VOL 2 (2018) NO 2

e-ISSN : 2549-9904
ISSN : 2549-9610

INTERNATIONAL JOURNAL ON INFORMATICS VISUALIZATION

Application of Genetic Algorithm and Personal Informatics in Stock Market

Khulood Albeladi[#], Salha Abdullah[#]

[#] *Department of Information Systems, Faculty of Computing and IT, King Abdul Aziz University, Jeddah, Saudi Arabia*
E-mail: khuloodsalem@yahoo.com, sbente@kau.edu.sa

***Abstract*— The financial market is extremely attractive since it moves trillion dollars per year. Many investors have been exploring ways to predict future prices by using different types of algorithms that use fundamental analysis and technical analysis. Many professional speculators or amateurs had been analysing the price movement of some financial assets using these algorithms. The use of genetic algorithms, neural networks, genetic programming combined with these tools in an attempt to find a profitable solution is very common. This study presents a prototype that utilizes genetic algorithms (GAs) and personal informatics system (PI) for short-term stock index forecast. The prototype works according to the following steps. Firstly, a collection of input variables is defined through technical data analysis. Secondly, GA is applied to determine an optimal set of input variables for a one-day forecast. The data is gathered from the Saudi Stock Exchange as being the target market. Thirdly, PI is utilised to create a smart environment, which enables visualisation of stock prices. The outcome indicates that this approach of forecasting the stock price is positive. The highest accuracy obtained is 64.67% and the lowest one is 48.06%.**

***Keywords*— genetic algorithms, personal informatics, stock market, Saudi Stock Exchange.**

I. INTRODUCTION

Stock market is very complex and challenging to grasp for many investors. Many investors find it hard to decide which stock to buy, or when to sell it, and when to buy more. In addition to this, the market is also volatile that no one can precisely tell when the price will rise, how much it will rise or how long the rise will continue. Over time, investors tried to reduce the risk and uncertainty by examining more about the market, the stocks and the relationships. They may be successful in curtailing the risk but they can never avoid it. Most of the investors use their money for long-term investment. Buy and hold (B&H) is the most preeminent approach for investors [3]. In B&H strategy, investors buy a stock of a primarily good company, and then hold it for a long period of time, wait for the price to rise and thus, make money. Usually the company that they invest will earn money consistently; thereby investors do not need to worry about the daily stock prices, which saves time and efforts.

Today, there are many efforts to create intelligent software in order to respond to the market evolution that is constantly changing. In particular, such matters as genetic algorithms, fuzzy systems, evolutionary computation, swarm intelligence, and neural networks are being widely used in financial markets in many different ways, for optimizing the collection of investment assets or prediction of the future direction of prices by analysing past market data. In addition, automated trading methods, usually referred as algorithmic trading, and soft computing techniques are becoming increasingly widespread applied to solve stock market problems [2].

Genetic algorithm is an optimization technique that simulates the process of natural selection of biological systems. GA works according to the following: evaluate the available chromosomes (string of bits) and giving each one of them fitness value. Repeating this process until reaching maximum number of cycles or there is no change in the fitness value. Then, select the two chromosomes with highest obtained fitness values. After that, the two selected chromosomes will be combined in order to produce a new solution that could be better than the previous [3].

Personal Informatics (PI) concept studies the interaction between humans and information together with the construction of interfaces, organizations, technologies and systems. It also develops its own conceptual and theoretical foundations and uses foundations developed in other areas. PI also supports data visualization, which makes interpretation easier. It provides a class of tools that aims to enhance the decision-making process through the use of visual representations such as charts, graphs, and tables to analyse information easily and quickly [3].

Trade decisions on a stock market are made on the basis of predicting trends, which is a challenging task. Investors should have some notions of the probability that the prices of stocks will increase or decrease on the subsequent day.

II. Market Analysis

Experts from many markets use technical indicators to assess the market. Many investors make investments in accordance with forecasts and news of the next day rumours and based on socioeconomic and even political environment in which the company in question falls [4]. Two important categories of stock analysis are technical analysis and fundamental analysis.

Fundamental Analysis (FA) is the type of analysis that studying factors that affect companies, industry groups, and economy. It aims to analysis the profit and forecast from future price movements. FA analyses the data according to levels. At the company level, FA includes examination of business concept, management, financial data, competitors' data, and competition. At the industry level, FA examines the demands and supplying for the products. At the economy level, FA examines on assessment of national economic data to conclude future direction of economy logically [4][5][6].

Technical analysis (TA) is a technique in financial markets that able to predict the future direction and behaviour of market or a particular stock, through the analysis of past market data, predominantly the actual price and its volume. Most of traders usually used TA when they want to decide which stock to buy or sell or to keep it [7]. Technical analysts consider stock trends based on their chart and stock market close each day while completing stock technical analysis [8].

The biggest advantage is that the technical analysis focuses only on the share price. If the main concern is to predict the future evolution of the stocks, it makes sense to analyse the historical price by identifying support and resistance lines that set limits on potential evolution of the price. It also sets out times of market entry and exit. The prices, in addition to being influenced by their past performance, are also affected by the market which they belong. Therefore, fundamental analysis cannot be ignored. The technical analysis normally does not anticipate the inversions occurring trends. Hence, the paper concludes both technical analysis and fundamental analysis are needed to come up with sound outcome [9][10].

III. Saudi Stock Exchange

Saudi Stock Exchange is the joint stock company and the sole exchange in Saudi Arabia. It is popularly known by its Arabic name Tadawul and was established in 2007. Currently, Tadawul lists 171 public listing traded companies divided between 15 sectors (as at 7th January 2018). The indices are called Tadawul All Shares Index (TASI) and its currency is SAR. Capital Market Authority supervised Tadawul.

IV. GAMA workflow

This study describes the implementation of a Genetic Algorithm for Market Analysis (GAMA) prototype, which is supported by PI, to deliver a short-term prediction of an index value. GAMA uses a suitable combination of the computational power-optimization capabilities of GA with flexibility and the visualization capabilities of a PI system.

GAMA is divided into three stages i.e. data extraction stage, genetic algorithm stage and visualisation stage. Fig.1 represents the GAMA workflow.

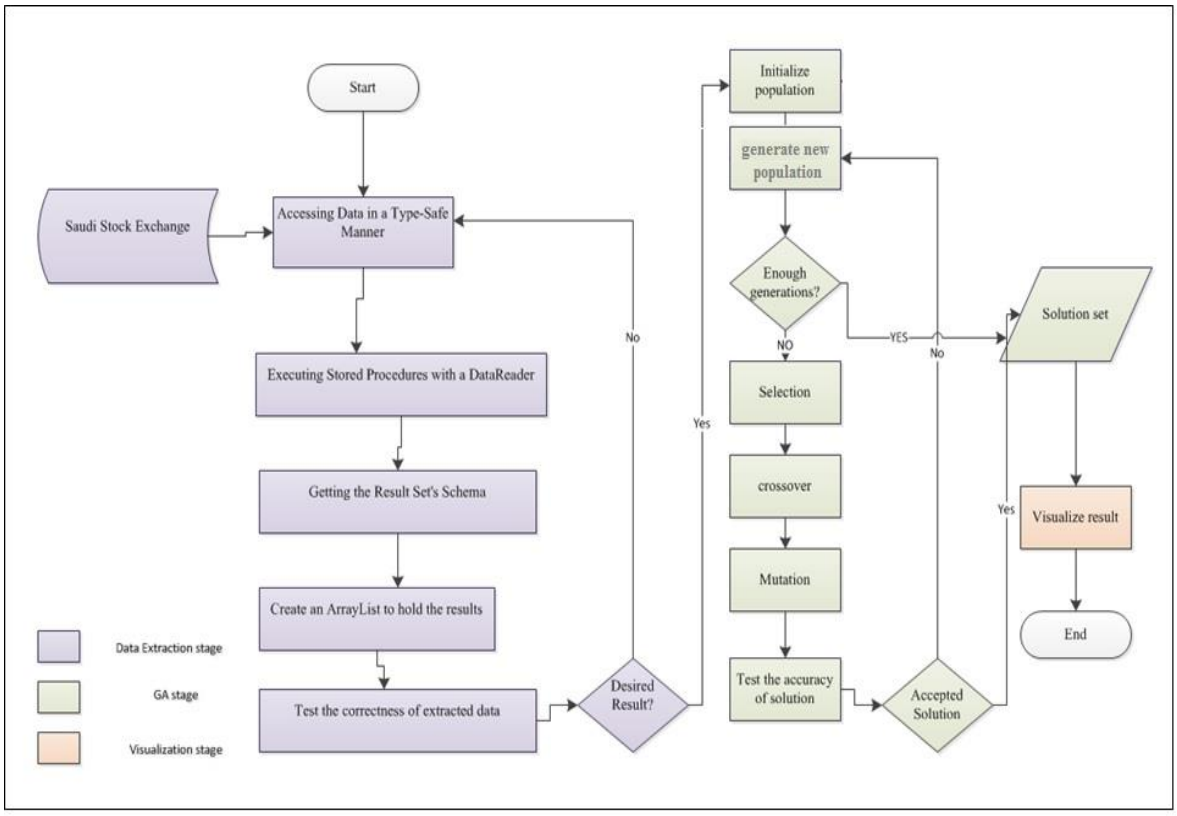

Fig. 1 GAMA Workflow

A. *Data Extraction Stage*

This stage starts by analysing raw values of index, defining oscillators and averages, and ends with creating pattern of extracted data. The source data, which are extracted from the Saudi Stock Exchange (Tadawul), which comprises opening, highest, lowest, closing price, volume and percentage of change of the index of the telecommunication and information technology sector on subsequent days.

The extraction of data used a C# .NET DataReader which is based on ADO.NET architecture. A simple user interface, as illustrated in Fig 2, has been created to make sure that the extracted data is correct and accurate by following the steps below:

1) Step 1: Accessing data in a type-safe manner.
2) Step 2: Executing stored procedures with a Data Reader.
3) Step 3: Getting the results set schema.
4) Step 4: Create an Array List to hold the results.
5) Step 5: Populate a List View with the Values from an Array List

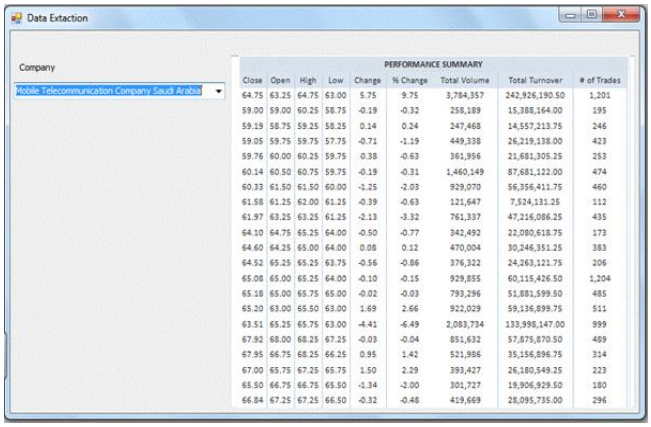

Fig. 2 List View with the Values from an Array List

B. *Genetic algorithm stage*

The GAMA GA stage has three major components. GA is very valuable because they are good at dealing with highly nonlinear problems and noisy domain such as stock market. The components are as the followings:

1) Create an initial population of *m* selected individuals. The population are the six attributes extracted in the data extraction stage.
2) Take *m* as an input and evaluate as output each individual from population based on fitness value. GAMA uses Roulette Wheel Selection (RWS) to select candidate for subsequent generation. The objective is to increase the number of trades because a solution with higher number of trades and equal return will be more robust in the test period.
3) The formulation of the next generation. The breeding processes include three basic operations, i.e. reproduction, crossover and mutation. GAMA uses uniform crossover that uses a fixed mixing ration between two parents [12].

The evolution process is repeated until the stop condition is triggered where there is no improvement in producing better results. Table 1 list the parameter set for GAMA

TABLE 1
LIST VIEW WITH THE VALUES FROM AN ARRAY LIST

Parameters	**Values**
Population Size	100
Crossover Probability	0.5
Crossover Type	Uniform
Mutation Probability	0.013
Selection	Roulette Wheel
Stopping Criteria	1000 generations

C. *Visualization stage*

Data visualization is the mechanism of taking data and conveying it in a human-friendly and visual way. GAMA uses ILNumeric. It is that is a mathematical class library for Common Language Infrastructure (CLI) developers. It makes the implementation process of an array of numerical algorithms easier and can create distribution-ready applications.

ILNumerics directly operates on the features of modern development environments and programming languages such as C# [12].

V. GAMA IMPLEMENTATION

GAMA is a Web based system. GAMA's architecture and interface are illustrated in Fig. 3 and Fig. 4 respectively as the following steps.

1) Analysis of historical data of the telecommunication sector of the Tadawul.
2) Iterations of GA.
3) Visualisation of results through PI.

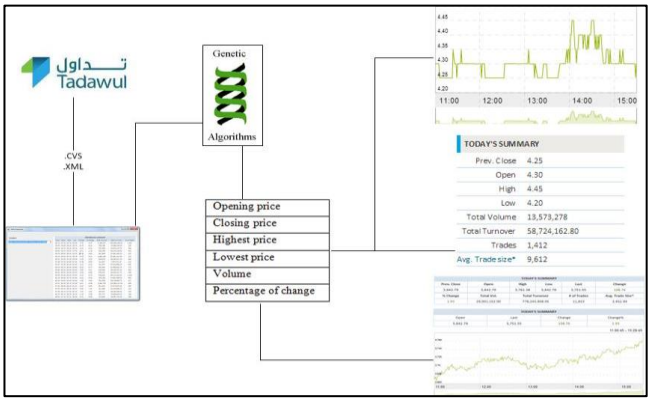

Fig. 3 GAMA Architecture

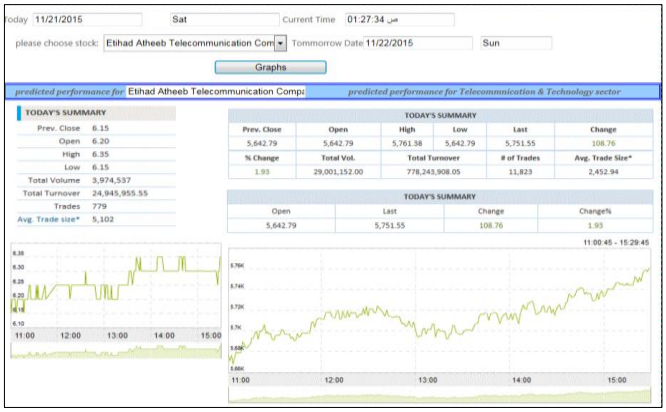

Fig. 4 GAMA Interface

VI. Testing and Evaluation

A. Data Extraction Testing

For the purpose of testing the accuracy of the data extracted from Tadawul, a tool with a simple user interface was created to make sure the extracted data (in Fig. 5) were correct and accurate.

The tool will give a result of a comparison between the data available on the Tadawul website and the data that were extracted. In Fig. 3, a different data set has been extracted 40 times. In the comparison between the source data and the data that were extracted, the accuracy result was 97.45%.

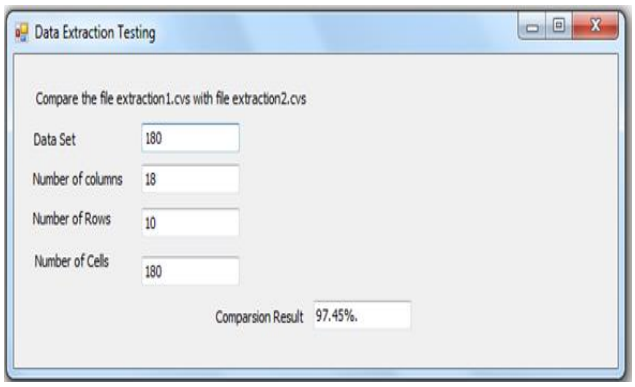

Fig 5: Extracted data accuracy

B. Functionality Testing

The dataset applied to experiment GAMA consists of extracted data from January 3, 2016, until March 17, 2016. The four companies selected were Etihad Atheeb Telecommunication Company, Etihad Etisalat Company, Mobile Telecommunications Company, and Saudi Telecom.

Dataset that used to test GAMA are divided to two categories: training dataset is used to find connection weights for each specified attribute. In addition, there is another testing dataset has been used in order to validate the result. Thus, it will check for any over fitting. The results obtained showed that there is no over fitting occurred.

After the real data has been extracted for each company of the Telecommunications sector, the data is normalized in order to reduce the experimental errors.

From Table 2 shows the optimal connection weight evaluated for each attribute.

TABLE 2
Connection weights for each company in GAMA

Company	Open price	Closing price	Highest price	Lowest price	Volume	% Change
Etihad Atheeb Telecom	23.1	23.1	23.1	22.45	3,849,544	9.95
Etihad Etisalat	4	3.91	4.1	3.9	3,120,114	0.51
Mobile Telecommunications	63.75	60.48	63.75	60	1,024,826	1.92
Saudi Telecom	6.3	6.01	6.3	6	6,211,350	0.5

Table 3 illustrate the accuracy of GAMA to predict the highest price. In some cases, the accuracy for prediction is higher for the testing data than the training data. The highest accuracy obtained using the GA is 64.67% and the lowest one is 48.06%.

TABLE 3
The accuracy of the predicted stock price for each company

Company	Fitness value using GA	
	Training data	Testing data
Etihad Atheeb Telecommunication	51.61%	51.29 %
Etihad Etisalat Company	48.06%	48.54%
Mobile Telecommunications	59.67%	61.12%.
Saudi Telecom Company	63.87%	64.67%

C. Usability Testing

GAMA was tested on 10 users. When a user launched GAMA, as shown in Figure 4.3, the main screen appears. After analysing the data, 93% of participants classified GAMA as easy to use, while 7% classified it as medium difficulty.

D. Performance Testing

The metric selected to evaluate the performance of the system is the response time, i.e. is the duration that a system or application takes to respond to the client. The system responds time is illustrated in Table 4.

TABLE 3
GAMA Response Time

Company	Response Time
Etihad Atheeb Telecommunication Company	0.4593 s
Etihad Etisalat Company	0.7884 s
Mobile Telecommunications	0.5616 s
Saudi Telecom Company	0.4549 s

VII. Conclusions

This study presented a prototype that utilizes personal informatics system (PI) and genetic algorithms (GAs) for one-day stock index prediction in an individual sector—i.e.

the telecommunication and information technology sector—with data taken from Saudi Stock Exchange. Using PI made a smart environment by enabling the GAMA to know user's interests, provide privacy, and display results in professional way. One of the limitations during implementing GAMA is limitation is the limited sources of personal informatics system as it considers new field. Another limitation is extracting data from Tadawul as it is in different formats in specific interval, to overcome this obstacle the DataAdapter had been used to build a bridge between source data and the needed data format in GAMA. Experimental results for GAMA performance show that the highest accuracy obtained is 64.67% and the lowest one is 48.06%. As future work, there is a plan to improve the GAMA performance and increase the per cent of accuracy and add more sectors to include all stock companies in Saudi Stock Exchange. We also plan to improve the visualization part of GAMA by adding dynamic graphs and exporting regular reports for each company in specific sector and each sector in the target market.

REFERENCES

[1] Hameed, A., Kang, W. and Viswanthan, S. (2010). Stock Market Declines and Liquidity. The Journal of Finance, 65(1), pp.257-293.

[2] Enke, D. and Thawornwong, S. (2005). The use of data mining and neural networks for forecasting stock market returns. Expert Systems with Applications, 29(4), pp.927-940.

[3] Cao, L., Luo, C., Ni, J., Luo, D. and Zhang, C. (2014). Stock Data Mining through Fuzzy Genetic Algorithm. Australian Research Council Discovery Grant, 06S3011(546).

[4] Einhorn, M. (2012). Trademark Valuation and Market Analysis. SSRN Journal.

[5] Sumantyo, R. (2014). Effect Analysis of Fundamental Factors Toward Cigarettes Companies Stock Price That Listed In Indonesia Stock Exchange (IDX) Period 2008-2013. SSRN Electronic Journal.

[6] Drakopoulou, V. (2016). A Review of Fundamental and Technical Stock Analysis Techniques. Journal of Stock & Forex Trading, 05(01).

[7] Yamamoto, R. (2012). Intraday technical analysis of individual stocks on the Tokyo Stock Exchange. Journal of Banking & Finance, 36(11), pp.3033-3047.

[8] Bradic-Martinovic, A. (2006). Stock market prediction using technical analysis. Economic. Annals, 51(170), pp.125-146.

[9] Federici S, Borsci S. 2010. Usability evaluation: models, methods, and applications. In: JH Stone, M Blouin, editors. International Encyclopedia of Rehabilitation. Available online: http://cirrie.buffalo.edu/encyclopedia/en/article/277/ [Accessed 12 April 2016].

[10] Buie, E. (2001). Whiteboard: usability design and testing. Interactions, 8(5), pp.13-17.

[11] Wang, Y., Yu, J. and Wen, S. (2014). Does Fundamental and Technical Analysis Reduce Investment Risk for Growth Stock? An Analysis of Taiwan Stock Market. International Business Research, 7(11).

[12] Budgetsimple.com. (2016). BudgetSimple - Easy Budget Software to Manage Your Personal Finances. [Online] Available at: https://www.budgetsimple.com/ [Accessed 12 April 2016].