

Effectiveness of Using Quantified Intermarket Influence for Predicting Trading Signals of Stock Markets

Chandima D. Tilakaratne^{1,2}

Musa A. Mammadov¹

Sidney A. Morris¹

¹ Center for Informatics and Applied Optimization
School of Information Technology and Mathematical Sciences
University of Ballarat, PO Box 663
Ballarat, Victoria 3353, Australia,

Email: {ctilakaratne@students., m.mammadov@, s.morris@}ballarat.edu.au

² Department of Statistics
University of Colombo,
PO Box 1490, Colombo 3, Sri Lanka.

Abstract

This paper investigates the use of influence from foreign stock markets (intermarket influence) to predict the trading signals, buy, hold and sell, of the of a given stock market. Australian All Ordinary Index was selected as the stock market whose trading signals to be predicted. Influence is taken into account as a set of input variables for prediction. Two types of input variables were considered: quantified (weighted) input variables and their un-quantified counterparts. Two criteria was applied to determine the trading signals: one is based on the relative returns while the other uses the conditional probability that a given relative return is greater than or equals zero. The prediction of trading signals was done by Feedforward neural networks, Probabilistic neural networks and so called probabilistic approach which was proposed in past studies. Results suggested that using quantified intermarket influence as input variables to predict trading signals, is more effective than using their un-quantified counterparts.

Keywords: Forecasting, Stock market, Intermarket Influence, Neural networks, Optimization

1 Introduction

Profitability of stock market trading is directly related to the prediction of trading signals. The majority of the past studies (Chenoweth et al., 1996; Fernando et al., 2000; Vanstone, 2006; Wood & Dasgupta, 1996; Yao et al., 1999) focused on classification of future values into two categories (up or down) which are considered to be buy and sell signals. Timely decisions must be made which result in buy signals when the market is low and sell signals when the market is high (Chapman, 1994). However, it is worth holding shares if there is no significant rise or drop in the price index. Therefore, from the practical point of view, it is important to consider the 'hold' category.

The literature (Bhattacharyya & Banerjee, 2004; Eun & Shim, 1989; Taylor & Tonks, 1989; Wu & Su, 1998; Yang et al., 2003) confirms that the world's major stock markets are integrated. Also some studies (Becker et al., 1990; Eun & Shim, 1989; Wu & Su,

1998) provide evidence that US stock markets have strong influence on the other major global markets. These studies confirm the existence of intermarket influence¹ among the global stock markets. Hence, one stock market can be considered as a part of a single global system (Tilakaratne et al., 2006). The influence from one stock market on a dependent market may include the influence from one or more stock markets on the former. This matter indicates that the intermarket influence (from a set on influential markets on a dependent market) needs to be quantified in order to use them effectively in applications such as prediction.

Although, some evidence found in the literature (Olson & Mossaman, 2001; Pan et al., 2005) for the possibility of improving the prediction accuracy by incorporating intermarket influence, none of these studies either aimed at predicting trading signals or incorporated quantified intermarket influence for predictions.

The aim of this paper is to investigate the effectiveness of applying quantified intermarket influence for predicting trading signals, buy, hold and sell, of stock markets. We chose the Australian All Ordinary Index (AORD) as the stock market to be studied. Following Yao & Tan (2000) this study also assumed the major blue chips in the stock basket are bought or sold, and the aggregate price of the major blue chips is the same as the index.

The remainder of this paper is presented as follows: The next section discusses the related work. The third section explains the technique used for quantifying intermarket influence on the AORD together with the corresponding optimization problem. This section also presents the quantified intermarket influence on the AORD. The fourth section defines the trading signals. The fifth section discusses the techniques (algorithms) applied for predicting trading signals and how these algorithms were trained. The input features used for these algorithms are also discussed in this section. The sixth section described how the prediction results were evaluated. The next section presents the results together with interpretations. Final section concludes the paper.

2 Related Work

In the last few decades, there has been a growing number of studies attempting to predict the trading

Copyright ©2007, Australian Computer Society, Inc. This paper appeared at the Sixth Australasian Data Mining Conference (AusDM 2007), Gold Coast, Australia. Conferences in Research and Practice in Information Technology (CRPIT), Vol. 70, Peter Christen, Paul Kennedy, Jiuyong Li, Inna Kolyshkina and Graham Williams, Ed. Reproduction for academic, not-for profit purposes permitted provided this text is included.

¹Intermarket Influence Analysis is defined as the study of relationships between the current price (or a derivative of price) of a dependent market with lagged price (or a derivative thereof) of one or more influential markets (Tilakaratne, 2006; Tilakaratne et al., 2006)

signals of financial market indices. Many past studies (for example, Chenoweth et al. (1996); Fernando et al. (2000); Vanstone (2006); Wood & Dasgupta (1996); Yao et al. (1999)) considered only two trading signals: buy and sell. Although not very common, some studies (for example, Chen et al. (2003); Chenoweth et al. (1996); Kohara et al. (1997); Kuo (1998); Leung et al. (2000); Mizuno et al. (1998)) considered a third signal: hold.

Feedforward neural networks (FNNs) and Probabilistic neural networks (PNNs) seem to be the most commonly used techniques in the literature (Chen et al., 2003; Fernando et al., 2000; Kohara et al., 1997; Leung et al., 2000; Mizuno et al., 1998; Wood & Dasgupta, 1996; Yao et al., 1999) to forecast the trading signals. Some studies (Chen et al., 2003; Leung et al., 2000) showed that the PNNs outperform alternative linear as well as non-linear models in terms of profitability. The literature (Kohara et al., 1997; Yao et al., 1999) reveals that the FNNs outperform the alternative linear models. Furthermore, Leung et al. (2000) found that the PNNs outperformed the FNNs in terms of profitability and predictability.

A PNN directly outputs the trading signals while a FNN outputs the value of the stock market index of interest or its derivative such as relative return. The predicted value is classified as a trading signal according to a certain criterion.

Different studies used different criteria for defining trading signals. Fernando et al. (2000); Kohara et al. (1997); Leung et al. (2000); Mizuno et al. (1998); Vanstone (2006) and Yao et al. (1999) determined the trading signals based on the value of index level or relative return. Chen et al. (2003) and Leung et al. (2000) used a criterion based on the probability to define the trading signals. Studies (such as, Fernando et al. (2000); Vanstone (2006); Yao et al. (1999)), which concern only two signals (buy and sell), considered only one threshold. On the other hand, the studies (for instance Kohara et al. (1997); Kuo (1998); Mizuno et al. (1998)), which considered three trading signals used two threshold criteria. Unlike others, Chen et al. (2003) applied a single threshold criterion as well as a two threshold criterion to determine the trading signals.

Vanstone (2006) suggested that the fundamental variables may be suitable as the input features, if the intention is to do long term forecasts. On the other hand, if the intention is to do short term predictions, technical variables may be more suitable.

Some studies, for example, (Chen et al., 2003; Kohara et al., 1997; Kuo, 1998; Leung et al., 2000; Vanstone, 2006), relied on both fundamental and technical variables for forecasting. Many published research (for instance, Chen et al. (2003); Chenoweth et al. (1996); Fernando et al. (2000); Kohara et al. (1997); Kuo (1998); Leung et al. (2000); Mizuno et al. (1998)) used technical indicators to predict trading signals. Some of these studies (Chenoweth et al., 1996; Fernando et al., 2000; Mizuno et al., 1998) relied only on technical indicators. The use of lagged price or derivatives of the price of the stock market whose trading signals to be predicted seems to be a common feature in many fast studies (Chen et al., 2003; Chenoweth et al., 1996; Fernando et al., 2000; Kohara et al., 1997; Leung et al., 2000; Mizuno et al., 1998).

Nowadays, experts argue that stock markets are influenced by many interrelated factors including the effects of economic, political and even psychological factors. These factors interact with each other in a complex fashion, and it is therefore, very difficult to find an exact set of factors which determine the behaviour of stock markets (Tilakaratne, 2004). The effect of these factors may reflect on the price index on global markets. Hence, it may be able to capture

the combine effect of these factors by using lagged price index to predict future price index of the same markets as well as the other global market.

It is noteworthy that the use of lagged prices or the derivatives of the prices of foreign stock markets to predict the trading signals of a selected market is very rare. The use of such information to predict trading signals may improve the predictability and profitability of the prediction.

3 Quantification of Intermarket Influences

This study selected the AORD as the stock market index whose trading signals to be predicted. In order to investigate the effectiveness of applying quantified intermarket influence for predicting trading signals of the AORD, the intermarket influence on the AORD needs to be quantified. This study adopts the quantification technique developed by Tilakaratne et al. (2007).

This technique quantifies the intermarket influences on a dependent market by finding the coefficients, ξ_i , $i=1, 2, \dots$ (see Section 3.1), which maximise the median rank correlation between the relative return of the Close price of day t of the dependent market and the sum of ξ_i multiplied by the lagged relative returns of the Close prices of a combination of influential markets over a number of small non-overlapping windows of a fixed size. ξ_i measures the contribution from the i th influential market to the combined influence which equals to the optimal correlation.

There is a possibility that the maximum value leads to a conclusion about a relationship which does not exist in reality. In contrast, the median is more conservative in this respect. Therefore, instead of selecting the maximum of the optimal rank correlation, the median was considered.

Spearman's Rank Correlation coefficient was used as the rank correlation measure. For two variables X and Y , Spearman's Rank Correlation Coefficient, r_s , can be defined as:

$$r_s = \frac{n(n^2 - 1) - 6 \sum d_i^2 - (T_x + T_y)/2}{\sqrt{(n(n^2 - 1) - T_x)(n(n^2 - 1) - T_y)}} \quad (1)$$

where n is the total number of bivariate observations of x and y , d_i is the difference between rank of x and rank of y in the i th observation, T_x and T_y are the number of tied observations of X and Y , respectively.

Since, influential patterns between markets may vary with time (Tilakaratne, 2006), the whole study period was divided into a number of moving windows of a fixed length. The correlation structure between stock markets also changes with time (Wu & Su, 1998). Therefore, each moving window was further divided into a number of small windows of length 22 days. 22 days of a stock market time series represent a trading month. Spearman's rank correlation coefficients (see (1)) were calculated for these smaller windows within each moving window.

The absolute value of the correlation coefficient was considered when finding the median optimal correlation. This is appropriate as the main concern is the strength rather than the direction of the correlation (that is either positively or negatively correlated).

The objective function to be maximised (Section 3.1 described below) is defined by Spearman's correlation coefficient. Spearman's correlation coefficient is a piece-wise constant function as it depends on the rank of the elements of the vectors used for the calculation. Solving this type of optimization

problems is extremely difficult. The majority of algorithms need smoothness or at least semi-smoothness of the objective functions to be minimised. Only a few algorithms, that can be used to solve optimization problems with discontinuous objective functions, are available.

In this study, the global optimization algorithm developed by Mammadov (2004) and Mammadov et al. (2005) was used. This algorithm uses a line search mechanism where the descent direction is obtained via a dynamical system approach. The performance of this algorithm has been demonstrated in solving different optimization problems with discontinuous objective functions (for example Koubor et al. (2006)).

3.1 Optimization Problem

Let $Y(t)$ be the relative return of the Close price of a selected dependent market at time t and $X_j(t)$ be the relative return of the Close price of the j th influential market at time t . Define $X_\xi(t-i)$ as:

$$X_\xi(t-i) = \sum_j \xi_j X_j(t-i) \quad (2)$$

where the coefficient $\xi_j \geq 0$, $j = 1, 2, \dots, m$, measures the strength of influence from each influential market X_j . We named these coefficients *quantification coefficients*. m is the total number of influential markets and i represents the time lag.

The aim is to find the optimal values of the quantification coefficients, $\xi = (\xi_1, \dots, \xi_m)$, which maximise the rank correlation $Y(t)$ and $X_\xi(t-i)$ for a given window and time lag i . In the calculations, $i = 0, 1, 2, 3, 4$, which represent influence within a week, were considered. $i = 0$ gives the same day correlation between the Close price of the dependent market and a selected combination of the Closes prices of influential markets. $i = 1$ gives the correlation between the Close price of day t of the dependent market and the Close prices of day $(t-1)$ of a combination of influential markets and this correlation is referred as the previous day's combined influence from the influential markets on the dependent markets. Other time lags can be defined in a similar manner.

The correlation can be calculated for a window of a given size. This window can be defined as;

$$T(t^0, l) = \{t^0, t^0 + 1, \dots, t^0 + (l-1)\} \quad (3)$$

where t^0 is the starting date of the window and l is its size (in days).

The correlation between the variables $Y(t)$, $X_\xi(t-i)$, $t \in T(t^0, l)$, defined on the window $T(t^0, l)$, will be denoted as;

$$\text{Corr}(Y(t), X_\xi(t-i) \parallel T(t^0, l)) \quad (4)$$

For a period of several years, the optimal correlation changes according to the starting point of the window. To define optimal weights for a long period, the following method is applied. Let $[1, T] = 1, 2, \dots, T$ be a given period (for instance a large window). This period is divided into n windows of size l . This study set $l=22$ days.

$$T(t_k, l), \quad k = 1, 2, 3, \dots, n \quad (5)$$

so that,

$$T(t_k, l) \cap T(t_{k'}, l) = \phi \quad \text{for } \forall k \neq k' \quad (6)$$

$$\bigcup_{k=1}^n T(t_k, l) = [1, T] \quad (7)$$

For given i , the correlation coefficient on a window $T(t_k, l)$ is denoted as;

$$C_k^i(\xi) = \text{Corr}(Y(t), X_\xi(t-i) \parallel T(t_k, l)), \quad k = 1, \dots, n. \quad (8)$$

To define an objective function over the period $[1, T]$, the median of the vector, $(C_1^i(\xi), \dots, C_n^i(\xi))$ is used. Now, the main optimization problem can be defined as:

$$\text{Maximise } f(\xi) = \text{Median } (C_1(\xi), \dots, C_n(\xi)); \quad (9)$$

$$\text{s.t. } \sum_j \xi_j = 1, \quad \xi_j \geq 0 \quad j = 1, 2, \dots, m. \quad (10)$$

The solution to (9) (10) is a vector, $\xi = (\xi_1, \dots, \xi_m)$, where ξ_j , $j = 1, 2, \dots, m$, denotes the strength of the influence from the j th influential market.

In this paper, the quantity, $\xi_j X_j$ is called the quantified relative return corresponding the j th influential market.

3.2 Quantification of Intermarket Influence on the AORD

Tilakaratne et al. (2006) revealed that the Close prices of the US S&P 500 Index (GSPC), the UK FTSE 100 Index (FTSE), French CAC 40 Index (FCHI), German DAX Index (GDAXI) as well as that of the AORD itself showed an impact on the direction of the next day's Close price of the AORD. Therefore, this study quantified the intermarket influence from the following two combinations of stock market indices: (i) the GSPC, FTSE, FCHI and the GDAXI; and, (ii) the GSPC, FTSE, FCHI, GDAXI and the AORD. Also Tilakaratne et al. (2007) found that only the Close prices of day $(t-1)$ of these market significantly ² influence the Close price of day t of the AORD. Hence it is sufficient to consider $i=1$ in (2). In other words, the relative returns of the Close prices of day $(t-1)$ of the above mentioned market combinations were considered for the quantification.

For this study, we consider the time series data corresponding to the relative returns of Close prices of the above mentioned five markets, from 2nd July 1997 to 30th December 2005. Since different stock markets are closed on different holidays, the regular time series data sets considered have missing values. If no trading took place on a particular day, the rate of change of price should be zero. Therefore, the missing values of the Close price were replaced by the corresponding Close price of the last trading day.

Relative Returns RR of the daily Close price of the stock market indices were used for the analysis.

$$RR(t) = \frac{P(t) - P(t-1)}{P(t-1)} \quad (11)$$

where $RR(t)$ and $P(t)$ are the relative return and the Close price of a selected index on day t , respectively. Returns are preferred to price, since returns for different stocks are comparable on equal basis.

It is worth noting that the opening and closing times for many of the various markets do not coincide. For example, the Australian, Asian, French and German markets have all closed by the time the US markets open.

²Optimal median rank correlation is significant at the 5% level

The whole study period was divided into six moving windows of three trading years (for stock market time series, 256 days is considered as a trading year). Each time the window was shifted forward by one trading year in order to get the starting point of the next window. For each window, the quantification coefficients, which maximise the median Spearman's rank correlation between the relative return of the Close price of day t of the AORD and the sum of the quantification coefficient multiplied by the relative returns of the Close prices day $(t - 1)$ of the potential influential markets, were derived.

Table 1 and 2 presents the quantification coefficient associated with each market for each window with the corresponding combined influence (optimal median Spearman's correlation) for market combinations (1) and (2), respectively.

Table 1: Optimal values of quantification coefficients (ξ) and the optimal median Spearman's correlations corresponding to market combination (1) for different moving windows

Mov. win. no.	Optimal values of ξ				Opt. median Spear. corr.
	GSPC	FTSE	FCHI	GDAXI	
1	0.57	0.29	0.12	0.02	0.578
2	0.61	0.18	0.08	0.13	0.548
3	0.77	0.09	0.13	0.01	0.568
4	0.79	0.06	0.15	0.00	0.579
5	0.56	0.17	0.03	0.24	0.590
6	0.66	0.06	0.08	0.20	0.5359

Table 2: Optimal values of quantification coefficients (ξ) and the optimal median Spearman's correlations corresponding to market combination (2) for different moving windows

Mov. win. no.	Optimal values of ξ					Opt. median Spear. corr.
	GSPC	FTSE	FCHI	GDAXI	AORD	
1	0.56	0.29	0.10	0.03	0.02	0.580
2	0.58	0.11	0.12	0.17	0.02	0.550
3	0.74	0.00	0.17	0.02	0.07	0.570
4	0.79	0.07	0.14	0.00	0.00	0.580
5	0.56	0.17	0.04	0.23	0.00	0.590
6	0.66	0.04	0.09	0.20	0.01	0.537

Optimal median correlations are significant at 5% level irrespective of the window number and the market combination (Table 1 to 2). The GSPC seems to be the most influential market on the AORD.

The quantification coefficients (ξ) presented in the above two tables (Table 1 to 2) were used when predicting the trading signals of the AORD (Section 5.1).

4 Defining Trading Signals

Most of the past studies (Fernando et al., 2000; Vanstone, 2006; Wood & Dasgupta, 1996; Yao et al., 1999) classified the future values into buy or sell signals based on the direction of the trend (upward or downward) of the future values. The studies (Chen et al., 2003; Chenoweth et al., 1996; Kohara et al., 1997; Kuo, 1998; Leung et al., 2000; Mizuno et al., 1998) aimed at predicting three trading signals (buy,

hold and sell) applied two threshold criteria. Since, this study also consider three signals, the following criterion, which uses two thresholds, was introduced to determine the trading signals.

Criterion A

$$\begin{aligned} \text{buy} & \text{ if } Y(t) \geq l_u \\ \text{hold} & \text{ if } l_l < Y(t) < l_u \\ \text{sell} & \text{ if } Y(t) \leq l_l \end{aligned}$$

where $Y(t)$ is the relative return of the Close price of day t of the AORD while l_u and l_l are two thresholds.

The values of l_u and l_l depend on the traders' choice. There is no standard criterion found in the literature how to decide the values of l_u and l_l and these values may vary from one stock index to another. A traders may decide the values for these threshold according to his/her knowledge and experience.

We tested a range of values for l_u and l_l . The selection of suitable pair of values was done on basis of the profitability. Detailed description is found in Section 5.2.

The other way to identify the trading signals is to consider the probability of the predicted return is in upward (or downward) trend (Chen et al., 2003; Leung et al., 2000). Chen et al. (2003) considered the corresponding trading signal is a buy signal if this probability is above 0.7 and a sell signal if its value is below 0.5. Otherwise, the corresponding trading signal was considered as a hold signal. However, these limits associate with the probability of the predicted return is in upward trend may vary according to the stock index. Following Leung et al. (2000), this study also employed a criterion based on probability (Criterion B) to identify the trading signals.

Criterion B

$$\begin{aligned} \text{buy} & \text{ if } P \geq p_2 \\ \text{hold} & \text{ if } p_1 < P < p_2 \\ \text{sell} & \text{ if } P \leq p_1 \end{aligned}$$

where P is the conditional probability that a given relative return of the Close price of day t of the AORD ≥ 0 . The choice of p_1 and p_2 is based on the profitability of trading and described in Section 5.4.1.

5 Predicting Trading Signals of the AORD

Since stock market time series are non-linear systems, the linear classification techniques (such as, linear regression, vector autoregressive models, linear discriminant analysis and ARIMA models) are not suitable for our prediction purpose. The literature (Section 2) shows that the non-linear classification techniques such as FNN, PNN and probabilistic approached (Section 5.4) proposed by Leung et al. (2000) performed well in predicting the trading signals of stock market movements. Therefore, this study also adopted these three techniques (algorithms) to predict the trading signals of the AORD.

5.1 Data Set Generation for Prediction Experiments

Two types of inputs sets were used as input features to the prediction algorithms (FNN, PNN and probabilistic approach); one set consists of the quantified relative returns while the other set contains the unquantified relative returns. The aim was to examine the effectiveness of applying quantified intermarket influence for the prediction of interest.

As previously mentioned in Section 3.2, the Close price on day t of the AORD is affected by those prices

on day $(t - 1)$ of the GSPC, FTSE, FCHI, GDAXI as well as the AORD itself, the two combinations (mentioned in Section 3.2) of stock markets were considered when forming the input sets. Therefore, the input sets used for algorithms are:

1. Four input features of the relative returns of the Close prices on day $(t - 1)$ of the market combination (1)
- $(\text{GSPC}(t - 1), \text{FTSE}(t - 1), \text{FCHI}(t - 1), \text{GDAXI}(t - 1))$; denoted as GFFG;
2. Four input features of the quantified relative returns of the Close prices on day $(t - 1)$ of the market combination (1)
- $(\xi_1^A \text{GSPC}(t - 1), \xi_2^A \text{FTSE}(t - 1), \xi_3^A \text{FCHI}(t - 1), \xi_4^A \text{GDAXI}(t - 1))$; denoted as GFFG-q;
3. Five input features of the relative returns of the Close prices of on day $(t - 1)$ the market combination (2)
- $(\text{GSPC}(t - 1), \text{FTSE}(t - 1), \text{FCHI}(t - 1), \text{GDAXI}(t - 1), \text{AORD}(t - 1))$; denoted as GFFGA;
4. Five input features of the quantified relative returns of the Close prices on day $(t - 1)$ of the market combination (2)
- $(\xi_1^A \text{GSPC}(t - 1), \xi_2^A \text{FTSE}(t - 1), \xi_3^A \text{FCHI}(t - 1), \xi_4^A \text{GDAXI}(t - 1), \xi_5^A \text{AORD}(t - 1))$; denoted as GFFGA-q;

$(\xi_1, \xi_2, \xi_3, \xi_4)$ and $(\xi_1^A, \xi_2^A, \xi_3^A, \xi_4^A, \xi_5^A)$ are the solutions to (9) (10) (in Section 3.1) corresponding to the two market combinations: the GSPC, FTSE, FCHI and the GDAXI, and the GSPC, FTSE, FCHI, GDAXI and the AORD, respectively. We note that it may be $\xi_i \neq \xi_i^A$, for $i=1, 2, 3, 4$.

As mentioned previously in Section 3, the influential patterns between markets may vary with time. Hence, to capture these varying patterns, the algorithms were trained for several moving windows. The same six moving windows employed for quantified intermarket influence on the AORD (Section 3.2), was used for training the algorithms. Hence, the respective optimal values of quantification coefficient presented in Table 1 to 2 can be used as the corresponding values of ξ_i , $i=1, 2, 3, 4$ and ξ_i^A , $i=1, 2, 3, 4, 5$, respectively.

Each moving window consists of 768 samples (relative returns of three trading years). The most recent 10% of data (76 samples) of each window was allocated for testing while the remaining 90% (692 samples) was allocated for training. The training set was further divided into two sets; the most recent 22.2% of data of each training set (20% of the full data set) was allocated for validation while the remaining 77.8% (70% of the full data set) was used for training.

5.2 Training PNNs

The six moving windows, which was mentioned in Section 5.1, were used for the experiments with PNN. The above mentioned four input sets (Section 5.1) were considered for network training. Networks output the class (buy, hold, or sell) of AORD according to Criterion A (Section 4). In Criterion A, different pairs of values of l_l and l_u were tested; $l_l = 0.003, 0.0035, 0.004, 0.0045, \dots, 0.007$ and $l_u = 0.003, 0.0035, 0.004, 0.0045, \dots, 0.007$. The aim was to find suitable pair of values for l_l and l_u , which yield higher profits. Section 5.2.1 below describes how these values were determined.

The lost incurred by misclassification, for each class was assumed to be equal. The joint distribution

of the input variables was assumed to be Gaussian. The parameters of the distribution were estimated by using the training data. When there were multiple inputs, the average standard deviation of the individual input variables was considered as the standard deviation of the joint distribution.

5.2.1 Choosing the Values for l_u and l_l

The validation set was used to determine the appropriate values for l_u and l_l in Criterion A. By varying the values of l_u and l_l , the corresponding trading signals for the validation was obtained. Trading simulations (described in Section 6.1) were performed on the trading signals corresponding to each pair of values considered. The pairs of values of l_u and l_l which gives the highest rate of return in each window for each input set are shown in Table 3.

Table 3: $(-l_l, l_u)$ which gives the maximum rate of return (relevant to validation sets) for the four input sets for different windows (Some windows have more than one pair of values for $(-l_l, l_u)$ and NA shows the cases where no trading (either buy or sell) took place

Input set	Window number	$(-l_l, l_u)$
GFFG	1	(-0.0030, 0.0030)
	2	(-0.0040, 0.0030)
	3	(-0.0030, 0.0030)
	4	(-0.0055, 0.0030)
	5	NA
	6	NA
GFFG-q	1	(-0.0035, 0.0030)
	2	(-0.0035, 0.0030)
	3	(-0.0030, 0.0030)
	4	(-0.0040, 0.0030)
	5	NA
	6	NA
GFFGA	1	(-0.0030, 0.0030)
	2	(-0.0035, 0.0035)
	3	(-0.0040, 0.0035)
	4	(-0.0045, 0.0035)
	5	(-0.0030, 0.0030)
	6	(-0.0035, 0.0030)
		(-0.0040, 0.0030)
		(-0.0045, 0.0030)
		(-0.0050, 0.0030)
		(-0.0055, 0.0030)
		(-0.0060, 0.0030)
		(-0.0065, 0.0030)
GFFGA-q	1	(-0.0050, 0.0035)
	2	(-0.0035, 0.0035)
	3	(-0.0030, 0.0035)
	4	(-0.0030, 0.0030)
	5	(-0.0030, 0.0030)
	6	(-0.0030, 0.0035)
		(-0.0035, 0.0035)
		(-0.0030, 0.0040)
		(-0.0035, 0.0040)

The value of l_l which gives the highest rate of return varies from 0.0030 to 0.0070 while that for l_u takes values from 0.0030 to 0.0040 (Table 3). Therefore, the middle values of the ranges, $[0.0030, 0.0070]$ and $[0.0030, 0.0040]$ (that is 0.0050 and 0.0035), were chose as the appropriate values for l_l and l_u , respectively.

To evaluate the prediction results, the trading simulation was performed on the trading signals obtained from the test results. Criterion A with $l_l=0.0050$ and $l_u=0.0035$ was applied to determined these trading signals.

5.3 Training FNNs

The same six moving windows, that considered for PNN experiments (Section 5.2), were used for experiments with FNN. Three-layered FNNs with one hidden layer were trained for each one of the six moving windows considered. In each window, FNN was trained for 500 times.

The same sets of inputs (which were used as inputs for PNN) were considered as the inputs to FNN. These networks output the relative return of the Close price of the AORD. The average value of the prediction (over 500) for each day was calculated and this average value subsequently classified into the three classes of interest according to Criterion A (Section 4).

For this study we chose the same values, which used for the PNN experiments (Section 5.2.1), as the corresponding limits of Criterion A. In other words, 0.0050 and 0.0035 were taken as l_l and l_u , respectively.

A tan-sigmoid function was used as the transfer function between the input layer and the hidden layer while the linear transformation function was employed between the hidden and the output layers. The slope of a sigmoid function approaches zero as the input gets large and therefore the gradient can have a very small magnitude. If the steepest descent algorithm is used, this causes small changes in the weights and biases, even though the weights and biases are far from their optimal values (Demuth et al., 2006). Resilient backpropagation training algorithm (Rprop) (Riedmiller & Braun, 1989) eliminates these harmful effects of the magnitudes of the partial derivatives. It uses the sign of the derivative to determine the direction of the weight update; the magnitude of the derivative has no effect on the weight update. Therefore, the networks were trained with the resilient backpropagation training algorithm.

Different number of neurons for the hidden layer and different values for learning rate as well as the momentum coefficient were tested. FNNs gave the best results when there were three neurons in the hidden layer and the learning rate and the momentum coefficient were 0.003 and 0.01, respectively.

5.4 Probability Based Approach for Forecasting Trading Signals

Let $Y(t)$ be the relative return of the Close price of day t of the AORD (the target variable). The data is classified into two classes using $Y(t)$ as below:

$$\text{Upward Trend if } Y(t) \geq 0 \quad (12)$$

$$\text{Downward Trend if } Y(t) < 0 \quad (13)$$

Suppose that C_i is the target class corresponding to the i -th observation of a selected set of input features. Also let:

$$C_i = 1 \text{ if } Y(t) \geq 0 \quad (14)$$

$$C_i = 0 \text{ if } Y(t) < 0 \quad (15)$$

Then the conditional probability (P) that a given observation X_i belongs to the upward trend class is;

$$\begin{aligned} P &= \Pr(C_i = 1 | X = X_i) \\ &= \frac{\Pr(X = X_i | C_i = 1) \times \Pr(C_i = 1)}{\sum \Pr(X = X_i | C_i = j) \times \Pr(C_i = j)} \quad (16) \end{aligned}$$

$\Pr(X = X_i | C_i = j)$, $j=0,1$, can be calculated assuming a Gaussian distribution.

$$\Pr(X = X_i | C_i = j) = \frac{1}{2\pi^{I/2}\sigma^I} \sum_{i=1}^{n_j} \exp \frac{-(X_i - j)'(X_i - j)}{2\sigma^2 n_j} \quad (17)$$

where I is the number of input features included in the input set and n_j is the number of training observations the class in which $C_i = j$, $j=0,1$.

The probability corresponding to each class (upward trend or downward trend) can be calculated as below:

$$\Pr(C_i = j) = \frac{n_j}{N_T} \quad (18)$$

where N_T is the total number of observations in the training set.

5.4.1 Parameter Estimation

The same six moving windows, which were used for training FNNs, PNNs were used for these experiments. Unlike Leung et al. (2000), this study considers a validation set in addition to the training and the test sets. The above mentioned four input sets (Section 5.1) were considered as the input variables.

As described by Leung et al. (2000), the parameters of the Gaussian distribution was estimated by using the training data set. This study assumes that the average standard deviations of the input variables (of the training sample) as the value of σ of the Gaussian distribution (see (17)). $\Pr(C_i = j)$, $j=0,1$ (see 18), was also estimated by using the training data.

Using the estimated Gaussian distribution and $\Pr(C_i = j)$, the conditional probability that a given observation X_i in the validation set belongs to the upward trend class, was derived. This probability associated with to each observation in the validation set was found.

Applying Criterion B (described in Section 4) on these probabilities relevant to the validation set, the corresponding trading signals (for the validation set) were determined. Different values for p_1 and p_2 were considered; $p_1=0.20, 0.25, 0.30, 0.35, \dots, 0.50$ and $p_2=0.50, 0.55, 0.60, 0.65, \dots, 0.80$. Practically, the conditional probability that a given observation X_i on an upward trend, P is below 0.5, the corresponding signal can not be considered as a buy signal. In contrast, if this probability is above 0.5, then the corresponding trading signal will not be a sell signal. Therefore, the upper limit for p_1 as well as the lower limit for p_2 should be 0.5. Leung et al. (2000) fixed the lower limit of p_1 at 0.254 and the upper limit of p_2 at 0.746. Therefore, we also chose closer values for the lower limit of p_1 and the upper limit of p_2 .

By varying the values of p_1 and p_2 , we aimed to find a suitable pair of values (for p_1 and p_2) which gives higher profits. Trading simulations (described in Section 6.1) were performed on the trading signals obtained by substituting different values of (p_1, p_2) in Criterion B. Values of (p_1, p_2) which yields highest rate of return (profit) for each window for each input set are shown in Table 4. These rates of returns were obtained from the trading simulations performed on the trading signals obtained from the validation set of each window.

According to trading simulations performed on the validation sets, the value of p_1 which yield the highest rate of return from the trading simulations varied between 0.4 and 0.5 (Table 4). The corresponding range for p_2 was [0.50, 0.70]. Therefore, the median value of the ranges of were taken as the values of p_1 and p_2 . In other words, it was assumed $p_1=0.45$ and $p_2=0.60$.

Table 4: (p_1, p_2) which gives the maximum rate of return (relevant to validation sets) for the four input sets for different windows (Some windows have more than one pair of p_1 and p_2)

Input set	Window number	(p_1, p_2)
GFFG	1	(0.50, 0.50)
	2	(0.50, 0.60)
	3	(0.50, 0.65)
	4	(0.40, 0.60)
	5	(0.50, 0.60)
	6	(0.50, 0.70)
GFFG-q	1	(0.50, 0.60)
	2	(0.40, 0.60), (0.45, 0.60)
	3	(0.45, 0.55), (0.50, 0.55)
	4	(0.45, 0.50), (0.50, 0.50)
	5	(0.50, 0.60)
	6	(0.40, 0.70)
GFFGA	1	(0.50, 0.50)
	2	(0.50, 0.65)
	3	(0.50, 0.60)
	4	(0.40, 0.60)
	5	(0.45, 0.60)
	6	(0.50, 0.70)
GFFGA-q	1	(0.50, 0.55)
	2	(0.50, 0.50)
	3	(0.50, 0.60)
	4	(0.40, 0.50)
	5	(0.50, 0.60)
	6	(0.40, 0.70)

These probability levels are different from the corresponding probability levels used in Chen et al. (2003) (Section 4).

The conditional probability that a given test observation X_i belongs to the upward trend class was found. Finally, Criterion B with $p_1=0.45$ and $p_2=0.60$ was applied to determine the trading signals of each test set.

6 Evaluation of Predictions

The prediction results were evaluated in terms of profitability. Profitability was measured by the rate of return obtained by performing trading simulations.

The rate of return is a measure that provides the net gain in assets as a percentage of the initial investment. Profit depends not only on the accuracy of the forecasts but also on the trading strategy.

Different past studies employed different trading strategies to assess the profitability of the forecasts (Thawornwong & Enke, 2004). This study adopted buy and sell strategy to form the trading simulation. As mentioned in Section 1, this study assumed the major blue chips in the stock basket of the Australian stock exchange are bought or sold, and the aggregate price of the major blue chips is the same as the AORD.

The speciality of the trading simulation proposed in this study is that it search for the proportion of money that a trader needs to invest and the proportion of shares that he/she needs to sell in order to maximise the profit. In this sense, the proposed simulation is very close to the reality.

6.1 Trading Simulations

This study assumes that at the beginning of each period, the trader has some amount of money as well as

a number of shares. Furthermore, it is assumed that the value of money in hand and the value of shares in hand are equal. Two types of trading simulations were used: (1) response to the predicted trading signals which might be a buy, hold or a sell signal; (2) do not participate in trading, and hold the initial shares and the money in hand until the end of the period. The second simulation was used as a benchmark.

6.1.1 First Trading Simulation (The Proposed Trading Simulation)

Let the value of the initial money in hand be M^0 and the number of shares at the beginning of the period be S^0 . $S^0 = M^0/P_0$, where P_0 is the Close price of the AORD on the day before the starting day of the trading period.

Also let M_t, S_t, P_t, VS_t be the money in hand, number of shares, Close price of the AORD, value of shares holding on the day t ($t=1, 2, \dots, T$), respectively. This simulation assumes that always a fixed amount of money is used in trading regardless of the trading signal is buy or sell. Let this fixed amount be denoted as F^0 and be equal to M^0/L , $L > 0$. In the calculations $L = 1, 2, \dots, 10$ is considered. When $L = 1$, F^0 equals to M^0 , when $L = 2$, F^0 equals to 50% of M^0 and so on. Let Δ_t^b and Δ_t^s be the number of shares buy and the number of shares sell at day t , respectively.

Suppose the trading signal at the beginning of the day t is a buy signal. Then the trader spends $F = \min\{F^0, M_{t-1}\}$ amount of money to buy a number of shares at a rate of the previous day's Close price.

$$M_t = M_{t-1} - F, \quad F = \min\{F^0, M_{t-1}\} \quad (19)$$

$$\Delta_t^b = \frac{F}{P_{t-1}} \quad (20)$$

$$S_t = S_{t-1} + \Delta_t^b \quad (21)$$

$$VS_t = S_t \times P_t \quad (22)$$

Suppose the trading signal is a hold signal, then:

$$M_t = M_{t-1} \quad (23)$$

$$S_t = S_{t-1} \quad (24)$$

$$VS_t = S_t \times P_t \quad (25)$$

Let the trading signal at the beginning of the day t is a sell signal. Then the trader sells $S' = \min\{(F^0/P_{t-1}), S_{t-1}\}$ amount of shares.

$$\Delta_t^s = S', \quad S' = \min\{(F^0/P_{t-1}), S_{t-1}\} \quad (26)$$

$$M_t = M_{t-1} + S' \times P_{t-1} \quad (27)$$

$$S_t = S_{t-1} - \Delta_t^s \quad (28)$$

$$VS_t = S_t \times P_t \quad (29)$$

It should be noted that a buy signal that immediately follows another buy signal will be treated as a hold signal. Also, if all shares have been sold, a sell signal is ignored.

6.1.2 Second Trading Simulation (The Benchmark Trading Simulation)

In this case the trader does not participate in trading. Therefore, $M_t = M^0$ and $S_t = S^0$ for all $t=1, 2, \dots, T$. However, the value of the shares changes with the time and therefore, the value of shares at day t , $VS_t = S^0 \times P_t$.

6.2 Rate of Return

At the end of the period (day T) the total value of money and shares in hand:

- for the first trading simulation

$$TC = M_T + S_T \times P_T \quad (30)$$

- for the second trading simulation

$$TC = M^0 + S^0 \times P_T \quad (31)$$

The rate of return ($R\%$) at the end of a trading period is calculated as below:

$$R\% = \frac{TC - 2M_0}{2M_0} \times 100 \quad (32)$$

7 Results and Interpretations

This sections presents the rates of returns corresponding to FNN, PNN and the probabilistic approach, with the interpretations.

The trading simulations showed that the highest rate of return was obtained when the full amount of money in hand is invested and the full amount of shares in hand is sold. This matter was true for all the input sets as well as all the windows used.

Table 5 shows the average rates of return obtained by performing the proposed trading simulation (described in Section 6.1) on the prediction results (corresponding to the test set) obtained by FNN, PNN and the probabilistic approach, for different input sets.

Table 5: Average (over the six windows) rates of return relating to three algorithms trained with the four input sets (The annual average rate of return relating to the benchmark simulation = 9.57%)

Algorithm	Input set	Rate of return for test period	Annual rate of return
PNN	GFFG	3.65%	12.31%
	GFFG-q	3.99%	13.44%
	GFFGA	6.43%	21.66%
	GFFGA-q	6.91%	23.29%
FNN	GFFG	8.23%	27.72%
	GFFG-q	7.61%	25.63%
	GFFGA	8.21%	27.65%
	GFFGA-q	8.50%	28.63%
Prob. approach	GFFG	5.92%	19.94%
	GFFG-q	9.69%	32.64%
	GFFGA	8.91%	30.01%
	GFFGA-q	9.25%	31.16%

Table 5 evidences that, irrespective of the input set used, a trader can gain higher profits by responding to the trading signals produced by any algorithm considered. The average rate of return, obtained from the probabilistic approach, is higher when the predictions are based on the quantified intermarket influence (that is, input sets GFFG-q and GFFGA-q) than when the predictions are based on un-quantified intermarket influence. The highest rate of return was obtained when the predictions are based quantified intermarket influence from the GSPC, the three European markets and the AORD (input set GFFG-q).

The rates of return relevant to PNN also suggests that quantified intermarket influence produced more

profitable trading signals, than their un-quantified counterparts. The highest rate of return was obtained when the quantified intermarket influence from the GSPC, the three European markets as well as the AORD itself was used as the input variables to predict the trading signals.

Results relating to FNN indicates that higher profits can be obtained when the quantified intermarket influence from the GSPC, the three European markets and the AORD were used as the input variables than using their un-quantified counterparts. However, the results relevant to the market combination of the GSPC and the three European markets suggests the opposite.

This exceptional behaviour of FNN may be due the inappropriateness of the Ordinary least squares (OLS) error function (used in the standard FNNs) for a classification problem. FNNs output the value of the prediction, but not the predicted class. OLS error function minimises the difference between the actual and predicted values irrespective of whether the predicted value in is the correct class or not.

8 Conclusions and Further Research

Probabilistic approached described in Section 5.4 seems to be a better technique to predict the trading signals of the AORD, than PNN and FNN. The criterion applied to determine the trading signals may also contributed to the effectiveness of this approach. This criterion uses the conditional probability that a relative return is in upward trend.

In general, the prediction results were better when the quantified intermarket influence on the AORD used as the input variables, than when their un-quantified counterparts were used as input variables. The exceptional behaviour of the FNN may be due to the inappropriateness of its error function for a classification problem.

Designing new neural network algorithm with appropriate error function, for predicting trading signals may be a good direction for future research. Such error function can be proposed by introducing a penalty to the Ordinary least squares error function, to deal with incorrectly predicted trading signals.

References

- Becker, K. G., Finnerty, J. E. and Gupta, M. (2003), 'The Intertemporal Relation Between the U. S. and Japanese Stock markets', *The Journal of Finance* **XLV**(4), 1297-1306.
- Bhattacharyya M. and Banerjee, A. (2004), 'Integration of Global Capital Markets: An Emphirical Exploration', *International Journal of Theoretical and Applied Finance* **7**(4), 385-405.
- Cao, L. and Tay, F. E. H. (2001), 'Financial forecasting using support vector machines', *Neural Computing and Applications* **10**, 184-192.
- Chapman, A. J. (1994), 'Stock market trading systems through neural networks: developing a model.', *International Journal of Applied Expert Systems* **2**(2), 88-100.
- Chen, A. S., Leung, M. K. and Daouk, H. (2003), 'Application of neural networks to an emerging financial market: forecasting and trading the Taiwan Stock Index', *Computers and Operations Research* **30**, 901-923.
- Chenoweth, T., Obradovic Z. and Stephenlee, S. (1996), 'Embedding Technical Analysis into Neural

- Network Based Trading Systems', *Applied Artificial Intelligence* **10**, 523-541.
- Demuth, H., Beale, M. and Hagan, M. (2006), *Neural Network Toolbox User's Guide*, The Math Works Inc..
- Eun C. S. and Shim, S. (1989), 'International Transmission of Stock Market Movements', *The Journal of Financial and Quantitative Analysis* **24**(2), 241-256.
- Fernando, F. R., Christian, G. M. and Simon, S. R. (2000), 'On the profitability of technical trading rules based on artificial neural networks: Evidence from the Madrid stock market', *Economics Letters* **69**, 89-94.
- Kaasra, I. and Boyd, M. (1996), 'Designing a neural network for forecasting financial and economic time series', *Neurocomputing* **10**, 215-236.
- Kim, S. H. and Chun S. H. (1998), 'Graded forecasting using an array of bipolar predictions: Application of probabilistic neural networks to a stock market index', *International Journal of Forecasting* **14**, 323-337.
- Koubor, S., Ugon, J., Mammadov, M. A., Rubinov, A. M. and Kruger, A. (2006), Coverage in WLAN: Optimization Model and Algorithm, in 'First IEEE International Conference on Wireless Broadband and Ultra Wideband Communications', Sydney, Australia.
- Kohara, K., Ishikawa, T., Fukuhara, Y. and Nakamura, Y. (1997), 'Price Prediction Using Prior Knowledge and Neural Networks', *Intelligent Systems in Accounting, Finance and Management* **6**, 11-22.
- Kuo, R. J. (1998), 'A Decision Support System for the Stock Market Through Integration of Fuzzy Neural Networks and Fuzzy Delphi', *Applied Artificial Intelligence* **12**, 501-520.
- Leung, M. T., Daouk, H. and Chen, A. S. (2000), 'Forecasting stock indices: a comparison of classification and level estimation models', *International Journal of Forecasting* **16**, 173-190.
- Mammadov, M. A. (2004), A new global optimization algorithm based on dynamical systems approach, in A. Rubinov and M. Sniedovich eds, 'The Sixth International Conference on Optimization: Techniques and Applications (ICOTA6)', Ballarat, Australia.
- Mammadov, M. A., Rubinov A. M. and Yearwood, J. (2005), Dynamical systems described by relational elasticities with applications to global optimization, in V. Jeyakumar and A. Rubinov, eds, 'Continuous Optimisation: Current Trends and Applications', Springer, pp. 365-387.
- Mizuno, H., Kosaka, M. and Yajima, H. (1998), 'Application of Neural Network to Technical Analysis of Stock Market Prediction', *Studies in Information and Control* **7**, 111-120.
- Nguyen, D. and Widrow, B. (1990), 'Improving the learning speed of 2-layer neural networks by choosing initial values of the adaptive weights', *Proceedings of the International Joint Conference on Neural Networks (IJCNN)*, **3**, 21-26.
- Olson D. and Mossaman, C. (2001), 'Cross-correlation and Predictability of Stock Returns', *Journal of Forecasting* **20**, 145-160.
- Pan, H., Tilakaratne, C., and Yearwood, J. (2005), 'Predicting the Australian Stock Market Index Using Neural networks Exploiting Dynamical Swings and Intermarket Influences', *SJournal of Research and Practice in Information Technology* **37**, 43-55.
- Riedmiller, M. and Braun, H. (1993), A direct adaptive method for faster backpropagation learning: The RPROP algorithm, in 'The IEEE International Conference on Neural Networks', California, USA.
- Taylor, M. P. and Tonks, I. (1989), 'The Internationalisation of Stock Markets and the Abolition of U.K. Exchange Control', *Review of Economics and Statistics* **71**, 332-336.
- Thawornwong, S. and Enke, D. (2004), Forecasting stock returns with artificial neural networks, in G. P. Zhang eds, *Neural Networks in Business Forecasting*, chapter 3, Idea Group Publishing.
- Tilakaratne, C. D. (2004), A Neural Network Approach for Predicting the Direction of the Australian Stock Market Index, MIT (by research), University of Ballarat, Australia.
- Tilakaratne, C. D. (2006), A Study of Intermarket Influence on the Australian All Ordinary Index at Different Time Periods, in 'Second Australian Business and Behavioural Sciences Association (ABBASA) International Conference', Adelaide, Australia.
- Tilakaratne, C. D., Mammadov M. A. and Hurst, C. P. (2006), Quantification of Intermarket Influence Based on the Global Optimization and Its Application for Stock Market Prediction, in K-L. Ong, K. Smith-Miles, V. Lee & W-K. Ng eds, 'International Workshop on Integrating AI and Data Mining (AIDM'06)', IEEE Computer Society, California, USA, pp. 42-49.
- Tilakaratne, C. D., Morris S. A., Mammadov M. A. and Hurst, C. P. (2007), Quantification of Intermarket Influence on the Australian All Ordinary Index Based on Optimization Techniques, in W. Read eds, 'Proceedings of the 13th Biennial Computational Techniques and Applications Conference (CTAC'06)', *ANZIAM Journal*, **48**, C104-C118.
- Vanstone, B. (2006), Trading in the Australian stock-market using artificial neural networks, Ph.D., Bond University, Australia.
- Wood, D. and Dasgupta, B. (1996), 'Classifying Trend Movements in the MSCI U.S.A. Capital Market Index - A Comparison of Regression, ARIMA and Neural Network Methods', *Computers & Operations Research* **23**(6), 611-622.
- Wu, C. and Su, Y. (1998), 'Dynamic Relations among International Stock Markets', *International Review of Economic and Finance* **7**(1), 63-84.
- Yang, J., Khan M. M. and Pointer, L. (2003), 'Increasing Integration Between the United States and Other International Stock Markets?: A Recursive Cointegration Analysis', *Emerging Markets Finance and Trade* **39**(6), 39-53.
- Yao, J., Tan, C. H. and Poh, H. L. (1999), 'Neural Networks for Technical Analysis: A Study on KLCI', *International Journal of Theoretical and Applied Finance* **2**(2), 221-241.
- Yao, J. and Tan, C. H. (2000), Time Dependent Directional Profit Model for Financial Time Series Forecasting, in 'IEEE-INNS-ENNS International Joint Conference on Neural Networks (IJCNN'00)', Como, Italy.

