Applied Soft Computing 68 (2018) 756–764

Contents lists available at ScienceDirect

Applied Soft Computing

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

Using intelligent computing and data stream mining for behavioral finance associated with market profile and financial physics

Chien-Cheng Lin, Chun-Sheng Chen, An-Pin Chen*

Department of Information Management and Finance, National Chiao Tung University, 1001 University Road, Hsinchu 300, Taiwan, ROC

ARTICLE INFO

Article history:
Received 6 February 2017
Received in revised form 14 July 2017
Accepted 2 August 2017
Available online 14 August 2017

Keywords:
Market profile theory
Financial physics
Neural networks
Taiwan futures exchange
Trading analysis
Data stream mining

ABSTRACT

Day trading has become an important topic of discussion in the last decades, especially with regard to computer program trading or the increasing trend of high-frequency transactions. However, due to the high level of complexity regarding the forecasting of day trading trends, the use of traditional financial analysis or technical indicators for the forecasting of short-term market trends is often ineffective. The main reason is that in addition to the technical analysis of market physical trends, financial market trading behaviors are also often affected by psychological factors such as greed and fear, which are emotions displayed by investors during the transaction process. For this reason, this study will use the neural network to integrate into the financial engineering technology analysis of the physical momentum behavior and market profile theory to quantify controlled learning. The goal is to be able to provide an empirical explanation of the discoveries related to trading behaviors by using trading strategies. Our experiments showed that trading behaviors in the financial market could be explained by the physical trends of a quantitative and technical analysis of the market profile theory. It has also been proven that the financial trading market follows the existence of a certain trading logic.

1. Introduction

In the past, most studies on instant financial transaction behaviors had often focused on changes in technical parameters or types of changes in behavior or innovation behavior analysis. As for the research tools, the supervised learning for neural networks was often used for the performance studies of learning under single or multiple technical indicators [16–19]. This study, which is more concerned with financial engineering physics, suggests a dynamic inertia behavior tracking and supervised learning of the eigenvalues of technical indicators by using the idea of occurrence and change of inertia momentum during the changes in eigenvalues following a certain timeline. And this can be further Newton's first law of motion which, assuming the absence of any external forces, states that a body at rest will remain at rest" while "a body in motion will remain in motion." It also has something to do with the Taylor expansion, which says that the change in a result is often influenced by the first order, second order, and nth order influence of all relevant characteristic factors.

These two theories take into account the formation of the momentum generated by the inertia and changes in eigenvalues in the time axis; they also made a strong assumption about the inevitable inertia of the direction of stock price movements. Because the focus of neural networks learning in the past was on the discovery and consistency of eigenvalues, it failed to take into account the physical changes in these eigenvalues. In fact every change in the result has something to do with those physical changes. Under this concept, each of which we believe is a non-linear relationship, only using a simple Taylor series expansion does not lead to a linear expansion of the first-order and second-order variables. Therefore, we used all the dependent variables of the neural network to learn about the result changes. We refer to all of that collectively as financial engineering physics learning.

Secondly, intraday trading behaviors are often influenced by such human emotional feelings as greed and fear. These factors cannot be described by technical indicators. Greed and fear as critical elements of the market profile theory. That is, the psychological and mobile behaviors of suppliers (and other time frame traders). Under an initial aggressive situation, the market tends to form a band trend, going up or down. On the contrary, if it is a defensive (response) on the part of the trader, then the market tends to display a series of stride intervals. Because these eigenvalues are often presented in the form of physical forces, this pattern is diffi-

* Corresponding author.
E-mail addresses: a0926295383@yahoo.com.tw (C.-C. Lin), sugardba@gamil.com (C.-S. Chen), apc888888@gmail.com (A.-P. Chen).

https://doi.org/10.1016/j.asoc.2017.08.008
1568-4946/

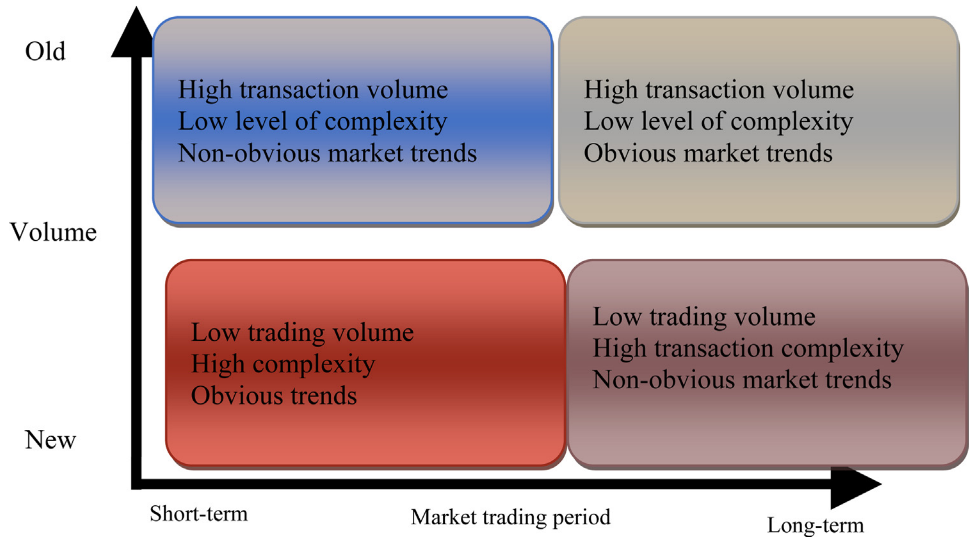

Fig. 1. Analysis of Transactional Roles and Trend Complexity.

cult to quantify. For this reason, this study suggests the use of the market profile theory for the trading market because during the process of obtaining the difference between the market price and value differences, a quantitative study is needed to understand the prices created by these psychological behaviors known as greed, fear, aggressiveness, and defensiveness.

Therefore, based on the above-mentioned argument needs, if the use of mathematical and statistical quantitative analysis is not easy, we hope that resorting to neural network learning will help find out the main characteristics of the market trend to determine whether the state of the movement is caused by result of suppliers' aggressive or defensive behaviors. If a clear distinction can be made, then the use of the neural network will make a great contribution to the future movement of the stock price. So we believe the fact that we can still highlight the trend directions of the market even in times of uncertain fluctuations is because these learning's help achieve better results compared to using the traditional methods. And the results of this study showed that after considering these two factors, the accuracy level for the short-term trend behavior forecasting was improved 20–30% compared to the traditional method.

Section 2 offers a review and summary of the literature, including the concepts of neural networks and market profile theory. Section 3 describes the research methods, sources, pretreatment, experimental model design, and performance evaluation. Section 4 compares the experimental results of different models and then analyzes their level of efficiency. Section 5 offers an analysis of the experimental results and goes on to offer a few suggestions for future research.

2. Literature review

2.1. Market profile

The Market Profile Theory was developed by Steidlmayer [23] as a mainstream market analysis that refutes both the Efficient-market Theory and the Random Walk Theory. Participants at different time intervals will always have different reactions to prices and values, which in turn will lead to the development of non-random price movements. Also, different participants have different opinions and behaviors with regard to the same price. As a result it impossible for the market to meet the individual needs of all participants at the same time. The Market Profile Theory claims that lengthening trading hours and enlarging trading volume can help provide a clear roadmap from which it is possible to see the entire trend of the market [27]. Using the trading role and complexity analysis chart, and taking into account the gradual change in trading time and volume, the market profile roadmap can be divided into four major parts. One of the most complex and difficult things to predict in day-trading is comprehending the trend of short-swing trading. To avoid the risks associated with overnight positions, we have deemed it important to conduct an in-depth analysis of day traders' trading patterns (Figs. 1–2).

By using a statistical bell curve, the Market Profile Theory found out that the majority of a day's total transactions occur in the middle of the curve, while the extremes of the curve in question account for only a small number of transactions that took place that same day. The widest part (the middle) of the bell curve represents the price region where the trader spent most of his time. It also represents the value range (Value = Price * Time) recognized by the trader. Market participants can be divided into the following two categories: (1) Day Traders (also called risk arbitrage): These include general market brokers. Their main objective is to engage in quick buy and sell transactions to obtain small profits. (2) Long-term traders (including other timeframe traders): These are participants whose transactions last more than one day. Long-term traders are the key players when it comes to controlling the market. They will usually enter the market only when the price deviates from the value zone, and respond rapidly to price changes. It is the sum of all the roles played in the market that represents the trend of the market transactions.

The basic framework of the Market Profile Theory – referred to as the TPO (Time Price Opportunities) – is expressed as an English letter every half hour on a trading day. If a price is traded at a certain time, the price is marked on the corresponding letters, each letter will be stacked up to form a complete market profile (Fig. 3).

In any given trading day, the market structure can lead to a certain pattern or form. And the importance of this form is decided by the behavior of the long-term participants. The correct analysis of this form can help understand the direction of the market trends and increase the chances of success in the market [28]. The importance of having different market patterns lies in the fact that they can lead to better judgment with regard to trend development. In this study, we expect to know more about pattern-related changes so we can use them to forecast the market trends.

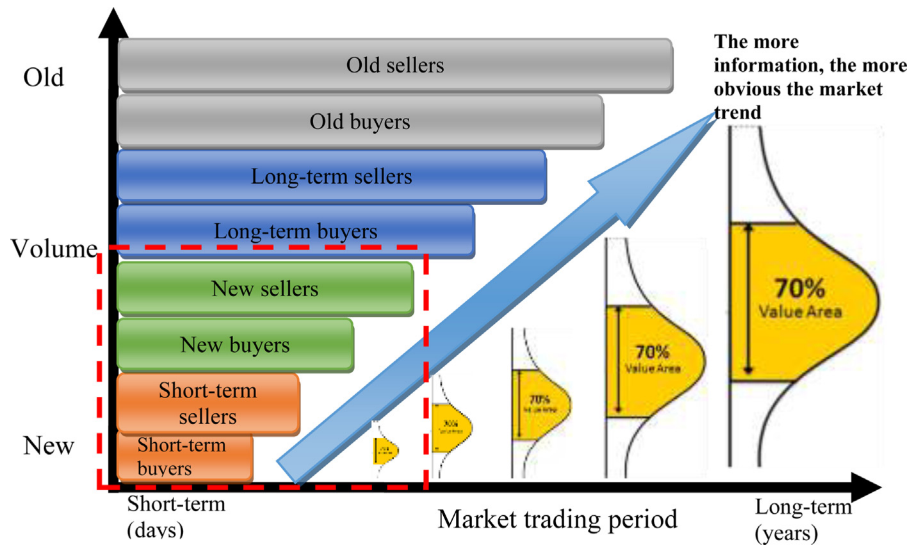

Fig. 2. Market Profile Trading Role Roadmap.

8271	K
8270	K
8269	JK
8268	JK
8267	IJK
8266	IJLM
8265	HJLM
8264	HJIJLM
8263	FGHIJLNO
8262	FGHJILNOPQRST
8261	EFGHJILNPQRSTUV
8260	ABCDEFGPQRSTU
8259	ABCDEFGPQRS
8258	ABCDEFGPQRS
8257	ABCDEF
8256	ABCD
8255	ABC
8254	BC
8253	C
8252	C
8251	C

Fig. 3. Form of the Market Profile.

2.2. Back-propagation neural networks

Proposed by McClelland, Rumelhart, Hinton and Williams in 1986, the Back-propagation Neural Network is a multi-layer supervised learning network that can adapt to new environments. Through the establishment of a problem-solving knowledge learning model based on self-learning, it has the ability to analyze, predict, and classify information. It can also repeat information in relation to past learning experiences or data. This process is similar to the human brain neural network learning process.

Compared with traditional statistical methods, the back-propagation neural network has a more powerful analytical ability and can be used to handle complex problems. It is also widely used in financial investment decision-making. In Japan, Kimoto and Asakawa [12] actually made use of the back-propagation neural network model, the multiple regression model, and the American Dow Jones index and applied them to the Nikkei Stock Market Index forecast. Their substitute variables were foreign exchange rates, interest rates, volume, turnover rates, etc. The results showed that the back-propagation neural network model is not only better than the multivariate regression model in terms of return rate; it also has a better ability to deal with nonlinear problems.

Harris [16] tested data of the New York Stock Exchange (NYSE) and found that there was, at the forty-fifth minute preceding the opening, a significantly exceptional remuneration together with the so-called U-Shaped pattern. In Japan, Kimoto and Asakawa [5] actually made use of the back-propagation neural network model, the multiple regression model, and the American Dow Jones index, foreign exchange rates, interest rates, trading volume, turnover rates and other indicators as input variables to analyze the Nikkei index. They used the TOPIX indices from 1987 to 1989 to predict the fluctuations and trading times of the next month. They also resorted to simultaneous use of mobile simulation methods for trading time simulation. The results showed that the back-propagation neural network model is not only better than the multivariate regression model in terms of return rate; it also has a better ability to deal with nonlinear problems.

Yoon & Swales [13] used the data of 98 companies selected by Fortune and Businessweek and analyzed nine of their variables, including confidence indices, economic factors, and growth rates. These variables were then used as the input variables for the back-propagation neural network model to forecast the performance of future stock prices and compare it with the empirical results of the statistical method known as MDA (Multiple Discriminant Analysis). The results showed that using the back-propagation neural network model is more accurate than using the Multiple Discriminant Analysis. Gencay [14] used the American Dow Jones index moving average as the input variables for the neural network. He especially used both the long-term moving average and the short-term moving average as a way to determine the trading signal. The results showed that the back-propagation neural network model has a more accurate predictive ability.

Paul [15] used the back-propagation neural network to predict the Straits Times Index (which is the most globally-recognised benchmark index and barometer for Singapore's stock market). The input variables were the daily opening price, the highest price, the lowest price, the closing price, the trading volume, the Dow Jones index, the NASDAQ index, the Hong Kong Hang Seng Index (HSI) and the Nikkei index. He then used the genetic algorithm to optimize the architecture and parameters of the neural network. The experimental results found that the predictive accuracy of the incoming neural network can be as high as 81% for 13–15 consecutive days. Applying it to Fortune and BusinessWeek, Yoon and Swales [6] managed to retrieve the data of 98 companies which they analyzed using more than ten variables, including confidence indices, economic factors, and growth rates. The results also showed that the back-propagation neural network model is better than the Multiple Discriminant Analysis. Barak and Modarres [7] proposed a composite architecture prediction model that is more accurate than a single model prediction. Chen [8] used an adaptive fuzzy neural network architecture to predict the trend of stock price changes and got a higher level of accuracy.

Hafezi [9] proposed a bat-neural network multi-agent system (BNNMAS), which predicts the stock price based on a fourth-order neural network structure that uses learning algorithm. Results of the experiment were very good. Liang [10] developed a composite system model and first used technical indices to find out the level of affiliation. Then, with the help of a stepwise regression and decision tree, he managed to improve the prediction accuracy of recurrent network (RNN). From the literature [1–5,24,25,26], it clearly appears that compared to traditional statistical methods, using the back-propagation neural network or deep neural networks not only leads to better accuracy; but it can also help solve complex nonlinear problems when it comes to predicting the complex financial market trends. Hence the rationale for adopting this model in this paper for the prediction of stock price trends.

3. Research methods and procedures

According to Chen [11], this study uses the Taiwan Capitalization Weighted Stock Index to explore day traders' greed and fear-related behaviors. Through quantification of traders' behaviors in the financial market, we used the market profile theory, financial physics, and the back-propagation neural networks model to assess and compare prediction performances and trading strategies. The goal was also to understand whether using the market profile theory with an empirical exploration of the factors that lead to greed and fear will result in better investment outcomes compared to random transactions.

Fig. 4 shows the experimental structure of this study. Experimental group 1 refers to the Taiwan Index Futures tick data. We used 1 min as the ordinal scale to find out the price within that timeframe. Then moving the window around, the data of every 25 min would come together and their combined shape would reflect the contours or profile of the market. The market profile price index was then extracted, and these raw values were later converted into financial physics data. Then, the first-order and second-order variables were normalized and computed, and finally the back-propagation neural network model was constructed.

Aside from using the same method, Experimental group 2 and experimental group 1 also included an analysis of the cumulative volume index. We used the Taiwan Futures tick data (with 1 min as a unit) to calculate the price level within that timeframe. We compared the control group model approach with the experimental groups to see whether trading signals were generated. As the control group is using a random trading strategy (a random change between 0 ∼1 happened), so the trading signals will set the probability of 0.5. That is, the probability of buying or selling is half of each. If the change was greater than 0.5 it was considered as a buying signal. If it was lower than 0.5, it was considered a selling signal. We used these signals to conduct real transactions. The experimental performance measurement was conducted based on two criteria, namely accuracy and average profitability (Fig. 4).

3.1. Calculating market profile input variables

Based on the Market Profile Theory, the price range of 70% of the control point is regarded as the value interval, and the price range and value range are calculated as follows:

(1) Control Point: It is the price with the most transactions in a working day.
(2) Value Range: in a given working day, it is the total turnover of the transaction that accounts for an interval of 70%.
(3) Price Range: in a trading day, it is the range or the difference between the upper price value and the lower price value.
(4) Tail: In a trading day, it is either the upper end or the lower end of a single TPO interval.

Each minute represents a TPO (Time Price Opportunity), and the result is a 25 min market profile that shows key information such as the POC price, a certain TPO volume, the Value Area, the Range Area, the Upper Tail, and the Lower Tail (Fig. 5).

The tick data of the Taiwan Weighted Index Futures is used to separate the data by moving the windows so as to make a clear distinction between the training group and the test group which are required for the back-propagation neural network. We first used the previous 25 min data (the training data) and then the 1 min data (the test data) to calculate the required information for every 25 min of the market profile.

3.2. First-order variables and second-order variables of the market profile index

In the Taylor (series) expansion equation, any change in the result is still influenced by the combined effects of factors such as the first-order variable, the second-order variable and Nth-order variable [29]. Due to the Taylor series is quite complex, according to the recommendations of Chen [6], the First-order and second-order Taylor series expansions are sufficient to represent the cause-and-effect relationships. Therefore, this study uses these two variables as experimental parameter. In order to increase the reliability of neural network learning, there is a need to go beyond learning about the technical indices and take into account the changes in the behaviors of the physical indices. The equations for the Taylor series first-order variables and second-order variables are expressed as follows:

$$f(\chi) = f(\chi_0) + \sum_{i=1}^{n} \frac{\partial f(\chi_0)}{\partial \chi_i}(\chi - \chi_{i0}) \tag{1}$$

$$f(\chi) = f(\chi_0) + \sum_{i=1}^{n} \frac{\partial f(\chi_0)}{\partial \chi_i}(\chi - \chi_{i0}) + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial^2 f(\chi_0)}{\partial \chi_i \partial \chi_j} \times (\chi - \chi_{i0})(\chi - \chi_{j0}) \tag{2}$$

In this study where time and efficiency computations are taken into account, the input of the back-propagation neural network will also focus on the relative first-order variables and the second-order variables of the input parameters. The goal is to better compre-

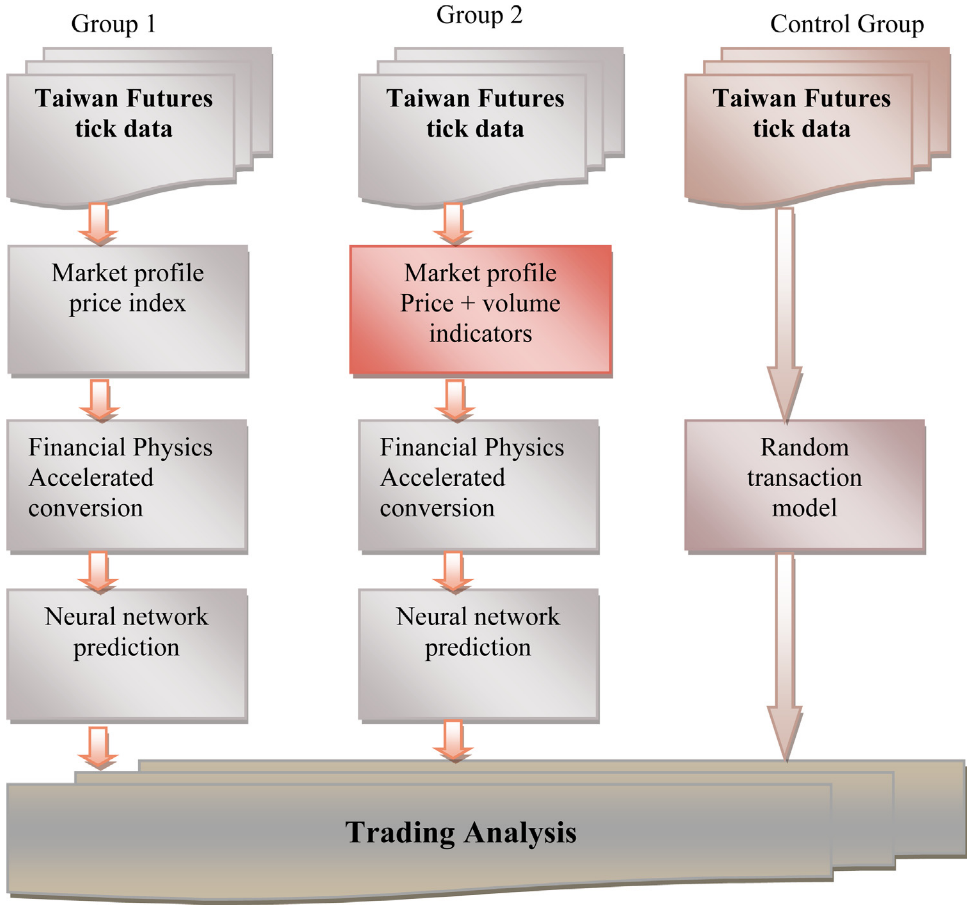

Fig. 4. Experimental Structure.

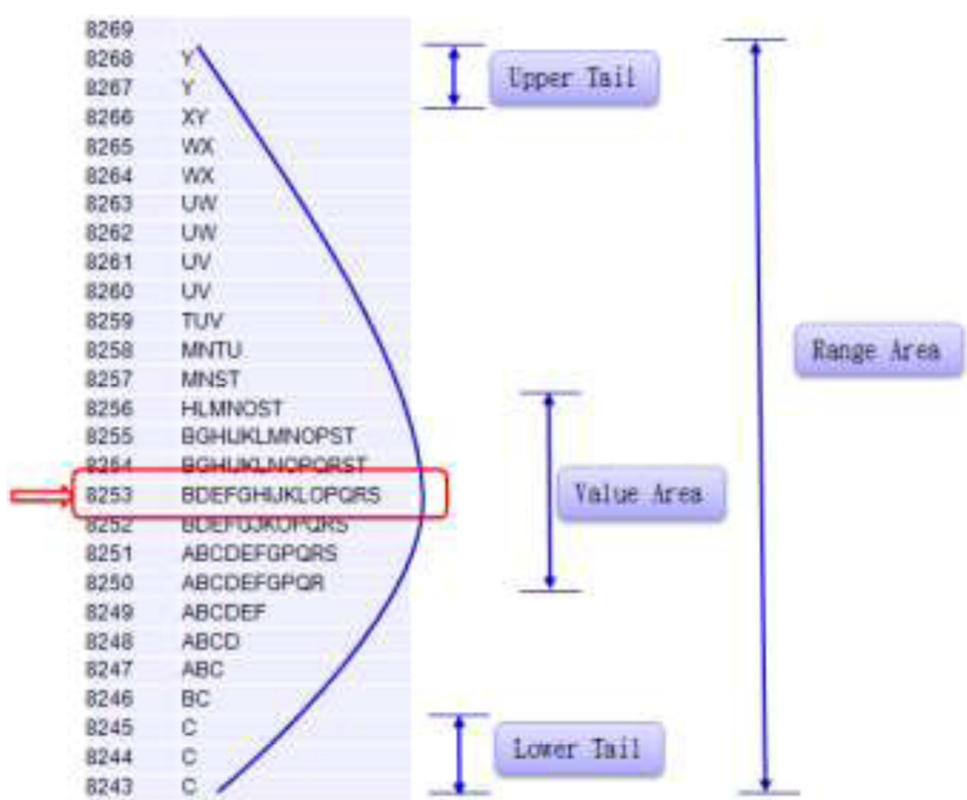

Fig. 5. Market Profile Map Partition.

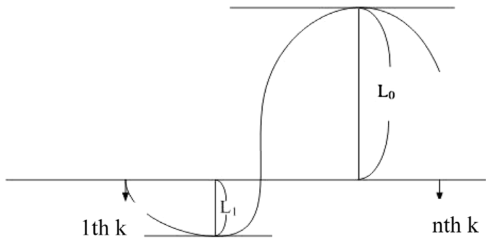

Fig. 6. Relative fluctuations in stock prices.

hend the trends of psychological behaviors that arouse as a result of trading in the stock market.

Therefore, by adding up all input variables and the values of first-order and second-order variables, the performance indicator appeared each time there was a change in trend during the calculation period. The computational formula is as follows:

$$\Delta X_t = \frac{X_t - X_{t-1}}{X_{t-1}} \quad (3)$$

ΔX_t: First-order variable value of input variables at time tX_t: Input variables at time tX_{t-1}: Value of input variables at time $t-1$

The first-order variables indicate the trend of the physical changes in each index. And as the first-order variables go through the process of the second-order (i.e., the difference of trend change) the trend of a continuous kinetic energy change is shown. The equation formula for the second-order variable is as follows:

$$\Delta = \Delta X_t - \Delta X_{t-1} \quad (4)$$

ΔX_t: First-order variable value of input variable at time $t\Delta X_{t-1}$First-order variable of input variable at time $t-1$

3.3. Calculating the output variables (Reward-penalty mechanism)

In this study, the changes of the future stock price (i.e., the rise and fall of stock prices) are considered as the output variables of the neural network. However, considering that errors could be made during the forecasting period, some penalties are necessary. To reflect that reality, appropriate adjustments were therefore made during the calculation of the output value of the neural network. The neural network was used to forecast the trends of 5 min, 15 min, and 25 min relative to the K-line representing the fluctuations in the closing price.

Where the nth fluctuation K-line is defined as follows: Value difference between the highest price and current closing price of the next nth K-line minus the value difference between the lowest price and current closing price of the next nth K-line (Fig. 6).

$$\text{Relativefluctuationsratesofthenth}K - \text{line} = L_0 - L_1 \quad (5)$$

Calculating the fluctuations can help accurately predict the future nth K-line as well as the changing power of both sides. If the strength of the rising side is greater than the strength of the falling side, then this rising trend is used as the trend of the next nth K-line.

3.4. Parameter setting for back-propagation neural networks

According to Vellido [20], the parameters of the neural network have no specific restrictions. They are usually determined by literature reviews, expert opinions or other methodologies. Zhang [21] suggests that the structure be made based on a single hidden layer, as long as such a layer can help achieve a more reliable level of learning. Therefore, the Back-propagation neural network in this paper was set to one hidden layer. The number of hidden nodes is obtained by calculating the average of input nodes and output nodes. The parameters of the neural networks were set as follows: Number of nodes in input layers: 51; number of nodes in output layers: 1; number of hidden layers: 1; number of nodes in hidden layers: 26; number of trainings: 2000.

3.5. Research restriction and trading strategy

1. In this study, we used the back-propagation neural network model to forecast the trends of the Taiwan Stock Exchange Capitalization Weighted Stock Index. Below is a detailed description of the limitations as far as the research process is concerned:
2. A futures contract transaction costs include a handling charge and a transaction tax; the handling charge is currently computed as 60 Taiwan dollars of the futures quotation price, and the transaction tax amounts to 0.002% of the quotation price. This study does not consider the impact of the changes in the index on the transaction tax. So the estimated transaction cost is set to 2 points.
3. To simplify the model, we did not take into account the impact of the changes in interest rates, price changes, exchange rates and futures margin.
4. In situations where the actual transaction includes a pending order price and the transaction price is different, we proceeded based on the assumption that the pending order price could be traded immediately.

In this study, all day transactions closed at 13:30 at the latest. A stop-loss order design, and each time the stopping point was reached, we assumed that the stop-loss price could be traded immediately. We then stopped at the stop-loss point in order to control the maximum loss per transaction as well as transaction costs.

An appropriate and more filtered use of the decision threshold helps attain a greater level of forecast accuracy. The output value for the neural network prediction was between 0 and 1. When a prediction mechanism with no threshold value was used, predicted values greater than 0.5 were considered multi-signals while values less than 0.5 were considered sorting signals. After using the mechanism with threshold values (with 0.2 set as the threshold), the predicted values needed to be greater than 0.7 (0.5 + 0.2) before they were considered as multi-signals, while they had to be less than 0.3 (0.5 – 0.2) before they were considered sorting signals. When we used the trial and error method to conduct the experiment on different threshold values, we found that the higher the threshold, the higher the level of accuracy, whereas the number of transactions decreased in the meanwhile. In order to avoid the above problems, the following screening rule was adopted: Based on the premise that on average, 2 of every 3 transactions are types of day trading, we selected the one with the highest rate of average profit as the basis for screening.

3.6. Performance evaluation model

In this study, two methods of assessment were used to measure and compare the performances of different simulation trading models. With the back-propagation neural network, trading strategies were based on the predicted value which itself was determined by the threshold value. At this point, if the profit of the transaction was greater than zero, then the transaction was said to be in the right direction. Based on that principle, accuracy is therefore defined as follows:

$$Accuracy = \frac{Number\ \ of\ \ correct\ \ predictions\ \ \ (price\ \ rise) + Number\ \ of\ \ correct\ \ predictions\ \ (price\ \ fall)}{Total\ \ number\ \ of\ \ forecasts} \quad (6)$$

Working under the assumption that there was sufficient margin, this study used the Taiwan Futures to conduct real transactions and then used the futures profit points to calculate the profit/loss. With regard to transaction costs, each futures contract included a bilateral transaction tax and a total processing fee of 2 points. The relevant formula is as follows:

$$\begin{aligned}&\text{Total Profit/Loss Points} = \text{Total Purchasing Points}\\&+\text{Total Selling Points} = \Sigma\text{Purchasing Points} + \Sigma\text{Selling Points}\\&= \Sigma(\text{Open Position Price} - \text{Closing Purchase Price}\\&-\text{Transaction Cost}) + \Sigma(\text{Clo singSellingPrice} - \text{Closing Price}\\&-\text{Transaction Cost})\end{aligned}$$

$$\begin{aligned}&Average\ \ income\ \ per\ \ transaction\\&= \frac{Total\ \ Income}{Total\ \ number\ \ of\ \ transactions}\end{aligned} \quad (7)$$

4. Results and analysis

This study is based on both the Taiwan Futures Exchange and the Taiwan Weighted Index Futures data. The sample was collected between January 01, 2013 and May 31, 2013. According to Kearns [22], the ideal training period data and testing data ratio is 8–2. For this study, a total of 26852 written articles were collected. Training-related data were collected between January 1, 2013 and April 30, 2013 for a total of 76 trading days. Testing-related data were collected between May 1, 2013 and May 31, 2013, which corresponds to a total of 22 trading days. The study consisted of two experimental groups and one control group. The training period and the test period were the same in the experimental group and the control group. Each transaction cost corresponds to 2 points. The following experimental values were averaged over 30 experiments. In experimental group 1, the market price index was taken as the input variable to predict the stock price trends of 5 min, 15 min and 25 min respectively. The results are as follows: Table 1.

Results of the experiment showed that the prediction accuracy of experimental group 1 was the highest in 25 min (it reached 76.1%); the second highest level of accuracy was obtained at 15 min. Prediction accuracy for the average profit per day was also highest at 25 min. In the experimental group, predictability at 25 min was higher than that of the control group. Also, both the level of accuracy and the average profit of each group were significantly better than the control group. In the second group, the market price index and the trading volume index were used as the input variables to predict the stock price trends of 5 min, 15 min and 25 min respectively. The results are as follows: Table 2.

Results of the experiment showed that the forecast accuracy for the second group was highest at 25 min (78.3%); and the second highest forecast accuracy was obtained at 15 min. Predictability for experimental group 2 at 25 min was higher than that of the control group. Also, the level of accuracy of the two groups was significantly higher than that of the control group. In the experimental group, the market price index was used as the input variable, while in the experimental group 2 the market price index and the volume index were used as input variables to forecast the rise and fall in stock prices after 5 min, 15 min and 25 min respectively. The experimental results are as follows: Table 3

4.1. Statistical test

In this study, the significant level of 0.05 was used as the basis for determining whether the variants were equal. The results of the statistical test are shown in Table 4. Only the test results of Experimental group 1 and Experimental group 2 did not reject H0. After determining whether the population variances were equal or not, the accuracy T-test of the control group and the experimental groups was conducted.

$$H_0 : \sigma_A^2 = \sigma_B^2$$

$$H_1 : \sigma_A^2 \neq \sigma_B^2$$

4.2. Accuracy testing of experimental groups and the control group

In this study, the results of the T-test were used to test the accuracy of the experimental groups and the control group. The results are shown in Table 5. The null hypothesis is expressed as:

H_0: $\mu_B \geq \mu_A{}^{\circ}$

From Table 5, it can be seen that at a significant level of 0.05, the null hypothesis is rejected at 5 min, 15 min, and 25 min prediction. This is an indication that both experimental group 1 and 2 were greater than the control group, and that the level of accuracy of experimental group 2 was greater than that of experimental group 1. This means that the input variables for the trading volume indicators of the market profile theory have an actual effect when applied to the neural network learning.

Using the market profile theory, the experimental results showed that the level of prediction accuracy is highest and significant at 25 min followed by that of 15 min. The level of prediction accuracy was lowest at 5 min. Since experimental group 2 fell into the trading volume index of the market profile principle, we compared its accuracy level with that of experimental group 1 and realized that there was improvement. In fact the prediction accuracy was especially significant at 5 min.

Studying the transaction volume indicators can better reflect the dynamics of the market, thus allowing for a more correct comprehension of the rules that govern short-term fluctuations. But in terms of average profit, no significant difference was observed at 5 min. It wasn't until 25 min that experimental group 2 showed

Table 1
Results for Experimental Group 1 and Control Group.

Forecast Time	Model	Accuracy	Total Profit Score	Number of Transactions	Average Profit Score per Transaction
5 min	Experimental group 1	65.5	259	15	17.2
	Control group	47.7	−91	15	−6
15 min	Experimental group 1	69.3	212	13	16.3
	Control group	43.9	−127	13	−9.8
25 min	Experimental group 1	76.1	390	17	22.9
	Control group	51.2	20	17	1.2

Table 2
Results for Experimental Group 2 and Control Group.

Forecast Time	Model	Accuracy	Total Profit Score	Number of Transactions	Average Profit Score per Transaction
5 min	Experimental group 2	70.8	220	13	16.9
	Control group	42.5	−171	13	−13.1
15 min	Experimental group 2	73.2	261	14	18.6
	Control group	47.1	−69	14	−4.9
25 min	Experimental group 2	78.3	488	15	32.5
	Control group	48.8	−14	15	−0.9

Table 3
Results for Experimental Group 1 and Experimental Group 2.

Forecast Time	Model	Accuracy	Total Profit Score	Number of Transactions	Average Profit Score per Transaction
5 min	Experimental group 1	65.5	259	15	17.2
	Experimental group 2	70.8	220	13	16.9
15 min	Experimental group 1	69.3	212	13	16.3
	Experimental group 2	73.2	261	14	18.6
25 min	Experimental group 1	76.1	390	17	22.9
	Experimental group 2	78.3	488	15	32.5

Table 4
Statistical test – Variations.

Forecast Time	Model A	Accuracy	Model B	Accuracy	P-Value	Test results
5 min	Experimental group 1	65.5	Control group	42.5	***	Reject H_0
	Experimental group 2	70.8	Control group	42.5	***	Reject H_0
	Experimental group 2	70.8	Experimental group 1	65.5	**0.258233**	**Non rejectH_0**
15 min	Experimental group 1	69.3	Control group	47.1	***	Reject H_0
	Experimental group 2	73.2	Control group	47.1	***	Reject H_0
	Experimental group 2	73.2	Experimental group 1	69.3	**0.178283**	**Non rejectH_0**
25 min	Experimental group 1	76.1	Control group	48.8	***	Reject H_0
	Experimental group 2	78.3	Control group	48.8	***	Reject H_0
	Experimental group 2	78.3	Experimental group 1	76.1	**0.440079**	**Non rejectH_0**

Table 5
Statistical T-test – Accuracy.

Forecast Time	Model A	Accuracy	Model B	Accuracy	P-Value	Test results
5 min	Experimental group 1	65.5	Control group	42.5	***	Reject H_0
	Experimental group 2	70.8	Control group	42.5	***	Reject H_0
	Experimental group 2	70.8	Experimental group 1	65.5	***	Reject H_0
15 min	Experimental group 1	69.3	Control group	47.1	***	Reject H_0
	Experimental group 2	73.2	Control group	47.1	***	Reject H_0
	Experimental group 2	73.2	Experimental group 1	69.3	***	Reject H_0
25 min	Experimental group 1	76.1	Control group	48.8	***	Reject H_0
	Experimental group 2	78.3	Control group	48.8	***	Reject H_0
	Experimental group 2	78.3	Experimental group 1	76.1	***	Reject H_0

obvious signs of difference in profit value. But in general, adding up the eigenvalues of the trading volume index can effectively improve the effect of the neural network prediction.

5. Conclusion

Financial engineering physics can really help indicate the movements of the stock market and the directions of the physical forces. Our experiments also helped clearly define the direction of the difference between price and value discussed in the Market profile theory. Based on our findings, the following conclusions can be drawn:

Contrary to the control group (the random trade strategy model), the experimental group used the back-propagation neural network to study the eigenvalues of the market profile. The levels of accuracy and profitability were both significantly better than the ones obtained with the random trading strategy model. This confirms the fact that the market profile theory is an effective technical analysis tool.

For experimental group 1 and experimental group 2, results were the best at 25 min, both in terms of forecast accuracy and profitability. This clearly shows that the market profile indicators have a better forecasting ability in the long-term.

Compared with experimental group 1, the input variables for experimental group 2 were added to the volume of the market profile eigenvalues, all of which have contributed to increasing the forecast performance. The improvement in forecast accuracy at 5 min was the most remarkable. The display of the volume indicators reflects the real level of greed and fear in the market dynamics. This, combined with the volume model, can result in better performance 20–30% compared to random models.

In this study, we used the market profile theory to describe the greed and fear-related behaviors of day traders. And to understand to what extent these psychological factors affect the “buy high sell low” behavior in the market, we looked at the trading volume of large as well as retail investors. The reasons behind the volume amplification were often the aggressive behaviors (greed) displayed by large investors and the sudden strong feeling of fear that prevented retail investors from making rational decisions. Although the market price index can help quantify the velocity of price changes by using the concepts of first-order variables and second-order variables in financial engineering physics, it is impossible, however, to grasp the trends and dynamics of traders' behavior when price volumes increase or decrease.

We therefore suggest that the tool of measurement be supplemented with a volume indicator to reflect any changes in the psychological factors (fear and greed) that affect traders' decisions. This will not only lead to a better understanding of day traders' greed and fear-related behaviors; it will also be an exemplary method for public investors in the stock market, helping them to reach a better forecast of intraday stock price trends, and better rewards.

In order to improve the empirical process and model, we offer the following three recommendations for future research on the same topic: First: The trading strategies adopted in this study are limited to single transaction contracts, venting, and stop-loss settings. We believe that including dynamic trading strategies will lead to better investment performances. Second: In the input variables section, it would be better to also make use of other pre-clustering methods to obtain the eigenvalues. This can increase the effectiveness of neural network learning and help better predict the trend of stock prices. Third: From the roadmap of the market profile theory it is possible to obtain a direction for a new study on behavior patterns. One may be for instance interested in exploring transaction volume changes of intraday stock alerts in the long run. Another possibility is analyzing the behavior of (old and new) buyers and sellers with regard to changes in transaction volume. This method can be used in practice for a better and more appropriate description of long-term traders' behaviors.

References

[1] L.S. Maciel, R. Ballina, Design a Neural Network for Time Series Financial Forecasting: Accuracy and Robustness Analysis, 2008.

[2] D. Marcek, Granular RBF NN approach and statistical methods applied to modelling and forecasting high frequency data, Int. J. Comput. Intell. Syst. 2 (4) (2009) 353–364.

[3] M. Tanaka-Yamawaki, S. Tokuoka, Adaptive use of technical indicators for the prediction of intra-day stock prices, Phys. A: Stat. Mech. Appl. 383 (1) (2007) 125–133.

[4] T. Takahama, S. Sakai, A. Hara, I. Noriyuki, Predicting stock price using neural networks optimized by differential evolution with degeneration, Int. J. Innov. Comput. Inf. Control 5 (12 (B)) (2009) 5021–5032.

[5] An-Pin Chen, Y.-C. Chen, W.-C. Tseng, Applying extending classier system to develop an option-operation suggestion model of intraday trading-an example of Taiwan index option, Lect. Notes AI (2005) 27–33.

[6] An-Pin Chen, Yu-Chia Hsu, Dynamic physical behavior analysis for financial trading decision support, IEEE Comp. Int. Mag. 5 (4) (2010) 19–23.

[7] S. Barak, M. Modarres, Developing an approach to evaluate stocks by forecasting effective features with data mining methods, Expert Syst. Appl. 42 (2015) 1325–1339.

[8] M.-Y. Chen, D.-R. Chen, M.-H. Fan, T.-Y. Huang, International transmission of stock market movements: an adaptive neuro-fuzzy inference system for analysis of TAIEX forecasting, Neural Comput. Appl. 23 (2013) S369–S378.

[9] R.J. Shahrabi Hafezi, E. Hadavandi, A bat-neural network multi-agent system (BNNMAS) for stock price prediction: case study of DAX stock price, Appl. Soft Comput. 29 (2015) 196–210.

[10] L.-Y. Wei, C.-H. Cheng, A hybrid recurrent neural networks model based on synthesis features to forecast the Taiwan stock market, Int. J. Innov. Comput. Inf. Control 8 (8) (2012) 5559–5571 (August).

[11] C.S. Chen, An-Pin Chen, Market Profile Behavior Analysis Based on 1 min Intervals of TPO: A Thesis Submitted to Institute of Information Management College of Management National Chiao Tung University, 2013.

[12] T. Kimoto, K. Asakawa, Stock market prediction system with modular neural networks, IEEE Int. Joint Conf. Neural Netw. 1 (1990) 1–6.

[13] Y. Yoon, G. Swales, Predicting stock price performance: a neural network approach, Mit Press 25 (11) (1991) 39–44.

[14] R. Gencay, Non-linear prediction of security returns with moving average rules, J. Forecast. 15 (1996) 165–174.

[15] Paul Kang, Hoh Phua, Daohua Ming, Weidong Lin, Neural network with genetically evolved algorithms for stocks prediction, Asia Pac. J. Operat. Res. Singapore 18 (2001) 103–107 (May).

[16] L. Harris, A transaction's data study of weekly and intradaily patterns in stock returns, J. Financ. Econ. 16 (1986) 99–117.

[17] A.-S. Chen, M.T. Leung, H. Daouk, Application of neural networks to an emerging financial market: forecasting and trading the Taiwan Stock Index, Comput. Operat. Res. 30 (6) (2003) 901–923.

[18] S. Chai, M. Kim, H.R. Rao, Firms' information security investment decisions: stock market evidence of investors' behavior, Decis. Support Syst. 50 (4) (2011) 651–661.

[19] Z. Tan, C. Quek, P.Y.K. Cheng, Stock trading with cycles: a financial application of ANFIS and reinforcement learning, Expert Syst. Appl. 38 (5) (2011) 4741–4755.

[20] A. Vellido, P.J.G. Lisboa, J. Vaughan, Neural networks in business a survey of applications (1992–1998), J. Expert Syst. Appl. 17 (1999) 51–70.

[21] G. Zhang, B.E. Patuwo, M.Y. Hu, Forecasting with artificial neural networks: the state of the art, Int. J. Forecast. 14 (1998) 35–62.

[22] M. Kearns, A bound on the error of cross validation using the 43 approximation and estimation rates, with consequences for the training-test split, Adv. Neural Inf. Process. Syst. 8 (1996) 183–189.

[23] J.P. Steidlmayer, Markets and Market Logic, Chicago: Porcupine Press, 1984.

[24] J.B. Heaton, N.G. Polson, J.H. Witte, Deep Learning in Finance arXiv preprint arXiv. (2016) 1602.06561.

[25] Eunsuk Chong, Chulwoo Han, Frank C. Park, Deep learning networks for stock market analysis and prediction: methodology, data representations, and case studies, Expert Syst. Appl. 83 (2017) 187–205.

[26] Akira Yoshihara, et al., Predicting stock market trends by recurrent deep neural networks, in: Pacific Rim International Conference on Artificial Intelligence, Springer. Cham., 2014, pp. 759–769.

[27] Chiu-Chin Chen, et al., Applying market profile theory to forecast Taiwan Index Futures market, Expert Syst. Appl. 41 (2014) 4617–4624 (10).

[28] Mei-Chen Wu, et al., Data Mining Application to Financial Market to Discover the Behavior of Entry Point–A Case Study of Taiwan Index Futures Market, Harmony Search Algorithm, Springer, Berlin Heidelberg, 2016, pp. 295–303.

[29] S. Alamelu Mangai, B. Ravi Sankar, K. Alagarsamy, Taylor series prediction of time series data with error propagated by artificial neural network, Int. J. Comput. Appl. 89 (1) (2014) 0975–8887.