Designing a Decision Support System Model for Stock Investment Strategy

Chai Chee Yong and Shakirah Mohd Taib

***Abstract*—Investors face the highest risks compared to other form of financial investments when they invest in stock market. Their involvement in stock trading is mostly based on speculation where their aim is more to obtaining capital gain rather than earning dividend as their investment return over the long run. Many people had tried to predict the movement of share prices and beat the market but no one can really accurately predict the movement of a particular share prices for company listed in the stock exchange. There has been attempt from Information Technology (IT) professionals to exploit the stock price prediction area through the Artificial Intelligence (AI) approach. This paper presents the continuous effort to explore stock price and trend prediction from finance perspective as well as from the combination of two major IT areas which are AI and Data Mining (DM). These areas have been explored to design a hybrid stock price prediction model with relevant techniques into the Stock Price Analysis and Prediction activities.**

***Index Terms*—Artificial Neural Network, Data Mining, Decision Support System Model, Stock price prediction.**

I. Introduction

The movements and advances in computing give implications to finance and economics today. New area of research known as financial engineering or computational finance is highly mathematical and cross-disciplinary field which relies on finance, mathematics and programming. Thus it enables financial analyst to analyze data efficiently. Much work can be done using mathematical model in spreadsheet program but the more demanding work needs one to write a computer program. This paper presents the study on designing a decision model for investment strategy by utilizing financial methods and computation techniques.

The movement of the share price is unpredictable. There was a study done to estimate and predict the movement of share prices by using statistics, math and forecasting methods based on the historical share price data and this named as Technical Analysis (TA). According to Rockefeller [8], TA has been around for 100 years and before the existence of computers, it was done only by professionals who had access to real time price on the stock exchange.

Manuscript received August 15, 2009. Chai Chee Yong was with Universiti Teknologi Petronas, MALAYSIA. He is now with the Department of Facilities Engineering, Development Division, Petronas Carigali Sdn. Bhd., Petronas Twin Towers, Kuala Lumpur, MALAYSIA (email: chai_cheeyong@petronas.com.my).

Shakirah Mohd Taib is with Department of Computer and Information Sciences, Universiti Teknologi Petronas, Tronoh 31750 Perak, MALAYSIA. (phone: 605-3687497, fax: 605-3656180, email: shakita@petronas.com.my).

There are tools and software created to automatically illustrates and generate various graphs and perform various complex calculations in the TA. These tools eliminate the haste of Financial Analyst in performing TA manually by paper and pen in the past. With computer capabilities, now investors also perform TA themselves with the help of appropriate software which help them in making their trading and investment decision.

Other than assisting and making the TA process more efficient and effective, IT field also brought a totally new perspective and thinking into stock analysis through the researches and studies on the implementation of AI and DM concepts in the stock analysis. For example, there were studies conducted on the combining the Candlestick analysis (One form of the TA) with the Genetic Algorithms in stock analysis [2]. Besides that, Quah [7] says that with the significant advancement in the field of Artificial Neural Network (ANN), the generalization ability of the ANN is able to create an effective and efficient tool for stock selection.

High Risk Investment

Investors face the highest risks compared to other form of financial investments such as bond, treasury bill and fixed deposit when they invest in the stock market. Stock price fluctuates almost every second and most people see the share price movement as unpredictable or in other term “random walk”. Some investors involve in this activity mostly based on speculation where their aim is more to obtaining capital gain rather than earning dividend as their investment return over the long run. Investing based on speculation sometimes caused many individual investors had their ‘hand burn’ where they lose their initial investment. Kamich [5] defined that stock market is a private or public market for the trading of company stocks at an agreed price. One of the main objectives of the company to be listed in the stock market is to raise capital to fund future expansion.

If the investors able to purchase the correct share when the price at its lowest range and it can be sold at a better price later, they will be able to earn big chunk of money from the stock market. Of course, this is not easy to be done as no one able to predict the future, as well as the movement of the share price accurately.

II. Finance Theories And Analysis In Stock Price Prediction

The attempt of analyzing the stock performance using financial theories can be traced back up to late 1800s, where the oldest approach to common stock selection and price

prediction, TA is created and used [4]. Edward and Magee [3], stated that throughout the years of stock market study, two distinct schools of thought has arisen each with radically different methods of obtaining answers on what share to buy and when to buy and sell the share. These two different thought in stock analysis are Fundamental Analysis (FA) and TA.

Jones [4] defined FA as a method of security valuation which involves analyzing basic financial variables such as sales, profit margin, depreciation, tax rate, sources of financing, asset utilization and other factors. Edward and Magee [3] notify that FA relies heavily on statistics and people who performing FA will be looking through the auditor's reports, the profit-and-loss statement, balance sheet, dividend records, and policies of the companies whose shares is under their observation. They will also analyze business activities and daily news to estimate the company's future business condition. The investors who use FA will purchase stocks that are viewed as underpriced by the analysts who believe that in the future, the share price will increase to the level that it should be as computed in FA.

Another theory, TA is defined as the study of how securities prices behave and how to exploit that information to make money while avoiding losses [8]. TA focuses on the share price to assess and evaluate the demand and the supply for the shares based on the market price itself and do not listen to chatter about securities. Jones [4] says that technical analyst believes that the market itself is its own best source of data. It is believed that all the investors' reactions towards all the information regarding the security already embedded in the share price.

According to Rockefeller [8], TA works because people constantly repeat behavior under similar circumstances and TA is a forecasting method where it uses past and current behavior (price) to predict future behavior (price). Concurrent to this behavior repetition concept, Edward and Magee [3] noted that share price move in trends and they tend to continue until something happens that will change the supply and demand balance. As a conclusion, TA charts and analyzes the historical share prices to reveal the pattern, formation and investor's behavior and interpret it to predict the possible future share price trend.

III. Artificial Intelligence (AI) and D ata Mining (DM) Implementation in Stock Price Prediction

Generally, prediction of stock is a very difficult task as it behaves like 'random walk' process and the prediction might be run off due to some unexpected news that have direct impacts on the respective company. The obvious complexity of the problem paved the way for the importance of intelligent prediction paradigms [1]. One of the most popular researches in this field is on the usage and implementation of ANN in the stock market analysis. Kamruzzaman, Begg and Sarker [6] said that stock analysis has been one of the most important applications of neural networks in finance where numerous researches had been done on ANN in stock market index prediction, stock performance/selection prediction, and stock risk prediction. Quah [7] further explained that ANN generalization ability is able to infer the characteristics of performing stocks from the historical pattern. ANN is used to discover the non-linear relationship between financial variables and stock price which cannot be done by the computerized rule-based expert system. In stock prediction using ANN, the input will be historical data such financial variable and the out of it will be performance of stock.

TA is about find the pattern in the historical data to predict the future price movement, therefore researchers found DM is a suitable technique to reveal the hidden pattern in the historical data and predict the price trend. Both Y.Y.Shi and Z.K.Shi [9] has conducted a study on clustering technique (one of DM techniques) in stock prediction. In their studies, they identified that clustering technique is one of the efficient methods for stock investment analysis and they also noticed the weakness of this technique in stock prediction where constructed cluster might not completely reflecting the characteristics of the stock. In the study, the idea of rules mining of stock market using the statistical pattern recognition is proposed in order to overcome the weakness in clustering techniques. From techno-index in TA, Shi et al. [9] developed an algorithm of fuzzy kern clustering and tested this idea on Shanghai and Shenzhen Stock Market historical data since 1997. The result of the study clearly indicate that statistic rules in essential trends of the stock market and implied rules can be recognized to a certain degree through the construction of proper clustering algorithm [9].

The authors' proposed model for stock price prediction is based on Decision Support System (DSS) model by Turban et al. [10]. AI and DM are included in one of the subsystem which is Knowledge based subsystem.

IV. DSS Model For Stock Investment Strategy

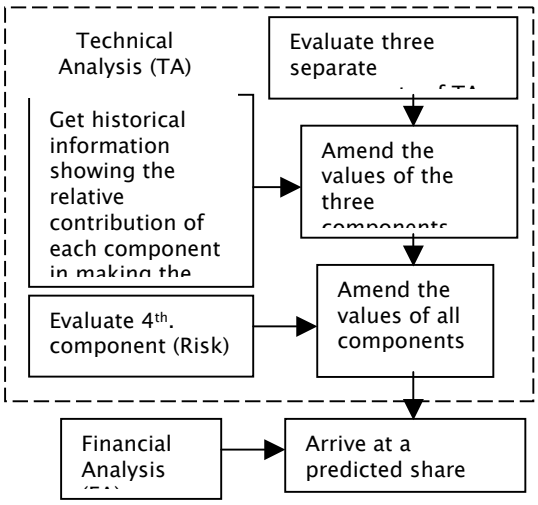

Fig. 1. The Flow of Stock Investment Strategy Decision Support System

The authors have designed DSS model and architecture with some minor modification for stock investment strategy. This model can be used to adapt AI and DM elements into the system and utilizing TA and FA methods. The flow of the system is shown in Fig. 1.

V. Architecture of Stock Investment Strategy Decision Support System

Based architecture that shown in Fig. 2, the flow of the system activities can be divided into two which reflect the

two main processes of the system. The two main processes of the system are analyzing the raw financial data from TA concept (TA process) and analyzing the financial data from FA concept (AI process) with both process have the same aim of predicting stock price and trends. The first flow of activities will be started from database where it provides raw financial data to DM component in knowledge-based subsystem, the output from DM component together with raw financial data from database will be sent to TA component in model management subsystem for further processing and the output of TA component will be the input for Recommendation Component. The second flow of activities will be from raw financial data to AI component in knowledge-based subsystem. Finally the output from AI component will be sent to Recommendation Component in model management subsystem.

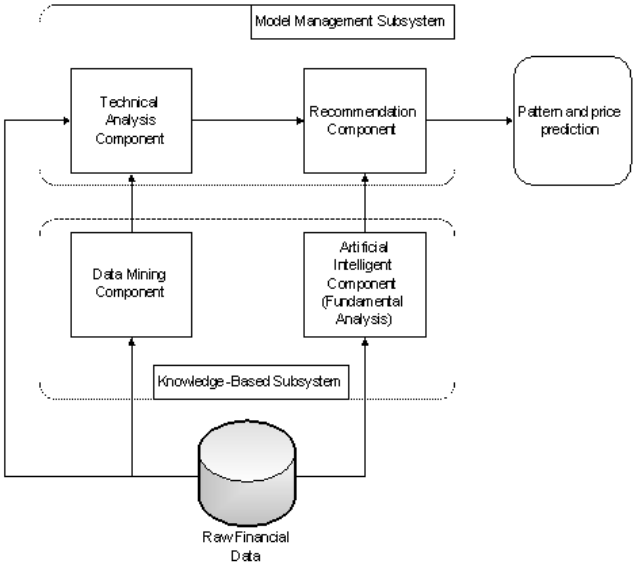

Fig 2. Architecture of Stock Investment Strategy Decision Support System

This component focuses on illustrating historical stock price data into the form of graph as well as other TA indicators to interpret investors and market sentiment in the effort of predicting the stock price movement and trend. Among the most popular charting techniques that being used in TA are Line Graph, Open-High-Low-Close chart and Japanese Candlestick chart. The mechanism proposed here is the system that able to indentify and recognized pattern directly from the raw financial data using rules generated from the underlying reasoning and interpretation in various TA techniques. In achieving this, TA component of the system will use rule-based TA model that is build up from underlying rules in various TA charting techniques.

A. DM Component

Rockefeller [8] highlighted that the same TA method would not work in the same stock analysis all the time. The same TA method that proven working on a particular stock does not necessary mean the same method will also work in other stocks. Therefore, instead of identifying and interpreting patterns from the raw financial based on the rules generated from one TA method only, the system will identify the patterns based on underlying rules in various TA method. This will reduce the dependency on a single TA method in analyzing a particular stock as well as provide better reasoning and analysis capability in estimating future trend of the stock as we compare the same pattern using different rules from different techniques which give higher accuracy of the prediction. This system will use the association rules mining in determining the accuracy and effectiveness of TA methods in predicting price movement and trend for a particular stock. Therefore, the flow of activities and information in analyzing the stock in TA concept will start from database and end at recommendation component as illustrated in Fig. 3.

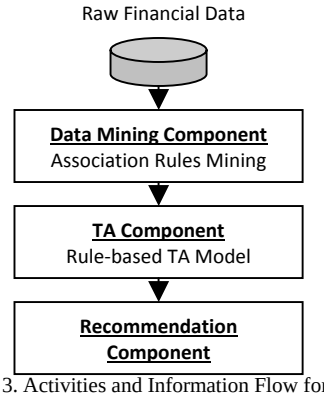

Fig. 3. Activities and Information Flow for TA process

The first phase of TA process will be the financial raw data being mined using the association rules mining. With association rules mining, all the underlying rules that reflecting the interpretation of TA in stock price and trend prediction will be mined by calculating their respective support and confidence values for a particular stock using the historical price of the stock. By using association rules mining, the system will be able to identify which association rules that actually works on that particular stock based on the historical data stored in the database based on the confidence and support. Those confidence and support value together with raw financial data will be the input for TA component.

In the rule-based TA model, the recent historical stock price movement and its related information will be analyzed by all association rules to identify pattern. If recent historical price information triggered more than one association rules, the output from the association rules will be taken into consideration based on weight that are determined based on the support and confidence level. If the triggered association rules have a confidence level that is more than 0.50 and it's the highest among other triggered association rules, that association rules will have the highest weight. On the other hand, if an association rules has the confidence level below than 0.50, it means that although the association has been triggered before in the past, but the actual result is different from the expected result based on the association rules and the rule has wrongly predict the result more than 50% based on historical data. This indicates that the association rules does not work well with the stock, hence this association rules will receive lower weight. Besides that, if an association rule is triggered has the support value of 0, it means the rule never being triggered before in the past and the system will assign a new confidence value of 0.5 for the associate rule

because since the rule never triggered before hence the rule will have a 50% chance that it will predict correctly. If all the triggered associate rules have confidence level below than 0.50 the system will notify the user that prediction might have higher risk to be incorrect and the system would still provide the stock trend prediction. If from all association rules, the predicted stock trend is a mixture of up and down, the system will compute the likelihood of the expected trend to be happening based on the expected trend that has the highest accumulated weight which this process will be done by the Inference Engine in the TA Component.

B. TA Component

There are six main sub-components included in TA component as shown in Fig. 4.

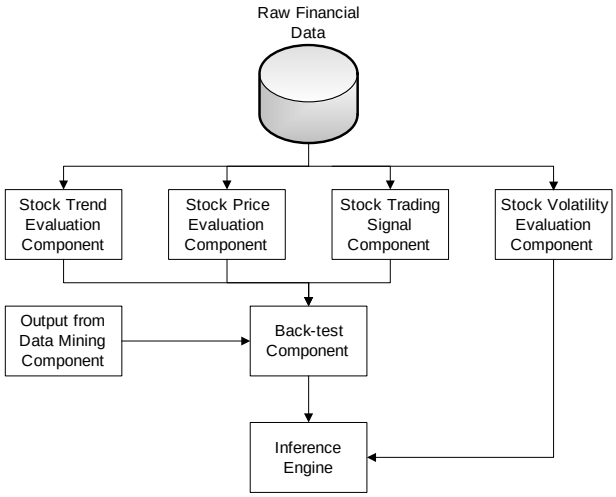

Fig. 4. Processes in TA Component

Stock Trend Evaluation Component

The first process would be 'Stock Trend Evaluation Component' where in this process, latest two weeks historical data is used to run through the underlying rules derived from the Open-High-Low-Close chart and Japanese Candlestick chart which is used to detect the pattern, gauge the investor's sentiment on the stock and discover the current trend of the particular stock. This component has the most association rules compared to other components. The output for this process will be the trend of the particular stock in the near future which will be uptrend, downtrend, indecision or no trend.

Stock Price Evaluation Component

This component will evaluate either the stock has been overvalue or undervalued using the Williams %R Indicator and Relative Strength Index. Williams %R indicator is developed with the aim to assess either the particular stock is being overbought or oversold. There will be two outputs from two indicators which the result might be contradicted with one to another. This contradiction issue will be sort out in back-test component. The output for this process is the evaluation on the stock either it's overvalued, undervalue or reasonable.

Stock Trading Signal Component

Another component is used to identify any buying or selling signal generated based on TA for that particular stock. There will be six TA methods being used to trace those trading signals which are Weighted Multiple moving Average, Momentum, Engulfing, Doji, Long Shadow and Gap. Association rules in this component do not being triggered all the time. There are high tendency that none of the association rules in this component is being triggered most of the time because trend reversal itself does not happen all the time and this component focuses on tracing trend reversal that lead to generation of trading signal. The output of this component will be trading signal and forecasted stock trend.

Stock Volatility Evaluation Component

The forth component in the TA component is used to evaluate the risks and the noise level of a particular stock. Volatility is a measure of price variation which is directly linked to risk level and noise level. Stock with high volatility means trading is riskier but has more profit potential, while low volatility means less immediate risk. Therefore, this component serves to evaluate the stock's volatility level and provide a warning if the volatility level is high because high volatility will cause the system's forecast and prediction less accurate due to the high noise level in the stock movement.

Back-test Component

Back-test component is one of the main parts in the TA component. Output from the various components explained earlier might have contradicted forecasts that even occurred within a particular component itself. Back-test component will weight each output from all of the TA indicators based on the performance of those indicators for that particular stock in the past where indicators that performed the best in the past will be given the highest weight. This component receives an important input from the DM component which is the support and confidence values of each of the association rules used in the three components explained earlier based on the stock historical price. All the outputs from Stock Trend Evaluation Component, Stock Price Evaluation Component and Stock Trading Signal Component together with its respective weight are sent to Inference Engine.

Inference Engine

Inference Engine will produce the final decision on the various different forecasts and predictions generated by other respective components. Inference engine works based on all the forecasts and their respective weights. If there are contradict forecasts incurred, those forecasts with higher weight will be chosen as the decision. All of the outputs from four other components will be included because each of the indicators in all the components are complementary to each other. The verifying a forecast done by indicators is very important to reduce the unprofitable trade resulted from false trading signal generated by a particular indicator. A final recommendation will be generated by Inference Engine that includes Trading Signal (Trading Signal is not generated everyday), Forecasted Stock Trend, Stock Price Evaluation and risks level. Provided together is the possibility level for the generated trading signal, forecasted stock trend and stock price evaluation which is derived from the confidence level from each indicators.

C. AI Component

TA techniques exclusively cannot recognize the non-linear relationship between financial variables and stock price. Therefore, AI component that particularly using ANN will be included in the system. This ANN is based on the concepts in FA. The inputs of ANN which are extracted from raw financial data, the system would be able to discover the non-linear relationship between financial variables and stock price. Furthermore, by integrating FA concept into ANN will enable the system to compare and back-test the stock price and trend prediction done by TA to achieve higher prediction accuracy than only use TA or FA in ANN alone.

The inputs for the ANN will be from company and economic factors. Inputs from company factor will be Return on Equity, Historical P/E Ratio, Prospective P/E ratio, Cash flow yield, Payout Ratio, Operating Margin, Current Ratio and Debt to Capital at book. As for economic factor, the inputs will be Inflation, Interest rate, GDP Growth. Hence, the output from the ANN in the AI component will be the forecasted stock price in the next four month. The forecasted stock price will be in the format of percentage of increase or decrease such as less than 5%, between 5 and 10% or above 10% from the base stock price which is the price of the day the forecast is computed by ANN.

D. Recommendation Component

Both outputs from TA component and AI component are the inputs for the Recommendation component. Forecasted Stock Trend, Stock Price Evaluation, risks level and Trading Signal are the inputs obtained from TA component while four-month target stock market price is the input obtained from AI component. The main reason that the output from AI does not undergo the back-test component like in TA component is because the ANN itself has been trained using historical financial data before producing the four month target stock price prediction. This component task is to assess either the four month target stock price is achievable or not based on the predicted near future stock trend, evaluation of the current stock price as well as trading signals. Hence, it actually integrates both long term and short term forecast for the stock and present it to the user. The final output from this component will be the predicted near future stock trend to enable user to know the current and possible trend of the stock with the historical accuracy that the prediction accurate, stock price evaluation on either the stock price is reasonable, undervalued or overvalued and with the historical accuracy, risks level to alert users on the possibility that the system's prediction is wrong due to high level of volatility, trading signal (buy or sell signal) if there are trading signal generated for that stock and with historical accuracy and finally the four month targeted stock price.

The recommendation will be provided together with all the outputs given by various TA indicators and their historical accuracy in presenting the fine-grain information to the users for their use in making trading decision making concurrent with the role of the system as a Decision Support System. The system does not dictate, but recommend. Any final decision is based on the user's judgment and evaluation themselves.

VI. Conclusion

A study on the FA and TA has been done in order to understand how stock price prediction works in Finance world as well as to create the TA Model and ANN based on FA. The further study on DSS model and architecture is done to ensure Stock Investment Strategy Decision Support System follows the actual basic framework and model of a DSS. The combination of AI and DM is applied in this hybrid model particularly in the knowledge-based subsystem to generate better analysis of the historical data. The output from the system will be the result analysis and recommendation to the user. The analysis will be generated by the TA component and AI component where it will be presented to the users to aid them in their trading decision making with the presentation of this fine grained information. However, it will only present the information and outputs derived from various TA and FA techniques used in the system. The recommendation will judge, evaluate and provide trading recommendation based on the same outputs that are presented to the user in the analysis.

References

[1] Abraham A, Nath B, and Mahanti P.K, "Hybrid Intelligent Systems for Stock Market Analysis," Springer-Verlag Berlin Heidelberg, 2001.
[2] Belford P, "Candlestick Stock Analysis with Genetic Algorithms," ACM Digital Library, 2006.
[3] Edward R.D, and Magee J, TA of Stock Trends, 8th ed., St. Lucie Press, 2001.
[4] Jones C.P, Investments, 14th ed., John Wiley & Sons, 2007.
[5] Kamich B.M, How TA Works, New York Institute of Finance, 2003.
[6] Kamruzzaman J, Begg R, and Sarker R, Artificial Neural Networks in Finance and Manufacturing, Idea Group Publishing, 2006.
[7] Quah T.S, "Improving Returns on Stock Investment through Neural Network Selection," Nanyang Technical University, 2006.
[8] Rockefeller B, TA for Dummies, Wiley Publishing, 2004.
[9] Shi Y, and Shi Z, "Clustering Based Stocks Recognition," ACM Digital Library, 2006.
[10] Turban E, Aronson J.E, Liang T.P, and Sharda R, Decision Support System and Intelligence Systems, Pearson Education, 2007.