

Analisis Pola Pembelian dan Penjualan Bisnis Menggunakan Algoritma Apriori dalam Studi Market Basket

Ketut Widya Kayohana¹, M Danang Samudra¹, Ni Komang Dewani², NI Luh Putu Febiyanti³, Eka Futjiastuti⁴

ketut.widya@universitasbumigora.ac.id¹, 2001010161@universitasbumigora.ac.id², 2001010163@universitasbumigora.ac.id³,
2001010170@universitasbumigora.ac.id⁴, 2001010172@universitasbumigora.ac.id⁵

^{1,2}Universitas Bumigora, Mataram, Indonesia

Article Info

Article history:

Received 20.07.2023

Revised 21.07.2023

Accepted 23.07.2023

ABSTRACT

Using the Apriori Algorithm to help optimize product placement, increase cross-selling opportunities, and improve inventory management and promotions. In-depth insights into purchase patterns can enhance sales, customer satisfaction, and operational efficiency. Enhancing cross-selling strategies: By understanding associations between items purchased together, this research provides insights for improving cross-selling strategies. Businesses can offer relevant product recommendations, whether through in-store placement or online suggestions, to boost sales and enhance the shopping experience. The Apriori method and algorithm are used in analyzing the buying and selling of market baskets.

The research findings revealed a significant association between customers who purchase Bread and Coffee together. Support "toast" and "coffee" is 0.023666, which means that 2.3666% of the total 30 transactions includes both itemsets, confidence "toast → coffee" is 0.704403, which means that if customers buy "toast", chances are they also buy "coffee" by 70.4403%, lift "toast → coffee" is 1.472431, which shows that the probability of buying "coffee" increases by 1.472431 times if the customer also buys "toast". This presents an opportunity for businesses to improve sales through strategic product placement, promotions or special offers involving the combination of Bread and Coffee, and optimizing inventory to meet the demand of customers who tend to purchase both items simultaneously.

Keywords:

Apriori,
Market Basket
analysis,
Businesses



Corresponding Author:

Lalu Yayan Ardiansyah
Faculty of Economics and Business, Digital Business, Universitas Bumigora
Jalan Ismail Marzuki No. 22, Mataram, NTB, Indonesia
Email: income@universitasbumigora.ac.id

1. INTRODUCTION

Dalam era digital dan persaingan bisnis yang semakin ketat, pemahaman yang mendalam tentang perilaku konsumen menjadi kunci keberhasilan bagi perusahaan. Salah satu aspek penting dalam memahami perilaku konsumen adalah analisis pola pembelian (Saragih, 2021). Dengan menganalisis pola pembelian, perusahaan dapat mengidentifikasi hubungan antara produk atau layanan yang dibeli oleh konsumen, sehingga dapat mengambil keputusan strategis yang lebih baik dalam hal pemasaran, penentuan harga, dan pengelolaan persediaan (Yuli, 2012).

Salah satu algoritma yang banyak digunakan dalam analisis pola pembelian adalah algoritma Apriori. Algoritma Apriori merupakan algoritma yang digunakan untuk menemukan asosiasi atau keterkaitan antara item atau produk dalam data transaksi (Riszky & Sadikin, 2019). Dengan menggunakan algoritma Apriori, perusahaan dapat mengidentifikasi pola pembelian yang signifikan, seperti kombinasi produk yang sering dibeli bersamaan oleh konsumen (Jha & Ragha, 2013).

Penelitian ini bertujuan untuk menerapkan algoritma Apriori dalam studi market basket, yaitu analisis pola pembelian berdasarkan data transaksi dari pelanggan. Dengan menggali wawasan dari data transaksi, penelitian ini diharapkan dapat membantu perusahaan dalam mengidentifikasi kecenderungan dan pola pembelian yang dapat meningkatkan efisiensi operasional, kepuasan pelanggan, dan pendapatan. Berdasarkan paparan di atas diharapkan dapat memberikan kontribusi dalam pemahaman tentang perilaku konsumen, serta memberikan wawasan yang berharga bagi perusahaan dalam mengoptimalkan strategi pemasaran dan pengelolaan persediaan mereka.

2. METHOD

Pengumpulan Data: Data transaksi yang digunakan dalam penelitian ini diambil dari kaggle market basket yang berisi data penjualan produk setiap harinya dan produk apa saja yang sering di beli secara bersamaan. Data ini mencakup daftar produk yang dibeli oleh pelanggan dalam setiap transaksi.

Preprocessing Data: Data yang dikumpulkan perlu melalui tahap preprocessing untuk membersihkan dan mengorganisasi data (Edastama et al., 2021). Langkah-langkah *preprocessing* yang akan dilakukan meliputi: Menghapus data yang tidak relevan atau duplikat, Melakukan konversi data ke dalam format yang sesuai untuk analisis dengan algoritma Apriori. Menyesuaikan format data transaksi menjadi bentuk "itemset" atau daftar produk yang dibeli dalam satu transaksi.

Menganalisis Pola Pembelian menggunakan Algoritma Apriori: Algoritma Apriori akan diimplementasikan dalam bahasa pemrograman atau menggunakan alat analisis data yang mendukung eksekusi algoritma ini, Tahap pertama dalam analisis adalah mengidentifikasi item-item yang sering muncul bersamaan dalam transaksi (*itemset*). Algoritma Apriori menggunakan konsep "support" untuk mengidentifikasi itemset yang sering muncul dalam dataset, Setelah mendapatkan itemset yang sering muncul, langkah selanjutnya adalah mengidentifikasi asosiasi antara item-item tersebut. Algoritma Apriori menggunakan konsep "confidence" untuk mengukur seberapa sering item B dibeli bersamaan dengan item A dibandingkan dengan item B dibeli secara mandiri, Hasil dari analisis Apriori akan berupa daftar pola pembelian yang signifikan, seperti kombinasi produk yang sering dibeli bersamaan atau urutan pembelian yang umum terjadi. Cara kerja sistem dalam penelitian ini dapat ditunjukkan pada gambar 1.

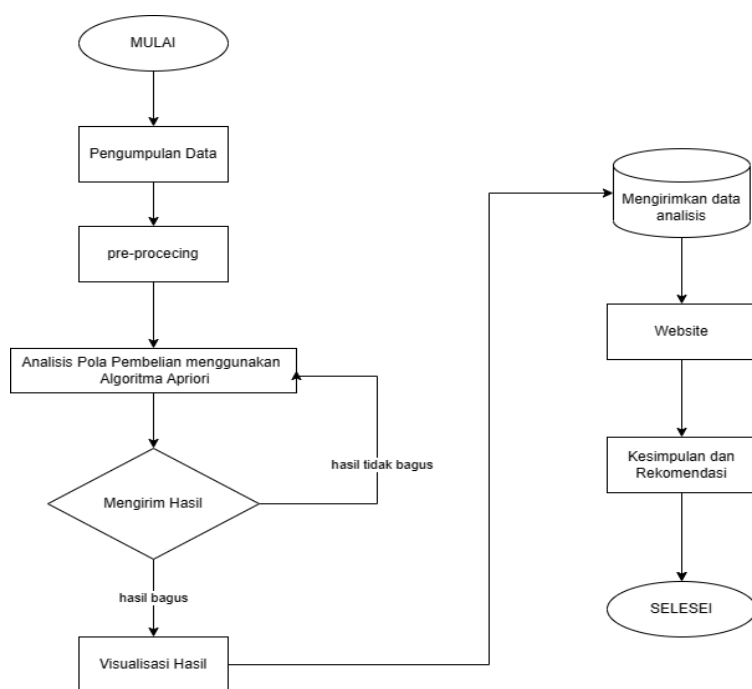


Figure 1. flowchart sistem

Pada Gambar 1 dapat dijelaskan bahwa proses *Market basket analysis* yang terjadi diawali dari pengumpulan data pada dataset setelah itu dilakukan *pre-procesing* analisis pada pola pembelian menggunakan algoritma apriori lalu mengirim hasil, jika hasil tidak bagus maka akan diulang hingga mengirim hasil yang sudah divisualisasi hasilnya kemudian mengirimkan data analisis dikirimkan ke website dengan menghasilkan kesimpulan dan rekomendasinya.

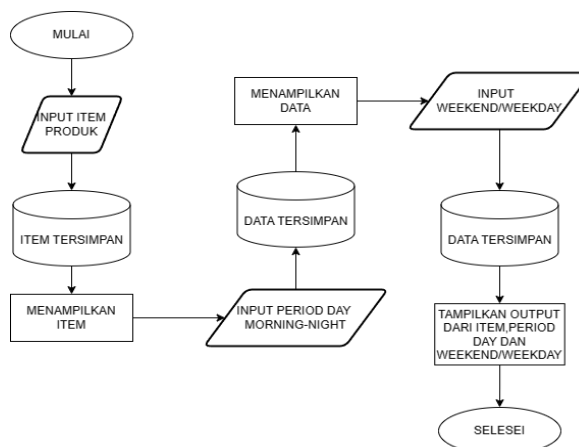


Figure 2. Flowchart Sistem website.

Pada gambar 2 flowchart dapat dijelaskan bahwa proses *Market basket analysis* yang terjadi diawali dari input item produk lalu item tersimpan yang akan menampilkan item diinputkan period *day morning-night* data tersimpan akan menampilkan data menginputkan *weekend/weekday* kemudian data tersimpan menampilkan/tampilkan output dari *item period day* dan *weekend/weekday*

3. RESULTS AND DISCUSSION

Algoritma Apriori

Tabel 1. Data penjualan market basket terlaris dari 2016-2017 yang diambil dari dataset

No	Produk	Jumlah
1	Coffe	5471
2	Bread	3325
3	Tea	1435
4	Cake	1025
5	Pastry	856
6	Sandwich	771
7	Medialuna	616
8	Hot chocolate	590
9	Cookies	540
10	Brownie	379
11	Farm house	374
12	Muffin	370
13	Alfajores	369
14	Juice	369
15	Sop	342
16	Scone	327
17	Toast	318
18	Scandinavian	277
19	Truffles	193
20	Coke	185

Transaction	item	adjustment	afternoon with the baker	alfajores	argentina night	art tray	bacon	baguette	bakewell	bare popcorn	basket	...	the bart	the nomad	tiffin	toast	truffles	tshirt	valenti
1	0	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0
11	0	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0
12	0	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0
13	0	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0
14	0	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0
15	0	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0
16	0	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0
17	0	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0
18	0	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0
19	0	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0
20	0	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0

Figure 3. Analisa data dengan Algoritma Apriori

Tabel 2. K = 1

NO	ITEMSET	JUMLAH	SUPPORT
1.	Tea	26	14.2%
2.	Coffe	6	47.8%
3.	Cake	4	10.3%

Tabel 2 merupakan persentasi jumlah item terpilih dan yang memenuhi syarat minimum *support* = 0.1 % dari data transaksi.

1. *Itemset: 'Tea'*

- Jumlah: 26
- *Support*: 14.2%

Itemset ini hanya terdiri dari satu item, yaitu *'Tea'*. Dalam dataset yang Anda analisis, itemset ini muncul dalam 26 transaksi. Supportnya dinyatakan sebagai persentase dari total transaksi, yaitu 14.2%.

2. *Itemset: 'Coffee'*

- Jumlah: 6
- *Support*: 47.8%

Itemset ini juga terdiri dari satu item, yaitu *'Coffee'*. Dalam dataset Anda, itemset ini muncul dalam 6 transaksi. Supportnya adalah 47.8%, yang berarti 47.8% dari total transaksi mengandung item *'Coffee'*.

3. *Itemset: 'Cake'*

- Jumlah: 4
- *Support*: 10.3%

Itemset ini juga terdiri dari satu item, yaitu *'Cake'*. Dalam dataset Anda, itemset ini muncul dalam 4 transaksi. Supportnya juga adalah 10.3%.

Tabel 3. K = 2

NO	ITEMSET	JUMLAH	SUPPORT
1.	<i>Coffe, cake</i>	42	5.4%
2.	<i>Coffee, tea</i>	55	4.9%
3.	<i>Coffe, pastry</i>	50	4.7%

Tabel 3 merupakan proses pembentukan 2 itemset dihitung minimum support = 0.1 %. Persentasi jumlah 2 itemset terpilih hanya kombinasi 2 itemset yang memenuhi syarat minimum support yang ada dalam data penjualan

Itemset: '*Coffee, cake*'

- Jumlah: 42
- Support: 5.4%

Itemset ini terdiri dari kombinasi dua item, yaitu '*Coffee*' dan '*cake*'. Dalam dataset yang dianalisis, itemset ini muncul dalam 42 transaksi. Supportnya dinyatakan sebagai persentase dari total transaksi, yaitu 5.4%.

2. Itemset: '*Coffee, tea*'

- Jumlah: 55
- Support: 4.9%

Itemset ini terdiri dari kombinasi dua item, yaitu '*Coffee*' dan '*tea*'. Dalam dataset, itemset ini muncul dalam 55 transaksi. Supportnya adalah 4.9%, yang berarti 4.9% dari total transaksi mengandung kombinasi itemset tersebut.

3. Itemset: '*Coffee, pastry*'

- Jumlah: 50
- Support: 4.7%

Itemset ini terdiri dari kombinasi dua item, yaitu '*Coffee*' dan '*pastry*'. Dalam dataset, itemset ini muncul dalam 50 transaksi. Supportnya adalah 4.7%.

Tabel 4. K = 3

NO	ITEMSET	JUMLAH	SUPPORT
1.	<i>coffee</i> , ' <i>pastry</i> ', ' <i>bread</i> '	59	1.11%
2.	<i>coffee</i> , ' <i>bread</i> ', ' <i>cake</i> '	58	1.3%
3.	<i>coffee</i> , ' <i>tea</i> ', ' <i>cake</i> '	60	1.3%

Tabel 4 merupakan proses pembentukan 3 itemset dihitung minimum support = 0.1 %. Persentasi jumlah 3 itemset terpilih hanya kombinasi 3 itemset yang memenuhi syarat minimum support yang ada dalam data penjualan

1. Itemset: '*coffee*', '*pastry*', '*bread*'

- Jumlah: 59
- Support: 1.11%

Itemset ini terdiri dari kombinasi tiga item, yaitu '*coffee*', '*pastry*', dan '*bread*'. Dalam dataset yang dianalisis, itemset ini muncul dalam 59 transaksi. Supportnya dinyatakan sebagai persentase dari total transaksi, yaitu 1.11%.

2. Itemset: '*coffee*', '*bread*', '*cake*'

- Jumlah: 58
- Support: 1.3%

Itemset ini terdiri dari kombinasi tiga item, yaitu 'coffee', 'bread', dan 'cake'. Dalam dataset, itemset ini muncul dalam 58 transaksi. Supportnya adalah 1.3%, yang berarti 1.3% dari total transaksi mengandung kombinasi itemset tersebut.

3. Itemset: 'coffee', 'tea', 'cake'

- Jumlah: 60
- Support: 1.3%

Itemset ini terdiri dari kombinasi tiga item, yaitu 'coffee', 'tea', dan 'cake'. Dalam dataset ini, itemset ini muncul dalam 60 transaksi. Supportnya juga adalah 1.3%.

Aturan asosiasi final

Tabel 5.

NO	ATURAN	SUPPORT	CONFIDENCE	LIFT
1.	Jika konsumen membeli <i>toast</i> , maka	2.3%	70.4%	1472
2.	cenderung akan membeli <i>coffee</i>	1.0%	59.8%	1251
3.	Jika konsumen membeli <i>Spanish Brunch</i> , maka cenderung membeli <i>Coffee</i> secara bersamaan	3.5%	56.9%	1189
	Jika konsumen membeli <i>Medialuna</i> , maka membeli <i>Coffee</i> secara bersamaan			

1. *Support*: *Support* menggambarkan sejauh mana itemset muncul dalam transaksi dataset. Rumus: $Support(X) = (\text{Jumlah transaksi yang mencakup } X) / (\text{Jumlah total transaksi})$

Dalam kasus ini, *support "toast"* dan "*coffee*" adalah 0.023666, yang berarti 2.3666% dari total 30 transaksi mencakup kedua itemset tersebut.

2. *Confidence*: *Confidence* mengukur sejauh mana itemset Y muncul dalam transaksi yang mencakup itemset X. Rumus: $Confidence(X \rightarrow Y) = (\text{Jumlah transaksi yang mencakup } X \text{ dan } Y) / (\text{Jumlah transaksi yang mencakup } X)$

Dalam kasus ini, *confidence "toast" → "coffee"* adalah 0.704403, yang berarti jika pelanggan membeli "toast", kemungkinan mereka juga membeli "*coffee*" sebesar 70.4403%.

3. *Lift*: *Lift* adalah pengukuran sejauh mana itemset X dan Y terkait secara asosiatif, dengan memperhitungkan seberapa sering keduanya muncul bersamaan dibandingkan dengan muncul secara independen. Rumus: $Lift(X \rightarrow Y) = (Support(X \rightarrow Y)) / (Support(X) \times Support(Y))$

Dalam kasus ini, *lift "toast" → "coffee"* adalah 1.472431, yang menunjukkan bahwa kemungkinan membeli "*coffee*" meningkat sebesar 1.472431 kali jika pelanggan juga membeli "*toast*".

Data import dan hasil asosiasi final

	antecedents	consequents	support	confidence	lift
30	(toast)	(coffee)	0.023666	0.704403	1.472431
28	(spanish brunch)	(coffee)	0.010882	0.598837	1.251766
18	(medialuna)	(coffee)	0.035182	0.569231	1.189878
22	(pastry)	(coffee)	0.047544	0.552147	1.154168
1	(alfajores)	(coffee)	0.019651	0.540698	1.130235
16	(juice)	(coffee)	0.020602	0.534247	1.116750
24	(sandwich)	(coffee)	0.038246	0.532353	1.112792
6	(cake)	(coffee)	0.054728	0.526958	1.101515
26	(scone)	(coffee)	0.018067	0.522936	1.093107
13	(cookies)	(coffee)	0.028209	0.518447	1.083723
14	(hot chocolate)	(coffee)	0.029583	0.507246	1.060311
4	(brownie)	(coffee)	0.019651	0.490765	1.025860
21	(muffin)	(coffee)	0.018806	0.489011	1.022193
3	(pastry)	(bread)	0.029160	0.338650	1.034977
10	(cake)	(tea)	0.023772	0.228891	1.604781
39	(tea, coffee)	(cake)	0.010037	0.201271	1.937977
32	(sandwich)	(tea)	0.014369	0.200000	1.402222
9	(hot chocolate)	(cake)	0.011410	0.195652	1.883874
38	(cake, coffee)	(tea)	0.010037	0.183398	1.285822
11	(tea)	(cake)	0.023772	0.166667	1.604781

Figure 4. data yang telah diimport

Pada gambar 4 merupakan data yang telah diimport kemudian diproses dan akan menampilkan hasil asosiasi final.

Tampilan website

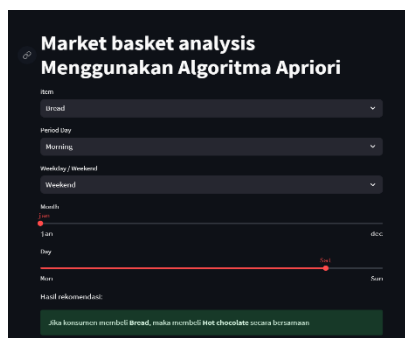


Figure 5. Tampilan website dari program yang sudah dibuat menggunakan *opensource* (streamlit), menampilkan hasil rekomendasi.

Menganalisis Hasil data penjualan market basket

algoritma Apriori mengurangi jumlah itemset yang diproses berdasarkan prinsip apriori, yaitu jika suatu itemset jarang muncul, maka itemset yang lebih panjang yang mengandung item tersebut juga jarang muncul. Dengan demikian, algoritma Apriori secara efisien memproses dan menyaring itemset berdasarkan frekuensi kemunculannya dalam dataset untuk menghasilkan pola asosiasi yang signifikan.

3. CONCLUSION

Penelitian ini memberikan informasi berharga bagi bisnis untuk memahami pola pembelian konsumen pada tingkat bulanan, harian, dan jam tertentu, termasuk hari libur. Informasi ini dapat

digunakan untuk mengoptimalkan pengelolaan stok secara efektif dan menentukan peningkatan stok pada barang yang diperlukan. Algoritma apriori mempunyai pengaruh positif dalam laporan penjualan market basket karena pemilik dapat mengetahui produk yang perlu diperbanyak atau dikurangi stoknya guna meningkatkan omset penjualan. *Support "toast" dan "coffee"* adalah 0.023666, yang berarti 2.3666% dari total 30 transaksi mencakup kedua itemset tersebut, *confidence "toast → coffee"* adalah 0.704403, yang berarti jika pelanggan membeli "toast", kemungkinan mereka juga membeli "coffee" sebesar 70.4403%, *lift "toast → coffee"* adalah 1.472431, yang menunjukkan bahwa kemungkinan membeli "coffee" meningkat sebesar 1.472431 kali jika pelanggan juga membeli "toast"

SUGGESTION

Setelah dilakukan penelitian dan penarikan kesimpulan, peneliti memiliki beberapa. Dimana didalam menganalisis setiap item set harus memperhatikan minat dari pembeli. Selain itu, di dalam proses penjualan dan pemilihan item tidak hanya cenderung pada pembelian yang disukai, tetapi memberikan variasi guna untuk meningkatkan jumlah penjualan setiap item pada market basket

REFERENCES

- [1] Edastama, P., Bist, A. S., & Prambudi, A. (2021). Implementation Of Data Mining On Glasses Sales Using The Apriori Algorithm. *International Journal of Cyber and IT Service Management*, 1(2), 159–172. <https://doi.org/10.34306/ijcitsm.v1i2.46>
- [2] Jha, J., & Raha, L. (2013). Educational Data Mining using Improved Apriori Algorithm. *International Journal of Information and Computation Technology*, 3(5), 411–418. <http://www.>
- [3] Riszky, A. R., & Sadikin, M. (2019). Data Mining Menggunakan Algoritma Apriori untuk Rekomendasi Produk bagi Pelanggan. *Jurnal Teknologi Dan Sistem Komputer*, 7(3), 103–108. <https://doi.org/10.14710/jtsiskom.7.3.2019.103-108>
- [4] Saragih, S. D. (2021). Analisa Pola Penjualan Alat Pancing Menggunakan Algoritma Apriori. *Jurnal Mahasiswa Aplikasi Teknologi Komputer Dan Informasi*, 3(3), 78–83.
- [5] Yuli. (2012). *Institut Teknologi Telkom Surabaya Fakultas Teknologi Informasi Dan Industri Program Studi S1 Teknologi Informasi Rencana Pembelajaran Semester*. 1–9.

