e-ISSN: 2349-9745
p-ISSN: 2393-8161
Scientific Journal Impact Factor (SJIF): 1.711
International Journal of Modern Trends in Engineering and Research
www.ijmter.com

Efficient Way To Forecast Share Market

Harshada Ashok Kore[1], SamataAjitkumar Gandhi[2] ,GaikwadNeha Pandharinath,[3] Prof.More Anjali[4]
[1]Student, S.B.Patil,college of Engineering, Indapur,India.
[2]Student,S.B.Patil college of Engineering, India
[3]Student,S.B.Patil college of Engineering, India
[4]Assistant Professor, S.B.Patilcollege of Engineering, Indapur, India.

Abstract— This paper proposes a framework that will give predictions about the share market, which will take after two principle stages, which are fragment based affiliation mining and further on more advancement and expectation will be given by genetic algorithm. The significant point of interest of utilizing fragment based mining is that, it aggregates all the traits once and performs operations gathering shrewd rather than single qualities which brings about more summed up principles which are further exceptionally advanced utilizing genetic algorithm as now is the right time space many-sided quality is not exactly whatever other algorithm and give forecast of little scale organizations based on exchange information of vast scale and in addition little scale organizations.

Keywords— Data mining, Fragment based mining, Association rule mining, optimization, Genetic algorithm.

I. INTRODUCTION

The prediction of the stock market is a fascinating subject. It changes the lives of investors on a daily basis based on the decisions they make whether to buy or sell shares. The stock market is most important resource for companies to raise money. This allows businesses to be publicly traded by selling shares of ownership of the company in a public market. Examination demonstrates that the cost of shares and different resources is a critical piece of the progress of financial action and can impact social mind-set.

Indeed, the stock exchange is regularly viewed as the essential sign for nation's monetary quality and improvement. Driving offer costs has a tendency to be connected with expanded business venture and the other way around. Along these lines, national banks have a tendency to watch the control and conduct of the stock exchange. Trades likewise go about as a clearing house for every exchange importance that they gather and convey the shares and guarantee instalment to the dealer of security.

This wipes out the danger to an individual purchaser or merchant that the counter party could default on the exchange. The smooth working of all these exercises encourages financial development, lower expenses; advance the creation of merchandise and administrations and occupation. Thus the monetary framework adds to expanded prosperity.[8]Stock exchange Price expectation utilizing information mining is a stand out amongst the most interesting issues to be researched and it is one of the essential issues of securities exchange inquire about over the previous decade. In any case, deciding the best time to purchase, offer or hold a stock stays exceptionally troublesome on the grounds that there are loads of components that may impact stocks costs like basic, Technical files, obscure factors.[6]

Information mining system like Association tenet mining (ARM) concentrates on discovering most successive thing sets and relating affiliation rules. Section based principle mining system by taking simply chronicled datasets as data is proposed. In any case, this calculation produces a greatly vast number of affiliation standards, regularly in hundreds or even thousands. Further, the affiliation standards are once in a while huge. It is about inconceivable for the end clients to see effortlessly. Subsequently, the clients need to utilize an attempt and -lapse way to get suitable number of standard.

II. LITERATURE REVIEW

Sr No	Name of Algorithm	Behavioural characteristics	Drawbacks
1	Apriori	Finds subsets common to minimum number confidence threshold of item-sets.	Costly to handle huge amount of candidate set.
2.	FITI	Introduces parameter called maxspan (w),used in the mining of association rules, and only rules spanning less than or equal to w transactions will be mined.	Increased time and space complexity
3.	Artificial Neural Network	Detects complex non-linear relationship between dependent and independent variables.	Black box nature Greater computational burden
4.	Fragment based mining Excluding Genetic algorithm	Divides the attributes into tiers	Creates large number of generalized rules.

III. IMPLEMENTED SYSTEM

Methodology-

As expressed in past focuses stock expectation in this proposed framework can be performed utilizing Fragment based mining and further on applying genetic algorithm to furnish the tenets with greatest fitness.[1] Distinctive routines which are accessible and utilized with the end goal of the finding of forecast tenets reasons disappointment in the expectation regarding constant execution or as the information increments.

Accordingly this paper proposes a framework which will be utilizing fragment based mining and genetic algorithm to give more fitting expectation. This method is partitioned in after steps and functions as takes after

1 .Data collection-

The principal thing which is obliged is up to date information rather say up to time information of vast scale organizations and little organizations ought to be kept up in the plain structure with its date, from online account locales like yahoo finance. The data of the organizations is maintained in an excel sheet.

2. Data Extraction

Second thing when forecast is require of some specific organizations then information of that organizations ought to be separated from the gathered information for the further handling. The separated even information will contain exchange information of expansive scale organizations and exchange information of little scale organizations.

3. Data preprocessing-

In share market the contrast between the two exchanges speaks to the status level of qualities. Let Δ be the distinction from characteristic qualities among inter-transactions Assume 1,0,-1 show expand ,unbiased and diminish separately .Each exchange speaks to the change in past date exchange.

4. Fragment based mining –

Fragment based approach divides the attributes in two tiers :

- Small scale and Large scale SUM functions.

After this we need to discover contrasts for property estimations among between exchange, Expect 1, 0 delineate build and abatement separately. Presently as indicated by our methodology we will consider just those exchanges whose both little scale and vast scale SUM is same i.e. both are 1, 1 or 0, 0 individually. This we do on the grounds that we are just keen on discovering the affiliation if both little and substantial scale organizations expand or diminish in the meantime. So the first exchange table will get minimized. This demonstrated that piece based mining calculation gets precise results with less time and space multifaceted nature when contrasted with FITI algorithm.[9]

As in granule mining, section based methodology sections the information sets into parts for handling consequently lessening the data size of information sets encouraged to the calculation. As opposed to granule mining, in section based mining the condition and choice traits are summed for acquiring summed up affiliation rules. In the wake of performing part based mining the yield summed up guidelines are given to Genetic Algorithm.

As selection of companies by user-

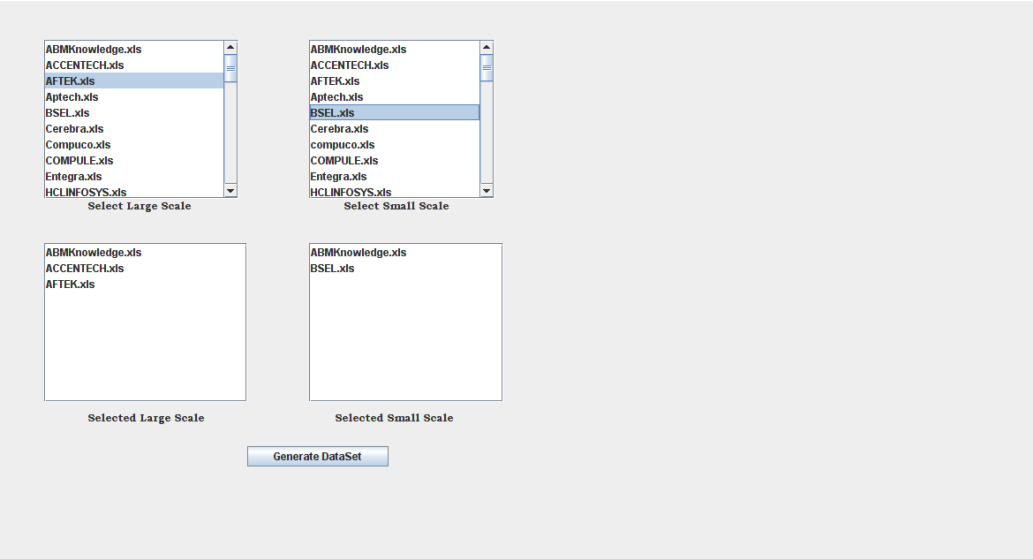
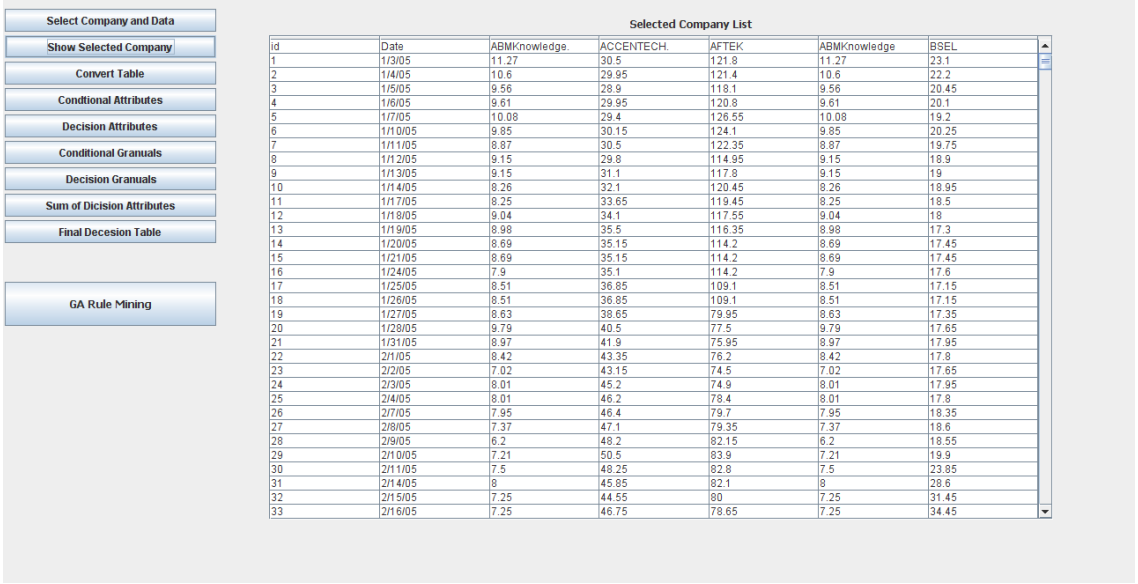


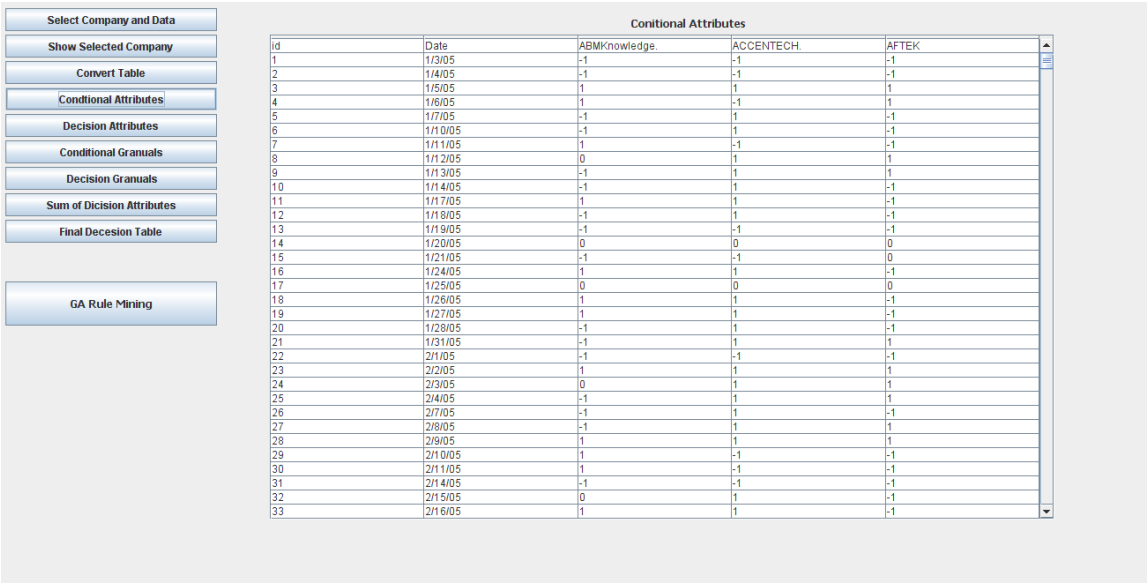
Fig.1 Selection of companies

The data of selected companies is extracted as shown in following figure-



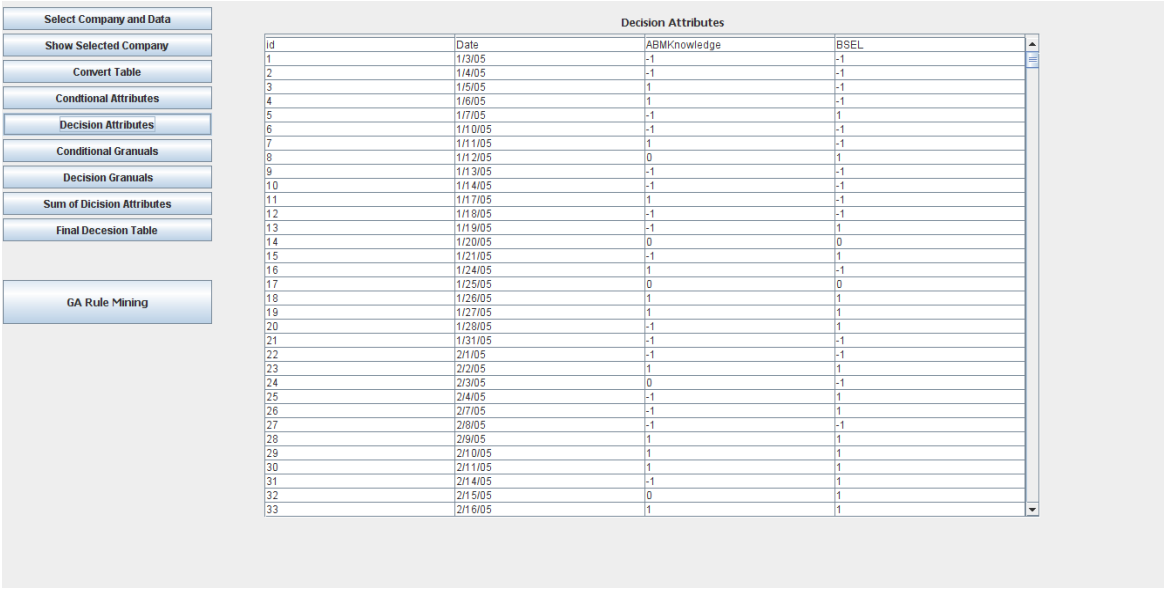
id	Date	ABMKnowledge.	ACCENTECH.	AFTEK	ABMKnowledge	BSEL
1	1/3/05	11.27	30.5	121.8	11.27	23.1
2	1/4/05	10.6	29.95	121.4	10.6	22.2
3	1/5/05	9.56	28.9	118.1	9.56	20.45
4	1/6/05	9.61	29.95	120.8	9.61	20.1
5	1/7/05	10.08	29.4	126.55	10.08	19.2
6	1/10/05	9.85	30.15	124.1	9.85	20.25
7	1/11/05	8.87	30.5	122.35	8.87	19.75
8	1/12/05	9.15	29.8	114.95	9.15	18.9
9	1/13/05	9.15	31.1	117.8	9.15	19
10	1/14/05	8.26	32.1	120.45	8.26	18.95
11	1/17/05	8.25	33.65	119.45	8.25	18.5
12	1/18/05	9.04	34.1	117.55	9.04	18
13	1/19/05	8.98	35.5	116.35	8.98	17.3
14	1/20/05	8.69	35.15	114.2	8.69	17.45
15	1/21/05	8.69	35.15	114.2	8.69	17.45
16	1/24/05	7.9	35.1	114.2	7.9	17.6
17	1/25/05	8.51	36.85	109.1	8.51	17.15
18	1/26/05	8.51	36.85	109.1	8.51	17.15
19	1/27/05	8.63	38.65	79.95	8.63	17.35
20	1/28/05	9.79	40.5	77.5	9.79	17.65
21	1/31/05	8.97	41.9	75.95	8.97	17.95
22	2/1/05	8.42	43.35	76.2	8.42	17.8
23	2/2/05	7.02	43.15	74.5	7.02	17.65
24	2/3/05	8.01	45.2	74.9	8.01	17.95
25	2/4/05	8.01	46.2	78.4	8.01	17.8
26	2/7/05	7.95	46.4	79.7	7.95	18.35
27	2/8/05	7.37	47.1	79.35	7.37	18.6
28	2/9/05	6.2	48.2	82.15	6.2	18.55
29	2/10/05	7.21	50.5	83.9	7.21	19.9
30	2/11/05	7.5	48.25	82.8	7.5	23.85
31	2/14/05	8	45.85	82.1	8	28.6
32	2/15/05	7.25	44.55	80	7.25	31.45
33	2/16/05	7.25	46.75	78.65	7.25	34.45

Fig.2 Show Selected companies data



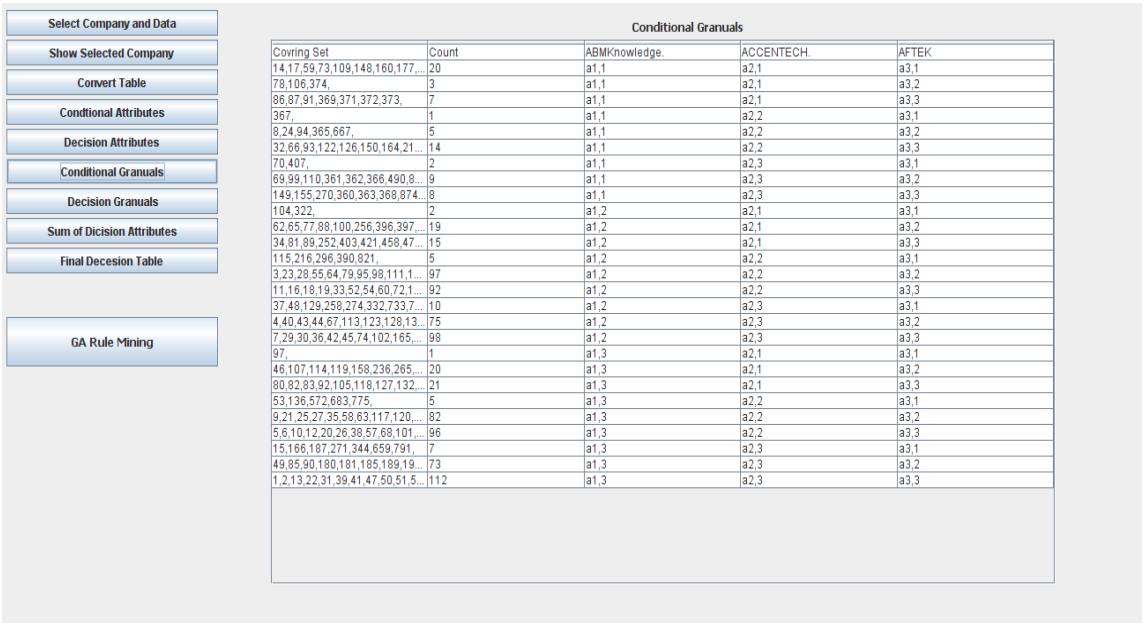
Id	Date	ABMKnowledge.	ACCENTECH.	AFTEK
1	1/3/05	-1	-1	-1
2	1/4/05	-1	-1	-1
3	1/5/05	1	1	1
4	1/6/05	1	-1	1
5	1/7/05	-1	1	-1
6	1/10/05	-1	1	-1
7	1/11/05	1	-1	-1
8	1/12/05	0	1	1
9	1/13/05	-1	1	1
10	1/14/05	-1	1	-1
11	1/17/05	1	1	-1
12	1/18/05	-1	1	-1
13	1/19/05	-1	-1	-1
14	1/20/05	0	0	0
15	1/21/05	-1	-1	0
16	1/24/05	1	1	-1
17	1/25/05	0	0	0
18	1/26/05	1	1	-1
19	1/27/05	1	1	-1
20	1/28/05	-1	1	-1
21	1/31/05	-1	1	1
22	2/1/05	-1	-1	-1
23	2/2/05	1	1	1
24	2/3/05	0	1	1
25	2/4/05	-1	1	1
26	2/7/05	-1	1	-1
27	2/8/05	-1	1	1
28	2/9/05	1	1	1
29	2/10/05	1	-1	-1
30	2/11/05	1	-1	-1
31	2/14/05	-1	-1	-1
32	2/15/05	0	1	-1
33	2/16/05	1	1	-1

Fig.3. Formation of conditional attributes



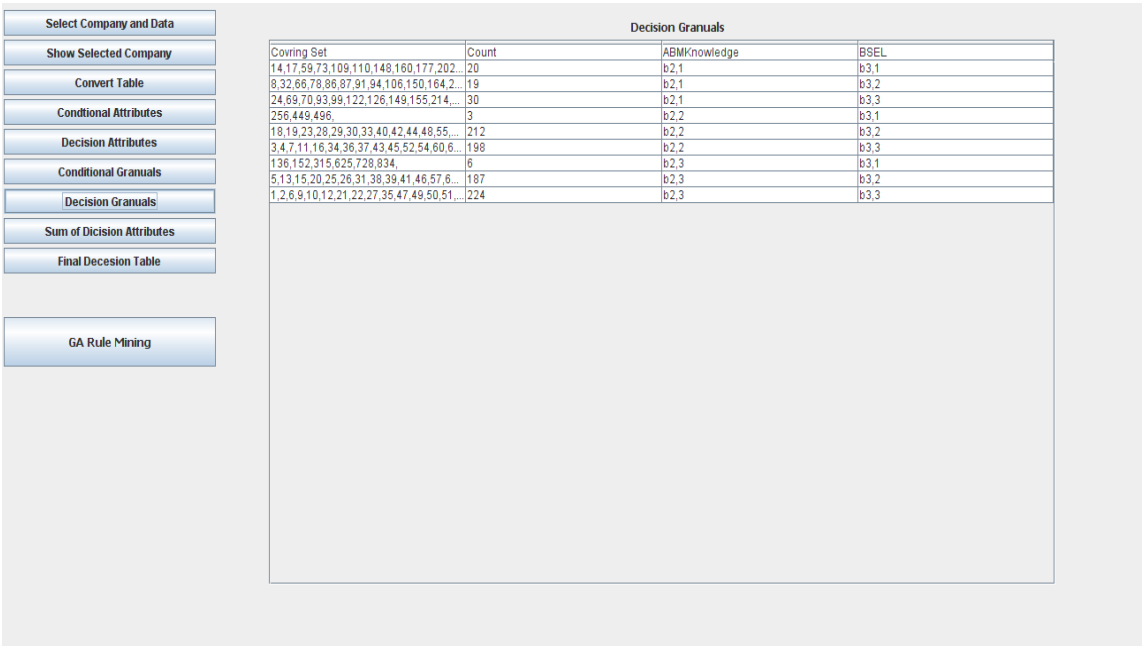
id	Date	ABMKnowledge	BSEL
1	1/3/05	-1	-1
2	1/4/05	-1	-1
3	1/5/05	1	-1
4	1/6/05	1	-1
5	1/7/05	-1	1
6	1/10/05	-1	-1
7	1/11/05	1	-1
8	1/12/05	0	1
9	1/13/05	-1	-1
10	1/14/05	-1	-1
11	1/17/05	1	-1
12	1/18/05	-1	-1
13	1/19/05	-1	1
14	1/20/05	0	0
15	1/21/05	-1	1
16	1/24/05	1	-1
17	1/25/05	0	0
18	1/26/05	1	1
19	1/27/05	1	1
20	1/28/05	-1	1
21	1/31/05	-1	-1
22	2/1/05	-1	-1
23	2/2/05	1	1
24	2/3/05	0	-1
25	2/4/05	-1	1
26	2/7/05	-1	1
27	2/8/05	-1	-1
28	2/9/05	1	1
29	2/10/05	1	1
30	2/11/05	1	1
31	2/14/05	-1	1
32	2/15/05	0	1
33	2/16/05	1	1

Fig.4. Formation of decision attributes



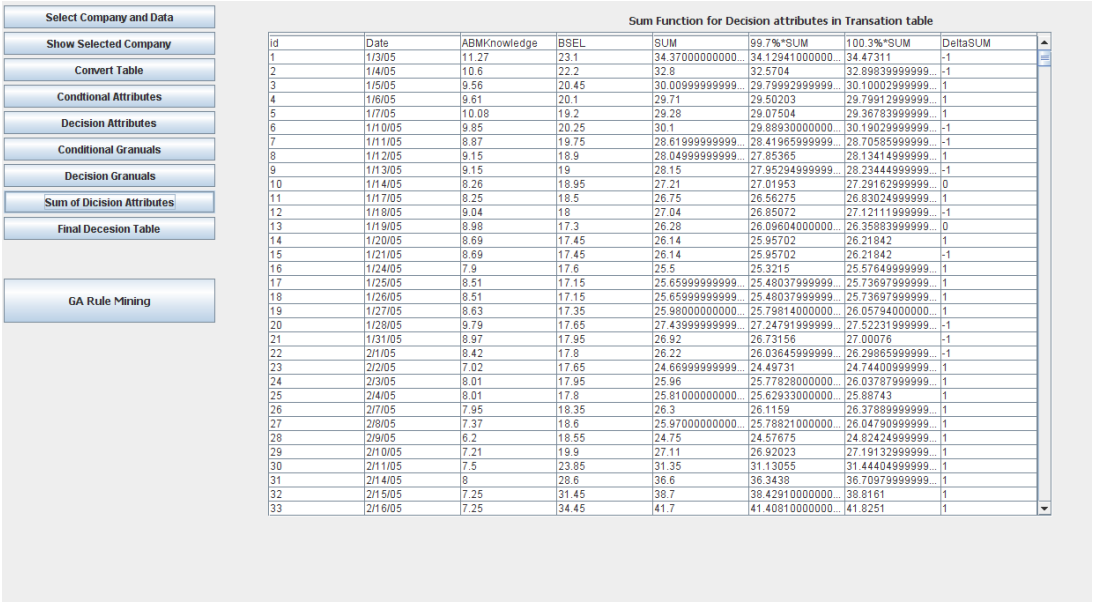
Covring Set	Count	ABMKnowledge.	ACCENTECH.	AFTEK
14,17,59,73,109,148,160,177,...	20	a1,1	a2,1	a3,1
78,106,374,	3	a1,1	a2,1	a3,2
86,87,91,369,371,372,373,	7	a1,1	a2,1	a3,3
367,	1	a1,1	a2,2	a3,1
8,24,94,365,667,	5	a1,1	a2,2	a3,2
32,66,93,122,126,150,164,21...	14	a1,1	a2,2	a3,3
70,407,	2	a1,1	a2,3	a3,1
69,99,110,361,362,366,490,8...	9	a1,1	a2,3	a3,2
149,155,270,360,363,368,874...	8	a1,1	a2,3	a3,3
104,322,	2	a1,2	a2,1	a3,1
62,65,77,88,100,256,396,397,...	19	a1,2	a2,1	a3,2
34,81,89,252,403,421,458,47...	15	a1,2	a2,1	a3,3
115,216,296,390,821,	5	a1,2	a2,2	a3,1
3,23,28,55,64,79,95,98,111,1...	97	a1,2	a2,2	a3,2
11,16,18,19,33,52,54,60,72,1...	92	a1,2	a2,2	a3,3
37,48,129,258,274,332,733,7...	10	a1,2	a2,3	a3,1
4,40,43,44,67,113,123,128,13...	75	a1,2	a2,3	a3,2
7,29,30,36,42,45,74,102,165,...	98	a1,2	a2,3	a3,3
97,	1	a1,3	a2,1	a3,1
46,107,114,119,158,236,265,...	20	a1,3	a2,1	a3,2
80,82,83,92,105,118,127,132,...	21	a1,3	a2,1	a3,3
53,136,572,683,775,	5	a1,3	a2,2	a3,1
9,21,25,27,35,58,63,117,120,...	82	a1,3	a2,2	a3,2
5,6,10,12,20,26,38,57,68,101,...	96	a1,3	a2,2	a3,3
15,166,187,271,344,659,791,	7	a1,3	a2,3	a3,1
49,85,90,180,181,185,189,19...	73	a1,3	a2,3	a3,2
1,2,13,22,31,39,41,47,50,51,5...	112	a1,3	a2,3	a3,3

Fig.5. Generation of conditional granules



Covring Set	Count	ABMKnowledge	BSEL
14,17,59,73,109,110,148,160,177,202...	20	b2,1	b3,1
8,32,66,78,86,87,91,94,106,150,164,2...	19	b2,1	b3,2
24,69,70,93,99,122,126,149,155,214,...	30	b2,1	b3,3
256,449,496,	3	b2,2	b3,1
18,19,23,28,29,30,33,40,42,44,48,55,...	212	b2,2	b3,2
3,4,7,11,16,34,36,37,43,45,52,54,60,6...	198	b2,2	b3,3
136,152,315,625,728,834,	6	b2,3	b3,1
5,13,15,20,25,26,31,38,39,41,46,57,6...	187	b2,3	b3,2
1,2,6,9,10,12,21,22,27,35,47,49,50,51,...	224	b2,3	b3,3

Fig.6. Generation of decision granules



id	Date	ABMKnowledge	BSEL	SUM	99.7%*SUM	100.3%*SUM	DeltaSUM
1	1/3/05	11.27	23.1	34.37000000000...	34.12941000000...	34.47311	-1
2	1/4/05	10.6	22.2	32.8	32.5704	32.89839999999...	-1
3	1/5/05	9.56	20.45	30.00999999999...	29.79992999999...	30.10002999999...	1
4	1/6/05	9.61	20.1	29.71	29.50203	29.79912999999...	1
5	1/7/05	10.08	19.2	29.28	29.07504	29.36783999999...	1
6	1/10/05	9.85	20.25	30.1	29.88930000000...	30.19029999999...	-1
7	1/11/05	8.87	19.75	28.61999999999...	28.41965999999...	28.70585999999...	-1
8	1/12/05	9.15	18.9	28.04999999999...	27.85365	28.13414999999...	1
9	1/13/05	9.15	19	28.15	27.95294999999...	28.23444999999...	-1
10	1/14/05	8.26	18.95	27.21	27.01953	27.29162999999...	0
11	1/17/05	8.25	18.5	26.75	26.56275	26.83024999999...	1
12	1/18/05	9.04	18	27.04	26.85072	27.12111999999...	-1
13	1/19/05	8.98	17.3	26.28	26.09604000000...	26.35883999999...	0
14	1/20/05	8.69	17.45	26.14	25.95702	26.21842	1
15	1/21/05	8.69	17.45	26.14	25.95702	26.21842	-1
16	1/24/05	7.9	17.6	25.5	25.3215	25.57649999999...	1
17	1/25/05	8.51	17.15	25.65999999999...	25.48037999999...	25.73697999999...	1
18	1/26/05	8.51	17.15	25.65999999999...	25.48037999999...	25.73697999999...	1
19	1/27/05	8.63	17.35	25.98000000000...	25.79814000000...	26.05794000000...	1
20	1/28/05	9.79	17.65	27.43999999999...	27.24791999999...	27.52231999999...	-1
21	1/31/05	8.97	17.95	26.92	26.73156	27.00076	-1
22	2/1/05	8.42	17.8	26.22	26.03645999999...	26.29865999999...	-1
23	2/2/05	7.02	17.65	24.66999999999...	24.49731	24.74400999999...	1
24	2/3/05	8.01	17.95	25.96	25.77828000000...	26.03787999999...	1
25	2/4/05	8.01	17.8	25.81000000000...	25.62933000000...	25.88743	1
26	2/7/05	7.95	18.35	26.3	26.1159	26.37889999999...	1
27	2/8/05	7.37	18.6	25.97000000000...	25.78821000000...	26.04790999999...	1
28	2/9/05	6.2	18.55	24.75	24.57675	24.82424999999...	1
29	2/10/05	7.21	19.9	27.11	26.92023	27.19132999999...	1
30	2/11/05	7.5	23.85	31.35	31.13055	31.44404999999...	1
31	2/14/05	8	28.6	36.6	36.3438	36.70979999999...	1
32	2/15/05	7.25	31.45	38.7	38.42910000000...	38.8161	1
33	2/16/05	7.25	34.45	41.7	41.40810000000...	41.8251	1

Fig.7. SUM function calculation

Further optimization is performed using genetic algorithm by selecting rules having maximum confidence.

5. Genetic algorithm for optimization

Genetic algorithm (GA) is a search heuristic that mimics the process of natural Evolution. This heuristic is routinely used to generate useful solutions to optimization and search problems. Genetic algorithms belong to the larger class of evolutionary algorithms, which generate solutions to optimization problems using techniques inspired by natural evolution, such as inheritance, mutation, selection, and crossover.[4]

The genetic algorithms are important when discovering association rules because they work with global search to discover the set of items frequency and they are less complex than other algorithms often used in data mining. The genetic algorithms for discovery of association rules have been put into practice in real problems.

Genetic Operation: The Genetic algorithm focus the inquiry ability and merging of the calculation. Genetic algorithm hold the determination hybrid and change on the populace and produce the new population. [5]

5.1 Select operation: Selection is a method of selecting an individual from a population of individuals in a genetic algorithm. Selection involves running several competitions among a few individuals chosen at random from the population.

5.2 Crossover operation: In Genetic algorithm, crossover is a genetic administrator used to fluctuate the programming of a chromosome or chromosomes starting with one generation then onto the next. It is comparable to multiplication and organic hybrid, whereupon genetic calculations are based. crossover is a procedure of taking more than one guardian arrangements and delivering a tyke arrangement from them.

5.3 Mutation operation: Mutation is a hereditary administrator used to keep up genetic assorted qualities from one era of a populace of genetic calculation chromosomes to the following. It is practically equivalent to natural change. Mutation changes one or more quality values in a chromosome from its starting state. In mutation, the arrangement may change completely from the past arrangement. Thus GA can come to better arrangement by utilizing mutation. Change happens amid advancement as per a client perceptible transformation likelihood.

IV. CONCLUSION AND FUTHER WORK

The approach of this project is to minimize the length of transaction table of stock market on the basis of common features among attributes indirectly to minimize processing complexity. In FITI (First Intra Then Inter) algorithm approach it is difficult to process an information table with many attributes and long intervals for inter transaction associations. This results into large amount of time and cost in processing the data. Fragment based mining groups all attribute once and perform operations group-wise instead of single attribute, thus fragment based produces rules but they are referred as generalized rules as their count more, so further genetic algorithm is applied for the purpose of optimization. The result can be made more promising, by checking different crossover and mutation rates and finally comparing and selecting the rates which gives more optimization.

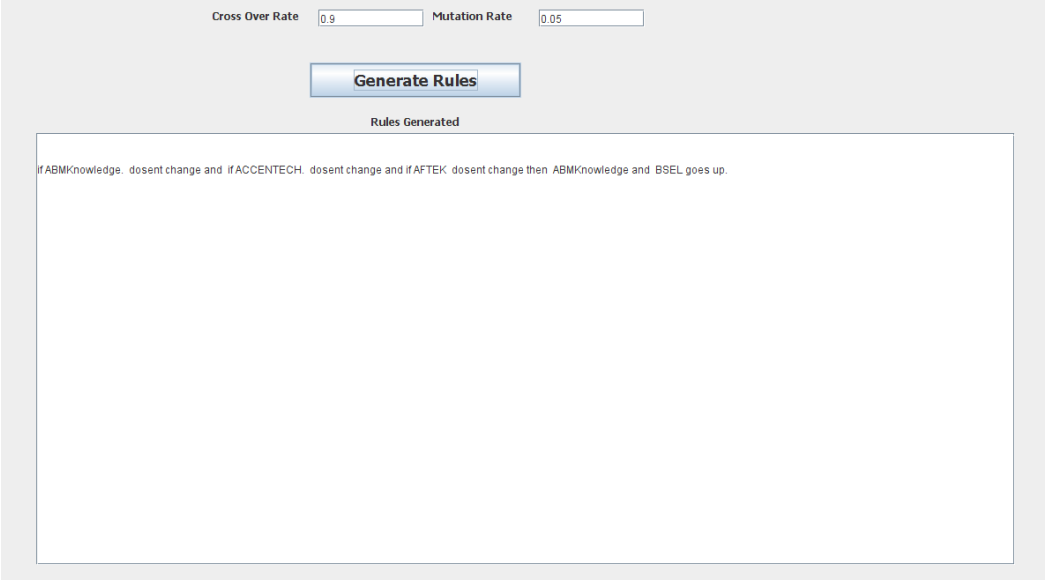


Fig.8. Final predicted rules

REFERENCES

[1] Harya Widiputra, Bagus Pahlevi "Inter-transaction Association Rule Mining in the Indonesia Stock Exchange Market" International Conference on Uncertainty Reasoning and Knowledge Engineering,2012.
[2] Rajesh V. Argiddi, S. S. Apte," Future Trend Prediction of Indian IT Stock Market using Association Rule Mining of Transaction data", IJCA, Volume 39– No.10, February 2012.
[3] Rajesh V. Argiddi, Sulabha S. Apte," Fragment Based Approach to Forecast Association Rules from Indian IT Stock Transaction Data" IJCSIT, Vol. 3 (2) , 2012
[4] Anubha Sharma, Nirupma Tivari, "A Survey of Association Rule Mining Using Genetic Algorithm "IJCAIT, Vol. I, Issue II, August 2012.
[5] Anandhavalli M., Suraj Kumar Sudhanshu, " Optimized association rule mining using genetic algorithm" Bioinfo Publications, Advances in Information Mining, ISSN: 0975–3265, Volume 1, Issue 2, 2009.
[6] Rajesh V. Argiddi, Sulabha S. Apte," A Study Of Association Rule Mining In Fragmented Item-Sets For Prediction Of Transactions Outcome In Stock Trading Systems", IJCET,IAEME, Volume 3, Issue 2, July- September (2012).
[7] Wanzhong Yang,"Granule Based Knowledge Representation for Intra and Inter Transaction Association Mining, Queensland University of Technology", July 2009.
[8] Rajesh V. Argiddi, Sulabha S. Apte," An Evolutionary Fragment Mining Approach to extract Stock market behaviour for investment portfolio" Volume 4, Issue 5, September - October (2013), © IAEME, July 2009.
[9] J. Hipp, U. Gntzer, and G. Nakhaeizade, "Algorithms for association rule mining a general survey and comparison", ACM SIGKDD Explorations Newsletter, New York, USA, 2000.
[10] Professor Lee "Apriori Algorithm Review for Finals" Spring 2007.
[11] Ole KristianFivelstad "Temporal Text Mining" Norwegian University of Science and Technology, June 2007.
[12] KannikaNirai Vaani M, E Ramaraj "An integrated approach to derive effective rules from association rule mining using genetic algorithm" IEEE2013 International Conference.

[13] Anandhavalli M, Suraj Kumar Sudhanshu, Ayush Kumar and Ghose M. K. "Optimized association rule mining using genetic algorithm", Advances in Information Mining, ISSN: 0975–3265, Volume 1, Issue 2, 2009 .
[14] Pei-Chann Chang , Chin-Yuan Fan , Jun-Lin Lin, "Trend discovery in financial time series data using a case based fuzzy decision tree" 2011 Elsevier ScienceDirect.
[15] R. Agrawal, T. Imielinski, and A. Swami. "Mining association rules between sets of items in large databases". In Proceedings of the ACM SIGMOD International Conference on Management of Data (ACM SIGMOD '93), pages 207216, Washington, USA, May 1993.
[16] Agarwal, Srikant April 2005. Fast Algorithm for Mining Association Rules.
[17] Rakesh Agrawal, Heikki Mannila, Ramakrishnan Srikant. Fast Discovery of Association.
[18] HongjunLu ,Jiawei Han , Ling Feng Jan 2003.Efficient Mining of Intertransaction Association Rules. Qing Li, Ling Feng, Allan Wong June 2005. From intra-transaction to generalized inter-transaction: Landscaping multidimensional contexts in association rule mining.
[19] M. Chen, C. Huang, H. Wu, M. Hsu, F. Hsu (2005). A Data Mining Technique to Grouping Customer Orders in Warehouse Management System. The Fourth IEEE International Workshop on Soft Computing as Trans disciplinary Science and Technology.

