Pattern Recognition for Stock Index Prediction Using Apriori Algorithm - The Case of NSE NIFTY

S.Thirupparkadal Nambi[1] , M.V.Subha[2]

Abstract—**Predicting stock market performance and identifying the influence of various factors on stock prices is an interesting as well as a rewarding activity. The ability to predict its direction will enhance the decision making capability of individual investors, institutional investors and regulating bodies. Hence financial managers always look for more sophisticated tools for prediction. This study is the application of machine learning technique to understand the trend (Bull/Bear) of the popular Indian stock market index NSE NIFTY. It is an attempt to understand the relationship among sectoral indices -which are exclusive performance indicators of various sectors of the industry- and the broad market index NSE NIFTY and to mine association rules using Apriori algorithm. The generated rules would be of immense help for the investors and financial managers to identify the patterns in the sectoral indices that results in the next day's up and down movement of NSE NIFTY.**

Keywords - **Apriori algorithm, Association rule mining, Predictive analysis, Stock Index Prediction, Stock market data mining.**

I. INTRODUCTION

Stock market changes over time (Barsky & De Long, 1992; Hendershott & Moulton, 2011). Stock market prediction is one of the greatest challenges for experts and researchers who work in the financial sector (Hussein A.S., Hamed I.M., Tolba M.F.,2015). High Performance Computing (HPC) techniques have been considered in order to have somewhat "real-time" predictions (Ahmad et al,2004). Fu et al used both the rule-based and template-based approaches for stock charts pattern detection, relying on the Perceptually Important Points (PIPs). Stock market participants across the globe try to extract previously unknown and potentially useful information from the past data. This enables them to predict future behavior of the markets. Exploration of various factors that influence and direct the prices of stock markets is carried out to understand the movement of stock prices. Many models are employed to model and predict the stock price.

[1] S.Thirupparkadal nambi is working with D.J. Academy for Managerial Excellence, Pollachi Main Road, Othakkalmandapam, Coimbaore- 641032, Tamilnadu. (e-mail: nambist@gmail.com)

[2] M.V.Subha is working with Department of Management Studies, Anna University Regional Center-Coimbatore. Maruthamalai main Road, Navavoor, Coimbatore-641046, Tamilnadu.(e-mail: subhamv@gmail.com)

With the advent of new technologies, new investment strategies are employed, which has made some of the old models obsolete. Association rule mining (ARM) is a popular technique for discovering interesting relations between variables in large databases. Mining the financial data is a challenging task. Modern financial analysis aims at identifying smarter ways to understand and visualize stock market data into useful information that will aid better investing decisions. The discovery of interesting patterns has become one of the most important data mining tasks, and it can be applied to many domains (Caraa-Valente and Lopez-Chavarrias, 2000; Lerner et al., 2004). Financial tasks are highly complicated and multifaceted; they are often stochastic, dynamic, nonlinear, time-varying, flexible-structured, and are affected by many economical & political factors (Tan TZ, Quek, C, and G.S. Ng, GS, 2007). One of the most important problems in the modern finance is finding efficient ways of summarizing and visualizing the stock market data that would allow one to obtain useful information about the behavior of the market (Boginski V, Butenko S, and Pardalos PM, 2005).

While many works are aimed at studying the predictability of markets, not many studies have attempted to analyze the movement of sectoral stock indices viz a vis the National stock index. With the Indian economy on the revival path, and the government initiatives to create vibrant Industrial set up in the country, studying the co movement of NSE NIFTY and its Sectoral indices will be largely beneficial. The outcome of this study will help in understanding the impact of various sectors on the overall stock market performance. As curiosity gets better of the investors, they want to identify the major sectoral indices that move along the country's stock market index. In this domain, ARM can be applied to discover the interesting behavior within a time series or the relationship among a set of time series so that investors can collect more useful information from the already available but huge amount of data. This study endeavors to find out the association between various sectoral indices alongside the movement of Indian stock index. It also generates statistically significant rules that would help to understand the market dynamics. The rest of the paper is organized as follows: Section II reviews the past literature, section III describes the data and sources of data, section IV dwells in detail about the methodology of association rule mining, section V deals with results & analysis and section IV concludes the study.

II. LITERATURE REVIEW

The association rule mining technique has been applied in various areas of finance. Most popularly referred as market basket analysis, the rule mining technique is applied for predicting banking crisis, financial defaults, financial risks in insurance etc. There have been attempts to employ the popular data mining technique of association rule mining in stock market analysis. The initial attempts were to study the predictive relationships among various financial variables. Liao, Chu and You (2001) implemented the association ruling approach to explore the co-movement between foreign exchange rates and category stock indices in Taiwan. They proposed several possible portfolio alternatives in the Taiwan financial capital market including foreign exchange currencies and stock investment under different circumstances. Lio, Ho and Lin (2005) investigated investment issues on Taiwan stock market using a two stage data mining approach, Association rule mining and k-means algorithm, Cluster analysis to explore the stock clusters for investment information. The study proposed several possible Taiwan stock market alternatives under different circumstances. Enke and Thawornwong (2005) introduced an information gain technique used in machine learning for data mining to evaluate the predictive relationships of numerous financial and economic variables. Jo Ting, Tak-chung Fu, and Fu-lai Chung(2006) propose a pattern-based stock data mining approach which transforms the numeric stock data to symbolic sequences, to carry out intra-stock and inter-stock association analysis and uses the mined rules for predicting the further price movements. Intra-stock mining which focuses on finding frequently appearing patterns for the stock time series itself and inter-stock mining which discovers the strong inter-relationship among several stocks.

Most of the research works on stock market analysis, view them as a time series problem, and there have been few studies trying to explore the cause and effect relationships among different stock categories or the influence of outside factors (Liao, Ho & Lin, 2008). The review of the above work reveals that studying predictive relationships of stock markets along with various financial factors is an interesting area of research. Sung and So(2011) proposed the analysis of association rule for predicting changes in the Korea Composite Stock Price Index (KOSPI) based on the time series data of various interrelated world stock market indices. Subha & Nambi (2013) studied the relationship between global economic cues and Indian stock markets. Eleven major time series comprising of daily close price of four global stock indices, daily exchange rates of four strong currencies, gold price, Brent oil price and London Interbank Offer Rate (LIBOR) are considered as the global cues for this study. The co-movement of all these global cues and the Indian Indices-BSE Sensex and NSE Nifty – are analyzed to generate strong association rules.

The review of literature also show that not many studies are undertaken in the Indian stock market, and how the Indian stock markets move alongside various sectors representing Indian industry are not fully explored. Hence an attempt is made to employ the association rule mining technique, the Apriori algorithm, to generate association rules, that will enable researchers to arrive at how and when the stock markets will move up or down in relation to sectoral indices.

III. DATA AND SOURCES OF DATA

The NIFTY is a well diversified 50 stock index and it represents important sectors of the Indian economy. The base period selected for NIFTY 50 index is the close of prices on November 3, 1995, which marks the completion of one year of operations of NSE's Capital Market Segment. The base value of the index has been set at 1000 and a base capital of Rs.2.06 trillion. The NIFTY 50 Index represents about 65% of the free float market capitalization of the stocks listed on NSE as on March 31, 2016. It is computed using free float market capitalization method and can be used for a variety of purposes such as benchmarking fund portfolios, launching of index funds, ETFs and structured products etc.

Sectoral index is designed to reflect the behavior and performance of a particular segment of the financial market. Eleven sectoral indices such as NIFTY Auto, NIFTY Bank, NIFTY Financial Services, NIFTY IT, NIFTY FMCG, NIFTY Media, NIFTY Metal , NIFTY Pharma, NIFTY Private Bank, NIFTY PSU Bank, NIFTY Realty are considered for this study. NIFTY Auto Index comprises 15 tradable, exchange listed companies. This index represents auto related sectors like Automobiles 4 wheelers, Automobiles 2 & 3 wheelers, Auto Ancillaries and Tyres. The NIFTY Bank Index comprises of the most liquid and large Indian Banking stocks. This Index has 12 stocks from the banking sector which trade on the National Stock Exchange of India Ltd. The NIFTY Financial Services Index is designed to reflect the behavior and performance of the Indian financial market which includes banks, financial institutions, housing finance and other financial services companies. It comprises of 15 stocks that are listed on the NSE. The NIFTY FMCG Index is designed to reflect the behaviour and performance of FMCGs (Fast

Moving Consumer Goods) which are non-durable, mass consumption products and available off the shelf. This Index comprises of 15 stocks from FMCG sector listed on the NSE. The NIFTY IT index captures the performance of the Indian IT companies. It comprises of 10 companies listed on the National Stock Exchange. The NIFTY Media Index reflects the behaviour and performance of the Media & Entertainment sector including printing and publishing. It comprises of maximum 15 stocks from Media & Entertainment sector that are listed on the NSE. The NIFTY Metal Index represents the Metals sector (including mining). It comprises of maximum 15 stocks that are listed on the National Stock Exchange. NIFTY Pharma Index captures the performance of the pharmaceutical sector. It comprises of 10 companies listed on NSE. The NIFTY Private Bank Index is the performance indicator of the banks from private sector. This Index consists of 10 stocks and is based on free float market capitalization method. The NIFTY PSU Bank Index captures the performance of the PSU Banks. The Index comprises of 11 companies listed on National Stock Exchange. NIFTY Realty Index exhibits the behavior and performance of Real Estate companies. The Index comprises of 10companies listed on NSE.

This study includes select thematic indices that reflect the performance of various broad investment themes such as NIFTY Commodities, NIFTY CPSE, NIFTY Energy, NIFTY MNC. The NIFTY Commodities Index captures the behaviour and performance of a diversified portfolio of companies representing the commodities segment which includes sectors like Oil, Petroleum Products, Cement, Power, Chemical, Sugar, Metals and Mining. It includes 30 companies that are listed on the Stock Exchange. NIFTY CPSE Index has been constructed to facilitate Government of India's initiative to disinvest some of its stake in Central Public Sector Enterprises (CPSEs) through ETF route. The index comprises of select 10 CPSEs. NIFTY Energy sector Index includes companies belonging to Petroleum, Gas and Power sectors. The Index comprises of 10companies listed on National Stock Exchange of India. The NIFTY MNC Index comprises 15 listed companies on National Stock Exchange in which the foreign shareholding is over 50% and / or the management control is vested in the foreign company.

This study also includes broad indices such as NIFTY Midcap50, NIFTY Smallcap 50 along with NSE NIFTY. NIFTY Midcap 50 includes top 50 companies based on full market capitalization from NIFTY Midcap 150 Index and on which derivative contracts are available on National Stock Exchange. The primary objective of the NIFTY Smallcap 50 Index is to capture the movement of the smallcap segment of the market. This index represents top 50 companies selected based on average daily turnover from the top 100 companies selected based on full market capitalization in NIFTY Smallcap 250 Index.

To understand the influence of several sectors of Indian industry on Indian Stock Market as a whole, the above-mentioned sectoral indices(11), thematic indices(4) and broad indices(2) along with NSE NIFTY is considered for the study. The daily close price of all these eighteen indices for the period from 01/04/2017 to 31/03/2018 with about 250 values of trading days are downloaded from www.nseindia.com.

IV. METHODOLOGY

The pioneering association rule mining algorithm, the Apriori method, was developed by Agrawal and Srikant in 1994. As the name implies, this algorithm is based on the fact that it uses prior knowledge of frequent itemset properties. The basic idea of this algorithm is to generate candidate itemsets of a given size and then scan the dataset to check if their counts are really large. A set of items in a transaction is called a market basket. Frequent Patterns are items sets that occur in a data set frequently. Frequent pattern mining searches for recurring relationships in a given data set and generate interesting associations and correlations between item sets in transactional and relational databases. Such analysis is called association rule mining (ARM) and it is also known as Market Basket Analysis or Affinity Analysis. The process is iterative. The working of the algorithm is illustrated in the following steps.

- All single itemsets are candidates in the first pass. Any item that has a support value less than the pre-specified minimum is dropped from the pool of candidate itemsets.
- Remaining single itemsets are combined to form two-member candidate itemsets. Support values of these candidates are then determined by scanning the dataset again. During this second pass, only the candidates above the pre-specified support value are retained.
- The next pass creates a three member candidate itemsets and the process is repeated.
- The process ends only when all the frequent itemsets are accounted for.

The frequent item sets are then used to generate association rules which have confidence values greater than or equal to the specified minimum confidence. The process of creation of large datasets is illustrated in the figure 1.

Transaction
ACD
BCE
ABCE
BE

1-pass →

C_1

1- itemset	Count
A	2
B	3
C	3
D	1
E	3

→

L_1

Large 1-itemset	Count
A	2
B	3
C	3
E	3

C_2

2- itemset
AB
AC
AE
BC
BE
CE

2-pass →

C_2

2- itemset	Count
AB	1
AC	2
AE	1
BC	2
BE	3
CE	2

→

L_2

Large 2-itemset	Count
AC	2
BC	2
BE	3
CE	2

C_3

3- itemset
BCE

3-pass →

C_3

3- itemset	Count
BCE	2

→

L_3

Large 3-itemset	Count
BCE	2

Result →

Large itemsets	Count
BCE	2
AC	2

Fig 1. Process of creating of large data sets

A transaction dataset (T) contains the items A, B, C and D. During the first pass, candidate itemset C_1 containing single items are generated with their support values. With the minimum support level of two, the first pass results in the large dataset L_1, eliminating item D.

During the second pass, set of two member candidate itemsets are generated, eliminating the item D, and the datasets are scanned to find their support level. As the itemsets AB and AE are lagging behind the minimum support level of two, they are eliminated from the candidate itemset C_2, resulting in larger dataset L_2.

C_3 is the three item candidate sets which again undergoes the third iteration resulting in L_3.

The results of the three passes are two large data itemsets, BCE and AC with a minimum support of two as shown in the above figure.The two step Apriori algorithms is illustrated by the following pseudopodia:

Pseudo code for Apriori algorithm *procedure* ***Apriori****(T, minSupport) { //T is the database and minSupport is the minimum support*

L1= {frequent items};

for*(k= 2; Lk-1 !=φ; k++)* ***{***

Ck= candidates generated from Lk -1

//that is Cartesian product Lk -1 x Lk -1 and eliminating any k -1 size //itemset that is not frequent

for each *transaction* ***t****in database* ***do{***

#increment the count of all candidates in Ck that are contained in t

Lk = candidates in Ck with minSupport

}//end for each

}//end for

Return $\boldsymbol{U_k L_k}$

}

From the large itemsets thus identified, strong association rules can be generated and are represented in the form 'A' => 'B'. This is an association rule and has to be read, if ‘A’ then ‘B’. Here A is called the antecedent and B is the consequent. This rule A => B has the ***support****s*, where *s* is the percentage of transactions in T that contains both A and B (i.e. A U B). This is the probability, P(A U B). This rule A => B has a ***confidence***, where *c* is the percentage of transactions in T containing A that also contain B. This is the conditional probability, P(B/A).

$$support(A \Rightarrow B) = P(AUB)$$
$$= \frac{number\ of t\ ransactions\ containing\ union\ of\ the\ sets\ Aand\ B}{total\ number\ of\ transactions} \quad (1)$$

$$confidence(A \Rightarrow B) = P(B/A) = \frac{P(AUB)}{P(A)}$$
$$= \frac{number\ of\ transactions\ containing\ union\ of\ the\ sets\ A\ and\ B}{number\ of\ transactionsc\ ontaining\ A} \quad (2)$$

Notice that the notation P(AUB) indicates the probability that a transaction contains the union of set A and set B (i.e. it contains every item in A and in B). This should not be confused with P(A or B) which indicates the probability that transactions contain either a or B. Support and confidence are important parameters in ARM as support measures the frequency of association, and confidence measures the strength of the association. Apriori algorithm usually starts with a minimum support of 100% of the data items and decreases this in steps of 5% until there are at least 10 rules with the required minimum confidence of 0.9 or until the support has reached a lower bound of 10%, whichever occurs first. ARM algorithms are so flexible that they can generate rules that satisfy desirable support and confidence levels.

V. Analysis and interpretation

To study the association between the Indian stock indices and sectoral indices, the daily close prices of Indian stock index NSE NIFTY, and the daily price of above mentioned seventeen indices are captured as a date-wise transactional database for the entire study period. The values in the transactional database are converted either into 0 or 1 depending upon whether today’s price is lesser than yesterday’s price.

$$t_s = \begin{cases} 0, & p_t < p_{t-1} \\ 1, & p_t > p_{t-1} \end{cases} \quad (3)$$

Where t_s is today’s state; p_t is today’s price, p_{t-1} is yesterday’s price. Weka’s implementation of Apriori algorithm is used to generate association rules. The top ten association rules based on the frequently occurring patterns for NSE NIFTY is listed below.

TABLE I

ASSOCIATION RULES FOR NSE NIFTY @ SUPPORT 0.1 AND CONFIDENCE 0.72

Rule	Antecedent		Consequent
1.	Bank=1 Commodity=1 CPSE=1 FMCG=1 FS=1 IT=1 (40)	==>	Nifty_nxt=1 (29) conf:(0.72)
2.	Commodity=1 CPSE=1 FMCG=1 FS=1 IT=1 Private Bank=1 (40)	==>	Nifty_nxt=1 (29) conf:(0.72)
3.	Bank=1 Commodity=1 CPSE=1 FMCG=1 FS=1 IT=1 Private Bank=1 (40)	==>	Nifty_nxt=1(29) conf:(0.72)
4.	Bank=1 Commodity=1 CPSE=1 FMCG=1 FS=1 IT=1 Midcap=1 (36)	==>	Nifty_nxt=1 (26) conf:(0.72)
5.	Bank=1 Commodity=1 CPSE=1 FMCG=1 IT=1 Midcap=1 Energy=1 (36)	==>	Nifty_nxt=1 (26) conf:(0.72)
6.	Commodity=1 CPSE=1 FMCG=1 FS=1 IT=1 Midcap=1 Private Bank=1 (36)	==>	Nifty_nxt=1 (26) conf:(0.72)
7.	Commodity=1 CPSE=1 FMCG=1 IT=1 MNC=1 Midcap=1 Energy=1 (36)	==>	Nifty_nxt=1(26) conf:(0.72)
8.	Bank=1 Commodity=1 CPSE=1 FMCG=1 FS=1 IT=1 Midcap=1 Private Bank=1 (36)	==>	Nifty_nxt=1 (26) conf:(0.72)
9.	Bank=1 Commodity=1 CPSE=1 FMCG=1 IT=1 Midcap=1 Private Bank=1 Energy=1 (36)	==>	Nifty_nxt=1 (26) conf:(0.72)
10.	Commodity=1 CPSE=1 FMCG=1 FS=1 IT=1 Energy=1 (39)	==>	Nifty_nxt=1(28) conf:(0.72)

The number within the bracket in the second column represents the total number of instances for which the antecedent is true and the number within the bracket in the third column represents the number of instances for which the consequent is true. Confidence is the ratio between the two. Value '1' represents the rise in today's value compared to yesterday's value (Bull) and '0' represents the fall in today's price compared to yesterday's price (Bear).

- Rule#1 identifies a frequently occurring antecedent pattern of rise in Bank, Commodity, CPSE, FMCG, FS and IT. The consequent is the rise in Nifty. As this is observed in 29 days out of 40 days, the rule has a confidence of 72%.
- Rule#2 identifies another frequently occurring antecedent pattern of rise in, Commodity, CPSE, FMCG, FS, IT and Private bank. The consequent is the rise in Nifty. As this is observed in 29 days out of 40 days, the rule has a confidence of 72%.
- Rule#3 identifies a frequently occurring antecedent pattern of rise in Bank, Commodity, CPSE, FMCG, FS, IT and Private bank. The consequent is the rise in Nifty. As this is observed in 29 days out of 40 days, the rule has a confidence of 72%.

Thus the other rules can be interpreted similarly. Thus the generated association rules can be used identify the patterns in the trend of sectoral indices that would help to predict the market movement of Nifty.

VI. CONCLUSION

In today's globally integrated economic scenario, stock markets are continually affected by stream of information emanating from various factors such as domestic, global, political, calamity, unforeseen events, investors psyche etc. As a result, it leaves a vast amount of informational trail. Data analytics ensure that you are not data rich and information poor. The famed Apriori algorithm of market basket analysis is used with the stock market data to generate association rules for Indian stock indices from sectoral indices**.** Seventeen such indices along with the daily close price of NSE NIFTY during the study period are captured as a date-wise transactional database. Apriori algorithm could successfully generate ten rules for NSE NIFTY at the minimum support and confidence level of 0.1 and 0.72, respectively. All patterns generated identify the bullish trend of NSE NIFTY and no pattern is generated for identifying the downtrend (Bear). It is an indication that it would be more difficult to predict the Bear than Bull. These association rules would help stock market analysts to better understand the effect of sectoral indices and also help to predict the trend of the Indian market in relation to sectoral indices. Thus, this study proves the usefulness of Apriori algorithm in generating actionable information from voluminous stock market data that would be of immense value to individual investors, institutional investors, and even regulating agencies.

REFERENCES

[1] K. Ahmad, T. Taskaya-Temizel,. D.Cheng., L. Gillam., S. Ahmad, H.Traboulsi and J. Nan-kervis, "Financial Information Grid –an ESRC e-Social Science Pilot". *In: 3rd UK e-Science Programme All Hands Meeting*, Nottingham, United Kingdom ,2004.

[2] R. Agrawal, and R. Srikant "Fast algorithms for mining association rules". In Proc. 20th Int. Conf. *Very Large Data Bases*, J. B. Bocca, M. Jarke, and C. Zaniolo, Eds., Santiago, Morgan Kaufmann, pp. 487-499, 1994.

[3] R. Agrawal, T. Imielinski and A. Swami, "Mining association rules between sets of items in large databases," *Proceedings of the ACM SIGMOD Conference on Management of Data*, Washington, D.C., pp. 207-216, May 1993.

[4] R.B. Barsky and J.B. De Long, " Why does the stock market fluctuate?" *s.l.:National Bureau of Economic Research*., 1992.

[5] J.P. Caraa-Valente and I. Lopez-Chavarrias,. "Discovering similar patterns in time series", In: Proc. *Sixth ACM SIGKDD International Conference on Knowledge Discovery and Data mining*, pp. 497–505, 2000.

[6] W.Y.Chiang, "To mine association rules of customer values via a data mining procedure with improved

model: An empirical case study", *Expert Systems with Applications* , vol. 38, no. 3, pp. 1716– 1722, 2011.

[7] S.H. Chun and Y.J. Park ,"Dynamic adaptive ensemble case-based reasoning: Application to stock market prediction", *Expert Systems with Applications*, 28, 435–443, 2005.

[8] E. David and S Thawornwong S, "The use of data mining and neural networks for forecasting stock market returns", *Expert Systems with Applications*, 29, 927-940, 2005.

[9] T. Fu, F. Chung., R. Luk and C. Ng, "Stock Time Series Pattern Matching: Template-based vs. Rule-based Approaches", *Engineering Applications of Artificial Intelligence* 20(3), 347–364, 2007.

[10] J. Han, and M.Kamber, *Data Mining: Concepts and Techniques*, Morgan Kaufman Publishers, San Francisco, 2006,pp. 227-283.

[11] T. Hendershott. and P.C. Moulton,." Automation, speed, and stock market quality", *The NYSE's hybrid, Journal of Financial Markets,* 14(4), pp. 568--604, 2011.

[12] A.S. Hussein, I.M. Hamed. and M.F. Tolba M.F, "An Efficient System for Stock Market Prediction." *Advances in Intelligent Systems and Computing*, vol 323. Springer, Cham,2015.

[13] Jo Ting, Tak-chung Fu, and Fu-lai Chun ," Mining of Stock Data: Intra- and Inter-Stock Pattern Associative Classification", Conference: Proceedings of the *International Conference on Data Mining, DMIN 2006*, Las Vegas, Nevada, USA, June 26-29, 2006

[14] S. Liao, H. Ho and H. Lin," Mining stock category association and cluster on Taiwan Stock exchange", *Expert Systems with Application*s, 35, 19-29. , 2008.

[15] S.Liao, P.Chu and T. TEng , "Mining co-movement in the Taiwan stock funds market", *Expert Systems with Applications*, 38, 5276-5288, 2001.

[16] S. Liao, P. Chu and Y. You, "Mining the co-movement between foreign exchange rates and category stock indices in the Taiwan financial capital market", *Expert Systems with Applications,* Vol.38, PP.4608-4617. 2001.

[17] S.R.Nanda., B.Mahanty and M.K. Tiwari, "Clustering Indian stock market data for Portfolio Management*", Expert Systems with Applications*, vol. 37, no. 12, pp. 8793-8798, 2010.

[18] M.V.Subha and S.T. Nambi , " Association Rule Mining using Apriori Algorithm – The case of Indian Stock market and global cues", In Proc. *International Conference on Business Analytics and Intelligence* ,IIM-B, 11–13 December 2013.

[19] H.N.Sung., Y. So, "Forecasting changes in Korea Composite Stock Price Index (KOSPI) using association rules", *Expert Systems with Applications*, Vol. 38, no. 7, 9046–9049, 2011.

[20] G. Shmueli,N.R, Patel N.R and Bruce P.C, *Data Mining for Business Intelligence*, Wiley, New Jersey, 2010, pp.263-278.

[21] I.H. Witten, E. Frank and M.A. Hall , *Data Mining: Practical Machine Learning Tools and Techniques*, Morgan Kauffman Publishers, Burlington USA, pp.116-123, 2011.

[22] D.Zhang. and L.Zhou,, "Discovering golden nuggests: Data mining in financial application," *IEEE Transaction s on Systems, Man, and Cybernetics - Part C: Applications and Re views*, vol. 34, no. 4, pp. 513522. 2004.