

Comparative Analysis of Association Rule Mining Algorithms Using Apriori and Eclat on Retail Transaction Data

Keisya Aqila Putri Kausara^{1*}, MY Teguh Sulistyono^{2*}

* Sistem Informasi, Universitas Dian Nuswantoro
112202206824@mhs.ac.id¹, teguh.sulistyono@mhs.dinus.ac.id²

Article Info

Article history:

Received 2025-12-28

Revised 2026-02-23

Accepted 2026-04-08

Keyword:

*Apriori Algorithm,
Association Rule Mining,
Eclat Algorithm,
Product Recommendation.*

ABSTRACT

Retail transaction data can be utilized to identify consumer purchasing patterns and product relationships. This study evaluates and compares the performance of the Apriori and ECLAT algorithms in generating association rules from a retail dataset comprising 30,000 transactions collected between January and April 2025. A grid search method was employed to determine the optimal minimum support and minimum confidence thresholds. Using a minimum support of 0.005 and a minimum confidence of 0.3, both algorithms generated 1,736 frequent itemsets and 78 association rules. The resulting rules were assessed using support, confidence, and lift measures, and all rules obtained lift values greater than 1, indicating meaningful positive dependencies among products. The strongest rule revealed that customers who purchased Bread and Milk were highly likely to also purchase Cereal, with a confidence of 60.5% and a lift value of 2.70. Although both algorithms produced comparable rule quality, ECLAT demonstrated superior computational efficiency due to its vertical data representation approach. These findings provide practical insights for retail decision-making, particularly in cross-selling strategies, product bundling, and shelf arrangement optimization.



This is an open access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.

I. PENDAHULUAN

Sektor ritel merupakan bidang usaha yang berinteraksi langsung dengan konsumen akhir dan secara berkelanjutan menghasilkan data transaksi dalam jumlah besar akibat penggunaan sistem penjualan terkomputerisasi. Setiap transaksi menyimpan informasi penting terkait perilaku belanja pelanggan, seperti ragam produk yang dibeli serta pola kombinasi item dalam satu keranjang belanja[1]. Namun, dalam praktiknya, data transaksi tersebut umumnya hanya dimanfaatkan sebagai laporan penjualan periodik tanpa dilakukan analisis lebih lanjut untuk mengidentifikasi pola pembelian konsumen. Padahal, pengolahan dan analisis data transaksi secara sistematis dapat memberikan wawasan strategis untuk mendukung penyusunan strategi penjualan, pengendalian persediaan, serta peningkatan kualitas layanan pelanggan secara lebih tepat[2].

Pelaku usaha ritel masih menghadapi tantangan dalam memahami keterkaitan antarproduk yang sering dibeli secara bersamaan. Akibatnya, keputusan terkait penataan produk, promosi, dan rekomendasi barang sering dilakukan secara

subjektif dan kurang berbasis data. Selain itu, meningkatnya volume dan keragaman transaksi menyebabkan pendekatan analisis konvensional menjadi kurang efektif dalam mengungkap pola pembelian secara menyeluruh. Oleh karena itu, diperlukan metode analisis berbasis data mining yang mampu mengidentifikasi pola tersembunyi dalam data transaksi guna mendukung pengambilan keputusan bisnis yang lebih akurat[3].

Market Basket Analysis (MBA) merupakan salah satu pendekatan yang umum digunakan untuk mengidentifikasi hubungan antarproduk melalui pembentukan aturan asosiasi[4]. Informasi yang dihasilkan dari MBA dapat dimanfaatkan untuk mendukung strategi *bundling* produk, penataan rak, serta perencanaan persediaan barang[5]. Dalam implementasinya, berbagai algoritma *Association Rule Mining* telah digunakan, di antaranya *Apriori*, *Equivalence Class Transformation* (ECLAT), dan *FP-Growth*.

Penelitian Wardhana (2024) menunjukkan bahwa algoritma ECLAT efektif dalam menemukan *frequent itemset* melalui representasi data vertikal sehingga tidak

memerlukan pemindaian basis data secara berulang [6]. Husain et al (2023) melaporkan bahwa ECLAT mampu menghasilkan aturan asosiasi yang setara dengan Apriori dan FP-Growth, meskipun performanya bergantung pada karakteristik dataset dan parameter yang digunakan [7]. Di sisi lain, algoritma Apriori dikenal mudah diinterpretasikan, namun memiliki keterbatasan efisiensi akibat proses pembangkitan kandidat *itemset* secara iteratif yang meningkatkan waktu komputasi pada dataset berskala besar [8][9]. Sementara itu, FP-Growth dikembangkan untuk mengatasi kelemahan Apriori dengan menghilangkan proses *candidate generation* melalui pembentukan struktur FP-tree, sehingga lebih efisien dalam pengolahan data berukuran besar.

Meskipun berbagai penelitian telah menerapkan algoritma *Association Rule Mining*, sebagian besar masih berfokus pada satu metode atau menggunakan dataset berskala terbatas. Studi yang secara simultan membandingkan kualitas aturan asosiasi dan kinerja komputasi pada dataset ritel berskala besar masih relatif sedikit, serta belum banyak yang mengintegrasikan metrik *support*, *confidence*, dan *lift* secara komprehensif.

Berdasarkan hal tersebut, penelitian ini bertujuan membandingkan algoritma Apriori dan ECLAT dalam mengidentifikasi pola keterkaitan produk pada data transaksi ritel. Kontribusi penelitian ini terletak pada evaluasi terintegrasi menggunakan *support*, *confidence*, dan *lift* serta analisis kinerja komputasi pada dataset berskala besar untuk memberikan gambaran efektivitas kedua algoritma dalam mendukung strategi penjualan ritel.

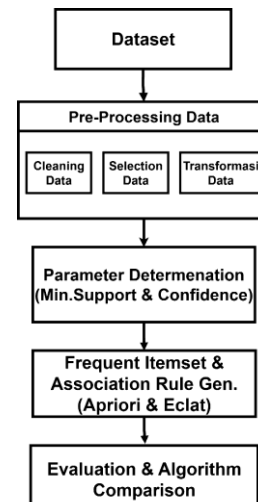
II. METODE

Penelitian ini menganalisis data transaksi ritel menggunakan pendekatan *Association Rule Mining* dengan menerapkan dan membandingkan algoritma Apriori dan ECLAT. Tahapan penelitian meliputi pengumpulan dataset, proses *pre-processing*, penentuan parameter minimum *support* dan *confidence*, pembentukan *frequent itemset* dan aturan asosiasi, penyaringan aturan, evaluasi kualitas aturan, serta analisis perbandingan kinerja algoritma untuk menghasilkan rekomendasi produk.

A. Dataset

Dataset yang digunakan dalam penelitian ini merupakan data transaksi ritel sekunder yang diperoleh dari platform Kaggle[10]. Dataset tersebut terdiri atas 30.000 transaksi dengan empat atribut utama, yaitu *TransactionID*, *CustomerID*, *Products*, dan *Timestamp*. Atribut *TransactionID* merepresentasikan identitas unik setiap transaksi, sedangkan *CustomerID* menunjukkan identitas pelanggan yang melakukan pembelian. Atribut *Products* memuat daftar produk yang dibeli dalam satu transaksi dan dituliskan dalam satu kolom dengan pemisah berupa tanda koma. Sementara itu, atribut *Timestamp* mencatat waktu terjadinya transaksi. Dataset ini mencakup transaksi pada

periode Januari sampai April 2025 dengan variasi produk yang beragam dalam setiap pembelian. Struktur transaksi yang memuat lebih dari satu *item* dalam satu keranjang belanja menjadikan dataset ini sesuai untuk penerapan *Market Basket Analysis* dan *Association Rule Mining* menggunakan algoritma Apriori dan ECLAT guna mengidentifikasi pola keterkaitan antarproduk.



Gambar 1. Metode Penelitian

B. Tahap Pre-Processing

Menurut Bhargavi(2022), *pre-processing* merupakan tahap awal yang bertujuan menyiapkan data mentah agar siap dianalisis dengan cara menangani kesalahan pencatatan, ketidakkonsistenan, serta nilai yang hilang sehingga kualitas data meningkat [11]. Pada penelitian ini, *pre-processing* dilakukan melalui tiga tahap utama, yaitu *cleaning*, *selection*, dan *transformasi* data, sehingga penerapan algoritma Apriori dan ECLAT dapat dilakukan secara objektif dan hasilnya dapat dibandingkan secara konsisten.

a. Cleaning Data

Tahap *cleaning* dilakukan untuk memastikan data berada dalam kondisi yang layak dan konsisten sebelum dianalisis [12]. Proses ini mencakup pemeriksaan dan penghapusan data duplikasi pada *TransactionID*, penyeragaman penulisan nama produk untuk menghindari perbedaan representasi *item*, serta pengecekan kelengkapan data pada atribut utama. Berdasarkan hasil pemeriksaan, seluruh atribut yang digunakan dalam analisis berada dalam kondisi lengkap tanpa nilai kosong yang signifikan. Dengan proses pembersihan ini, kualitas dataset meningkat sehingga pembentukan *frequent itemset* oleh algoritma Apriori dan ECLAT dapat dilakukan secara lebih akurat dan konsisten.

b. Selection Data

Tahap *selection* bertujuan memilih atribut yang relevan dengan analisis pola asosiasi sehingga data yang digunakan lebih terfokus dan efisien[13]. Pada penelitian ini, atribut yang digunakan dalam proses pembentukan aturan asosiasi adalah *TransactionID* dan *Products*, karena keduanya secara

langsung merepresentasikan kombinasi *item* dalam setiap transaksi. Sementara itu, atribut *CustomerID* dan *Timestamp* tidak dilibatkan dalam proses pembentukan *frequent itemset* karena tidak berpengaruh langsung terhadap analisis keterkaitan antarproduk. Pemilihan atribut ini dilakukan untuk menyederhanakan struktur data serta meningkatkan efisiensi proses komputasi.

c. Transformasi Data

Tahap transformasi data bertujuan mengubah struktur data transaksi ke dalam format yang sesuai untuk proses *Association Rule Mining*[14]. Pada tahap ini, daftar produk yang tercantum dalam satu kolom dipisahkan berdasarkan pemisah koma sehingga setiap transaksi direpresentasikan sebagai kumpulan *item* individual. Selanjutnya, data dikonversi ke dalam format transaksi (*basket format*) agar dapat diproses oleh algoritma Apriori dan ECLAT dalam pembentukan *frequent itemset* dan aturan asosiasi. Proses transformasi ini memastikan setiap *item* memiliki representasi yang konsisten dan siap digunakan dalam analisis pola keterkaitan antarproduk.

C. Parameter Determenation (Minimum Support & Confidence)

Penentuan nilai minimum support dan minimum confidence dilakukan untuk menyaring *frequent itemset* dan aturan asosiasi yang signifikan[15]. Pada penelitian ini, enentuan parameter tidak dilakukan secara langsung, melainkan melalui pengujian beberapa kombinasi nilai menggunakan pendekatan grid search. Nilai minimum support diuji pada rentang 0,005 hingga 0,03, sedangkan minimum confidence diuji pada rentang 0,2 hingga 0,6 untuk memperoleh konfigurasi yang menghasilkan aturan asosiasi yang optimal.

Berdasarkan hasil pengujian, nilai minimum support sebesar 0,005 dan minimum confidence sebesar 0,3 dipilih karena mampu menghasilkan jumlah aturan yang terkontrol serta tetap mempertahankan kualitas asosiasi yang signifikan berdasarkan evaluasi lift. Pemilihan konfigurasi tersebut bertujuan untuk menjaga keseimbangan antara jumlah aturan yang dihasilkan dan relevansi pola yang terbentuk. Nilai parameter yang terlalu rendah cenderung menghasilkan aturan dalam jumlah berlebihan dengan tingkat signifikansi yang lemah, sedangkan nilai yang terlalu tinggi berpotensi menghilangkan pola penting dalam dataset.

a) Support

$$Support(X \rightarrow Y) = \left(\frac{\text{Jumlah transaksi yang mengandung } (X \cup Y)}{\text{Jumlah total transaksi}} \right) \quad (1)$$

b) Confidence

$$Confidence(X \rightarrow Y) = \left(\frac{Support(X \cup Y)}{Support(X)} \right) \quad (2)$$

Keterangan:

- a) X = *item antecedent* (syarat)

- b) Y = *item consequent* (hasil)

- c) $X \cup Y$ = gabungan *item X* dan *Y*

- d) Penyebut = jumlah keseluruhan transaksi dalam dataset

D. Frequent Itemset and Association Rule Generation (Apriori and Eclat)

Setelah nilai *minimum support* dan *confidence* ditetapkan, tahap berikutnya adalah menghasilkan *frequent itemset* dan membentuk aturan asosiasi dengan menerapkan algoritma Apriori dan ECLAT. Algoritma Apriori bekerja dengan menghasilkan kandidat *itemset* secara bertahap berdasarkan prinsip *frequent itemset*, yaitu hanya kombinasi *item* yang memenuhi nilai minimum support yang dipertahankan. Proses ini dilakukan secara berulang mulai dari 1-*itemset* hingga k-*itemset*, hingga tidak ditemukan lagi kombinasi *item* yang memenuhi nilai ambang yang telah ditentukan[16].

Sebaliknya, algoritma ECLAT menggunakan pendekatan representasi data vertikal dalam bentuk *Transaction ID List* (TID-list). Pencarian *frequent itemset* dilakukan melalui operasi irisan antar TID-list tanpa melakukan pemindaian basis data secara berulang, sehingga secara teoritis lebih efisien dibandingkan metode berbasis kandidat seperti Apriori[17].

Frequent itemset yang terbentuk merupakan kombinasi *item* yang memiliki nilai *support* yang memenuhi ambang *minimum*. Nilai *support* menunjukkan proporsi kemunculan suatu *item* atau kombinasi *item* dalam keseluruhan transaksi dan dirumuskan sebagai berikut[18]:

$$Support(X) = \frac{\text{Jumlah transaksi yang mengandung itemset } X}{\text{Jumlah total transaksi}} \quad (3)$$

Berdasarkan *frequent itemset* tersebut, selanjutnya dibangun aturan asosiasi dengan memperhatikan nilai *confidence*. Nilai *confidence* digunakan untuk mengukur tingkat kepercayaan bahwa *item Y* akan muncul apabila *item X* terdapat dalam suatu transaksi, yang dirumuskan sebagai berikut[18]:

$$Confidence(X \rightarrow Y) = \frac{Support(X \cup Y)}{Support(X)} \quad (4)$$

Keterangan:

- a) X = *item* atau *itemset antecedent*

- b) Y = *item* atau *itemset consequent*

- c) $X \cup Y$ = gabungan *item X* dan *Y*

- d) Penyebut = total jumlah transaksi dalam dataset

E. Evaluation and Algorithm Performance Comparison

Tahap evaluasi dilakukan untuk menilai kualitas aturan asosiasi yang dihasilkan serta membandingkan kinerja algoritma Apriori dan ECLAT. Meskipun aturan telah memenuhi nilai *minimum support* dan *confidence*, diperlukan metrik tambahan untuk memastikan bahwa hubungan antaritem memiliki tingkat ketergantungan yang signifikan. Oleh karena itu, penelitian ini menggunakan metrik *lift* sebagai ukuran utama untuk mengevaluasi kekuatan asosiasi.

Nilai *lift* menunjukkan tingkat peningkatan peluang kemunculan *item Y* ketika *item X* muncul dalam suatu transaksi dibandingkan dengan kondisi independen. Secara matematis, nilai *lift* dirumuskan sebagai berikut:

$$Lift(X \rightarrow Y) = \frac{Confidence(X \rightarrow Y)}{Support(Y)} \quad (5)$$

Keterangan:

$Lift > 1$ → hubungan positif, X meningkatkan peluang Y

$Lift = 1$ → X dan Y bersifat independen

$Lift < 1$ → hubungan negative, X menurunkan peluang Y

Selain kualitas aturan, perbandingan performa algoritma dilakukan berdasarkan jumlah frequent itemset, jumlah aturan asosiasi, serta waktu eksekusi. Pendekatan ini memungkinkan evaluasi yang komprehensif terhadap efektivitas dan efisiensi kedua algoritma dalam mengidentifikasi pola keterkaitan produk.

III. HASIL DAN PEMBAHASAN

A. Dataset

Penelitian ini menggunakan transaksi ritel yang diperoleh melalui platform Kaggle[19]. Dataset tersebut berisi 30.000 transaksi yang terjadi pada periode Januari hingga April 2025. Setiap transaksi memuat informasi berupa *TransactionID* sebagai identitas unik transaksi, *CustomerID* sebagai identitas pelanggan, *Products* yang berisi daftar item yang dibeli dalam satu transaksi, serta *Timestamp* yang menunjukkan waktu terjadinya transaksi. Struktur data ini memungkinkan analisis kombinasi produk dalam satu keranjang belanja, sehingga sesuai untuk penerapan algoritma Apriori dan ECLAT dalam mengidentifikasi pola asosiasi antarproduk.

TABEL 1.
DATASET TRANSAKSI RITEL 2025

Transaction ID	Customer ID	Products	Timestamp
1	C546	Dish, Sponge, Flatbread with Meat, Chips, Orange Juice	2025-02-18
2	C385	Onion, Juice, Flatbread with Meat, Chicken	2025-04-26
3	C292	Egg, Flatbread with Meat, Banana, Pizza	2025-04-25
...
30000	C750	Egg, Apple, Fish, Cereal	2025-01-27

Untuk memperoleh gambaran awal mengenai karakteristik dataset, dilakukan analisis distribusi frekuensi kemunculan setiap produk. Analisis ini bertujuan mengidentifikasi produk dengan tingkat permintaan tertinggi serta memberikan indikasi awal terhadap potensi pembentukan frequent itemset. Produk dengan frekuensi

tinggi memiliki peluang lebih besar untuk memenuhi nilai minimum support dan membentuk aturan asosiasi yang signifikan.

Hasil menunjukkan bahwa sepuluh produk dengan frekuensi tertinggi didominasi oleh kebutuhan konsumsi sehari-hari. Produk Cereal memiliki frekuensi tertinggi sebanyak 6.982 transaksi atau sekitar 23% dari total transaksi, yang menunjukkan nilai support relatif besar. Distribusi frekuensi yang tidak terlalu timpang antarproduk teratas menunjukkan bahwa dataset cukup representatif untuk analisis Association Rule Mining. Tabel 2 menyajikan sepuluh produk dengan frekuensi kemunculan tertinggi pada dataset transaksi ritel.

TABEL 2.
DISTRIBUSI FREKUENSI 10 PRODUK TERLARIS

Peringkat	Produk	Frekuensi Muncul
1	Cereal	6.982
2	Ice Cream	5.038
3	Chicken	4.981
4	Soda	4.974
5	Juice	4.959
6	Cheese	4.958
7	Soap	4.952
8	Beans	4.949
9	Orange	4.945
10	Sausage	4.932

B. Hasil Pre-Processing

Tahap pra-pemrosesan dilakukan untuk memastikan data transaksi ritel berada dalam kondisi yang layak sebelum penerapan algoritma Apriori dan ECLAT. Tahapan ini meliputi *cleaning data*, *selection data*, dan *transformasi data* yang bertujuan menyiapkan data agar rapi, konsisten, dan sesuai untuk proses pembentukan aturan asosiasi pada tahap analisis selanjutnya.

a. Hasil *Cleaning Data*

Hasil proses *cleaning* menunjukkan bahwa dataset berada dalam kondisi baik tanpa *missing value* maupun duplikasi data. Pembersihan difokuskan pada atribut *products* untuk menyeragamkan penulisan *item* melalui penghapusan spasi berlebih, penyesuaian huruf besar/kecil, serta koreksi variasi nama produk yang merujuk pada *item* yang sama, seperti penyeragaman penulisan "Flatbread with Meat" menjadi "Flatbread With Meat". Proses ini memastikan seluruh produk tercatat secara konsisten dan siap digunakan pada tahap analisis berikutnya, sebagaimana ditunjukkan pada perbandingan data sebelum dan sesudah *cleaning*. Gambar 2 dan Tabel 3 menunjukkan perbandingan data sebelum dan sesudah proses *cleaning*.

```
df_clean = df.copy()
df_clean['Products'] = df_clean['Products'].astype(str)

def clean_item_list(product_string):
    items = product_string.split(',')
    cleaned_items = [item.strip().title() for item in items]
    return ", ".join(cleaned_items)

df_clean['Products'] = df_clean['Products'].apply(clean_item_list)
```

Gambar 2. Cleaning Data

TABEL 3.
HASIL PRE-PROCESSING ITEM

No	Item Sebelum Cleaning	Item Setelah Cleaning
1	Dish Sponge, Flatbread with Meat, Chips, Orange...	Dish Sponge, Flatbread With Meat, Chips, Orange...
2	Onion, Juice, Flatbread with Meat, Chicken	Onion, Juice, Flatbread With Meat, Chicken
3	Egg, Flatbread with Meat, Banana, Pizza	Egg, Flatbread With Meat, Banana, Pizza
4	Ice Cream, Soda, Orange, Potato, Cereal, Chocolate...	Ice Cream, Soda, Orange, Potato, Cereal, Chocolate...
5	Ice Cream, Soap, Shampoo, Chicken, Banana, Beans...	Ice Cream, Soap, Shampoo, Chicken, Banana, Beans...

b. Hasil Selection Data

Tahap selection dilakukan untuk mempertahankan atribut yang relevan dengan pembentukan pola asosiasi serta menyederhanakan struktur data. Dari empat atribut awal, yaitu TransactionID, CustomerID, Products, dan Timestamp, hanya TransactionID dan Products yang digunakan karena secara langsung merepresentasikan kombinasi item dalam setiap transaksi.

Atribut CustomerID dan Timestamp tidak dilibatkan karena tidak berpengaruh terhadap analisis keterkaitan antarproduk. Setelah tahap ini, jumlah transaksi tetap 30.000, namun dataset menjadi lebih ringkas dan efisien untuk penerapan algoritma Apriori dan ECLAT. Gambar 3 dan Tabel 4 menunjukkan hasil data setelah tahap selection.

```
df_sel = df_cleaned[['TransactionID', 'Products']].copy()
```

Gambar 3. Selection Data

TABEL 4.
HASIL SELECTION DATA

Transaction ID	Products
1	Dish Sponge, Flatbread With Meat, Chips, Orange...
2	Onion, Juice, Flatbread With Meat, Chicken
3	Egg, Flatbread With Meat, Banana, Pizza

...	...
29998	Apple, Fish, Cereal, Cookie, Pizza, Water, Tom...
29999	Chickpeas, Shampoo, Onion, Yogurt, Cereal, Fla...
30000	Egg, Apple, Fish, Cereal

c. Hasil Transformasi Data

Tahap transformasi bertujuan menyesuaikan struktur data agar kompatibel dengan kebutuhan algoritma Apriori dan ECLAT. Data produk yang semula tersimpan dalam satu kolom dengan pemisah koma diuraikan menjadi item-item individual dan direpresentasikan dalam format transaksi (basket format). Setiap transaksi kemudian disimpan dalam kolom *ProductList* sebagai kumpulan item yang dapat diproses lebih lanjut.

Proses ini memastikan konsistensi struktur data serta mendukung perhitungan support dan pembentukan frequent itemset secara tepat. Hasil transformasi ditunjukkan pada Gambar 4 dan Tabel 5.

```
def split_products(products_str):
    return [p.strip() for p in str(products_str).split(',') if p.strip() != ""]
df_sel['ProductList'] = df_sel['Products'].apply(split_products)
```

Gambar 4. Transformasi Data

TABEL 5.
HASIL TRANSFORMASI DATA

Trans.ID	Products	ProductList
1	Dish Sponge, Flatbread with Meat, Chips, Orange...	['Dish Sponge', 'Flatbread with Meat', 'Chips', 'Orange', ...]
2	Onion, Juice, Flatbread with Meat, Chicken	['Onion', 'Juice', 'Flatbread with Meat', 'Chicken']
3	Egg, Flatbread with Meat, Banana, Pizza	['Egg', 'Flatbread with Meat', 'Banana', 'Pizza']
4	Ice Cream, Soda, Orange, Potato, Cereal, Chocolate...	['Ice Cream', 'Soda', 'Orange', 'Potato', 'Cereal', 'Chocolate']
5	Ice Cream, Soap, Shampoo, Chicken, Banana, Beans...	['Ice Cream', 'Soap', 'Shampoo', 'Chicken', 'Banana', 'Beans']

C. Hasil Parameter Determenation (Minimum Support & Confidence)

Pada tahap ini dilakukan penentuan nilai *minimum support* dan *minimum confidence* sebagai parameter dalam proses pembentukan *frequent itemset* dan aturan asosiasi. Berbeda dengan pendekatan penetapan satu nilai tetap, penelitian ini menggunakan pendekatan metode *grid search* untuk mengevaluasi berbagai kombinasi nilai *support* dan *confidence* guna memperoleh konfigurasi parameter yang menghasilkan aturan asosiasi yang optimal. Sebagaimana ditunjukkan pada Gambar 4, nilai minimum support diuji pada rentang 0,005 hingga 0,03, sedangkan minimum confidence diuji pada rentang 0,2 hingga 0,6. Pada algoritma

ECLAT, parameter maksimum ukuran itemset (max-k) juga divariasikan pada nilai 2, 3, dan 4 untuk mengendalikan kompleksitas kombinasi item.

```

# 3.4 GRID NILAI
# =====
support_grid = [0.005, 0.01, 0.015, 0.02, 0.03]
conf_grid    = [0.2, 0.3, 0.4, 0.5, 0.6]
max_k_grid   = [2, 3, 4] # Eclat saja

MIN_RULES = 50
MAX_RULES = 5000
MIN_LIFT_GT1_RATIO = 0.60 # minimal 60% rules lift>1

```

Gambar 4. Konfigurasi Grid Parameter Penentuan Min.Support dan Confidence

Evaluasi setiap kombinasi parameter dilakukan dengan mempertimbangkan jumlah aturan yang dihasilkan serta proporsi aturan dengan nilai lift lebih dari 1. Untuk menjaga keseimbangan antara kuantitas dan kualitas aturan, ditetapkan kriteria bahwa jumlah aturan berada dalam rentang yang dapat diinterpretasikan (50 hingga 5000 aturan) dan minimal 60% aturan memiliki nilai lift lebih dari 1, sehingga menunjukkan hubungan positif yang signifikan antaritem[20]. Hasil pengujian menunjukkan bahwa kombinasi minimum support sebesar 0,005 dan minimum confidence sebesar 0,3 memberikan keseimbangan terbaik antara jumlah aturan dan kualitas asosiasi berdasarkan nilai lift. Konfigurasi ini kemudian digunakan secara konsisten pada kedua algoritma untuk memastikan perbandingan yang objektif. Rincian parameter terbaik hasil grid search disajikan pada Tabel 5.

TABEL 5.
PARAMETER MINIMUM TERBAIK HASIL GRID SEARCH

Algoritma	Minimum Support	Minimum Confidence
Apriori	0,005	0,3
ECLAT	0,005	0,3

D. Hasil Frequent Itemset dan Association Rule

Penerapan algoritma Apriori dan ECLAT pada dataset transaksi ritel dengan *minimum support* 0,005 dan *minimum confidence* 0,3 menghasilkan 1.736 *frequent itemset* dan 78 aturan asosiasi untuk masing-masing algoritma. Kesamaan jumlah tersebut menunjukkan bahwa kedua algoritma mampu mengidentifikasi kombinasi *item* yang memenuhi ambang parameter secara konsisten meskipun menggunakan pendekatan pencarian yang berbeda.

a. Hasil Frequent Itemset Algoritma Apriori

Berdasarkan konfigurasi parameter terbaik, algoritma Apriori menghasilkan 1.736 *frequent itemset* dan 78 aturan asosiasi dengan waktu eksekusi sebesar 0,92 detik. Jumlah frequent itemset yang dihasilkan menunjukkan bahwa dataset memiliki variasi kombinasi produk yang memenuhi ambang minimum support.

Produk dengan nilai support tertinggi adalah Cereal sebesar 0,2327, yang berarti produk tersebut muncul pada

sekitar 23% dari seluruh transaksi. Nilai ini jauh di atas ambang minimum support yang ditetapkan (0,005), sehingga berkontribusi signifikan dalam pembentukan kombinasi itemset. Produk lain seperti Chicken, Cheese, Beans, Bread, dan Banana juga memiliki nilai support yang relatif tinggi dan berdekatan, yang mengindikasikan pola konsumsi rutin pada kategori kebutuhan sehari-hari.

Dominasi produk dengan tingkat kemunculan tinggi ini menunjukkan adanya stabilitas perilaku pembelian konsumen. Item-item tersebut berpotensi menjadi komponen utama dalam pembentukan aturan asosiasi karena memiliki probabilitas kemunculan yang konsisten dalam berbagai transaksi.

Tabel 6 menyajikan ringkasan hasil penerapan algoritma Apriori, sedangkan Tabel 7 menampilkan sepuluh 1-itemset dengan nilai support tertinggi.

TABEL 6.
HASIL PENERAPAN APRIORI

Keterangan	Nilai
Waktu Eksekusi (detik)	0,92
Jumlah Frequent Itemset	1.736
Jumlah Aturan Asosiasi	78

TABEL 7.
FREQUENT ITEMSET (1-ITEMSET) PADA APRIORI

No	Item	Support
1	Cereal	0,232733
2	Chicken	0,166033
3	Cheese	0,165267
4	Beans	0,164967
5	Bread	0,162333
6	Banana	0,163033
7	Butter	0,162033
8	Chickpeas	0,163700
9	Chips	0,161700
10	Apple	0,160467

b. Hasil Frequent Itemset Algoritma Eclat

Hasil penerapan algoritma *Equivalence Class Transformation* (ECLAT) pada data transaksi ritel menunjukkan bahwa algoritma ini mampu mengidentifikasi pola pembelian konsumen secara efektif. Dengan parameter terbaik yang telah ditentukan sebelumnya, algoritma ECLAT menghasilkan sebanyak 1.736 *frequent itemset* dan 78 aturan asosiasi dengan waktu eksekusi sebesar 0,80 detik. Jumlah itemset dan aturan yang dihasilkan identik dengan Apriori, menunjukkan bahwa kedua algoritma mampu mengidentifikasi kombinasi item yang sama pada konfigurasi parameter yang setara. Distribusi nilai support pada 1-itemset juga menunjukkan pola yang konsisten. Produk Cereal memiliki nilai support tertinggi sebesar 0,2327, yang berarti muncul pada sekitar 23% dari seluruh transaksi dan jauh melampaui ambang minimum support yang ditetapkan. Produk lain seperti Ice Cream, Chicken,

Soda, dan Juice memiliki nilai support yang relatif berdekatan, yang mencerminkan pola konsumsi harian yang stabil dalam dataset.

Kesamaan hasil frequent itemset antara ECLAT dan Apriori menunjukkan bahwa perbedaan utama kedua algoritma tidak terletak pada kualitas itemset yang dihasilkan, melainkan pada mekanisme pencarian dan efisiensi komputasi. Dengan pendekatan representasi data vertikal (TID-list), ECLAT tetap mampu menghasilkan kombinasi itemset yang sama tanpa kehilangan informasi pola yang signifikan.

Tabel 8.
HASIL PENERAPAN ECLAT

Keterangan	Nilai
Waktu Eksekusi (detik)	0,80
Jumlah Frequent Itemset	1.736
Jumlah Aturan Asosiasi	78

Tabel 9.
FREQUENT ITEMSET (1-ITEMSET) PADA ECLAT

No	Item	Support
1	Cereal	0,232733
2	Ice Cream	0,167933
3	Chicken	0,166033
4	Soda	0,165800
5	Juice	0,165300
6	Cheese	0,165267
7	Soap	0,165067
8	Beans	0,164967
9	Orange	0,164833
10	Sausage	0,164400

c. Hasil Pembentukan Association Rule Generation

Berdasarkan penerapan algoritma Apriori dan ECLAT dengan minimum support sebesar 0,005 dan minimum confidence sebesar 0,3, kedua algoritma menghasilkan 78 aturan asosiasi yang memenuhi ambang parameter. Seluruh aturan yang dihasilkan memiliki nilai lift lebih dari 1, yang menunjukkan adanya hubungan positif dan ketergantungan yang bermakna antarproduk.

Pada algoritma Apriori, aturan dengan nilai confidence tertinggi menunjukkan hubungan antara Bread dan Milk terhadap Cereal, dengan confidence sebesar 60,5%. Hal ini mengindikasikan bahwa lebih dari setengah transaksi yang mengandung Bread dan Milk juga mencakup Cereal. Tingginya nilai confidence dan lift pada aturan ini menunjukkan adanya pola pembelian komplementer antara produk-produk tersebut, yang berpotensi dimanfaatkan dalam strategi bundling atau penempatan produk secara berdekatan.

Sementara itu, algoritma ECLAT menghasilkan aturan dengan pola yang serupa, terutama pada keterkaitan antara Cereal dan Milk. Namun, kombinasi antecedent yang dihasilkan cenderung lebih bervariasi, seperti Cookie dan Cereal terhadap Milk, serta Banana dan Cereal terhadap Milk. Variasi ini menunjukkan bahwa hubungan antara Cereal dan Milk tidak terbatas pada satu kombinasi produk

tertentu, melainkan muncul dalam berbagai konteks transaksi.

Secara keseluruhan, kedua algoritma menghasilkan pola asosiasi yang konsisten, dengan dominasi produk konsumsi harian sebagai komponen utama dalam aturan. Kesamaan pola ini menunjukkan bahwa kualitas aturan yang dihasilkan relatif setara, sedangkan perbedaan utama antara kedua algoritma lebih terletak pada mekanisme pencarian dan efisiensi komputasi. Tabel 10 dan Tabel 11 menyajikan contoh aturan asosiasi dengan nilai confidence tertinggi yang dihasilkan oleh masing-masing algoritma.

Tabel 10.
ATURAN ASOSIASI HASIL ALGORITMA APRIORI

No	Antecedent	Consequent	Support (%)	Confidence (%)
1	Bread, Milk	Cereal	1.61%	60.50%
2	Tomato, Milk	Cereal	1.64%	60.47%
3	Orange, Milk	Cereal	1.57%	60.20%
4	Bread, Milk	Cereal	1.61%	60.50%
5	Tomato, Milk	Cereal	1.64%	60.47%

Tabel 11.
ATURAN ASOSIASI HASIL ALGORITMA ECLAT

No	Antecedent	Consequent	Support (%)	Confidence (%)
1	Cookie, Cereal	Milk	1.68%	43.94%
2	Banana, Cereal	Milk	1.61%	43.09%
3	Shampoo, Cereal	Milk	1.60%	42.87%
4	Tomato, Cereal	Milk	1.64%	42.73%
5	Egg, Cereal	Milk	1.58%	42.63%

G. Hasil Evaluation and Algorithm Comparison

Hasil evaluasi menunjukkan bahwa algoritma Apriori dan ECLAT menghasilkan jumlah aturan asosiasi yang sama, yaitu 78 aturan, ketika menggunakan parameter minimum support sebesar 0,005 dan minimum confidence sebesar 0,3. Kesamaan jumlah aturan ini menunjukkan bahwa kedua algoritma memiliki kemampuan yang setara dalam mengidentifikasi pola asosiasi pada dataset yang sama, meskipun menggunakan pendekatan pencarian yang berbeda.

Secara kuantitatif, nilai support rata-rata yang dihasilkan oleh kedua algoritma adalah sebesar 1,73%, dengan rentang antara 1,39% hingga 9,51%. Hal ini menunjukkan bahwa aturan yang terbentuk muncul secara konsisten dalam data transaksi. Nilai confidence rata-rata mencapai 49,32%, dengan nilai maksimum sebesar 60,5%, yang mengindikasikan bahwa aturan asosiasi yang dihasilkan memiliki tingkat kepastian yang cukup baik dalam memprediksi kemunculan item consequent.

Evaluasi berdasarkan metrik lift menunjukkan bahwa seluruh aturan memiliki nilai lift lebih dari 1, dengan rata-rata sebesar 2,49 dan nilai maksimum mencapai 2,70. Nilai lift yang secara konsisten lebih besar dari 1 menunjukkan adanya hubungan positif antaritem, di mana kemunculan antecedent meningkatkan probabilitas kemunculan consequent dibandingkan kondisi independen. Hal ini menegaskan bahwa aturan yang dihasilkan tidak hanya sering muncul, tetapi juga memiliki ketergantungan yang signifikan.

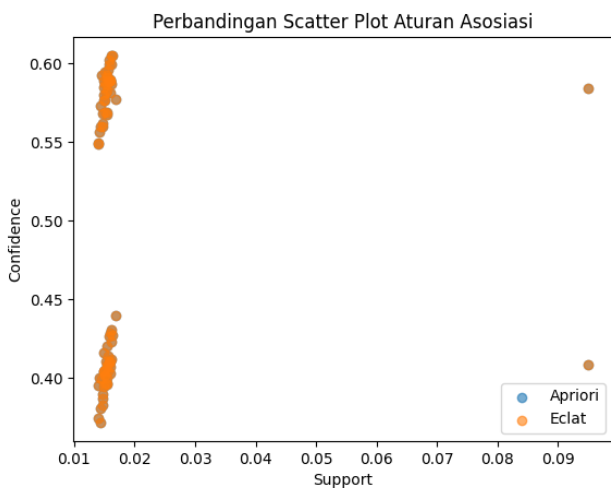
Ringkasan hasil evaluasi kedua algoritma ditampilkan pada Tabel 12, sedangkan distribusi aturan berdasarkan metrik evaluasi divisualisasikan pada Gambar 5.

Selain kualitas aturan, perbandingan efisiensi komputasi juga dilakukan berdasarkan waktu eksekusi. Hasil pengujian menunjukkan bahwa algoritma Apriori memerlukan waktu eksekusi sebesar 0,9245 detik, sedangkan algoritma ECLAT membutuhkan waktu 0,7954 detik. Meskipun selisih waktu relatif kecil, ECLAT menunjukkan kinerja yang lebih cepat dibandingkan Apriori.

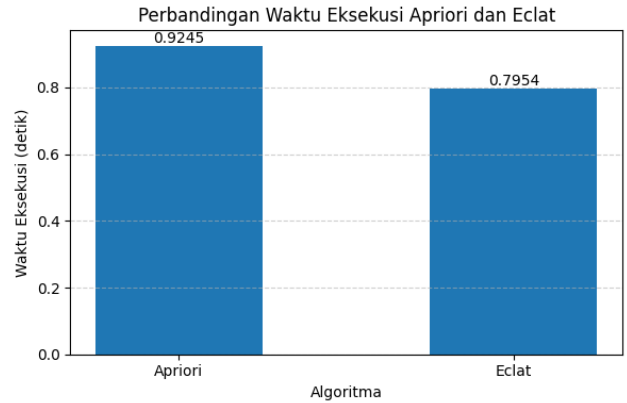
Perbedaan ini disebabkan oleh mekanisme pencarian pola yang digunakan. Apriori melakukan pencarian frequent itemset melalui proses candidate generation secara iteratif, sedangkan ECLAT menggunakan pendekatan representasi data vertikal (TID-list) yang memungkinkan pencarian dilakukan melalui operasi irisan transaksi. Dengan demikian, meskipun kedua algoritma menghasilkan kualitas aturan yang setara, ECLAT menunjukkan keunggulan dari sisi efisiensi komputasi pada dataset yang digunakan dalam penelitian ini.

Tabel 12. Hasil Evaluasi Apriori dan Eclat

Algoritma	Jumlah Rules	Rata-rata Lift	Lift Minimum	Lift Maksimum
Apriori	78	2.4940	2.2841	2.7013
Eclat	78	2.4940	2.2841	2.7013



Gambar 5. Scatter Plot Apriori dan Eclat



Gambar 6. Perbandingan Waktu Eksekusi Apriori vs Eclat

Hasil pembentukan dan evaluasi aturan asosiasi menggunakan algoritma Apriori dan Eclat, diperoleh sejumlah aturan yang memiliki nilai *confidence* dan *lift* yang tinggi, sehingga mencerminkan adanya hubungan yang kuat antarproduk dalam data transaksi. Aturan asosiasi tersebut dapat dimanfaatkan sebagai dasar dalam penyusunan rekomendasi produk guna meningkatkan peluang terjadinya pembelian secara bersamaan (*cross-selling*).

Hasil analisis menunjukkan bahwa produk Milk dan Cereal merupakan item yang paling dominan, baik sebagai *antecedent* maupun *consequent*, dalam aturan asosiasi dengan kualitas terbaik. Salah satu aturan dengan nilai *confidence* tertinggi adalah {Bread, Milk} → {Cereal}, yang memiliki *confidence* sebesar 60,5% dan lift sebesar 270,13%. Temuan ini mengindikasikan bahwa konsumen yang membeli Bread dan Milk memiliki kecenderungan yang tinggi untuk juga membeli Cereal dalam satu transaksi.

Selain itu, beberapa aturan asosiasi lain seperti {Tomato, Milk} → {Cereal} dan {Orange, Milk} → {Cereal} juga menunjukkan pola pembelian yang serupa. Sementara itu, algoritma Eclat menghasilkan aturan seperti {Cookie, Cereal} → {Milk} dan {Banana, Cereal} → {Milk}, yang menegaskan adanya hubungan timbal balik antara produk Cereal dan Milk dengan berbagai produk lainnya. Berdasarkan pola-pola tersebut, produk Milk dan Cereal direkomendasikan untuk ditempatkan secara berdekatan pada rak penjualan atau dikombinasikan dalam bentuk paket promosi bersama produk lain seperti Bread, Tomato, Orange, dan Cookie. Strategi ini diharapkan mampu meningkatkan nilai transaksi serta efektivitas penjualan dengan memanfaatkan pola pembelian konsumen yang telah teridentifikasi.

Tabel 13. REKOMENDASI PRODUK BERDASARKAN ATURAN ASOSIASI

No	Aturan Asosiasi	Pernyataan Jika-Maka	Strategi Rekomendasi
1	Bread, Milk → Cereal	Jika konsumen membeli <i>Bread</i> dan <i>Milk</i> , maka konsumen cenderung membeli <i>Cereal</i> .	Cross-selling

2	Tomato, Milk → Cereal	Jika konsumen membeli <i>Tomato</i> dan <i>Milk</i> , maka <i>Cereal</i> berpotensi ikut dibeli.	Penempatan rak
3	Orange, Milk → Cereal	Jika konsumen membeli <i>Orange</i> dan <i>Milk</i> , maka terdapat peluang tinggi bahwa <i>Cereal</i> juga dibeli.	Paket promosi
4	Cookie, Cereal → Milk	Jika konsumen membeli <i>Cookie</i> dan <i>Cereal</i> , maka konsumen cenderung membeli <i>Milk</i> .	Cross-selling
5	Banana, Cereal → Milk	Jika konsumen membeli <i>Banana</i> dan <i>Cereal</i> , maka <i>Milk</i> berpotensi dibeli sebagai produk pelengkap.	Penempatan rak

IV. KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan, algoritma Apriori dan Eclat mampu mengidentifikasi pola asosiasi pada transaksi ritel secara efektif dengan menggunakan parameter *minimum support* dan *minimum confidence* yang sama. Kedua algoritma menghasilkan jumlah aturan asosiasi yang sama, yaitu sebanyak 78 aturan, serta menunjukkan kualitas aturan yang sebanding berdasarkan nilai *support*, *confidence*, dan *lift*. Seluruh aturan yang dihasilkan memiliki nilai *lift* lebih besar dari 1, yang mengindikasikan adanya hubungan antaritem yang bersifat positif.

Dari aspek efisiensi waktu pemrosesan, algoritma Eclat memiliki kinerja yang lebih baik dibandingkan algoritma Apriori karena mampu menyelesaikan proses pencarian *itemset* dalam waktu yang lebih singkat. Selain itu, hasil analisis menunjukkan bahwa produk Milk dan Cereal memiliki tingkat keterkaitan yang kuat dengan produk lainnya. Pola asosiasi tersebut dimanfaatkan untuk menyusun rekomendasi produk dalam bentuk pernyataan jika-maka yang dapat mendukung strategi *cross-selling*, penataan produk, dan penyusunan paket promosi.

Secara keseluruhan, meskipun kedua algoritma menghasilkan kualitas aturan asosiasi yang setara, algoritma Eclat lebih unggul dari sisi efisiensi waktu, sehingga lebih sesuai untuk digunakan dalam analisis pola pembelian pada transaksi ritel.

DAFTAR PUSTAKA

- I. F. Rahman and D. Riana, "Market Basket Analysis untuk Penjualan Retail: Perbandingan Akurasi Algoritma Apriori dan FP-Growth Berbasis CRISP-DM," *J. Algoritma*, vol. 22, no. 1, pp. 468–479, May 2025, doi: 10.33364/algoritma/v.22-1.2303.
- A. Setiawan, V. Kurniawan, and R. Novita, "Application of the Eclat Algorithm to Find Relationship Patterns Between Items in Sales Transaction Data," 2024.
- I. P. S. Handika and I. K. S. Satwika, "Perbandingan Kinerja Algoritma Apriori Dan Equivalence Class Transformation (Eclat) Dalam Menemukan Pola Pembelian Pada Data Transaksi Minimarket," vol. 9, no. 2, 2024.
- Dr. Y. Arora and Dr. Y. Arora, "Market Basket Analysis using Apriori Algorithm," *Int. J. Innov. Res. Comput. Sci. Technol.*, pp. 62–66, May 2022, doi: 10.55524/ijircst.2022.10.3.12.
- S. Aulia Miranda, F. Fahrullah, and D. Kurniawan, "Implementasi Association Rule Dalam Menganalisis Data Penjualan Sheshop dengan Menggunakan Algoritma Apriori," *METIK J.*, vol. 6, no. 1, pp. 30–36, Jul. 2022, doi: 10.47002/metik.v6i1.342.
- D. B. Wardana, "Analisa Data Mining Menggunakan Algoritma Apriori Dan Algoritma Eclat Di PT Astra International BMW Semarang," vol. 17, no. 1, 2024.
- Yusuf Husain, Enny Dwi Oktaviyani, and Sherly Christina, "Analisis Perbandingan Algoritma Apriori, FP-Growth, Dan Eclat dalam Menemukan Pola Pembelian Konsumen," *KONSTELASI Konvergensi Teknol. Dan Sist. Inf.*, vol. 3, no. 2, pp. 231–243, Dec. 2023, doi: 10.24002/konstelasi.v3i2.7007.
- A. N. Saputri, A. F. Setiawan, and J. D. Irawan, "Analisis Pola Pembelian Di Toko Berkah Afi Menggunakan Metode Algoritma Apriori," vol. 9, no. 6, 2025.
- M. Mariko, "Perbandingan Algoritma Apriori Dan Algoritma Fp-Growth Untuk Rekomendasi Item Paket Pada Konten Promosi," vol. 11, no. 2, 2021.
- "Retail Transaction Dataset".
- Bhargavi Konda, "The impact of data preprocessing on data mining outcomes," *World J. Adv. Res. Rev.*, vol. 15, no. 3, pp. 540–544, Sep. 2022, doi: 10.30574/wjarr.2022.15.3.0931.
- T. Prasetya, J. E. Yanti, A. I. Purnamasari, A. R. Dikananda, and O. Nurdian, "Analisis Data Transaksi Terhadap Pola Pembelian Konsumen Menggunakan Metode Algoritma Apriori," *Inform. Educ. Prof. J. Inform.*, vol. 6, no. 1, p. 43, Apr. 2022, doi: 10.51211/itbi.v6i1.1688.
- I. W. Supriana, "Implementasi Algoritma Apriori sebagai Association Rule Learning untuk Mengidentifikasi Pola Item Dataset Penjualan".
- L. Aprita and A. Perdana, "Penerapan Metode Data Mining terhadap Data Transaksi Penjualan Menggunakan Algoritma Apriori pada Toko Metro Akustik," *J. Teknol. Inform. Dan Komput.*, vol. 9, no. 1, pp. 274–283, Mar. 2023, doi: 10.37012/jtik.v9i1.1381.
- M. Brilliantino and A. Perdana, "Penerapan Algoritma Apriori pada Analisis Data Transaksi penjualan UMKM Banyu Burgerbar," *J. Teknol. Inform. Dan Komput.*, vol. 9, no. 1, pp. 61–71, Mar. 2023, doi: 10.37012/jtik.v9i1.1339.
- E. Alma'arif, E. Utami, and F. W. Wibowo, "Implementasi Algoritma Apriori Untuk Rekomendasi Produk Pada Toko Online," *Creat. Inf. Technol. J.*, vol. 7, no. 1, p. 63, Mar. 2021, doi: 10.24076/citec.2020v7i1.241.
- Nurul Nisa Habibah, Widodo, and Bambang Prasetya Adhi, "Proses Association Rule Untuk Mengetahui Kecenderungan Belanja Pada Kopma Unj Menggunakan Algoritma Apriori Dan Eclat," *PINTER J. Pendidik. Tek. Inform. Dan Komput.*, vol. 6, no. 1, pp. 30–39, Jun. 2022, doi: 10.21009/pinter.6.1.5.
- L. Veronika and F. Syakti, "Perbandingan Algoritma Apriori dan Algoritma Eclat dalam penentuan pola peminjaman buku pada Perpustakaan Stikes Abdi Nusa Palembang".
- E. Halimatussa'diyah and S. K. Dini, "Implementasi Metode Content-Based Filtering Dengan Pendekatan Euclidean Distance Dalam Sistem Rekomendasi Produk Skincare," *Rabit J. Teknol. Dan Sist. Inf. Univrab*, vol. 10, no. 2, pp. 763–783, Jul. 2025, doi: 10.36341/rabit.v10i2.6371.