The Association and Prediction of Index of the Stock Exchange of Thailand (SET) between the Index of Foreign Exchange by using Data Mining Techniques

Sayan Tepdang   Tortanut Choosang   Suvijak Sothanakul

Department of Information Systems

Faculty of Business Administration and Information Technology

Rajamangala University of Technology Tawan-Ok Chakrabongse Bhuvanarth Campus

122/41 Vibhavadi Rangsit Road Dindeang Sub-District, Dindeang District Bangkok 10400 Thailand

E-mail: Sayan.te@cpc.ac.th

Abstract: This paper presents the association and prediction of the Stock Exchange of Thailand (SET) by using data the index of foreign exchange from Bloomberg news agency during September 8, 2014 to September 7, 2015. The first stage of the research is to find the association of information which has an increase or decrease of the Stock Exchange of Thailand with the composite index of Shanghai, China, Nikkei 225 Stock index of Japan, the index of the New York Stock Exchange, USA and the index of the London Stock Exchange, UK. Are they relevant after the close of trading on the stock market every day? The method is FP-Growth Algorithm. The results of the research showed that the Support value is greater than 10% of data. This can create association rules related to Stock Exchange of Thailand index which has a total of 56 law relationships; for example: when the exchange composite, Stock exchange Nikkei 225, New York Stock Exchange and the London Stock Exchange are at the close of trading and show the rise of stock market indices all four. This indicates that the close of trading on the Stock Exchange of Thailand will have to increase the index. The confidence is up to 70% and the lift is at 1.443. At the second stage of the research, after association rules related Stock Exchange of Thailand, the selected rule related with the confidence more than 60%. To create a model for prediction the index of SET, by means of Ensemble consisting of three ways: Decision Tree, K-Nearest Neighbors (K-NN) and Neural Networks. The results of the research showed that modeling of the London Stock Exchange has a value of up to 63.48% in Accuracy, followed by the two equal models: Modeling relationship rules of the Shanghai Stock Exchange and the London Stock Exchange composite. The second model is law relationship Stock Exchange Nikkei 225 and the London Stock Exchange with a value of Accuracy equal to 63.38% and the minimum is to create a model law relationship Stock Exchange composite of Shanghai and New York Stock Exchange is worth at 58.43% of Accuracy.

Keywords-- **Artificial Intelligence, Data Mining, Ensemble, Neural Networks, FP-Growth Algorithm**

1. Introduction

Stock market plays an important role in the economy of the countries concerned in the interest of investors, entrepreneurs, business owner and the general public regularly. They can follow news from several types of media such as television and radio. There will be a status report on security of trading and analysis of the factors that affect the trading of the stock traded on the stock market on a daily basis. To continuouly inform people, Stock exchange or the stock market is a place for exchange-traded security short and long term of the company limited. The company will sell issued shares to investors or the public in general. The returns for investor will be the profits from investment or the dividends from holdings of the issued securities. The stock exchange market are available all over the world. The largest stock exchange in the world's top 10 are sorted by market value of the shares of domestic companies as on 31 January 2015 (The monthly report by the Federation of Exchanges). This is shown in Table 1 of the Stock Exchange.

No	The Global Stock Market	Value (Billion USD)
1.	New York Stock Exchange Euronext	19,223
2.	Exchange Dniester Riksdag LOM X.	6,831
3.	London Stock Exchange	6,187
4.	Tokyo Stock Exchange	4,485
5.	Shanghai Stock Exchange	3,986
6.	The Stock Exchange of Hong Kong	3,325
7.	Stock Exchange Euronext	3,321
8.	Shenzhen Stock Exchange	2,285
9.	Stock Exchange, TMX Group.	1,939
10.	The Leaders Exchange Churchill's Nursery.	1,762

Table 1. World Stock Exchange Ranking

The Table1. shows that the stock exchange investment has been used in a lot of number around the world. It is regarded as a major factor in the economy of a certain country. SET is Market, which is a source of many companies to be listed and to allow investors to co-invest. Investors will be one of the shareholders of the company or joint the ownership in a company. The Thai exchange market establishes under the Securities Exchange Act of 1974 with the objective or joint ownership in companies. The objectives are to provide a central source for buying securities, to promote savings and to raise funds in the country. It is opened to trading the official on April 30, 1975 by the then English name is "Securities Exchange of Thailand". Then its name had been changed as "The Stock Exchange of Thailand (SET)" on January 1, 1991 onwards. Currently, investors, entrepreneurial businesses as well as the general public who invest in the stock market will try to find different techniques to help investors to profit from

investing in the stock market. Investors are: therefore; essential to find information or techniques affecting the increase or decrease of the index on a daily basis. Thus, the research for the law of relationship and prediction of the Securities Exchange of Thailand and the Securities exchange abroad with data mining techniques is interesting. This is to guide the investors who can apply to invest in the stock market.

2. RELATED RESEARCHES

Data mining is a process that is done with the data to find patterns and relationships hidden in the data series in the statistical analysis, recognition, machinery learning, mathematics, etc. At present, data mining has been applied in several categories in terms of business, scientific and medical as well as economic and social study. There are processes that act on the information in Data Mining. To find formula , guidelines and relationships hidden in the data set that many forms such as data clustering. In medical study, data clustering shares data with similar characteristics into groups of patients with the same disease as the symptoms. To be useful in analyzing the causes of the disease based on patients with similar symptoms.

Data classification is to find a specific type of object properties of objects such as finding the relationship between the results of a physical examination of the body and the disease by using patient data and diagnostic medical stored to help diagnose the patient or medical research. For the business area, data classification is used to view the properties of the debt or bad debt to consider and approve the loan and so on. The use of association rule shows the relationship of objects or events occurring simultaneously. An example of the application of the rule is the analysis of the sale of goods with considering buyers' purchase the goods at the same time and to investigate the potential relationship. For example, it was found that people often buy a video cassette tape with glue.

Research concerning the relationship related to the Stock Exchange of Thailand: Research in economics used on the Stock Exchange of Thailand index, such as the relationship between the reserve currency countries. A Test of the relationship between the stock exchange of Thailand index and foreign reserve of Thailand [3]. A Test of relationship between the exchange rate and the stock price of energy sector in the Stock Exchange of Thailand [4]. Analysis of the relationship between the stock market indices and the exchange rates of some selected G 20 countries by panel co-integration method [5]. This research stated that it has gained popularity and it was used in the main branch of economics, such as the use of co-integrated applications and using panel Co- Integrated functions and so on.

Researches related to data mining: The use of rules, relationships and forecasting such as Socio-Economic Household data analysis using the Clustering and Association Technique for Data Mining is research using the rules in relation to the relationship of each group. To find the variables which are related [6]. The associations rule mining by using top weight of complete symmetric digraph is a research data mining association rules from transaction to transaction, promotional items are

recommended to buy with repeated [7]. The data mining for finding the factors of poverty by using class association rules technique is the study of the techniques of data mining. Class Association Rules is a research study to share samples with individualized cost less than the poverty line. [8] Question Classification for Answer Searching Using Semantic Web and Data Mining shows a connection between the main questions and answers using the Apriori algorithm. The processing of results from the answers to questions linked to Thai in a semantic keyword of questions and answers [9]. The Recommender System for Adaptive Learning Based on Students' Competency Profile by Using FP-Growth Techniques is the research correlation technique to be used in education[11]. For the land price forecasting Data Mining Techniques is the neural networks technique which resulted in accuracy to 97%, and the decision tree accuracy was 86%, which was a method that was effective in forecasting. Research on modeling predictions for the fund using data mining and neural networks NAV Forecasting Model Using Data Mining and Artificial Neural Network is a research technique using Neural Networks in forecasting such figures can forecast more effectively.

From the above Researches, this paper will use the rules associated with the algorithm. The algorithm FP-Growth which is way faster than the algorithms presented in Apriori Algorithm used to read data from the database to only two times faster. After getting the law relationship and then to create a model using a technique called Ensemble Methods using a technique called Vote Ensemble, which uses three techniques include techniques Decision Tree, techniques, K-Nearest Neighbours (K-NN) and Neural techniques Networks for prediction of prophecy.

3. EXPERIMENTS

How to conduct research on five stages. (As shown in Figure 1. Process of Research)
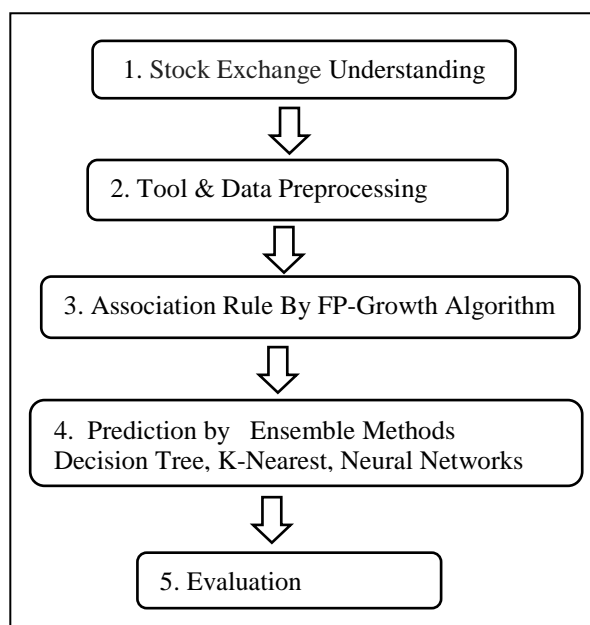

Figure 1. Process of Research

1. Stock Exchange Understanding: understands and seeks for data related to the stock market, both domestic and foreign to be used in research.

2. Tool & Data Preprocessing: is data analysis with data mining using software RapidMiner Studio 6.

The data is from the use of stock in exchange for each of the followings: The Shanghai Stock Exchange Composite (CHINA) Nikkei 225 stock (JAPAN) New York Stock Exchange (USA) London Stock Exchange (UK), researchers from the Bloomberg news agency. On September 8, 2014 until September 7, 2015 which is the one-year period for using in data mining architecture if the market closes each day. If the market is covered by the index increased, data storage is set to "1". If the market is closed, then the index would drop to store a value of "0" will convert the data into Figure 2. The Data Collection

ID	THAI	CHINA	JAPAN	USA	UK
1	1	0	0	1	1
2	0	0	0	0	0
3	1	1	1	0	1
4	1	1	1	1	1
5	1	0	0	0	1
6	0	0	0	0	0
7	0	0	0	0	0
8	0	0	0	0	0
9	1	1	0	0	0
10	0	0	0	0	0

Figure 2. The Data Collection

3. To find Association Rules ,information will be used to with FP-Growth Algorithm.

4. After getting rules related to information then the rules associated with the confidence (more than 60%)will be selected in order to create a model for forecasting the index SET. By means of a so-called Vote Ensemble consisting of three ways: Decision Tree, K-NN and Neural Networks.

5. Evaluation is to evaluate the effectiveness of the results of the modeling.

4. RESULTS

The result showed that the support of the Stock Exchange of Thailand in relation to the rules of The Shanghai Stock Exchange Composite, Stock exchange Nikkei New York and stock Exchange London stock Exchange as shown on Figure 3. the support on the Stock Exchange of Thailand.

Size	Support	Item 1	Item 2	Item 3
1	0.483	THAI		
2	0.325	UK	THAI	
2	0.315	CHINA	THAI	
2	0.300	JAPAN	THAI	
2	0.281	USA	THAI	
3	0.222	CHINA	UK	THAI
3	0.217	UK	USA	THAI
3	0.212	JAPAN	UK	THAI
3	0.202	CHINA	JAPAN	THAI

Figure 3. The Support on the Stock Exchange of Thailand

The Figure 3. Shows the support of the Stock Exchange of Thailand. It found that during the period from September 8, 2014 until September 7, 2015 the Stock Exchange of Thailand. There were off the market each day of the year. SET index has risen to 0.483 or 48.3% on the London Stock Exchange and the Stock Exchange of Thailand. When the stock market's index rose simultaneously in one year was at 0.325, or 32.5%, and the Shanghai Stock Exchange Composite and SET index rose simultaneously at .315 or 31.5%, for example.

After getting the support, the next step was to find the value of confidence. The Confidence was the number expressed confidence that the rules of the relationship of the form on the left of the rule which is between [0-1]. If the value is almost 1, it means there is confidence in the legal relationships. And the Lift was an indication that the rules were interrelated or not. If the correlation value Lift was less than 1, it indicates that the rule was not a relationship or independent. But if it is greater than one, it indicates that the legal relationship between the two rules as well (as shown in Figure 4. the Confidence and Lift.)

Premises	Conclusion	Confidence	Support	Lift
chaina, japan, uk, usa	Thai	0.700	0.104	1.443
chaina, japan, uk	Thai	0.681	0.158	1.403
chaina, uk, usa	Thai	0.660	0.153	1.360
chaina, uk	Thai	0.643	0.223	1.325
japan, uk, usa	Thai	0.643	0.134	1.325
japan, uk	Thai	0.642	0.213	1.323
chaina, usa	Thai	0.617	0.183	1.271
uk	Thai	0.611	0.327	1.260
chaina, japan, usa	Thai	0.605	0.114	1.248
uk, usa	Thai	0.603	0.218	1.242

Figure 4. The Confidence and Lift

The Figure 4. shows the confidence and Lift will be found to explain the rules or assumptions that were obtained by processing the data to create a rule or assumption of up to 56 rules, for example

1. [CHINA JAPAN UK USA] > [THAI] (Confidence: 0.700) Rules or Premises: If The Shanghai Stock Exchange Composite, Stock exchange Nikkei London stock Exchange , the New York Stock Exchange closed of trading on that day and if there was a rise in the index, the fourth stock exchange indicated that the Stock Exchange of Thailand at the close of trading would increase the index by the confidence of 0.7 or 70%, and the Lift was at 1.443

2. [CHINA JAPAN UK] > [THAI] (Confidence: 0.681) Rules or Premises: If the Shanghai Stock Exchange Composite, Stock exchange Nikkei, London stock Exchange closed of trading on that day and the increased it mean that the stock exchange of Thailand transaction would increase the index at the closing. The Confidence was at 0. 681, or 68.1%, and the Lift was at 1.403.

3. [CHINA USA] > [THAI] (Confidence: 0.617) Rules or Premises: The Shanghai Stock Exchange Composite, the New York Stock Exchange, the close of trading on that day and rose of the index, it showed that the primary market of

Thailand would increase at closing of the transaction. The Confidence was at 0.617 or 61.7%, and the Lift was at 1.271.

At the next stage of research, after getting the relationship rules related to Stock Exchange of Thailand the researcher has selected the rules associated with the Confidence (which was more than 60%). There were 10 relationship rules to create a model for forecasting the index of SET. By means of a so-called Vote Ensemble consisting of three ways: Decision Tree, K-NN and Neural Networks. The measurement accuracy is a measure of the Accuracy (Acc) of the model by including all classes. And to measure by the F-measure (F1), which can be obtained. Precision (P) and Recall (R) which are determined by the following formula.

$$F1 = \frac{2 * P * R}{P + R} \qquad (1)$$

The results can be displayed as a table.

Name	P	R	F1	Acc
Chaina, Japan, Uk, Usa	61.40	66.27	63.74	59.93
Chaina, Japan,Uk	63.28	67.36	65.25	62.45
Chaina, Uk, Usa	60.60	62.45	61.51	58.88
Chaina, Uk	66.49	59.64	62.87	**63.38**
Japan,Uk, Usa	63.20	63.64	63.41	61.29
Japan, Uk	66.49	59.64	62.87	**63.38**
Chaina, Usa	57.99	75.91	65.75	**58.43**
Uk	65.69	59.64	62.51	**63.48**
Chaina,Japan, Usa	59.79	69.09	64.10	59.33
Uk,Usa	65.42	59.73	62.44	63.29

Table 2. The results of the research

The Table 2. The results of the research showed that the predictive modeling of rules associated with the London Stock Exchange, the highest was at 63.48% of accuracy. A minor had the same two models was to create a model law relationship of the Shanghai Stock Exchange Composite and the London Stock Exchange. The second model was to create a model law relationship and the London Stock Exchange Nikkei Accuracy was equal to 63.38%. And the minimum was to create a model relationship rule Stock Exchange of New York Stock Exchange and the Shanghai Composite was at 58.43% Accuracy.

5. CONCLUSIONS AND FUTURE WORKS

The study found that the accuracy in forecasting a great extent. But the value was not very high. The increase or decrease of the index was influenced by many factors. Therefore, future research should be studied new form of the algorithms or to compare or to find other factors that effect of changes in stock market indexes such as currency, international reserve, the investment of foreign investors in the stock market, the forecast economic expansion of the remnants of the country import-export, political uncertainty in the country and abroad, and so on. Or collect data from one year to five years researches in order to assist in the enhancement of the high accuracy.

References

[1] Eakasit Pacharawongsakda, Introduction to Business Analytics with RapidMiner Studio 6, ASIA DIGITAL PRESS, 2014.

[2] Eakasit Pacharawongsakda, An Introduction to Data Mining Techniques, ASIA DIGITAL PRESS, 2014.

[3] Jintanan Kunchai, A Test of the relationship between the stock exchange of Thailand index and foreign reserve of Thailand, Chiangmai University, 2007.

[4] Jutaporn Vejmanat, A Test of relationship between the exchange rate and the stock price of energy sector in the Stock Exchange of Thailand, Chiangmai University, 2009.

[5] Chadarat Phipathananunth, Sirithorn Jalearnrat, Sasithorn Mongkolsripattana, Socio-Economic Household data analysis using the Clustering and Association Technique for Data Mining, The University of Thai Chamber of Commerce, 2013.

[6] Pramool Suksakaophong, Phayung Meesad, Sasithorn Mongkolsripattana, "Associations Rule Mining By Using Top Weight of Complete Symmetric Digraphs," The Eleventh National Conference on Computing and Information Technology 2015 ,pp.352-358, July 2015.

[7] Surayos Sricharoen and Nuanwan Soonthornphisaj, "Data Mining for Finding the Factors of Poverty by Using Class Association Rules Technique,"National Graduate Research Conference,pp.423-431, Februay 2009.

[8] Supote Bualerng and Wararat Songpan, "Question Classification for Answer Searching Using Semantic Web and Data Mining," The Tenth National Conference on Computing and Information Technology 2014,pp.101-108, May 2014.

[9] Uraiwan Inyaem, "The Recommender System for Adaptive Learning Based on Students' Competency Profile by Using FP-Growth Techniques," The Eleventh National Conference on Computing and Information Technology 2015,pp.364 - 370, July 2015.

[10] Kannika Nutchomphu And Maleerat Sodanil, "Land Price Forecasting using Data Mining Techniques," The Tenth National Conference on Computing and Information Technology 2014,pp.671- 676, May 2014.

[11] Thararat Udomchaibanjerd Pudsadee Boonrawd, "NAV Forecasting Model Using Data Mining and Artificial Neural Network," The Tenth National Conference on Computing and Information Technology 2014, pp.890- 895, May 2014.

[12] Bloomberg News, available at: http://www.bloomberg.com

[13] Wikipedia, available at: https://th.wikipedia.org/wiki