

Stock Movement Modeling Based on the Analysis of Negative Correlation

Kacha Chansilp^{1*}, Kittisak Kerdprasop^{1,2}, Paradee Chuaybamroong³, Nittaya Kerdprasop^{1,2}

¹ School of Computer Engineering, Suranaree University of Technology, Nakhon Ratchasima, Thailand.

² Data and Knowledge Engineering Research Unit, School of Computer Engineering, Suranaree University of Technology, Thailand.

³ Department of Environmental Science, Thammasat University, Rangsit Campus, Thailand.

* Corresponding author. Email: kacha@sut.ac.th

Manuscript submitted July 27, 2019; accepted December 8, 2019.

doi: 10.17706/ijeeee.2020.10.2.125-134

Abstract: This research presents the data-driven modeling method to derive a combined trading model from the analysis of negative correlations among the top-five active stocks from each sector of the Thailand stock market. The negative movements are computed from the closing price direction of major stocks in the eight biggest sectors. The highly negative correlated stocks among market groups are then used to build predictive trading models with three algorithms: regression analysis, generalized linear modeling, and chi-square automatic interaction detection. An ensemble from the combination of the best two models is then created. The experimental results reveal that the proposed method of trading based on negative movement analysis can accurately predict closing price of the active stock with low error rate around 1.86%.

Key words: Negative correlation analysis, stock trading, regression, chi-squared automatic interaction detection, CHAID.

1. Introduction

Equity security such as stocks that have monetary value and can be traded is a primary financial instrument for companies to gain new capital for driving organizations' operations. Investors traditionally earn profit from their belonging securities by selling them in the stock exchange market at higher price than the bought value. Buying low and selling high is a common strategy used by all investors to earn profit, but to apply successfully such tactic the traders need complicate analysis to manage trading plan.

Some analysts may use knowledge regarding fundamental aspects such as financial statement of a company to evaluate intrinsic value of that company's security. A stock value measurement method based on a company's performance is known as a fundamental analysis. Another measurement paradigm used by many short and medium term traders is called technical analysis that evaluates stock value based on trading statistics such as prices and volumes in the past. Our proposed work in this paper is in the line of technical analysis.

On technically analyzing historical stock price movement patterns, visualization tools such as candle sticks and bar charts are commonly used to identify movement trends. There are many movement indicators used for graph plotting, for instance, moving average (MA), moving average convergence divergence (MACD), relative strength index (RSI), and many more. The trend identification is mostly based

on the movement pattern of a single stock of interest. Trend forecasting of a specific stock over consecutive time intervals is analyzed through the serial correlation to determine the future price of a stock. Such correlation is based on univariate analysis.

Another kind of correlation analysis recently appeared in the literature is the cross correlation between markets or among different sectors in the same market [1]-[7]. The positive cross-correlation between markets or among sectors is called co-movement because the index changes or the price changes are in the same direction. In this work, we propose the opposite study of stock movement; our analysis is based on the negative correlation between market sectors. We call this kind of study as *reverse movement analysis*.

Our intuitive idea is based on the observation of Thailand stock movement that major stocks in the same sector are positively correlated with each other and move in the same direction. Investing on the correlated stocks is vulnerable to a great loss. Managing portfolio to include some negatively correlated stocks can more or less be a good protection against market volatility. After the literature review section, we explain our methodology to analyze the reverse movement of stocks across sectors. We then illustrate our methodology implementation and experimentation results. In conclusion, we discuss the advantage of our proposed approach as well as the future modification of this research to overcome some current limitations.

2. Literature Review

Predicting accurately direction of stock movement is the main focus of both fundamental and technical analysts; they only differ in data used for the analysis. Fundamental analysts make their prediction based on performance of the industry and other economic factors, whereas technical analysts evaluate their investments based on statistics generated by market activity such as historical stock prices and trading volumes. The majority of research work in the literature focuses on the technical analysis indicators to derive trading rules [8] and to forecast stock price direction [9]-[12].

Stock direction forecasting techniques that have been proven by many researchers in the past decade yielding satisfactory results are those based on machine learning and soft computing techniques [13]-[16]. Machine learning tasks can be categorized into three main groups: classification (on categorical target) or prediction (on numerical target), clustering, and association. The most applicable machine learning techniques in finance domain are in the group of prediction task and the extensively adopted learning algorithm is artificial neural network (ANN). Tilakaratne et al. [17] applied ANN to predict trading signals of the Australian index. Ticknor [18] introduced a Bayesian regularized network that assigns probability to the network weights and the proposed ANN had been tested with the Microsoft and Goldman Sachs stocks. Chen and Seneviratna [19] used feed forward back propagation ANN to predict price index of Colombo stock exchange, Sri Lanka.

Researchers also adopted other learning algorithms to predict stocks including decision tree induction [20], [21], genetic algorithm [22]-[24], state space modeling [25], and optimization techniques [26]. To improve predictive accuracy of the induced models, many researchers considered a fusion approach that combined results from model ensemble such as a hybrid adaptive neuro inference system [27], a bagging of tree-based classifiers [28], [29], an integrated forecasting system using wavelet neural network and artificial bee colony [30], a deep learning approach incorporated with two-directional principal component analysis [31], and support vector machine integrated with probabilistic AdaBoost [32].

Clustering and association analyses are two machine learning tasks that have also been applied to the finance domain. Renugadevi et al. [33] applied hierarchical agglomerative clustering to rank Indian stocks based on the difference between close price and open price. Ta and Lin [34] adopted k-means and hierarchical clustering methods to visualize patterns of the Vietnamese stock market. Peachavanish [35] also detected Thailand stock trend using cluster analysis. Hsieh et al. [36] applied the concept of association

analysis by discovering closed itemsets among numerous profit rules with the main purpose of rapid search for high profit rules for the real-time trading. Bhoopathi and Rama [37] also applied association analysis to derive the efficient trading rules.

Stock clustering based on similarity in price movement can help investors making decision on managing their port folios. Association among stocks is also a supplement on port folio management in the sense that it reveals movement direction that tends to be helpful for trading decision. Cross-correlation between stock markets and interactions among market sectors have been extensively studied recently [38]. The study of cross-correlation and market movement are mostly based on the correlation analysis and the application of random matrix theory [39]-[42].

In this work, we focus our study on cross-correlation among market sectors. Instead of performing comovement analysis as other researchers had been done, we propose to anticipate stock price movement with the observed inverse correlation of stocks from different sectors. The details of our proposed research framework are explained in the next section.

3. Methodology for Trading Model Creation

Our trading model is data-driven such that model is learned through the systematic analysis of historical trading events. Our model creation process consists of seven steps as shown in Fig. 1.

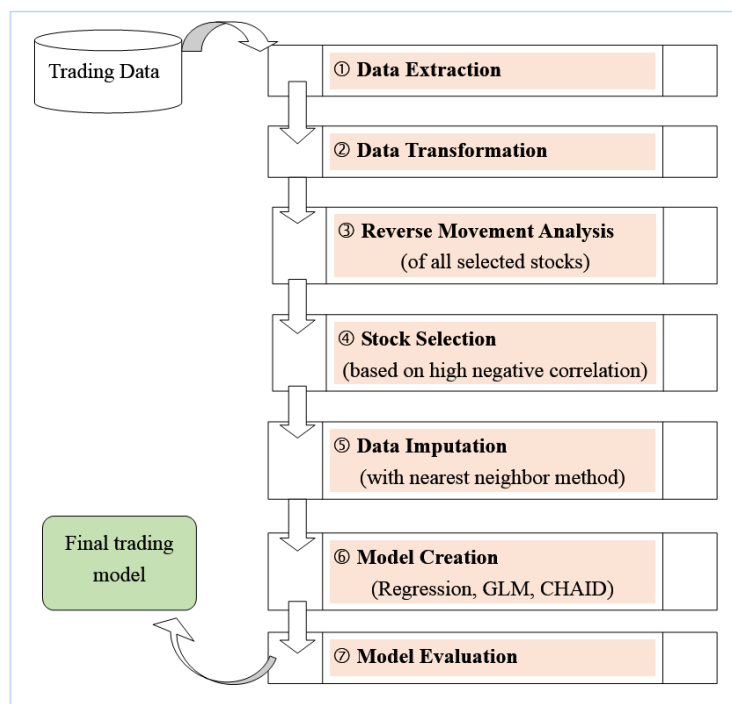


Fig. 1. Trading model creation based on the reverse movement analysis.

Data collection, preparation and analysis steps can be explained as follows:

Step 1 Data Extraction: Investors specify market sectors of interest and trading indicators such as opening price, closing price, trading volume, and so on. The proposed systematic process will extract data from each sector the top five stocks based on historical trading volumes.

Step 2 Data Transformation: The extracted data are ordered in such a way that rows are consecutive trading dates and indicators of all selected stocks are in the same row. For example, suppose traders select only two trading indicators: opening price and closing price. If they specify only two market sectors, then there will be ten stocks selected from step 1. We name the selected stocks as S_1, S_2, \dots, S_{10} . Thus, the ordered

data in this step will be:

Day1, OpenS₁, CloseS₁, OpenS₂, CloseS₂, ..., OpenS₁₀, CloseS₁₀
Day2, OpenS₁, CloseS₁, OpenS₂, CloseS₂, ..., OpenS₁₀, CloseS₁₀
...
DayN, OpenS₁, CloseS₁, OpenS₂, CloseS₂, ..., OpenS₁₀, CloseS₁₀

Step 3 Reverse Movement analysis: At this step, traders identify the target indicator. For instance, if we are interested in predicting closing price of a specific stock, then closing price of that stock will be the focus of our movement analysis. Based on the example in step 2, negative correlation of all the selected ten stocks' closing prices are computed.

Step 4 Stock Selection: Stocks from different market groups showing highly negative correlation are selected for further processing. We set the default threshold of negative correlation to be at least 0.7 in its strength to be considered as high. This threshold can be adjusted according to the traders' decision. The output of this step is a set of negatively correlated stocks from different market sectors.

Step 5 Data Imputation: This step is to be performed only if there exists missing items in the set of negatively correlated stocks. We impute the missing attributes with the nearest neighbor method.

Step 6 Model Creation: From the obtained set of negatively correlated stocks, the traders pick one stock as the target for the creation of predicting models. In our proposed methodology, we adopt three numerical prediction methods: linear regression, generalized linear model (GLM), and chi-square automatic interaction detector (CHAID).

GLM is the extension of linear modeling framework to allow variable that are not normally distributed [43]. The GLM framework includes a broad class of models such as ANOVA, Poisson regression, log-linear models, and many more. Generalization power of GLM is achieved via a modification of link function to handle different kinds of data distribution.

CHAID [44] is a tree-based learning algorithm that recursively partitioning large dataset of mixed distributions into smaller data subsets with homogeneous distribution at the terminal (or leaf) node. For continuous response (or target) variable, each terminal node of the CHAID tree represents regression-type predictor.

At the end of this step, our analysis methodology generates three predictive models to forecast stock price movement.

Step 7 Model Evaluation: The generated three predictive models have to be assessed on their predicting accuracy. We use the hold out method for model evaluation. The set of negatively correlated stocks that has been outputted from step 5 is to be split into two subsets with proportion around 70:30. The large portion (70% of data) is to be used for training and generating the three predictive models as explained in step 6, whereas the remaining (30% of data) is to be used as model tester in this step. Among the candidate three predictive models, the one with the highest accuracy (or the lowest predicting error) will be selected as the final result of our methodology. If the accuracy assessed from the three models is not distinctive, we propose a combination method by using a fusion of two or three models as an ensemble.

4. Experimentation and Model Evaluation Results

To demonstrate our proposed methodology through real-world data experimentation, we use daily trading data of Thailand stock exchange market acquired from the online source "<http://siamchart.com/stock/>". The stock trading data in our experiments are between January to December of the year 2016, which are comprising of 244 records. We explain the implementation and experimentation details of all seven steps as the following:

Step 1 Data Extraction. We extract the top-five most active stocks from each market sector. There are eight major sectors in Thailand stock market. The sector name and the most active stocks in each sector are summarized and shown in Table 1.

Table 1. The Top-5 Active Stocks Extracted for Reverse Movement Analysis

SECTOR	Top-5 Active Stocks				
Industries	PTTGC	IVL	TCCC	VNT	STANLY
Agricultures and foods	CPF	MINT	TUF	M	TF
Consumption	KYE	SUC	ICC	TR	SABINA
Finance and banks	SCB	KBANK	BBL	BAY	KTB
Properties and constructions	SCC	CPN	LH	SCCC	TPIPL
Natural resources	PTT	PTTEP	TOP	GLOW	EGCO
Goods and services	AOT	CPALL	BIGC	BIC	MAKRO
Technologies	ADVANC	TRUE	INTUCH	DELTA	DTAC

Step 2 Data Transformation. For all extracted 40 stocks, we select five indicators and the trading date for further analysis. The selected indicators are opening price (opn), highest value of the day (hgh), lowest value of the day (low), closing price (cls), and trading volume (vol). All the 40 stocks, $S_1 \dots S_{40}$, with their attributes are rearranged in the tabular format as shown in Fig. 2.

$$5 \text{ indicators} \times 40 \text{ stocks} = 200 \text{ attributes}$$

Date	opn S_1	hgh S_1	low S_1	cls S_1	vol S_1	...	opn S_{40}	hgh S_{40}	low S_{40}	cls S_{40}	vol S_{40}
...

Fig. 2. Layout of trading records to be used for movement analysis.

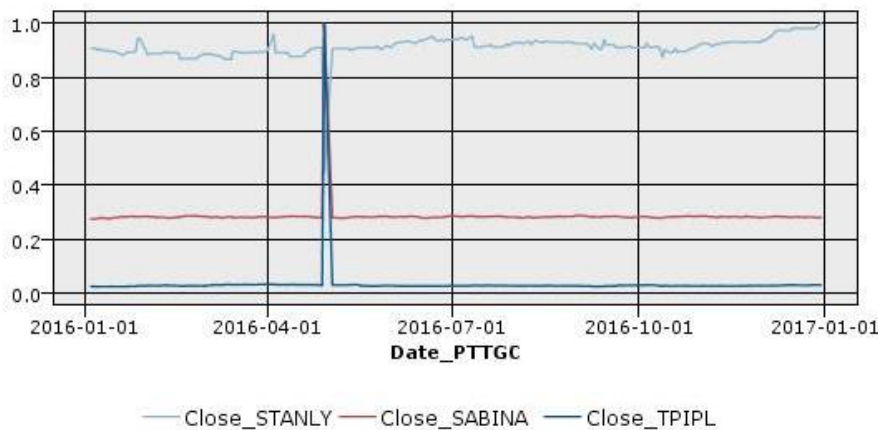


Fig. 3. Reverse movement of closing prices between STANLY (top line), SABINA (middle red graph), and TPIPL (bottom blue graph).

Step 3 Reverse Movement Analysis. The correlations of all 40 stocks are then computed based on their closing price attribute. We focus our analysis on the industrial sector and we have found only one stock in this sector, that is STANLY, showing cross-negative relations to SABINA in the Consump sector (coefficient = -0.75) and TPIPL in the Propcon sector (coefficient = -0.76). The negative movements between STANLY, SABINA, and TPIPL are graphically shown in Fig. 3.

Step 4 Stock Selection. In our experimentation, we set the correlation strength to be at least 0.75 and the direction has to be negative. Based on the correlation analysis results, the set of negatively correlated stocks from cross sectors is {STANLY, SABINA, TPIPL}.

Step 5 Data Imputation. The target of our prediction is the closing price of the STANLY stock. But the STANLY trading data from step 4 contain missing values for 39 records out of the total 244 records in the three attributes: Open, High, and Low. Thus, we need to perform data imputation with the nearest neighbor method.

Steps 6-7 Model Creation and Evaluation. To predict closing price of STANLY stock based on the cross-sector reverse movement of SABINA and TPIPL stocks, we apply three numerical prediction algorithms: linear regression, GLM, and CHAID. For GLM algorithm, we experiment with log and exponential link functions for the normal, Gaussian, and binomial distributions. The experimental results of GLM modeling are as good as the linear regression. Both GLM and linear regression models yield lower error than the CHAID model. The performances of linear regression, GLM, and CHAID models are summarized in Table 2.

Table 2. Predictive Performances of Linear Regression, GLM, and CHAID Models

MODEL	CORRELATION COEFFICIENT	TARGET ATTRIBUTE	PREDICTORS	RELATIVE ERROR
Regression	0.756	STANLY	SABINA, TPIPL	0.429
GLM	0.756	STANLY	SABINA, TPIPL	0.429
CHAID	0.349	STANLY	TPIPL	0.878

Final Trading Model

Based on the model evaluation results, the best models are linear regression and GLM and the second best one is CHAID. We thus perform another step of modeling by combining the two models: linear regression and CHAID. To reduce model assessment bias, we split the 244 data records obtained from step 5 into two separate data subsets. The large subset (167 records, around 68%) is used as a training dataset to create a combined linear regression + CHAID model. The hold-out data subset (77 records, 32%) is used as a test set to evaluate predictive performance of the combined model. Evaluation results of the combined regression + CHAID as compared to the single linear regression and CHAID models are illustrated in Table 3.

Table 3. Comparative Performances of a Combined Regression+CHAID Model against the Single Models

ASSESSMENT METRIC	MODEL					
	REGRESSION+CHAID		LINEAR REGRESSION		CHAID	
	TRAIN	TEST	TRAIN	TEST	TRAIN	TEST
Min. error	-40.60	-9.28	-9.06	-8.99	-81.24	-9.57
Max. error	15.48	12.04	16.03	12.66	15.26	11.42
Mean error	0.21	-0.47	0.22	-0.48	0.21	-0.46
MAE	3.77	3.19	3.93	3.40	4.13	3.38
SD	5.79	4.26	5.11	4.48	8.01	4.33
Correlation	0.82	0.33	0.80	0.12	0.36	0.33
Data instances	167	77	167	77	167	77

It can be seen from the results that a combined regression + CHAID model predict closing price of a target

stock more accurate than a single model. Mean absolute error (MAE) on out-of-sample data is as low as 3.19, whereas the linear regression and CHAID models predict the same set of test data with higher errors at 3.40 and 3.38, respectively. The combined regression + CHAID model is shown in Fig. 4.

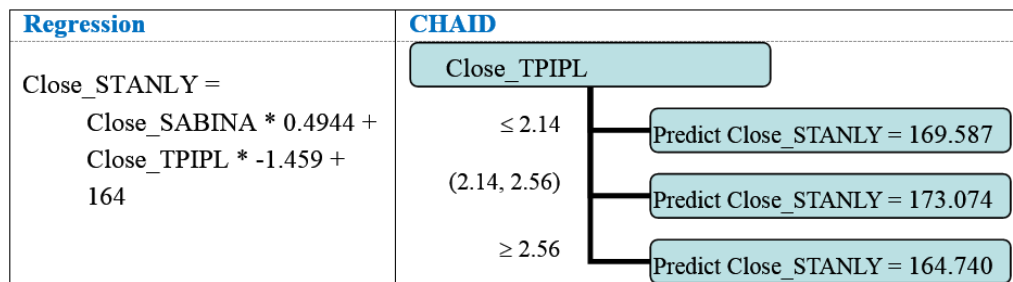


Fig. 4. The combined regression + CHAID model to predict the closing price of STANLY stock.

On applying the combined model to predict value of the target variable, both regression and CHAID predictions are to be applied and scores from both predictions are finally computed as a single model ensemble. The advantage of multiple scoring models is to avoid the limitation of each single model and thus results in a higher predictive accuracy.

5. Conclusion

In this work, we propose a systematic approach to analyze reverse movement of stocks among different market sectors and extract only highly negative correlated stocks to build potential predictive models. Based on the sample data of Thailand stock exchange market during the year 2016, we found that linear regression and CHAID are two accurate models. Then, we combine these statistical-based and tree-based modeling methods to generate the final trading model ensemble. Based on our experimental result of predicting the closing price of STANLY stock using the combined model, we observed the mean absolute error of predicting 77 out-of-sample trading data records to be as low as 3.199. Taking into consideration the mean closing price of STANLY stock at 171.891, the error rate of our combined model is around 1.86%. This percentage error is computed from $(3.199/171.891)*100$.

The proposed method creates a trading model ensemble from a single statistical approach (linear regression) and a tree-based learning algorithm (CHAID). As the result is promising, we plan to further our study on investigating other learning methods such as support vector machine and deep learning. We also plan to extend our work to cover lagged correlation on the model building process to assist the long-term trade forecasting.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

The first author is responsible for research design and manuscript preparation. The second and third authors are responsible for experimentation and result validation. The fourth author is responsible for algorithm design.

References

- [1] Forbes, K. J., & Rigobon, R. (2002). No contagion, only interdependence: Measuring stock market comovements. *The J of Finance*, 57(5), 2223–2261.

- [2] Tilakaratne, C. D., Morris, S. A., Mammadov, M. A., & Hurst, C. P. (2007). Quantification of intermarket influence on the Australian all ordinary index based on optimization techniques. *The ANZIAM J*, 48, C104–C118.
- [3] Shi, W. B., & Shang, P. J. (2013). Cross-sample entropy statistics as a measure of synchronism and cross-correlation of stock market. *Nonlinear Dynamics*, 73, 539–554.
- [4] Jiang, X. F., Chen, T. T., & Zheng, B. (2014). Structure of local interactions in complex financial dynamics. *Scientific Reports*, 4(5321).
- [5] Lin, A., Shang, P., & Zhou, H. (2014). Cross-correlations and structures of stock markets based on multiscale MF-DXA and PCA. *Nonlinear Dynamics*, 78(1), 485–494.
- [6] Han, R. Q., Xie, W. J., Xiong, X., Zhang, W., & Zhou, W. X. (2017). Market correlation structure changes around the great crash: A random matrix theory analysis of the Chinese stock market. *Fluctuation and Noise Letters*, 16(2), 1750018.
- [7] Kerdprasop, K., & Kerdprasop, N. (2017). Trading decision models based on the index movement analysis of East Asian stock exchange market. *Review of Integrative Business and Economics Research*, 6, 92–101.
- [8] Arévalo, R., García, J., Guijarro, F., & Peris, A. (2017). A dynamic trading rule based on filtered flag pattern recognition for stock market price forecasting. *Expert Systems with Applications*, 81, 177–192.
- [9] Patel, J., Shah, S., Thakkar, P., & Kotecha, K. (2015). Predicting stock and stock price index movement using trend deterministic data preparation and machine learning techniques. *Expert Systems with Applications*, 42(1), 259–268.
- [10] Jabbarzadeh, A., Shavvalpour, S., Khanjarpanah, H., & Dourvash, D. (2016). A multiple-criteria approach for forecasting stock price direction: Nonlinear probability models with application in S&P 500 index. *Int J of Applied Engineering Research*, 11(6), 3870–3878.
- [11] Kampouridis, M., & Otero, F. E. B. (2017). Evolving trading strategies using directional changes. *Expert Systems with Applications*, 73, 145–160.
- [12] Nelson, D., Pereira, A., & Oliveira, R. (2017). Stock market's price movement prediction with LSTM neural networks. *Proceedings of Int. Joint Conf. on Neural Networks* (pp. 1419–1426). Anchorage, Alaska, U.S.A.
- [13] Atsalakis, G. S., & Valavanis, K. P. (2009). Surveying stock market forecasting techniques – Part II: Soft computing methods. *Expert Systems with Applications*, 36(3), 5932–5941.
- [14] Azevedo, J. M., Almeida, R., & Almeida, P. (2012). Using data mining with time series data in short-term stocks prediction: A literature review. *Int J of Intelligence Science*, 2(4), 176–180.
- [15] Dash, R., & Dash, P. K. (2016). A hybrid stock trading framework integrating technical analysis with machine learning techniques. *The J of Finance and Data Science*, 2(1), 42–57.
- [16] Kamley, S., Jaloree, S., & Thakur, R. S. (2016). Performance forecasting of share market using machine learning techniques: A review. *Int J of Electrical and Computer Engineering*, 6(6), 3196–3204.
- [17] Tilakaratne, S. D., Mammadov, M. A., & Morris, S. A. (2009). Modified neural network algorithms for predicting trading signals of stock market indices. *J of Applied Mathematics and Decision Sciences*, 2009, 125308.
- [18] Ticknor, J. L. (2013). A Bayesian regularized artificial neural network for stock market forecasting. *Expert Systems with Applications*, 40(14), 5501–5506.
- [19] Chen, D., & Seneviratna, D. (2014). Using feed forward BPNN for forecasting all share price index. *J of Data Analysis and Information Processing*, 2(4), 87–94.
- [20] Iskrich, D., & Grigoriev, D. (2017). Automatic creation of stock trading rules on the basis of decision trees. *Proceedings of the 2nd Conf. on Software Engineering and Information Management* (pp. 25–30).

Saint Petersburg, Russia.

- [21] Saluja, M. S., & Shaikh, Y. (2017). Decoding investment pattern of FIIs and DIIs in Indian stock market using decision tree. *Int J of Advanced Research in Computer Science*, 8(3), 911–916.
- [22] Pinto, J. M., Neves, R. F., & Horta, N. (2015). Boosting trading strategies performance using VIX indicator together with a dual-objective evolutionary computation optimizer. *Expert Systems with Applications*, 42(19), 6699–6716.
- [23] Huang, C. F., Li, H. C., & Chang, B. R. (2016). An intelligent computational finance model for microstructure-based trading system. *Proceedings of Int. Conf. on Applied System Innovation* (pp. 1–4). Okinawa, Japan.
- [24] Macedo, L. L., Godinho, P., & Alves, M. J. (2016). A comparative study of technical trading strategies using a genetic algorithm. *Computational Economics*, pp. 1–33.
- [25] Nakano, M., Takahashi, A., & Takahashi, S. (2017). Creating investment scheme with state space modeling. *Expert Systems with Applications*, 81, 53–66.
- [26] Karathanasopoulos, A. (2016). Modelling and trading the English stock market with novelty optimization techniques. *Economics and Business Letters*, 5(2), 50–57.
- [27] Chakraborty, A., Mukherjee, D., Dutta, A., Chakraborty, A., & Kole, D. K. (2016). An effective stock price prediction technique using hybrid adaptive neuro fuzzy inference system based on grid partitioning. *Intelligent Systems Technologies and Applications*, 275–284. Springer.
- [28] Wang, J. L., & Chan, S. H. (2006). Stock market trading rule discovery using two-layer bias decision tree. *Expert Systems with Applications*, 30(4), 605–611.
- [29] Barak, S., Arjmand, A., & Ortobelli, S. (2017). Fusion of multiple diverse predictors in stock market. *Information Fusion*, 36, 90–102.
- [30] Khuat, T. T., Le, Q. C., Nguyen, B. L., & Le, M. H. (2016). Forecasting stock price using wavelet neural network optimized by directed artificial bee colony algorithm. *J of Telecommunications and Information Technology*, 2, 43–52.
- [31] Singh, R., & Srivastava, S. (2017). Stock prediction using deep learning. *Multimedia Tools and Applications*, 76(18), 18569–18584.
- [32] Zhang, X., Li, A., & Pan, R. (2016). Stock trend prediction based on a new status box method and AdaBoost probabilistic support vector machine. *Applied Soft Computing*, 49, 385–398.
- [33] Renugadevi, T., et al. (2016). Stock market prediction using hierarchical agglomerative and k-means clustering algorithm. *Indian J of Science and Technology*, 9(48).
- [34] Ta, V. D., & Liu, C. M. (2016). Stock market analysis using clustering techniques: the impact of foreign ownership on stock volatility in Vietnam. *Proceedings of 7th Symp. on Information and Communication Technology* (pp. 99–106). Ho Chi Minh City, Vietnam.
- [35] Peachavanish, R. (2016). Stock selection and trading based on cluster analysis of trend and momentum indicators. *Proceedings of Int. MultiConf. of Engineers and Computer Scientists* (pp. 317–326). Hong Kong.
- [36] Hsieh, Y. L., et al. (2016). Efficient mining of profit rules from closed inter-transaction itemsets. *J of Information Science & Engineering*, 32(3), 575–595.
- [37] Bhoopathi, H., & Rama, B. (2017). A novel framework for stock trading analysis using casual relationship mining. *Proceedings of 3rd Int. Conf. on Advances in Electrical, Electronics, Information, Communication and Bio-Informatics* (pp. 136–141). Chennai, India.
- [38] Marti, G., et al. (2017). A review of two decades of correlations, hierarchies, networks and clustering in financial markets. *ArXiv:1703.00485*.
- [39] Rosenow, B., Gopikrishnan, P., Plerou, V., & Stanley, H. E. (2003). Dynamics of cross-correlations in the

stock market. *Physica A: Statistical Mechanics and Its Applications*, 324, 241–246.

- [40] Wilcox, D., & Gebbie, T. (2007). An analysis of cross-correlations in an emerging market. *Physica A: Statistical Mechanics and Its Applications*, 375, 584–598.
- [41] Oh, G. (2014). Grouping characteristics of industry sectors in financial markets. *Physica A: Statistical Mechanics and Its Applications*, 395, 261–268.
- [42] Jiang, X. F., Zheng, B., Ren, F., & Qiu, T. (2017). Localized motion in random matrix decomposition of complex financial systems. *Physica A: Statistical Mechanics and Its Applications*, 471, 154–161.
- [43] Madsen, H., & Thyregod, P. (2011). *Introduction to General and Generalized Linear Models*. Chapman & Hall/CRC.
- [44] Kass, G. V. (1980). An exploratory technique for investigating large quantities of categorical data. *Applied Statistics*, 29(2), 119–127.

Copyright © 2020 by the authors. This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited ([CC BY 4.0](https://creativecommons.org/licenses/by/4.0/)).



Kacha Chansilp is an assistant professor at the School of Computer Engineering, Suranaree University of Technology (SUT), Thailand. He received his bachelor degree in computer science from Queens College, U.S.A., in 1993, the master degree in graphic design from New York Institute of Technology, U.S.A., in 1995 and the doctoral degree in interactive multimedia technologies from Edith Cowan University, Australia, in 2003. His current research includes web databases, educational systems, environmental modeling, and computational applications in engineering.



Kittisak Kerdprasop is an associate professor and chair of the School of Computer Engineering, SUT. He received his bachelor degree in mathematics from Srinakarinwirot University, Thailand, in 1986, the master degree in computer science from the Prince of Songkla University, Thailand, in 1991 and the doctoral degree in computer science from Nova Southeastern University, Florida, U.S.A., in 1999. His current research includes machine learning and artificial intelligence.



Paradee Chuaybamroong is currently an associate professor in environmental science with the Department of Environmental Science, Thammasat University, Thailand. She received her bachelor degree in Public Health from Mahidol University, Thailand, in 1990, the master degree in environmental science from Colorado School of Mines, U.S.A. in 1997 and the doctoral degree in environmental science from University of Florida, U.S.A., in 2002. Her current research includes environmental science and engineering, geophysics, bioaerosols, and environmental photocatalysis.



Nittaya Kerdprasop is an associate professor and the head of Data and Knowledge Engineering Research Unit, School of Computer Engineering, SUT. She received her bachelor degree in radiation techniques from Mahidol University, Thailand, in 1985, the master degree in computer science from the Prince of Songkla University, Thailand, in 1991 and the doctoral degree in computer science from Nova Southeastern University, U.S.A., in 1999. Her research of interest includes data mining, artificial intelligence, logic programming, and machine intelligence.