Knowledge-Based Systems 23 (2010) 626–633

Contents lists available at ScienceDirect

Knowledge-Based Systems

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

MISMIS – A comprehensive decision support system for stock market investment

Vincent Cho *

Department of Management and Marketing, The Hong Kong Polytechnic University, Hung Hom, Hong Kong

ARTICLE INFO

Article history:
Received 18 August 2008
Received in revised form 14 April 2010
Accepted 14 April 2010
Available online 20 April 2010

Keywords:
Stock market intelligent trading system
Decision support system
Time series analysis
Hang Seng index
Multi-level system

ABSTRACT

Nowadays, stock market is becoming a popular investment platform for both institutional and individual investors. The current financial information systems serve to provide latest information. However, they lack sophisticated analytical tools. This paper proposes a new architecture for financial information systems. The developed prototype is entitled as the Multi-level and Interactive Stock Market Investment System (MISMIS). It is specially designed for investors to build their financial models to forecast stock price and index. The performance of the financial models can be evaluated on a virtual trading platform. There are other features in MISMIS that are tailor-made to handle financial data; these include synchronized time frame, time series prediction techniques, preprocessing and transformation functions, multi-level modeling and interactive user interface. To illustrate the capability of MISMIS, we have evaluated strategies of trading the future options of Hang Seng Index (HSI). We find that historical HSI, Dow Jones Index, property price index, retailing sales figure, prime lending rate, and consumer price index in Hong Kong are essential factors affecting the performance of the trading of HSI's future option. Also there are some feedbacks from the in-depth interviews of six financial consultant upon how they perceived the prototype MISMIS.

1. Introduction

In the stock market, many investors attempt to earn extra gain or to protect their savings from shrinking due to inflation. However, the existing financial information systems (FIS) are too much focusing on dissemination of information such as financial news or latest stock value. They may equip with simple stock trading heuristics. For example, a buy signal is horned if the curve of the 7-days moving average crosses and is above the 50-days moving average. They are indeed lack of analytical tools to do sophisticated analyses. Nowadays, some statistical tools such as Clementine have armed with advanced techniques particularly on time series forecasting such as Autoregressive Integrated Moving Average (ARIMA) and artificial neural network. Nevertheless, they provide a platform for general applications and do not have a specific financial domain. For example, they do not provide adequate support on virtual trading environment for investors to try out their trading strategies. In order to facilitate investors for their decision making in stock trading, this paper proposes a new architecture of FIS which integrates the advanced time series forecasting techniques with a virtual trading platform under a GUI environment. This development would fit into the trend of using advanced techniques such as data mining method and knowledge-based decision supporting tool to predict or evaluate business performance [32,33,35].

This paper describes the prototype of a Multi-level and Interactive Stock Market Investment System (MISMIS) which combines different area – financial economics, prediction techniques, and dynamical systems theory, in handling financial data. MISMIS enables investors to build their prediction models interactively by dragging various pre-defined functions. The prediction outcomes can also be fed into another prediction model as input. Let us say, the Hang Seng Index can be predicted by feeding the prediction outcomes of those composite stock prices. It may result in having a model that produces more reliable predictions. The prediction will subsequently be fitted into a trading strategy and be evaluated on a virtual trading platform. Performance metrics will be reported to users who would adjust the model parameters for further enhancement on the prediction model. Hence, MISMIS assists on the decision making of an investor in the stock trading. This prototype MISMIS has also been evaluated by some consultants in financial investment. They found this prototype once being launched would be much welcome by most individual investors.

The presentation of this paper is as follows. We first describe the theoretical base of financial forecasting. There is a section to elaborate on the system feature of MISMIS, then followed by a case study of the stock market in Hong Kong. In the case study, we would explore the underlying factors that have significant influence on the prediction of Hang Seng Index and compare different models on their performance of trading HSI's future option. Also there is a section on how investment consultant feedback to this prototype. In the final section, we discuss the limitations of MISMIS and conclude this paper.

* Tel.: +852 27666339; fax: +852 27743679.
E-mail address: msvcho@polyu.edu.hk

0950-7051/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.knosys.2010.04.009

2. Related framework for financial forecasting

This section reviews on the fundamentals of why predicting financial marketing is sensible. According to Fama [12], an 'efficient' market, securities will be appropriately priced upon all available information. If a market is efficient, no information or analysis can be expected to result in outperformance of an appropriate benchmark. Fama believed that an 'efficient' market is defined as a market where there are large numbers of rational, profit-maximizers actively competing, with each trying to predict future market values of individual securities, and where important current information is almost freely available to all participants. In other words, in an efficient market at any point in time the actual price of a security will be a good estimate of its intrinsic value. Most importantly, Efficient Market Hypothesis [12] states that at any given time, security prices fully reflect all available information. It implies that stock price value could be calculated by suitable indictors. Thus prediction of stock market only makes sense if the market is an efficient one. In this regard, building a financial information system would help investors in their decision making.

For a less efficient market, we will usually find a delay in response to the market information, but still will be settled in a longer term perspective. Ritchie [26], based on fundamental analysis, designed to assess important economic, political, and social dynamics as well as company data, and suggested their potential impact on the capital markets and security values. This approach attempts to determine the expected return of an investment given a certain amount of financial risk including purchasing power risk, interest rate risk and business risk. It involves an in-depth study of the economy and its implications for industries and companies, valuation of security based on its future earning power and dividend paying expectations utilizing the valuation approaches. Both qualitative and quantitative (financial statements including the income statement and balance sheet) information and analyses are used for stock value estimation. The derived value of the stock is then compared with its current market value to determine whether to invest on this stock or not. In brief, fundamental analysis is suitable for long term forecasting and analysis. It could be used to predict the overall trend of movement in future.

Along with the short term movement, Murphy [23] attempted to study the market action primarily through the use of charts. The technical approach is based on three assumptions: market action discounts everything; prices move in trend; and history repeats itself. The types of chart available include daily bar chart, long range weekly and monthly chart, shorter term intra-day chart, major reversal patterns like the head and shoulders pattern and the double top or bottom, moving averages, oscillators and contrary opinion, point and figure chart, etc. In short, technical analysis is mainly for short term forecasting while fundamental analysis is for long term prediction. The prototype MISMIS can handle both short term and long term depending on what data the user input into the system and the time frame for his/her prediction. For short term forecasting, a user may use the past historical stock price to predict the future stock price using Autoregressive Integrated Moving Average (ARIMA) or artificial neural network (ANN). For long term prospective, a user may predict the stock index movement by considering the underlying fundamental economic data using regression or ANN.

3. The prototype – MISMIS

We propose a prototype of the Multi-level and Interactive Stock Market Investment System (MISMIS) which can provide both flexibility and interactivity to formulate a prediction model on time series data. Its interface looks like Clementine which is an advanced module of SPSS and user can build their prediction model by putting different functional blocks together.

As shown in Fig. 1, the MISMIS have seven essential features, namely specific transformation functions, multi-level modeling, flexible time frame alignment, sophisticated tools for financial prediction, virtual trading environment, tabulated and graphical outputs, and interactive user interface. The following sections describe these features in details.

3.1. Data and time frame specification

For the processing, all data are time-stamped and synchronized. The user can specify the time unit, e.g. 10 min, 1 h, 1 day, 1 week or 1 month for the synchronization of various time series data. Data with time-stamp in a longer period are interpolated according to some pre-defined functions. For instance, if a user wants to predict the closing Hang Seng Index of tomorrow, the time unit for this application is in terms of days, the other time series data are synchronized according to this time unit. Let us say, the interest rate, once unchanged within a period of half a year, is interpolated to be the same value at different time points within the period. On the other hand, if the user wants to analyze the short term fluctuation in the stock market, for example the movement of stock price for every 10 min, then the time unit for this application is measured in 10-min and every time series data is synchronized according to this time unit.

3.2. Specific transformation function for financial data preprocessing

Transformation functions, which are specifically for economic data manipulation, such as logarithm or time lag, are developed in MISMIS for handling the financial data. The followings highlight those essential transformation functions.

Log – logarithm function is available for some models that have errors proportional to variables, or for transforming a multiplicative model to an additive model. It also applies to data with long term exponential trend. For examples, log(liquidity) is used for economic indicator forecasting [8].

Difference – the difference of the previous value at time (t–k) with respect to the current value at time (t). This is used for making a time series to become stationary. Non-stationary time series (variance depends on time) often cause problems in prediction. Differencing is a common method to solve such kind of problems. For example, change in the monthly average of the spot price of crude oil per barrel, which is a stationary time series, is measured by differencing the monthly average of spot prices [21]. These changes of monthly average are helpful to predict the trend of the economic movement.

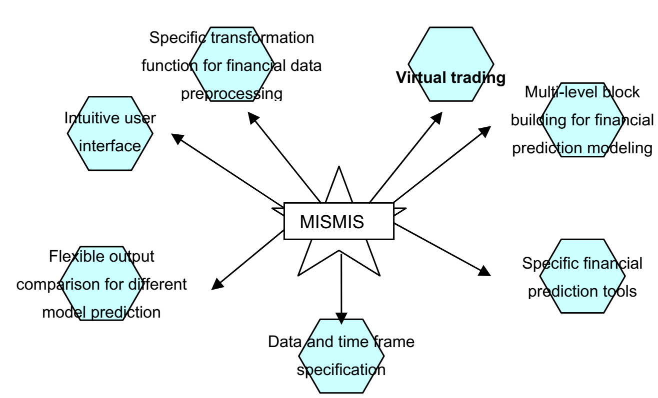

Fig. 1. The architectural diagram of MISMIS.

Normalization – some financial data have a big difference in scale. For example, Hang Seng Index (HSI) ranges from 17,000 to 31,000 while deposit rate ranges from 0.00 to 0.05. If these data used directly, the indicators with large value, e.g. HSI, may dominate the prediction. Usually these data need to be normalized in same scale before analysis so as to have the same weighting for forecasting. All value will be in a range of 0–1 after normalization. It helps to analyze the relative importance of various indicators. The normalization is formulated as follows:

$$X_{normalized} = \frac{(X_i - X_{\min})}{(X_{\max} - X_{\min})}$$

Time lagging – the time series is shifted backward/forward in time frame. Time lagging is mainly used for leading indicators and lagging indicators, which they either signal future events or follow past events. To adjust the time point that these data are referring in a model, time shifting is a necessary step in data preparation. For instance, money supply (M2) is commonly considered as a leading indicator of economy [28]. The parameter for the time lag can be controlled by double click the lag icon and the user can set the parameter (the number of time unit to be shifted) in a pre-defined dialogue.

Percentage change – for financial analysis, the percentage of change is much more important than the real value. In order to compare the relative fluctuation of stock markets, percentage change is a good indicator for comparison. For instance, the daily percentage change of Dow Jones Index on 31 January 2007 was 0.78%. This value would have an impact on other stock markets and as a result, the percentage change of Hang Seng Index would be speculated to be of similar order on 1 February 2007. The percentage change formula is as follows:

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

3.3. Advanced prediction techniques

There is a growing evidence that macroeconomic series such as money flow contain non-linearities [5] but linear models such as Autoregressive Integrated Moving Average (ARIMA) [6] and regression are still widely used as they provide higher explanation power and are statistical reliable. ARIMA models are used for homoskedastic stationary series. It is assumed that the expected value of squared errors is the same across the entire series. This assumption is called homoskedasticity. However, most of financial series do not meet this assumption, the expected values of squared errors may be larger at some ranges than others, and this is called heteroskedasticity. Autoregressive Conditional Heteroskedasticity (ARCH) [11] and Generalized Autoregressive Conditional Heteroskedasticity (GARCH) are applicable to heteroskedastic series with high volatilities. In the other vein, neural network models can be used to solve non-linear problems.

In contrast to traditional statistical methods, which are usually linear-based, artificial neural networks (ANNs) cater for non-linear systems. They are computing techniques that were inspired by the function of nerve cells in a brain. These networks are composed of many parallel, interconnected computing units. Each of these units performs a few simple operations and communicates the results to its neighbouring units. ANNs are also very effective in learning cases that contain noisy, incomplete or even contradictory data. The ability to learn and the capability to handle imprecise data make ANNs very suitable for handling financial and business information. However, a main limitation of ANNs is that they lack explanatory capabilities. ANNs do not provide users with reasons how particular conclusions are deducted. Past studies favor neural network in the sense it gives slightly better prediction accuracy [14,18]. Moreover, increasingly people believe that finance market price movement is highly non-linear and dynamic [1,4,7]. Hence, the MISMIS encloses the above essential techniques.

Parameter estimates: for each prediction technique, there is a pre-defined dialogue for user to specify their ways for estimating parameters. For instance, coefficients of a regression model can be estimated by various techniques, like least square method, maximum likelihood, Monte Carlo simulation, etc. On the other hand, the parameters in the ARMIA time series model are estimated using the following algorithm as highlighted in Fig. 2.

3.4. Evaluation and output presentation

A user can plot the predicted outcome graphically and can evaluate the prediction results using the following evaluation matrices.

Accuracy function: the prediction accuracy on both the training and testing data will be reported. Among the error functions, the most common ones are root mean square error (RMSE) and mean absolute percentage error (MAPE) [13,pp. 13–14].

RMSE is defined as

$$RMSE = \sqrt{\frac{\sum e_t^2}{n}}$$

MAPE is defined as

$$MAPE = \frac{\sum |e_t|/Y_t}{n}$$

where e_t is the forecast error in the time period t; Y_t is the actual value in time period t; and n is the number of forecast observations in the estimation period.

Prediction power is one of important issues that should be considered in the model building process. In common, a prediction model is evaluated using the testing data. Once the testing result is not satisfactory, the model would be fine-tuned with different training heuristics. The process is iterative until a satisfactory model is built. Other than simply dividing data into two static portions, Jackknife or Window Shifting techniques may also be available in MISMIS for model training and testing. Different prediction models are generated in which each trial uses different portions of data for testing. In MISMIS, a user can easily reiterate the process by adjusting the training and testing data set; changing the prediction algorithm or adding/dropping factors for prediction. As a result, a user can understand which factors, algorithms or training timeframe would generate a more reliable prediction.

The graphical output functions: the predicted results and the collected data can be plotted graphically. This would give users a visual outcome on how the prediction performs. Even without doing any forecasting, a user can plot the charts based on the data collected for some preliminary analysis. For examples, head and shoulder pattern and the double tops or bottoms are meaningful pattern to technical analysts but these patterns are not easily analyzed by quantitative models. In MISMIS, there are other technical charts such as moving averages, oscillators and contrary opinion, point and figure chart that are useful for preliminary study.

3.5. Multi-level modelling

Different prediction models can be combined by either aggregating them together or making a consensus judgement. The consensus function enables the users to specify how to come up with a final agreement from different predictions of a variable. On the other hand, the predictions of several different variables would also be used to formulate the value of another variable, e.g. a user may like to predict the Hang Seng Index by aggregating the predictions of its composite stock prices.

Aggregated function – the derived outputs of different prediction models can be aggregated to produce an output. For instance, when making the prediction of Hang Seng Index (HSI), it would be estimated by the weighted average of the composite indices.

Consensus function – when there are several predictions on HSI, e.g. a user applies different techniques to predict HSI, it is better to have a consensus among different predictions. This function is available in MISMIS and the mechanism behind is using regression, or using neural network with reinforcement technique. For the regression, we propose a combining model as follows:

$$f_{combine} = \sum_{i=1}^{k} w_i f_i$$

where w_i is the weight assigned to classification method i, $\sum_i^k$

such as neural networks, normalization to scale between -1 and 1 is usually required.

The time unit of this case study was in terms of days. For those spot data, say stock price and crude oil price, changes in every minute, closing value was recorded. On the grounds that some economic data were released monthly or quarterly, their daily values were the same within the whole period. For example, CPI, released quarterly by the Hong Kong government, would be of the same value for the whole quarter. In other words, CPI value in the first quarter of year 2003 was $-$

Table 1
Estimates of the prediction models (ARIMA, ANN, regression).

	ARIMA Coefficient estimates	ANN Relative importance	Regression Coefficient estimates
Constant	−**3491.974****		−**643.6****
MA1	−**1.018*****		
MA2	−**0.494*****		
Seasonal AR1	**0.790*****		
Seasonal MA1	**0.237*****		
lag1_HSI		**0.836069**	**0.942*****
lag1_DJI	**0.766*****	**0.0746794**	**0.0603*****
Property price index	**52.136*****	**0.195885**	4.48
Retailing sales	**10.741***	**0.137019**	0.261
Prime lending rate	**931.289*****	**0.107814**	**65.801*****
CPI in Hong Kong	**-214.214*****	**0.0832492**	-17.304
Gold price	−0.441	0.0565497	0.677
Crude oil price	−1.208	0.0558185	−2.133
Unemployment rate	29.166	0.0550763	0.003
Euro exchange rate	68.821	0.055073	−8.87
JPY exchange rate	17,426.075	0.0511798	0.003
GDP in HK	0.000	0.0173758	0.0004

* For $p < 0.05$.
** For $p < 0.01$.
*** For $p < 0.001$.

important in the ANN model. From the pattern, it is shown that historical HSI, Dow Jones Index, property price index, retailing sales figure, prime lending rate, and CPI in Hong Kong are essential factors affecting HSI. In contrast, the less important factor would be the GDP in Hong Kong.

On the other hand, the predicted HSI was further evaluated in terms of its accuracy and its performance in a virtual trading. The related model was formed by dragging the two icons: accuracy and virtual trading to the icon of ARIMA as shown in Fig. 4. The trained model was then fitted into the testing data and the result of accuracy is shown in Table 2. In the virtual trading, we have specified the trading rules as follows:

$(\text{HSI}_{\text{forecasting model},t} - \text{HSI}t_1) > 0.8\%^* \text{HSI}t_1$ then call HSI option at the beginning of the trading day t and settle the option at the end of the trading day.

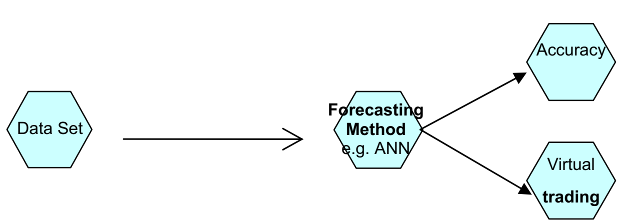

Fig. 4. Forecasting using MISMIS.

$(\text{HSI}_{\text{forecasting model},t} - \text{HSI}t_1) < -0.8\%^* \text{HSI}t_1$ then put HSI option at the beginning of the trading day t and settle the option at the end of the trading day.
$-0.8\% \leqslant (\text{HSI}_{\text{forecasting model},t} - \text{HSI}t_1) \leqslant 0.8\%^* \text{HSI}t_1$ then no action.

This trading strategy was applied to the prediction models for evaluating their financial performance. In terms of the transaction fee, there is a parameter which allows the user to specify. Here we used \$500 for a transaction. For a trading to be settled, two transactions (buy and sell or vice versa) are involved and the total transaction fee is \$500 * 2 = \$1000. As our models were on daily forecasts of HSI, for simplicity, the settlement of a call/put option need to be completed in the same day. We assumed a trader would put or call in the morning using the opening index and settle the transaction using the closing index in the same day. The financial return will be the product of HK\$50 times the difference between the closing and opening indices. Whether it is a gain or loss is determined by the correctness of prediction direction. In the mean time, the transaction fee was taken into account. The performance outputs of the virtual trading in terms of the annualized returns and Sharpe ratios are shown in Table 2.

4.4. Consensus prediction of HSI

In order to make a more reliable prediction, the prediction outcomes from ARIMA, ANN and regression were combined together to form a consensus prediction. The consensus model, which was based on regression, is shown in Fig. 5 and the results are shown in Table 3. Table 2 also indicates the performance results when fitted into our virtual trading platform.

With respect to prediction accuracy in terms of MAPE and RMSE, ARIMA and regression are more out-performed. Consequently, as indicated in Table 3, our consensus model is highly dependent on these two predictions. Our result shows that the consensus model has higher accuracy (with lower RMSE) than other models. For the financial performance in the virtual trading, the annualized returns of the ARIMA model are the highest for both the training and the testing data. By considering the risk factor, the ARIMA model, with its high Sharpe ratio for the testing period, is a good predictor. Our consensus model comes in the second place and is higher than that of ANN and regression model for the testing period.

With this case study on the HSI, we show that MISMIS facilitates users to build their models with specialized preprocessing transformation, sophisticated techniques for financial prediction, and interactive and multi-level modeling environment. Users can interactively fit data with different models as well as their consensus for comparison in a single platform. Once this prototype is web-based enabled and is attached to some kind of e-banking facilities, we expect it will be over-whelming and be well-received by the public.

On the other hand, MISMIS can support the concept of shifting window and allows users to train the model recursively on a shifting time frame which is very important for financial forecasting as the financial market is very volatile. The trained model would be

Table 2
Performance of the ARIMA model.

	ARIMA		ANN		Regression		Consensus	
	Training	Testing	Training	Testing	Training	Testing	Training	Testing
MAPE	1.43	1.00	1.28	2.87	1.07	**0.90**	**1.06**	0.92
RMSE	242.3	156.7	244.1	433.0	195.9	147.4	**191.7**	**147.3**
Annualized return (%)	**59.9**	**19.6**	56.9	4.4	9.0	5.2	27.1	5.1
Sharp ratio	0.898	**3.616**	**1.081**	0.143	0.691	0.127	0.781	0.392

The lowest MAPE and RMSE and the highest Sharp ratio are bolded.

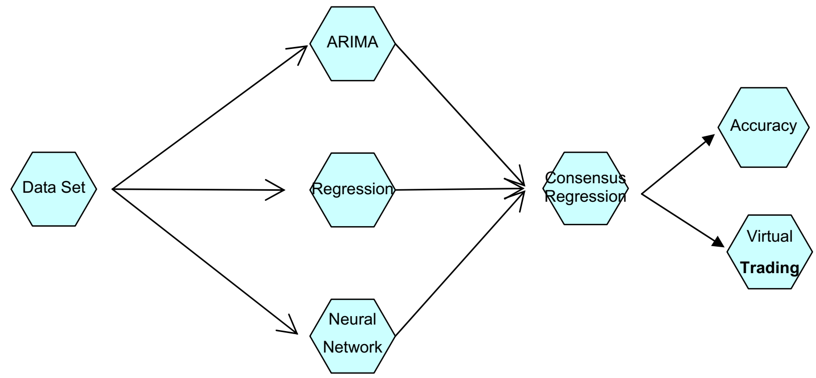

Fig. 5. Modeling for the consensus making.

Table 3
Consensus prediction.

	Coeff.
Constant	−46.32
Prediction by ARIMA	**0.217***
Prediction by regression	**0.752***
Prediction by NN	0.034

*** $p < 0.001$

more relevant as the training data is more up-to-date. The challenge for the shifting window scheme is on the analysis of these underlying factors. In each trial of a training time frame, it may come out of a model with several significant factors which would be different for different trails. Hence, it is more difficult to analyze which underlying factors are important to the prediction of HSI.

5. Evaluation feedbacks

The prototype MISMIS was separately presented to 6 investment consultants who are working in large financial institutes involving stock trading and portfolio management in June 2008. A small sample size was appropriate in this stance given the preliminary nature of the system [10,16]. That is, the focus of the evaluation was not on its general usability but on its design and possible enhancement. This should also enable us to make sense beyond its design issues [22]. The interviews were not rigidly structured, that is a flexible approach was adopted when asking respondents questions. In this way, issues which arose during the interview could be explored in depth and respondents were encouraged to elaborate in order to enhance the richness of their answers [24,31]. After a brief description of the research project, each interviewee was asked about the following two issues.

(1) After seeing the demo of the MISMIS, what are their perceptions on the usefulness and ease of use of this package? Will they adopt this package in their workplace?
(2) Will there be any possible enhancement on the MISMIS?

Through the in-depth interviews, with each one around an hour, they all (six of them) agreed that MISMIS would be very useful to most individual investors. Four consultants also agreed that this package is easy to use. In terms of adopting this package in their workplace, while three consultants recommended that its launching would be through an e-banking platform, the other three consultants suggested a pilot trial within their firms. In the current situation of Hong Kong, most individual investors trades their stocks in banks, which offer an integrated service of deposits, equity management and mortgages. Four of them perceived the most important barrier coming from the know-how of those sophisticated techniques. These four consultants said that they do not understand the prediction mechanism behind MISMIS. Thus, they suggested strengthening the MISMIS by adding more sample applications (e.g. trading of stocks, options, and warrants), a comprehensive on-line help facility, and some on-line training for those individual investors who are usually laymen on ARIMA, regression and neural network. Three respondents perceived some other practical difficulties while a bank launches the MISMIS. The database of the financial data will be a huge one and may not be easily maintainable. Two consultants puzzled that this application may generate a lot of Internet traffic and may demand for some intensive computing power in the servers of a bank. Nevertheless, they all (six of them) thought that it would a good move even a bank just put MISMIS on a trial basis for some selected customers.

6. Limitations

The MISMIS is still bound to some limitations. First, textual data is ignored in this system. Qualitative data, says company's new strategy and the quit of CEO, cannot be easily converted into quantitative format. Second, the investment environment is also ignored. Investment atmosphere may be affected so easily and changed in various period of time. For instance, atmosphere becomes tense during at war, president voting period, Severe acute respiratory syndrome (SARS) outbreaks, etc. Last but not the least, there are many different types of data mining method could be applied to financial forecasting such as genetic algorithm [15,25], support vector machine (SVM) [17] and adaptive reinforcement learning [9]. Kuo et al. [20] developed a genetic algorithm based fuzzy neural network to combine both qualitative data and quantitative data for stock market prediction. Dempster and Leemans [9] introduced adoptive reinforcement learning for a fully automated foreign exchange trading system.

7. Conclusion

By integrating several approaches from literature in handling financial data, a prototype of Multi-level and Interactive Stock Market Investment System (MISMIS) is developed. The system allows investors to build their prediction models level by level. Prediction outcomes from a model can be fed into another prediction model as input for a multi-level prediction. This approach is best suited to build a model that is complicated. The interactivity of the MISMIS enables users to know the prediction instantly. While building a complicated model using traditional approaches, it is very difficult to link up input and output variables. Moreover a complicated model is very difficult to diagnose when the prediction outcome is not satisfactory. By applying a multilevel and interactive approach, the complicated model can be broken down into some sub-models, which are much easy to analyze and the prediction model can be readjusted instantly. Finally, our virtual trading platform facilitates users to evaluate different trading strategies which are based on the predicted outcomes of stocks or stock market indices.

In a case study on the option trading of Hong Kong Hang Seng Index, we have identified that historical HSI, Dow Jones Index, property price index, retailing sales figure, prime lending rate, and CPI in Hong Kong are essential factors affecting HSI. In contrast, the less important factor would be the GDP in Hong Kong.

Upon our in-depth interviews with six financial consultants, they all agreed that MISMIS is a very useful decision support system for stock market investment. It would be further enhanced when the help facilities and on-line training with sample applications are developed. It is a good move if a bank would launch this as part of its e-banking services for some selected customers.

Appendix A. Details of selected data

This appendix lists out the details of data such as attributes, sources and their symbols used in the database.

A.1. Global financial factor

Gold: the daily closing value of Gold Bullion price $US/Troy Ounce. Source: DataStream.

Crude: the daily closing value of brent crude oil price of current day, fob U$/BBL. Source: DataStream.

EUR: the daily closing value of exchange rate of HK dollar against Euro dollar. Source: DataStream.

JPY: the daily closing value of exchange rate of HK dollar against Japan Yen. Source: DataStream.

DJI: the daily closing value of Dow Jones Industrials price index. Source: DataStream.

A.2. Local economic data

GDP: Year-on-year rate of change in real terms (%) of HK gross domestic product. Source: HK Census & Statistics Department.

Retail sales: Year-on-year rate of change in real terms (%) of HK retail sales. Source: HK Census & Statistics Department.

Property price index: Price and rental movements of private sector property (% change during the period) of HK residential price. Source: HK Census & Statistics Department.

Unemployment rate: Seasonally adjusted unemployment rate (%). Source: HK Census & Statistics Department.

CPI: Year-on-year rate of change (%) of HK composite consumer price index. CPI also represents the local inflation rate. Source: HK Census & Statistics Department.

Prime: local annually prime lending rate. Source: DataStream.

A.3. Hang Seng Index

∧HSI: the daily closing value of Hang Seng Index. Source: DataStream and HSI service Ltd.

References

[1] Y.S. Abu-Mostafa, A.F. Atiya, Introduction to financial forecasting, Applied Intelligence 6 (1996) 205–213.
[2] R. Aggarwal, P. Rivoli, Seasonal and day-of-the-week effects in four emerging stock markets, Financial Review 24 (1989) 541–550.
[3] G. Armano, M. Marchesi, A. Murru, A hybrid genetic-neural architecture for stock indexes forecasting, Information Sciences 170 (2005) 3–33.
[4] S.C. Bank, Chaos in futures market? A nonlinear dynamical analysis, Journal of Futures Markets 11 (1991) 711–728.
[5] J.M. Binner, R.K. Bissoondeeal, T. Elger, A.M. Gazely, A.W. Mullineux, A comparison of linear forecasting models and neural networks: an application to Euro inflation and Euro Divisia, Applied Economics 37 (2005) 665–680.
[6] G. Box, G. Jenkins, Time Series Analysis, Forecasting, and Control, Holden Day, San Francisco, 1976.
[7] G.P. DeCoster, W.C. Labys, D.W. Mitchell, Evidence of chaos in commodity futures prices, Journal of Futures Markets 12 (1992) 291–305.
[8] Demirgüç-Kunt, H. Huizinga, Market Discipline and Financial Safety Net Design, Policy Research Paper No. 2183, World Bank, 1999.
[9] M.A.H. Dempster, V. Leemans, An automated FX trading system using adaptive reinforcement learning, Expert Systems with Applications 30 (2006) 543–552.
[10] N.K. Denzin, The art and politics of interpretation, in: N.K. Denzin, Y.S. Lincoln (Eds.), Handbook of Qualitative Research, Sage Publication, London, 1994, pp. 500–515.
[11] R.F. Engle, Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation, Econometrica 50 (1982) 987–1007.
[12] E.F. Fama, The behavior of stock-market prices, Journal of Business 38 (1) (1965) 34–105.
[13] P.E. Gaynor, R.C. Kirkpatrick, Introduction to time-series modeling and forecasting in business and economics, McGraw-Hill, Inc., NY, 1994.
[14] R. Gencay, The predictability of security returns with simple technical trading rules, Journal of Empirical Finance 5 (1998) 347–359.
[15] F.H. Grupe, S. Jooste, Genetic algorithms: a business perspective, Information Management & Computer Security 12 (3) (2004) 288–297.
[16] E.G. Guba, Y.S. Lincoln, Competing paradigms in qualitative research, in: N.K. Denzin, Y.S. Lincoln (Eds.), Handbook of Qualitative Research, Sage Publications, 1994, pp. 105–117.
[17] W. Hung, Y. Nakamori, S.Y. Wang, Forecasting stock market movement direction with support vector machine, Computers and Operations Research 32 (2005) 2513–2522.
[18] H. Ince, Non-parametric regression methods, CMS 3 (2006) 161–174.
[19] J. Jaffe, R. Westerfiled, The weekend effect in common stock returns: the international evidence, Journal of Finance 40 (1985) 433–454.
[20] R.J. Kuo, C.H. Chen, Y.C. Hwang, An intelligent stock trading decision support system through integration of genetic algorithm based fuzzy neural network and artificial neural network, Fuzzy Sets and Systems 118 (2001) 21–45.
[21] T. Mateus, The risk and predictability of equity returns of EU accession countries, Emerging Markets Review 5 (2004) 241–266.
[22] M.B. Miles, A.M. Huberman, Qualitative Data Analysis, second ed., Sage, London, 1994.
[23] J. Murphy, Technical Analysis of the Financial Markets, Prentice Hall, 1998.
[24] J. Macfarlane Smith, Interviewing in Market and Social Research, Routledge Kegan Paul, London, 1972.
[25] A.K. Nag, A. Mitra, Forecasting daily foreign exchange rates using genetically optimized neural networks, Journal of Forecasting 21 (7) (2002) 501–511.
[26] J.C. Ritchie, Fundamental Analysis: A Back-to-the-Basics Investment Guide to Selecting Quality Stocks, revised ed., third ed., McGraw-Hill Companies, 1996.
[27] R.J. Rogalski, S.M. Tinic, The January size effect: anomaly or risk mismeasurement, Financial Analysts Journal (1986) 63–70.
[28] R.M. Rogers, Handbook of Key Economic Indicators, second ed., McMraw-Hill, NY, 1998.
[29] L. Sarno, G. Valente, Modeling and forecasting stock returns: exploiting the futures market, regime shifts and international spillovers, Journal of Applied Economics 20 (2005) 345–376.
[30] W.F. Sharpe, Adjusting for risk in portfolio performance measurement, Journal of Portfolio Management (1975).
[31] H.W. Smith, Strategies of Social Research, Prentice Hall, London, 1975.
[32] J. Sun, H. Li, Data mining method for listed companies' financial distress prediction, Knowledge-Based Systems 21 (1) (2008) 1–5.
[33] J. Vanicek, I. Vrana, S. Aly, Fuzzy aggregation and averaging for group decision making: a generalization and survey, Knowledge-Based Systems 22 (1) (2009) 79–84.
[34] S. Walczak, An empirical analysis of data requirements for financial forecasting with neural network, Journal of Management Information Systems 17 (4) (2001) 203–222.
[35] W. Wen, Y.H. Chen, I.C. Chen, A knowledge-based decision support system for measuring enterprise performance, Knowledge-Based Systems 21 (2) (2008) 148–163.