EXTRACTION AND ANALYSIS ASSOCIATION RULE ON THE DATA INDONESIA STOCK EXCHANGE WITH THE APRIORI ALGORITHM

Deden Prayitno 1), Riska Hanifah 2)

1) Prodi Sistem Komputer, FTI, Perbanas Institut.Jl. Perbanas, Karet Kuningan, Setiabudi, Jakarta, Indonesia
2) Prodi Sistem Informasi, FTI, Perbanas Institut.Jl. Perbanas, Karet Kuningan, Setiabudi, Jakarta, Indonesia

riska_hanifah@yahoo.com, dedenp2011@gmail.com

Abstract

This study was done to identify the relationships between issuers on the movement of stock prices and also something to do with the composite stock price index (IHSG). These relationships can form a positive-valued (p), negative (n) or equal (e). All issuers who becomes the object of the research was recorded BEI (Indonesia stock exchange) were classified into nine sectors of shares according to the classification of industries. The population used is the movement of stock prices that are on a three-year period, namely 2011 – 2013. The sample used was 18 companies with sample withdrawal methods are purposive sampling method was performed (judging sampling). Analysis techniques using Association Rule Mining (ARM) with the algorithm Judgmental and data analysis using a program applications WEKA 6.3.10. Apriori algorithms can describe the pattern of stock price movements and their relationship are inter-related. The process of measurement for the amount of stock value by looking at the relationship of large or small values on support and confidence level. Research results showed the existence of a relationship between shares traded during a period of three years and also something to do with the composite stock price index (IHSG), where this type of relationship is shown through the categories of values, either a direct relationship (positive), reverse (negative) or the absence of changes to the change from the other stock value (equal).

Keyword — IHSG, Association Rule Mining, Stock, Data Mining, Apriori Algorithm

1. INTRODUCTION

The development of a wide range of business activities that exist today, happened so rapidly and in large quantities, can be seen many mega projects that have been built and being built, particularly in big cities. Such development requires a large capital course for long periods of time. For a public company (Tbk) which in General has many subsidiaries, to get a large capital one of the stock sale proceeds. The capital markets [7] which maintained the Indonesia Stock Exchange (BEI) is a means of funding for the company as well as other institutions (e.g. Government), and as a means for investing activities. Prices of shares traded of course relates to the health of the company in managing finance functions, when the company was growing well, usually followed by a positive share price. Stock buyers (investors) will be highly selective and careful in choosing forms and kinds of businesses that will become his business partner, given not least many types of businesses that failed with major losses occur every time. Investing in stocks is certainly quite seductive because it gives a considerable advantage in the future [8], but the impact of the risk is big enough anyway. Investors need complete information from businesses that want to become partners, needs the reliable analyzer capable of describing the technical complexity of a business, to be easy to understand and interesting, so investors become confident in investing. Many types of businesses that offer cooperation in investment capital by providing a large enough profit expectations is an investment in the form of shares. More than 500 public company [11] already listed on the Indonesia stock exchange (BEI).

Current information systems have grown rapidly, the various activities of the business world is almost entirely using computer aid, with the power of data processing speed, accuracy, and ease of being able to adapt as required, then the computer has become one of the major factors to support the success of a business. The ability of a particular application system through computer, capable of analyzing data summarizing and extracting knowledge
from large amounts of data, which will not be able to when treated conventionally. Disciplines studied methods to extract knowledge or find a pattern of data to obtain important information is implicit and previously unknown is data mining [5].

The technique for extracting data from a data mining Association rules, one of which is, by the rules of associative relationship can be found between the data and the rules of the existing data to bring up the value of the degree of confidence (confidence level) of the relationship. The recorded data is a fact and does not carry the meaning while knowledge is patterns, rules or models that emerge from the data. As an example of the data pattern is movement up and down the stock price change at any time, the pattern can be extracted in a period 3 to 5 years that has happened (the past). Techniques of data mining algorithm using the method of association rule mining with WEKA 3.6.10 program application [6]. Data extraction results can give you the basic knowledge that can be used for the selection of investment in stocks is good.

1.1. The formulation of the problem
With respect to the background of the above then problem formulation in this research are:

a. How to look at the pattern of rise and fall of the IHSG share price value and stock value of 9 sector shares on IDX for 3 years (2011-2014).

b. How to perform measurements of the stock value by using WEKA 3.6.10 and a priori algorithm that can show the level of trust relationships between the indicated stock with a value of confidence level.

c. How should investors assess the IHSG movement and various companies stock prices as well as knowing the relationships between existing stock on the stock sector 9 in BEI.

1.2. Research objectives
a. Build a knowledge base for investment fields shares by way of stock price movements on the extraction of the stock exchange of Indonesia by using association rule mining and apriori algorithm that can be used by investors as one of the parameters of a consideration in buying shares in the future.

b. To understand the movement of a stock by knowing the influence of IHSG shares with the 9 stock sector.

c. To determine confidence level of any relationships between stocks on stock and sector 9 IHSG in the BEI, using WEKA 3.6.10 and apriori algorithm.

2. THE RESULTS OF PREVIOUS STUDIES AND STAGES OF RESEARCH
2.1. Previous research results
The capital market by definition of Suad husnan 2005 is research done by Azhari and Anshori from the Faculty of MATHEMATICS and NATURAL SCIENCES, the University of Gajah Mada that is doing research on some of the companies whose shares are BEI in 2009, had concluded that the use of data mining techniques with ARM and apriori algorithm can be used to look at the relationship of the price movement of stocks from the stock trading that has taken place. Other research conducted by Ernestasia Siahaan of ITB, by doing research on the stock market, had concluded that data mining can be used to predict stock prices by using a neural network. Further research was carried out by Silvia Jane and the Subanas conducted on the Jakarta Stock Exchange in 2007, obtained the conclusion that data mining can analyze the data in the database is large with many and complex variables, especially for searching for patterns in the data, while the research conducted by k. Senthamarai Kannan Monday on global markets in 2010 concluded that data mining algorithms are able to predict with the closing price of the stock the next day whether increased or decreased, so that it can be used for technical analysis of stock data. Other research conducted by Tampubolon anymore on system supplies health equipment declared concluded the relationship between selling by those types of specific health tools, so that the inventory can be managed by grouping types of goods sold by a positive relationship.

3. RESULTS AND DISCUSSION
Testing using methods of Description (the Description Methods) and function Analysis Association (Association Rule Mining) and the apriori algorithms [4]. Application to conduct done with WEKA 3.6 [12] generating associative rule shown with two parameters, support (supporting values) that is the percentage of such items in the combination of database and confidence (the value of certainty) that the strong relationships between items in the associative rules.

3.1 Association Rule Mining (ARM) Apriori Algorithm
Data testing performed on the existing stock sector 9 in BEI [14] for the movement of the share price during the last 3 years (2011-2014). In testing data is implemented with 2 (two) testing. First use the data stock that has the biggest stake in each sector coupled with the value of the IHSG, next the second stock using data that has the smallest stock value on each sector coupled with the value of the IHSG.

3.1.1 Data Set 1
Testing the value of the largest stocks in each sector coupled with the value of the stock price index (IHSG), to produce optimal confidence values is done by changing the value of two variables is vulnerable in the a priori algorithm i.e. Minimum Support and Minimum Metric (Confidence). as can be seen in the table below:

<table>
<thead>
<tr>
<th>No</th>
<th>Name Stock</th>
<th>Code Share</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Astra Internasional Tbk.</td>
<td>ASII</td>
</tr>
<tr>
<td>2</td>
<td>MNC Investama Tbk.</td>
<td>BHIT</td>
</tr>
<tr>
<td>3</td>
<td>Bank Internasional Indonesia Tbk.</td>
<td>BNII</td>
</tr>
<tr>
<td>4</td>
<td>PP London Sumatra Indonesia Tbk.</td>
<td>LSIP</td>
</tr>
<tr>
<td>5</td>
<td>Energi Mega Persada Tbk.</td>
<td>ENRG</td>
</tr>
<tr>
<td>6</td>
<td>Kalbe Farma Tbk.</td>
<td>KLBF</td>
</tr>
<tr>
<td>7</td>
<td>Pakuwon Jati Tbk.</td>
<td>PWON</td>
</tr>
<tr>
<td>8</td>
<td>Sekawan Inti Pratama Tbk.</td>
<td>SIAP</td>
</tr>
<tr>
<td>9</td>
<td>Telekomunikasi Indonesia Tbk.</td>
<td>TLKM</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 3.2 Test of Data Test 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimum support: 0.1 &amp; Minimum metric (confidence): 0.65 &amp; Rules: 100</td>
</tr>
<tr>
<td>No</td>
</tr>
<tr>
<td>----</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>4</td>
</tr>
<tr>
<td>5</td>
</tr>
<tr>
<td>6</td>
</tr>
<tr>
<td>7</td>
</tr>
<tr>
<td>8</td>
</tr>
</tbody>
</table>

9. LSIP=N ENRG=N PWON=N 107 ==> ASII=N 84 | 0.78
10. ASII=N PWON=N SIAP=E 111 ==> ENRG=N 87 | 0.78
11. ENRG=N KLBF=N PWON=N 105 ==> ASII=N 82 | 0.78
12. ENRG=E 105 ==> SIAP=P 87 | 0.77
13. BHIT=P PWON=P 115 ==> ASII=P 89 | 0.77
14. ASII=N PWON=N TLKM=N 101 ==> ENRG=N 78 | 0.77
15. ASII=N PWON=N 161 ==> ENRG=N 124 | 0.77
16. BHIT=N PWON=N 134 ==> ENRG=N 103 | 0.77
17. BHIT=N KLBF=N 133 ==> ASII=N 102 | 0.76
18. LSIP=N PWON=N 140 ==> ENRG=N 107 | 0.76
19. TLKM=E 163 ==> SIAP=E 124 | 0.76
20. BNII=N PWON=N 111 ==> ENRG=N 84 | 0.75
21. ENRG=N PWON=N SIAP=E 116 ==> ASII=N 87 | 0.75
22. KLBF=P PWON=P 131 ==> ASII=P 98 | 0.75
23. ASII=N BHIT=N ENRG=N 110 ==> PWON=N 82 | 0.74
24. BHIT=P ENRG=P 117 ==> ASII=P 87 | 0.74
25. ENRG=N PWON=N 168 ==> ASII=N 124 | 0.74
26. PWON=N TLKM=N 137 ==> ASII=N 101 | 0.74
27. ASII=N LSIP=N ENRG=N 114 ==> PWON=N 84 | 0.74
28. JKSE=N BNII=E 136 ==> SIAP=E 100 | 0.73
29. KLBF=N PWON=N 132 ==> ASII=N 97 | 0.73
30. BHIT=N BNII=N 113 ==> ENRG=N 83 | 0.73
31. BHIT=P KLBF=P 128 ==> ASII=P 94 | 0.73
32. BHIT=E PWON=E 116 ==> SIAP=E 85 | 0.73
33. JKSE=N BHIT=E 118 ==> SIAP=E 86 | 0.73
34. BNII=E PWON=E 114 ==> SIAP=E 83 | 0.73
35. BHIT=E BNII=E 128 ==> SIAP=E 93 | 0.73
36. ASII=P KLBF=P TLKM=P 106 ==> SIAP=E 77 | 0.73
37. JKSE=N PWON=P 113 ==> SIAP=E 82 | 0.72
38. JKSE=P BNII=P 116 ==> SIAP=E 84 | 0.72
39. BHIT=N LSIP=N 134 ==> ENRG=N 97 | 0.72
40. JKSE=P LSIP=N 151 ==> SIAP=E 109 | 0.72
41. BHIT=N KLBF=N 133 ==> ENRG=N 96 | 0.72
42. BNII=N PWON=N 111 ==> ASII=N 80 | 0.72
43. KLBF=P TLKM=P 157 ==> SIAP=E 113 | 0.72
44. ASII=N ENRG=N KLBF=N 114 ==> PWON=N 82 | 0.72
45. ENRG=P KLBF=P 135 ==> ASII=P 97 | 0.72

161
Table 3.2 is the testing Data from Test 1 that performed a few times so get good confidence value, where the minimum support that used 0.1, the minimum metric (confidence) that are used and the number of rules that 0.65 sought as many as 100 shares for a period of three years, namely the years 2011 to 2013. To the number of rule 100 was able to describe the relationship of all companies (stocks) and stock price index (IHSG).

Number of the rule that is displayed in the table above by as much as 50 rule as an overview of the relationship of each of the different categories of shares, with a value of confidence level. The number one rule in the table 4.2 shows if the stock shares, LSIP ASII and the stock fell as much as 98 times the PWON, shown with the requirement of N, will be followed by the stock down as much as 84 times ENRG, thus generating confidence value of 0.86, meaning that as many as 86% of the occurrences of stock movements that descend from ENRG contains stock ASII, LSIP and PWON, also went down. For another example that illustrates the rule that contains the value of the stock price index (IHSG) (stock code JKSE) against other stock shown in rule 37, i.e.

If the stock goes down, the indicated JKSE with N categories and shares rose PWON (P) that appear 113 times, will be followed by categories with equal SIAP stock (E) appearing as much as 82 times with confidence, resulting in a value of 0.72 confidence, that means as many as 72% of occurrences equal from the movement SIAP stock containing stock and stock price index (IHSG) on certain days experiencing the value goes down. The reality varies inversely with the stock rose with the incident emerged PWON the same time as many as 113 times and followed by shares SIAP with equal category appearing as many as 82 times.

Another example is again in rule 38 , that is if JKSE shares rose, shown with p and category BNII shares rose as many as 116 times, SIAP to be followed by shares worth equal shown with category E times as many as 84 , thus producing value 0.72 confidence.

3.1.2 Data Set 2

The stock value of the testing done by the smallest value of stocks in each sector coupled with the value of the stock price index (IHSG), to produce optimal confidence values is done by changing the value of two variables is vulnerable in the a priori algorithm i.e. Minimum Support and Minimum Metric (Confidence). as can be seen in the table below:

Table 3.4 Test of Data Test 2

<table>
<thead>
<tr>
<th>No</th>
<th>Name Stock</th>
<th>Code Share</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ARTA=E</td>
<td>KICI=E</td>
</tr>
<tr>
<td>2</td>
<td>RIMO=E</td>
<td>KARW=E</td>
</tr>
<tr>
<td>3</td>
<td>ARTA=E</td>
<td>KICI=E</td>
</tr>
<tr>
<td>4</td>
<td>KICI=E</td>
<td>RIMO=E</td>
</tr>
<tr>
<td>5</td>
<td>ARTA=E</td>
<td>KICL=E</td>
</tr>
<tr>
<td>6</td>
<td>KARW=E</td>
<td>KICI=E</td>
</tr>
<tr>
<td>7</td>
<td>KARW=E</td>
<td>KICL=E</td>
</tr>
<tr>
<td>8</td>
<td>ARTA=E</td>
<td>KICI=E</td>
</tr>
<tr>
<td>9</td>
<td>AALI=F</td>
<td>RIMO=E</td>
</tr>
</tbody>
</table>
Table 3.4 is the testing Data from Test 1 committed several times so get good confidence value, where the minimum support that used 0.1, the minimum metric (confidence) that are used and the number of rules that 0.65 sought as many as 100 shares for a period of three years, namely the years 2011 to 2013. To the number of rule 100 was able to describe the relationship of all companies (stocks) and IHSG. Number of the rule that is displayed in the table above by as much as 50 rule as an overview of the relationship of each of the different categories of shares, with a value of confidence level.

The number one rule in table 4.4 shows if shares ARTA, stocks and shares worth KICI equal with the indicated categories of E as much as 286 times, shown with the requirement of N, will be followed by a total of 281 KARW equal times, resulting in a value of 0.98 means confidence value equal as much as 98% of the emergence of the movement equal KARW stock containing shares ARTAKICI, and RIMO are also equal.

For another example that illustrates the rule that contains the value of the stock price index (IHSG), (stock code Jkke) against other stock shown in rule 5, that is, if the stock rises JKSE, indicated by P and categories of shares and RIMO KARW equal as much as 177 times, will be followed by KICI equal with the indicated categories of E as much as 160 times, thus generating confidence value is 0.9.

4. CONCLUSION:
Association rule mining with the a priori algorithm can build a knowledge base of stock movement database by showing the relationships between stocks. The rise and fall of stock values on the stock sector 9 no effect with a value of IHSG. The movement of the stock value of its assets has little confidence level is higher than the value of its assets stock movement is great. Information current confidence level indicated by every rule of relationships between stocks, can give you the ease to investors as one of the primary considerations for determining the investment on a stock.

REFERENCES


Referensi online:
[10] The names of the sector and code shares BEI (http://www.idx.co.id/), access date October 10, 2014

