

Analysis of Determination of Strategy Promotion using Apriori Algorithm

Rinna Rachmatika*, and Kecitaan Harefa

¹Faculty of Information Technology, Universitas Pamulang, Tangerang, Indonesia

*rinnarachmatika@unpam.ac.id

Abstract. Promotion is an activity or marketing idea commonly carried out by business actors to attract the attention of consumers either to introduce an item or to increase consumers so they want to buy the item. Apriori is an algorithm that can perform several itemset required by knowledge to determine the rules in making information related to determining candidates who will emerge as a result based on the support and trust determined by the analyst. The results of the analysis using the Apriori Algorithm in this study are that the 2 itemset rules that have been made can be used as a strategy for the promotion of goods based on sales transaction data. Because what is produced is high, the highest level of trust is 100%, and the lowest is 84%. Goods that have the highest trust value can be used as a promotional strategy.

1. Introduction

The development of information technology is currently developing very rapidly, especially in the business world or business or trade. Developers and business people to be able to sell their merchandise and get profits will do many ways in attracting consumer interest, one of which is by promotion. Most consumers, when they want to buy the goods they want, they will definitely look for goods at cheap prices or are promoting in advance. Promotion is an activity or marketing idea that is usually done by business actors to attract the attention of consumers both to introduce an item and to influence consumers so they feel they want to buy the item.

Development and business competition in trading based on advances in information technology bring companies to a level of competition that is getting tougher and more open in meeting the demands of customers who are also increasingly high. Businesses in this information technology era must also compete in markets with rapid, complex, global, and customer-focused changes [8]. So that the Supermarket developers and business people must find solutions and think of strategies that can continue to provide benefits so they can maintain and develop this Supermarket business [6].

In competition in the business world requires developers or business people to be able to determine a strategy that can increase sales and find ways to make consumers interested in shopping. There are still many business actors who have not implemented a strategy based on the pattern of sales transactions due to the large amount of data, so the determination of the promotion is still not in accordance with what is commonly purchased by consumers.

Data mining is defined as a process to find meaningful relationships, patterns and new trends by filtering very large data, which is stored in storage, using pattern recognition techniques such as statistical and mathematical techniques [2]. Data mining is the process of extracting information or something important or interesting from the data in a database so as to produce valuable information.



Market basket analysis technique is a technique that adapts the science of data mining. This technique is used to design a strategy for selling and marketing goods through the process of finding associations or relationships between data items from a relational database [9].

From the description above, the author uses the Apriori Algorithm in order to help in forming several combinations of item sets, because this Apriori analysis is famous for using association rule techniques that can analyze shopping basket contents based on transaction data, then testing whether the combination meets the minimum Support and Confidence is the threshold value given by the user. If you meet the Support and Confidence, the results can help in determining the promotion strategy, so consumers can be interested in buying the item when shopping.

2. Method

Sales strategy is very important for companies where the sales strategy is a way to achieve the goals of a company. In designing and developing products, both in the form of services and goods, can not be separated from the marketing concept that aims to meet the needs that satisfy customers. The marketing concept is to describe how a company must operate to plan, determine prices, promote, and distribute goods and services that can satisfy the needs of buyers in achieving company goals [12].

The sales strategy believes that if consumers and business people are left alone, they will not buy adequate quantities of the company's products. Companies must conduct sales and promotion efforts that understand all the needs and desires of all consumers within the scope or limits of existing resources. The process of meeting the needs and desires of consumers is the sales strategy that consists of 4P, namely [12] :

2.1 Product

A product is designed to satisfy the needs of consumers. The product strategy includes a number of decisions regarding the use of quality, trademark packaging, design, and other types of models.

2.2 Price

In addition to the price set for a product sold to consumers, pricing includes several management policies regarding discounts, prices, and so on.

2.3 Place

Placing a product means providing the product in the right place and at the right time. The product distribution sales strategy includes a number of decisions such as location and store area.

2.4 Promotion

Promotion intends to inform and persuade target consumers in terms of the value of the product being sold. The main means of promotion are through advertising, and mass sales.

Apriori's algorithm uses knowledge about frequent known itemsets, to process further information. In the Apriori algorithm to determine candidates that might appear by paying attention to minimum support. The two main processes carried out in the Apriori algorithm, namely [4] :

- Join

In this process each item is combined with other items until no more combinations are formed.

- Prune

In this process, the results of the combined items are then trimmed using the minimum support specified by the user.

Apriori is an algorithm that is very well known in conducting frequent itemset searches using association rule techniques [3]. Association rules are often called Market Basket Analysis. Association rule is a data mining method that aims to find a set of items that often appear together. Generally this association rule is analogous to a shopping basket. From the shopping basket the Supermarket visitors will be able to know what items are often bought together and which items are not. Association rules generally take the form of IF-THEN which combines several items into one, for example: IF A and B THEN C. In theory, several things are used to measure whether a set of items (an item set) often appears together or not, is a Support of an item set, Confidence of an association rule, and several rule selection methods. Support of an item set is a percentage of all transactions that occur that contain the item set [1].

Association rule is a data mining technique for finding associative rules between combinations of items. An example of an associative rule from a purchase analysis at a supermarket is knowing how likely a customer is to buy bread together with milk. With this knowledge, supermarkets can arrange the placement of goods or design marketing campaigns by cursing discount coupons or combinations of certain items.

Association analysis is also known as one of the data mining techniques that is the basis of various other data mining techniques. In particular, one of the stages of association analysis that attracts the attention of many authors to produce efficient algorithms is analysis of high frequency patterns (frequent pattern mining). The importance of an associative rule can be determined by two parameters, namely Support and Confidence. Support (supporting value) is the percentage of the combination of these items in the database, whereas Confidence (certainty value) is the strength of the relationship between items in the association rules [6].

3. Result and discussion

For testing the author will use a sample of 38 consumer purchase transactions. Transactions to be tested are based on secondary data datasets that the author gets from the primary source or primary data. Each transaction has a specific list of items that are translated into variables, namely :

Table 1. List items in variable form

| Variable | |
|---------------------------------------|-------------------------|
| A = HOLDER PRONTO CE | L = SWAB ALCOHOL |
| B = HOLDER STANDARD CE | M = SYR ABG PRESET 3 |
| C = LANCET CONT PINK | N = TUBE CIT PLH 2.7 |
| D = NEEDLE CNV 22 CE | O = TUBE EDTA 3.0 |
| E = NEEDLE FLASHBACK 22 | P = TUBE MICRO EDTA LAV |
| F = NITRIL GLOVES | Q = TUBE SST 3.5 |
| G = POLYSORB 3-0 UND 75CM C13 TSL822 | R = VENFLON SAFETY 18GA |
| H = POLYSORB 4-0 75CM C-13 TSL691 | S = VENFLON SAFETY 20GA |
| I = SOFSILK 1 75CM GS-21 TCS425 | T = WINGSET 23 |
| J = SURGIPRO 4-0 90CM 2XCV-23 TVP557X | U = WINGSET 25 |
| K = SURGIPRO 7-0 75CM CV-1 TVP713X | |

The first stage in this analysis, the authors analyze sales transaction data and then do a percentage of support to find patterns of support that have support that is more than 10%.

Table 2. Support List and Support Percentage

| Variable | Item | Support | Support (%) |
|----------|----------------------------------|---------|-------------|
| A | HOLDER PRONTO CE | 2 | 5.27% |
| B | HOLDER STANDARD CE | 3 | 7.89% |
| C | LANCET CONT PINK | 2 | 5.27% |
| D | NEEDLE CNV 22 CE | 2 | 5.27% |
| E | NEEDLE FLASHBACK 22 | 6 | 15.78% |
| F | NITRIL GLOVES | 4 | 10.50% |
| G | POLYSORB 3-0 UND 75CM C13 TSL822 | 1 | 2.63% |

| Variable | Item | Support | Support (%) |
|----------|-----------------------------------|---------|-------------|
| H | POLYSORB 4-0 75CM C-13 TSL691 | 1 | 2.63% |
| I | SOFSILK 1 75CM GS-21 TCS425 | 1 | 2.63% |
| J | SURGIPRO 4-0 90CM 2XCV-23 TVP557X | 1 | 2.63% |
| K | SURGIPRO 7-0 75CM CV-1 TVP713X | 1 | 2.63% |
| L | SWAB ALCOHOL | 4 | 10,52% |
| M | SYR ABG PRESET 3 | 2 | 5.27% |
| N | TUBE CIT PLH 2.7 | 1 | 2.63% |
| O | TUBE EDTA 3.0 | 11 | 28.90% |
| P | TUBE MICRO EDTA LAV | 3 | 7.80% |
| Q | TUBE SST 3.5 | 13 | 34.21% |
| R | VENFLON SAFETY 18GA | 5 | 13,15% |
| S | VENFLON SAFETY 20GA | 9 | 23.68% |
| T | WINGSET 23 | 4 | 10.50% |
| U | WINGSET 25 | 6 | 15.78% |

Furthermore, items that have more than 10% support are used to determine the frequent itemset process. The itemset that will be obtained is 2 itemset items.

Table 3. List Items with more than 10% support

| Variable | Item | Support | Support (%) |
|----------|---------------------|---------|-------------|
| E | NEEDLE FLASHBACK 22 | 6 | 15.78% |
| F | NITRIL GLOVES | 4 | 10.50% |
| L | SWAB ALCOHOL | 4 | 10.52% |
| O | TUBE EDTA 3.0 | 11 | 28.90% |
| Q | TUBE SST 3.5 | 13 | 34.21% |
| R | VENFLON SAFETY 18GA | 5 | 13.15% |
| S | VENFLON SAFETY 20GA | 9 | 23.68% |
| T | WINGSET 23 | 4 | 10.50% |
| U | WINGSET 25 | 6 | 15.78% |

Frequent Itemset shows itemset that has more than the specified minimum value (minimum). Here the author gives a minimum value of $\phi = 10\%$ of the value of support. So we get the results like the table below :

Table 4. Candidate 2 itemset

| Item | Support | Support (%) |
|--|---------|-------------|
| NEEDLE FLASHBACK 22, NITRIL GLOVES | 0 | 0 |
| NEEDLE FLASHBACK 22, SWAB ALCOHOL | 0 | 0 |
| NEEDLE FLASHBACK 22, TUBE EDTA 3.0 | 2 | 5.26% |
| NEEDLE FLASHBACK 22, TUBE SST 3.5 | 0 | 0 |
| NEEDLE FLASHBACK 22, VENFLON SAFETY 18GA | 0 | 0 |
| NEEDLE FLASHBACK 22, VENFLON SAFETY 20GA | 0 | 0 |

| Item | Support | Support (%) |
|--|----------|---------------|
| NEEDLE FLASHBACK 22, WINGSET 23 | 0 | 0 |
| NEEDLE FLASHBACK 22, WINGSET 25 | 0 | 0 |
| NITRIL GLOVES, SWAB ALCOHOL | 0 | 0 |
| NITRIL GLOVES, TUBE EDTA 3.0 | 0 | 0 |
| NITRIL GLOVES, TUBE SST 3.5 | 0 | 0 |
| NITRIL GLOVES, VENFLON SAFETY 18GA | 0 | 0 |
| NITRIL GLOVES, VENFLON SAFETY 20GA | 0 | 0 |
| NITRIL GLOVES, WINGSET 23 | 0 | 0 |
| NITRIL GLOVES, WINGSET 25 | 0 | 0 |
| SWAB ALCOHOL, TUBE EDTA 3.0 | 0 | 0 |
| SWAB ALCOHOL, TUBE SST 3.5 | 0 | 0 |
| SWAB ALCOHOL, VENFLON SAFETY 18GA | 0 | 0 |
| SWAB ALCOHOL, VENFLON SAFETY 20GA | 0 | 0 |
| SWAB ALCOHOL, WINGSET 23 | 0 | 0 |
| SWAB ALCOHOL, WINGSET 25 | 0 | 0 |
| TUBE EDTA 3.0, TUBE SST 3.5 | 5 | 13.15% |
| TUBE EDTA 3.0, VENFLON SAFETY 18GA | 0 | 0 |
| TUBE EDTA 3.0, VENFLON SAFETY 20GA | 0 | 0 |
| TUBE EDTA 3.0, WINGSET 23 | 0 | 0 |
| TUBE EDTA 3.0, WINGSET 25 | 0 | 0 |
| TUBE SST 3.5, VENFLON SAFETY 18GA | 0 | 0 |
| TUBE SST 3.5, VENFLON SAFETY 20GA | 0 | 0 |
| TUBE SST 3.5, WINGSET 23 | 0 | 0 |
| TUBE SST 3.5, WINGSET 25 | 0 | 0 |
| VENFLON SAFETY 18GA, VENFLON SAFETY 20GA | 3 | 7.80% |
| VENFLON SAFETY 18GA, WINGSET 23 | 0 | 0 |
| VENFLON SAFETY 18GA, WINGSET 25 | 0 | 0 |
| VENFLON SAFETY 20GA, WINGSET 23 | 0 | 0 |
| VENFLON SAFETY 20GA, WINGSET 25 | 0 | 0 |
| WINGSET 23, WINGSET 25 | 4 | 10.52% |

From the above data, if the value $\phi = 10\%$ is set, the combination of the 2 itemset rules generated is :

Table 5. Combination 2 itemset

| Item | Support | Support (%) |
|------------------------------------|---------|-------------|
| TUBE EDTA 3.0, TUBE SST 3.5 | 11 | 28.90% |
| WINGSET 23, WINGSET 25 | 6 | 15.78% |

So the association rules are [6]:

$$Support(A) = \frac{Jumlah\ transaksi\ mengandung\ A}{Total\ Transaksi} \quad (1)$$

Table 6. Association Rules

| Association 2 Itemset | Confidence | |
|--|------------|------|
| If buy Tube Edta 3.0, then will buy Tube SST 3.5 | 11/11 | 100% |
| If buy Tube SST 3.5, then will buy Tube Edta 3.0 | 11/13 | 84% |
| If buy Wingset 23, then will buy Wingset 25 | 6/4 | 100% |
| If buy Wingset 25, then will buy Wingset 23 | 6/6 | 100% |

From table 6 it can be seen that the association rules can be based on previous analyzes, where the minimum Support is 10% and the minimum Confidence is 50%. With the highest Confidence value of 100% and the lowest is 84%. So the strategies that can be used to determine the item promotional items are adjusted to the highest Confidence value, which is 100%.

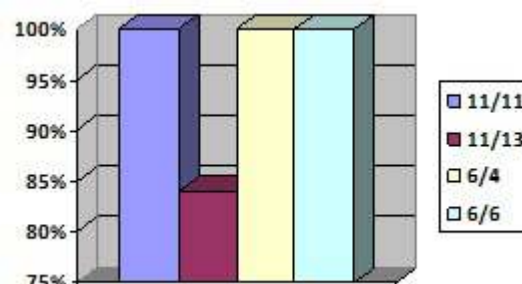


Figure 1. Association Rules Results Diagram

4. Conclusion

Based on the results of research and discussion on how to determine the promotion strategy using the Apriori Algorithm, conclusions are drawn :

- This a priori algorithm can be used to determine the promois strategy, because what is produced has the highest confidence level of 100%, and the lowest is 84%.
- The promotion strategy that can be used is to look at the association rules that have the highest confidence value.

Acknowledgments

The researcher is very aware that in this research process there are things that make us difficult, but because of the encouragement and enthusiasm from various parties, the researcher finally can make this research even though there are still many shortcomings and limitations. Researchers will continue to try to develop this research and learn in order to have broader and better insights.

References

- [1] Agusta, Y. (2008, 4 4). Retrieved 2019, from yudiagusta.wordpress.com: <https://yudiagusta.wordpress.com/2008/08/04/association-rules/>
- [2] Amiruddin, K. E. (2007). *Penerapan Association Rule Mining Pada Data Nomor Unik Pendidik Dan Tenaga Kependidikan Untuk Menemukan Pola Sertifikasi Guru*. Teknik Elektro FTI, Surabaya.
- [3] Annie, L. C. (2012). Market Basket Analysis for a Supermarket based on Frequent Itemset Mining. *IJCSI International Journal of Computer Science Issues*, Vol. 9, Issue 5, No 3 , 257-264.
- [4] Gunaldi, G. (2012). Penerapan Metode Data Mining Market Basket Analysis Terhadap Data Penjualan Produk Buku dengan Menggunakan Algoritma Apriori dan Frequent Pattern Growth (Fp-Growth) :Studi Kasus Percetakan PT Gramedia. *Jurnal Telematika M.Kom Vol.4* , 1.
- [5] Han, J. (2012). *Data Mining: Concepts and techniques*, Morgan Kaufmann. 225 Wyman Street, Waltham, MA 02451, USA: Morgan Kaufmann.
- [6] Kusriani. (2009). *Algoritma Data Mining*. Yogyakarta. Yogyakarta.
- [7] Listriani, D. (2016). Penerapan Metode Asosiasi Menggunakan Algoritma Apriori Pada Aplikasi Analisa Pola Belanja Konsumen. *Jurnal Teknik Informatika Vol.9, No.2* , 10-127.
- [8] Santoso, H. (2016). Data Mining Analisa Pola Pembelian Produk Dengan Menggunakan Metode Algoritma Apriori. *Seminar Nasional Teknologi Informasi dan Multimedia* , 3.7-23.
- [9] Setiawati, D. D. (2009). *Penggunaan Metode Apriori Untuk Analisa Keranjang Pasar Pada Data Transaksi Penjualan Minimarket Menggunakan Java & MySql*. Teknik Informatika, Universitas Gunadarma, Depok.
- [10] Suyanto. (2005). *Pengantar Teknologi Informasi Untuk Bisnis*. Yogyakarta: Andi.
- [11] Syaripudin, G. A. (2017). Implementasi Algoritma Apriori Dalam Menentukan Persediaan Obat. *Jurnal Informatika dan Komputer (JIKO) – Vol. 2, No. 1* , 10-14.
- [12] Tama, B. A. (2010). Penetapan Strategi Penjualan Menggunakan Association Rules dalam Konteks CRM. *Jurnal Generic Vol.5, No.1* , 35-38.
- [13] Tampubolon, K. (2013). Implementasi Data Mining Algoritma Apriori Pada Sistem Persediaan Alat-Alat Kesehatan. *Informasi dan Teknologi Ilmiah (INTI) Volume : 1, Nomor : 1* , 93-106.