

## Penerapan Market Basket Analysis dengan Algoritma Apriori pada Data Transaksi Pos Retail untuk Mendukung Strategi Penjualan

Dendi Winardi<sup>1</sup>, Gilang Ramadhan Pratama<sup>2</sup>, Nandar Brayun Syah Lubis<sup>3</sup>

<sup>1,2,3</sup>Program Studi Sistem Informasi, Universitas Pamulang

Email: [dendiwinardi777@gmail.com](mailto:dendiwinardi777@gmail.com)

### Article Info

#### Article history:

Received June 01, 2026

Revised June 10, 2026

Accepted June 20, 2026

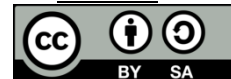
#### Keywords:

Apriori Algorithm, Data Mining, Interactive Dashboard, Market Basket Analysis, Retail.

### ABSTRACT

*Increasing competition in the retail industry requires businesses to understand customer purchasing patterns to support more effective decision-making. This study applies Market Basket Analysis (MBA) using the Apriori algorithm to identify product associations and generate data-driven marketing recommendations. The research follows the CRISP-DM methodology, consisting of business understanding, data understanding, data preparation, modeling, evaluation, and deployment stages. The dataset was obtained from the Retail POS Basket Data on Kaggle, comprising 1,991 unique transactions and 68 product categories. The analysis results indicate that a minimum support of 0.2% and a minimum lift of 1.0 produced 2,370 frequent itemsets and 1,974 association rules. The strongest rule identified was {Energy Drink, Fish} → {Potato}, with a confidence value of 66.7% and a lift value of 9.69. Furthermore, the resulting association rules were implemented in an interactive Streamlit-based dashboard to facilitate dynamic product recommendation and pattern exploration. The findings demonstrate that Market Basket Analysis can provide valuable insights to support marketing strategies and improve sales performance in the retail sector.*

*This is an open access article under the CC BY-SA license.*



### Article Info

#### Article history:

Received June 01, 2026

Revised June 10, 2026

Accepted June 20, 2026

#### Kata kunci:

Algoritma Apriori, Dashboard Interaktif, Data Mining, Market Basket Analysis, Ritel.

### ABSTRAK

Persaingan industri ritel yang semakin ketat menuntut pelaku usaha untuk memahami pola pembelian konsumen guna mendukung pengambilan keputusan yang lebih efektif. Penelitian ini menerapkan metode Market Basket Analysis (MBA) menggunakan algoritma Apriori untuk mengidentifikasi hubungan antarproduk dan menghasilkan rekomendasi strategi pemasaran berbasis data. Metodologi yang digunakan adalah CRISP-DM yang mencakup tahapan business understanding, data understanding, data preparation, modeling, evaluation, dan deployment. Dataset penelitian berasal dari Retail POS Basket Data di Kaggle yang terdiri atas 1.991 transaksi unik dan 68 jenis produk. Hasil analisis menunjukkan bahwa parameter minimum support 0,2% dan minimum lift 1,0 menghasilkan 2.370 frequent itemset serta 1.974 aturan asosiasi. Aturan terkuat yang ditemukan adalah {Energy Drink, Fish} → {Potato} dengan nilai confidence 66,7% dan lift 9,69. Selanjutnya, hasil analisis diimplementasikan ke dalam dashboard interaktif berbasis Streamlit yang mendukung eksplorasi aturan asosiasi dan rekomendasi produk secara dinamis. Temuan penelitian menunjukkan bahwa penerapan Market Basket Analysis mampu menghasilkan informasi yang bermanfaat untuk mendukung strategi pemasaran dan peningkatan penjualan pada sektor ritel.



**Corresponding Author:**

Dendi Winardi  
Universitas Pamulang  
Email: dendiwinardi777@gmail.com

## PENDAHULUAN

Perkembangan teknologi informasi telah mendorong transformasi signifikan pada sektor bisnis ritel. Tingkat persaingan yang semakin kompetitif menuntut pelaku usaha untuk memanfaatkan data sebagai aset strategis dalam mendukung pengambilan keputusan bisnis [1]. Sistem Point of Sale (POS) yang digunakan dalam operasional toko ritel menghasilkan data transaksi dalam jumlah besar setiap harinya. Namun, data tersebut sering kali hanya berfungsi sebagai arsip digital dan belum dimanfaatkan secara optimal untuk menghasilkan informasi yang bernilai bagi perusahaan [2].

Data transaksi penjualan menyimpan informasi mengenai pola perilaku belanja konsumen yang dapat digunakan untuk mengidentifikasi hubungan antarproduk dalam suatu transaksi. Informasi tersebut penting bagi pengelola bisnis karena dapat digunakan sebagai dasar dalam penyusunan strategi pemasaran, seperti promosi bundling, cross-selling, dan penataan produk pada rak penjualan untuk meningkatkan nilai transaksi pelanggan [3], [4]. Salah satu teknik data mining yang banyak digunakan untuk menemukan pola hubungan antarproduk adalah Market Basket Analysis (MBA). Market Basket Analysis merupakan metode analisis yang bertujuan mengidentifikasi keterkaitan antar item dalam suatu transaksi sehingga dapat diketahui produk-produk yang sering dibeli secara bersamaan [5], [6]. Teknik ini termasuk ke dalam kategori association rule mining yang berfokus pada pencarian pola asosiasi dari kumpulan data transaksi.

Dalam penerapannya, salah satu algoritma yang paling banyak digunakan adalah Apriori. Algoritma Apriori bekerja dengan mencari frequent itemset berdasarkan nilai minimum support, kemudian membentuk aturan asosiasi menggunakan metrik support, confidence, dan lift [7], [8]. Prinsip utama algoritma ini menyatakan bahwa jika suatu kombinasi item memenuhi nilai minimum support, maka seluruh subset dari kombinasi tersebut juga harus memenuhi nilai minimum support. Pendekatan tersebut memungkinkan proses pencarian pola dilakukan secara sistematis dan efisien [9].

Beberapa penelitian sebelumnya telah menerapkan algoritma Apriori untuk menganalisis pola pembelian konsumen. Hermawan dkk. memperoleh aturan asosiasi dengan nilai lift tertinggi sebesar 9,689 pada data transaksi supermarket [10]. Brighton dan Hariyanto menerapkan pendekatan CRISP-DM pada data transaksi toko ritel elektronik menggunakan minimum support 20% dan minimum confidence 60% [11]. Purwati dan Karnila menerapkan Market Basket Analysis pada data penjualan produk frozen food menggunakan perangkat lunak Orange dengan minimum support 30% dan minimum confidence 90% [12]. Sementara itu, Rizaldi dan Adnan melakukan analisis pola pembelian pada data transaksi 212 Mart Pekanbaru dan menghasilkan aturan asosiasi dengan nilai lift tertinggi sebesar 2,665 [13].

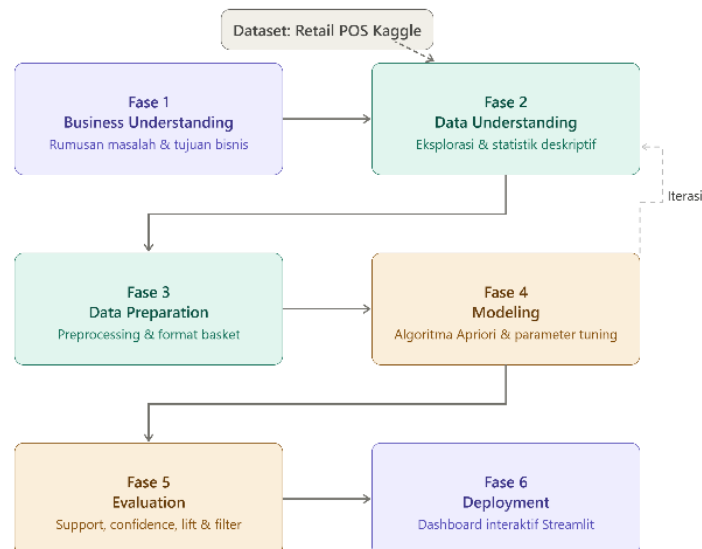
Meskipun penelitian mengenai Market Basket Analysis telah banyak dilakukan, masih terdapat peluang pengembangan dalam proses analisis dan penyajian hasil. Penelitian ini menawarkan beberapa kontribusi, yaitu melakukan eksplorasi parameter minimum support melalui lima skenario pengujian yang berbeda (0,001; 0,002; 0,003; 0,005; dan 0,01), menerapkan filter redundansi untuk menghilangkan aturan asosiasi simetris sehingga diperoleh aturan yang lebih representatif, serta menyajikan hasil analisis dalam bentuk dashboard interaktif berbasis Streamlit yang memungkinkan pengguna melakukan eksplorasi data secara dinamis.

Berdasarkan permasalahan tersebut, penelitian ini bertujuan menerapkan algoritma Apriori pada data transaksi ritel publik untuk menemukan pola asosiasi antarproduk, menghasilkan aturan asosiasi berdasarkan nilai support, confidence, dan lift, serta merumuskan rekomendasi strategi pemasaran yang dapat digunakan untuk mendukung peningkatan penjualan melalui pendekatan bundling dan cross-selling.

## METODE

### 1. Kerangka CRISP-DM

Penelitian ini mengadopsi kerangka *Cross-Industry Standard Process for Data Mining* (CRISP-DM) sebagai panduan metodologis yang bersifat iteratif dan terstruktur. Kerangka ini dipilih karena telah diakui secara luas sebagai standar de facto dalam pelaksanaan proyek data mining dan terbukti efektif dalam penelitian sejenis [1, 2]. CRISP-DM terdiri dari enam fase yang saling berkaitan, yakni *business understanding*, *data understanding*, *data preparation*, *modeling*, *evaluation*, dan *deployment*.



Pemetaan setiap fase terhadap penelitian ini adalah sebagai berikut. Fase *business understanding* diwujudkan dalam perumusan masalah dan penetapan tujuan penelitian yang telah diuraikan pada Bab 1. Fase *data understanding* dan *data preparation* dijelaskan secara rinci pada sub-bab 3.2 dan 3.3. Fase *modeling* dijabarkan pada sub-bab 3.4, sedangkan fase *evaluation* dan *deployment* disajikan pada Bab 4. Sifat iteratif kerangka ini memungkinkan peninjauan ulang antar fase apabila hasil evaluasi menunjukkan kebutuhan penyesuaian parameter atau pendekatan pemodelan.

## 2. Deskripsi Dataset

Dataset yang digunakan dalam penelitian ini adalah *Retail POS Dataset for Market Basket Analysis* yang bersumber dari platform Kaggle [18]. Dataset tersebut merepresentasikan transaksi penjualan ritel yang disimulasikan selama periode satu tahun, yaitu dari 1 Januari hingga 31 Desember 2024.

Dataset terdiri atas 10.000 baris data yang mencakup 1.991 transaksi unik dengan 68 jenis produk yang dikelompokkan ke dalam 17 kategori produk. Setiap baris data merepresentasikan satu item produk yang dibeli dalam suatu transaksi.

Atribut yang digunakan dalam penelitian meliputi `order_id` sebagai identitas transaksi, `user_id` sebagai identitas pelanggan, `order_date` sebagai tanggal transaksi, `time` sebagai waktu transaksi, `order_hour_of_day` sebagai jam transaksi, `product_name` sebagai nama produk, `quantity` sebagai jumlah pembelian, `price` sebagai harga satuan produk, `category` sebagai kategori produk, serta `product_id` sebagai identitas unik produk [18].

Dataset ini digunakan sebagai sumber data utama untuk proses analisis pola pembelian konsumen menggunakan algoritma Apriori. Analisis karakteristik dan statistik deskriptif dataset disajikan pada Bab 4 sebagai bagian dari hasil penelitian.

## 3. Persiapan dan Preprocessing Data

Tahap persiapan data bertujuan untuk memastikan data berada dalam kondisi yang memenuhi persyaratan teknis algoritma Apriori. Proses ini mencakup tiga langkah utama: pembersihan data, penanganan duplikat, dan transformasi data ke dalam format yang sesuai.

- Pembersihan data dilakukan melalui pemeriksaan terhadap seluruh atribut untuk mengidentifikasi kemungkinan adanya *missing value* maupun data yang tidak valid. Tahap ini bertujuan memastikan bahwa data yang digunakan dalam proses pemodelan memiliki kualitas yang memadai dan tidak menimbulkan bias pada hasil analisis [1].
- Penanganan duplikasi dilakukan pada kombinasi atribut `order_id` dan `product_name`. Pemeriksaan ini diperlukan karena terdapat kemungkinan suatu produk tercatat lebih dari satu kali dalam transaksi yang sama. Apabila ditemukan data duplikat, dilakukan proses agregasi menggunakan fungsi `groupby` dengan penjumlahan nilai `quantity`, sehingga setiap kombinasi transaksi dan produk hanya direpresentasikan oleh satu baris data [2].
- Transformasi data dilakukan dengan mengubah struktur data tabular menjadi format *basket transaction* menggunakan teknik *one-hot encoding*. Dalam format ini, setiap baris merepresentasikan satu transaksi berdasarkan `order_id`, sedangkan setiap kolom merepresentasikan produk. Nilai biner digunakan untuk menunjukkan keberadaan atau ketidakterdapatnya suatu produk pada transaksi tertentu. Hasil transformasi ini menghasilkan matriks transaksi yang siap digunakan dalam proses pembentukan *frequent itemset* menggunakan algoritma Apriori [4].

## 4. Penerapan Algoritma Apriori

Algoritma Apriori bekerja secara iteratif untuk menemukan *frequent itemset* dan membangkitkan aturan asosiasi [9, 10]. Pada setiap iterasi, kandidat itemset dibentuk dari kombinasi yang lebih kecil menuju yang lebih besar, dan itemset yang tidak memenuhi nilai *minimum support* dieliminasi melalui mekanisme *pruning*. Pendekatan ini memungkinkan algoritma untuk hanya memproses kombinasi item yang relevan pada iterasi berikutnya [12, 13]. Metrik evaluasi yang digunakan — *support*, *confidence*, dan *lift* — telah diuraikan secara lengkap beserta rumus matematis pada sub-bab 2.4.

Dalam penelitian ini dilakukan eksplorasi terhadap beberapa kandidat nilai *minimum support*, yaitu 0,001 (0,1%), 0,002 (0,2%), 0,003 (0,3%), 0,005 (0,5%), dan 0,01 (1%). Pengujian beberapa nilai parameter dilakukan untuk memperoleh konfigurasi yang paling sesuai dengan karakteristik dataset serta mampu menghasilkan aturan asosiasi yang relevan. Selain parameter *minimum support*, penelitian ini juga menerapkan *minimum confidence* sebesar 30% dan *minimum lift* sebesar 1,0 sebagai kriteria pembentukan aturan asosiasi. Penentuan parameter akhir dilakukan berdasarkan hasil evaluasi yang diperoleh pada tahap pemodelan dan dibahas pada hasil dan pembahasan.

## 5. Filter Redundansi dan Perancangan Dashboard

Setelah aturan asosiasi berhasil dibangkitkan, diterapkan prosedur filter redundansi guna mengeliminasi aturan-aturan simetris. Secara teknis, aturan  $A \rightarrow B$  dan  $B \rightarrow A$  mengandung informasi pasangan item yang identik, namun arah kausalitasnya berbeda. Dari perspektif rekomendasi bisnis, mempertahankan kedua aturan tersebut secara bersamaan tidak memberikan nilai tambah dan berpotensi menimbulkan ambiguitas dalam interpretasi [2]. Oleh karena itu, untuk setiap pasangan item, hanya satu aturan unik yang dipertahankan berdasarkan nilai *confidence* tertinggi.

Selanjutnya, seluruh hasil pemodelan diimplementasikan ke dalam *dashboard* interaktif berbasis Streamlit. Platform ini dipilih karena kemampuannya menghasilkan antarmuka berbasis web yang dapat dioperasikan tanpa keahlian pemrograman khusus [7]. Dashboard yang dikembangkan dilengkapi dengan panel filter dinamis pada *sidebar* untuk pengaturan parameter *support*, *confidence*, dan *lift* secara real-time, fitur pencarian rekomendasi produk berdasarkan *antecedent*, serta visualisasi aturan asosiasi dalam bentuk *bar chart*, *scatter plot*, dan *heatmap*.

## HASIL DAN PEMBAHASAN

### 1. Statistik Deskriptif Dataset

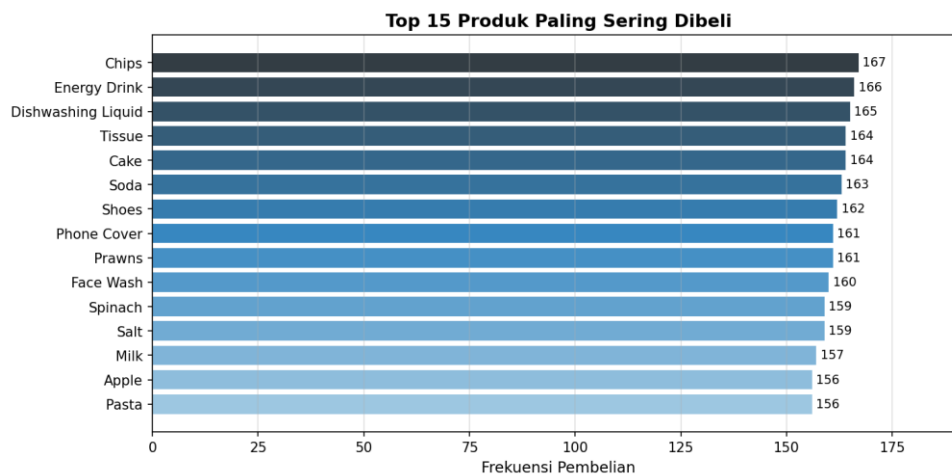
Berdasarkan hasil eksplorasi data yang telah dilakukan pada tahap preprocessing, diperoleh ringkasan statistik dataset yang disajikan pada Tabel 1. Dataset yang digunakan dalam penelitian ini terdiri dari 10.000 baris data transaksi yang merepresentasikan 1.991 transaksi unik dengan 68 produk yang terdistribusi ke dalam 17 kategori produk. Periode transaksi mencakup rentang waktu satu tahun penuh, yaitu dari tanggal 1 Januari 2024 hingga 31 Desember 2024.

Table 1 Ringkasan Statistik Dataset

| Metrik                       | Nilai                     |
|------------------------------|---------------------------|
| Total Transaksi              | 1.991                     |
| Total Produk Unik            | 68                        |
| Total Kategori Produk        | 17                        |
| Rentang Tanggal              | 2024-01-01 s/d 2024-12-31 |
| Minimum Item per Transaksi   | 2                         |
| Rata-rata Item per Transaksi | 5,02                      |
| Maksimum Item per Transaksi  | 7                         |

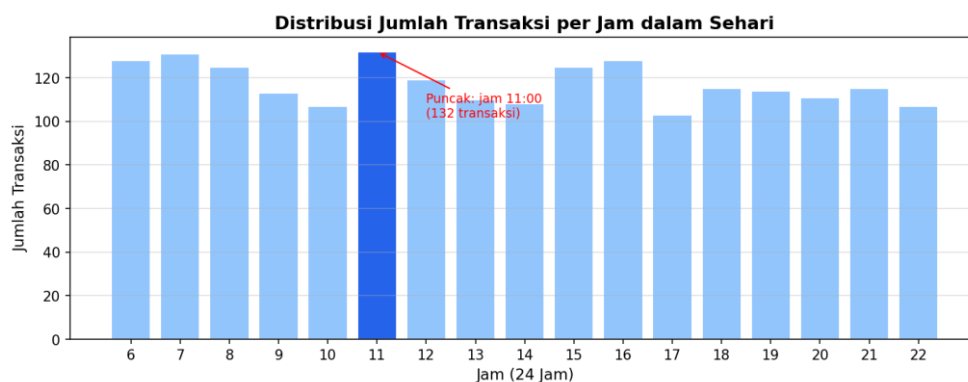
Berdasarkan Tabel 1, diketahui bahwa rata-rata jumlah item per transaksi adalah 5,02 item dengan rentang antara 2 hingga 7 item per transaksi. Kondisi ini mengindikasikan bahwa dataset memiliki karakteristik *sparse* (tersebar), di mana kombinasi item yang sama cenderung jarang muncul secara berulang. Karakteristik tersebut menunjukkan bahwa kombinasi item yang sama relatif jarang muncul secara berulang pada transaksi yang berbeda. Kondisi ini umum ditemukan pada dataset ritel yang memiliki variasi produk cukup tinggi.

Distribusi frekuensi pembelian produk ditampilkan pada Gambar 1. Produk dengan frekuensi pembelian tertinggi adalah *Dishwashing Liquid* dengan 8,19% dari total transaksi, diikuti oleh *Chips* (8,09%), *Cake* (7,94%), *Energy Drink* (7,94%), dan *Soda* (7,94%). Hasil ini menunjukkan bahwa beberapa produk rumah tangga, makanan ringan, dan minuman memiliki frekuensi pembelian yang relatif lebih tinggi dibandingkan produk lainnya dalam dataset.



Gambar 1 Top 15 Produk Paling Sering Dibeli

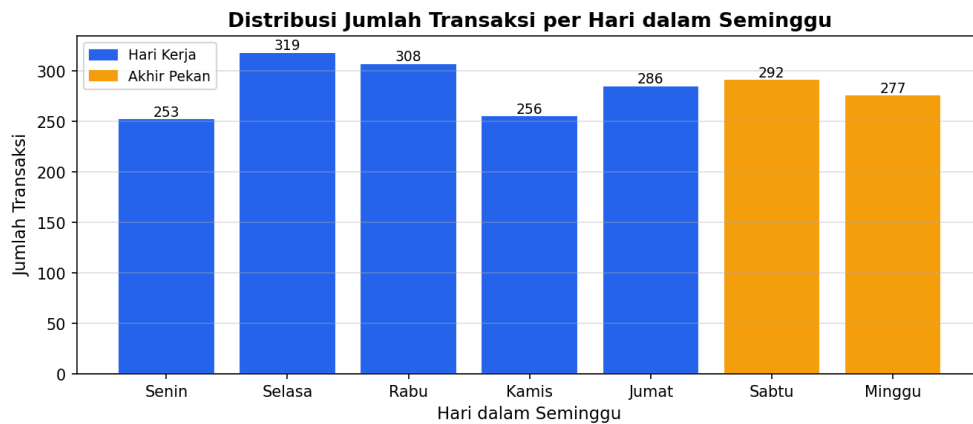
Analisis distribusi waktu transaksi menunjukkan pola perilaku belanja konsumen berdasarkan jam dan hari. Gambar 2 menyajikan distribusi transaksi per jam dalam sehari. Puncak transaksi terjadi pada pukul 15.00-16.00 dengan 130 transaksi, diikuti oleh pukul 12.00-13.00 dan 19.00-20.00. Temuan ini menunjukkan bahwa aktivitas transaksi relatif lebih tinggi pada rentang waktu siang hingga sore hari dibandingkan periode waktu lainnya.



Gambar 2 Distribusi Jumlah Transaksi per Jam dalam Sehari

Gambar 3 menampilkan distribusi transaksi berdasarkan hari dalam seminggu. Hari Sabtu mencatatkan jumlah transaksi tertinggi (328 transaksi), diikuti oleh Jumat (316 transaksi) dan Kamis (290 transaksi). Temuan ini menunjukkan bahwa aktivitas transaksi meningkat

menjelang akhir pekan, dengan jumlah transaksi tertinggi tercatat pada hari Sabtu, yang dapat menjadi pertimbangan strategis dalam penjadwalan promosi dan pengelolaan stok barang.



Gambar 3 Distribusi Jumlah Transaksi per Hari dalam Seminggu

## 2. Data Preparation

Tahap Data Preparation dilakukan untuk memastikan kualitas data sebelum diterapkan pada proses pembentukan frequent itemset menggunakan algoritma Apriori. Proses pembersihan data diawali dengan pemeriksaan duplikasi pada kombinasi atribut *order\_id* dan *product\_name*. Hasil pemeriksaan menunjukkan adanya 299 entri duplikat yang kemudian ditangani melalui proses agregasi nilai *quantity*, sehingga setiap kombinasi transaksi dan produk hanya muncul satu kali dalam dataset.

Selain itu, dilakukan pemeriksaan terhadap kelengkapan data untuk memastikan tidak terdapat nilai yang hilang (*missing value*) pada atribut yang digunakan dalam penelitian. Hasil pemeriksaan menunjukkan bahwa seluruh atribut utama telah terisi dengan baik sehingga tidak diperlukan proses imputasi maupun penghapusan data.

Setelah proses pembersihan selesai, data ditransformasikan ke dalam bentuk *transaction basket* yang merepresentasikan kumpulan produk pada setiap transaksi. Selanjutnya dilakukan proses *one-hot encoding* untuk mengubah data transaksi menjadi matriks biner yang dapat diproses oleh algoritma Apriori. Hasil transformasi ini digunakan sebagai input pada tahap pemodelan dan pembentukan aturan asosiasi.

## 3. Parameter Tuning

Sebelum menerapkan algoritma Apriori, dilakukan eksplorasi parameter untuk menentukan nilai minimum *support* yang optimal. Lima nilai minimum *support* diuji, yaitu 0,001 (0,1%), 0,002 (0,2%), 0,003 (0,3%), 0,005 (0,5%), dan 0,01 (1%). Hasil eksplorasi parameter disajikan pada Tabel 2.

Table 2 Hasil Eksplorasi Parameter Minimum Support

| min_support | frequent_itemsets | association_rules |
|-------------|-------------------|-------------------|
| 0,001       | 7.166             | 32.202            |
| 0,002       | 2.370             | 1.974             |
| 0,003       | 2.113             | 1.542             |
| 0,005       | 1.034             | 1.454             |
| 0,010       | 68                | 0                 |

Berdasarkan Tabel 2, diketahui bahwa pada nilai minimum *support* 0,01 (1%) tidak dihasilkan aturan asosiasi sama sekali. Hal ini disebabkan oleh karakteristik dataset yang *sparse* dengan rata-rata hanya 5 item per transaksi. Nilai minimum *support* 0,002 (0,2%) dipilih karena menghasilkan jumlah aturan asosiasi yang cukup (1.974 aturan) tanpa menghasilkan terlalu banyak *noise* seperti pada nilai 0,001 yang menghasilkan 32.202 aturan. Pemilihan ini sejalan dengan penelitian [15] yang menggunakan nilai *support* rendah (0,1%-0,2%) pada dataset ritel dengan keragaman produk tinggi.

#### 4. Frequent Itemset

Penerapan algoritma Apriori dengan nilai minimum *support* 0,002 menghasilkan 2.370 *frequent itemset* yang terdiri dari 68 \*1-itemset\* (produk tunggal), 2.230 \*2-itemset\* (kombinasi 2 produk), dan 72 \*3-itemset\* (kombinasi 3 produk). Tabel 3 menyajikan 10 \*1-itemset\* dengan nilai *support* tertinggi.

Table 3 Top 10 Frequent 1-Itemset (min\_support = 0,002)

| No | Produk             | Support |
|----|--------------------|---------|
| 1  | Dishwashing Liquid | 8,19%   |
| 2  | Chips              | 8,09%   |
| 3  | Cake               | 7,94%   |
| 4  | Energy Drink       | 7,94%   |
| 5  | Soda               | 7,94%   |
| 6  | Prawns             | 7,89%   |
| 7  | Tissue             | 7,89%   |
| 8  | Phone Cover        | 7,84%   |
| 9  | Milk               | 7,79%   |
| 10 | Salt               | 7,73%   |

Tabel 4 menyajikan 10 kombinasi 2 produk (\*2-itemset\*) dengan nilai *support* tertinggi. Kombinasi *Detergent* + *Earphones* dan *Butter* + *Milk* memiliki nilai *support* tertinggi sebesar 0,95%.

Table 4 Top 10 Frequent 2-Itemset (min\_support = 0,002)

| No | Kombinasi Produk      | Support |
|----|-----------------------|---------|
| 1  | Detergent + Earphones | 0,95%   |
| 2  | Butter + Milk         | 0,95%   |
| 3  | Shoes + Spices        | 0,90%   |
| 4  | Energy Drink + Sugar  | 0,90%   |
| 5  | Cooking Oil + Saree   | 0,90%   |
| 6  | Energy Drink + Tomato | 0,90%   |
| 7  | Baby Food + Milk      | 0,90%   |
| 8  | Bread + Pastry        | 0,90%   |
| 9  | Cheese + Chips        | 0,90%   |
| 10 | Energy Drink + Potato | 0,85%   |

Tabel 5 menyajikan 10 kombinasi 3 produk (\*3-itemset\*) dengan nilai *support* tertinggi. Kombinasi *Baby Food* + *Spices* + *Tissue*, *Chicken Breast* + *Coffee* + *Shirt*, dan *Carrot* + *Phone Cover* + *Soap* memiliki nilai *support* tertinggi sebesar 0,251%.

Table 5 Top 10 Frequent 3-Itemset (min\_support = 0,002)

| No | Kombinasi Produk                      | Support |
|----|---------------------------------------|---------|
| 1  | Baby Food + Spices + Tissue           | 0,251%  |
| 2  | Chicken Breast + Coffee + Shirt       | 0,251%  |
| 3  | Carrot + Phone Cover + Soap           | 0,251%  |
| 4  | Apple + Cheese + Chips                | 0,201%  |
| 5  | Beef + Energy Drink + Potato          | 0,201%  |
| 6  | Belt + Fish + Lentils                 | 0,201%  |
| 7  | Apple + Cheese + Salt                 | 0,201%  |
| 8  | Banana + Chips + Spinach              | 0,201%  |
| 9  | Biscuits + Chicken Breast + Face Wash | 0,201%  |
| 10 | Belt + Fish + Tea                     | 0,201%  |

## 5. Association Rules

Setelah *frequent itemset* ditemukan, tahap selanjutnya adalah pembentukan aturan asosiasi (*association rules*). Algoritma Apriori menghasilkan 1.974 aturan asosiasi dengan metrik *lift* sebagai parameter evaluasi. Untuk menghilangkan aturan redundan (aturan simetris seperti  $A \rightarrow B$  dan  $B \rightarrow A$ ), diterapkan filter redundansi yang hanya mempertahankan satu aturan unik untuk setiap pasangan item. Setelah proses filter, diperoleh 138 aturan asosiasi unik dengan nilai *confidence*  $\geq 30\%$ .

Tabel 6 menyajikan 15 aturan asosiasi dengan nilai *lift* tertinggi. Aturan terkuat yang ditemukan adalah *Energy Drink + Fish*  $\rightarrow$  *Potato* dengan nilai *support* 0,201%, *confidence* 66,7%, dan *lift* 9,6886. Nilai *lift* sebesar 9,6886 menunjukkan bahwa peluang ketiga produk dibeli bersamaan adalah 9,7 kali lebih besar dibandingkan dengan pembelian secara acak.

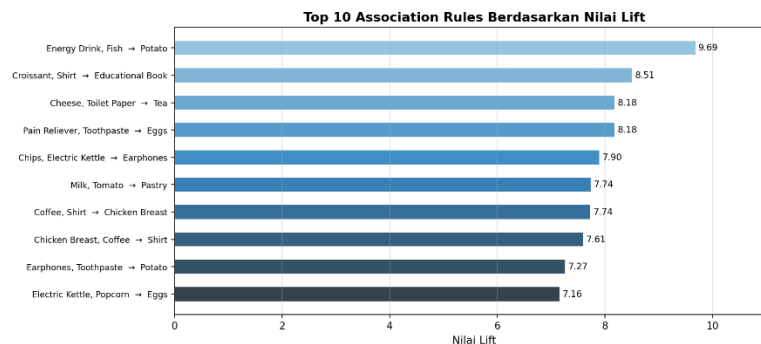
Table 6 15 Association Rules dengan Nilai Lift Tertinggi

| No | Antecedent (Pemicu)       | Consequent (Ikutan) | Support | Confidence | Lift   |
|----|---------------------------|---------------------|---------|------------|--------|
| 1  | Energy Drink, Fish        | Potato              | 0,201%  | 66,7%      | 9,6886 |
| 2  | Croissant, Shirt          | Educational Book    | 0,201%  | 50,0%      | 8,5085 |
| 3  | Cheese, Toilet Paper      | Tea                 | 0,201%  | 57,1%      | 8,1850 |
| 4  | Pain Reliever, Toothpaste | Eggs                | 0,201%  | 57,1%      | 8,1850 |
| 5  | Chips, Electric Kettle    | Earphones           | 0,201%  | 57,1%      | 7,9008 |
| 6  | Milk, Tomato              | Pastry              | 0,201%  | 57,1%      | 7,7396 |
| 7  | Coffee, Shirt             | Chicken Breast      | 0,251%  | 55,6%      | 7,7350 |
| 8  | Chicken Breast, Coffee    | Shirt               | 0,251%  | 45,5%      | 7,6050 |
| 9  | Earphones, Toothpaste     | Potato              | 0,201%  | 50,0%      | 7,2664 |
| 10 | Electric Kettle, Popcorn  | Eggs                | 0,201%  | 50,0%      | 7,1619 |
| 11 | Baby Food, Spices         | Tissue              | 0,251%  | 55,6%      | 7,0453 |
| 12 | Baby Food, Tissue         | Spices              | 0,251%  | 50,0%      | 7,0106 |
| 13 | Energy Drink, Tea         | Spices              | 0,201%  | 50,0%      | 7,0106 |
| 14 | Onion, Soap               | Notebook            | 0,201%  | 50,0%      | 7,0106 |
| 15 | Educational Book, Shirt   | Croissant           | 0,201%  | 50,0%      | 6,9132 |

Berdasarkan hasil evaluasi, aturan asosiasi yang dihasilkan memiliki nilai *confidence* antara 30% hingga 66,7% dengan nilai *lift* seluruh aturan terpilih lebih besar dari 1. Kondisi ini

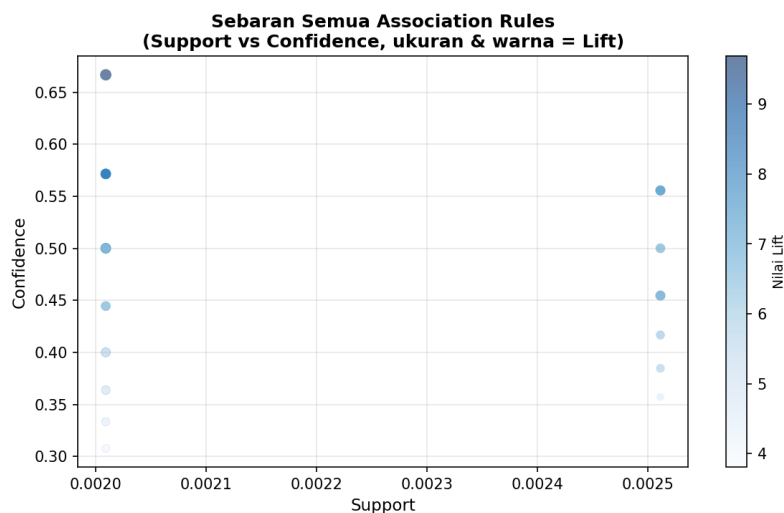
menunjukkan bahwa hubungan antar produk yang ditemukan merupakan asosiasi positif dan tidak terjadi secara kebetulan. Meskipun demikian, beberapa aturan dengan nilai *lift* tinggi memiliki nilai *support* yang relatif rendah sehingga interpretasi hasil perlu mempertimbangkan ketiga metrik, yaitu *support*, *confidence*, dan *lift*, secara bersamaan agar rekomendasi yang dihasilkan tetap relevan secara bisnis.

Visualisasi aturan asosiasi disajikan pada Gambar 4 hingga Gambar 6. Gambar 4 menyajikan *bar chart* 10 aturan asosiasi teratas berdasarkan nilai *lift*. Aturan *Energy Drink + Fish → Potato* menempati posisi tertinggi dengan nilai *lift* mendekati 10, diikuti oleh aturan *Croissant + Shirt → Educational Book* dan *Cheese + Toilet Paper → Tea*.



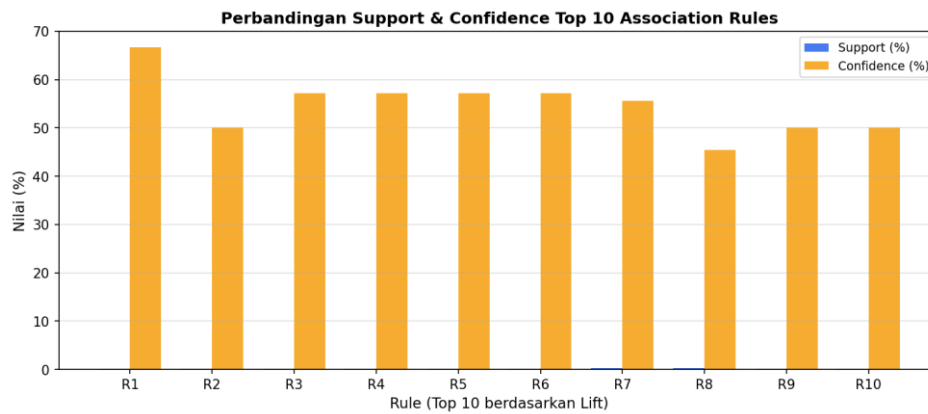
Gambar 4 Bar Chart Top 10 Association Rules Berdasarkan Nilai Lift

Gambar 5 menyajikan *scatter plot* yang memvisualisasikan hubungan antara nilai *support* dan *confidence* dari seluruh aturan asosiasi. Ukuran dan warna titik menunjukkan nilai *lift*, di mana titik yang lebih besar dan lebih gelap mengindikasikan nilai *lift* yang lebih tinggi. Sebaran titik menunjukkan bahwa sebagian besar aturan memiliki nilai *support* yang rendah (0,2%-0,3%) dengan variasi *confidence* yang cukup beragam.



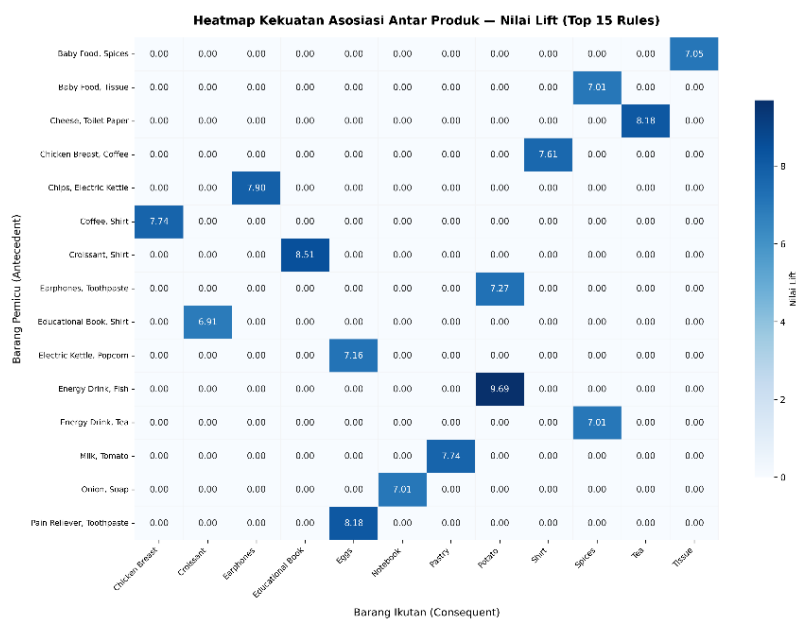
Gambar 5 Scatter Plot Support vs Confidence

Gambar 6 menyajikan perbandingan nilai *support* dan *confidence* untuk 10 aturan asosiasi teratas. Grafik ini menunjukkan bahwa aturan dengan nilai *lift* tertinggi belum tentu memiliki nilai *confidence* tertinggi, sehingga evaluasi aturan asosiasi perlu mempertimbangkan ketiga metrik secara simultan.



Gambar 6 Perbandingan Support & Confidence Top 10 Association Rules

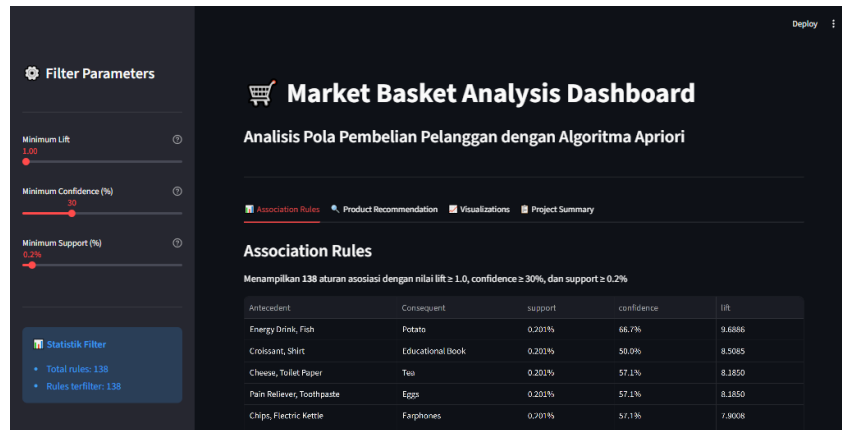
Gambar 7 menyajikan *heatmap* yang memvisualisasikan kekuatan asosiasi antar produk berdasarkan nilai *lift* untuk 15 aturan asosiasi teratas. Sel yang lebih gelap mengindikasikan nilai *lift* yang lebih tinggi, menunjukkan hubungan asosiasi yang lebih kuat antar produk. *Heatmap* ini memudahkan identifikasi produk-produk yang memiliki keterkaitan erat secara visual.



Gambar 7 Heatmap Kekuatan Asosiasi Antar Produk — Nilai Lift (Top 15 Rules)

## 6. Implementasi Dashboard

Hasil analisis Market Basket Analysis diimplementasikan ke dalam dashboard interaktif berbasis Streamlit sebagai tahap Deployment dalam framework CRISP-DM. Implementasi ini bertujuan agar hasil analisis dan aturan asosiasi yang diperoleh dapat dimanfaatkan secara langsung oleh pengguna bisnis tanpa perlu menjalankan proses analisis secara manual. Dashboard dilengkapi panel filter yang memungkinkan pengguna menyesuaikan parameter minimum support, confidence, dan lift secara dinamis.



Gambar 8 Tampilan Dashboard Market Basket Analysis

Dashboard ini memiliki beberapa fitur utama: (1) filter parameter real-time yang langsung mempengaruhi tabel aturan asosiasi, (2) fitur pencarian rekomendasi produk berdasarkan *antecedent*, (3) visualisasi interaktif aturan asosiasi dalam bentuk *bar chart*, *scatter plot*, dan *heatmap*, serta (4) ringkasan proyek dan kesimpulan. Dengan adanya dashboard ini, pemilik toko dapat mengeksplorasi pola pembelian konsumen secara mandiri tanpa harus memahami kode pemrograman.

## 7. Rekomendasi Bisnis

Berdasarkan aturan asosiasi yang ditemukan, khususnya aturan terkuat *Energy Drink + Fish* → *Potato* dengan nilai *lift* 9,6886 dan *confidence* 66,7%, disusun rekomendasi strategis untuk mendukung peningkatan penjualan. Tabel 7 menyajikan *action plan* rekomendasi bisnis yang dapat diimplementasikan oleh pengelola toko.

Table 7 Action Plan Rekomendasi Bisnis

| No | Strategi                              | Implementasi                                                                                            | Dasar Rule                                       | Target KPI                                |
|----|---------------------------------------|---------------------------------------------------------------------------------------------------------|--------------------------------------------------|-------------------------------------------|
| 1  | Bundling Produk (Cross-Selling)       | Buat paket bundling: Energy Drink + Fish + Potato dengan diskon 10-15%                                  | Energy Drink+Fish→Potato (Lift=9,69, Conf=66,7%) | Kenaikan nilai transaksi rata-rata 10-20% |
| 2  | Tata Letak Toko (Cross-Merchandising) | Letakkan rak Energy Drink berdekatan dengan display Fish & Potato                                       | Energy Drink+Fish→Potato (Lift=9,69)             | Kenaikan penjualan Potato 15-25%          |
| 3  | Promosi Otomatis via POS              | Program sistem kasir (POS) agar otomatis menawarkan Potato saat pelanggan membeli Energy Drink dan Fish | Energy Drink+Fish→Potato (Conf=66,7%)            | Konversi upsell 30-40%                    |
| 4  | Targeted Recommendation               | Gunakan aturan asosiasi sebagai mesin rekomendasi di aplikasi belanja                                   | Seluruh aturan asosiasi                          | Peningkatan item per transaksi 1.974      |

Selain aturan dengan nilai lift tertinggi, pengelola toko juga dapat memanfaatkan aturan asosiasi lain yang memiliki nilai support dan confidence yang memadai untuk mendukung strategi cross-selling, bundling produk, maupun penataan tata letak toko. Dengan demikian, keputusan bisnis tidak hanya bergantung pada satu pola pembelian, tetapi mempertimbangkan keseluruhan pola asosiasi yang ditemukan selama proses analisis.

## KESIMPULAN

Penelitian ini berhasil menerapkan metode Market Basket Analysis menggunakan algoritma Apriori pada 1.991 data transaksi ritel untuk menemukan pola pembelian konsumen dan menghasilkan aturan asosiasi antarproduk. Hasil pengujian menunjukkan bahwa nilai minimum support sebesar 0,2% merupakan parameter yang paling optimal untuk karakteristik data yang bersifat sparse. Setelah dilakukan proses evaluasi dan penyaringan aturan redundan, diperoleh aturan asosiasi terkuat yaitu Energy Drink dan Fish → Potato dengan nilai confidence sebesar 66,7% dan lift sebesar 9,69, yang menunjukkan adanya hubungan pembelian yang kuat antarproduk tersebut. Selain menghasilkan aturan asosiasi, penelitian ini juga berhasil mengimplementasikan hasil analisis ke dalam dashboard interaktif berbasis Streamlit yang dapat menampilkan informasi aturan asosiasi dan rekomendasi produk secara dinamis untuk mendukung pengambilan keputusan bisnis. Untuk penelitian selanjutnya, analisis dapat dikembangkan dengan membandingkan kinerja algoritma Apriori dengan algoritma frequent pattern lainnya seperti FP-Growth dan ECLAT, menggunakan dataset dengan periode transaksi yang lebih panjang, serta menambahkan visualisasi dan fitur analisis yang lebih komprehensif guna meningkatkan kualitas rekomendasi yang dihasilkan.

## DAFTAR PUSTAKA

- [1]. A. Hermawan, B. Wicaksono, T. Ahmadjayadi, B. S. Prakasa, and J. D. Aruan, "Implementasi Algoritma Apriori pada Market Basket Analysis terhadap Data Penjualan Produk Supermarket," *ALGORITMA: Jurnal Matematika, Ilmu Pengetahuan Alam, Kebumihan dan Angkasa*, vol. 2, no. 5, pp. 95–105, Sep. 2024.
- [2]. K. Brighton and S. Hariyanto, "Penerapan Metode Market Basket Analisis Dengan Algoritma Apriori Pada Toko Ritel Elektronik," *bit-Tech*, vol. 7, no. 1, pp. 37–46, 2024.
- [3]. W. Nugraheni and A. Nugroho, "Penerapan Metode Market Basket Analysis (MBA) dengan Algoritma Apriori Untuk Menganalisis Pembelian Jajanan Khas Lebaran Pada Warung Sembako di Toko Win," *Jurnal JTIK (Jurnal Teknologi Informasi dan Komunikasi)*, vol. 7, no. 4, pp. 636–643, Oct. 2023.
- [4]. N. Purwati and S. Karnila, "Strategi Peningkatan Penjualan Produk Menggunakan Market Basket Analysis," *Jurnal Sistem Informasi Bisnis*, vol. 13, no. 2, pp. 96–103, 2023.
- [5]. A. N. Ridho, A. P. A. Masa, and P. P. Widagdo, "Implementasi Market Basket Analysis pada Data Penjualan CV. XYZ Menggunakan Algoritma FP-Growth," *JITET (Jurnal Informatika dan Teknik Elektro Terapan)*, vol. 12, no. 3, pp. 1985–1994, 2024.
- [6]. A. A. Widodo and I. A. Fachrudin, "Prediksi Itemset Promosi Penjualan Menggunakan Metode Market Basket Analysis," *KONVERGENSI*, vol. 18, no. 1, pp. 18–24, Jul. 2022.
- [7]. E. Umar, D. Manongga, and A. Iriani, "Market Basket Analysis Menggunakan Association Rule dan Algoritma Apriori Pada Produk Penjualan Mitra Swalayan Salatiga," *Jurnal Media Informatika Budidarma*, vol. 6, no. 3, pp. 1367–1377, Jul. 2022.
- [8]. N. Ramadhani, A. W. Syahroni, A. Supikar, and W. Zumann, "Penerapan Market Basket Analysis Menggunakan Metode Multilevel Association Rules dan Algoritma ML\_T2L1 Pada Data Order PT. Unirama," *InfoTekJar: Jurnal Nasional Informatika dan Teknologi Jaringan*, vol. 4, no. 2, pp. 261–274, Mar. 2020.

- [9]. E. D. Putra, M. H. Rifqo, and D. Hardianto, "Implementasi Algoritma Apriori Pada Market Basket Analysis (MBA) Aksesoris Telepon Seluler," *Jurnal Komputer, Informasi dan Teknologi (J-KOMITEK)*, vol. 2, no. 2, 2022.
- [10]. R. Fadillah, L. Qadriah, and M. Rizal, "Market Basket Analisis Data Mining Untuk Mengetahui Pola Penjualan Pada Cerry Mart Beureunueun Menggunakan Algoritma Apriori," *Jurnal Ilmiah Teknologi Informasi dan Komunikasi*, vol. 1, no. 1, pp. 1–6, 2023.
- [11]. S. R. Bagaskara and D. H. Bangkalang, "Analisis dan Implementasi Market Basket Analysis (MBA) Menggunakan Algoritma Apriori dengan Dukungan Visualisasi Data," *Jurnal Sistem Komputer dan Informatika (JSON)*, vol. 4, no. 4, pp. 612–620, Jun. 2023.
- [12]. S. G. Setyorini, E. K. Sari, L. R. Elita, and S. A. Putri, "Market Basket Analysis with K-Means and FP-Growth Algorithm as Citra Mustika Pandawa Company," *Journal of Applied Informatics and Computing*, vol. 5, no. 1, pp. 1–8, 2021.
- [13]. A. Lewis, M. Zarlis, and Z. Situmorang, "Penerapan Data Mining Menggunakan Task Market Basket Analysis Pada Transaksi Penjualan Barang di AB Mart dengan Algoritma Apriori," *Jurnal Media Informatika Budidarma*, vol. 5, no. 2, pp. 676–681, Apr. 2021.
- [14]. Ismasari, M. Ramadhan, and W. Hadikristanto, "Analisis Tingkat Pembelian Konsumen dengan Algoritma Apriori," *Jurnal Teknologi Informasi dan Komunikasi*, ol. 8, no. 1, pp. 75–84, 2020.
- [15]. R. Fatahillah and A. Wibowo, "Analisis Data Transaksi (Market Basket Analysis) Penjualan Vape Menggunakan Algoritma Apriori Berbasis Website," in *Prosiding Seminar Nasional Mahasiswa Fakultas Teknologi Informasi (SENAFTI)*, Jakarta, Indonesia, Aug. 2023, vol. 2, no. 2, pp. 901–908.
- [16]. A. Thariq, "Implementasi Market Basket Analysis Menggunakan Algoritma Apriori pada Data Penjualan Buku," *Jurnal Kolaboratif Sains (JKS)*, vol. 6, no. 3, pp. 154–163, Mar. 2023.
- [17]. A. C. Simanjuntak, M. E. Sitanggang, M. I. Cahyani, and N. Syahputri, "Penerapan Metode Data Mining Market Basket Analysis Terhadap Data Penjualan Produk Pada Toko Iblite Luxury Menggunakan Algoritma Apriori," *BRIDGE: Jurnal Ilmiah Teknik*, vol. 2, no. 3, pp. 62–74, Aug. 2024.
- [18]. N. M. Mustofa, A. M. Alfari, and A. Tholib, "Penerapan Algoritma Apriori Untuk Menentukan Pola Pembelian Konsumen," *JSITIK: Jurnal Sistem Informasi dan Teknologi Informasi Komputer*, vol. 4, no. 1, pp. 11–30, Des. 2025.
- [19]. D. L. Safitry, N. Rosianti, E. Divyaning, H. Zidan, Z. J. Arnecia, I. Paryudi, I. Veritawati, and S. R. C. Nursari, "Analisis Pola Pembelian Konsumen Menggunakan Algoritma Apriori untuk Menentukan Strategi Pemasaran Produk di Toko Retail X," *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 9, no. 1, pp. 505–511, Feb. 2025.
- [20]. D. Rizaldi and A. Adnan, "Market Basket Analysis Menggunakan Algoritma Apriori: Kasus Transaksi 212 Mart Soebrantas Pekanbaru," *Jurnal Statistika dan Aplikasinya*, vol. 5, no. 1, pp. 31–40, Jun. 2021.
- [21]. A. N. K. S. Workspace, "Retail POS Dataset for Market Basket Analysis," *Kaggle*, 2025. [Online]. Available: <https://www.kaggle.com/datasets/arunworkspace/retail-pos-dataset-for-market-basket-analysis>