Implementasi Data Mining menggunakan Algoritma Aprirori untuk Meningkatkan Pola Penjualan Produk UMKM

Implementation of Data Mining using Aprirori Algorithm to Improve Sales Patterns of MSME Products

[1]Riza Akhsani Setyo Prayoga*, [2]Adzanil Rachmadhi Putra
[1]Bisnis Digital, Fakultas Teknologi Informasi dan Bisnis, Institut Teknologi Telkom Surabaya
[2]Sistem Informasi, Fakultas Teknologi Informasi dan Bisnis, Institut Teknologi Telkom Surabaya
Jl. Ketintang 156 Surabaya Indonesia
*e-mail: *rizayoga30@ittelkom-sby.ac.id*

Abstrak

Perkembangan bisnis semakin cepat dan penuh saingan apalagi pada dunia UMKM. Pengusaha UMKM perlu mengembangkan strategi pemasaran agar produk mereka bisa terjual banyak sehingga bisa memperoleh keuntungan. Salah satu cara untuk membantu dalam strategi pemasaran bisa menggunakan algoritma apriori sehingga bisa membantu dalam pola penjualan dengan menggunakan riwayat data transaksi konsumen. Algoritma apriori ini memiliki beberapa tahapan dalam proses perhitungan mulai dari proses perhitungan support serta penentuan minimum support untuk mengevaluasi nilai data yang berada dibawah minimum support. Lalu perhitungan confident untuk membantu dalam pembentukan aturan asosiasi sehingga ada langkah evaluasi nilai data yanh berada dibawah minimum confident. Kemudian ada lift untuk memberikan informasi secara valid dan menunjukkan proses transaksi. Hasil yang diperoleh dari algoritma apriori ini membentuk suatu aturan asosiasi "Jika membeli stick singkong balado, maka membeli stick jagung balado" maka dari aturan asosiasi ini pengusaha bisa terbantu dalam menjual produk yang cenderung banyak dibeli oleh konsumen. Tahapan terakhir dilakukan pengujian dari perhitungan manual mulai dari support, confident hingga lift dengan perhitungan yang ada di web. Hasil pengujian tersebut diperoleh bahwa perhitungan manual maupun perhitungan yang ada di web memiliki nilai yang sama.

Kata kunci: Algoritma Aprirori, Data Mining, Strategi Pemasaran, UMKM

Abstract

Business development is getting faster and full of rivals, especially in the world of MSMEs. MSME entrepreneurs need to develop marketing strategies so that their products can sell a lot so that they can make a profit. One way to assist in marketing strategies can use a priori algorithms so that they can help in sales patterns by using historical consumer transaction data. After testing the results of this system gives good results. This a priori algorithm has several stages in the calculation process starting from the process of calculating support and determining minimum support to evaluate data values that are below minimum support. Then the confident calculation to assist in the formation of association rules so that there are steps to evaluate the value of data that is below the minimum confident. Then there is an elevator to provide information validly and show the transaction process. The results obtained from this a priori algorithm form an association rule "If you buy a cassava balado stick, then buy a balado corn stick" then from this association rule entrepreneurs can be helped in selling products that tend to be bought by many consumers. The last stage is tested from manual calculations ranging from support, confident to elevator with calculations on the web. The test results obtained that manual calculations and calculations on the web have the same value.

Keywords: Algoritma Aprirori, Data Mining, Marketing Strategies, MSME
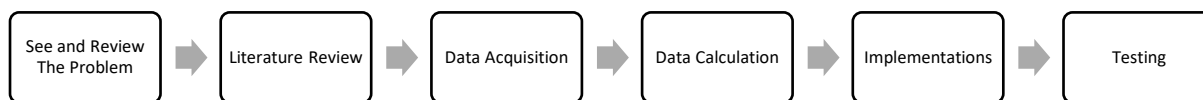
1   Introduction

Competition around busniess, especially in MSMEs, is quite high. MSME sellers need to find ways to create the right marketing strategy so that sales of products from MSMEs can be sold in large quantities [1]. One of the ways to increase sales of MSME products so that they are sold is to use data on trnasactions data and sold products data [2]. The data can be processed to find a pattern of sales of MSME products so that it indirectly helps enterpreneurs to sell their products in large amounts of quantities. The sales data can also be used as information to support decision so that MSME products can be controlled properly [3]. One method that can be applied to determine the sales pattern is by utilizing the Apriori algorithm. The Apriori algorithm is an algorithm used to retireve data by applying association rules to determine the combination of data [4]. In association rules there are calculations related to support, confidence, lift from relationship data. Support is frequently used by to see how often a dataset appears in a transaction. Confident is a indicator that shows the possibility of an item being purchased on the condition that other items are also purchased [6]. Lift is an indicator to find interesting rules besides that list also helps in showing the transaction process and providing valid information [7]. If the lift is equal to 1 (one) then the rule cannot be formed from both [8]. Then if the lift is less than 1 (one), then the data and rules can replace each other [9]. If the lift is more than 1 (one), then the data and rules depend each other [10]. ssociation rules can get interesting predicates if the support value is greater than the minimum support and the confidence value is greater than the minimum confidence value [11]. The a priori algorithm can be used to analyze a pattern so that it can find a tendency, especially a tendency in purchasing an item or consumer behavior [12]. The benefits obtained from this research can help in determining the sales pattern of an MSME product. In addition, it helps in marketing strategies related to products that tend to be purchased by consumers so that MSME entrepreneurs can prepare a sufficient number of products to meet consumer needs.

2   Literatur Review

There are several previous studies related to data mining especially utilizing apriori algorithm, including application of pizza sales data mining using apriori method, which utilizes the apriori algorithm to support pizza sales patterns in order to find pizza products that are in demand enough to be purchased by consumers [13]. There are more research related to the apriori algorithm, namely Analysis of Publication Articles title patterns Based on Google Scholar Using the Apriori Algorithm, which utilizes the apriori algorithm to create publication of article title patterns, which refer to google scholar. In addition, this study produced 23 rules to find pattenrns of word pairs that are often associated with the title of lecturer publications [14]. Then there is another research related to the apriori algorithm, namely the implementation of Data Mining Using Apriori Algorithm to obtain student graduation patterns in this study resulting in 12 rules to form a student graduation patterns [15]. In This study, there is still a relevance with previous research which is related to the steps and calculation methods, but there is a slight difference that this research is carried out to the implementation and testing stages related to the results of manual calculations and calculations conducted on the web. In Addition, this reseearch also focuses on the calculation of lifts where previous research only focused on support and confidence.

3   Research Methodology

In this research process, there are several stages in order to obtain maximum results related to data mining using the apriori algorithm. The research process begins with determining the problem. Researchers see the problems found from the object of research. Then look for supporting literature to help solve this problem thus can help in building a framework of thinking besides that it also helps in contributing to a study. Furthermore, researchers collected data obtained from observations and interviews so that they could help researchers for the next process. The data obtained contains 20 transactions fro 6 types of MSME products. After obtaining the data, calculations are carried out using the apriori algorithm. Then the implementation is carried out and ends with manual testing. The framrwork of thinking as in Figure 1 below

```
┌──────────────┐   ┌──────────────┐   ┌──────────────┐   ┌──────────────┐   ┌──────────────┐   ┌──────────────┐
│ See and Review│ ▶ │  Literature  │ ▶ │    Data      │ ▶ │    Data      │ ▶ │Implementations│ ▶ │   Testing    │
│  The Problem │   │    Review    │   │ Acquisition  │   │ Calculation  │   │              │   │              │
└──────────────┘   └──────────────┘   └──────────────┘   └──────────────┘   └──────────────┘   └──────────────┘
```

Figure 1. The Framework of Thinking

In figure 1 above, explains the framework that starts from looking at existing problems. Then proceed with looking at existing problems. Then proceed with looking for literature to support and provide solutions to existing problems. After that, collect data from 20 transactions from 6 types of MSME products. Then proceed with data calculation with the apriori algorithm and implement In the form of web. After all has been passed, testing is carried out both in terms of calculations both from manual side and web side.

4   Result and Discussion

The first step need to be done is doing a calculation to get pattern frequency. This step is used to search an item combination which qualified the minimun requirements as support marks. Hence, need to utilize the first equations as follows

$$Support\ (A) = \frac{number\ of\ transaction\ contains\ A\ value}{Total\ transactions} \tag{1}$$

Later for the calculation of the support value involving 2 values, you can use the following equation 2 (two) as follows

$$Support\ (A, B) = \frac{number\ of\ transaction\ contains\ A\ and\ B\ value}{Total\ Transactions} \tag{2}$$

Then enter to form a candidate from 6 item divided by 3 snacs and 3 drinks, with 20 transaction sample from such MSME product sales. List of 6 item candidates as presented in Table 1 below

Table 1. MSME Products

Num	Product Code	Product Name
1.	SKB	Stick Kentang Balado
2.	SSB	Stick Singkong Balado
3.	SJB	Stick Jagung Balado
4.	TM	Teh Manis
5.	KS	Kopi Susu
6.	SM	Susu Murni

Table 1 above descirbe product name and product code from 6 candidate. There are 20 transaction sample on product transaction data. This transaction acquired from visitor purchased MSME products. These 20 transactions presented in Table 2 below.

Table 2. Transaction Data

Num	Transaction Code	Transaction Name
1.	TRX1	SSB, SJB, TM
2.	TRX2	SSB, SJB, SM
3.	TRX3	SJB, TM, SM
4.	TRX4	SSB, SJB, KS
5.	TRX5	SSB, KS, SM
6.	TRX6	SKB, SSB, SJB
7.	TRX7	SJB, KS, SM
8.	TRX8	SSB, SJB, SKB
9.	TRX9	SJB, SSB, SM
10.	TRX10	SM, SSB, SJB
11.	TRX11	TM, KS, SM
12.	TRX12	SM, KS, SSB
13.	TRX13	SJB, KS, TM
14.	TRX14	SSB, TM, SM
15.	TRX15	SKB, SSB, SM
16.	TRX16	SJB, SSB, TM
17.	TRX17	KM, SKB, SSB
18.	TRX18	KM, SJB, SKB
19.	TRX19	TM, SKB, SSB
20.	TRX20	SSB, SJB, SKB

After displaying 20 sample of product trasaction data, the procedure is proceeds to with the calculation of support for candidate 1 itemset from each data. The calculations of support refer to the transaction data in Table 2 using equation 1. The calculation of support is presented in Table 3 below.

Table 3. Result of support calculation for 1 itemset

No	Product Name	Calculation Value	Support Marks
1.	Stick Kentang Balado	5	25%
2.	Stick Singkong Balado	15	75%
3.	Stick Jagung Balado	13	65%
4.	Teh Manis	7	35%
5.	Kopi Susu	8	40%
6.	Susu Murni	9	45%

After displaying in Table 3, then proceed with the pruning procees by applying minimum support. The minimum support value is 30% so that there is 1 (one) item that is pruned because it is not included in the minimum support, i.e Stick Singkong Balado. It has a support value of 25% which is below the minimum support value that has been detemined with a value of 30%. The temporary results after pruning are presented in Table 4 below

Table 4. Result of support calculation after pruning

No	Product Name	Calculation Value	Support Value
1.	Stick Singkong Balado	15	75%
2.	Stick Jagung Balado	13	65%
3.	Teh Manis	7	35%
4.	Kopi Susu	8	40%
5.	Susu Murni	9	45%

Proceed with the next step are merge process with calculte how often 2 items show simultaneously thus later generate 2 itemset. The Result of support for 2 itemset is presented in Table 5 below.

Table 5. Result of Support Calculation for 2 itemset

No	Product Name	Calculation Value	Support Marks
1.	SSB, SJB	9	45%
2.	SSB, TM	3	15%
3.	SSB, KS	4	20%
4.	SSB, SM	6	30%
5.	SJB, TM	4	20%
6.	SJB, KS	4	20%
7.	SJB, SM	4	20%
8.	TM, KS	2	10%
9.	TM, SM	3	15%
10.	KS, SM	4	20%

After Table 5 is presented, the pruning process is carried out by applying a minimum support of 30% so that the support value more than equal to 30% is only two, i.e SSB, SJB and SSB, SM. The results after pruning are presented in Table 6 below.

Table 6. Result of Support Calculation Pruning

No	Product Name	Calculation Result	Support Marks
1	SSB, SJB	9	45%
2	SSB, SM	6	30%

Next, perform the process of forming association rules which in this rule formation uses confidence. Still using minimum confidence to help in the pruning process. The pruning results of the confident calculation are presented in Table 7 below.

Table 7. Result of Confident Calculation

No	Asociation Rules	Support Marks	Confident Marks
1	SSB, SJB	45%	60%
2	SSB, SM	30%	40%

The results of Table 7 above can already be known association rules that have exceeded the minimum confidence, namely If you buy SSB, then buy SJB with a confidence value of 60%. Then the last step is to calculate the lift to help show the transaction process and provide valid information. The elevator calculation formula is presented in equation 3 (three) below.

$$Lift\ (A,B) = \frac{support\ A\ dan\ b\ value}{support\ value\ A\ ,support\ value\ B} \qquad (3)$$

From equation 3 (three) above can be applied to obtain the lift value. If the lift is equal to 1 (one) then the rule cannot be formed from both. Then if the lift is less than 1 (one) then the data and rules can replace each other. If the lift is more than 1 (one) then the data and rules are interdependent The results of the elevator calculation are presented in Table 8.

Table 8. Result of lift Calculations

No	Asociation Rules	Support Value	Confident Value	Lift Value
1	SSB, SJB	45%	60%	0,92

Refer to the results of the lift calculation in table 8 above can be concluded that, the lift value is less than 1 (one), so the data and rules can alternate with each other. Association rules can alternate to "if you buy SJB, then buy SSB". Next, enter the implementation stage where the apriori algorithm is applied on the web. The product list display is presented in Figure 2 below.

Figure 2. MSME Product Data View

Next step is the list of transactions on this MSME product with a total of 20 sample transactions which are related to the product data that has been presented in Figure 2. Product transaction data is presented in Figure 3 below.

Figure 3. Product Data Transaction View

The next steps is enter the stage of calculating the support of 1 itemset which applies a minimum support with a value of 30%. There is a pruning process that is carried out because it does not meet the minimum support with a value of 30%. There are previously 6 data which becomes 5 data. The results of the calculation of support 1 itemset are presented in Figure 4 below.

Figure 4. Result of Support Calculation for 1 itemset

The next step is to enter the calculation of the 2 itemset support calculation where this calculation involves 2 combinations of items. Later it still applies a minimum support with a value of 30% so that later there will be some pruning, of course, data that does not meet the minimum support. The results of the 2 itemset support calculation are presented in Figure 5 below.

Figure 5. Result of Hasil perhitungan support 2 itemset

Then the last stage is to form association rules and carry out elevator calculations. In the application of the rules of this association has a minimum confident with a value of 50%. In addition, elevator calculations are used to assist in showing the transaction process and providing information validly. Then the last stage is to form association rules and carry out elevator calculations. In the application of the rules of this association has a minimum confident with a value of 50%. In addition, elevator calculations are used to assist in showing the transaction process and providing information validly. So that the final result obtained is the association rule that is valued "If you buy SSB, then buy SJB" with a confident value of 60% and an elevator value of 0.92. The results of confident and elevator calculations are presented in Figure 6 below.

Hasil perhitungan confident

Dashboard / Hasil perhitungan confident

Berikut adalah tabel hasil perhitungan confident dengan minimum support 50% .

⊞ Berikut adalah tabel hasil perhitungan confident dengan minimum support 50%

10 ∨ entries per page

Search...

No	Aturan Asosiasi	Hasil Support	Nilai Confident	Nilai Lift
1	Jika membeli SSB, maka membeli SJB	45%	60%	0,92

Showing 1 to 1 of 1 entries

Figure 6. Result of confident and lift calculation

From the result shown, it can be concluded that the sales patterns of MSME products focus more on balado potato sticks (SKB) and balado corn sticks (SJB) because they have high support and confidence values than other products. After that there is a test of the results of the apriori algorithm by comparing the results of manual calculations and results from the web. The results of the comparison of manual and web calculations are presented in Table 9 below.

Tabel 9. Result of calculation comparison

No	Test Checklist	Manual Calculation	Result From Web	Status
1	Result of Calculation with 1 itemset	Pruning SKB Product because they are below minimum support value	Cutting SKB products because they are below the minimum confident value	match
2	Result of calculation with 2 itemset	Found 2 itemsets with a value greater than the minimum support, namely SSB, SJB and SSB, SM.	Finds 2 itemsets with values greater than or equal to the minimum support	match

5   Conclusions

Based on the results presented above, it can be concluded that using this a priori algorithm can help entrepreneurs of MSME products to sell their products appropriately, especially in their marketing strategies. Because this a priori algorithm can help entrepreneurs to determine the pattern of buyer tendencies in buying MSME products offered. In the future, this research can be developed by adding the types of products available and can add the number of transactions available. In addition, it can use other algorithms as a comparison and complement so that the results are better, such as using the FP (Frequent Pattern) Growth algorithm.

Reference

[1]   K. Widianto, I. Alfarobi, and A. H. Prameswara, "Metode Algoritma Apriori untuk Mendukung Penjualan 212 Mart Batan Indah Tangerang Selatan," *Information System for Educator and Professional*, vol. 05, no. 01, pp. 1–10, 2020, doi: https://doi.org/10.51211/isbi.v5i1.1297.

[2]   R. Saputra and A. J. P. Sibarani, "Implementasi Data Mining menggunakan Algoritma Apriori untuk Meningkatkan Pola Penjualan Obat," *Jurnal Teknik Informatika dan Sistem Informasi*, vol. 7, no. 2, pp. 262–272, 2020, doi: https://doi.org/10.35957/jatisi.v7i2.195.

[3]   Sutrisno, "Penerapan Algoritma Apriori untuk Mencari Pola Penjualan Produk Dana pada PT Bank Rakyat Indonesia (Persero) Tbk Kanca Jakarta Pasar Minggu," *SIMIKA*, vol. 3, no. 1, pp. 12–26, 2020, doi: https://doi.org/10.47080/simika.v3i1.834.

[4]   N. Hadinata and Kurniawan, "Analisis Pola Pembelian Produk Makanan Ringan menggunakan Algoritma Apriori," *Jurnal SISFOKOM*, vol. 09, no. 01, pp. 1–7, 2020, doi: 10.32736/sisfokom.V9.I1.623.

[5]   F. Mansyur and F. Ridho, "Analisis Pola Permintaan Publikasi Data Badan Pusat Statistik menggunakan Association Rule Apriori," *Kumpulan jurnaL Ilmu Komputer (KLIK)*, vol. 07, no. 02, pp. 187–198, 2020, doi: http://dx.doi.org/10.20527/klik.v7i2.322.

[6]   A. Firmansyah and N. Merlina, "Prediksi Pola Penjualan Tiket Kapal PT. Pelni Cabang Makassar menggunakan Metode Algoritma Apriori," *Jurnal Ilmu Pengetahuan dan Teknologi Komputer*, vol. 5, no. 2, pp. 183–189, 2020, doi: https://doi.org/10.33480/jitk.v5i2.1123.

[7]   F. Adhinda, K. Wardani, and T. Kristiana, "Implementasi Data Mining Penjualan Produk Kosmetik pada PT. Natural Nusantara menggunakan Algoritma Apriori," *Paradigma - Jurnal Informatika dan Komputer*, vol. 22, no. 1, pp. 85–90, 2020, doi: 10.31294/p.v21i2.

[8]   M. T. Anwar, Hi. D. Purnomo, M. Novita, and C. H. Primasari, "Implementasi Metode Asosiasi Apriori untuk Mengetahui Pola Beli Konsumen dan Rekomendasi Penempatan Produk pada Swalayan XYZ," *Dinamik*, vol. 25, no. 01, pp. 29–38, 2020, doi: https://doi.org/10.35315/dinamik.v25i1.7747.

[9]   A. Riyanto and M. Susanti, "Analisa Data Mining terhadap Penjualan Food dengan Metode Apriori pada Kopsyahira," *Bianglala Informatika*, vol. 8, no. 1, pp. 44–48, 2020, doi: https://doi.org/10.31294/bi.v8i1.8148.g4135.

[10]   R. A. Saputra, S. Wasiyanti, and R. Nugraha, "Penerapan Algoritma Apriori untuk Analisa Pola Penempatan Barang berdasarkan Data Transaksi Penjualan," *Jurnal SWABUMI*, vol. 08, no. 02, pp. 160–170, 2020, doi: http://dx.doi.org/10.31294/swabumi.v8i2.9031.

[11]   A. B. Kusdinar, D. Riyadi, and A. Asriyanik, "Implementasi Algoritma Apriori pada Penyusunan Menu Makanan Rumah Makan Prasmanan," *Jurnal Teknik Informatika dan Sistem Informasi*, vol. 6, no. 2, pp. 391–399, Aug. 2020, doi: 10.28932/jutisi.v6i2.2742.

[12]   A. Setiawan and R. Mulyanti, "Market Basket Analysis dengan Algoritma Apriori pada Ecommerce Toko Busana Muslim Trendy (Market Basket Analysis with Apriori Algorithms in Ecommerce Trendy Muslim Clothing Stores)," *JUITA*, vol. 8, no. 1, pp. 11–18, 2020, doi: http://dx.doi.org/10.30595/juita.v8i1.4550.

[13]   Rusdiansyah, N. Suharyanti, Triningsih, and Murniyati, "Application Of Pizza Sales Data Mining using Apriori Method," *SinkrOn*, vol. 4, no. 2, pp. 1–5, Mar. 2020, doi: 10.33395/sinkron.v4i2.10500.

[14]   I. Dwi Ulumiyah and H. Yuliansyah, "Analisis Pola Asosiasi Judul Artikel Publikasi berdasarkan Data Google Scholar menggunakan Algoritma Apriori," *Jurnal Sarjana Teknik Informatika*, vol. 10, no. 3, pp. 140–148, 2022, doi: 10.12928/jstie.v8i3.xxx.

[15]   A. Anas and A. J. Zebua, "Implementasi Algoritma Apriori untuk Mendapatkan Pola Kelulusan Mahasiswa," *Jurnal Ilmiah Media Sisfo*, vol. 16, no. 1, pp. 54–61, Apr. 2022, doi: 10.33998/mediasisfo.2022.16.1.1173.