

Simulating the Investor's Behavior in Stock Market with an Intelligent Agent

FengYu Chang; DenYiv Chiu

Institute of Information Management

Chung Hua University

Hsinchu, Taiwan

phungyu@mi.chu.edu.tw; Chiuden@mi.chu.edu.tw

Abstract

In this paper, from the investor's perspective, we will use intelligent agents to simulate the investor's behaviors. The intelligent agent will learn the investor's behavior according to investor's historical data. It will learn from the decisions making for the investment. And we will build a virtual stock market in which intelligent agents can invest according to their decisions. Intelligent agents' investment decisions are different due to different consideration and conclusion for the same message. Then, every intelligent agent will deal with these messages. We will train intelligent agent with data mining technology and artificial neural network to make the intelligent agent's behaviors similar to investor's behaviors in the real world. The intelligent agent makes decisions and invests in the virtual stock market.

1. Introduction

The investors will expect satisfactory benefit from the investment by applying best planning. There are many financial management tools about the investment to serve the investors. Some investors may trust the tools and some may doubt of these tools. They may believe in their own judgments or investment consultants. They may have their habits to invest stocks in the stock market. This research is from another point of view to propose a new architecture in which prediction can be performed from investor's perspectives.

An investor is a member in the stock market. Investors make their investment decisions in stock market. Stock price is influenced by their buying or selling stocks, so we can say that investor's supply and demand decide stock price. Because the investor's supply and demand decide the stock price, we can think about stock price from the investor's perspectives.

From the investor perspective of prediction, because investor's supply and demand will form stock price and they decide the TWSP (Taiwan weighted stock price index). The investors sell and buy stocks to decide the stock price. The investor is an entity of the stock market. We can simulate the investor's investment decision or portfolio and put the decision to the virtual stock market. It can decide the trend of the TWSP or the stock price.

So, in this project, we will use intelligent agents to simulate the investor's behaviors. We will use intelligent agents to learn the investors' behaviors according to investor's historical data. Then, agents will find the rules about the investors' behaviors. These rules help agents to

make decisions. And we will construct a virtual stock market that is a virtual place for agent to commit its transactions. If we can learn investor's decision-makings, we can predict the result of weighted stock prices. I hope that the research can provide valuable reference for real investors.

We will discuss background, problem definition and motivation about this research in the following sections.

1.1 Background

Some of researches [16,10] use artificial intelligent technologies to predict the tendency of stock market. Others are studies about interest rates [8], and financial data [11, 12]. The others are studies about behaviors [14, 15, 9]. But the price fluctuations of the stocks are the result of individual operations of many individual investors in the financial markets [21]. Most investors believe that messages about stock affect stock prices, and also they have significantly influence on investor's behaviors. Investors invest in the stock market to behave in various manners depending on their individual preference, expectations, decision environment, strategy, and individual difference [6]. So, this paper explores how the investor's behaviors are influenced by the messages about the stock, and also uses investor's historical data to simulate the decision process. We use those approaches to predict investor's behaviors and decisions.

1.1.1 Messages and Information Affecting Investors

The stock price is decided by the supply and demand, and it is also formed by the behaviors of the investors mutually. So, the messages rather affect the stock price than affect the investors. Then, investors' behavior affects the stock price. An investor will anticipate the trend of all investors' trading in the stock market, and the investor may follow the tendency to invest. This tendency is the foundation of stock price [6].

The messages and information that affect the stock market actually affect individual investors or institutional investors. We simulate the investors' decision-makings and predict it according to how investors are influenced by financial messages.

The factors affecting the investors can be classified into following:

- (1) Politics: this includes the war, Cross-Strait relationship and situation of politics. In October of 1998, sent the TAIEX plummeting, where the weighted index had taken a free-fall from its high 10,200 points straight

down to bottom at an all-time low of 5,421 points, this can be the immediate cause of the missile threats by mainland Chinese authorities or the ripple effects of the Asian financial crisis, or a long-time effect of over-specified stock valuations [6]. And some crash situations about politics will also influence the stock market [19].

- (2) Economics: this includes the exchange rate, interest rate, inflation, money supply, international economics, and internal economics. The fluctuation of exchange rate had obvious influence [5]. In the 1997, the Thailand's currency went straight down to 80% in one day. It started the Asia financial crisis. The stock price fell or the foreign money flew out. Those factors triggered the starting time of the financial crisis [3]. The government diminished the interest rate to make increment of money supply. And investors were encouraged to invest their money in the stock market. The international and the internal economics influence investor's behaviors.
- (3) The risk of company: this includes the technical analysis, economic condition, financial structure, cash flow, debt paying ability, buying on margin, short sales, and so on. The market to book ratio has a positive relationship with prior 1,2,3 and 4 period's return rate. Book to market ratio and size effect exists constantly in Taiwan's electronic stocks. This is a positive relationship between return rate and systematic risk [4].
- (4) Others: this includes earthquake and others. The earthquake influences the stock because it will cause much danger for the electronic industry.

1.1.2 Investor's Behaviors

Investor's behaviors may change because of the relative messages, technology analysis, base analysis and so on. We can explore the investor's behaviors to analysis his possible investment path. In the common theory proposed by Kenyes, he had done the research about the mentality of the herding people. He pointed out that investors might not make decisions from the information he collected, but investors would follow the herding.

Herding and feedback trade can be used to explain potential finance situations, including excess volatility, momentum and reversals in stock prices.

Herding means that a group of investors follow mutually to trade in the same way in the some period. Feedback trade is a relation between the herding and lag return.

Feedback trade has two types as below.

- (1) Positive feedback trade strategy: The decision threshold is the average performance in the pass. If the stock's performance is higher than the threshold, it can be highly recommended to purchase. If the stock's performance is lower than the threshold, it should be sold.
- (2) Negative feedback trade strategy: It is a negative strategy to the positive feedback trade strategy. Buy the stock if the performance of the stock is lower than the threshold performance, and sell the stock if the

performance of the stock is higher than the standard performance.

The investors follow the same messages to form the herding. The individual investor sells the stock that he earned a few profits, but a loser holds the stock for a long time [18]. The individual investing stocks may follow his own behavior patterns.

The intelligent agent is simulating investor's behaviors in the research. The individual investor may be influenced by messages, individual preferences and wealthy structure. The intelligent agents receive messages about stock to decide buy or sell the stock. These messages include the noise of stock market, the operating situations of companies, and other relative information. The intelligent agent makes different decisions because of the individual preferences. It has also the different influence to stock market because of individual wealthy structure. Every intelligent agent may make different investment decisions. Major Institutional Investors have the huge capital that can be used to buy a lot of stock, and they have strong influence on the stock market. But most of individual investors may have similar tendency about buying or selling, this situation also can make strong influence on the stock market. So, each of investors can play an important role in the market.

1.2 Motivation

As discussed above, we can know that an investor is the most important role in the stock market. Because stock price is influenced by investors' buying or selling stocks from economic view according to some researcher's proposition. And then, if we can simulate the investors' behaviors and predict them, we can predict the tendency of Taiwan stock market. So, we will use intelligent agents to simulate the investor's behaviors.

1.3 Purpose

Our purpose is to simulate the investor's behaviors with intelligent agent. Intelligent agents will have themselves rules, and we will let the agents to make their decisions according to the rules in the virtual stock market. So we would construct the virtual stock market according to real working operations in the Taiwan stock market. The intelligent agents can buy stock on margin, cash, or short sales in the virtual stock market. Then, we will compare the real trade investors' situations with agents' decisions and evaluate the performance.

1.4 Assumption

We use the concept of classification technique of Chinese news proposed by Kun-Po Chuang's thesis, and he had made several assumptions. Therefore, we also have to observe these assumptions as followings.

- (1) We focus on processing Chinese document which contains a caption and a content. (This type of documents includes newspapers, magazines and some reports.)

- (2) Each classification keyword contains at least two characters and the main keywords of a document appear frequently in the document.
- (3) All Chinese news that happens each day can be obtained by every investor.
- (4) Chinese news is affecting the Taiwan Stock Market Weighted Price Index (TSWPI).

And we also have one assumptions in our framework:

- (1) Every investor would obtain all macroeconomic variables and technological analysis variables that happen each day.

The construction in this paper is described as following. In the section two, we will review related literatures and related technologies. Next, we will give a complete description of our model in the section three. In the section four, we will describe the experimental results of our model and analysis the performance of our model. The last is the conclusion and the future works.

2. Literature Review

We would introduce technologies that be used in this research.

2.1 Artificial Stock Market Technology Review

Shu-Heng Chen and Chia-Hsuan Yeh proposed an architecture to study artificial stock markets. This architecture rests on mechanism called “school” which is a procedure to map the phenotype to the genotype [20]. It’s an agent-based model of “school”, and considers “school” as an evolving population driven by single-population GP [20]. An agent is a trader who is an investor. Traders would learn the investment’s strategies in the business school and trade with these strategies in artificial stock market. This model gives us an ideal about simulating individual investors. In this project, we propose a framework in which we use agents to simulate individual investor’s behaviors. And, we will build the virtual stock market as the ‘school’ in which an agent acts as a trader to make decisions and to invest.

2.2 Data Mining Technology Review

Data mining is the process of extracting potentially useful information from raw data [7]. Data mining can mine large amounts of data. This makes data mining a fast and inexpensive way of summarizing, exploring, understanding, and analyzing data [7].

We will find rules from historical data. We will use data mining to find rules. Data mining often mentions finding patterns, rules, and facts in raw data. And we introduce two mining n-dimensional inter-transaction algorithm:

2.2.1 EH-Apriori

EH-Apriori is a method to mine N-dimensional inter-transaction rules. EH-Apriori is the extension of Apriori based algorithm. EH-Apriori adopts a similar technology of hashing to filter out unnecessary candidate

2-itemsets. When the support of candidate C_1 is counted by scanning the database, EH-Apriori accumulates information about candidate 2-itemsets in advance. Such resulting hash table can be used to greatly reduce the number of 2-itemsets in C_2 [10].

2.2.2 OP-Apriori

OP-Apriori is used to reduce the total mining times [5].

OP-Apriori scans the database to find out the L_1 and L_2 in the first time. It doesn’t use C_2 to produce L_2 . It used to need C_2 to produce L_2 before the OP-Apriori method. OP-Apriori modified this way to perform better than EH-Apriori.

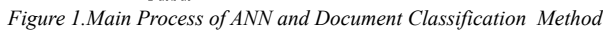
Mining n-dimensional inter-transaction association rules is more complex than mining classical association rules. Investor’s behaviors are also complex. We try to use n-dimensional inter-transaction to mine investor’s investment data, and try to find rules about investor’s behaviors.

2.3 ANN and Document Classification Used to Deal with News

In the Chunag’s research [13], he combines neural network with document classification techniques to deal with news. Chinese historical news are classified into several fixed categories. And he uses document classification and keywords table to classify these Chinese historical news. He also collects the Taiwan Weighted Stock Price Index (TWSPI) and computes the raising or falling rate every day. And he maps this result to map the Chinese history news. He quantifies the information of Chinese historical news by transformation process. Finally, he feeds the quantitative data and qualitative data into artificial neural network and use artificial neural network to train the weights as the degree of influence.

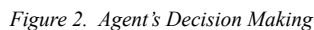
First, he uses these keywords to classify the Chinese history news collected from WWW, as shown in figure 1. After classifying, this component of classification mechanism saves and output some keywords. Secondly, he uses clustering methods to deal with each category. This method would group the similar Chinese history news into the same type for each category. In this framework, he used hierarchical clustering to group the similar news into the same type.

Then, he can use raising or falling rate of each historical news to compute average raising or falling rate for each type. We also compute how many times raising or falling occurs for each type. We use ANN-document classification to deal with news and try to find the news that has influence on investors.



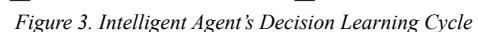
We would introduce our methodology in this chapter. We divide this chapter into two parts. The first part is our architecture. The second part is agent’s decision framework.

We use agents to simulate real investors' behaviors, and each agent represents an individual investor. Then, each agent has his own decisions and it can make decisions in the virtual stock market. These decisions can affect every stock's price in the virtual stock market. And weighted stock price index is composed of stocks' prices in the virtual stock market, as shown in figure 2.



(1) **Historical Data:** Investor's investment transactions will be saved to historical data every day. The historical data consists of stock name, stock shares, times, and

The agent's learning cycle is shown in figure 3. First, this model will save the investor's historical investment data and use historical data to analyze investor's behaviors. But events also may affect investor's behaviors. So, we use analytic information and events to decide agent's portfolio.



(1) *Individual Wealth Structure*: Individual wealth structure means how much money an investor can invest in the stock market. Individual investor has himself individual wealth structure. Individual wealth structure would record the wealthy distribution about bonds, common fund, stocks and so on. In the record, we can find how much money investor can invest in the stock market. If an agent

has ability to buy more stocks, he has more influence to the stock market. An agent's wealthy structure would change while an agent gains or loses from investments, and he will leave from the virtual stock market if he loses all of his money. An agent also can enter the market again if he has money. Wealthy structure decides whether an agent is active or not in the virtual stock market.

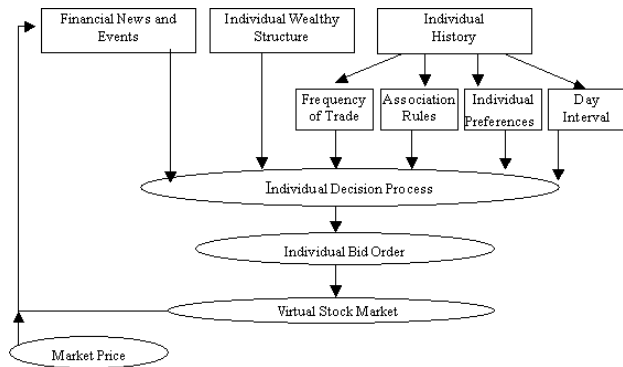


Figure 4. The Framework of Simulation

- (2) *Individual Investor Historical Data*: This component stores data that investors invest every day. The data of this component includes: 1. stock names, stock prices, and stock shares that investors buy or sell. 2. information about investor's buying stocks on margin sales or short sales.
- (3) *Individual Decision Process*: Individual decision process uses the frequency of trade, individual preference, association rules, and day interval form the component two. And individual decision process will produce individual bid order of an agent.
- (4) *Individual Bid Order*: Individual bid order is the trading list that an agent will trade next week. The intelligent agent will make decisions by his bid order next trading week in the virtual stock market.
- (5) *Virtual Stock Market*: We will construct the virtual stock market to simulate the real stock market in Taiwan. The virtual stock market includes three parts.
 - (a) *An Intelligent Agent*: An intelligent agent represents an investor. An agent can trade according to his bid order in the virtual stock market.
 - (b) *Model of Price Determination*: The mechanism is to decide the stock price at which agents buy or sell. And the mechanism is similar to the real Taiwan stock market. The agents can buy the stocks by cash, margin sales, and short sales. Virtual stock market may decide the stock price according to the agents' investment decisions, buying or selling the stock.
 - (c) *Taiwan Weighted Stock Price Index (TWSPi)*:

The virtual stock can decide the TWSPi according to all stocks' prices. So we can find that the supply and demand from the intelligent agents in the virtual stock market can predict the Taiwan weighted stock price index.

3.2 Agent's Decision Framework

This part will introduce the technologies used in the agent's decision framework and we will explain how to construct our framework step by step. Our framework is shown in the figure 5 and it has several elements as followings:

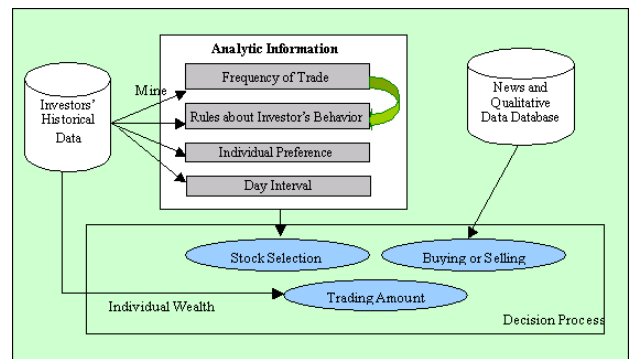


Figure 5. Agent's Decision Framework

- (1) *Buying or Selling Stocks*: We will collect all kinds of events including news or qualitative data, and use the events to decide whether investor would buy or sell stocks or not. We will also construct database to save news and qualitative data, and use ANN-document classification methods to deal with news and qualitative data. Then, we can obtain the result and use the result to decide whether investor would buy or sell stocks or not.
- (2) *Trading Amount*: Trading amount means how much money an individual investor can invest in stock market. We can compute investor's wealth from the transactions of investor's investment in the stock market and infer individual investor's trading amount that may investe in the stock.
- (3) *Stocks Selection*: We will use analytic information to select stocks that an individual investor may invest in the stock market. We can find analytic information from investor's historical data. Analytic information includes rules about investor's behaviors, individual preferences, and frequency of trades. We use n-dimensional inter-transaction (*OP-Apriori*) and Apriori to mine associate rules. The frequency of trades will be used as scale of sliding windows' dimensions. If an investor has higher investment frequency in the same stock, we can get the higher value of individual preference. Then the higher value of individual preference means that the investor has better opportunity to invest in the stock. Day interval is the duration of two successive transactions for the same stock committed by an investor. So we can use the analytic information to select stocks that investor may

invest next week.

3.2.1 Buying or Selling Stocks

We can use events to decide whether investor would buy or sell stocks or not. The events include news, technology analysis data, economic variable, and EPS etc. We will use ANN-document classification to deal with events that may influence investor's decisions. Every investor has different decision about the same events. We use ANN-document classification to deal with events and try to find degree of influence for each agent. Every agent has his own ANN-document classification to deal with news that may influence the investor and trains his ANN-document classification according to his buying or selling stocks every day. We describe the process in the followings.

(1) Inputs: ANN's input includes following elements. And following element would be save in news and qualitative data database in figure 5.

- (a) Qualitative Data: this part is responsible to search daily Chinese news, including internal politics, political situations, internal economic situations, international political situations, international economic situations, company situations, cross-Strait relationships, general investment environments, and so on.
- (b) Macroeconomic Variables: this part is responsible to search several economic indicators consisting of daily change in exchange rate, daily change in interest rate, monthly change in unemployment rate, annual change in economic growth rate, annual change in consumer price index, annual change in Gross Domestic Product (GDP), monthly change in inflation rate, and monthly change in money supply rate. These data can be collected from several major countries, including American, Japan, Hong Kong, Singapore, Korea, England, France, and Germany.
- (c) Technological Analysis Variables: this part is responsible to search Taiwan Weighted Stock Price Index (TWSPI) and compute several technological indicators, including Twenty-Four Days Moving Average (MV), Relative Strength Index (RSI), Bias, Moving Average Convergence/Divergence (MACD), Volume Moving Average, and Twenty-Four Days On Balance Volume (OBV).

(2) Outputs: Individual investors may buy or sell their stocks. Events influenced whether the investors would buy or sell stock or not. We assign one to buying stock and assign -1 to selling stock to train the ANN-document classification model.

(3) Prediction: We will predict whether the investors would buy or sell stock or not next week.

3.2.2 Trading Amount

We may not have individual wealthy structure. But we can compute individual wealthy through an investor investment in stock market. We will find how much money an investor can invest in the stock market. This can influence how many stocks that an investor can invest. We will use following equation to compute trade size, as shown in figure 6

An investor may own following stocks, and we can compute his wealthy.

Asset	Shares	Prices
Stock ₁	Share ₁	P ₁
Stock _n	Share _n	P _n

$$\text{Trading Amount} = \text{Share}_1 * P_1 + \dots + \text{Share}_n * P_n$$

Figure 6. Trading Amount

3.2.3 Stocks Selection

Analytic information includes frequency of trade, rules about an investor's behaviors, individual preferences, and day interval as shown in figure 5. We use analytic information to select stocks.

(1) Frequency of Trade

This part will compute the frequency of trade that an investor may trade daily, weekly, monthly, or a long time from his historical data. An investor has his own habits to trade stocks. We will use following equation to compute average frequency of trade, as shown in figure 7. We can use frequency of trade to be the size of each dimension for the sliding window.

T_1 : Trade date that an investor invest the first time

T_2 : Trade date that an investor invests the next time

T_n : Trade date that an investor invests Nth time

$$\text{Frequency of Trade} = \frac{\langle T_2 - T_1 \rangle + \langle T_3 - T_2 \rangle + \dots + \langle T_n - T_{n-1} \rangle}{N-1}$$

One scale of Sliding Window's "Time" Dimension = Frequency of Trade

Figure 7. Frequency of Trade

(2) Mining Rules about Investor's Behaviors

We use OP-Apriori and Apriori algorithm to find out rules about investor's behaviors from individual historical data.

When algorithm scans the database, it will produce C_1 and C_2 directly the first time. And we will record the support of C_1 and C_2 in an array after scanning the database. We will mine association rules for investors' behaviors with following steps.

- (a) Transforming to Star Schema: We will transform our individual historical data into star schema and we will transform all

transactions to star schema in the next step. Table 1 shows the original data and the table 3 shows its star schema. And table 1 shows the meaning of the sliding window's dimension.

Table1 Original Data in Database

Investor	Stock-Number	Buy-or-Sell	Price-Difference	Time
A	2806	sell	2	4/22
A	2409	sell	3	4/23
A	2409	buy	-2	4/24
A	2325	buy	-3	4/25
A	2806	buy	-1	4/26

Dimension one: dimension one is the field of buying-or-selling and its value can be sell or buy. Then we transform “buy” to “1” and “sell” to “2”.

Dimension two: dimension two is the field of price-difference and its value can be -3, -2, -1, 2, or 3. Then we transform “-3” to “1”, “-2” to “2”, “-1” to “3”, “2” to “4” and “3” to “5”.

Dimension three: dimension three is the field of time and its value can be 4/22, 4/23, 4/24, 4/25, or 4/26.

$$\text{Frequency of trade} = \frac{(4/23 - 4/22) + (4/24 - 4/23) + (4/25 - 4/24) + (4/26 - 4/25)}{5 - 1} = 1$$

The scale of dimension three is one.

Then we transform “4/22” to “1”, “4/23” to “2”, “4/24” to “3”, “4/25” to “4” and “4/26” to “5”.

Table 2. Meaning of Sliding Window's Dimension

		Value	Meaning of Value
Events		a, b, c	a is {2806}, b is {2409}, c is {2325},
Dimension One	Buy-or-Sell	1, 2	1 is {buy}, 2 is {sell}
Dimension Two	Price-Difference	1,2,3,4,5,6	1 is {-3}, 2 is {-2}, 3 is {-1}, 4 is {2}, 5 is {3}
Dimension Three	Time	1, 2, 3, 4, 5	1 is {4/22}, 2 is {4/23}, 3 is {4/24}, 4 is {4/25}, 5 is {4/26}

Table 3. Comparison of Sliding Window's Dimension

Investor	Stock Number	Event	Buy or Sell	Dimension One	Price Difference	Dimension Two	Time	Dimension Three
A	2806	a	Sell	2	2	4	4/22	1
A	2409	b	Sell	2	3	5	4/23	2
A	2409	b	Buy	1	-2	2	4/24	3
A	2325	c	Buy	1	-3	1	4/25	4
A	2806	a	Buy	1	-1	3	4/26	5

(b) Scanning the Database: OP-Apriori will scan the investor's star schema. There are two steps as followings. First, it scans each point in the N-dimensional database with sliding windows to find base address. Secondly, it will infer the relative address according to base address and it will move the sliding window to scan the n-dimension database.

By using the data in table 3, we can draw a 2-dimensional sliding window as shown in figure 8.

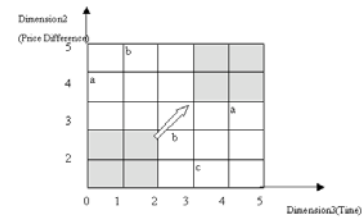


Figure 3-8. A Sliding Window Example

Figure 8. A Sliding Window Example

(c) Saving the Association Rules: We will record these association rules, and an agent will use these rules to select stocks.

(d) Saving the Association Rules: We will record these association rules, and an agent will use these rules to select stocks.

(3) Individual Preferences

An investor may interest in some kind of stocks. If someone likes to invest in stocks of electronics industry, he may be familiar with Taiwan Semiconductor Manufacturing Company and United Microelectronic Corporation. If he invests in Taiwan Semiconductor Manufacturing Company and United Microelectronic Corporation, he may interest in investing in other stocks of electronics industry. An investor may interest in several kinds of stock. So, we can find the stocks that he prefers to invest.

(4) Day Interval and Density Rate

Day interval is N's trading date minus N-1's trading date of the same stock for the same investor. *Density rate* is how many times that an investor invests in the same stock the same week.

There is an example in the following table 4. First, we would compute the day interval that investors invest in the same stock. Investor A invests in stock A on 2001/3/18, 2001/3/19, 2001/3/26 and 2001/3/28. We can compute the day interval which are 0, (2001/3/19 - 2001/3/18), (2001/26 - 2001/3/19), and (2001/3/28 - 2001/3/26). And we can get the day interval, 0, 1, 7, and 2. The one represents that investor A may trade stock A next day. The two represents that he may trade stock A next two days. We use following rules to compute *density rate* as shown in figure 9.

Table 4. Investor's Data

Auto_Number	Investor's Name	Stock Name	Date
1	Investor A	Stock A	2001/3/18
2	Investor A	Stock A	2001/3/19
3	Investor A	Stock A	2001/3/26
4	Investor A	Stock A	2001/3/28

```

Day Interval= Dates - Daten-1
If 0 < Day Interval < 7 then
    Count=Count + 1
endif

Density Rate= Count / Sum of Trading Times

```

Figure 9. Day Interval

We will compute the sum of trading times of the investor A. Because we have four data for investor A, its sum of trading times is four. There are two data with day interval between 0 and 7. So, the density rate is fifty percent, two of four.

Fifty percent is high rate of density and it has high correlation this week. But it may not have any correlation between two weeks. If the density rate is higher than fifty percent, it has higher correlation on one week. And we can use the stocks transacted in the preceding half of the week to predict the stocks that investors will trade in the next half of the week.

(5) Main Algorithm to Select Stocks

We have following steps to select stocks according to the result obtained from previous steps:

Step 1: Searching the association rule database

In this step, we will use the following algorithm to search the association rule as shown in figure 10.

```

IF support value > threshold, then
    association rules will be selected
    If the number of appropriate stocks > 2, then
        compare the stock of association rules, and find the popular stock
        then output the stock name
    else output the stock name
else
    Search the preference database and day interval database.

```

Figure 10. Selecting Rules

Step 2: Searching preference database and day interval database

We use ANN-document classification to decide whether an investor would buy or sell stocks or not. If we know an investor who will buy the stock next week, we find the preference of his buying stock in the past. And we select the most popular stocks in the buying preference data. If we know an investor who will sell the stock, we find the preference of his selling stock in the past. We also would use the final n trading data to find day interval of these

stocks. Finally, we select the most popular stock in the selling preference data and select the most popular stocks in the day interval database.

3.2.4 Individual Decision Process

Agents will use analytic information to select stocks which an individual investor may invest in the stock market. An agent will select the stock that an agent will invest according to the step of stocks selection. An agent computes the average trading amount of each trading, and he gets the stock quantity that an agent will invest in the virtual stock market. An agent uses ANN-document classification to decide whether an investor would buy or sell stocks or not. And the agent will take the action, buy or sell. Therefore, an agent issues a bid order showing his investment data. An agent will use his bid order to trade in the virtual stock market.

4. Experiment

In this section, we will introduce data preparation, experimental design, and experimental results.

4.1 Data Preparation

We would simulate an investor's behaviors, and we need to collect individual investors' investment data. But, it is not easy to obtain individual investor's investment data due to personal privacy. There is an opportunity to get simulated data. In a class, the instructor teaches students how to operate stock's relative softwares and let students simulate investing in stocks. These students work for semiconductor manufactures located at the Hsinchu Science-Base Industrial Park and they are familiar with stock market. They usually trade in the stock market. The instructor assists us to achieve students' investment data. So, we use these data as our experimental data.

4.2 Experimental Design

We will perform the experiment in the three sections. In first section, we use an investor's preference stocks to predict stocks' name that an investor will trade next week for each agent. In second section, we combine the individual preference stocks and the day interval to predict stocks' name that an investor will trade next week for each agent. In third section, we use individual preference stocks, the day interval, and association rules to predict stocks' name that an investor will trade next week for each agent. The result of first section and second section will be compared with the result of third section.

4.3 Experimental Result

We trained our agents with the simulated data. The result of experiment is shown in table 5. Prediction of 54.3 percent of investors is better than other predictions. Prediction of 14.3 percents of investors is not better than other prediction. Prediction of 11 percents of investors is

the same percentages as other predictions. There is a better result than other predictions in the whole.

We would integrate the result in the table 6. We have prediction rate of 37.3 percent in total average. There are some investors that almost never trade in the same stocks and we cannot predict their behaviors. If we drop these investors, we can get the prediction of 42.19 percent. It would have 40 percent of investors that has high prediction rate, 58.57%, and standard deviation is 0.0875. So, the accuracy rate of each investor is approximate. Therefore, our result has high representation. This is the best result for the experiments. There would be 11.4 percent investors of investors that we cannot predict and their prediction rate is zero. It is not the best prediction, but it also can be accepted.

Table 5. Integrated Prediction of Experiment

Investor	Preference Average	Day Interval Average	Density Average	Average in proposed Model
A001	20%	20%	20%	25%
A002	60%	80%	80%	60%
A005	35%	47.5%	47.5%	50%
A009	60%	60%	60%	60%
A011	20%	20%	20%	30%
A015	22.5%	22.5%	22.5%	20%
A017	50%	50%	50%	50%
A027	0	0	50%	50%
A035	45%	70%	70%	65%
A039	25%	22.5%	22.5%	25%
A043	42.5%	70%	70%	80%
A049	20%	20%	20%	25%
A051	31.4%	38.5%	38.5%	40%
A053	0	35%	35%	35%
A057	0	0	0	0
A061	45%	45%	45%	50%
A063	20%	10%	10%	20%
A069	0	0	20%	20%
A071	35%	35%	35%	40%
A073	20%	20%	20%	20%
A077	60%	60%	60%	65%
A081	12.5%	12.5%	12.5%	15%
A085	65%	65%	65%	65%
A093	0	10%	0	0
A095	32.5%	32.5%	32.5%	35%
A099	10%	32.5%	32.5%	33%
A101	0	0	0	0
A105	65%	65%	65%	40%
A109	40%	60%	60%	65%
A113	60%	60%	60%	60%
A115	25%	50%	50%	45%
A121	34.3%	48.6%	48.6%	50%
A123	12.5%	25%	25%	20%
A129	60%	40%	40%	50%
A131	0	0	0	0
Average	29.4%	35.1%	36.7%	37.3%

Table 6. Total Average Prediction Rate

	Percentage Data	Average Prediction	Standard Deviation	Notation
Prediction Rate between 50%-100%	40%	58.57%	0.0875	Good
Prediction Rate between 30%-50%	22.9%	37.25%	0.0452	
Prediction between 1-30%	25.7%	20.63%	0.02997	
Prediction Rate = 0	11.4%	0	0	Bad
Total average		37.3%		
Net Total Average*		42.19%		

* Net Total Average excludes the records with prediction rates = 0.

5. Conclusion

In this research, an agent would simulate an investor's behaviors according to his historical data. An agent can find the analytic information including association rules, preference stocks of an investor, and so on. The agent would use the analytic information to predict the stock that agent would trade next week. In experiment, we would have a satisfactory prediction rate.

This research would provide another perspective for investors to think how to predict the trend of stock. The stock price is decided by the supply and demand. Every investor plays an important role in the stock market. If we can predict the investors' behavior more accurately, we can predict the trend of weight stock price in Taiwan stock market better.

The following work can be continued in the future:
The following work can be continued in the future:

- (1) Classifying the Investors according to Investors' Intentions: Because there are so many investors in stock market, we can't simulate every investor. So we can classify the investors according to investors' intention. For example, some of investors will invest the other financial market including bonds, mutual funds and treasury bills etc for *avoiding risk*. Another investors will accept high risk because high-risk assets may offer higher expected returns than low-risk assets may offer. The others may not care risk of stocks at all, and they have their investment principles like herding, positive feedback trade strategy, negative feedback trade strategy and etc. So, we can classify investors to several classifications in the future work.
- (2) Getting Real Investors' Investment Data: Our data is simulated data, not real investors' investment data. There may be differences between simulated data and real investors' investment data. If we can collect the real investors' investment data, we can predict better. The followings can be considered in the future.
 - (a) Hundreds or Thousands of Investor's Data: If we can get these data, the model will perform better. If we can get hundreds of investors' data or thousands of investors' data, use our agents to simulate. It can produce plenty of trading data in the virtual stock market and agents can trade with the other agents. This architecture can

perform better than results that have done in our experiments.

- (b) Long Data Period: Experimental data should be collected for a period longer than six months, a year or a cycle period to train our agents. Enough data can train our agent that can be similar to real investors.

- (3) Getting Data from Three Major Institutional Investors (TMII): Major Institutional Investors have the huge capital that can be used to buy plenty of stock, and they have strong influence on the stock market. So, we can collect the investment data of TMII and use our agents to simulate the behaviors of TMII. We can modify our agents' architecture to fit with the behaviors of TMII. It can be a right direction for the future works.

References

- [1] 李治平, 2000, 「亞洲金融風暴與國外機構投資人之行為與影響」, 國立台大財務金融研究所碩士論文。
- [2] 林介勝, 2000, 「影響台灣電子股報酬之共通因素研究—以LISREL探討各影響因素間的可能關係」, 國立台大財務金融研究所碩士論文。
- [3] 周欣倫, 2000, 「台灣股票市場外匯暴露性之探討」, 國立東華大學國際經濟研究所碩士論文。
- [4] 馬武三, 2001, 「台灣上市公司各產業之類股報酬型態分析—亞洲金融風暴期間之驗證」, 國立交通大學經營管理研究所碩士論文。
- [5] 許長裕, 2001, 「多維度跨界性資料挖掘—以股市為例」, 國立中央大學資訊管理研究所碩士論文。
- [6] 謝榮記, 2000, 「台中地區個別投資人對訊息心理反應與加權股價指數關聯性之研究」, 私立朝陽科技大學財務金融研究所碩士論文。
- [7] Ashraf Azmy(1998), "SuperQuery; Data Mining for Everyone", <http://www.azmy.com>
- [8] B. Bouqata; A. Bensaid; R. Palliam, (1999), "Recurrent Artificial Neural Networks for Forecasting of Forward Interest Rates", Neural Networks, 1999, IJCNN 1999, International Joint Conference on Volume: 6, Pages: 3894–3897.
- [9] F.M. Werner; Bondt De (1998), "A Portrait of the Individual Investor", European Economic Review 42, Pages: 831-844.
- [10] H. Lu; J. Han; L. Feng (1998), "Stock Movement Prediction and N-Dimensional Inter-Transaction Association Rules." In Proceeding ACM SIGMOD Workshop on Research Issues on Data Mining and Knowledge Discovery, Pages: 12:1-12:7, Seattle, Washington, June 1998.
- [11] H. Iba; T. Sasaki, (1999), "Using Genetic Programming to Predict Financial Data", Evolutionary Computation, CEC 1999, Proceedings of the 1999 Congress, Pages: 245-251.
- [12] J. Roman; A. Jameel, (1996), "Backpropagation and Recurrent Neural Networks in Financial Analysis of Multiple Stock Market Returns.", System Sciences, 1996, Proceedings of the Twenty-Ninth Hawaii International Conference on Volume: 2, Pages: 454–460.
- [13] Kun-Po Chuang; Deng-Yiv Chiu (2001), "Predicting the Tendency of Stock Market in Taiwan Using Neural Network and Document Classification Technique", Chinese Institute of Industrial Engineers 2001, A16-1.
- [14] Lee, Charles M.C.; Radhakrishna, Balkrishna (2000), "Inferring Investor Behavior: Evidence from TORQ Data", Journal of Financial Markets, Volume: 3, Issue: 2 Pages: 83-111.
- [15] M. E. Malliaris, (1994), "Modeling The Behavior of the S&P 500 Index: A Neural Network Approach", Artificial Intelligence for Applications, 1994, Proceedings of the Tenth Conference, Pages: 86–90.
- [16] N. Baba; N. Inoue; H. Asakawa, (2000), "Utilization of Neural Networks and Gas for Nonstructuring Reliable Decision Support Systems to Deal with Stocks.", Neural Networks, 2000, IJCNN, Proceedings of the IEEE-INNS-ENNS International Joint Conference on Volume: 5, Pages: 111–116.
- [17] Odean; Terrance (1998), "Are Investors Reluctant to Realize Their Loses?", Journal of Finance 53, Pages: 1775-1798.
- [18] Rast, M. Fuzzy (1999), "Forecasting with Fuzzy Neural Networks: A Case Study in Stock Market Crash Situations", Information Processing Society, 1999. NAFIPS. 18th International Conference of the North American, Pages: 418–420.
- [19] Rakesh Agrawal; Ramakrishnan Srikant (1994), "Fast Algorithms for Mining Association Rules," In Proceeding 20th Int. Conf. Very Large Data Bases, VLDB, Pages: 478-499.
- [20] Shu-Heng Chen; Chia-Hsuan Yeh (1999), "Genetic Programming in the Agent-based Artificial Stock Market", Evolutionary Computation, 1999. CEC 99. Proceedings of the 1999 Congress on Volume: 2, Pages: 834-841.
- [21] Sorin Solomom (1999), "Behaviorly Realistic Simulations of Stock Market Traders with a Soul", Computer Physics Communications, Pages: 161-167.