0.18	1.50
RSI(3)	0.00	100.00	47.78	26.21
RSI(15)	20.23	80.00	49.31	10.64
RSI(3) / RSI(15)	0.00	2.00	0.93	0.43
MACD	-4.88	4.99	0.00	0.15
LPR	0.01	1.00	0.73	0.22
HPR	0.01	1.00	0.69	0.25

Table 4-18 TNN(Allshare): Characteristics of input data

Table 4-19 shows the same summary information for the training target variable. This target is the maximum percentage change in price over the next three days, computed for every element i in the input series as:

$$\text{Target} = \left(\frac{\left(\max(close_{i+3}, close_{i+2}, close_{i+1}) - close_i \right)}{close_i} \right) \times 100$$

Equation 4-3 TNN(Allshare): Computing the target variable

Effectively, this target allows the neural network to focus on the relationship between the input technical variables, and the expected forward price change.

Training Target	Min	Max	Mean	Stddev
Output Target	-85.71	200.00	2.99	9.84

Table 4-19 TNN(Allshare): Characteristics of training target

4.4.2 Architecture selected

Each ANN was trained until no new error low had been reached for at least 2000 epochs. The architecture to be carried forward to out-of-sample testing is selected by reference to the in-sample metrics previously detailed. Table 4-20 displays the computed in-sample metrics.

ANN Architecture	Expectancy	Average Profit/Loss % (3 days)
4 hidden nodes	0.76	9.67%
5 hidden nodes	0.92	12.22%
6 hidden nodes	0.61	8.57%

Table 4-20 TNN(Allshare): In-sample metric values

From the above table, the architecture with 5 hidden nodes is selected. In this research, the neural network with this architecture proceeds on to trading system development.

4.4.3 Derivation of Trading System parameters

4.4.3.1 Signal Strength Threshold

By inspection of the function profile graph below, a signal strength threshold of 40 must be exceeded to initiate trades. The value of 40 is chosen as it is the point at which the neural network begins to signal trades with net positive results. This can be seen in the graph as it represents the signal strength value for which the expected returns are greater than zero.

The trading rules component of the trading system for this neural network consists of the following buy/sell rules:

Buy: Buy tomorrow when neural signal output(today) > 40, and neural signal output(today) > neural signal output(yesterday)

Sell: Sell tomorrow when neural signal output(today) <= 40, and neural signal output(today) < neural signal output(yesterday)

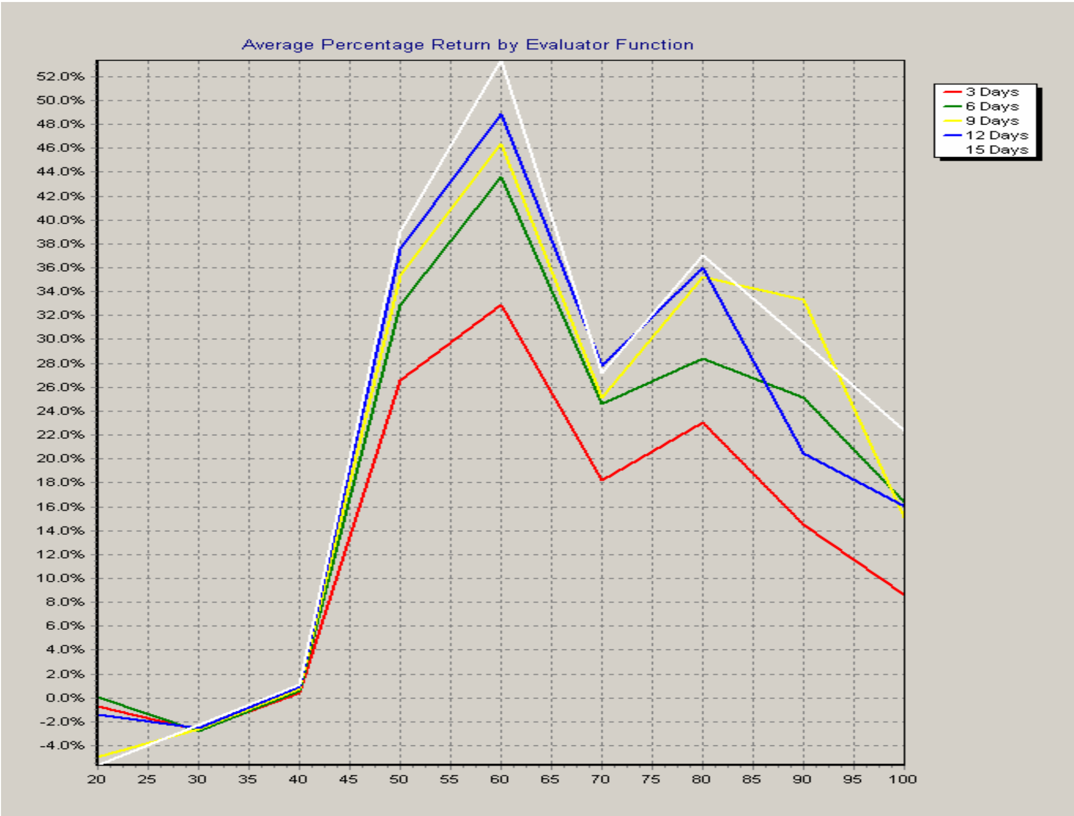

Figure 4-7 TNN(Allshare): In-sample output signal strength graph

4.4.3.2 Stop-Loss threshold

The following figures show the MAE for the in-sample trades.

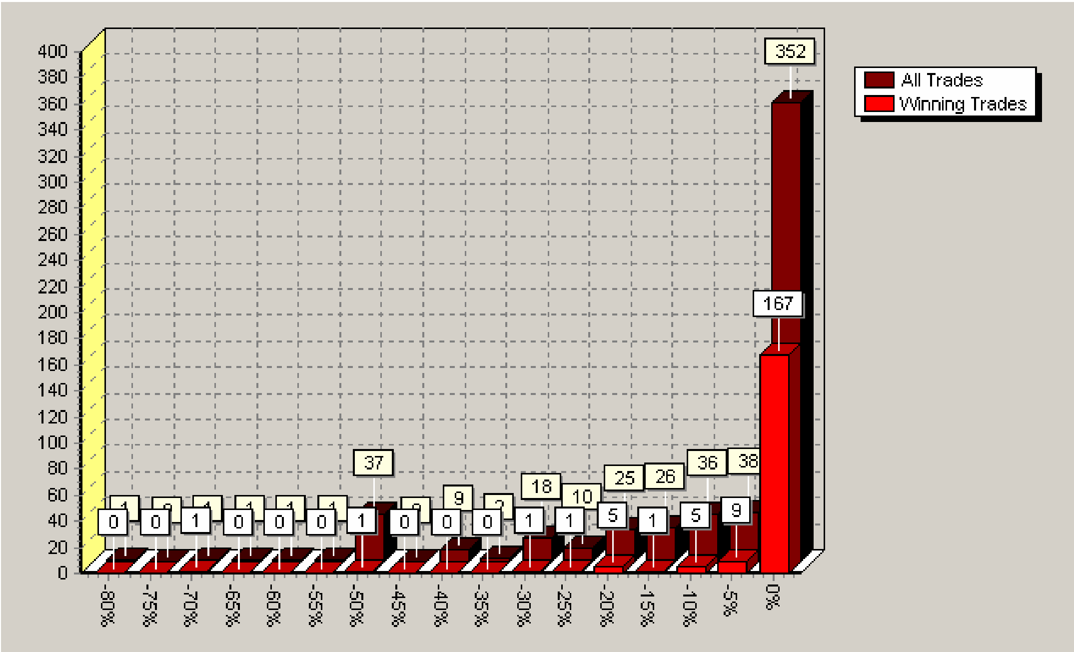

Figure 4-8 TS-TNN(Allshare): In-sample MAE

The MAE shows that very few of the in-sample trades which go on to become profitable incur losses of 5% or more during the trades' lifetime (the period between opening and closing the trade).

By inspection of the MAE graph, a stop-loss threshold of 5% can be chosen. This value will then be used as the stop-loss threshold for the out-of-sample testing.

4.4.4 Out-of-Sample Results

The following tables report the out-of-sample results of the trading system TS-TNN(Allshare), built around the TNN(Allshare) neural network. The out-of-sample period is two years, from the start of trading in January 2002, to the end of trading in December 2003. In this section, results are displayed for the dataset including delisted stocks, as well as for the dataset excluding delisted stocks. As the realistic case is to include delisted stocks, only these results will be included in the discussions. The results are presented for the dataset which excludes delisted stocks for interest only.

Metric	TS-TNN(Allshare) includes delisted	TS-TNN(Allshare) excludes delisted	Buy-and-Hold includes delisted
Net Profit (Loss)	$ 66,119.18	$ 65,712.45	$ 32,379.12
Net Profit (Loss) %	66.12 %	65.71 %	32.38 %
Annualized Gain %	29.00 %	28.84 %	15.11 %
Exposure %	6.39 %	6.28 %	96.71 %
All Trades: Number of trades	1,028	1,021	1,367
All Trades: Average Profit (Loss), Average Profit (Loss) %	$ 64.32 6.53 %	$ 64.36 6.55 %	$ 23.69 32.74 %
All Trades: Average Bars Held	2.85	2.72	472.4
Winning Trades: Number of Trades	281	278	682
Winning Trades: Winning %	27.33 %	27.23 %	49.89 %
Winning Trades: Average Profit, Average Profit %	$ 545.18 41.90 %	$ 548.09 42.21 %	$ 82.28 113.07 %
Winning Trades: Average Bars Held	3.47	3.28	493.13
Winning Trades: Max consecutive wins	5	5	10
Losing Trades: Number of Trades	747	743	685

Metric	**TS-TNN(Allshare) includes delisted**	**TS-TNN(Allshare) excludes delisted**	**Buy-and-Hold includes delisted**
Losing Trades: Losing %	72.67 %	72.77 %	50.11 %
Losing Trades: Average Loss, Average Loss %	$ 116.57 6.78 %	$ 116.63 6.79 %	$ 34.65 47.24 %
Losing Trades: Average Bars Held	2.61	2.51	451.75
Losing Trades: Max consecutive losses	20	19	11
Max Drawdown, Max Drawdown %, Maximum Drawdown Date	$ 8,775.68 6.43 % 17/04/2003	$ 8,602.29 6.32 % 17/04/2003	$ 22,390.84 20.86 % 13/03/2003
Profit Factor (PF)	1.7593	1.7583	2.3641
Recovery Factor (RF)	7.5344	7.6389	1.4461
Payoff Ratio (PR)	6.1824	6.2134	2.3925
Sharpe Ratio (SR)	2.96	2.95	0.7623
Ulcer Index	1.8853	1.8624	11.3687
Luck Coefficient (LC)	2.3164	2.2997	9.6076
Pessimistic Rate of Return (PRR)	2.1097	2.108	2.2075

Table 4-21 TS-TNN(AllShare): Out-of-sample trading metrics

Table 4-22 shows a statistical breakdown of the out-of-sample trades from the TS-TNN(Allshare). The means of the net profit/loss are tested against the mean of the distribution curve that a random trading strategy would produce, which is assumed to be zero under the null hypothesis of no excess returns.

The hypotheses for the t-tests will be:

H_0: $\mu_{profit} = 0$,

H_1: $\mu_{profit} > 0$

Metric	**TS-TNN(Allshare) includes delisted**
Sample size	1028
Sample Mean	64.3183
Sample Standard Deviation (SD)	446.8143
Standard Error of the Mean	13.9357
t-statistic (P/L > 0)	4.615
Degrees of freedom (df)	1027
t-statistic (5%,df) 1-tailed to exceed	1.646
Lower 95% confidence interval of the Mean	36.9725
Upper 95% confidence level of the Mean	91.6641

Table 4-22 TS-TNN(Allshare): Statistical analysis of mean profit/loss

The t-statistic for the TS-TNN(Allshare) trades allows us to reject the null hypothesis and conclude that the mean profit of the trading system is significantly greater than zero. Specifically, $t(1027) = +4.615$, $p = 0.00$ ($p < 0.05$), one tailed.

The results of Vince's runs test for dependency is shown in Table 4-23 below.

Metric	**TS-TNN(Allshare) includes delisted**
Total Cases	1028
Number of Runs	402
Z score	-0.54

Table 4-23 TS-TNN(Allshare): Runs test results

In the Runs Test, we can conclude that there is insufficient evidence to accept dependency amongst the trades. According to Vince (1990), a negative Z-score implies positive dependency, meaning fewer streaks than the normal probability function would imply, and hence that wins beget wins and losses beget losses.

Finally, the set of trades from the TS-TNN(Allshare) strategy are compared to the set of trades from the buy-and-hold approach, using the ANOVA technique. The results are reported in Table 4-24 below.

	TNN Allshare incl delisted	**Buy-and-Hold incl delisted**
Mean	64.318	23.686
Standard Deviation	446.8143	95.9704

Table 4-24 TS-TNN(Allshare): Trades compared to Buy-and-Hold trades

The analysis of variance reveals that the trades selected using TS-TNN(Allshare) are significantly different from the trades from the buy-and-hold approach, specifically $F(1,2395) = 10.652$, $p = 0.001$ ($p < 0.05$).

Finally, a brief presentation of the distribution of monthly returns from the strategy is included. This gives additional insight into how the trading system can be expected to perform into the future, and can also be used by traders to establish confidence intervals on monthly returns. This is helpful in identifying whether a model is performing within expectations, and can also aid in detecting when in the future a model has ceased to provide an advantage, and needs to be retrained.

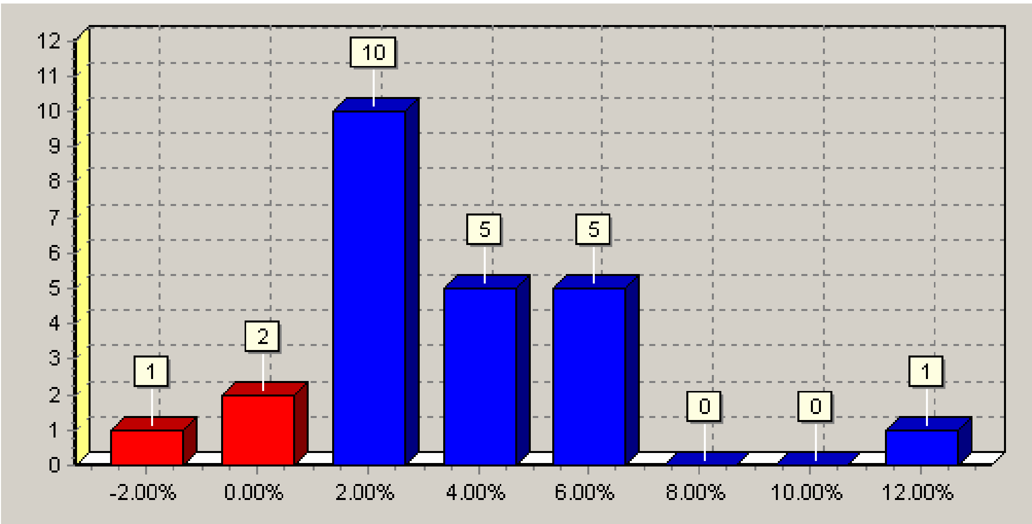

Figure 4-9 TS-TNN(Allshare): Distribution of monthly returns

The distribution presented above shows significant positive skewness. For a normal distribution the skewness value is zero, whilst a positively skewed distribution has a long right tail. The standard method of assessing skewness is to compare the degree of skewness to the standard error of skewness. A skewness value more than twice its standard error is taken to indicate a departure from normality. For the distribution presented above, the skewness value is +1.067, with a standard error of skewness of +0.472. As the skewness value is more than twice its standard error, it is concluded that the distribution of monthly returns from this trading strategy is positively skewed.

4.4.5 Summary

Overall, returns from the trading system, TS-TNN(Allshare) are excellent. As shown by the results discussed above, it can be concluded that the mean profit of the trades selected by the ANN at the heart of this trading system, TNN(Allshare), are significantly different from the mean profit of the trades which constitute the buy-and-hold strategy. Therefore, we can confidently state that the trading system which embodies this neural network has outperformed the buy-and-hold approach.

4.5 Trading in the S&P ASX200 using neural networks and technical data

The neural network TNN(ASX200) has been trained to signal abnormal short-term trading opportunities. It was trained with technical data, using a 3 day lookahead period. Results from key stages in the development of the neural network are presented below. To test the neural network properly, it must be sited within a valid trading context. The trading system developed around this neural network is named TS-TNN(ASX200), and this trading system is thoroughly tested out-of-sample, and the results are also presented below.

4.5.1 Training data

Due to the large volume of training data available, and the need to fit the training data within main memory, a systematic sampling strategy was implemented. The sampling strategy was to select every 20th row in each training dataset as a candidate for a neural input row. Every 20th row was then passed to the neural network as long as it did not contain outlier values for any of the technical variables. Rows that contained outlier values were dropped. Outliers were easily identified as function profiles for each technical variable were created, similar to those already presented for the TNN(Allshare) in Appendix A. Due to the amount of space required to present this data for the ASX200 dataset, it was decided to exclude these results from this thesis. They are available separately if required. It was interesting to see that the same technical variables which showed promise in the Allshare dataset also showed promise in the ASX200 dataset.

Also, the three technical variables which did not prove useful in the Allshare datset also did not prove useful in the ASX200 dataset. Table 4-25 shows the acceptable range, and outlier range as determined for each technical variable. These were selected from the study of the function profiles; all category rows in the function profiles that did not contain at least 1% of the available observations were deemed outliers. In practice, from a study of the rows rejected by this process, it was found that if one technical variable in the row was an outlier, it was likely that several of the other technical variables in the same row were also outliers.

Technical Variable	Acceptable Range	Outliers
SMA(close,3) / SMA(close,15)	> 0.5 to <= 1.5	<= 0.5 or > 1.5
ATR(3) / ATR(15)	> 0 to <= 2	<= 0 or > 2
SMA(Volume,3) / SMA(Volume,15)	> 0 to <= 3	<= 0 or > 3
ADX(3)	> 10	<= 10
ADX(15)	> 0 to <= 60	<= 0 or > 60
STOCHK(3)	ANY	none
STOCHK(15)	ANY	none
STOCHK(3) / STOCHK(15)	<= 20	> 20
MOM(3)	> -10 to <= 10	<= -10 or > 10
MOM(15)	> -10 to <= 10	<= -10 or > 10
MOM(3) / MOM(15)	> -25 to <= 25	<= -25 or > 25
RSI(3)	> 0	<= 0
RSI(15)	> 20 to <= 80	<= 20 or > 80
RSI(3) / RSI(15)	> 0 to <= 2	<= 0 or > 2
MACD	> -2.5 to <= 2.5	<= -2.5 or > 2.5
LPR	ANY	none
HPR	ANY	none

Table 4-25 TNN(ASX200): Identification of outliers for technical variables

Using this approach, a total of 6,857 rows were input to the neural network for training purposes. The characteristics of these input rows are shown in Table 4-26.

Technical Variable	Min	Max	Mean	Stddev
SMA(close,3) / SMA(close,15)	0.88	1.46	1.02	0.03
ATR(3) / ATR(15)	0.01	2.47	0.98	0.26
SMA(Volume,3) / SMA(Volume,15)	0.02	2.99	0.97	0.54
ADX(3)	10.36	100.00	52.65	18.46
ADX(15)	6.05	59.81	23.93	9.50
STOCHK(3)	0.00	100.00	59.61	34.56
STOCHK(15)	0.00	100.00	72.02	22.21
STOCHK(3) / STOCHK(15)	0.00	7.00	0.84	0.57
MOM(3)	-3.21	3.77	0.05	0.24
MOM(15)	0.01	9.41	0.26	0.45
MOM(3) / MOM(15)	-17.00	24.00	0.21	1.79
RSI(3)	1.01	100.00	61.82	23.20
RSI(15)	26.44	79.91	58.35	7.90
RSI(3) / RSI(15)	0.02	1.93	1.05	0.35
MACD	-1.09	2.33	0.06	0.14
LPR	0.09	0.99	0.78	0.14
HPR	0.05	1.00	0.89	0.13

Table 4-26 TNN(ASX200): Characteristics of input data

Table 4-27 shows the same summary information for the training target variable. This target is the maximum percentage change in price over the next three days, computed for every element i in the input series as:

$$\text{Target} = \left(\frac{\left(\max(close_{i+3}, close_{i+2}, close_{i+1}) - close_i\right)}{close_i} \right) \times 100$$

Equation 4-4 TNN(ASX200): Computing the target variable

Effectively, this target allows the neural network to focus on the relationship between the input technical variables, and the expected forward price change.

Training Target	Min	Max	Mean	Stddev
Output Target	-13.76	50.36	1.37	3.16

Table 4-27 TNN(ASX200): Characteristics of training target

4.5.2 Architecture selected

Each ANN was trained until no new error low had been reached for at least 2000 epochs. The architecture to be carried forward to out-of-sample testing is selected by reference to the in-sample metrics previously detailed. Table 4-28 displays the computed in-sample metrics.

ANN Architecture	Expectancy	Average Profit/Loss %
4 hidden nodes	-0.41	-1.54%
5 hidden nodes	-0.35	-1.47%
6 hidden nodes	-0.48	-1.79%

Table 4-28 TNN(ASX200): In-sample metric values

All architectures tested showed negative expectancy and negative Average Profit/Loss %. However, the ANNs are being tested without a stop-loss, so it remains to be seen if the best architecture will become profitable when stop-losses are introduced. From the above table, the architecture with 5 hidden nodes is selected. In this research, the neural network with this architecture proceeds on to trading system development.

4.5.3 Derivation of Trading System parameters

4.5.3.1 Signal Strength Threshold

By inspection of the function profile graph below, a signal strength threshold of 22 must be exceeded to initiate trades. The value of 22 is chosen as it is the point at which the neural network begins to signal trades with net positive results. This can be seen in the graph as it represents the signal strength value for which the expected returns are greater than zero and increasing. By studying the slope of the 3-day returns shown in the function profile, it is clear that increasing values of the ANN signal correspond to increasing expected returns. However, the amount of those returns is extremely low.

The trading rules component of the trading system for this neural network consists of the following buy/sell rules:

Buy: Buy tomorrow when neural signal output(today) > 22, and neural signal output(today) > neural signal output(yesterday)

Sell: Sell tomorrow when neural signal output(today) <= 22, and neural signal output(today) < neural signal output(yesterday)

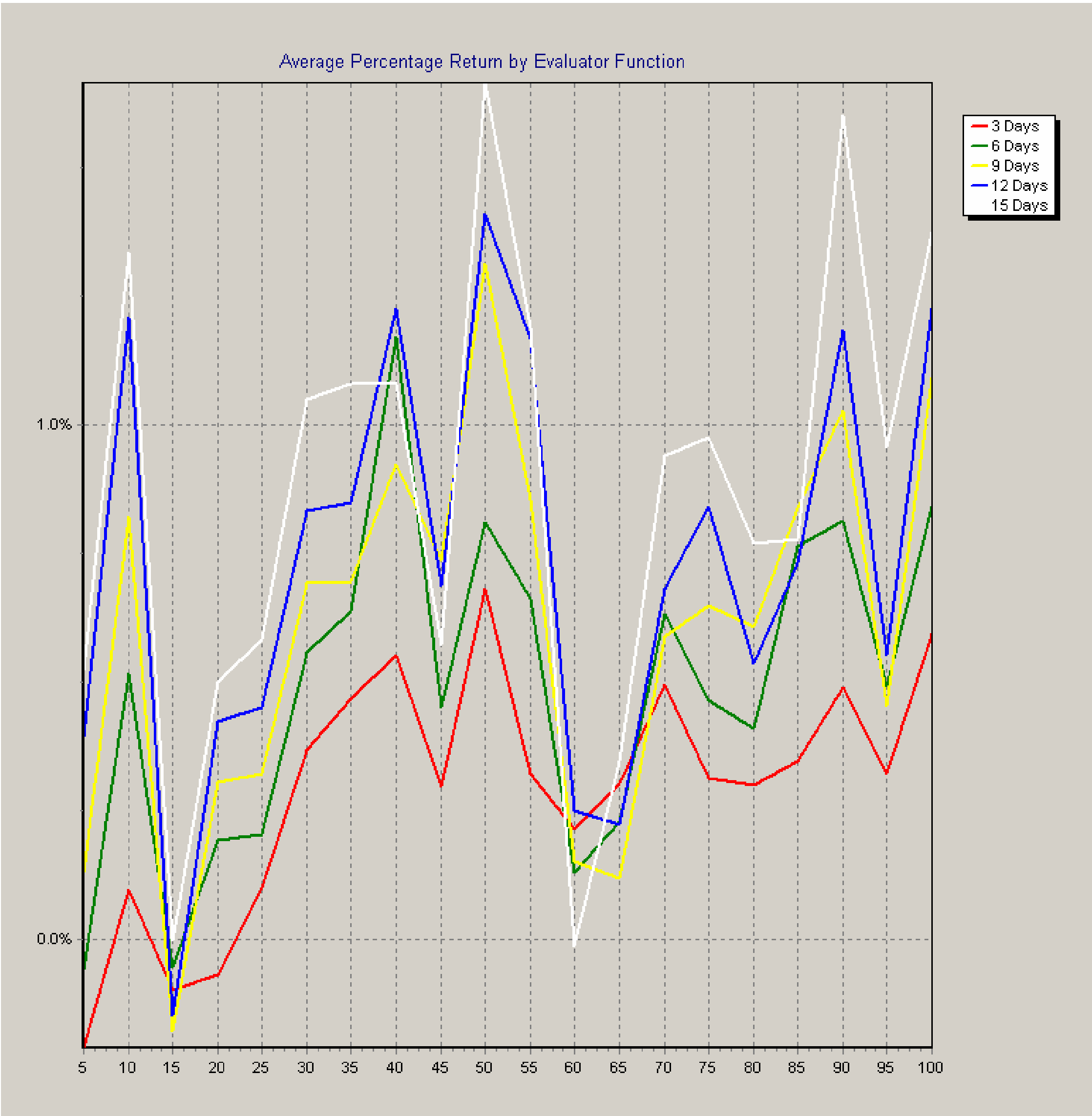

Figure 4-10 TNN(ASX200): In-sample output signal strength graph

4.5.3.2 Stop-Loss threshold

The following figure shows the MAE for the in-sample trades.

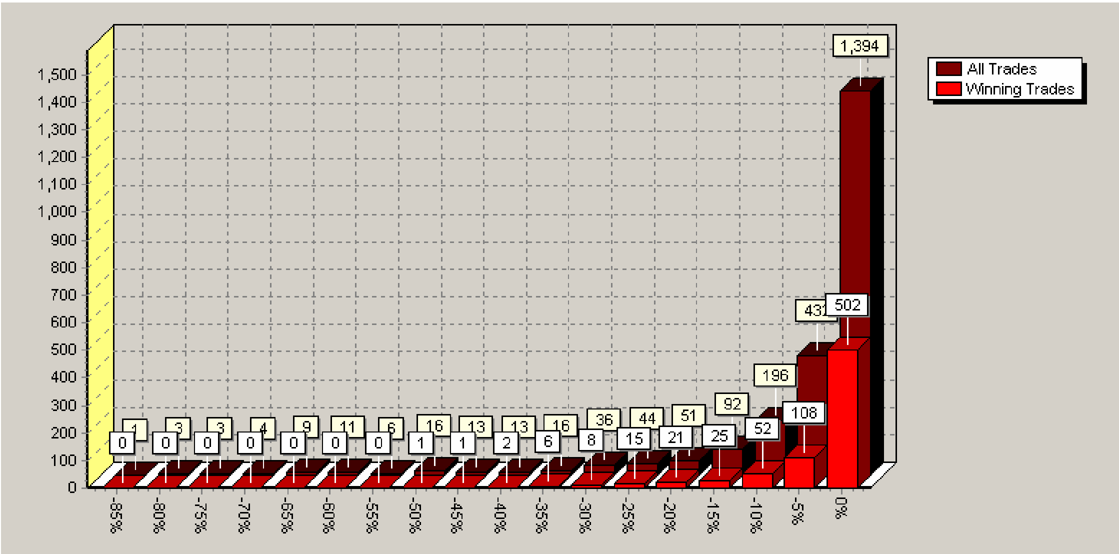

Figure 4-11 TS-TNN(ASX200): In-sample MAE

The MAE shows that few of the in-sample trades which go on to become profitable incur losses of 5% or more during the trades' lifetime (the period between opening and closing the trade).

By inspection of the MAE graph, a stop-loss threshold of 5% can be chosen. This value will then be used as the stop-loss threshold for the out-of-sample testing.

4.5.4 Out-of-Sample Results

The following tables report the out-of-sample results of the trading system TS-TNN(ASX200), built around the TNN(ASX200) neural network. The out-of-sample period is two years, from the start of trading in January 2002, to the end of trading in December 2003. In this section, results are displayed for the dataset including delisted stocks, as well as for the dataset excluding delisted stocks. As the realistic case is to include delisted stocks, only these results will be included in the discussions. The results are presented for the dataset which excludes delisted stocks for interest only.

Metric	**TS-TNN(ASX200) includes delisted**	**TS-TNN(ASX200) excludes delisted**	**Buy-and-Hold includes delisted**
Net Profit (Loss)	-$ 13,670.08	- $ 12,894.12	$ 1,100.54
Net Profit (Loss) %	-13.67 %	- 12.89 %	1.10 %
Annualized Gain %	-7.11 %	- 6.69 %	0.55 %
Exposure %	92.65 %	92.62 %	100 %
All Trades: Number of trades	847	829	251
All Trades: Average Profit (Loss), Average Profit (Loss) %	- $ 16.14 - 1.65 %	- $ 15.55 - 1.48 %	$ 4.38 1.01 %
All Trades: Average Bars Held	45.1	46.64	397.76
Winning Trades: Number of Trades	143	153	125
Winning Trades: Winning %	16.88 %	18.46 %	49.80 %
Winning Trades: Average Profit, Average Profit %	$ 215.17 26.14 %	$ 191.70 23.93 %	$ 147.50 36.71 %
Winning Trades: Average Bars Held	156.54	151.01	418.66
Winning Trades: Max consecutive wins	4	9	7

Metric	TS-TNN(ASX200) includes delisted	TS-TNN(ASX200) excludes delisted	Buy-and-Hold includes delisted
Losing Trades: Number of Trades	704	676	126
Losing Trades: Losing %	83.12 %	81.54 %	50.20 %
Losing Trades: Average Loss, Average Loss %	$ 63.12 7.29 %	$ 62.46 7.23 %	$ 137.59 34.41 %
Losing Trades: Average Bars Held	22.46	23.02	377.03
Losing Trades: Max consecutive losses	42	40	7
Max Drawdown, Max Drawdown %, Maximum Drawdown Date	$ 30,654.99 30.57 % 13/03/2003	$ 31,776.27 31.73 % 13/03/2003	$ 17,495.25 17.37 % 13/03/2003
Profit Factor (PF)	0.6924	0.6946	1.0635
Recovery Factor (RF)	0.4459	0.4058	0.0629
Payoff Ratio (PR)	3.5833	3.3088	1.0667
Sharpe Ratio (SR)	-0.8213	-0.7844	- 0.0152
Ulcer Index	18.5497	19.0181	8.5491
Luck Coefficient (LC)	6.9198	5.862	6.0283
Pessimistic Rate of Return (PRR)	0.6428	0.6628	0.8847

Table 4-29 TS-TNN(ASX200): Out-of-sample trading metrics

4.5.5 Summary

Considering the poor outcome achieved by this trading system, TS-TNN(ASX200), the usual post out-of-sample diagnostics are omitted in favor of a more detailed analysis of the outcome.

Overall, there are several possibilities for this poor result. Although each of these possibilities is considered separately below, they are not mutually exclusive.

4.5.5.1 Non-Causal inputs

One possible explanation for the poor result from this trading system, TS-TNN(ASX200), is that the inputs supplied to the neural network, TNN(ASX200) were not useful in explaining / predicting abnormal returns.

Although the function profile of the ANNs performance over the in-sample set showed that higher returns were associated with higher network outputs, it was noted earlier that the average returns were still very poor. Therefore, it could be concluded that the inputs supplied to the ANN did not help in identifying stocks which were likely to outperform.

4.5.5.2 Lack of movement in the market

As stated above, over the in-sample period, returns were very poor. The same is true of the out-of-sample period; indeed, returns were even lower, as the S&P/ASX200 stagnated and was listless during much of the out-of-sample period, as evidenced by the trivial returns to the buy-and-hold approach.

Therefore, another possible explanation is that although the ANN was capable of detecting stocks that increased in value, as evidenced by the function profile, there were not enough stocks achieving significant increases during the out-of-sample period to allow the trading system to prosper.

Figure 4-12 and Figure 4-13 show the equity curve and the monthly returns for the trading system TS-TNN(ASX200) during the out-of-sample period. The blue line represents the standardized buy-and-hold returns.

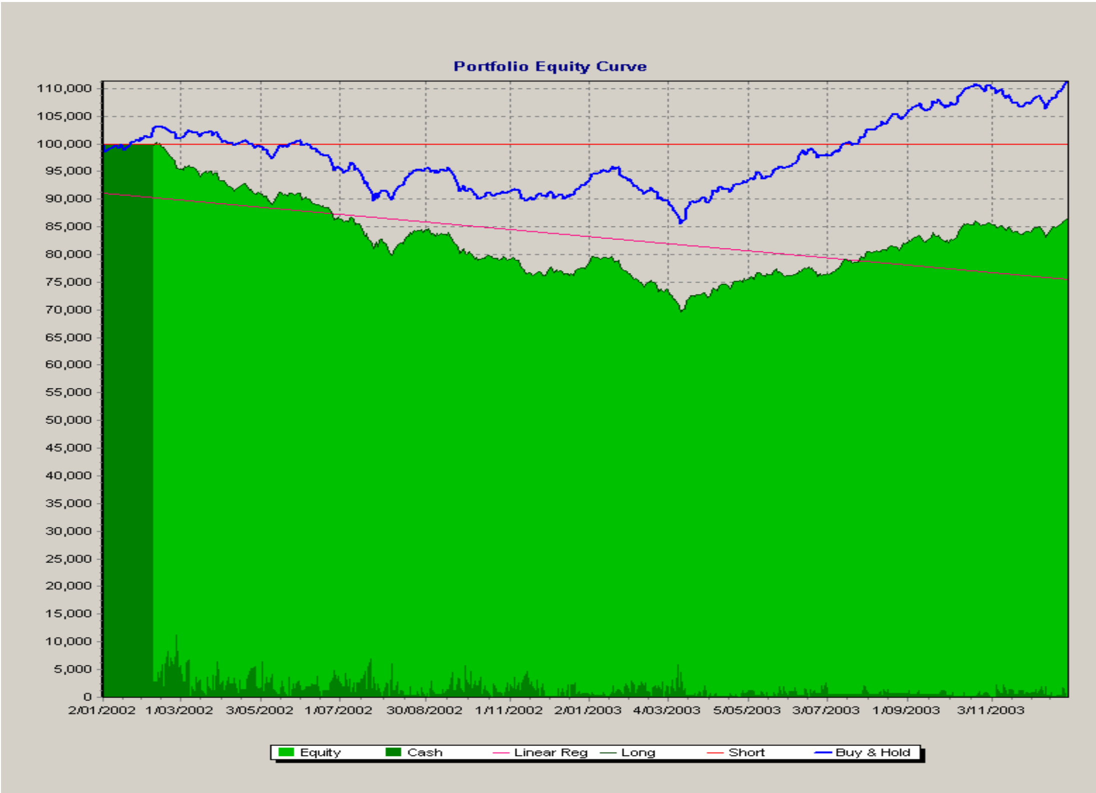

Figure 4-12 TS-TNN(ASX200) Equity Curve

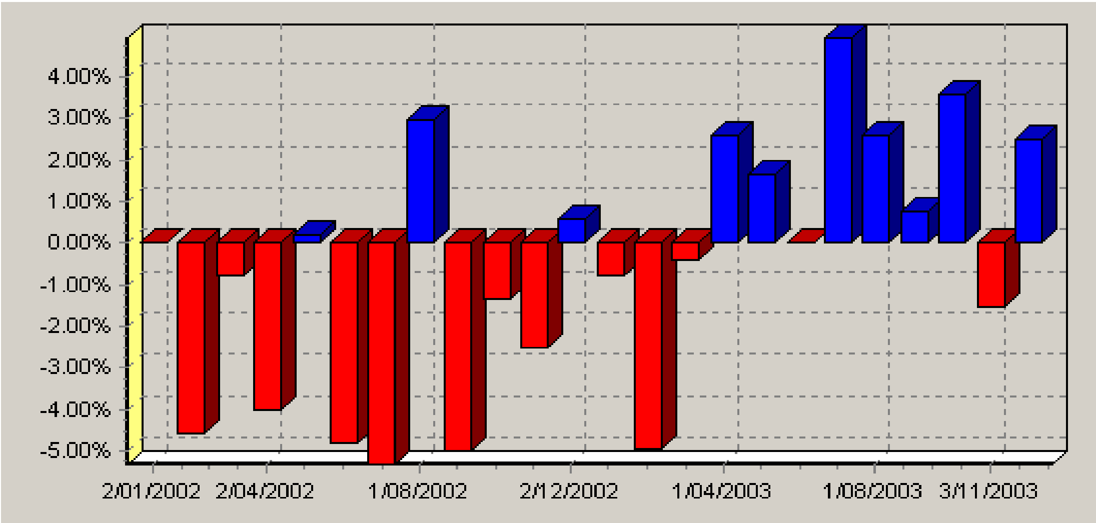

Figure 4-13 TS-TNN(ASX200) Monthly Returns

By comparing the returns to the trading system on a month-by-month basis to the standardized buy-and-hold returns, it is clear that the trading system has some good months and some bad months, and that these align quite closely to the actual performance of the market. In those months when the market is going down/sideways, the trading system has negative returns. In those months when the market is going up, the trading system has positive returns. Overall, as already mentioned, the market was quite listless, and the trading system was poorly equipped to deal with this.

4.5.5.3 Market Efficiency

The S&P/ASX200 is the Investable Benchmark for Australia. As such, it is extremely heavily traded, and therefore very liquid. It is reasonable to conclude that the S&P/ASX200 is much more efficient than the ASX Allshare, and therefore, the neural network was not able to consistently detect and identify short term anomalies.

4.5.5.4 Change in Market Dynamics

It is accepted by academics that the underlying dynamics of markets change over time. It could be the case that the actual dynamics of the S&P/ASX200 underwent a shift during the out-of-sample period.

A brief analysis of the out-of-sample trades shows that of the 847 trades taken by the trading system, 640 were closed out due to hitting the stop-loss. This suggests that the stop-loss is much too tight, although it was chosen correctly from the in-sample trades. This could be indicative of a behavioral shift in the ASX200.

Overall, the out-of-sample ‘Average Profit/Loss %’ is -1.65 %, compared to -1.47 % observed during the in-sample period. For this reason, it is unlikely that there has actually been a behavioral shift in the market, it is more likely that the market was simply ‘more listless’ during the out-of-sample period compared to the in-sample period.

4.5.5.5 Conclusion

In essence, the trading system, TS-TNN(ASX200) has performed poorly, and some possible reasons for this are discussed above. The main information to come out of this trading system is a confirmation of the in-sample result. The ANN was tested using in-sample metrics before it proceeded to trading system development. The results from the in-sample metrics clearly showed that this network would not perform adequately out-of-sample. The out-of-sample results confirmed this.

In essence, the trading system is attempting to detect stocks that will have higher than average returns. It is well known that major market players tend to ‘fade’ against breakouts. That is, when a stock makes a sudden break upward in price, many retail investors will back this breakout, however, institutional investors will bet on a return to normality. As the S&P/ASX200 is so heavily traded by institutional investors, it is likely that the style of trading system which looks for short-term gains will not be particularly successful in this type of market.

Chapter 5 Conclusions and implications

5.1 Introduction

A great deal of the research cited in the literature review focused on training ANNs to predict changes in stock market prices and returns. Although ANNs were generally found to be suitable for this task, virtually no research attempted to site the completed ANNs in the context of a trading system, and actually determine whether the ANNs were economically viable.

This thesis has attempted to do just that. In doing so, it has laid down a structured approach to creating trading systems using neural networks, and has defined the key characteristics necessary to benchmark these trading systems tested on out-of-sample data.

Four neural networks were created during this thesis, focusing on different sets of input variables, and they were trained and tested over different portions of the market.

The purpose of this section of the thesis is to provide a discussion on why the neural networks achieved the results they did, and then to formally draw conclusions about the research question and thesis hypotheses from the results.

5.2 Discussion of Results

Neural networks are valued for their ability to determine non-linear relationships in noisy datasets. According to Refenes et al. (1993), there is a general acceptance in the academic community that many of the relationships concerning security prices (and returns) are most likely non-linear.

Three of the four ANNs trained in this thesis demonstrated this ability very well. These ANNs successfully determined relationships inherent in their underlying datasets. These ANNs were successful in approximating an underlying function that embodied the relationship of the input variables to the underlying structural mechanics of their respective markets. Where this relationship continued to hold, these ANNs performed robustly out of sample.

The neural network trained using technical data from the S&P ASX200 performed poorly, despite using the same set of inputs as the neural network trained using technical data from the Allshare. A number of possible explanations for this were proposed, and these mainly related to the listlessness of the S&P ASX200 during the out-of-sample period.

Another possible explanation is that the neural network did not perform poorly, and that it gave good signals, however, these signals were not focused enough to allow a trading system to profit. It is generally accepted in the trading community that 'a rising tide lifts all the boats'. This implies that when the market is rising, many stocks tend to benefit and participate in this rise. Naturally, the converse is also true. By implication, when the overall market is listless for a prolonged period, any strategy which relies on exploiting abnormal returns has limited its scope for success.

Function profiles are excellent tools to visually inspect the profile of any mathematical function. They allow a deep and visual understanding of a functions characteristics, and provide the ability to visually assess the function profile in varying timeframes simultaneously. Their main limitation becomes apparent when considering the function profile for the poorly performing ANN. This limitation is that the function profile is not 'contextual'. In other words, for a given output strength forecast by the ANN, the average return displayed in the function profile is the average return of all the observations for that output strength. However, due to the non-linear nature of the ANN functions, it is quite feasible that the interpretation of output strength is sensitive to the prevailing conditions inherent in the market when that output strength is forecast.

One guideline consistently presented in technical analysis is that traders increase their chances of success when they trade in the direction of the long-term trend. Technical Analysts use a rule-of-thumb that a stock is in an uptrend when prices move above the 30 week ema (see for example, Bedford (2004)). The exponential moving average (ema) is similar to a simple moving average (sma), except that when using an ema, higher weightings are applied to the more recent data elements. A simple moving average applies the same weighting to every data element.

A neural network which signals when a stock is likely to increase is more likely to be reliable when the stock itself is already in a sustained uptrend. It is less likely that the neural network signal is reliable when the stock is already in a sustained downtrend, or drifting listlessly. In other words, the 30-week ema rule-of-thumb can be used to add context to the ANN signals.

Using this rule as an additional filter, the signals from the poorly performing ANN can be retested. This allows us to determine whether the quality of the signals from the ANN were poor quality, or whether the signals were of good quality, but without context.

Originally, the trading system TS-TNN(ASX200) signaled 847 trades. Filtering these to only act on those signals where the stock was already above its 30 week ema cuts down the number of trades to 630. Taking only these trades changes the overall outcome of the trading system. With this additional filter, the trading system immediately becomes profitable. The analysis of variance reveals that the trades selected by applying the ema filter to the neural network signals are significantly different from the initial set of trades from the neural network alone, specifically $F(1,1475) = 4.922$, $p<0.05$. Ignoring the neural network signal entirely and taking only trades where the price is greater than the

30 week ema triggers 1,406 trades, and that system is also a net loser. Brief summary information is provided in Table 5-1.

Metric	TS-TNN(ASX200)	TS-TNN(ASX200) with EMA filter	Only EMA filter
Number of Trades	847	630	1,406
Annualized Gain	- 7.11 %	1.02 %	- 4.84 %

Table 5-1 TS-TNN(ASX200) with Filter Rule applied

The implication of this result is that the TS-TNN(ASX200) neural network was capable of providing good quality signals, however, the signals did not take into account the prevailing market conditions.

This observation leads to a number of possible directions for future study.

One possibility is to add an additional variable to the neural networks input set which would allow the neural network to assess overall trend direction. From the observation above, this additional variable could be simply set to a zero or a one depending on whether a stock was trading above or below its 30 week ema. This would allow the neural network to incorporate the long-term trend direction into its forecasts automatically.

Another appropriate new research direction is to investigate using neural networks to enhance existing technical or fundamental strategies. Effectively, this would involve testing and isolating strategies which were successful without the use of neural networks, and then enhancing those strategies by the addition of neural networks. From the observation above, a simple technical strategy which only acquired long-term positions when a stock rose above its 30-week ema could be enhanced by coupling such a strategy with the neural network described above.

Further comments regarding future work will be delayed until section 5.5.

5.3 Conclusions regarding Research Question and Hypotheses

The results of the trading systems created using the four neural models are now used to answer the question posed at the start of the thesis:

“Can ANNs be used to develop economically significant stockmarket trading systems?”.

The two hypotheses are also restated, and each is then considered in light of the thesis results.

Hypothesis 1
Trading strategies created using neural networks trained with fundamental data can outperform the buy-and-hold returns from their respective markets.

Hypothesis 2
Trading strategies created using neural networks trained with technical data can outperform the buy-and-hold returns from their respective markets.

5.3.1 Conclusions about Hypothesis 1

Conclusions about hypothesis 1 can be drawn after considering the results of both of the trading systems TS-FNN(Allshare) and TS-FNN(ASX200).

The return from TS-FNN(Allshare) was double the buy-and-hold return for Allshares, and the return from TS-FNN(ASX200) was triple the buy-and-hold return for the S&P/ASX200. In both cases, the ANOVA test showed that the trades selected by the ANNs at the heart of each system were significantly different from the buy-and-hold trades.

We can therefore confidently conclude that the trading systems trained with fundamental data significantly outperformed the buy and hold returns from their respective markets.

5.3.2 Conclusions about Hypothesis 2

Conclusions about hypothesis 2 can be drawn after considering the results of both of the trading systems TS-TNN(Allshare) and TS-TNN(ASX200).

The return from TS-TNN(Allshare) was just under double the buy-and-hold return for Allshares. However, the return from TS-TNN(ASX200) was much worse than the buy-and-hold return for the S&P/ASX200.

For TS-TNN(Allshare), the ANOVA test showed that the trades selected by the ANN at the heart of the system were significantly different from the buy-and-hold trades.

For TS-TNN(ASX200), it was immediately clear that the trading system has failed. However, an important lesson in trading system development was learnt. The in-sample metrics for the TNN(ASX200) neural network indicated failure was likely long before the trading system was ever developed. Therefore, an important early warning was signaled that the ANN was unable to predict abnormal returns from the in-sample data. In hindsight, and from previous discussion is section 5.2, it is clear that these important early warning signals should not be ignored. The early warning is sufficient to indicate that additional input variables may be required to allow the network greater discriminating ability.

Therefore, we can only confidently conclude that the trading system trained with technical data in the Allshares portion of the market significantly outperformed the buy and hold returns.

5.3.3 Conclusions about the research problem

The research problem was:

“Can ANNs be used to develop economically significant stockmarket trading systems?”.

The conclusion is that ANNs can be used to develop economically significant stockmarket trading systems. Three out of the four trading systems developed significantly outperformed their respective buy-and-hold results.

For the trading system which did not succeed, it is important to note that failure was signaled long before the final trading system was developed. The in-sample metrics clearly showed that the ANN was not capable of predicting excess returns. In many ways, this result can be seen as an affirmation of the methodology presented to develop trading systems.

When creating a trading system, it is important to know as early as possible whether using the end product is likely to result in financial success or failure. For each of the trading systems developed, the methodology created in this thesis clearly signaled the expected outcome. This is a particularly satisfying result, and it will provide future developers of neural trading systems with an early warning of whether the inputs being tested for a given market are likely to eventually result in an economically viable trading system. It is also a reassuring confirmation of Vinces (1995) statement that no trading approach with a negative mathematical expectancy can be expected to be successful in the longer term.

5.4 Implications for theory

It is clear that the outcomes documented in this thesis do not support the Efficient Market Hypothesis (EMH). The fact that both trading systems which used a neural network trained with fundamental data convincingly outperformed the buy-and-hold returns is inconsistent with the semi-strong form of the EMH. The fact that one of the trading systems which used a neural network trained with technical data convincingly outperformed the buy-and-hold returns is inconsistent with the weak form of the EMH.

A great many other researchers have already documented a huge amount of evidence which weighs heavily against the EMH. This thesis has clearly added to it.

It was initially Azoff (1994) who pointed out that a neural network with a high degree of predictive accuracy may not translate into a successful trading system. This point has been made many times since, most recently by Thawornwong and Enke (2004). There has been no real support for this idea except from comments made by practitioners, such as Chande (1997). This thesis finds support for this concept. Indeed, of the three successful trading systems developed, two had a lower winning percentage of trades than the buy-and-hold approach, yet they both provided significantly greater benefits.

Azoff (1994) also suggested that the use of direct price values (and raw data input) is preferred to price differences, to prevent destruction of fragile structure inherent only in the original time series. Longo (1996) stated that training ANNs for financial prediction appeared to be much more successful when the ANNs were trained using raw data as opposed to changes in the variables. Following on from their earlier work, this thesis used raw data as opposed to such transformations as variable differences or logs. The success of the ANNs trained in this thesis supports Longo's statement. The most logical reason for this would be that investors use raw data when making investment decisions. As an example, Longo states that an analysis of P/E ratios is the most widely used relative valuation technique, even though dividend discount models are theoretically superior.

It is hoped that the trading system development methodology presented in this thesis might encourage other academics to pursue the area of trading systems research, an area currently regarded by many in academia as something of a 'black art', or, at best, a 'trade secret'. Much of this suspicion is most likely related to the early suggestion that traditional measures of forecasting performance may not be strongly related to profits from trading. It was primarily for this reason that this thesis introduced 'traders metrics', as these are practical measures used by practitioners.

There is a great distance to go before a useful model of stock market pricing behavior can be defined. However, a great deal of depth can be added to the financial debate about the nature of such a model by detecting and documenting persistent anomalies that exist, as these will effectively help describe the model.

5.5 Future Research

This thesis has effectively attempted to document and exploit the prediction of stockmarket return anomalies by using separate ANNs trained with fundamental and technical data.

It was noted in the literature review that there is a growing trend towards using the ensemble approach to analysis amongst soft-computing researchers. According to Pan et al (2005), the probability ensemble of neural networks is one of the most promising directions.

From the work presented in this thesis, particularly in the discussion of results (section 5.2), it is clear that this synergy effect also applies when considering the combination of neural networks with existing well-defined and well understood strategies. This thesis finds that neural networks can be used to develop economically significant stockmarket trading systems. This finding, coupled with the implications of this synergy effect, implies that it may not be optimal to use neural networks as the sole signal generators for buy and sell signals. It may well be more appropriate to allow existing non-neural strategies, or statistical techniques to provide the context rules for initiating trades, and see neural network forecasts as providing a level of additional confidence to those signals.

For example, long-term technical trend following systems frequently suffer from a 'whipsaw' effect. That is, a long-term entry signal is identified, and a position is established. However, in a very short time, the trend breaks down, and the position must be liquidated. The 'whipsaw effect' is the reason that trend following systems will typically have more losing trades than winning trades. In financial terms, this still leads to a viable system, as long as the value of losing trades is quite low, and/or the value of winning trades is high. A neural network which gave quality signals regarding the likely price appreciation over the longer-term would have obvious benefit when combined with a long-term trend following system.

Also, this idea of combining neural networks trained with fundamental data and neural networks trained with technical data, as mentioned by Gately (1996), is an open area for future research. It will be interesting to study the synergy effects of combining the ANNs created during this thesis, and this will be pursued in future work.

Other researchers, for example Reinganum (1988) and Longo (1996), took a different approach to this same goal of detecting anomolies. They both studied groups of top-performing stocks, and then attempted to classify the fundamental and technical variables that were common to this group. They then used this subset of variables to build portfolios, again with excellent results. As was jokingly noted by Reinganum, ‘there is more than one way to skin the investment cat’.

There are also a great many different technical and fundamental variables in common usage. A good deal of scope exists for studying these different variables, and using them in construction of new ANNs. This thesis relied on two published US studies of fundamental variables to select its fundamental variables, and relied on published work and practitioner journals to source its technical variables. It will be interesting, in future research, to expand the range of variables considered.

It became clear during the course of this thesis that fixing the money management parameter at 1% for the entire thesis most likely worked against the final profit outcome of each trading system developed. The money management parameter was fixed at 1% to enable the returns from each trading system to be correctly attributed to the neural networks developed. This was appropriate given the goal of this thesis, and the research question. However, in terms of trading system returns, it can now be seen as a serious potential cap on returns. In general, each of the ANNs signaled a higher output value when higher returns were expected and a lower output value when lower returns were expected. In hindsight, it seems logical to relate the amount of risk capital placed for each trade to the ANN output signal, effectively increasing the amount of capital placed on a trade when the ANN signals a higher likely expected outcome.

The traditional ways of allocating capital, such as fixed amount allocation, or percentage of overall equity have recently given way to more advanced techniques like Monte-Carlo optimization. It will be interesting, in future work, to study the money management outcomes of using a Monte-Carlo approach, and compare these to the outcomes from relating risk-capital to the ANN output signal. There are also likely relationships between the ANN output signals and the stop-loss thresholds implemented for each strategy. Again, this will make for a fascinating area for future research.

Finally, there is a growing movement towards day-trading within the practitioner community. All ANNs in this thesis were developed using daily observations, as this timeframe is consistently used in academic studies. However, by its very nature, a great deal of activity that takes place during the trading day is excluded from end-of-day analysis. The function profiles created for the technical variables showed that the effects of most variables tested were consistent across all the timeframes tested, from 3 to 15 bars (days). It will be interesting to determine whether these effects are also present in data at much smaller intervals, such as within hourly, or minutely bars. Some work in this area has already been undertaken, generally following along the lines of the fractal principle of self-similarity, see for example, Yakuwa et al(2003).

The results documented in this thesis are an important early step on the long journey to understanding stockmarket pricing behavior. Early steps on such a journey should focus on identifying functions which help describe aspects of the underlying behavior. Only by identifying functions which describe this behavior can we ever hope to understand it.