PREDICTION OF STOCK PRICE DIRECTION BY ARTIFICIAL NEURAL NETWORK APPROACH

DOĞAÇ ŞENOL

BOĞAZİÇİ UNIVERSITY

2008

PREDICTION OF STOCK PRICE DIRECTION BY ARTIFICIAL NEURAL NETWORK APPROACH

by

Doğaç Şenol

Submitted to

The Institute for Graduate Studies in Social Sciences

in partial fulfillment of the requirements

for the degree of

Master of Management Information Systems

Boğaziçi University

2008

PREDICTION OF STOCK PRICE DIRECTION BY ARTIFICIAL NEURAL NETWORK APPROACH

The thesis of Doğaç Şenol

has been approved by:

Prof. Dr. Meltem Özturan ______________________________
(Thesis Advisor)

Dr. Ali Tükel ______________________________

Prof. Dr. Fikret Gürgen ______________________________

July 2008

Thesis Abstract

Doğaç Şenol, "Prediction of Stock Price Direction by Artificial Neural Network Approach"

The stock market has always been an attractive area for researchers since no method has been found yet to predict the stock price behavior precisely. It carries a higher risk than any other investment area, due to its high rate of uncertainty and volatility, thus making the stock price behavior difficult to forecast. For years, conventional methods have been developed but they have succeeded partially or have completely failed to deal with the nonlinear and complex behavior of stock prices. Artificial neural networks approach is a relatively new, active and promising field on the prediction of stock price behavior. Artificial neural networks (ANNs) are mathematical models simulating the learning and decision making processes of the human brain. Because of their nature of easy adaptation to noisy data, and solving complex and nonlinear problems, they fit into the area of stock price behavior prediction.

The Istanbul Stock Exchange (ISE) is the only stock market in Turkey, which has an emerging economy. The market situations and economic fluctuations in Turkey create more uncertainty and volatility in the stock market when compared to emerged markets. This study tries to reduce the effect of this uncertainty and volatility by modeling the change in stock price direction of stocks, identifying the theory and steps involved in applying ANN in financial markets and developing a software package to be used for predicting directional daily stock price behavior. It also discusses the appropriate ways to use this process in developing trading systems, further discussing the superiority of ANN over traditional methodologies.

Tez Özeti

Doğaç Şenol, "Yapay Sinir Ağları Yaklaşımı ile Hisse Fiyat Yönü Tahmini"

Borsa, henüz hisse fiyat davranışlarını tam olarak tahmin edebilecek herhangi bir yöntem bulunmadığından dolayı, araştırmacılar için her zaman çekici bir alan olmuştur. Hisse fiyat davranışlarının tahminini zorlaştıran yüksek belirsizlik ve volatilite nedeniyle, diğer tüm yatırım alanlarından çok daha fazla risk taşır. Yıllarca, geleneksel yöntemler geliştirilmiş ama bunlar, hisse fiyatlarının doğrusal olmayan ve karmaşık davranışlarını nedeniyle, kısmen başarılı ya da tamamen başarısız olmuşlardır. Yapay sinir ağları yaklaşımı, hisse fiyat davranışlarının tahmininde göreceli olarak yeni, faal ve umut veren bir alandır. Yapay sinir ağları, insan beyninin öğrenme ve karar verme işlemlerini taklit eden matematiksel modellerdir ve gürültülü veriye, karmaşık ve doğrusal olmayan problemlere kolay uyum sağlamaları nedeniyle, hisse fiyat davranışını tahmin etmeye uygundurlar.

İstanbul Menkul Kıymetler Borsası, gelişmekte olan bir ekonomiye sahip olan Türkiye'deki tek hisse senedi piyasasıdır. Türkiye'deki piyasa durumları ve ekonomik dalgalanmalar, gelişmiş ülkelere göre hisse senedi piyasalarında daha çok belirsizliğe ve volatiliteye neden olmaktadır. Bu çalışmada amaç; bu belirsizlik ve volatiliteyi düşürmek için, hisse fiyatlarındaki değişimi modellemek, yapay sinir ağlarının finansal piyasalara uygulanmasındaki teori ve adımlarını incelemek ve günlük hisse fiyat yönü değişimlerini tahmin eden bir yazılım geliştirmektir ve aynı zamanda da, bu edinimleri hisse ticareti yapan bir sistem geliştirmekte kullanarak, yapay sinir ağlarının geleneksel tekniklere olan üstünlüğünü tartışmaktır.

Acknowledgements

I would like to thank Prof. Dr. Meltem Özturan for her guidance, support and belief and Dr. Ali Tükel for his valuable ideas and experience in this field that enlightened this thesis. I also want to thank my family, my beloved wife Evrim Ece Şenol and my parents Ali and Nilgün Şenol for their continuous support and never ending belief. Also, I would like to thank TradeSoft Business Services, the company that I work for, for their allowance and patience during my study. Finally, I would like to thank Canan Danışman for editing my thesis.

TABLE of CONTENTS

CHAPTER I

INTRODUCTION

Artificial intelligence (AI) has been one of the most attractive research areas in the field of information systems. Computers do the calculations blazingly fast and flawless. Why would they not make the decisions instead of human beings?

Being one of the most important points in business life, financial operations have been the target of various AI models. Especially stock market trend prediction, because of its time series nature, is a popular field of AI applications.

The uncertainty in the stock market has forced the researchers to find a way to estimate the effect of this uncertainty to the flow of stock prices. A field of AI, called Artificial Neural Network (ANN), being a popular way to distinguish unknown and hidden patterns in data, is suitable in predicting stock market trends.

The stock market not only has known inputs and outputs but is also affected by external information causing uncertainty. The ANN model tries to predict the hidden relationship between inputs and outputs by simulating the stock market by taking a small subset of known information to reduce the effect of this uncertainty and outperform the market in profit making. Information not contained in this subset is considered *noise*. Representation of the stock market and the model constructed for it using ANN approach are given in Figure 1 and Figure 2 respectively.

The objective of this study is to reduce the effect of uncertainty and noise in the stock market by using the ANN approach. After feeding stock data and providing predictions for the direction of the change in value of stocks, the usability and feasibility of the model are discussed.

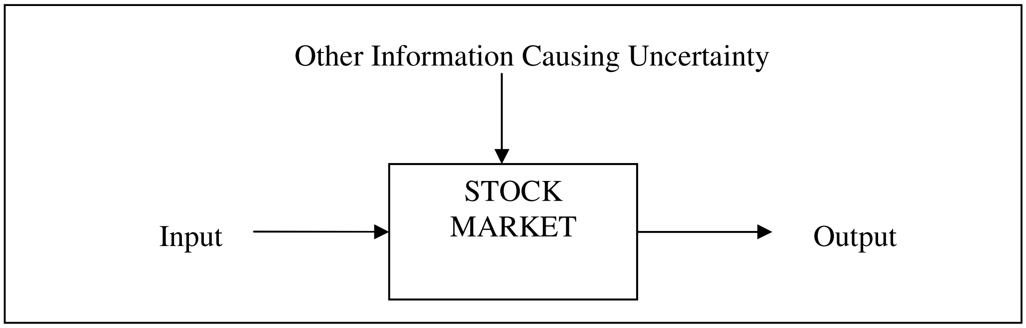

Figure 1. Representation of the stock market

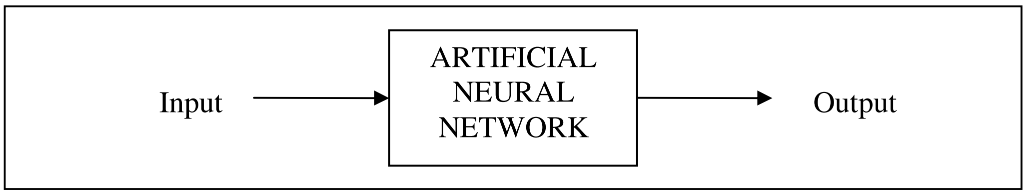

Figure 2. Representation of stock market model using ANN approach

Literature survey on the subject is done through journals, books and the web. Information on stocks, stock markets, data, statistical measures and indicators have been acquired from experts of the subjects working in a private brokerage firm. Rather than focusing on a single stock market index value, individual stocks included in ISE-30 index of the İstanbul Stock Exchange (ISE) that consists of top 30 non-volatile and high-liquidity stocks of reputable companies, are considered.

The study is divided into six chapters.

Chapter I is an overview of the study and also a guide to read through the whole text.

Chapter II is a literature review of previous studies on prediction of stock price and stock price behavior.

Chapter III describes the methodology of the stock price direction prediction system and ANN models that have been developed and implemented based on the literature review. It identifies the theory and steps involved in determining the indicators of stock price direction and in performing ANN in financial markets. The accuracy of the computer learning process and the appropriate ways to use this process in developing trading strategies are also discussed in this chapter.

Chapter IV gives the ANN approach results of the prediction system models developed for stocks in ISE-30. The results are compared with and benchmarked against logistic regression. The outcomes of the software developed for the ANN model is also compared to a commercial ANN software package.

Chapter V finalizes the study with a summary of the outcomes and yields to a conclusion.

CHAPTER II

LITERATURE REVIEW

The Stock Market

Stock

Encyclopedia of Britannica Online (2006) defines the term "Stock" as:

"The subscribed capital of a corporation or limited-liability company, usually divided into shares and represented by transferable certificates. The certificates may detail the contractual relationship between the company and its stockholders, or shareholders, and set forth the division of the risk, income, and control of the business."

A stock owner owns a part of the company and by buying or selling the shares, the rate of ownership might be increased or decreased. If the decision making power decides to distribute a portion of the company's profits rather than reusing them, an owner is paid a profit share, called "dividend".

The first company to issue shares of stock was the Dutch East India Company (of the Netherlands Empire) in 1602.

Stock Pricing

A stock is priced at a static price when issued; then, it may be traded at any rate. For publicly traded companies, the stocks are traded in the stock market, where prices are determined by the supply/demand equilibrium.

As stated by Brown (2002), a company that misses profit estimations can experience severe stock price drops due to the fear of investors.

Investment in the Stock Market

There are various ways for one to invest money. The stock market is probably the most variable, uncertain, unsteady and risky way to go. The general rule applies to the stock market. The more risk the strategy has, the more profit it carries. The stock market is where the company stocks are traded according to the defined rules. In the early stages, wealthy people or businesses have been the ones who trade, but with the emergence of technology individual investors also become able to invest in stock markets. In modern markets, the participants range from individuals to companies and funds.

An investment in the stock market may be short or long term. Short-term investments are generally based on speculations. Long-term investments are generally based on the company's overall performance, expectations and potential to increase its earnings and profit.

Each corporate entity has a ticker code, representing a short form of its name (like EREGL for "Ereğli Demir Çelik Fabrikaları T.A.Ş."). The effective price of a stock is defined by the market makers who quote a buy or sell price of a stock aiming to make profit in the market according to some rules and restrictions. An example of such kinds of restrictions is that, in the İstanbul Stock Exchange (ISE) the price cannot exceed or fall below the opening price more than 10% in a single session.

Investors have always been in the need of stock investment strategies that are going to create profit in a non-distant future. As defined above and denoted by various researchers like Avmarov (2002), the stock market has an important characteristic: *uncertainty*. This uncertainty has to be eliminated (or at least minimized) to predict the near future and make profit.

<u>Investment Theories</u>

While a deep understanding of the investor and capital theory is beyond the purpose of this study, at least some entry level knowledge is required.

According to Goetzmann (1997), the common question in investment is "why investors invest". Probably the simplest answer is the desire to save, preserve money in the future and increase wealth. For whatever reason the investment was done, the key measure is called the *return on investment*. It is defined in a simple way: the difference between the future value where the profit is realized (when a stock which exceeds the purchase price is sold) and the value where the stock has been acquired, $R_t = P_{t+d} - P_t$, where d represents the number of days the stock has been held and P_t represents the price at t.

The predictability of the price of a stock in a near or distant future has been under discussion from the start of trading history. Various researchers have been polarized in two different opinions: The ones who think that the prices are not predictable in any way, and the others who, even though they do not reject this opinion, try to find ways to predict future prices.

Traditionally, the investment community has used one of the two approaches to asset valuation as stated by Malkiel (1999):

i. Firm-Foundation Theory: This theory argues that each corporate stock has a value based on an underlying perception called *intrinsic value*, which can be determined through an analysis of present conditions and future prospects as stated by Sornette (2000). When market prices fall below this value, a buying opportunity arises and when the prices rise above this value, a selling opportunity arises. This theory is very straight-forward, as the effective price (the instant price the stock can be traded at that time) is expected to fluctuate around the intrinsic price. Potential problems arise when determining the intrinsic price, as the investor should estimate dividends, growth moving averages rates and cash flows.

ii. Castles-in-the-Air Theory: This theory states that success in investment can be gained not by estimating the intrinsic value, but by analyzing how the crowd of investors will tend to build their hopes (into castles in the air). The successful investor tries to beat the gun by buying before the crows and selling before the bubble bursts.

These traditional theories have been under discussion for a long time. The former has had a more logical point of view while the latter has had a psychological one. In spite of the polarization, there is one thing common in both theories, information.

Information in the Stock Market

The first question has been “why investors invest”. Now, there comes a second question: “how investors invest”; in other words “how investors make the decision to invest”. Investors invest according to the information they have, or they think that they have. The information becomes an investment decision. The investment decision may also be information to another investor. Different types of information which are used in different kinds of prediction techniques can be classified as *technical*, *fundamental* and *other* information.

Technical information is the short-term or long term performance values of the stock, like *effective price* (when an active trading session is present), *opening and closing prices* (at the beginning and end of the day or trading session) and *total stocks traded* (volume).

Fundamental information is the information related to the intrinsic value of the company, including macroeconomic effects, like *net profit, assets* and *future payables and receivables.*

Other information is any information about the stock that cannot be fully classified into the above mentioned two types. Because of the psychological pressure and great ambition of profit, effects of such types of information should not be underestimated. Examples are *word of mouth*, *insider information* and *exaggerated expectations*.

Efficient Market Hypothesis

Stock price behavior has been a widely questioned and not a mutually agreed area of researchers, where the main question is whether stock price behaviors are predictable or not. Researchers who believe that stock prices do not follow a trend, act in a "random walk" and cannot be predicted, are usually followers of a hypothesis called "The Efficient Market Hypothesis (EMH)".

EMH has been a widely accepted theory which claims that the prices are defined in a random walk procedure, making price behavior completely unpredictable. It also suggests that it is not possible for any kind of prediction algorithm to outperform a buy-and-hold strategy (a long term trading strategy based on the concept that in the long run financial markets give a good rate of return) consistently for a long period of time. This hypothesis has been further discussed, expanded and deepened by various researchers.

Research Supporting EMH

Reilly and Brown (1997) define an efficient market as one in which stock prices adjust rapidly when new information arrives and, therefore, the current prices of stocks have already reflected all information about the stock". For a market to be efficient, three assumptions must hold:

- Large numbers of competing profit-maximizing participants analyze and value stocks independently of each other
- New information regarding stocks comes to the market in a random fashion and the announcements are independent of each other
- Competing investors attempt to adjust stock prices rapidly to reflect the effect of new information

The combined effect of information coming in a random fashion, that is numerous competing investors analyzing that information independently and adjusting stock prices rapidly, means that one would expect price changes to be independent of each other and random. In other words the changes in prices of the stocks are expected to follow a random walk. For the prices to adjust rapidly and for the market to be informationally efficient, some minimum number of competing investors who analyze the information and trade the stocks is required. The higher the number of competing investors, the faster the price adjustments and as a consequence, the stock market is more efficient.

Fama (1970) defines an efficient market as a market in which prices always reflect the recent available information and states that three different levels of efficiency exist:

- *Weak form* assumes that the price of a stock at any point in time fully reflects all the market information of that stock, such as its past prices, returns and trading volumes. Because it assumes that the current market prices already reflect all the past returns and all the past information about the stock, the hypothesis implies that these past returns should have no connection with the future returns, so, future rates of the return should be independent. That implies that investors can gain little from any trading rule that decides when to buy or sell a stock based solely on the information of its past rates of return or past market prices.

- *Semi-strong form* implies that a stock price reflects all the publicly available information. The semi-strong form also encompasses the weak-form hypothesis, since all the information of the market information of the past returns, prices and trading volumes of the stock is public. Additionally to the market information, the public information includes all non-market information such as earnings, dividend announcements, other ratios and news about the overall economy. Investors, who base all their decisions on the information that becomes public, cannot gain above-average returns. The reason is that in such a market, all the prices of stocks have already reflected this information.
- *Strong form* presumes that the stock prices fully reflect all the available information, both from public and private sources. This means that no group of investors has monopolistic access to some information relevant to the stock, so no group of investors should be able to consistently make above-average returns. In this context, the strong form encompasses both the weak form and the semi-strong form, and assumes that prices adjust rapidly to the release of any new public information where all of this information is cost-free and available to everyone at the same time.

The weakest form is redefined by Malkiel (1987) who asserts that *"prices fully reflect the information contained in the historical sequence of prices. Thus investors cannot devise an investment strategy to yield abnormal profits on the basis of past price patterns."* This form of EMH also states that even if it is possible to predict an opportunity that is going to yield a profit, this opportunity becomes useless due to the higher adjustment rate of the market to this opportunity compared to individual investors. This research reminds one of Simon's (1955) who considers that humans are limited in ability to process information, therefore emergence of new technology (such as neural network based methods) may provide profit opportunities.

Dutt and Ghosh (1999) also redefine the weak form of EMH as the non-correlation between spot (current, recent, known) and forward (estimated, future) prices.

Dietrich et al. (2001) redefine the semi-strong market efficiency as the form in which all publicly available information (information open to the general public excluding other information that cannot be widely known, like insider information) is reflected in current price.

According to Dissanaike (2001), given the information set I_t at time t, an efficient market can be expressed as:

In price form: $E\left[P_{it+1} + D_{it+1} \mid I_t\right]/(1 + r_j) = P_{it}$

In return form: $E\left[R_{it+1} \mid I_t\right] = r_i$

In the excess return form: $E\left[R^*_{it+1} \mid I_t\right] = 0, \qquad R^*_{it+1} = R_{it+1} - r_i$

where P_{it}, D_{it}, R_{it} and r_i are the i-th stock's price, dividend, return and the cost of capital at time t respectively.

The efficiency of a market is important and has a number of implications. According to Fama (1970) and Baumol (1965), it is important for markets to be efficient, since if the markets are not efficient, a sub-optimal resource allocation can arise in the economy between firms and industries leading to a lower national income.

Another significance of market efficiency is that with respect to a certain information set, it is impossible to make economic profits by trading on the basis of that set. So, the informationally efficient market implies that technical analysis based solely on past prices of stock and thus usage of machine intelligence in investment decisions has no value. Thus an efficient market contradicts all forms of analysis as it is impossible to beat the market with technical, fundamental or time series analysis because they are considered to have no better performance than random guessing.

Research Contradicting EMH

Contradicting EMH, various methodologies have been put to use for predicting P_{t+d} where P_t is the price at time *t* and *d* is the number of days to a near or distant future.

Opponents of EMH say that human emotions are the strongest influence on the validity of market efficiency. If investors base all their decisions on logic, various forms of analysis can be the sole indicator on trading decisions. Their strongest evidences are stock market crashes. According to Ramon (1997), "stock market crashes contradict EMH because they are not based on randomly occurring information but arise in times of overwhelming investor fear". It is also stated that "many market observers tend to believe in its weaker forms, and thus are often unwilling to share proprietary investment systems."

Because EMH has been strictly defended for a long time and dominated the era of economics, researches done on new ideas, methods and criticisms against this hypothesis have been kept limited.

The Post-Autistic Economics Network (PAE) is a significant group that is against these conventional methods that prevent futuristic growth of modern economics.

Frankfurter (2006) states that *"Efficiency is a myth, invented and promoted by those who want capital, and capital alone rule their society and eventually the whole world. (…) It is only a belief rooted in ideology. When unchecked and unperturbed markets are manipulable by those who control capital and, de facto, the political process."*

Even though not directly related, the famous manifesto of Ph.D. students at Cambridge University (2001) signifies the importance of supporting new eras in economics as follows:

"By restricting research done in economics to that based on one approach only, the development of competing research programs is seriously hampered or prevented altogether. (…) in the current situation an economist who does not do economics in the prescribed way finds it very difficult to get recognition for her research."

These statements support the fact that the issue of market efficiency (or inefficiency) is still under argument and not strictly demonstrated in capital markets, leading to the need of further research.

Predicting the Stock Market

According to Hellstrom (1997), there are four main ways to predict the stock market; *fundamental analysis*, *technical analysis*, *time series forecasting* and *machine learning.*

Fundamental Analysis

In fundamental analysis, the value per share of the company is determined by several acts like measuring efficiency of management, examining its balance sheet (assets, profitability and cash flow), predicting competitive advantages, industrial and macro-economical values. Fundamental analysis is generally preferred when the investor is in the need of investing the money for a medium to long period. Analysis of financial statements is the most important part of the fundamental analysis.

Warren Buffett, known as the world's greatest investor and second richest man in the world as of the end of 2005 (and was number one before Bill Gates) has devoted himself to fundamental analysis and made his money from this approach since the 1960s. Lynch (1999) says that *"Warren Buffett stresses that the critical investment factor is determining the intrinsic value of a business and paying a fair or bargain price. He doesn't care what the general stock market has done recently or will do in the future"*. Warren Buffett does not invest on stock quote prices but on the company itself as an owner and tries to find investment opportunities where it is always going to be (or for a sufficiently long time) feasible and profitable to hold.

Tromsett (1998) states that, straightforward information for fundamental analysis in a practical format is difficult to find as it is easy for an individual investor (not an investment or financial expert) to lose himself among the reports prepared by accountants and analysts. He questions the *random walk* and *efficient market* theories and implies that fundamental analysis is the most common and logical way to make a decision on company performance.

According to Tan (1997), fundamental analysis is the study of the effect of supply and demand on price. He states that the relevant factors come together and define a value called the *intrinsic* value, which determines the actual price of a stock. If the stock has been undervalued by the market, it can be a worthy investment. He also implies that the biggest problem of fundamental analysis is that it is generally relevant only in predicting very long trends.

Technical Analysis

Historical data of stocks such as stock prices and trade volumes are used to find specific points of buy and sell in technical analysis and this analysis is generally used when the return for investment is expected in a short-term. The stock market is a place where human beings exchange their money with stocks or vice-versa. So, the trend in the stock market reflects human behavior. As it goes with the anonymous saying *"history repeats itself"*, each minor or major event that has a reflection on the market performance is going to be similar each time, showing recurrent figures. Technical analysis assumes that the stock market moves in trends which can be captured and used for forecasting. The technical analysts believe that there are recurring patterns in the market behavior and use techniques like Elliot waves, Fibonacci series and Gann lines. These techniques are used to monitor changes in stock price and trade volume. Traders and analysts like to use rules like "when the 14-day moving average crosses below the 37-day moving average and both are directed downward, it is time to sell" even though most of these rules have not been shown to be statistically valid and do not have any rational explanation.

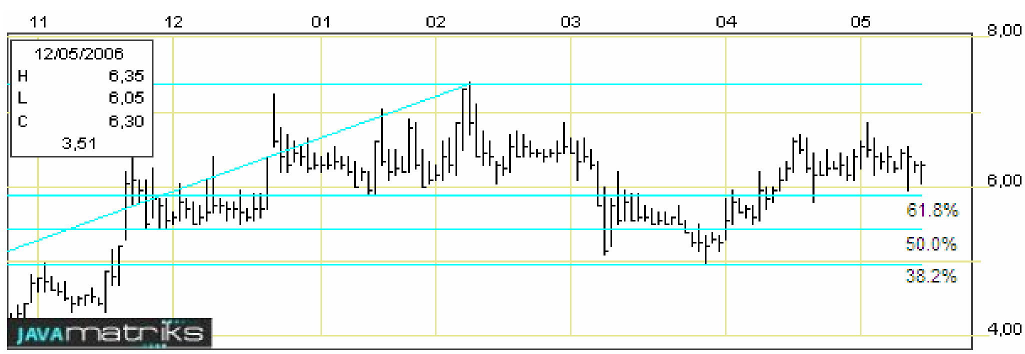

Figure 3. Analysis example of Fibonacci correction levels

It is a known fact that several times speculators in the market try to fit the prices into some rules. Figure 3 shows an example of technical analysis of a stock by using *Fibonacci correction levels*. According to this approach, the rising stocks have to do some correction/retraction to continue rising. Correction levels are at Fibonacci ratios 38.2%, 50%, 61.8%. It can be seen that the correction level at the price of 5 TRY (correction level at 61.8%) has functioned as a support line as expected and the stock hits the line twice to continue its rising trend.

Murphy (1986) defines technical analysis as the study of market action for the purpose of forecasting future price trends. He also summarizes the basis for technical analysis into three premises; *market action discounts everything (all relevant information is reflected by the prices), prices move in trends* and *history repeats itself*. He states that the shifts in supply and demand are sole bases for determining the prices and technical analysis focus on neither the fundamental changes nor the reasons behind the changes in price movements. He also implies that technical analysis tries to identify trends and price changes with respect to the previous trend patterns.

On the other hand, Fama (1970), who claims that technical analysis in efficient markets has no value, has summarized the studies on this subject as follows:

- According to Alexander (1964), price changes follow a random walk over time, but a trend, once initiated, tends to persist.
- Study of Fama and Blume (1966) has indicated negative returns after considering transaction costs.
- Result of the study of Pinches (1970), that examined trading rules based on past market data other than stock prices, also does not find excess returns.

Time Series Forecasting

This type of forecasting tries to create linear models using historical financial data such as stock prices, index values and interest rates. The historic values are spaced equally over time and represent several types of data, like daily closing prices, monthly changes in average. According to Kalekar (2004), this technique assumes that a time series is a combination of a pattern and some error. The goal is to separate the pattern from the error by understanding the *trend* of the pattern and its *seasonality* (the change caused by seasonal factors such as fluctuations in demand).

Several methods are used in time series forecasting; *moving averages*, *linear regression with time* and *exponential smoothing* being the most popular ones. Time series forecasting generally is confused with technical analysis. The main difference is that time series forecasting works on samples (values spaced equally over time) and treats the values as non-chaotic time series.

In their studies where they have applied time series analysis on ISE stock data, Muradoğlu and Ünal (1994) investigate the properties of time series distribution functions on individual ISE stock price data between 1988 and 1991 and state that there is an evidence for ISE stock prices not following a random walk. In a similar study, Aksoy and Sağlam (2001) use a classifier system on ISE index values to build a time series function which is used to calculate expected return and risk at ISE, and present empirical evidence where it is shown that ISE is not efficient in weak form.

Machine (Artificial) Learning

As Zhang (2005) suggests, machine learning technique uses a set of samples to look for patterns to produce the hidden functions and relations. Emergence of new technology has yielded to faster processing speeds, making it easy and feasible for machine learning applications to be applied on several types of pattern recognition problems. The most popular machine learning method is Artificial Neural Networks.

<u>Artificial Neural Networks</u>

There are various proposals on using Artificial Neural Networks (ANNs) on stock price behavior prediction. According to Hassan et al (2006), the research of White (1988) is important, being the first one to use ANNs to forecast and evaluate the stock values with the aim of finding empirical evidence to reject EMH.

White has searched for empirical evidence to reject the EMH using two models. The first one is a linear autoregressive model of order p, presented as AR(p) (which corresponds to a very simple two-layer feed-forward network) and the second one is an ANN model with a three-layer feed-forward network with five inputs and five hidden units.

The study contains two periods from 1972 to 1974 and from 1978 to 1980 (sample period of 500 and a pre-sample period of another 500 days) for both methodologies and the daily return has been computed as:

$$r_t = \frac{p_t - p_{t-1} + d_t}{p_{t-1}}$$

where p_t is the *closing price* on day t and d_t is the *dividend paid* on day t for IBM stock. He uses an input vector of $[r_{t-1}, r_{t-2}, ..., r_{t-p}]$ and computes the output as:

$$\widehat{r}_t = \widehat{w}_0 + \widehat{w}_1 r_{t-1} + ... + \widehat{w}_p r_{t-p}$$

where $\widehat{w}_0 ... \widehat{w}_p$ are the network weights after the learning/training procedure.

Regarding the ANN model, the neural network structure has 3 layers (input-hidden-output) and the topology is 5-5-1 (5 neurons in the input layer, 5 neurons in the hidden layer and 1 neuron in the output layer). Training algorithm is backpropagation and output function is sigmoid.

The squared correlation coefficient R^2:

$$R^2 = 1 - \frac{n^{-1}\sum_{t=1}^{n}(r_t - \hat{r}_t)^2}{n^{-1}\sum_{t=1}^{n}(r_t - \bar{r})^2}$$

where $\bar{r} = n^{-1}\sum_{t=1}^{n} r_t$ and n is the number of observations, is used for the discussion of the results.

For the AR(p) model, R^2 for the first period is found to be 0.0996 which means that since the value is close to zero, the nominator and denominator are very close to each other, and so the model fails in capturing patterns (if any) in p_t. For the second period it is -0.207, which is a significantly negative result indicating that the nominator is much larger than the denominator, and so still meaning that the model is unsuccessful.

For the ANN model, R^2 for the first model is found as 0.0751 and for the second period as -0.0699. These results are similar to the ones found in the AR(p) model.

The outputs of the model, of which the author has been strictly sure of the correctness of its design, show no evidence to reject EMH, but the appropriateness of the values selected and/or the use of simple neural network for the detection of non-linearity in this model are subject to discussion.

Dutta and Shekbar (1988) apply ANNs to bond rating using *bond prices*. They conclude that neural networks outperform classical statistical methods like linear regression models. This study helps to understand that linear models do not provide enough explanation to bond rating models. Also they conclude that even though it provides a better output for the error function (total square error in this case), increasing number of hidden layers in the ANN topology does not improve prediction significantly. More layers improve fit to the training data (also may cause an undesired situation called over-fitting) without changing the power of prediction.

Research of Kimoto et al (1990) seems to be the first research where a system based on neural networks has been tried in a real environment (Tokyo Stock Exchange Prices Indexes) and has succeeded in beating the market. They use five inputs; *vector curve*, *turnover*, *interest rate*, *foreign exchange rate* and *Dow Jones average index*. The approach followed is the modular network approach, in which different networks learn for different data items. Each expert module has its own input domain and preprocessing unit. A final post-processing unit has combined the results to an overall output. The research has been further funded by an investment firm. Each modular network has one hidden layer, uses standard sigmoid as an output function and is trained using back propagation algorithm.

Phua et al (2001) use ANNs with genetic algorithms to do predictions on the Stock Exchange of Singapore. 360 samples (between August 1998 and January 31, 2000), with daily *opening*, *daily high*, *daily low* and *closing* prices with the trading volume of the index, have been examined in this approach. The result is promising: a rate of 81% in predicting market direction.

Fernandez-Rodriguez et al (2000) build an ANN that takes nine inputs as difference of *Nikkei index values* between consecutive days corresponding to the returns in the previous nine days, so the input vector is $[r_{t-1}, r_{t-2}, ..., r_{t-p}]$ where p = 9. It has been stated that values around this number give similar results and this has been the most robust choice. The network has been built with one hidden layer with four neurons. The network is defined as:

$$y_t = G\left(a_0 + \sum_{j=1}^{4} a_j F(b_{0j} + \sum_{i=1}^{9} b_{ji} r_{t-i}) \right)$$

where F is the hidden transfer function, b_{ji} is defined as the weight of a hidden layer's connection from *i*th input unit to the *j*th hidden layer unit, G is the output transfer function, a_j is the weight of connection from the *j*th hidden layer unit. The final output is a number in (-1, 1). A value greater than 0 represents a buy signal while a value less than 0 represents a sell signal.

In this study, the economic significance (profit) gained from the ANN is calculated as: $\widehat{R}_t = \sum_{t=n+1}^{n+p+1} \widehat{y}_t r_t$ where p is the out-of-sample horizon, $\widehat{y}_t$ is the buy (+1) or sell (-1) position and *n* is the number of observations.

The researchers make a comparison of their results with EMH which states that no strategy can outperform buying and holding. The return of a simple buy-and-hold strategy is given by: $R_b = \log\left(\frac{P_{t+p}}{P_t}\right)$ where p indicates the holding period and P_x is the price of equity at time x. The results as given in Table 1 are promising.

Table 1. Findings of the Research of Fernandez-Rodriguez et al. (2000)

	Bear Market	Stable Market	Bull Market
Total Return of the Model	0.48	0.27	0.29
Buy and Hold Return	-0.40	0.0019	0.44

It can be seen that in all types of markets the model has succeeded in creating a profit. In a falling (bear) market, where the stock price has lost 40% of its value, the total profit rate generated by the trades using this model has been 48%. In a stable market, though the prices have not changed significantly, still the model has promised a 27% profit. In a bull market, where the prices are rising, it is known that the buy and hold strategy generally wins. This has also been valid for this case and the buy and hold strategy has brought a profit of 44% where the model has determined a profitability of 29%. The authors also state that the results are in line with their previous research, which is a work of applying non-linear predictors to the Nikkei Index. Therefore, this study is a good indicator that the ANN model has succeeded in capturing the non-linear patterns in the stock trend.

Yao et al (1999) use *moving average (MA)*, *momentum (M)*, *relative strength index (RSI)*, *stochastic (%K)*, and *moving average of stochastic (%D)* to predict the Malaysian stock index using an ANN model for 303 trading days in 1990-1991. Significant profits are obtained compared to interest rates and other investment techniques.

Versace et al (2004) use an ANN model to predict the next day's *closing price* of an exchange traded fund, DIA, tracking the 30 corporations of Dow Jones Industrial Average, achieving a 75.2% success rate.

Egeli et al (2003) indicate that there has been no specific research on ISE stock market values and build an ANN model that uses previous day's index value, exchange rate and simple interest rate as input to forecast ISE price fluctuations. They construct a model with *the previous day's index value*, *the previous day's Turkish Lira/USD exchange rate*, *the previous day's overnight interest rate* and *5 dummy variables each representing the working days of the week*. They try three different numbers of hidden layers (1, 2 and 4) and acquire the lowest error rate and highest accuracy using a single hidden layer. They conclude that ANN models have been superior to the 5-day/10-day moving averages model.

Yumlu et al (2004) have studied 12 years of financial data (*a set of ISE index close value*, *USD value* and *two interest rates*) using a modular ANN model and have concluded that the model outperforms the conventional autoregressive model used for comparison. The authors state that the model introduces a powerful way to predict the volatility of financial time series data, contradicting EMH.

One of the other studies again contradicting EMH, is the study of Tsibouris and Zeidenberg (1995) where they have built several ANN models, compare them to conventional techniques and show that using only past stock prices as inputs have predictive ability. On the other hand, Chiang et al (1996) use ANNs to forecast end-of-year asset value of mutual funds using prices of financial instruments composing the mutual fund portfolio. Traditional linear and non-linear methods have been greatly outperformed by ANNs when the amount of data is limited. Van Eyden (1996) models the performance of 45 selected stocks from the Johannesburg Stock Exchange and trading decisions based on 63 indicators show superior performance and has been able to predict stock market directions seven days ahead.

CHAPTER III

RESEARCH METHODOLOGY

In this chapter, first, the stock price direction prediction system and ANN models of this study, which are designed under the guidance of literature survey, are explained. Secondly, data set that is used in these models are summarized. Furthermore the ANN software that is designed and developed in this study is introduced and finally, evaluation methodologies used for testing the outputs of the study are discussed.

Model Development

In the model development phase, models of both stock price direction prediction system (from now on is going to be called Prediction System – PS) and ANN are described and discussed in general and then the models designed in this study are presented in detail.

Model Descriptions

Since PS and ANN models are the main focuses of this study, definition and description of these models together with the previous studies related to them are discussed clearly and briefly in this section.

Prediction System Models

There have been a variety of indicators that are used for prediction of stock price behavior. Tsibouris and Zeidenberg (1995) state that using only *past stock prices* as inputs have predictive ability. Studies like Egeli et al (2003) and Yumlu et al (2004) use *stock market index*, *foreign exchange* and *interest rate* values. Fernandez-Rodriguez et al (2000) use *stock market index value* only. Kimoto et al (1990) use *vector curve*, *turnover*, *interest rate*, *foreign exchange rate* and *index value*. Yao et al

(1999) use *moving average (MA)*, *momentum (M)*, *relative strength index (RSI)*, *stochastic (%K)*, and *moving average of stochastic (%D).* Kim and Han (2000) suggest 12 different derived values called indicators to use as inputs for stock price behavior prediction by the review of experts of the subject as shown in Figure 4.

Selected features and their formulas (C is the closing price, L the low price, H the high price, MA the moving average of price, M_t: $(H_t + L_t + C_t)/3$, SM_t: $(\sum_{i=1}^{n} M_{t-i+1})/n$, D_t: $(\sum_{i=1}^{n} |M_{t-i+1} - SM_t|)/n$, Up the upward price change, Dw the downward price change)

Name of feature	Formulas
Stochastic %K	$C_t - L_n/H_n - L_n \times 100$
Stochastic %D	$\sum_{i=0}^{n-1} \%K_{t-i}/n$
Stochastic slow %D	$\sum_{i=0}^{n-1} \%D_{t-i}/n$
Momentum	$C_t - C_{t-4}$
ROC (rate of change)	$C_t/C_{t-n} \times 100$
LW %R (Larry William's %R)	$H_n - C_t/H_n - L_n \times 100$
A/D Oscillator (accumulation/ distribution oscillator)	$H_t - C_{t-1}/H_t - L_t$
Disparity 5 days	$C_t/\mathrm{MA}_5 \times 100$
Disparity 10 days	$C_t/\mathrm{MA}_{10} \times 100$
OSCP (price oscillator)	$\mathrm{MA}_5 - \mathrm{MA}_{10}/\mathrm{MA}_5$
CCI (commodity channel index)	$(M_t - SM_t)/(0.015 \times D_t)$
RSI (relative strength index)	$100 - 100/1 + \sum_{i=0}^{n-1} \mathrm{Up}_{t-i}/n / \sum_{i=0}^{n-1} \mathrm{Dw}_{t-i}/n$

Figure 4. Indicators used by Kim and Han (2000)

Vanstone and Finnie (2006) listed mostly cited indicators listed in *Technical Analysis of Stocks & Commodities Magazine*, which is a magazine for financial traders as given in Table 2.

Table 2. Mosty Cited Indicators as Specified by Vanstone and Finnie (2006)

Moving averages
Volatility based variables
Volume based variables
ADX (Average Directional Index)
Stochastics
Momentum (both price and volume)
RSI (Relative Strength Index)
Variety of miscellaneous indicators and oscillators (eg MACD, Intermarket indicators, Money Flow, TRIN (Traders Index), etc)

Versace et al (2004) used various kinds of indicators (which they call operators) as given in Figure 5.

Security	Operators	Equation
DIA	% ROC Close	$[(\text{Close}_t - \text{Close}_{t-1})/\text{Close}_{t-1}]100$
DIA	% Diff. Open-Close	[(Open − Close)/Open]100
DIA	% Diff. High-Low	[(High − Low)/Low]100
DIA	% Diff. Open and Mob. Avg. 10 dd	[(Open − Mavg10 dd)/Mavg10 dd]100
DIA	% Diff. High and Mob. Avg. 10 dd	[(High − Mavg10 dd)/Mavg10 dd]100
DIA	% Diff. Low and Mob. Avg. 10 dd	[(Low − Mavg10 dd)/Mavg10 dd]100
DIA	% Diff. Close and Mob. Avg. 10 dd	[(Close − Mavg10 dd)/Mavg10 dd]100
DIA	% Diff. Volume and Mob. Avg. 10 dd	[(Volume − Mavg10 dd)/Mavg10 dd]100
DIA	Chaikin Volatility	H − L Average = Exponential moving average of(High − Low) $\left(\frac{\text{H − L Average}) - (\text{H − L Average n − periods ago}}{\text{H − L Average n − periods ago}}\right)100$
DIA	MACD	$\text{MACD}_t = \text{EMavg}_1 - \text{EMavg}_2$
DIA	MOMENTUM	$\text{Momentum} = \text{Close}_t - \text{Close}_{t-n}$
DIA	RSI	RS = (Avg. price change on up days − Avg. Price change on down days) RSI = 100 − (100/1 + RS)
Nasdaq	% Diff. Open and Mob. Avg. 10 dd	[(Open − Mavg10 dd)/Mavg10 dd]100
S.Poor 500	% Diff. High and Mob. Avg. 10 dd	[(High − Mavg10 dd)/Mavg10 dd]100
Dow Jones Ind	% Diff. Low and Mob. Avg. 10 dd	[(Low − Mavg10 dd)/Mavg10 dd]100
Dow Jones Trasp	% Diff. Close and Mob. Avg. 10 dd	[(Close − Mavg10 dd)/Mavg10 dd]100
Dow Jones Util		
Dow Jones Comp		
Nikkei Bovespa		
Dax		
Ftse 100		
Dollar/yen		
Dollar/Swiss Frank		
T-Bond 30 years		
T-Bond 10 years		
T-Bond 5 years		
T-Bond 13 weeks		
Eurobond	% Diff. Open and Mob. Avg. 10 dd %	[(Open − Mavg10 dd)/Mavg10 dd]100
Gold	Diff. High and Mob. Avg. 10 dd	[(High − Mavg10 dd)/Mavg10 dd] × 100
	% Diff. Low and Mob. Avg. 10 dd	[(Low − Mavg10 dd)/Mavg10 dd] × 100
	% Diff. Close and Mob. Avg. 10 dd	[(Close − Mavg10 dd)/Mavg10 dd] × 100
	% Diff. Volume and Mob. Avg. 10 dd	[(Volume − Mavg10 dd)/Mavg10 dd] × 100
DJIA and T Bond 30 years	% Diff. between Normalized DJIA and Normalized T Bond	Norm. DJIA-Norm. Tbond

On the left column, the securities for which each operator is calculated are listed. The second column contains the name of the operator and the third column contains the equation by which the operator is calculated. ROC, rate of change; Diff., difference; MAvg, mobile average; EMAvg, exponential mobile average. Normalization: if $\min(x) > 0$, $x_{\text{norm}} = [(x\_\max - x\_\min)\cdot(x - \min(x))]/[(\max(x) - \min(x)]$, where $x\_\max$ and $x\_\min$ are the maximum and minimum value of the required mapping, respectively. If $\min(x) \le 0, \min(x) = -|\max(x)|, x_{\text{norm}} = [(x\_\max - x\_\min)\cdot(x - \min(x))]/[(\max(x) - \min(x)]$.

Figure 5. Indicators used by Versace et al (2004)

The act of selecting true indicators for stock price behavior prediction, in other words designing a PS model is not easy and varies from market to market and even stock to stock. There are no strictly defined periods to use in these indicators so sometimes they are selected arbitrarily or according to rules of thumb.

ANN Models

An artificial neural network is a set of interconnected units called neurons, which "learns" from a set of previous observations without being aware of relations between them, to capture the patterns in the learning set.

Hornik et al (1989) indicate that multilayer feed-forward networks are pattern classifiers and universal approximators. This means that, after a sufficient amount of training (learning), the network is guaranteed to find a relation between any set of dependent and independent variables. According to Frankfurter (2006), "A system which is apparently random could have significant deterministic components embedded in its data". The author also believes that a trading system based entirely upon neural networks is capable of outperforming professional traders. White (1988) says that artificial neural networks are a way to mimic the way in which the human brain processes information.

The basic structure in the human brain is called *neuron*. A neuron contains *dendrites* (small branched extensions of nerve cells that receive signals from other cells) for receiving input and axons that carry output to the other neurons. A neuron carries a potential that is collected as signals from dendrites. When the signal strength exceeds a certain threshold value, the neuron sends out an impulse (a transformation of the original input signal) which is called an *action potential* as stated by Blakemore and Frith (2005).

Like the biological structure defined above, artificial neurons receive inputs and produce an output but they do not accurately model their biological counterparts. A biological neuron stores knowledge in a memory bank, while in an artificial neuron the data or information is distributed through the network and stored in the form of weighted interconnections. The simulation of transformational behavior of biological neurons is done by a nonlinear function. The interconnections between artificial neurons are called *weights*. Artificial neurons reside in *layers*. Figure 6 shows a graphical representation of an artificial neuron, where x_i represents the inputs to the neuron and w_i represents the weights of the neuron. The overall input to the neuron is calculated by $a = \sum_{i=0}^{n} w_i x_i$. To normalize this sum into a standard range, functions called *threshold functions*, (*sigmoid functions* being the most widely preferred one) are used. A sigmoid function is defined as $f(a) = \frac{1}{1+e^{-a}}$. The output of this function is guaranteed to be in (0, 1).

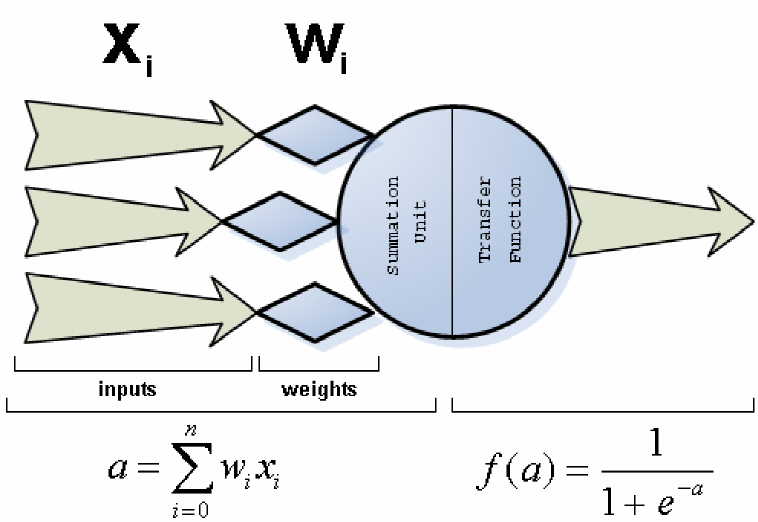

Figure 6. A graphical representation of an artificial neuron

ANNs generally have at least three layers containing the artificial neurons: input, hidden (or middle), and output. Artificial neurons are arranged in these layers. The input layer takes the inputs and passes to the middle layer. Even though it depends on the implementation, generally there occurs no data processing at the input layer. The middle (hidden) layer is where all the complexity resides and the computation is done. Even though there can be many, most of the time one is sufficient. The number of neurons in this layer is one of the most important things to focus on in the design of an ANN. This layer takes the input from the previous (input) layer and after some calculation passes to the next (middle or output) layer. The output layer takes the derived values from the previous (middle) layer and according to the output function an output value is generated.

Figure 7 shows these layers in a graphical representation. Small circles in the figure are the artificial *neurons* discussed above.

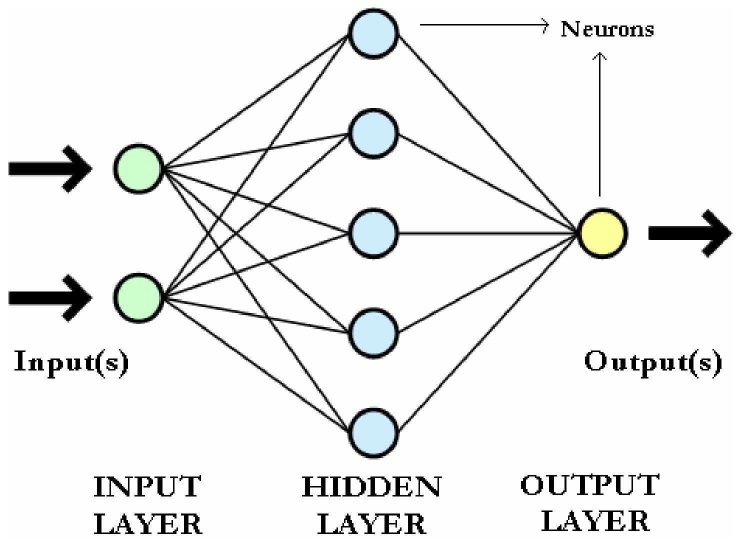

Figure 7. A graphical representation of an artificial neural network

As Zekic (1998) states in her literature research for stock price prediction with ANNs, networks trained with a back-propagation algorithm with one hidden layer perform better. This idea is supported by many other sources. According to Heaton (2005) there is no theoretical reason to use more than two hidden layers and for most of the problems where the function being approximated contains a continuous mapping from one finite space to another, one hidden layer is enough. Chenoweth & Obradovic (1996) prove that networks with large number of inputs do not necessarily perform better and generally lead to worse performance because of the amount of noise, while Steiner and Wittkemper (1996) state that networks with fewer inputs underperform compared to models that have more. As opposed to the myth that neural networks are best used including every data that seems to be handy and let it train for a large amount of time, a moderate amount of time with selectively chosen data using proper termination of training is known to perform better. These studies show that number of inputs should be balanced so that they should not be numerous to degrade the performance of the network while being enough in numbers to contain information that is going to lead the network to find the hidden patterns.

Determining the number of hidden layers, the number of neurons in the hidden layer(s), number of input(s) and output(s) are the most important issues in designing an ANN. There are several approaches or rules of thumb for choosing the number of neurons in hidden layer(s) while designing an ANN topology. The ones retrieved from literature are:

1. $\frac{(i+o)}{2}$, as defined by Man-Chung et al (2000)
2. $2i+1$, as defined by Azoff (1994)
3. $\frac{2}{3}(i+o)$, as defined in the NeuralWare Software Manual
4. $i(i+o)-1$, as defined in the NeuralWare Software Manual
5. $\frac{2}{3}i+1$, as defined by Heaton (2005)
6. $ki-1$, as defined by Freisleben (1992)
7. $\sqrt{i+o}$, as defined by Freisleben (1992)
8. $\ln(i)$, as defined by Gencay (1998)

where i is the number of *inputs*, o is the number of outputs and k is the number of hidden layers.

After an ANN model has been built, two operations are held; *training* and *learning* as classified by Svozil et al (1997).

i. Training Mode: This phase is where neural networks "learn" the training data (set of past observations). Weights of connections between neurons are set to small random values. The training algorithm processes the training set iteratively to adjust weights to reach a specified goal (generally minimization of the mean error) where each iteration is called an *epoch*. For each epoch, weights converge to the optimal values.

Research of Zekic (1998) shows that in most relevant papers on stock market predictions backpropagation algorithm has a higher accuracy rate than other training algorithms, no matter which data model is used.

Backpropagation algorithm was first defined by Werbos (1974) and further developed by Rumelhart (1986). The main idea of the backpropagation algorithm is to minimize the error, which is the difference between the expected value and the output of the model. Weights between neurons are adjusted until the error reaches an acceptable value. Training of a neural network with a backpropagation algorithm is a two-way operation; inputs are propagated forward through the input layer where sigmoid activation function is used to threshold the resultant product (weights and inputs) while errors are propagated backwards by calculating the error each unit is responsible for.

Tagliarini (2001) has defined the pseudocode for the backpropagation training algorithm as:

- Randomly choose the initial weights
- While error is too large and for each training pattern (presented in random order)

    - Apply the inputs to the network
    - Calculate the output for every neuron from the input layer, through the hidden layer(s), to the output layer
    - Calculate the error at the outputs
    - Use the output error to compute error signals for pre-output layers
    - Use the error signals to compute weight adjustments
    - Apply the weight adjustments
- Periodically evaluate the network performance

Derivation of the backpropagation training algorithm, as simplified by Venkataraman (1999) can be seen in Appendix B.

Quaha and Srinivasan (1999) state that error converging is generally slow after 10,000 cycles and it does not converge beyond 15,000 cycles, which is an indicator that the network is over-trained and there is degration in the ability of generalization of the network.

ii. Testing Mode: Also called the prediction mode, testing mode is a state where the network is questioned with input value(s), which yields a result having output(s). The output of the network is calculated by taking the sum of interconnected artificial neurons previously stated. In the testing mode, information flows from the neurons in the first layer to the ones in the hidden layer and finally to the ones in the output layer. Output of each neuron in the output layer represents one of the outputs of the whole network.

Model Designs

Various PS models are going to be designed according to the descriptions specified before and various ANN models are going to be built according to these PS models where the one performing the best is going to selected to be used as the PS model of this study.

Modeling of Prediction Systems

PSs can be modeled using many combinations of numerous financial variables. In this study, the following financial indicators based on previous studies are chosen to be the inputs of the PS models:

1. Moving Average of 14 days (MA14)
2. Moving Average of 37 days (MA37)
3. Stochastic Indicator for 14 days (%K14)
4. Stochastic Moving Average (%D3)
5. Relative Strength Index of 14 days (RSI14)

These indicators have been derived from raw price data where derivation processes are presented in Appendix C. As stated by Technical Analysis from A to Z (2004), an interval of 14 days is the standard for RSI and stochastic indicator, as it is the lowest boundary for *short term* period. Also, an interval of 37 days is the mean number of days for *minor intermediate term* period.

As indicated before, five inputs are considered to be suitable for this study; RSI14, MA14, MA37, %K14 and %D3. From these indicators, only RSI can stand itself, so when the others are used, they should be used in combination with their counterparts: MA14 should be used with MA37 and %K14 should be used with %D3. Considering the sets of these indicators, seven different PS models have been built in this study:

Model 1, M1: $f_O = f_{Model1}(RSI_{14})$
Model 2, M2: $f_O = f_{Model2}(K_{14}, D_3)$
Model 3, M3: $f_O = f_{Model3}(MA_{14}, MA_{37})$
Model 4, M4: $f_O = f_{Model4}(RSI_{14}, K_{14}, D_3, MA_{14}, MA_{37})$
Model 5, M5: $f_O = f_{Model5}(RSI_{14}, K_{14}, D_3)$
Model 6, M6: $f_O = f_{Model6}(RSI_{14}, MA_{14}, MA_{37})$
Model 7, M7: $f_O = f_{Model7}(K_{14}, D_3, MA_{14}, MA_{37})$

Modeling of ANNs

The first step in ANN modeling is determining the topology. ANN topology is defined by the number of input neurons, number of hidden layers, number of neurons in the hidden layer and number of output neurons.

The number of input neurons is the same as the number of inputs, therefore according to the PS model used, the number of neurons is the same as the number of financial indicators used as inputs.

The researches of Dutta and Shekbar (1988), Kimoto et al (1990), Zekic (1998), Fernandez-Rodriguez et al (2000), Egeli et al (2003) and Heaton (2005) conclude that there is no reason to use more than a single hidden layer as it increases complexity without an apparent reason. Therefore in this research, one hidden layer is used in all ANN models.

To determine the possible number of hidden neurons in the hidden layer, using the number of inputs of each PS model and the various rules for determining the number of hidden neurons presented before, Table 3 is constructed where columns contain the rules and the rows PS models. The last column contains the possible number of hidden neurons for the appropriate combination of the rule and number of inputs. Numbers in this column are used to build ANN models for each corresponding PS model, which are going to be tested to find which one of them produces the smallest error rate.

Table 3. Application of Hidden Neuron(s) Rules to PS Models for ANN Modeling (k=1 since one hidden layer is used)

System Model	Number of Inputs	$\frac{(i+o)}{2}$	$2i+1$	$\frac{2}{3}(i+o)$	$i(i+o)-1$	$\frac{2}{3}i+1$	$ki-1$	$\sqrt{i+o}$	$\ln(i)$	Possible # of Hidden Neurons in Single Hidden Layer
M1	1	1	3	1.33	1	1.66	0	1.4	0	1, 2, 3
M2	2	1.5	5	2	5	2.33	1	1.73	0.69	1, 2, 5
M3	2	1.5	5	2	5	2.33	1	1.73	0.69	1, 2, 5
M4	5	3	11	4	29	4.33	4	2.45	1.6	1, 2, 3, 4, 11, 29
M5	3	2	7	2.6	11	3	2	2	1.09	1, 2, 3, 7, 11
M6	3	2	7	2.6	11	3	2	2	1.09	1, 2, 3, 7, 11
M7	4	2.5	9	3.3	19	3.66	3	2.24	1.39	1, 2, 3, 9, 19

Neural networks are known to perform best when the output layer contains a single neuron. The model is going to follow Yao and Poh (1995) who assign boundary values to outputs (0, 1) and train the network as:

$$
\begin{aligned}
&\textit{if}\,(\hat{x}_{t+1} - \hat{x}_t) > 0 \text{ then } \textit{buy} \text{ else} \\
&\textit{if}\,(\hat{x}_{t+1} - \hat{x}_t) = 0 \text{ then } \textit{hold} \text{ else} \\
&\textit{sell} \\
&\textit{where } \hat{x}_t \text{ is the closing price at day t}
\end{aligned}
$$

In other words, the network is going to have one output, which represents if the stock price is going to go up or down in the next day. Table 4 shows the values assigned to change in direction of stock price for the next day, and possible response of the trader.

Table 4. Interpretation of Output Signals of the Network

Output Value	Direction of Stock Price	Possible Response
0.0	Go down	Sell
0.5	Remain the same	Hold
1.0	Go up	Buy

Due to the nature of the threshold function, output of the neural network is going to be between zero and one. This interpretation is similar to the one indicated by Fernandez-Rodriguez et al (2000). The more output is close to zero (one), the more the signal is a reliable sell (buy) recommendation. Due to the nature of neural networks, the output is not expected to be exactly at those boundary values. So, another rule should be set to determine whether output of the network represents a *buy* or a *sell* signal.

Classically, as seen in Figure 8, a critical point is set (the mid-point) and for values greater than the critical point, output is treated as the upper bound and for values smaller than the critical point, it is treated as the lower bound.

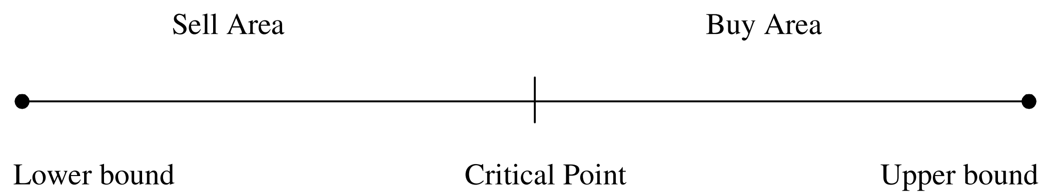

Figure 8. Critical point approach on signal determination

As suggested by the research of Zekic (1998), the training algorithm used is backpropagation and the activation (threshold) function is sigmoid. The common parameters of the neural networks built for the research are given in Table 5.

Table 5. ANN Parameters

Parameter	Value
Learning Rate	0.2
Activation Function	Sigmoid
Initial Weights	Randomized
Momentum Factor	0.5
Stopping Factor	As defined by Quaha and Srinivasan (1999)
Number of Hidden Layers	1

When the network predicts an upward (downward) direction in the stock value and the trading rule defined above generates a BUY (SELL) signal, the stock is bought (sold). As stated by Gencay (1998), the system assumes that a new position is opened at the beginning of every single day and these positions are closed at the end of the day.

Data Set

The target of the model is the İstanbul Stock Exchange (ISE) which is the only corporation in Turkey for the exchange of stocks. It was founded in 1986 as an autonomous organization. As of November 2006, stocks of 309 companies are actively traded. ISE National-100 index is used as a main indicator of the national market. As stated by Yılmaz and Alper (2004), the Turkish economy is regarded as an emerging (developing) market where volatility is high (price levels change rapidly) and liquidity is low (there are not enough willing buyers and sellers to create a deep market, where orders cannot strongly influence prices). This leads to a risky environment where returns of long term investments can be lower than expected because of a high rate of inflation, but also creates an opportunity to make satisfactory profits in a short time, when trading is done properly and timely.

The individual stocks building the most non-volatile stocks index, ISE-30 have been chosen for the data set of this study.

Daily *closing* prices of each stock in ISE-30 for each day were acquired from a private data feeder company. These prices are used to calculate the inputs defined in the prediction system model. Since the number of available trading dates for stocks listed under the name of DENIZ, DOAS and VAKBN of ISE-30 are less than 50% of the average number of days they are not included in this study due to insufficient amount of data.

The period used in the training data sets are between January 5, 1998 (first trading date of 1998) and December 29, 2005 (last trading date of 2005). The data used is given in Appendix A and a statistical summary of this data is given in Table 6.

The period used in the testing data sets are between January 6, 2006 (first trading date of 2006) and August 31, 2007 (last trading date of available data). The average number of days for each stock is 2263.

Table 6. Statistical Summary of the Data Set

Stock Code	Beginning of the Data Set (dd/mm/yyyy)	# of Days	Min (TRY)	Max (TRY)	Stdev. (TRY)
AKBNK	05/01/1998	2525	0.11	9.80	2.44
ARCLK	05/01/1998	2525	0.23	11.97	3.25
DENIZ	01/10/2004	810	2.35	17.10	4.51
DOAS	17/06/2004	735	2.84	10.99	1.82
DOHOL	05/01/1998	2525	0.05	3.53	0.89
DYHOL	06/08/1998	2521	0.13	7.50	1.82
EREGL	05/01/1998	2255	0.10	10.25	2.17
FINBN	05/01/1998	2524	0.02	5.80	1.88
FORTS	05/01/1998	2524	0.02	3.74	0.92
GARAN	05/01/1998	2525	0.07	9.55	1.98
GSDHO	11/11/1999	1949	0.25	3.83	0.70
HURGZ	05/01/1998	2525	0.04	5.66	1.47
ISCTR	05/01/1998	1929	0.17	9.16	2.26
ISGYO	09/12/1999	2525	0.37	2.76	0.65
KCHOL	05/01/1998	2519	0.27	6.40	1.46
MIGRS	05/01/1998	2525	0.58	22.10	4.83
PETKM	05/01/1998	2505	0.71	20.33	2.66
PTOFS	05/01/1998	2527	0.24	7.40	1.57
SAHOL	05/01/1998	2525	0.13	7.95	1.84
SISE	05/01/1998	2515	0.17	6.30	1.74
SKBNK	05/01/1998	2518	0.09	4.62	0.92
TCELL	11/07/2000	2525	0.73	9.65	2.29
THYAO	05/01/1998	1782	1.95	18.25	2.58
TOASO	05/01/1998	2510	0.11	6.85	1.54
TSKB	05/01/1998	2506	0.04	3.13	0.82
TUPRS	05/01/1998	2525	0.79	34.25	8.16
ULKER	05/01/1998	2521	0.06	6.96	2.02
VAKBN	18/11/2005	451	2.45	4.34	0.35
VESTL	05/01/1998	2525	0.25	7.00	1.64
YKBNK	05/01/1998	2517	0.11	4.14	0.90

From time to time, scaling of the price of a stock changes when companies pay dividend to their shareholders or announce stock splits. While causes and necessities of stock splits are beyond the scope of this study, it should be noted that the price data should always be split-adjusted (in other words, normalized according to the current price); otherwise homogeneity of the data is going to be compromised. Stock prices carry bias and may fluctuate in a wide range, so all input data is normalized as:

$$y = \frac{2x - (\max + \min)}{\max - \min}$$

where x is the data before normalizing and y is the data after normalizing.

Design and Development of ANN Software

The software is developed using C# .NET on a Microsoft Windows environment. Reasons for selecting this environment are familiarity, the portability nature of C# .NET and ease of interconnectivity with other platforms. Object oriented methodology, in which abstraction is used to model real world objects, is followed to ensure easy integration to other programs.

Figure 9 shows the class diagram where classes are the definitions of objects representing real world entities of the ANN software package developed. Objects in the software are *Neuron, Connection, Layer* and *NeuralNetwork*, representing the virtual objects that are used.

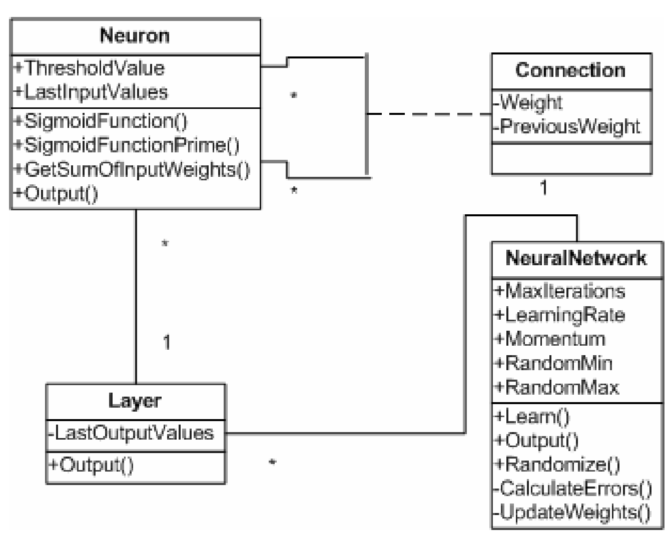

Figure 9. Class diagram of the developed software

Neuron class represents the artificial neuron, the basic structure of a neural network. It carries a *threshold value*, has *input values*, uses a *sigmoid function* and combination of those properties are used to calculate the overall *output*. Neurons are connected to each other with connection lines which carry *weights*, represented by the *Connection* association class. Neurons reside in layers, so the *Layer* class represents the artificial layer where a layer contains one to many neurons and a neuron can only reside in one layer. Therefore, the relationship between a *neuron* and a *layer* is one-to-many. The *neural network,* which is represented by the *NeuralNetwork* class, has parameters like *MaxIterations, LearningRate, Momentum* and consists of several *layers*. Therefore, the relationship between *NeuralNetwork* and *Layer* classes is one-to-many.

The training algorithm is *backpropagation* and the threshold function used in artificial neurons is the *sigmoid function*.

The program is developed as a library package so that it can easily be integrated to other software. Figure 10 shows a screenshot of the user interface that has been developed only to test the software.

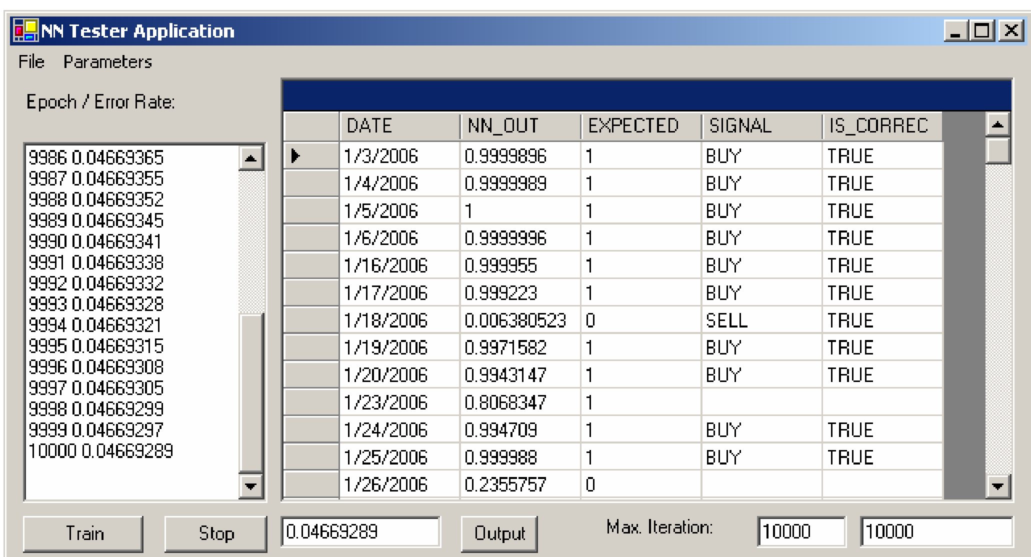

Figure 10. User interface, designed for testing the developed software

The implementation of the algorithm is tested by simulating a simple logical OR and AND switches. As they are very simple and well-defined operations, the network topology for this test is chosen as simple as it can be (two input neurons, two hidden layers and an output neuron, 2-2-1) as seen on Figure 11.

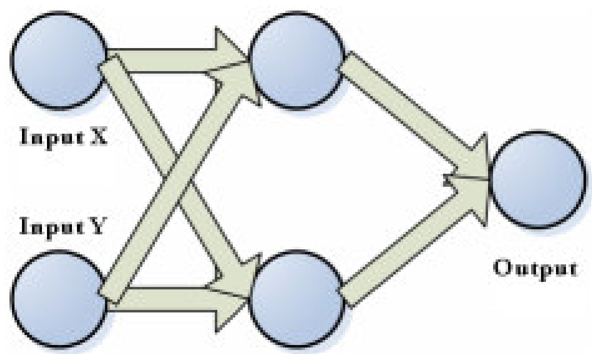

Figure 11. Testing network for the simulation of logical OR and AND switches

Table 7 and Table 8 show the output of the logical AND and OR test, with respect to expected output. In all cases, the network gives perfect results, with a correlation coefficient of r = 0.999. Output, error rate and correlation of the network signify that the network has passed this test successfully.

Table 7. Results of the Logical OR Switch Test

Input X	Input Y	Expected	NN Output	Error
0	0	0	0.0165719651	0.0165719651
0	1	1	0.991398036	0.008601964
1	0	1	0.9903082	0.0096918
1	1	1	0.9953286	0.0046714

Table 8. Results of the Logical AND Switch Test

Input X	Input Y	Expected	NN Output	Error
0	0	0	0.000433221	0.000433221
0	1	0	0.010115318	0.010115318
1	0	0	0.008959988	0.008959988
1	1	1	0.9851726	0.0148273706

Evaluation of Outputs of the Models

After the results have been achieved, they are compared with the actual outcomes (if any) using the following parameters:

i. Correlation: Correlation indicates the strength of a linear relationship between two sets of variables. It is used to understand the correlation between the expected values and the outputs of the network and is suitable for testing of a network which works on time series data sets. The correlation coefficient is calculated as:

$$r = \frac{\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \overline{x})^2 \sum_{i=1}^{n}(y_i - \overline{y})^2}}$$

where

x_i and y_i are ith elements of two different sets

$\overline{x}$ and $\overline{y}$ are sample means of x_i and y_i respectively

The more r gets closer to one, the more the two data sets are correlated perfectly. As the aim of all of the prediction system models proposed in this study is to predict the direction of the stock price rather than to do time series forecasting, the correlation between the outputs do not directly reflect the overall performance of the network. Therefore, it may be misleading on nonlinear models and complex classification problems where there is no linear correlation to look for. Even when no linear correlation has been targeted, correlation coefficient should always be stated as it is a significant statistical variable.

ii. Correctness: It is generally defined as a percentage value and is suitable for classification and boundarized prediction problems. As the correctness value approaches one, the model becomes more accurate in predicting target values, therefore a correctness value of one means perfect prediction. The correctness value is calculated as:

$$Correctness = \frac{\text{Total number of correctly predicted data}}{\text{Total number of testing data}}$$

ANN outputs can also be compared with the results of statistical methods, generally regressive models. Autoregressive models are used by White (1988), Weigend et al (1990), Bernd and Klaus (1996). Linear regression models are used by Dutta and Shekbar (1988) and Chiang et al (1996). Models which are used in these studies are targeted on forecasting a future stock or index value. This study focuses on predicting stock price direction which is represented by a binary number, therefore a regressive model with a binary output is appropriate for comparison of the outcomes.

Logistic regression is a statistical method used when the dependent variable is desired to be interpreted as binary. Logistic regression is very different from linear least-squares regression, in terms of underlying mathematics and computational details. Unlike a linear least-squares regression equation which can be solved explicitly, logistic regression equations are solved iteratively. A trial equation is fitted and tweaked over and over in order to improve the fit. Iterations stop when the improvement from one step to the next is suitably small.

The response variable that characterizes logistic regression is what makes it special. With linear least squares regression, the response variable is a quantitative variable. With logistic regression, the response variable is an indicator of a characteristic, that is, a binary variable. Logistic regression is used to determine whether other measurements are related to the presence of some characteristic, for example, whether people at a certain age are more addicted to smoking. As it is a suitable method for predicting the classifications, it is favored by medical applications.

Logistic regression is an efficient way to measure the accuracy and performance of ANN models if the output is going to be classified as binary variables.

Schumacher et al (1996) compare ANN and logistic regression methodologies and analyze the similarities between the two methods. They conclude that there are significant similarities with respect to their structures and state that they see logistic perceptron as the most accurate ANN equivalent where regression coefficients correspond to weights in ANN.

Dreiseitl and Ohno-Machado (2002) state that logistic regression and artificial neural networks have different origins but they share many similarities and common roots in statistical pattern recognition.

Bell et al (1990) compare logistic regression with ANN to predict bank failures and conclude that ANN outperforms logistic models.

Hwang et al (1994) compare ANN with logistic regression to predict the insolvency of life insurers and conclude that ANN outperforms the conventional techniques including the logistic model.

Luther (1998) compares the performances of ANN and logistic regression to predict the bankruptcy outcomes and states that the ANN approach is more robust than logistic regression in predicting the outcome.

Logistic regression has not been widely used in stock market prediction as the expected outcomes are generally numerical rather than binary in those analyses and predictions. But it has been used in many areas of research where binary variables are modeled and therefore is an appropriate model for comparing ANN outcomes of this study.

In this study, the outcomes of the ANN approach are compared with the outputs of logistic regression method. For that purpose, financial indicators are used as independent variables and buy and sell signals are used as the dependent variable in this methodology. SPSS software package is used for running up logistic regression and the significance of input variables are used to choose the prediction system model whereas correctness and correlation factors of the models are used for the comparison of ANN and logistic regression outputs.

For that purpose, similar to the neural network model, *MA14*, *MA37*, *RSI14*, *K14*, *D3* as independent variables and *EXPECTED* as the dependent variable are used in logistic regression methodology.

The logistic regression is executed by the SPSS command:

logistic regression EXPECTED with MA14 MA37 RSI14 K14 D3.

This command computes the statistical results and displays them in a tabular format as given in Figure 12.

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	MA14	,436	,348	1,568	1	,211	1,547
	MA37	-,172	,133	1,677	1	,195	,842
	RSI14	,052	,008	37,077	1	,000	1,053
	K14	,276	,013	444,259	1	,000	1,318
	D3	-,281	,014	415,574	1	,000	,755
	Constant	-1,650	,272	36,793	1	,000	,192

a. Variable(s) entered on step 1: MA14, MA37, RSI14, K14, D3.

Figure 12. SPSS output of logistic regression analysis

Selection of the independent variables that explain the dependent variable significantly are determined by the logistic regression methodology. Variables with p-values (represented by the *Sig.* column) higher than 0.05 are considered to be out of 95% confidence interval and therefore statistically insignificant. So, the model is rerun only with significant variables using the following SPSS command:

logistic regression EXPECTED with RSI14 K14 D3.

Output from the rerun model shows that only statistically significant variables are used in the operation as seen on Figure 13.

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	RSI14	,055	,008	49,909	1	,000	1,057
	K14	,274	,013	450,287	1	,000	1,315
	D3	-,280	,014	415,865	1	,000	,756
	Constant	-1,842	,232	63,206	1	,000	,158

a. Variable(s) entered on step 1: RSI14, K14, D3.

Figure 13. SPSS output of logistic regression analysis for the significant variables

In the next step, the equation is written as follows:

$$y_i = \beta_0 + \beta_1 x_{1,i} + \ldots + \beta_k x_{k,i}$$

where

i is the number of observation

β_k is the coefficient of kth variable

$x_{k,i}$ is the value of kth variable in the ith observation

For the sample table above, the equation is written as:

$$y_i = -1.842 + (RSI14_i \times 0.55) + (K14_i \times 0.274) + (D3_i \times -0.28)$$

For each y_i, the p_i value is calculated as:

$$p_i = \frac{e^{y_i}}{1 + e^{y_i}}$$

to be used as the probability of a buy signal. Therefore, if p = 0.50, it is interpreted as a HOLD signal. If p > 0.50 then it is interpreted as a BUY signal, otherwise a SELL signal.

On the other hand, other commercial packages are also a reference for comparison of the results of the software developed. Several commercial neural network packages are widely used in ANN studies, one of which is NeuroSolutions 5, developed by NeuroDimension, Inc. and the results of this study is benchmarked with this software. Figure 14 shows a screenshot of the proposed model built using NeuroSolutions.

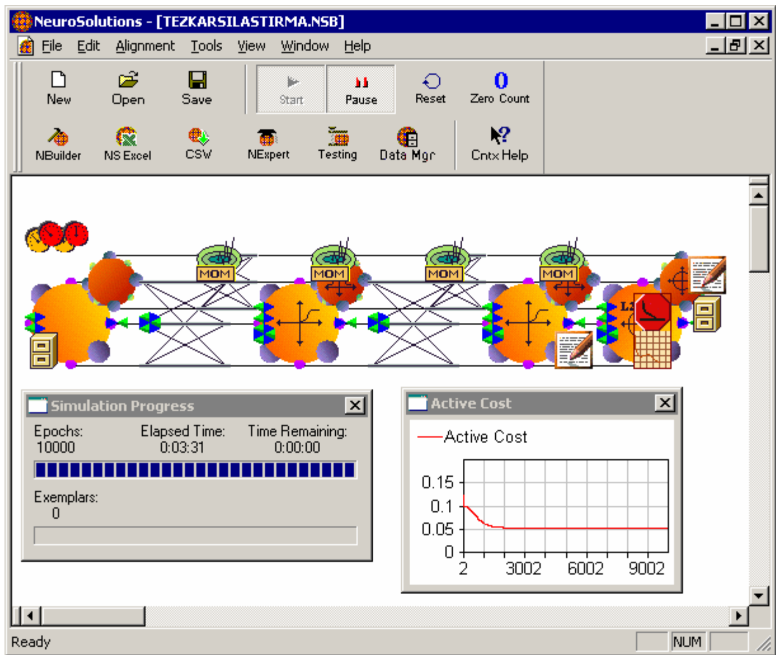

Figure 14. Proposed model built using NeuroSolutions

CHAPTER IV

RESULTS AND FINDINGS

This chapter gives the overall findings and the results of the application of ANN models to each of the proposed PS models described in the research methodology.

Table 9 is constructed to give an overall view of the average mean squared errors of ISE-30 stocks where each ANN model is applied to each PS model after running the software for each stock. Values presented in the *Possible Number of Hidden Neurons* column in Table 3 are used as columns of Table 9 while PS models are given as rows. If the rules given before do not yield a possible combination of number of inputs of the PS model with the corresponding number of hidden neurons, the corresponding cells are labeled as N/A (not applicable).

As seen on Table 9, PS Model 5 with a network topology of one hidden layer, three inputs (R14, K14 and D3), eleven hidden neurons in the hidden layer and one output (ANN model 3-11-1) yields to the lowest error rates and performs better than the others in all cases by the means of average. Therefore, it is chosen as the preferred model with the specified system topology for this research.

Table 9. Average Mean Squared Errors for Training of Each PS Model with Each ANN Model for ISE-30 Stocks

PS Model	Number of Hidden Neurons									
	1	2	3	4	5	7	9	11	19	29
M1	0.0995	0.0990	0.0990	N/A	N/A	N/A	N/A	N/A	N/A	N/A
M2	0.0549	0.0539	N/A	N/A	0.0510	N/A	N/A	N/A	N/A	N/A
M3	0.1018	0.1018	N/A	N/A	0.1018	N/A	N/A	N/A	N/A	N/A
M4	0.0551	0.0538	0.0513	0.0507	N/A	N/A	N/A	0.0464	N/A	0.573
M5	0.0547	0.0535	0.0523	N/A	N/A	0.0482	N/A	0.0364	N/A	N/A
M6	0.0991	0.0986	0.0986	N/A	N/A	0.0982	N/A	0.0928	N/A	N/A
M7	0.0553	0.0543	0.0534	N/A	N/A	N/A	0.0500	N/A	0.584	N/A

The mean square errors of each PS model for each stock are given in Table 10. It can be seen that for every stock, PS Model 5 performs the best and yields the lowest mean square error.

Table 10. Lowest Mean Square Errors of Each Prediction System Model and Stock

	M1	M2	M3	M4	M5	M6	M7	Lowest	Best System Model
AKBNK	0.0991	0.0599	0.0971	0.0456	0.0365	0.0854	0.0475	0.0365	M5
ARCLK	0.0956	0.0536	0.1158	0.0507	0.0394	0.0967	0.0576	0.0394	M5
DOHOL	0.1273	0.0525	0.0922	0.0432	0.0373	0.1173	0.0536	0.0373	M5
DYHOL	0.0976	0.0569	0.1089	0.0461	0.0400	0.0993	0.0431	0.0400	M5
EREGL	0.0992	0.0539	0.0913	0.0400	0.0373	0.0886	0.0491	0.0373	M5
FINBN	0.0985	0.0598	0.1284	0.0500	0.0282	0.1024	0.0511	0.0282	M5
FORTS	0.0965	0.0535	0.1101	0.0466	0.0268	0.1145	0.0517	0.0268	M5
GARAN	0.0969	0.0577	0.0910	0.0468	0.0399	0.0960	0.0519	0.0399	M5
GSDHO	0.0983	0.0541	0.1088	0.0444	0.0349	0.1100	0.0491	0.0349	M5
HURGZ	0.1599	0.0522	0.1181	0.0501	0.0393	0.1029	0.0593	0.0393	M5
ISCTR	0.0984	0.0491	0.0990	0.0391	0.0386	0.0872	0.0491	0.0386	M5
ISGYO	0.0960	0.0525	0.0923	0.0422	0.0380	0.1026	0.0493	0.0380	M5
KCHOL	0.0953	0.0584	0.0932	0.0455	0.0374	0.0992	0.0491	0.0374	M5
MIGRS	0.1140	0.0514	0.1294	0.0462	0.0393	0.1002	0.0573	0.0393	M5
PETKM	0.0963	0.0515	0.1123	0.0445	0.0429	0.0993	0.0577	0.0429	M5
PTOFS	0.0979	0.0544	0.1004	0.0524	0.0407	0.0914	0.0510	0.0407	M5
SAHOL	0.0957	0.0430	0.0974	0.0401	0.0368	0.0848	0.0645	0.0368	M5
SISE	0.0908	0.0517	0.0986	0.0551	0.0402	0.0996	0.0535	0.0402	M5
SKBNK	0.0991	0.0511	0.1080	0.0419	0.0372	0.0872	0.0493	0.0372	M5
TCELL	0.0990	0.0501	0.0998	0.0539	0.0417	0.0965	0.0473	0.0417	M5
THYAO	0.0982	0.0437	0.0915	0.0496	0.0427	0.0879	0.0469	0.0427	M5
TOASO	0.0971	0.0515	0.0910	0.0510	0.0401	0.0987	0.0523	0.0401	M5
TSKB	0.1088	0.0462	0.1126	0.0429	0.0229	0.0901	0.0494	0.0229	M5
TUPRS	0.0985	0.0537	0.1165	0.0483	0.0429	0.0968	0.0461	0.0429	M5
ULKER	0.0997	0.0559	0.1088	0.0381	0.0351	0.0842	0.0502	0.0351	M5
VESTL	0.0999	0.0497	0.1004	0.0410	0.0390	0.0848	0.0510	0.0390	M5
YKBNK	0.0954	0.0504	0.1021	0.0511	0.0400	0.0946	0.0457	0.0400	M5

The software developed is tested against NeuroSolutions 5, a commercial ANN software package. Comparison of correlations between ANN outputs and actual values for the PS model that outperformed others, Model 5, are given in Table 11 for each stock, processed in both software packages.

Table 11. Correlations of Outputs of the Software Developed and NeuroSolutions with Respect to Actual Outcomes

Stock	Custom Software	NeuroSolutions
AKBNK	0.795942372	0.782258381
ARCLK	0.752443223	0.747107897
DOHOL	0.784575611	0.763939481
DYHOL	0.75943713	0.773711962
EREGL	0.75629375	0.72417282
FINBN	0.651118596	0.616810748
FORTS	0.772380506	0.709701169
GARAN	0.802322539	0.770800673
GSDHO	0.721552783	0.707884941
HURGZ	0.792904571	0.769071176
ISCTR	0.752953771	0.771269508
ISGYO	0.752108926	0.758957729
KCHOL	0.779509645	0.76306179
MIGRS	0.758797487	0.738209417
PETKM	0.741909199	0.733373276
PTOFS	0.744319288	0.721479948
SAHOL	0.808670601	0.785064681
SISE	0.786549946	0.774592048
SKBNK	0.703136272	0.673685488
TCELL	0.75823256	0.778350839
THYAO	0.762541713	0.753298816
TOASO	0.723767644	0.728473442
TSKB	0.733671056	0.659202109
TUPRS	0.754812096	0.748625193
ULKER	0.754490611	0.735967861
VESTL	0.768570727	0.769658901
YKBNK	0.779596555	0.750856765

As shown in Table 12, a t-test applied to these correlations shows that one-tailed probability is much higher than 0.05. Therefore in the 95% confidence interval, there is statistically no significant difference between these sets, proving the reliability of the software developed.

Table 12. t-Test for Correlations Shown in Table 11

t-Test: Two-Sample Assuming Unequal Variances		
	R NN	*R NS*
Mean	0.734484	0.733614
Variance	0.007236	0.002467
t Stat	0.04836	
P(T<=t) one-tail	0.480817	
t Critical one-tail	1.677927	
P(T<=t) two-tail	0.961634	
t Critical two-tail	2.01174	

The results of the ANN approach are also compared to the outcomes of the logistic regression method which has been described in the previous chapter. As seen on Table 13, the significant PS model suggested by logistic regression comes out to be the same as the best performing PS model in the ANN approach for each stock (Model 5). Comparison of success rates and correlations of outputs with actual values of two methods, ANN and logistic regression, for each stock are also given in Table 13.

Table 13. Significant Logistic Regression Models, Success Rates and Correlations for Each Stock

Stock	Significant Variables	Corresponding Model	Success Rates		Correlations	
			ANN	LR	ANN	LR
AKBNK	RSI14, K14, D3	Model 5	78.47%	74.16%	0.795942	0.675952
ARCLK	RSI14, K14, D3	Model 5	77.03%	74.88%	0.752443	0.672036
DOHOL	RSI14, K14, D3	Model 5	76.79%	73.68%	0.784576	0.701115
DYHOL	RSI14, K14, D3	Model 5	76.79%	77.99%	0.759437	0.69535
EREGL	RSI14, K14, D3	Model 5	78.47%	71.05%	0.756294	0.616921
FINBN	RSI14, K14, D3	Model 5	78.71%	45.56%	0.651119	0.356088
FORTS	RSI14, K14, D3	Model 5	78.47%	75.60%	0.772381	0.6874
GARAN	RSI14, K14, D3	Model 5	78.71%	75.36%	0.802323	0.667539
GSDHO	RSI14, K14, D3	Model 5	78.71%	72.73%	0.721553	0.626456
HURGZ	RSI14, K14, D3	Model 5	77.75%	76.56%	0.792905	0.680636
ISCTR	RSI14, K14, D3	Model 5	78.23%	76.32%	0.752954	0.683173
ISGYO	RSI14, K14, D3	Model 5	79.19%	74.88%	0.752109	0.675791
KCHOL	RSI14, K14, D3	Model 5	79.19%	72.25%	0.77951	0.639041
MIGRS	RSI14, K14, D3	Model 5	78.95%	73.21%	0.758797	0.624054
PETKM	RSI14, K14, D3	Model 5	78.95%	69.14%	0.741909	0.637336
PTOFS	RSI14, K14, D3	Model 5	79.43%	68.90%	0.744319	0.595209
SAHOL	RSI14, K14, D3	Model 5	78.71%	73.92%	0.808671	0.673203
SISE	RSI14, K14, D3	Model 5	79.19%	75.60%	0.78655	0.679835
SKBNK	RSI14, K14, D3	Model 5	77.75%	70.81%	0.703136	0.59477
TCELL	RSI14, K14, D3	Model 5	77.99%	75.36%	0.758233	0.695597
THYAO	RSI14, K14, D3	Model 5	78.71%	70.33%	0.762542	0.635524
TOASO	RSI14, K14, D3	Model 5	79.19%	73.68%	0.723768	0.652795
TSKB	RSI14, K14, D3	Model 5	78.95%	74.16%	0.733671	0.660993
TUPRS	RSI14, K14, D3	Model 5	78.23%	67.39%	0.754812	0.624211
ULKER	RSI14, K14, D3	Model 5	78.23%	68.66%	0.754491	0.600728
VESTL	RSI14, K14, D3	Model 5	78.47%	72.97%	0.768571	0.69734
YKBNK	RSI14, K14, D3	Model 5	78.47%	75.36%	0.779597	0.667241

As shown in Table 14, a t-test applied to success rates shows that two-tailed probability is much lower than 0.05. Therefore, statistically speaking, in the 95%

confidence interval, the neural network method has scored significantly more than the logistic regression method in terms of successful signals.

Table 14. t-Test for Success Rates Shown in Table 13

t-Test: Two-Sample Assuming Unequal Variances		
	Neural Network	*Logistic Regression*
Mean	0.784334574	0.722412838
Variance	5.00586E-05	0.003561086
t Stat	5.354298634	
P(T<=t) one-tail	5.88599E-06	
t Critical one-tail	1.703288423	
P(T<=t) two-tail	1.1772E-05	
t Critical two-tail	2.051830493	

As shown in Table 15, a t-test applied to the correlations shows that two-tailed probability is much lower than 0.05. Therefore in the 95% confidence interval, the neural network method has scored more than the logistic regression method in terms of correlation with actual values.

Table 15. t-Test for Correlations Shown in Table 13

t-Test: Two-Sample Assuming Unequal Variances		
	ANN	*LR*
Mean	0.734484	0.64399
Variance	0.007236	0.00402
t Stat	4.671858	
P(T<=t) one-tail	1.01E-05	
t Critical one-tail	1.673565	
P(T<=t) two-tail	2.02E-05	
t Critical two-tail	2.004879	

CHAPTER V

SUMMARY AND CONCLUSIONS

The main objective of this study has been to show that, with a well chosen ANN topology, PS model and buy/sell decision parameter; the ANN approach can outperform conventional techniques like regressive methods in predicting the changes in daily stock price directions. For this purpose, individual stocks in ISE-30 index are used in this study as the data set.

After the literature review, it is seen that there was enough academic backing for testing the predictability of stock price behavior. Raw price data carries a high amount of bias; therefore several indicators derived from prices are used as inputs to analyze the daily direction changes, rather than predicting the next day's closing price.

Software used in this study is developed by the author himself in a software library format to be used easily with external systems. The outcomes of the software developed are compared with the outputs of a commercial ANN software package called *NeuroSolutions 5*. Correlations of outputs and actual values are statistically compared and found to have no significant difference, implying the reliability of the software.

The optimal PS model has been chosen from several models constructed using a set of rules retrieved from literature through inputting various combinations of financial indicators to various ANN models. This PS model has been used for the comparison of ANN and logistic regression models.

Statistically, comparison of the outcomes of the ANN model with the logistic regression model shows that the ANN model outperforms logistic regression significantly. The results presented in this study clearly establish that the ANN model performs better than the logistic regression model at capturing the underlying decision structure of this non-linear process and accurately predicting the daily change in stock price direction. The implications of this study suggest that the decision making process in directional stock price behavior is more accurately modeled through a neural networks approach than logistic regression models.

In conclusion, the system outperforms traditional methods in the specified interval, contradictory to the Efficient Market Hypothesis. The results provide empirical evidence that ISE-30 is not weak form efficient.

As future work, economic significance of this study can be evaluated using a day to day portfolio management simulation system, taking into consideration the transaction costs resulting from trades (buy and sells). More PS and ANN models can be built using different network parameters and combinations of indicators, or simply just the raw price data. Training and testing date periods can also be changed and several different models can be built for these date periods. Also, the results showing that the best PS models for both the ANN and the LR models are the same can be more thoroughly examined to check if it is a matter of coincidence or there is a statistical significance. After that, the best PS model can be found with a large number of financial variables and indicators to be filtered by significance using logistic regression, rather than the time consuming trial-and-error method used in this study.

APPENDIX A: STOCK PRICES AND ANN RESPONSES

Due to large amount of data, stock prices and responses of ANN models are given in the CD attached.

APPENDIX B: DERIVATION OF BACKPROPAGATION ALGORITHM

Legend

$\vec{x}_j$ = input vector for neuron j (x_{ji} = ith input to the jth unit)
$\vec{w}_j$ = weight vector for unit j (w_{ji} = weight on x_{ji})
$z_j = \vec{w}_j . \vec{x}_j$ (weighted sum of inputs for unit j)
$o_j = \sigma(z_j)$, output of unit j
t_j = target for unit j
Downstream(i) = set of units whose immediate inputs include the output of j
Outputs = set of output units in the final layer
E = Error

At first, what was searched for is the change in error with respect to change in weights, that is $\frac{\partial E}{\partial w_{ji}}$. By use of the chain rule, it can be expanded as:

$$\frac{\partial E}{\partial w_{ji}} = \frac{\partial E}{\partial z_j} \frac{\partial z_j}{\partial w_{ij}} \quad (\text{where } z_j = \vec{w}_j . \vec{x}_j)$$

Also, $\frac{\partial z_j}{\partial w_{ij}}$ can be expanded as $\frac{\partial}{\partial w_{ij}} \sum_{k=0}^{i} w_{ik} x_{jk} = x_{ji}$

So that $\frac{\partial E}{\partial w_{ji}} = \frac{\partial E}{\partial z_j} x_{ji}$

As $\frac{\partial E}{\partial z_j}$ would be the same for any input weight for neuron j, it is labeled as δ_j.

Case 1: Output Layer

$$E = \frac{1}{2} \sum_{k \in Outputs} (t_k - \sigma(z_k))^2$$

Since the outputs of all units $k \neq j$ are independent of w_{ji}, other elements of the summation are dropped and the contribution to E is considered by j only.

$$\delta_j = \frac{\partial E}{\partial z_j} = \frac{\partial}{\partial z_j}\frac{1}{2}(t_j - o_j)^2$$
$$= -(t_j - o_j)\frac{\partial o_j}{\partial z_j}$$
$$= -(t_j - o_j)\frac{\partial}{\partial z_j}\sigma(z_j)$$
$$= -(t_j - o_j)(1-\sigma(z_j))(\sigma(z_j))$$
$$= -(t_j - o_j)(1-o_j)o_j$$

Thus

$$\Delta w_{ji} = -\eta\frac{\partial E}{\partial w_{ij}} = \eta\delta_j x_{ji}$$

Case 2: Hidden Layer

- For each unit k downstream from j, z_k is a function of z_j.
- The contribution to error by all units $l \neq j$ in the same layer as j is independent of w_{ji}.

What we look for is $\frac{\partial E}{\partial w_{ji}}$ for each input weight w_{ji} for each hidden unit j.

The influence sequence is:

$$w_j \rightarrow z_j \rightarrow o_j \rightarrow z_k \in \forall_k \in Downstream(j) \rightarrow E$$

So,

$$\frac{\partial E}{\partial w_{ji}} = \sum_{k \in Downstream(j)} \frac{\partial E}{\partial z_k}\frac{\partial z_k}{\partial o_j}\frac{\partial o_j}{\partial z_j}\frac{\partial z_j}{\partial w_{ji}} = \sum_{k \in Downstream(j)} \frac{\partial E}{\partial z_k}\frac{\partial z_k}{\partial o_j}\frac{\partial o_j}{\partial z_j}x_{ji}$$

It can be seen that all terms except x_{ji} in the above product are the same regardless of which input weight of unit j we are trying to update. As before, this common quantity is going to be denoted by δ_j. It is also known that:

$$\frac{\partial E}{\partial z_k} = \delta_k, \frac{\partial z_k}{\partial o_j} = w_{kj} \text{ and } \frac{\partial o_j}{\partial z_j} = o_j(1-o_j)$$

So, substituting them:

$$\delta_j = \sum_{k \in Downstream(j)} \frac{\partial E}{\partial z_k} \frac{\partial z_k}{\partial o_j} \frac{\partial o_j}{\partial z_j} = \sum_{k \in Downstream(j)} \delta_k w_{kj} o_j (1 - o_j)$$

Thus,

$$\delta_k = o_j (1 - o_j) \sum_{k \in Downstream(j)} \delta_k w_{kj}$$

To define the backpropagation algorithm formally:

Each training example is of the form $(\vec{x}, \vec{t})$ where $\vec{x}$ is the input vector and $\underline{t}$ is the target vector, η is the learning rate, n_i, n_h and n_o are the number of input, hidden and output nodes respectively. Input from input *i* to unit *j* is denoted x_{ji} and its weight by w_{ji}.

- Create a feed-forward network
- Initialize all weights to small random values
- Until termination condition is met, do
    - For each training sample vector, do
        - Compute the output o_u for every unit of $\vec{x}$
        - For each output unit k, calculate

          $$\delta_k = o_k (1 - o_k)(t_k - o_k)$$

        - For each hidden unit h, calculate

          $$\delta_h = o_h (1 - o_h) \sum_{k \in Downstream(h)} \delta_k w_{kh}$$

        - Update each network weight w_{ji} as follows:

          $$w_{ji} \leftarrow w_{ji} + \Delta w_{ji}$$

          where $\Delta w_{ji} = \eta \delta_j x_{ji}$

APPENDIX C: DERIVATION OF INPUT VARIABLES

The input variables listed are calculated as:

$MA_X(t) = \frac{1}{n}\sum_{i=t-x}^{t} p_i$ where p_i is the price at time i

$$\%K_X = \frac{CCP - L_X}{H_X - L_X}100$$

where

L_x is the lowest low price of the past x days

H_x is the highest high price of the past x days

CCP is the closing price of the last day

$$\%D_X = \frac{\%K_Y + \%K_{Y-1} .... + \%K_{Y-X+1}}{X}$$

where K_y is the current stochastic indicator and X is the number of days to average

$$RSI_X = 100 - \frac{100}{1 + \frac{\sum \text{Positive Changes in x days}}{\sum \text{Negative Changes in x days}}}$$

REFERENCES

Aksoy, H., Sağlam, İ. (2002). Observing a New Anomaly in ISE With the Use of a Classifier System. Retrieved May 5, 2008 from http://ssrn.com/abstract=302650

Alexander, S. (1964). Price Movements in Speculative Markets: Trends or Random Walks. *Industrial Management Review, 2(2),* 7-26.

Avramov, D. (2002). Stock Return Predictability and Model Uncertainty. *Journal of Financial Economics, 64*.

Azoff, E. M. (1994). *Neural network time series forecasting of financial markets*. Chichester; New York, Wiley.

Baumol, W. (1965). The Stock Market and Economic Efficiency. *New York: Fordham University Press.*

Bell, T. B., Ribar, G. S., and Verchio, J. (1990). Neural Nets Versus Logistic Regression: A Comparison of Each Model's Ability to Predict Commercial Bank Failures, *Proceedings of the 1990 Deloitte and Touche/University of Kansas Symposium on Audit Problems*: (pp 29-52). Kansas, United States.

Bernd F, Klaus R. Volatility Estimation with Neural Network. (1996). *Proceedings of the IEEE/IAFE Conference on Computational Intelligence for Financial Engineering.* (pp *177-181).* New York, United States.

Blakemore, S. and Frith, U. (2005). *The Learning Brain: Lessons for Education.* Blackwell Publishing.

Brown, D. (2002). *All About Stock Market Strategies: The Easy Way To Get Started*. McGraw-Hill, 1st edition.

Chiang, W.C., Urban, T.L. and Baldridge, G.W. (1996). A Neural Network Approach to Mutual Fund Net Asset Value Forecasting. *Omega, 24*, 205-215.

Dietrich, J.R., Kachelmeier, S.J., Kleinmuntz, D.N., Linsmeier, T.J. (2001). Market Efficiency, Bounded Rationality, and Supplemental Business Reporting Disclosures. *Journal of Accounting Research, Volume 39: Issue 2*, 243-268.

Dissanaike, G. (2001). *Topics in Empirical Finance: Lecture Notes*. JIMS, University of Cambridge.

Dreiseitl, S. and Ohno-Machado, L. (2002). Logistic regression and artificial neural network classification models: a methodology review. *Journal of Biomedical Informatics*. 35(5-6), 352-359

Dutt, S.D., Ghosh, D. (1999) A Note on the Foreign Exchange Market Efficiency Hypothesis. *Journal of Economics and Finance. 23(2),* 157-161

Dutta, S. and Shekhar, S. (1888). Bond Rating: A Non-Conservative Application of Neural Network, *Proc. IEEE Intl. Conf. on Neural Networks, San Diego*.

Egeli, B., Özturan, M. and Badur, B. (2003). Stock Market Prediction Using Artificial Neural Networks. *Proceedings of the 3rd International Conference on Business (Honolulu, Hawaii June 18-21, 2003).*

Fama E. and Blume, M. (1966). Filter Rules and Stock Market Porfolios. *Journal of Business, 39,* 226-241.

Fama, Eugene F. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work, *Journal of Finance, American Finance Association, 25(2),* 383-417.

Fernandez-Rodriguez, F., Gonzalez-Martel, C., Sosvilla-Rivero, S. (2000). On the profitability of technical trading rules based on artificial neural networks: Evidence from the Madrid stock market. *Economics Letters, Elsevier. vol. 69(1)*, 89-94.

Fernandez-Rodriguez, F., Sosvilla-Rivero, S., Garcia-Artiles, M.D. (1999). Dancing with Bulls and Bears: Nearest Neighbour Forecast for the Nikkei Index. *Japan and the World Economy 11*, 395–413.

Frankfurter, G.M. (2006). The Theory of Fair Markets (TFM) Toward a New Finance Paradigm. *International Review of Financial Analysis, 2006, vol. 15, issue 2,* 130-144.

Freisleben, B. (1992). Stock market prediction with backpropagation networks, Industrial and Engineering Applications of Artificial Intelligence and Expert System. *5th International Conference, Paderborn, Germany,* 451-460.

Gencay R. (1998). Optimization of technical trading strategies and the profitability in security markets, *Economics Letters* 59, 249-254.

Goetzmann, W.N. (1997). An Introduction to Investment Theory. MPPM 540 Course Notes. Yale School of Management, 1997. http://viking.som.yale.edu/will/finman540/classnotes/class1.html, retrieved on 2006.

Hagstrom, G.H. (1999). *The Warren Buffett Way*. Wiley.

Hassan R., Nath B. and Kirley M. (2006). A Fusion Model of HMM, ANN and GA for Stock Market Forecasting. *Expert Systems with Applications.* 33 (1), pp. 171-180.

Heaton J.T. (2000). *Introduction to Neural Networks with Java*. Heaton Research, Inc.

Hellstrom, T. and Holmstrom, K. (1997). Predicting the Stock Market. *Technical Report Series IMa-TOM-1997-07, Malardalen University, Vasteras, Sweden.*

Hornik, K., Stinchcombe, M. and White, H. (1989). Multilayer feedforward networks are universal approximators. *Neural Networks, 2(5)*, 359-366.

Hwang, C., Dorsey, R and Boose, M.A. (1994). Life Insurer Financial Distress Prediction, *Journal of Insurance Regulation, 12(2),* 131-167.

Jingtao, Y., Chew, L.T. and Hean-Lee, P. (1999). Neural Networks for Technical Analysis: a Study on KLCI. *International Journal of Theoretical and Applied Finance, 2(2)*, 221-241.

Kalekar, P.S. (2004). Time series Forecasting using Holt-Winters Exponential Smoothing. *Seminar Paper, Kanwal Rekhi School of Information Technology*.

Kim K.J. and Han I.. (2000). Genetic algorithms approach to feature discretization in artificial neural networks for the prediction of stock price index. *Expert Systems with Applications 19(2),* 125-132.

Kimoto, T., Asakawa, K., Yoda, M. and Takeoka, M. (1990). Stock market prediction system with modular neural networks. *In Proceedings of IJCNN-90, San Diego*.

Luther, R.K. (1998). An Artificial Neural Network Approach in Predicting the Outcome of Chapter 11 Bankrupcy. *Journal of Business and Economic Studies, 4(1),* 57-73.

Malkiel, B.G. (1999). *A Random Walk Down Wall Street*. W.W. Norton & Company, 6th edition.

Man-Chung C, Chi-Cheong W, Chi-Chung L. (2000). Financial Time Series Forecasting By Neural Network Using Conjugate Gradient Learning Algorithm and Multiple Linear Regression Weight Initialization. *Computing in Economics and Finance 61*, Society for Computational Economics.

Murphy, J.J.(1986). Technical Analysis of the Futures Market*, NYIF: New York,* 2-4.

Phua, P.K.H., Ming, D., Lin, W. (2000). Neural Network with Genetic Algorithms for Stocks Prediction, *Fifth Conference of the Association of Asian-Pacific Operations Research Societies, 5th - 7th July, Singapore*.

Quah T.S. and Srinivasan B. (1999). Improving returns on stock investment through neural network selection, *Expert Systems with Applications 17(4):* 295-301.
Pinches, G.E. (1970). The Random Walk Hypothesis and Technical Analysis. *Financial Analysts Journal 26*, pp.104.

Ramon, L. (1997). Using Neural Networks to Forecast Stock Market Prices. Retrieved November 2, 2006, http://www.cs.uiowa.edu/~rlawrenc/research/nn.pdf

Reilly, D.L., Collins, E., Scofield, C. and Ghosh S. (1991). Risk Assessment of Mortgage Applications with a Neural Network System: An Update as the Test Portfolio Ages. *Proceedings of the IEEE International Conference of Neural Networks, II,* 479-482.

Reilly, F. and Brown, K. (1997). Investment Analysis and Portfolio Management, *Philadelphia, San Diego: Dryden Press.*

Ripley, B. D. (1994). Neural Networks and Related Methods for Classification. *Journal of the Royal Statistical Society, 56(3),* 409–456.

Rumelhart, D., Hinton, G., and Williams, R. (1986). *Learning internal representations by error propagation*. MIT Press, Cambridge, MA.

Sarle, W. S. (1994). Neural networks and statistical models. *In Proceedings of the 19th annual SAS Users Group International Conference.*

Schumacher, M., Robner, R., and Vach, W. (1996). Neural Networks and Logistic Regression: Part I. *Computational Statistics and Data Analysis, 21(6),* 661–682.

Simon, H.A. (1955). A behavioral model of rational choice. *Quarterly Journal of Economics*, 69:99--118, 1955.

Sornette, D. (2000). Stock Market Speculation: Spontaneous Symmetry Breaking of Economic Valuation. *Physica A, Volume 284, Number 1*, 355-375(21).

Steiner M. and Wittkemper H. G. (1996). Neural networks as an alternative stock market model. *Neural networks in the capital markets. England: John Wiley & Sons*, 137-149.

Stock. (2006). In Encyclopædia Britannica. Retrieved November 2, 2006, from Encyclopædia Britannica Online: http://www.britannica.com/eb/article-9379610

Svozil, D., Kvasnicka, V. and Pospichal, J. (1997). Introduction to multi-layer feed-forward neural networks. *Chemometrics and Intelligent Laboratory Systems, Volume 39, Number 1*, 43-62(20).

Tagliarini A. Introduction to Backpropagation. CSC 415-515 Course Notes. Department of Computer Science. University of North Carolina at Wilmington, 2001. http://people.uncw.edu/tagliarinig/, retrieved on 2006.

Tam, K.Y. and Kiang, M.Y. (1992). Managerial Applications of Neural Networks: The Case of Bank Failure Predictions. *Management Science, Vol. 38, No. 7*, 926-947.

Tan, C. N. W. (1997). An Artificial Neural Networks Primer with Financial Applications, *Examples in Financial Distress Predictions and Foreign Exchange Hybrid Trading System.*

Technical Analysis from A to Z. (2004). Retrieved November 17, 2007, from http://www.marketscreen.com/help/atoz/default.asp?Num=67

The Cambridge 27, (2001). "Opening Up Economics". Retrieved November 2, 2006, http://www.btinternet.com/~pae_news/review/issue7.htm

Theil, H. (1971). *Principles of Econometrics*, Wiley.

Tromsett, M.C. (1998). *Mastering Fundamental Analysis*. Kaplan Business.

Tsibouris, G. and Zeidenberg, M. (1995). Neural networks as an alternative stock market model. *Neural Networks in the Capital Markets.* John Wiley & Sons.

Van Eyden, R. J. (1996). The Application of Neural Networks in the Forecasting of Share Prices. *Finance and Technology Publishing.*

Vanstone B. and Finnie G. (2006). Combining Technical Analysis and Neural Networks in the Australian Stockmarket. *Proceedings of the 10th IASTED International Conference on Artificial Intelligence and Soft Computing (ASC2006).* Palma de Mallorca, 28-30th August 2006.

Venkataraman, A. (1999). *Machine Learning Course Notes*. Retrieved November 2, 2006, from http://www.speech.sri.com/people/anand/771/html/index.html

Versace, M., Bhatt, R., Hinds, O., Shiffer, M. (2004). Predicting the exchange traded fund DIA with a combination of genetic algorithms and neural networks. *Expert Systems with Applications.* 27 (3), pp. 417-425.

Weigend, A.S., Rumelhart, D.E., and Huberman, B.A. (1990). Generalization by Weight-elimination with Application to Forecasting. In Lippmann, R.P., Moody, J.E. and Touretzky, D.S. (Eds.), *Proceedings of the 3th Conference on Advances in Neural information Processing Systems 3,* (pp. 875-882). Denver, Colorado, United States.

Werbos, P.J. (1974). *Beyond regression: New Tools for Prediction and Analysis in the Behavioral Sciences*. Ph.D. dissertation, Committee on Applied Mathematics, Harvard University.

White, H. (1988). Economic prediction using neural networks: The case of IBM daily stock returns. *IEEE International Conference on Neural Networks, San Diego*, 451-459.

Yao, J. and Poh, H.L. (1995). Forecasting the KLSE Index Using Neural Networks. *IEEE International Conference on Neural Networks 2*.

Yilmaz K. and C. Emre Alper .(2004). Volatility and contagion: Evidence from the Istanbul stock exchange, *Economic Systems, Volume 28, Issue 4,* 353-367.

Chenoweth T. and Obradovic (1996). A multi-component nonlinear prediction system for the S&P 500 index. *Neurocomputing*. *Vol. 10(3),* 275-290.

Yümlü, S., Gürgen, F. And Okay, N. (2004). A comparison of global, recurrent and smoothed-piecewise neural models for Istanbul stock exchange (ISE) prediction. *Pattern Recognition Letters, Volume 26, Issue 13*.

Zekić, M. (1998). Neural Network Applications in Stock Market Predictions - A Methodology Analysis, *Proceedings of the 9th International Conference on Information and Intelligent Systems '98*, 255-263.

Zhang, D. (2005). Machine Learning Applications in Software Engineering. *World Scientific Publishing*.