Expert Systems with Applications 37 (2010) 1814–1820

Contents lists available at ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

A hybrid forecast marketing timing model based on probabilistic neural network, rough set and C4.5

Jao-Hong Cheng [a,*], Huei-Ping Chen [a], Yi-Min Lin [b]

[a] Department of Information Management, National Yunlin University of Science and Technology, 123, Section 3, University Road, Douliou, Yunlin 64002, Taiwan
[b] Department of Information Management, Transworld Institute of Technology, 1221, Jen-Nang Rd., Doulin, Yunlin, Taiwan

ARTICLE INFO

Keywords:
Forecasting
Supply chain
Probabilistic neural network (PNN)
Decision tree_C4.5
Rough set theory (RST)

ABSTRACT

One of the major difficulties in investment strategy is to integrate supply chain with finance for controlling the marketing timing. The present study uses not only the different indexes in fundamental and technical analysis, but also the rough set theory and artificial neural networks inference system to construct three investment market timing classification models. This includes probabilistic neural network classification model, rough set classification model and hybrid classification model combining probabilistic neural network, rough sets and C4.5 decision tree. We use the forecasting accuracy and investment return to evaluate the efficacy of these three classification models. Empirical experimentation shown hybrid classification model help construct a better predictive power trading system in terms of stock market timing analysis.

1. Introduction

In the initial stages of Taiwan securities market, there are only a limited number of listed companies in the issue market each year. And since securities transaction is in bad conditions, the desire for companies to finance through the capital market is not high. However, with the fast growth of the industry and commerce and the booming development of the whole economy both based on strong knowledge and capital intensity in Taiwan, the desire for domestic companies to finance through the capital market has been greatly strengthened. After a big sweeping revision of the law of Securities and Exchange in 1989, the revised law has relaxed restrictions of securities-dealer business and allowed them to transact business over the counter and finance through the capital market, which results in the fast growth of the numbers of securities dealers as well as their overall scale. The market has become more effective after new participants join in it, and the financing costs of companies have gradually reduced which results from the competition. Besides, in the course of pre-underwriting counselling and the underwriting business, the companies can gain better financial consultancy services than ever. By the end of December 2007, in securities market the listed companies have amounted to 698, over-the-counter companies 547, and the securities floatation has been on a large scale.

Many researchers have predicted price movements in the stock market using several approaches during the past decades. Use the time series analysis techniques to address stock price prediction (Cheng, Chen, Teoh, & Chiang, 2008; Kendall & Ord, 1990). Select artificial neural networks (ANNs) to stock market prediction (Choi, Lee, & Rhee, 1995; Trippi & Desieno, 1992) or a novel Adaptive resonance theory-Counterpropagation neural network for solving a forecasting problem (Liu & Li, 2005). Use genetic algorithms to choose optimal portfolio (Bauer, 1994) and predicting the S&P 100 index using rough sets (Skalko, 1996). Recently, research in this area tends to hybridize several artificial intelligence techniques (Tsaih, Hsu, & Lai, 1998; Tae, 2007).

Although there exists a large amount of paper addressing the predictability of stock market returns, most of the proposed models rely on accurate forecast level of the underlying stock index or its return. Forecasts methods based on minimizing the forecast error may not be adequate for matching researchers' objectives. However, detecting market timing seems to be more important. To a certain extent when to buy and sell stocks are more important than predicting daily price movement. Some recent studies have showed that trading strategies guided by forecasts on the direction of price change may be more effective and have higher profits. Aggrwal and Demaskey (1997) showed that the performance of cross-heading could be improved significantly if the direction of changes in exchange rate is considered. Maberly (1986), for this reason, explored the relationship between the direction of interday and intraday price change based on the S&P 500 futures.

In this paper, the problems which concern detecting stock market timing of the Taiwan Stock Exchange Index using a rough sets approach discussed, such as converting time series to rough sets objects, discretization and rough sets model modification. Rough set theory is quite valued for extracting trading rules. Compared

* Corresponding author. Tel.: +886 5 5342 601x5324; fax: +886 5 5312 077.
E-mail addresses: jhcheng@yuntech.edu.tw, g9623802@yuntech.edu.tw (J.-H. Cheng), yimmin.lin@msa.hinet.net (Y.-M. Lin).

0957-4174/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2009.07.019

to other methods used in financial area, multivariate statistical methods for example, the rough sets method has several advantages. First, it not only handles noise well but also eliminates irrelevant factors, thus it can generate profitable rules of market timing (Ruggiero, 1997). Second, by the means of using the rough sets, we can appropriately detect market timing when the market pattern is uncertain to generate a signal for trade. In addition, it does not make any assumption about the distribution of data. Also it is a tool suitable for analyzing quantitative and qualitative attributes.

For the monthly Taiwan Stock Exchange Index, we consider three models that are evaluated according to their out-of-sample forecasting robustness. Our forecast evaluation is based on the predictive capability of market timing and trading simulation performance. In order to predict the Taiwan Stock Index, our first model use probabilistic neural network (PNN), our second model use rough set theory, and the third model combines PNN, rough sets and C4.5 decision tree and blends them into a new model. C4.5 decision tree is proposed by Quinlan (1993) and extended the original ID3 algorithm by accounting for unknown values, continuous attribute value range, tree pruning, and rule deviation (Quinlan, 1993). The main objective of hybrid model is to construct a better predictive capability model.

The paper is organized as follows. In Section 2, the background of three theories that probabilistic neural network (PNN), decision tree_C4.5, rough set theory will be discussed. In Section 3, input variables in the forecasting model are discussed and the forecasting models built on Taiwan Stock Index are presented and addressed. We then explain our results in Section 4 and they are included in Section 5.

2. Literature review

There are a lot of tool to use in classification for prediction. Neural networks and decision trees are competitive techniques which are considered to be the most efficient tools in many pattern classification applications (Tsujino & Nishida, 1995; Lee & Oh, 1996). Neural networks are generally preferred for their generalization ability (Hansen & Salamon, 1990; Zhou, Wu, & Tang, 2002). However, decision trees obviously outperform neural networks in terms of interpretability (Zhou & Jiang, 2004). Prior researches use the rough set theory to solve the data which is unclear relation and uncertain or inconsistent (Ahn, Cho, & Kim, 2000; Beynon & Dri1eld, 2005).

2.1. Probabilistic neural network (PNN)

The power of neural network (NN) is come to life when a pattern that had no output associated with it was given as an input. Specht (1988) designs a very efficiency probabilistic neural network (PNN) that adapted well to manipulate classification problem. The probabilistic neural network (PNN) is a special type of neural network using a kernel-based approximation to form an estimate of the probability density function of categories in a classification problem. This particular type of ANN provides a general solution to pattern classification problems by following the probabilistic approach based on the Bayes decision theory. The experiment reveals it was excellent in efficiency and performance.

PNN uses a supervised training set to develop probability density functions within a pattern layer. The main advantages of PNN are the fast training process, an inherently parallel structure guaranteed to converge to an optimal classifier as the size of the representative training set increases and that training samples can be added or removed without extensive retraining (Gorunescu, Gorunescu, El-Darzi, Gorunescu, & Revett, 2005).

On the classification problem, probabilistic neural network had been used widely in all kinds of fields. In the medical science aspect, Gorunescu et al. (2005) had applied probabilistic neural network in the diagnosis of cancer. About civil engineering, Ni, Wang, and Ko (2000) had applied PNN to identify the damage type and location in the cable-stayed Ting Kau Bridge from the simulated noisy modal data. And in e-commerce Web pages, Anagnostopoulos, Anagnostopoulos, Loumos, and Kayafas (2004) had used probabilistic neural network estimating the population of specific e-commerce Web pages. Potential applications involved surveying Web activity in commercial servers, as well as Web page classification in largely expanding information areas like e-government or news and media. Selekwa, Kwigizile, and Mussa (2005) presented a systematic procedure for setting up a probabilistic neural network that can classify the globally nonseparable population of highway vehicles.

2.2. Decision tree

Decision trees are also fast and easy to use. Although their limitations are sometimes criticized (Dombi & Zsiros, 2005; Duch, Setiono, & Zurada, 2004), the rules generated by decision trees are simple and accurate (Duch et al., 2004) for most problems. Therefore, decision trees are very popular and powerful tools in data mining (Duch et al., 2004; Tsujino & Nishida, 1995) and several significant works confirm their efficiency (Tsujino & Nishida, 1995; Lee & Oh, 1996). Indeed, in many applications, the structural description of the knowledge is as important as the ability to perform well on new examples (Fiordaliso, 2000; Andrews, Diederich, & Tickle, 1995). The fact that decision trees are efficient in terms of performance and easily interpretable is a major argument for their use in our problem. Each node in a decision tree contains a question relative to a particular attribute. Leaf nodes are groups of instances that receive the same class label. An unknown (or test) instance is routed down the tree according to the values of the attributes in the successive nodes. When the instance reaches a leaf, it is classified according to the label assigned to the corresponded leaf. Leaves have to be homogeneous as much as possible.

2.2.1. Decision trees induction: C4.5 algorithm

Many algorithms for decision tree induction exist. ID3 and C4.5 (Quinlan, 1986, 1993) are the most widely used with the CART algorithm (Breiman, Friedman, Olshen, & Stone, 1984). The CART (classification and regression tree) algorithm is suitable for problems with numerical classes, thus, it is not appropriate in our context for which the classes are the label of the forecasting prototypes. We chose the C4.5 algorithm because of its predictive accuracy (Last & Maimon, 2004) and the numerous possibilities in term of pruning, treatment of numerical and nominal attributes.

C4.5 algorithm is an extension of ID3 (interactive dichotomizer) algorithm and the divide-and-conquer approach (Quinlan, 1993; Winston, 1992) which main improvements included the pruning methodology and the processing of numeric attributes, missing values and noisy data. The splitting node strategy is based on the computation of the information gain ratio. The basic idea is that each node should hold a question concerning the attribute which is the most informative amongst the set of attributes not yet considered in the path from the root to that node. Information value also called entropy measures how informative is the association of an attribute with a node (Gray, 1990). The notion of gain ratio (Quinlan, 1993) is useful to rank attributes. The decision tree induction purpose is the classical overfitting problem can be addressed via pruning strategies. There are two strategies can be adapted to prune a decision tree: pre-pruning or post-pruning. Pre-pruning involves trying to decide when to stop developing sub-trees or branches during the tree building. In general, the stop criterion associated with the Information Theory relies on the Minimum Descriptive Length principle (Fayyad & Irani, 1993; Rissanen, 1985). Post-pruning consists in building the complete tree and pruning it afterward. In general, the second method seems to be more attractive since it enables to build very informative sub-trees composed of attributes which are individually low informative (Quinlan, 1996). Two rather different operations have been considered for post-pruning: sub-tree replacement and sub-tree raise. At each node, the learning scheme might decide whether to perform sub-tree replacement, sub-tree raising or leave the sub-tree unpruned.

2.3. Rough set theory (RST)

The concept of a rough set, introduced by Pawlak (1982), proved to be an effective tool for the analysis of information tables (financial information tables) describing a set of objects (firms) by a set of multi-valued attributes (financial ratios).

Rough set theory applies the unclear relation and data pattern comparison based on the concept of an information system with indiscernible data, where the data is uncertain or inconsistent. The data is grouped into classes called elementary sets. Feature/attribute selection is crucial in data processing that consists of relevant (or maybe irrelevant) object patterns, but it may be redundant in data pattern recognition. More detailed information regarding attributes can be found in the works of Swiniarski and Skowron (2003), Polkowski (2002), and Inuiguchi and Tanino (2004). The objects in a class may have a relationship with the corresponding features/attributes, and expert knowledge is used to process attribute extraction. Each elementary set is independent of the others. We can extract knowledge from each elementary set used in the real-world.

Rough set theory is a mathematical approach to managing vague and uncertain data or problems related to information systems, indiscernibility relations and classification, attribute dependence and approximation accuracy, reduct and core attribute sets, and decision rules. In the insurance marketing, Shyng, Wang, Tzeng, and Wu (2007) had applied rough set theory (RST) to demonstrate that the redefined combination values of attributes can contribute to the precision of decisions in insurance marketing. About discuss the business failure prediction, (Ahn et al., 2000) had used a hybrid intelligent system that predicts the failure of firms based on the past financial performance data, combining rough set approach and neural network. And discussion of customer complains, Yang, Liu, and Lin (2007) had applied rough set theory to discover important attributes leading to complaints and induce decision rules based on the data of a Taiwanese IC packaging foundry that ranks one of the largest in the world.

3. Research method

This section firstly establishes each variable definition describes research samples, and data resources then compare the research methods which include PNN classifier, rough set classifier and the hybrid classifier.

3.1. Input variables in the forecasting model

3.1.1. Definition of variables

In general, techniques used in economic forecasting are the fundamental and technical analysis. The fundamental indicators are used for long-term analysis while technical indicators are used for short-term analysis. Fundamental analysis involves predicting stock price from other factors, such as money supply (M1B), government consumption level (GC), gross national products (GNP), gross domestic products (GDP), consumer price index (CPI), wholesale products index (WPI), and rate of exchange (RE). Studies related to predictability of stock market timing and its determinant factors are plentiful and easy to find in the literature. Hence, the discussion of economic rationale has not been addressed in detail in this study.

Technical analysis mainly involves several market indicators. For instance, there is moving average convergence/divergence (MACD), price rate of change (ROC), stochastic %K, stochastic %D, relative strength index (RSI), stochastic oscillator, and directional indicator (DI). These technical indicators scrutinize the trend of price indices and individual securities. The theory underlying these indicators is that once a trend is in motion, it will continue in that direction (Achelis, 1995). Technical analysis attempts to determine the strength and direction of the trend. However, each technical indicator has its character and applying limitation. Almost no indicator can be absolutely suitable for the various situations in the stock market. This study extracts the original component, such as opening price, closing price, the highest price, the lowest price, volume, and turnover rate, in each technical indicator.

The lagged macroeconomic variables M1B, GC, GNP, GDP, CPI, WPI, RE, and six original technical indicators which are used by the forecasting models analyzed in this study. In addition, this study will define the decision table that indicates future direction of the research data set. The decision attribute is constructed as follows:

$$Mr(j) = \text{sign}\{[(P(t+j) - P(t))/P(t)] * 100 - TB(j)\},$$
$$j = 1, 3, 6, 12 \text{ months}$$

where $P(t)$ is the closing price of the stock index traded at period t and $TB(j)$ is the Taiwan short-term risk-free interest rate which j months deposit rate at the First Commercial Bank in Taiwan. In this study, the $Mr(j)$ specifies three values, +1, 0, and −1. A value +1 indicates the market closed higher in $t+j$ of the future than in this period t. Similarly, a value −1 indicates the market closed lower $t+j$ of the future than in this period t. Also, if the value of $Mr(j)$ is +1 then generate a buy signal, 0 to hold and −1 to sell. Table 1 provides a brief description of the macroeconomic variables, original technical indicators and forecasted output variables.

3.1.2. Data description

The monthly Taiwan weighted stock index data using in this study was collected from the Taiwan Economic Journal Database (TEJ), developed by Taiwan Economic Journal. TEJ found in April, 1990, Taipei, Taiwan, it is to provide the data and information about financial market, except the real-time information. The total number of samples was 208 trading months, from January 1988 to April 2005. The data set distinguish between estimation (training) period and ex post (holdout) forecasting. In an estimation (training) data observations from January 1988 to December 1999, a total have 144 trading months, been used for rule-construction with rough sets. In an ex post (holdout) forecast observations from January 2000 to April 2005, a total 64 trading months, provide a mean of evaluating forecasting performance.

3.2. Forecasting models used for comparison

The propose here is to run the PNN classifier model, rough set classifier model, and the hybrid classifier model then compare with them.

3.2.1. Probabilistic neural network (PNN) for classification

PNN is a classifier conceptually built on Bayesian method. Our trading model based on PNN is used to predict the sign of index return with respect to a set of input variables. Using PNN is proved

Table 1
Variables included in analysis.

Input variables	Description
RE1, RE3	The price of one country's currency expressed in another country's currency. In other words, the rate at which one currency can be exchanged for another
M1B1	Money supply means the entire quantity of a country's bills, coins, loans, credit, and other liquid instruments in the economy. In Taiwan, money supply is divided into three categories, M1A, M1B, and M2, according to the type and size of account the instrument is kept in
	Monetary aggregate M1B = currency held by the public + deposit money, or M1B = M1A + passbook savings deposits of individuals (includes non-profit organizations) in Monetary Institutions
	Continuously compounded lagged 1 month of the M1B to the period being forecasted
GC3, GC6, GC12	Government consumption level
	Continuously compounded lagged 3, 6, and 12 months annualized growth rates of government consumption to the period being forecasted
GNP3, GNP6, GNP12	The gross national product (GNP) is the total dollar value of all final goods and services produced for consumption in society during a particular time period
	Continuously compounded: lagged 3, 6, and 12 months annualized growth rates of the gross national product and gross domestic product previous to the period being forecasted
GDP3, GDP6, GDP12	The gross domestic product (GDP) measures output generated through production by labor and property which is physically located within the confines of a country
	Continuously compounded: lagged 3, 6, and 12 months annualized growth rates of the gross national product and gross domestic product previous to the period being forecasted
CPI3	An index of prices of goods and services typically purchased by urban consumers
	Continuously compounded lagged 3 month annualized growth rates of the consumer price index to the period being forecasted
WPI1	An index of the prices paid by retail stores for the products they ultimately resell to consumers
	Continuously compounded lagged 1 month of the WPI to the period being forecasted
Opening index	The range of index at which the first bids and offers were made or first transactions were completed
Closing index	Closing index generally refers to the last price at which a stock trades during a regular trading session
Highest index	The highest index at which the stock has traded during a month
Lowest index	The lowest index at which the stock has traded during a month
Volume	The number of shares or contracts traded in a security or an entire market during a given period
Turnover ratio	The percentage of a mutual fund's holdings that have been sold over the past year. A high turnover rate, also called turnover ratio, means the fund has sold a high percentage of its investments
Output variables	Description
$Mr(1), Mr(3)$ $Mr(6), Mr(12)$	The probabilities estimated that direction of future (1, 3, 6 and 12 months) excess return in period t by the forecasting models

to be promising in financial forecasting (e.g., Kim, 1998; Yang, 1999).

> *By using PNN, when an input is presented, the first layer computes distances from the input vector to the training input vectors, and produces a vector whose elements indicate how close the input is to a training input. The second layer sums these contributions for each class of inputs to produce as its net output a vector of probabilities. Finally, a competing transfer function on the output of the second layer picks the maximum of these probabilities, and produces a 1 for that class and a 0 for the other classes (*Demuth & Beale, 2000*).*

Our PNN training and validation involves dividing the first period data into two segments. The first segment is used to train the network whereas the second segment is used to validate the model specification. In our PNN model of using the complete set of 20 attributes as input (i.e. explained in Section 3.1). The output of PNN indicates the Bayesian probability of class affiliation and thus the direction of index return.

3.2.2. Rough set for classification

Our trading model based on rough sets comprises three main steps, namely pre-processing, rough sets analyzer, and trading decision rule extraction.

In the pre-processing step, the data availability and data reliability should be considered. Then these data are pre-processed to construct the information table, which is the raw knowledge or data gleaned from the Taiwan stock market is stored and subsequently forwarded to the rough sets model for feature extraction. Here pre-processing includes the converting time series to rough sets objects, rescaling the attributes for nominal values, removing the outliers, missing values, and discretization of the continuous attributes.

When applying rough sets, these data gleaned from the Taiwan stock market, which cannot be put into directly. Baltzersen (1996) recommend two methods. One is the mobile window method which moves a window along the time series, the data points falling into the window are transferred to a rough sets object. The other is columnizing method which time series are organized in columns, such that each row represents a different point in time and each column is an economic indicator. The following experiment will adopt the columnizing method.

Although at discretizaion step, there are a lot of supervised discretization and unsupervised automatic discretization methods (Dougherty, Kohavi, & Sahami, 1995; Nguyen & Nguyen, 1998). However, definitely the experience of experts can give more reasonable cut points than the automatic discretization method. But sometimes due to a lack of experts' supervisions, we had to look for the help of automatic discretization methods. The individuality converts each continuous attributes will be called local method is chosen as our discretization tool.

In the rough sets analyzer step, it carries out three sub-tasks, namely called consistency check, concept forming, and approximation. It scans the data obtained from the pre-processing step and checks its consistency. Once an inconsistency is spotted, the concept partitioner and the approximation operator will be activated to use rough sets theory to carry out the analysis. The concept partitioner performs set operations for each concept. The approximation operator employs the variable precision rough sets (VPRS) model to seek subset of the attributes. These subsets of attributes are termed -reducts or approximate reducts. Our study uses the VPRS model because real-world data is always filthy with noise and defects. The is set to $[0.5, 1]$ in our study.

By using the ROSE 2 system, an microcomputer software designed to analyze data by means of the rough sets theory, suggest a -reducts for $Mr(1)$ which includes opening index, closing index, volume, turnover ratio, lowest index, WPI1, M1B1, RE1, GC3, GC12, GDP3, and GNP3 as the explanatory input variables. The $Mr(3)$ -reducts includes highest index, closing index, volume, turnover ratio, WPI1, M1B1, RE1, RE3, GC6, GC12, GNP3 and GNP6 as the explanatory input variables. The $Mr(6)$ -reducts includes highest index, closing index, lowest index, volume, GC6, GC12 and GNP6 as the explanatory state variables. Finally, the $Mr(12)$ -reducts includes highest index, closing index, lowest index, CPI3, GC6 and GC12 as the explanatory state variables. These -reducts are showed in Table 2.

The trading decision rule extraction is applying rough sets to generate decision rule and check the predictive accuracy on the testing data sets. A set of strong decision trading rules are extracted from the core of attributes or the largest strength reducts . The largest strength means that many objects of historical data support the extracted rules. Our study uses a modified version of the LEM2 algorithm to execute decision rules generation. The LEM2 method is the generation of a minimal set of rules covering all objects.

3.2.3. A Hybrid for classification

No one classification technique can always provide the best result for a classification problem. In this study, we apply the bagging technique to create an improved composite classifier. The bagging technique combines a series of learned classifiers and makes them into a new approach.

Owing to the PNN classifier is proved to be a useful tool for predicting the market time in Taiwan Stock Index (Chen, Leung, & Hazem, 2003). The classifier based on rough sets approach is the focus of this study. C4.5 decision tree is often treated as the benchmark for predicting classification problem, and have the advantages of dealing with missing data, continuous data, pruning, and generate rule from the decision tree which is built. Therefore, the combination of PNN, rough sets and C4.5 decision tree is very natural for yields the better predictive capability. And the combination is called hybrid classifier. Our hybrid classifier comprises two major phases, namely combine vote and trading market timing prediction.

In the combine vote phase, three learned classifiers are employed, including the PNN model, rough sets model, and C4.5 decision tree. We use rough sets approach as the pre-processing tool for the classifier of PNN and C4.5, and the approach is to filter irrelevant and redundant attributes from the information table. The PNN, rough sets and C4.5 classifiers yields its own prediction of the direction of excess return from the same training data, in Taiwan Stock Index from January 1988 to December 1999 of monthly trading.

In the trading market timing prediction phase, a voting strategy is applied to decide the trading market timing prediction for a given unknown sample. To classify an unknown sample, the PNN, rough sets and C4.5 classifier returns it's the trading market timing prediction, which counts as one vote. The phase counts the votes positive, but actually it's negative. Type4 error (type 4) is the stock excess return predicted to be negative, but actually it's positive. In this study, higher ratio is better, lower type2 and 3 are better, and higher relative to types 2 and 3 is tolerant because all the investors' evaluation guideline is the least loss.

Table 3 provides the forecasting result from PNN classifier, rough sets classifier and hybrid classifier at accuracy, type1, type2, type3, and type4. The correct predicting rate of the PNN classifier, $Mr(1),\ Mr(3),\ Mr(6)$, and $Mr(12)$ was 0.4688, 0.7344, 0.7344, and 0.7188. The correct predicting rate of the rough sets classifier, $Mr(1),\ Mr(3),\ Mr(6)$, and $Mr(12)$ was 0.6094, 0.5781, 0.6406, and 0.7500. The correct predicting rate of the hybrid classifier, $Mr(1),\ Mr(3),\ Mr(6)$, and $Mr(12)$ was 0.5781, 0.7500, 0.7344, and 0. 7656.

Also, Table 3 provides the column of the No. The No. is the numbers of times to correctly predict whether the market time is going to buy or sell. The hybrid and PNN classifier is able to correctly predict the directions of excess returns more than 50% times at the 10% level of statistical significance beside $Mr(1)$. The rough sets classifier is able to correctly predict the directions of excess returns more than 50% times at the 5% level of statistical significance beside $Mr(3)$. The PNN classifier performs a little better as the hybrid besides $Mr(1)$, but rough sets classifier does not perform well besides $Mr(12)$.

From the view points of accuracy index and type1, type2, type3, and type4, the hybrid classifier has better predictive capability. The hybrid classifier employing a composite classifier of the PNN, rough sets and C4.5 decision tree shows an improved classification performance.

In addition, this study basing on PNN, rough sets and hybrid models performs trading simulation to forecast the excess return percent, which evaluates the stock investing performance.

In this study, the basic assumptions of trading simulation strategy are as follows:

1. The testing period runs from January 2000 to April 2005 which has 64 trading months for validating the performance.
2. 1, 3, 6, and 12 months investment horizons are implemented by the forecasting of buy and hold, PNN, rough sets, and hybrid models, respectively.
3. When a transaction takes place every time, the total investing money in the stock market is constant.
4. When a transaction takes place every time, the stock buying index is the closing index of the last test period.
5. When a transaction takes place every time, the stock selling index is the closing index of the right test period.
6. When evaluating investment performance of 1, 3, 6, and 12 months horizons, consider the transaction cost and non-transaction cost. If it has the transaction cost, consider commissions as 1.425% when buying. Consider commissions as 1.425% and stock tax as 3% when selling.
7. The strategy of trading simulation is based on the classifier which forecasts the direction of future (1, 3, 6 and 12) excess return to buy the stock index, or allocate the assets to the certificate of deposit in 1, 3, 6, and 12 months. On the first day, invest the constant assets in the stock index, and sell out on the last day when the classifier predicting signal is positive in the investment horizon for each test period. Otherwise, invest the constant assets in risk-free time deposit when the predicting is negative. When the predicting signal is not judged, keep his asset allocation decision the same as the last test period.

Table 4 provides the simulation performance result to the transaction cost and non-transaction cost for testing period from January 2000 to April 2005 in the Taiwan Stock Index by buy and hold, PNN, rough sets, and hybrid strategy. From the view point of the investing performance, when the predicting time is less than or just three months, the hybrid strategy performs best, Rough sets strategy is better, and PNN strategy is worse. When the predicting time is over three months, PNN strategy is a good choice. However buy and hold strategy is negative of stock index excess return in 1, 3, 6, and 12 months investment horizons.

5. Conclusion and limitation

Recently, scholars are frequently using artificial neural networks and fuzzy theory to study the issues relevant to stock, especially in stock market timing and predicting the trend of stock price. Though the ways used vary in different studies, their findings in general cannot show a better predicting ability. This study primarily uses Rough sets and integrates fundamental analysis and technical analysis to build up the trading model of stock market timing. Through the detailed analysis in Taiwan Stock Index, how to convert time series data to rough sets object, predicting variables selection, removal of uncertainly from information table, and trading rules have been presented and discussed in this study.

The good performance of the trading model built by rough sets suggests that the usefulness of rough sets for predicting the direction of Taiwan Stock Index returns. This can be seen from forecasting market timing and simulation performance compared with PNN trading model.

The results of Tables 3 and 4 both show that the rough sets method determines group of input variables, C4.5 filters noisy attributes from the above input variables, and also uses PNN, rough sets and C4.5 classifiers to generate trading rule sets, which is helpful to construct a better predictive power trading system for stock market timing analysis.

Acknowledgements

This research was supported by the National Science Council of Taiwan, ROC, under Contract NSC 98-2410-H-224-003. We thank all anonymous reviewers for their valuable comments and advice.

References

Achelis, S. B. (1995). *Technical analysis from A to Z*. Chicago, IL: Probus Publishing.

Aggrwal, R., & Demaskey, A. (1997). Using derivatives in major currencies for cross-hedging currency risks in Asian emerging markets. *Journal of Futures Markets, 17*, 781–796.

Ahn, B. A., Cho, S. S., & Kim, C. Y. (2000). The integrated methodology of rough set theory and artificial neural network for business failure prediction. *Expert Systems with Applications*(18), 65–74.

Anagnostopoulos, I., Anagnostopoulos, C., Loumos, V., & Kayafas, E. (2004). Classifying Web pages employing a probabilistic neural network. *IEE Proceedings-Software, 151*(3), 139–150.

Andrews, R., Diederich, J., & Tickle, A. B. (1995). A survey and critique of techniques for extracting rules from trained artificial neural networks. *Knowledge Based Systems, 8*, 373–389.

Baltzersen, J. K. (1996). *An attempt to predict stock market data: a rough sets approach*. Diploma thesis, knowledge systems group. Department of Computer Systems and Telematics. The Norwegian Institute of Technology, University of Trondheim.

Bauer, R. J. Jr., (1994). *Genetic algorithms and investment strategies*. New York: Wiley.

Beynon, M. J., & Dri1eld, N. (2005). An illustration of variable precision rough sets model: An analysis of the findings of the UK monopolies and mergers commission. *Computers and Operations Research, 32*, 1739–1759.

Breiman, L., Friedman, J., Olshen, R., & Stone, C. (1984). *Classification and Regression Trees*. Belmont, CA: Wadsworth.

Cheng, C. H., Chen, T. L., Teoh, H. J., & Chiang, C. H. (2008). Fuzzy time-series based on adaptive expectation model for TAIEX forecasting. *Expert Systems with Applications, 34*(2), 1126–1132.

Chen, A. S., Leung, M. T., & Hazem, D. (2003). Application of neural networks to an emerging financial market: Forecasting and trading the Taiwan Stock Index. *Computer and Operations Research, 30*, 901–923.

Choi, J. H., Lee, M. K. & Rhee, M. W. (1995). Trading S&P 500 stock index futures using a neural network. In *Proceeding of the third annual international conference on artificial intelligence applications on wall street* (pp. 63–72). New York: Software Engineering Press.

Demuth, H. & Beale, M. (2000). *Neural network toolbox user's guide*, The Mathworks, Natick.

Dombi, J., & Zsiros, A. (2005). Learning multicriteria classification models from examples: Decision rules in continuous space. *European Journal of Operational Research, 160*(3), 663–675.

Dougherty, J., Kohavi, R., & Sahami, M. (1995). Supervised and unsupervised discretization of continuous features. In: *Proceedings of the 12th international conference of machine learning, Morgan Kaufmann, San Francisco, CA* (pp. 194–202).

Duch, W., Setiono, R., & Zurada, J. M. (2004). Computational intelligence methods for rule-based data understanding. *Proceedings of the IEEE, 92*(5), 771–805.

Fayyad, U. M., & Irani, K. B. (1993). Multi-interval discretization of continuous valued attributes for classification learning. In: *Proceeding of the 13th international joint conference on artificial intelligence, Chambery, France, Morgan Kauffmann, San Francisco*.

Fiordaliso, A. (2000). Autostructuration of fuzzy systems by rules sensitivity analysis. *International Journal for Fuzzy Sets and Systems, 118*(2), 281–296.

Gorunescu, F., Gorunescu, M., El-Darzi, E., Gorunescu, S., & Revett K. (2005). A cancer diagnosis system based on rough sets and probabilistic neural networks. In: *Proceedings of the first european conference on health care modelling and computation, University of Medicine and Pharmacy of Craiova* (pp. 149–159).

Gray, R. M. (1990). *Entropy and information theory*. New York: Springer-Verlag.

Hansen, L. K., & Salamon, P. (1990). Neural network ensembles. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 12*(10), 993–1001.

Inuiguchi, M., & Tanino, I. (2004). Fuzzy linear programming with interactive uncertain parameters. *Reliable Comput, 10*(5), 357–367.

Kendall, S. M., & Ord, J. K. (1990). *Time series* (3rd ed.). Sevenoaks: Edward Arnold.

Kim, S. (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.

Last, M., & Maimon, O. (2004). A compact and accurate model for classification. *IEEE Transactions on Knowledge and Data Engineering, 16*(2), 203–215.

Lee, K. C., & Oh, S. B. (1996). An intelligent approach to time series identification by a neural network-driven decision tree classifier. *Decision Support Systems, 17*, 183–197.

Liu, T. C., & Li, R. K. (2005). A new ART-counterpropagation neural network for solving a forecasting problem. *Expert Systems with Applications, 28*(1), 21–27.

Maberly, E. D. (1986). The informational content of the interday price change with respect to stock index futures. *Journal of Futures Markets, 6*, 785–795.

Nguyen, H. S., & Nguyen, S. H. (1998). Discretization methods in data mining. In L. Polkowski & A. Skowron (Eds.). *Rough sets in knowledge discovery* (Vol.1, pp. 451–482). Heidelberg: Physica-Verlag.

Ni, Y. Q., Wang, J. Y., & Ko, J. M. (2000). Modal interaction in cable-stayed Ting Kau Bridge. In: *Proceedings of the international conference on advances in structural dynamics, Hong Kong*.

Pawlak, Z. (1982). Rough sets. *International Journal of Computer and Information Sciences, 11*(5), 341–356.

Polkowski, L. (2002). *Rough sets: Mathematical foundations*. Heidelberg: Physica-Verlag.

Quinlan, J. R. (1986). Induction of decision trees. *Machine Learning, 1*, 81–106.

Quinlan, J. R. (1993). *C4.5: Programs for machine learning*. San Francisco: Morgan Kauffman.

Quinlan, J. R. (1996). Improved use of continuous attributes in C4.5. *Journal of Artificial Intelligence Research, 4*, 77–90.

Rissanen, J. (1985). The minimum description length principle. In S. Kotz & N. L. Johnson (Eds.). *Encyclopedia of statistical sciences* (Vol. 5, pp. 523–527). New York: John Whiley.

Ruggiero, M. A. (1997). *Cybernetics trading strategies: Developing profitable trading systems with state-of-the-art technologies*. New York: Wiley.

Selekwa, M. F., Kwigizile, V., & Mussa, R. N. (2005). Setting up a probabilistic neural network for classification of highway vehicles. *International Journal of Computational Intelligence and Applications, 5*(4), 411–423.

Shyng, J. Y., Wang, F. K., Tzeng, G. H., & Wu, K. S. (2007). Rough set theory in analyzing the attributes of combination values for the insurance market. *Expert Systems with Applications, 32*, 56–64.

Skalko, C. (1996). Rough sets help time the OEX. *Journal of Computational Intelligence in Finance, 4*(6), 20–27.

Specht, D. F. (1988). Probabilistic neural network for classification, map, or associative memory. In: *Proceeding of the IEEE International Conference on Neural Network* (Vol. 1, pp. 525–532).

Swiniarski, R. W., & Skowron, A. (2003). Rough set methods in feature selection and recognition. *Pattern Recognition Letters, 24*, 833–849.

Tae, H. R. (2007). Forecasting the volatility of stock price index. *Expert Systems with Applications, 33*(4), 916–922.

Trippi, R. R., & Desieno, D. (1992). Trading equity index futures with a neural network. *Journal of portfolio Management, 19*, 27–33.

Tsaih, R., Hsu, Y., & Lai, C. C. (1998). Forecasting S&P 500 stock index futures with a hybrid AI system. *Decision Support Systems, 23*(2), 161–174.

Tsujino, K., & Nishida, S. (1995). Implementation and refinement of decision trees using neural networks for hybrid knowledge acquisition. *Artificial Intelligence in Engineering, 9*, 265–275.

Winston, P. H. (1992). *Artificial intelligence* (3rd ed.). Boston: Addison-Wesley.

Yang, Z. (1999). Probabilistic neural networks in bankruptcy prediction. *Journal of Business Research, 44*, 67–74.

Yang, H. H., Liu, T. C., & Lin, Y. T. (2007). Applying rough sets to prevent customer complaints for IC packaging foundry. *Expert Systems with Applications, 32*, 151–156.

Zhou, Z. H., & Jiang, Y. (2004). NeC4.5: Neural ensemble based C4.5. *IEEE Transactions on Knowledge and Data Engineering, 16*(6), 770–773.

Zhou, Z. H., Wu, J., & Tang, W. (2002). Ensembling neural networks: Many could be better than all. *Artificial Intelligence, 137*(1–2), 239–263.