

Analisis Perilaku Pembelian Dan Prediksi Konsumen Pada Usaha Retail Dengan Metode Data Mining

Iskandar¹, Adi Muhamad Muhsidi^{2*}, Qona'ah El Hasan³,
Dadan Darmawan Muttaqien⁴, Aah Sumiah⁵

¹Sekolah Pascasarjana, Magister Manajemen, Universitas Kuningan, Kuningan, Indonesia

^{2,3,4}Fakultas Ekonomi dan Bisnis, Bisnis Digital, Universitas Kuningan, Kuningan, Indonesia

⁵Fakultas Ilmu Komputer, Manajemen Informatika, Universitas Kuningan, Kuningan, Indonesia

E-mail: ¹iskandar@uniku.ac.id, ^{2*}adi.muhamad@uniku.ac.id, ³qonaah.elz@uniku.ac.id,

⁴dadan.darmawan@uniku.ac.id, ⁵aah.sumiah@uniku.ac.id

^{*}) Email Penulis Utama

Abstrak—Usaha retail khususnya skala kecil dan menengah, menghadapi tantangan dalam memahami perilaku pembelian konsumen secara mendalam dari data transaksi yang melimpah. Ketidakmampuan untuk menganalisis pola ini secara efektif sering kali menyebabkan inefisiensi dalam manajemen inventori, penataan produk di dalam toko, serta perancangan strategi pemasaran yang kurang terarah. Penelitian ini bertujuan untuk menyajikan solusi atas permasalahan tersebut dengan menganalisis dan memprediksi perilaku pembelian konsumen di Azzam Mart melalui penerapan metode data mining. Pendekatan yang digunakan adalah *Association Rule Mining* dengan mengimplementasikan algoritma Apriori untuk mengungkap hubungan antar produk. Data transaksi penjualan, yang terdiri dari 300 transaksi selama periode Oktober 2024, diolah menggunakan bahasa pemrograman Python beserta pustaka *Pandas* untuk manipulasi data dan *Seaborn* untuk visualisasi. Algoritma Apriori bekerja dengan mengidentifikasi *frequent itemsets*, yaitu kombinasi produk yang sering dibeli bersamaan, berdasarkan tiga metrik utama: *Support*, yang mengukur frekuensi kemunculan sebuah item atau set item dalam seluruh transaksi; *Confidence*, yang menunjukkan probabilitas pembelian produk konsekuen ketika produk anteseden telah dibeli; dan *Lift*, yang mengukur seberapa besar peningkatan probabilitas pembelian bersama dibandingkan jika kedua produk dibeli secara independen. Hasil analisis berhasil mengidentifikasi 20 *frequent itemsets* teratas, dengan "Telur Ayam", "Minuman Soda" dan "Susu UHT" sebagai produk yang paling sering muncul dalam transaksi konsumen. Lebih lanjut, penelitian ini mengungkap aturan asosiasi yang kuat dan signifikan secara bisnis. Salah satu temuan utamanya adalah hubungan erat antara pembelian "Roti Tawar" (sebagai anteseden) dan "Minyak Goreng 1L" (sebagai konsekuen), yang menunjukkan nilai *confidence* sebesar 41,7% dan nilai *lift* 2,69. Nilai *lift* yang lebih besar dari 1 ini mengindikasikan bahwa kedua produk tersebut 2,69 kali lebih mungkin untuk dibeli bersama-sama daripada secara kebetulan, menandakan asosiasi yang positif dan kuat. Temuan ini memberikan wawasan strategis yang dapat ditindaklanjuti bagi manajemen ritel untuk mengoptimalkan penempatan produk, merancang program promosi yang efektif seperti *bundling* dan meningkatkan manajemen stok untuk kepuasan pelanggan dan peningkatan penjualan.

Kata Kunci: Data Mining, Algoritma Apriori, Perilaku Konsumen, Analisis Keranjang Belanja, Strategi Ritel.

Abstract—Retail businesses particularly small and medium-sized enterprises, face the challenge of deeply understanding consumer purchasing behavior from their vast amounts of transaction data. The inability to effectively analyze these patterns often leads to inefficiencies in inventory management, in-store product layout, and the design of unfocused marketing strategies. This research aims to provide a solution to this problem by analyzing and predicting consumer purchasing behavior at Azzam Mart through the application of data mining methods. The approach utilizes *Association Rule Mining* by implementing the Apriori algorithm to uncover relationships between products. Sales transaction data, comprising 300 transactions from the period of October 2024, were processed using the Python programming language along with the *Pandas* library for data manipulation and *Seaborn* for visualization. The Apriori algorithm functions by identifying frequent itemsets, which are combinations of products often purchased together, based on three primary metrics: *Support*, which measures the frequency of an item or itemset's appearance across all transactions; *Confidence*, which indicates the probability of purchasing a consequent product when an antecedent product has been bought; and *Lift*, which measures the extent to which the probability of purchasing them together increases compared to if they were purchased independently. The analysis successfully identified the top 20 frequent itemsets, with 'Chicken Eggs', 'Soft Drinks', and 'UHT Milk' emerging as the products most frequently appearing in consumer transactions. Furthermore, this study revealed strong and commercially significant association rules. One of the main findings is the strong relationship between the purchase of 'Bread' (as the antecedent) and '1L Cooking Oil' (as the consequent), which demonstrated a confidence value of 41.7% and a lift value of 2.69. This lift value, being greater than 1, indicates that these two products are 2.69 times more likely to be purchased together than by random chance, signifying a strong positive association. These findings provide actionable strategic insights for retail management to optimize product placement, design effective promotional programs such as *bundling*, and improve stock management, ultimately leading to enhanced customer satisfaction and increased sales.

Keywords: Data Mining, Apriori Algorithm, Consumer Behavior, Market Basket Analysis, Retail Strategy.

1. PENDAHULUAN

Dalam dunia retail, memahami pola pembelian konsumen merupakan faktor penting untuk meningkatkan efektivitas penjualan dan kepuasan pelanggan. Perusahaan retail dihadapkan pada tantangan untuk mengelola inventori secara efisien, mengoptimalkan tata letak produk, serta merancang strategi pemasaran yang lebih terfokus pada kebutuhan dan preferensi konsumen[1]. Dengan kemajuan teknologi informasi, data transaksi penjualan yang dikumpulkan sehari-hari memberikan peluang besar untuk memahami perilaku konsumen melalui analisis data[2]. Pendekatan *data mining*, khususnya melalui *Association Rule Mining* menggunakan algoritma Apriori, telah banyak digunakan untuk mengidentifikasi pola hubungan antara produk yang sering dibeli bersama, yang disebut dengan *frequent itemsets*. [3]

Pertumbuhan data transaksi yang eksponensial membuka peluang sekaligus tantangan bagi pelaku usaha retail untuk dapat menganalisis dan memprediksi perilaku pembelian konsumen secara lebih akurat[4]. Penelitian ini dilatarbelakangi oleh fenomena tersebut yang terjadi di Azzam Mart, sebuah toko ritel yang berlokasi di Desa Mekarjaya, Kecamatan Ciawigebang, Kabupaten Kuningan. Toko ini telah beroperasi selama enam tahun dan telah menjadi salah satu penyedia kebutuhan sehari-hari yang terpercaya bagi masyarakat sekitar. Meskipun memiliki data transaksi yang melimpah, Azzam Mart menghadapi tantangan dalam menganalisis data ini untuk mengembangkan strategi bisnis yang lebih efektif. Oleh karena itu, solusi yang ditawarkan dalam penelitian ini ditujukan kepada manajemen Azzam Mart untuk membantu mereka dalam pengambilan keputusan berbasis data. Data mining sebagai salah satu cabang ilmu yang fokus pada ekstraksi informasi berharga dari kumpulan data besar, menawarkan solusi yang menjanjikan untuk menganalisis pola pembelian konsumen.

Metode *Association Rule Mining* berfungsi untuk menemukan aturan-aturan asosiasi atau hubungan antar produk dalam dataset transaksi, sedangkan algoritma Apriori adalah algoritma yang efisien untuk menemukan kombinasi produk yang paling sering muncul secara bersamaan (*frequent itemsets*). Studi-studi sebelumnya telah menunjukkan efektivitas metode ini. Sebagai contoh, penelitian oleh Elisa (2018) menerapkan algoritma Apriori pada minimarket untuk menganalisis kecenderungan pembelian konsumen. Demikian pula, penelitian oleh Mulya, et al. (2019) mengimplementasikan algoritma Apriori untuk meningkatkan penjualan kantin universitas[5].

Data mining memungkinkan ekstraksi informasi berharga dari data besar untuk mendukung strategi pemasaran, perencanaan stok, hingga personalisasi layanan pelanggan. Namun, banyak usaha retail skala kecil hingga menengah yang belum sepenuhnya mengadopsi metode ini, sehingga mereka kehilangan peluang untuk bersaing secara efektif di pasar [6] [7]. Penelitian ini mengadopsi algoritma apriori sebagai metode analisis utama, dengan pertimbangan khusus pada kualitas hasil yang dihasilkan. Studi komparatif yang dilakukan Nofriansyah D. dan Nurcahyo G. menunjukkan bahwa walaupun algoritma ini membutuhkan durasi pemrosesan yang lebih ekstensif dibandingkan FP-Growth, apriori mampu menghasilkan rules asosiasi yang lebih kuat dan dapat diandalkan[8]. Tantangan utama dalam implementasi algoritma Apriori terletak pada aspek efisiensi komputasi saat menganalisis dataset skala besar. Untuk meningkatkan performa algoritma ini, penggunaan Python dan library Pandas menjadi penting karena menyediakan framework yang powerful untuk optimasi proses data mining[9]. Kedua tools ini tidak hanya mempercepat proses analisis, tetapi juga memungkinkan penanganan dataset yang lebih besar dengan konsumsi sumber daya yang lebih efisien.

Dalam konteks ini, beberapa masalah yang dapat diidentifikasi meliputi kompleksitas dalam menganalisis data transaksi retail yang besar dan beragam, kesulitan dalam menentukan pola pembelian konsumen secara akurat, tantangan dalam memprediksi perilaku konsumen untuk pengambilan keputusan bisnis, serta kurangnya pemanfaatan teknologi data mining dalam analisis retail[10]. Permasalahan-permasalahan tersebut mengarah pada pertanyaan penelitian tentang bagaimana mengimplementasikan metode data mining untuk menganalisis perilaku pembelian konsumen pada usaha retail, bagaimana memanfaatkan Python, pandas dan seaborn untuk mengolah dan memvisualisasikan data transaksi retail, serta seberapa akurat prediksi perilaku konsumen yang dihasilkan dari analisis data mining.

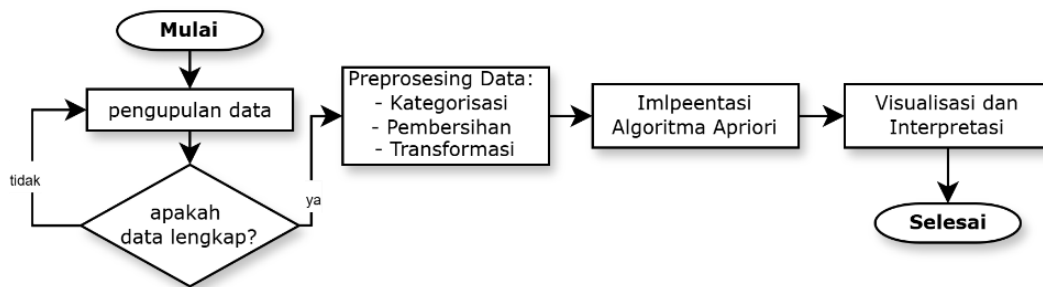
Manfaat analisis keranjang belanja (*Market Basket Analysis*) sangat beragam dan berdampak langsung pada optimalisasi strategi ritel. Analisis ini membantu pengecer menetapkan harga yang lebih strategis dengan mengidentifikasi produk-produk yang sering dibeli bersama [11]. Dengan pengetahuan ini, mereka dapat memberikan diskon pada produk tertentu untuk meningkatkan volume penjualan barang terkait. Selain itu, analisis ini membantu menentukan penempatan barang dengan menempatkan produk dengan afinitas tinggi secara berdekatan, sehingga mendorong pembelian impulsif. Strategi promosi juga dapat disesuaikan berdasarkan pola belanja individu pelanggan, misalnya dengan menawarkan kombinasi produk tertentu dengan harga diskon[12]. Selain itu, analisis ini memungkinkan pengecer memahami pengaruh antar produk dalam mendorong atau menurunkan penjualan, sehingga dapat menyusun langkah-langkah yang lebih terarah dalam menghadapi perubahan pasar.

Penelitian ini bertujuan untuk mengembangkan model analisis perilaku pembelian konsumen menggunakan metode data mining, mengimplementasikan dan mengevaluasi efektivitas penggunaan Python dengan library pandas dan seaborn dalam analisis data retail, serta menghasilkan sistem prediksi perilaku konsumen yang dapat membantu pengambilan keputusan bisnis. Manfaat yang diharapkan dari penelitian ini mencakup kontribusi

teoritis berupa pengembangan ilmu data mining dalam konteks retail dan memperkaya literatur tentang analisis perilaku konsumen menggunakan Python. Dari sisi praktis, penelitian ini dapat membantu pelaku usaha retail dalam mengoptimalkan strategi bisnis, meningkatkan efisiensi pengambilan keputusan berbasis data, serta memberikan pemahaman lebih baik tentang pola pembelian konsumen[13]. Untuk memastikan fokus dan kedalaman analisis, penelitian ini dibatasi pada data transaksi retail periode bulan Oktober 2024 dengan menggunakan Python versi 3.8 dengan library pandas versi 1.5.0 dan seaborn versi 0.12.0 untuk visualisasi. Data yang dianalisis mencakup transaksi penjualan, histori pembelian. Batasan ini diperlukan untuk memastikan hasil penelitian yang mendalam dan tepat sasaran dalam konteks yang spesifik.

2. METODE PENELITIAN

Penelitian ini menggunakan metode kuantitatif dengan pendekatan *data mining* untuk menganalisis pola pembelian konsumen. Teknik yang digunakan adalah *Association Rule Mining* dengan algoritma Apriori untuk menemukan hubungan atau pola antara barang-barang yang sering dibeli secara bersamaan oleh konsumen[14].



Gambar 1. Flowchart Tahapan/Metodologi Penelitian

a) Pengumpulan Data

Langkah pertama adalah mengumpulkan data transaksi dari usaha retail. Data yang dikumpulkan biasanya mencakup daftar barang yang dibeli konsumen dalam setiap transaksi. Data ini sering berupa data mentah (*raw data*) dari sistem POS (Point of Sales)[15].

b) Persiapan Data

Data yang dikumpulkan kemudian dibersihkan dan diolah agar siap untuk dianalisis. Tahapan ini meliputi penghapusan data duplikat, menangani data yang hilang (*missing values*) dan mengatur data dalam format yang sesuai untuk analisis menggunakan Python dan pustaka *Pandas*.

Tabel 1. Sampel *data set* yang digunakan dalam penelitian

No	ID Transaksi	Tanggal Transaksi	ID Pelanggan	Nama Produk	Jumlah Produk	Harga Produk
1	TR045	06/10/2024	C048	Susu UHT [Brand A]	5	15000
2	TR068	04/10/2024	C010	Apel Fuji (kg)	10	35000
3	TR088	09/10/2024	C071	Detergen	3	22000
4	TR066	08/10/2024	C040	Popok Bayi [Ukuran S]	1	60000
5	TR082	10/10/2024	C038	Pasta Gigi	1	15000
6	TR010	04/10/2024	C021	Minuman Soda	5	10000
...
299	TR049	02/10/2024	C057	Mie Instan [Rasa Ayam]	2	3000
300	TR053	08/10/2024	C033	Popok Bayi [Ukuran S]	3	60000

Data disusun dalam bentuk matriks transaksi di mana setiap baris mewakili transaksi dan setiap kolom mewakili item, dengan nilai 1 menunjukkan item yang dibeli dalam transaksi tertentu dan 0 jika tidak dibeli.

c) Penerapan Algoritma Apriori

Algoritma Apriori adalah algoritma utama yang digunakan dalam penelitian ini untuk menemukan *frequent itemsets* atau kombinasi item yang sering muncul bersama dalam transaksi. Tahapan penerapan algoritma Apriori meliputi:

- Pencarian Frequent Itemsets: Algoritma Apriori akan menghitung *support* untuk setiap item atau kombinasi item untuk menemukan pola yang paling sering muncul. *Support* adalah nilai probabilitas item muncul dalam transaksi.

- Penentuan Association Rules: Setelah menemukan *frequent itemsets*, aturan asosiasi antara item ditentukan berdasarkan *confidence* dan *lift*. *Confidence* menunjukkan seberapa sering aturan ini benar, sedangkan *lift* menunjukkan seberapa besar peningkatan probabilitas membeli item kedua jika item pertama dibeli[16].

Kedua parameter ini berperan vital dalam menyaring dan menghasilkan aturan-aturan asosiasi yang bermakna dari dataset transaksi. Pemilihan nilai yang tepat untuk kedua metrik ini akan menentukan kualitas hasil analisis yang dihasilkan. Nilai support mencerminkan frekuensi kemunculan kombinasi item dalam keseluruhan dataset. Penetapan ambang batas support yang tepat sangat krusial karena akan mempengaruhi relevansi pola yang ditemukan. Support yang terlalu rendah dapat menghasilkan pola yang tidak signifikan, sementara nilai yang terlalu tinggi berisiko mengabaikan pola yang potensial namun jarang muncul[17].

Parameter confidence mengindikasikan kekuatan hubungan antar item dalam suatu aturan asosiasi. Metrik ini dihitung sebagai probabilitas kemunculan item konsekuen ketika item antedeceden telah ada dalam transaksi. Penentuan nilai minimum confidence membantu memastikan bahwa aturan yang dihasilkan memiliki tingkat keyakinan yang memadai untuk dijadikan dasar pengambilan keputusan. Dengan menetapkan ambang batas support dan confidence yang sesuai, jumlah kombinasi item yang perlu diproses dapat berkurang drastis, sehingga meningkatkan efisiensi waktu pemrosesan terutama pada dataset berskala besar. Hal ini pada akhirnya menghasilkan aturan asosiasi yang tidak hanya akurat tetapi juga relevan untuk analisis bisnis.

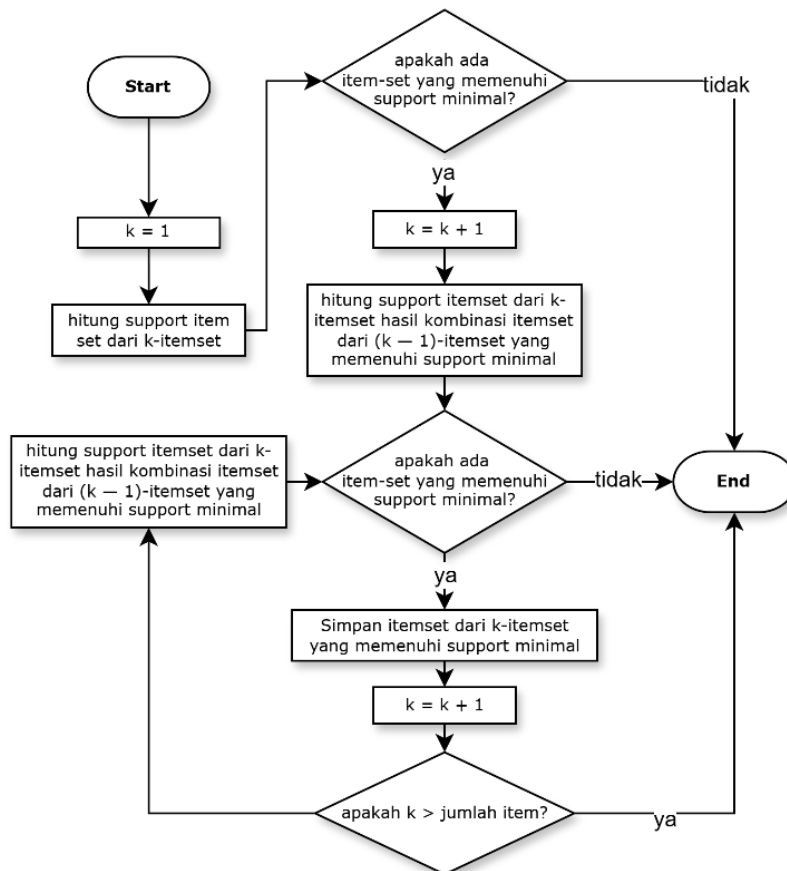
d) Evaluasi dan Interpretasi Hasil

Setelah aturan asosiasi terbentuk, hasilnya dievaluasi untuk menilai keakuratan dan relevansi aturan yang ditemukan. *Frequent itemsets* yang memiliki *support*, *confidence* dan *lift* tinggi diinterpretasikan untuk menentukan pola pembelian konsumen. Temuan yang relevan diidentifikasi sebagai dasar untuk strategi pemasaran dan pengelolaan stok produk.

e) Visualisasi Hasil

Untuk memudahkan pemahaman hasil analisis, hasil *frequent itemsets* dan aturan asosiasi divisualisasikan menggunakan grafik atau diagram, seperti matriks panas (*heatmap*) untuk menunjukkan pola pembelian atau diagram jaring untuk menampilkan hubungan antar produk.

f) Diagram Proses Algoritma Apriori



Gambar 2. Flowchart Algoritma Apriori

Berikut adalah penjelasan dari ilustrasi diagram alur penerapan algoritma Apriori dalam analisis data transaksi:

- 1) Input Data Transaksi: Pengguna memasukkan data transaksi ke dalam sistem untuk dianalisis.
- 2) Penghitungan Support untuk Item Tunggal: Algoritma menghitung *support* untuk setiap item tunggal, menyingkirkan item yang *support*-nya di bawah nilai minimum.
- 3) Pencarian Kombinasi Item yang Lebih Kompleks: Algoritma mencari kombinasi dua item, tiga item dan seterusnya, dengan mengulangi proses perhitungan *support* hingga menemukan semua *frequent itemsets* yang memenuhi syarat minimum.
- 4) Pembentukan Aturan Asosiasi: Setelah menemukan *frequent itemsets*, algoritma membentuk aturan asosiasi dengan menghitung *confidence* dan *lift* untuk setiap aturan.
- 5) Evaluasi dan Interpretasi Hasil: Hasil aturan asosiasi diinterpretasikan untuk menemukan pola pembelian yang paling relevan.

Berikut adalah beberapa rumus yang digunakan dalam algoritma Apriori:

- **Support:**

$$Support (Item) = \frac{\sum \text{transaksi mengandung item}}{\sum \text{transaksi}} \quad (1)$$

Nilai Support (Item) merupakan persentase jumlah kasus untuk kombinasi item tertentu. Dimana jumlah transaksi yang berisi item set dibagi total jumlah seluruh transaksi. Nilai support menjadi ukuran yang sangat penting dalam aturan asosiasi karena aturan yang sangat lemah nilai support-nya berarti asosiasi tersebut sangat jarang terjadi dalam dataset (seluruh data transaksi)[18].

- **Confidence:**

$$Confidence (A \rightarrow B) = \frac{\sum \text{transaksi mengandung A \& B}}{\sum \text{transaksi mengandung B}} \quad (2)$$

Nilai Confident (A dan B) merupakan persentase keakurasian dari aturan asosiasi yang dihasilkan. Nilai confident yang tinggi menggambarkan banyaknya B yang muncul dalam transaksi yang berisi A[19].

- **Lift:**

$$Lift Ratio = \frac{Confidence}{Expected confidence} \quad (3)$$

Lift adalah ukuran yang menunjukkan apakah probabilitas terjadinya kejadian B meningkat atau menurun jika kejadian A terjadi. Nilai lift digunakan untuk menentukan apakah aturan asosiasi valid atau tidak.

3. HASIL DAN PEMBAHASAN

a) Deskripsi Dataset

Dataset yang digunakan dalam penelitian ini terdiri atas 300 transaksi, yang diperoleh dari sistem pencatatan penjualan retail Azzam Mart. Dataset mencakup empat atribut utama, yaitu:

- ✓ ID Transaksi: Identitas unik setiap transaksi.
- ✓ ID Pelanggan: Identitas unik pelanggan.
- ✓ Tanggal Transaksi: Waktu terjadinya transaksi.
- ✓ Daftar Item yang Dibeli: Daftar produk yang dibeli dalam setiap transaksi.

Pra-pemrosesan data dilakukan untuk mengonversi data transaksi menjadi format matriks dengan menggunakan pendekatan binary (1 untuk produk yang dibeli, 0 untuk produk yang tidak dibeli). Transformasi ini memungkinkan penerapan algoritma Apriori untuk analisis pola pembelian.

b) Penerapan Algoritma Apriori

Analisis dilakukan dengan menggunakan Python 3.8.0 dan library pandas. Parameter yang diterapkan dalam algoritma Apriori adalah:

- ✓ Minimum Support: 0.05 (artinya pola harus muncul pada minimal 5% transaksi).
- ✓ Minimum Confidence: 1 (pola dianggap valid jika memiliki tingkat kepercayaan 100%).

Berikut adalah kode program yang digunakan untuk melakukan analisis dataset tersebut:

```

import pandas as pd
from mlxtend.frequent_patterns import apriori, association_rules
# Load dataset
df = pd.read_csv('dataset_perilaku_pembelian.csv')
# Data Preprocessing
basket = df.groupby(['ID Transaksi', 'Nama Produk'])['Jumlah
Produk'].sum().unstack().fillna(0)
basket = basket.applymap(lambda x: x > 0)

# Apply Apriori algorithm
frequent_itemsets = apriori(basket, min_support=0.05, use_colnames=True)
frequent_itemsets = frequent_itemsets.sort_values(by="support", ascending=False)
# Limit frequent itemsets to the top 20
top_itemsets = frequent_itemsets.head(20)
# Generate association rules
rules = association_rules(frequent_itemsets, metric="lift", min_threshold=1)
rules = rules.sort_values(by="confidence", ascending=False)
# Limit association rules to the top 10
top_rules = rules[['antecedents', 'consequents', 'support', 'confidence', 'lift']].head(10)

# Display results
print("Top 20 Frequent Itemsets:")
print(top_itemsets)
print("\nTop 10 Association Rules:")
print(top_rules)
# Save frequent itemsets to CSV
top_itemsets.to_csv('frequent_itemsets.csv', index=False)
# Save association rules to CSV
top_rules.to_csv('association_rules.csv', index=False)
# Save frequent itemsets and association rules to TXT
with open('analysis_results.txt', 'w') as f:
    f.write("Top 20 Frequent Itemsets:\n")
    f.write(top_itemsets.to_string(index=False))
    f.write("\n\nTop 10 Association Rules:\n\n")
    f.write(top_rules.to_string(index=False))

```

Gambar 3. Kode Program Python Untuk Analisis *Dataset*

Gambar 3 diatas menampilkan kode program yang ditulis menggunakan bahasa pemrograman Python. Program ini dirancang untuk melakukan analisis perilaku pembelian konsumen di Azzam Mart dengan metode *data mining* dan algoritma Apriori. Pada bagian awal, program mengimpor pustaka yang diperlukan, yaitu pandas untuk manipulasi data dan mlxtend.frequent_patterns yang berisi fungsi apriori dan association_rules. Selanjutnya, program membaca dataset dari file 'dataset_perilaku_pembelian.csv', melakukan pra-pemrosesan data untuk mengubah format transaksi menjadi matriks biner, dan menerapkan algoritma Apriori. Parameter yang digunakan adalah *min_support* sebesar 0.05 untuk menemukan *frequent itemsets* dan *min_threshold* 1 untuk metrik *lift* guna menghasilkan aturan asosiasi. Terakhir, program mengurutkan dan membatasi hasil menjadi 20 *frequent itemsets* teratas dan 10 aturan asosiasi teratas, lalu menampilkannya serta menyimpannya ke dalam file CSV dan TXT."

```

===== RESTART: C:\Python38\DataMining\DataMining.py =====
Top 20 Frequent Itemsets:
   support  itemsets
19  0.206106  (Telur Ayam)
 7  0.206106  (Minuman Soda)
18  0.195876  (Susu UHT [Brand B])
 0  0.175258  (Apel Fuji (kg))
17  0.175258  (Susu UHT [Brand A])
11  0.164948  (Popok Bayi [Ukuran M])
12  0.164948  (Popok Bayi [Ukuran S])
 1  0.154639  (Beras 10kg)
 4  0.154639  (Jeruk Mandarin (kg))
 8  0.154639  (Minyak Goreng 1L)
10  0.154639  (Pasta Gigi)
 3  0.134021  (Detergen)
14  0.134021  (Sabun Mandi)
13  0.123711  (Roti Tawar)
 9  0.123711  (Minyak Goreng 2L)
 2  0.123711  (Beras 5kg)
15  0.113402  (Sampo)
16  0.103093  (Snack Kentang)
 5  0.082474  (Mie Instan [Rasa Ayam])
 6  0.061856  (Mie Instan [Rasa Kari])

Top 10 Association Rules:
   antecedents  consequents  support  confidence  lift
9  (Roti Tawar)  (Minyak Goreng 1L)  0.051546  0.416667  2.694444
6  (Sabun Mandi)  (Minuman Soda)  0.051546  0.384615  1.865385
3  (Beras 10kg)  (Minuman Soda)  0.051546  0.333333  1.616667
5  (Jeruk Mandarin (kg))  (Minuman Soda)  0.051546  0.333333  1.616667
8  (Minyak Goreng 1L)  (Roti Tawar)  0.051546  0.333333  2.694444
10 (Minyak Goreng 1L)  (Susu UHT [Brand B])  0.051546  0.333333  1.701754
0  (Telur Ayam)  (Minuman Soda)  0.061856  0.300000  1.455000
1  (Minuman Soda)  (Telur Ayam)  0.061856  0.300000  1.455000
11 (Susu UHT [Brand B])  (Minyak Goreng 1L)  0.051546  0.263158  1.701754
2  (Minuman Soda)  (Beras 10kg)  0.051546  0.250000  1.616667
>>> |

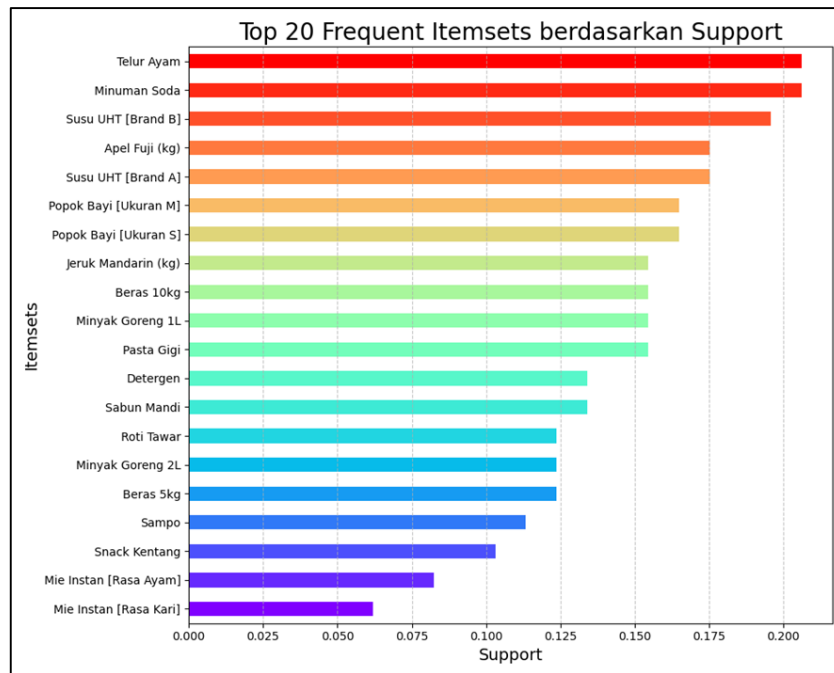
```

Gambar 4. Output IDLE Python

Tangkapan layar pada Gambar diatas menampilkan hasil langsung dari eksekusi skrip Python di lingkungan IDLE. Output ini terbagi menjadi dua bagian utama. Bagian pertama adalah "Top 20 Frequent Itemsets" , yang menyajikan daftar 20 produk dengan nilai support tertinggi, menunjukkan barang-barang yang paling sering dibeli oleh konsumen. Sebagai contoh, "Telur Ayam" dan "Minuman Soda" menempati posisi teratas dengan nilai support yang identik yaitu 0.206186. Bagian kedua adalah "Top 10 Association Rules", yang menampilkan 10 aturan asosiasi terkuat yang ditemukan. Setiap aturan disajikan lengkap dengan metrik evaluasinya: antecedents (produk pemicu), consequents (produk yang ikut terbeli), serta nilai support, confidence dan lift. Data inilah yang menjadi dasar untuk analisis dan visualisasi lebih lanjut dalam penelitian.

c) Hasil Analisis Pola Asosiasi

Hasil implementasi algoritma Apriori dapat dirangkum sebagai berikut, Hasil analisis mencakup 20 frequent itemsets dan 10 aturan asosiasi utama.



Gambar 5. Top 20 Frequent Itemsets berdasarkan Support

Gambar di atas menyajikan visualisasi data dalam bentuk diagram batang horizontal yang mengilustrasikan 20 item atau produk yang paling sering dibeli oleh konsumen (frequent itemsets), diurutkan berdasarkan nilai support tertinggi. Sumbu vertikal (Y) pada diagram ini mencantumkan nama-nama produk, sedangkan sumbu horizontal (X) merepresentasikan nilai support untuk setiap produk tersebut, yang menandakan frekuensi kemunculan produk dalam total transaksi. Dari visualisasi ini, dapat dengan jelas terlihat bahwa "Telur Ayam" dan "Minuman Soda" memiliki nilai support tertinggi, menempatkannya sebagai produk yang paling populer di kalangan pembeli. Penggunaan gradasi warna dari merah (nilai support tinggi) hingga ungu (nilai support rendah) membantu mempertegas peringkat popularitas setiap item secara visual.

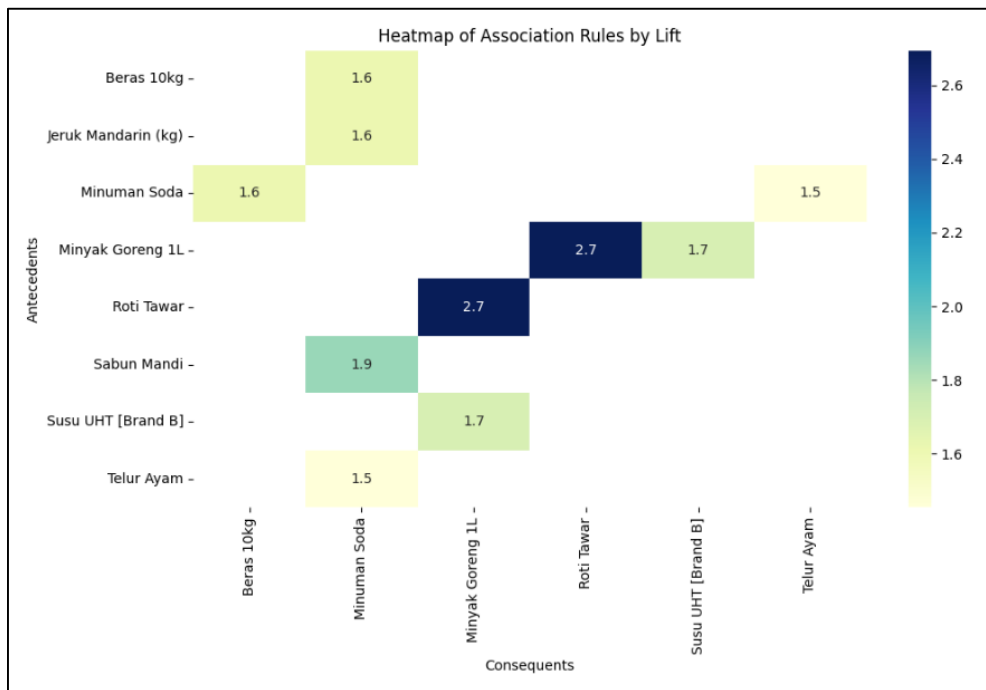
Tabel 2. Aturan Asosiasi (10 aturan utama)

Antecedents	Consequents	Support	Confidence	Lift
Roti Tawar	Minyak Goreng 1L	0.051546	0.416667	2.694444
Sabun Mandi	Minuman Soda	0.051546	0.384615	1.865385
Beras 10kg	Minuman Soda	0.051546	0.333333	1.616667
Jeruk Mandarin (kg)	Minuman Soda	0.051546	0.333333	1.616667
Minyak Goreng 1L	Roti Tawar	0.051546	0.333333	2.694444
Minyak Goreng 1L	Susu UHT [Brand B]	0.051546	0.333333	1.701754
Telur Ayam	Minuman Soda	0.061856	0.300000	1.455000
Minuman Soda	Telur Ayam	0.061856	0.300000	1.455000
Susu UHT [Brand B]	Minyak Goreng 1L	0.051546	0.263158	1.701754
Minuman Soda	Beras 10kg	0.051546	0.250000	1.616667

Frequent itemsets menunjukkan bahwa produk Telur Ayam dan Minuman Soda adalah dua produk yang paling sering muncul dalam transaksi. Sementara itu, aturan asosiasi menunjukkan bahwa pembelian Roti Tawar memiliki kecenderungan kuat untuk disertai pembelian Minyak Goreng 1L, dengan confidence sebesar 41,67% dan lift 2,69. Dari temuan ini, korelasi positif yang signifikan antara pembelian 'Roti Tawar' dan 'Minyak Goreng 1L' dapat menjadi dasar bagi manajemen Azzam Mart untuk mengambil keputusan strategis. Nilai *lift* sebesar 2,69 mengindikasikan bahwa kedua produk tersebut 2,69 kali lebih mungkin dibeli bersamaan daripada secara kebetulan. Temuan ini memberikan wawasan strategis untuk mengoptimalkan penempatan produk dengan menempatkan 'Roti Tawar' dan 'Minyak Goreng 1L' berdekatan, sehingga mendorong pembelian impulsif dan meningkatkan volume penjualan. Selain itu, data ini juga dapat digunakan untuk merancang program promosi yang efektif, seperti promosi *bundling* di mana kedua produk tersebut ditawarkan dalam satu paket dengan harga khusus.

d) Visualisasi dan Interpretasi data

Untuk mempermudah interpretasi hasil analisis aturan asosiasi yang kompleks, penelitian ini memanfaatkan visualisasi dengan diagram panas (*heatmap*) yang dibuat menggunakan pustaka Seaborn pada Python. Heatmap merupakan visualisasi data berbasis warna, dimana intensitas warna menggambarkan besaran nilai dalam data, seperti warna biru tua untuk nilai tinggi dan kuning muda untuk nilai rendah. Dalam aturan asosiasi ini, terdapat komponen anteseden yang berperan sebagai pemicu atau kondisi awal dan konsekuen sebagai hasil atau konsekuensi dari kondisi tersebut.



Gambar 6. Heatmap of Association Rules by Lift

Diagram *heatmap* tersebut memvisualisasikan hubungan antara produk-produk dalam aturan asosiasi, dengan nilai *Lift* direpresentasikan oleh intensitas warna. Semakin gelap warnanya, semakin tinggi nilai *Lift*-nya, yang mengindikasikan asosiasi yang lebih kuat dan signifikan antara dua produk. Berikut adalah poin-poin interpretasi utama dari diagram tersebut:

- a) Asosiasi terkuat ditunjukkan oleh kotak berwarna biru tua dengan nilai *Lift* 2.7. Terdapat dua asosiasi yang memiliki nilai *Lift* tertinggi ini:
 - Roti Tawar (sebagai *antecedent*) dan Minyak Goreng 1L (sebagai *consequent*). Ini menunjukkan bahwa pelanggan yang membeli Roti Tawar cenderung kuat juga membeli Minyak Goreng 1L.
 - Minyak Goreng 1L (sebagai *antecedent*) dan Roti Tawar (sebagai *consequent*). Ini adalah kebalikan dari aturan di atas, yang juga menunjukkan hubungan timbal balik yang sangat kuat. Nilai *Lift* sebesar 2.7 mengindikasikan bahwa probabilitas pembelian kedua produk ini bersamaan 2.7 kali lebih tinggi daripada jika dibeli secara acak.
- b) Selain asosiasi terkuat, terdapat beberapa asosiasi lain dengan nilai *Lift* yang juga signifikan:
 - Sabun Mandi dan Minuman Soda memiliki nilai *Lift* 1.9.
 - Minyak Goreng 1L dan Susu UHT [Brand B] memiliki nilai *Lift* 1.7.
 - Susu UHT [Brand B] dan Minyak Goreng 1L juga memiliki nilai *Lift* 1.7.

- Asosiasi antara Beras 10kg dan Minuman Soda memiliki nilai *Lift* 1.6.
 - Asosiasi antara Jeruk Mandarin (kg) dan Minuman Soda memiliki nilai *Lift* 1.6.
 - Minuman Soda dan Beras 10kg memiliki nilai *Lift* 1.6.
- c) *Heatmap* ini memberikan wawasan yang mudah dipahami bagi manajemen untuk membuat keputusan bisnis. Misalnya, manajemen dapat menempatkan "Roti Tawar" dan "Minyak Goreng 1L" berdekatan di rak toko atau membuat paket promosi (*bundling*) untuk meningkatkan penjualan. Nilai *Lift* yang lebih besar dari 1 menunjukkan bahwa semua aturan yang ditampilkan dalam *heatmap* ini adalah valid dan tidak terjadi secara kebetulan.

e) **Prediksi Perilaku Konsumen**

Berdasarkan analisis pola asosiasi, prediksi perilaku konsumen adalah:

- 1) Produk Roti Tawar dan Minyak Goreng 1L cenderung dibeli bersama. Strategi promosi seperti *bundling* dapat meningkatkan penjualan kedua produk.
- 2) Produk dengan support tinggi seperti Telur Ayam dan Minuman Soda sebaiknya ditempatkan di lokasi strategis di toko untuk mendorong pembelian impulsif.
- 3) Kombinasi produk seperti Susu UHT [Brand B] dan Minyak Goreng 1L memiliki potensi untuk digunakan dalam kampanye promosi tematik (misalnya, paket hemat).

Hasil ini menunjukkan bahwa algoritma Apriori dapat mengidentifikasi pola pembelian yang relevan untuk mendukung strategi bisnis retail. Hasil ini sejalan dengan studi sebelumnya oleh (Nofriansyah D, Nurcahyo G) yang menyoroti keunggulan Apriori dalam mengidentifikasi asosiasi antar produk. Namun, terdapat keterbatasan dalam penelitian ini, seperti ukuran dataset yang relatif kecil dan tidak mencakup informasi demografis pelanggan. Untuk penelitian selanjutnya, disarankan untuk menggunakan dataset yang lebih besar dengan atribut tambahan seperti usia atau preferensi pelanggan untuk analisis yang lebih mendalam.

4. KESIMPULAN

Penelitian ini berhasil menjawab permasalahan fundamental yang dihadapi usaha retail, yaitu kesulitan dalam menganalisis data transaksi untuk menentukan pola pembelian konsumen secara akurat. Melalui implementasi sistematis algoritma Apriori dengan bahasa pemrograman Python, penelitian ini menyajikan solusi praktis untuk mengubah data mentah menjadi wawasan bisnis strategis. Jawaban atas pertanyaan penelitian mengenai implementasi teknis ditunjukkan melalui alur kerja yang jelas, data transaksi diolah menggunakan pustaka Pandas untuk mengatasi kompleksitasnya, dianalisis dengan algoritma Apriori untuk menemukan pola dan hasilnya divisualisasikan dengan Seaborn. Kerangka kerja ini terbukti efektif dan dapat direplikasi, menjawab tantangan kurangnya pemanfaatan data mining di tingkat retail kecil dan menengah.

Secara spesifik, penelitian ini berhasil menentukan pola pembelian konsumen dengan akurat, terbukti dari identifikasi frequent itemsets seperti 'Telur Ayam' dan 'Minuman Soda' yang memiliki support tertinggi. Lebih penting lagi, penelitian ini menghasilkan aturan asosiasi yang dapat diandalkan untuk prediksi perilaku, seperti hubungan kuat antara 'Roti Tawar' dan 'Minyak Goreng 1L' dengan nilai confidence 41,7% dan lift 2,69. Tingginya nilai lift ini mengonfirmasi bahwa pola tersebut sangat valid dan bukan terjadi secara kebetulan, sehingga memberikan dasar yang kuat untuk pengambilan keputusan bisnis, seperti strategi *bundling* produk atau optimasi tata letak barang. Dengan demikian, studi ini secara komprehensif menunjukkan bagaimana penerapan data mining dapat memberikan keuntungan kompetitif yang nyata. Meskipun berhasil, penelitian ini memiliki keterbatasan seperti ukuran dataset yang relatif kecil. Oleh karena itu, penelitian selanjutnya disarankan untuk menggunakan dataset yang lebih besar, mengeksplorasi algoritma yang lebih efisien seperti FP-Growth dan mengembangkan model prediktif dengan machine learning untuk analisis yang lebih mendalam dan prediksi yang lebih akurat di masa mendatang.

UCAPAN TERIMAKASIH

Penulis mengucapkan terima kasih yang sebesar-besarnya kepada Universitas Kuningan atas dukungan dan pendanaan melalui hibah internal yang memungkinkan terlaksananya penelitian ini. Tak lupa apresiasi dan rasa terima kasih kepada pemilik usaha ritel Azzam Mart yang dengan baik hati telah memberikan izin dan menyediakan data transaksi sebagai *dataset* utama untuk penelitian ini. Semoga hasil dari penelitian ini dapat memberikan manfaat bagi pihak universitas, dunia akademis dan industri retail di Indonesia.

REFERENCES

- [1] E. Elisa, "Market Basket Analysis Pada Mini Market Ayu Dengan Algoritma Apriori," vol. 2, no. 2, pp. 472–478, 2018, [Online]. Available: <http://jurnal.iaii.or.id>

- [2] E. Alma, E. Utami, and F. Wahyu Wibowo, "Implementasi Algoritma Apriori untuk Rekomendasi Produk pada Toko Online Implementation of Apriori Algorithms for Product Recommendations at Online Stores," *Citec Journal*, vol. 7, no. 1, 2020.
- [3] M. Hasanul Huda and M. Syafrullah, "PENERAPAN METODE ASSOCIATION RULES DENGAN ALGORITME APRIORI UNTUK Mencari Pola Keterkaitan Setiap Item Lewat Perilaku Pelanggan Pada Gavio Boutique Collection Jakarta Pusat."
- [4] M. F. Mulya, N. Rismawati, and R. R. Alifi, "ANALISIS DAN IMPLEMENTASI DATA MINING MENGGUNAKAN ALGORITMA APRIORI UNTUK MENINGKATKAN PENJUALAN PADA KANTIN UNIVERSITAS TANRI ABENG," *Faktor Exacta*, vol. 12, no. 3, p. 210, Nov. 2019, doi: 10.30998/faktorexacta.v12i3.4541.
- [5] A. N. Rahmi and A. Mikola, "IMPLEMENTASI ALGORITMA APRIORI UNTUK MENENTUKAN POLA PEMBELIAN PADA CUSTOMER (STUDI KASUS : TOKO BAKOEL SEMBAKO)."
- [6] M. Sholeh and D. Andayati, "Penerapan Metode Clustering dengan Algoritma K-Means Pada Pengelompokan Indeks Prestasi Akademik Mahasiswa," *SKANIKA: Sistem Komputer dan Teknik Informatika*, vol. 6, no. 1, pp. 51–60, 2023.
- [7] A. Naufal Waliyus Zain and A. Tholib, "Klasifikasi Data Mining di Tingkat Kepuasan Mahasiswa Terhadap Pelayanan Sistem Informasi Fakultas Teknik Universitas Nurul Jadid," vol. 7, no. 2, pp. 204–213, 2024.
- [8] D. Nofriansyah and G. W. Nurcahyo, *Algoritma data mining dan pengujian*. Deepublish, 2015.
- [9] G. Gunadi and D. Indra Sensuse, "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," 2012.
- [10] D. Anggraini, S. A. Putri, and L. A. Utami, "Implementasi Algoritma Apriori Dalam Menentukan Penjualan Mobil Yang Paling Diminati Pada Honda Permata Serpong," *JURNAL MEDIA INFORMATIKA BUDIDARMA*, vol. 4, no. 2, p. 302, Apr. 2020, doi: 10.30865/mib.v4i2.1496.
- [11] A. H. Priyanto and A. B. Arifa, "Implementation of Market Basket Analysis with Apriori Algorithm in Minimarket," *Jurnal Teknik Informatika (Jutif)*, vol. 3, no. 5, pp. 1423–1429, 2022.
- [12] T. D. Putra, "Analisis Keranjang Belanja dengan Algoritma Apriori Klasik pada Data Mining," *Jurnal Kajian Ilmiah*, vol. 20, no. 1, pp. 59–66, 2020.
- [13] "Penerapan Asosiasi Produk Berbasis AI untuk Rekomendasi Bisnis di Industri 4.0 dengan Pertimbangan Analisis Perilaku Beli Konsumen".
- [14] R. Arista, A. Nugroho, and N. T. Kurniadi, "PENERAPAN DATA MINING UNTUK PREDIKSI POLA PEMBELIAN PELANGGAN MENGGUNAKAN ALGORITMA APRIORI (STUDI KASUS: TOKO JIHAN)," vol. 14, 2023.
- [15] A. M. Muhsidi and A. Sumiah, "ANALISIS IMPLEMENTASI APLIKASI BIG DATA PADA INDUSTRI KESEHATAN, KEUANGAN DAN PENDIDIKAN," *Digital Business and Entrepreneurship Journal*, vol. 3, no. 1, pp. 24–35, Jan. 2025, doi: 10.25134/digibe.v3i1.256.
- [16] H. Mawengkang and E. Budhiarti Nababan, "PENGEMBANGAN ALGORITMA APRIORI UNTUK PENGAMBILAN KEPUTUSAN THE DEVELOPMENT APRIORI ALGORITHM FOR DECISION-MAKING," 2015.
- [17] E. Elisa, K. Handoko, T. dan Komputer, and U. Putera Batam, "PENERAPAN DATA MINING METODE APRIORI DALAM ANALISIS KECENDRONGAN PEMBELIAN KONSUMEN GROCERY SHOP," *AJurnal TEKINKOM*, vol. 6, no. 1, 2023, doi: 10.37600/tekinkom.v6i1.817.
- [18] G. Soepriyono and A. Triayudi, "Implementasi Data Mining dengan Algoritma Apriori dalam Menentukan Pola Pembelian Aksesoris Laptop," *JURNAL MEDIA INFORMATIKA BUDIDARMA*, vol. 7, no. 4, p. 2087, Oct. 2023, doi: 10.30865/mib.v7i4.6555.
- [19] F. Achmad, O. Nurdiawan, and Y. A. Wijaya, "ANALISA POLA TRANSAKSI PEMBELIAN KONSUMEN PADA TOKO RITEL KESEHATAN MENGGUNAKAN ALGORITMA FP-GROWTH," 2023. [Online]. Available: www.researchgate.net