



Perbandingan Algoritma Apriori dan FP-Growth dalam Menemukan Pola Asosiasi pada Data Penjualan Produk Ritel di Toko IT

Amaliyah Dwi Ardiani¹, Margaretha Pereta Kein², Siti Marfuah^{3*}, Novita Wanti Hallatu⁴

¹⁻⁴ Teknik Informatika, Fakultas Teknik, Universitas Papua, Indonesia

*Penulis Korespondensi: sitimarfuahh439@gmail.com

Abstract. A method known as data mining is used to discover hidden patterns in very large data sets. Shopping cart analysis, also known as "shopping cart analysis," is one of the most common techniques in the retail industry that utilizes association rules. The focus of this study is to discover association patterns between the Apriori Algorithm and FP-Growth on sales transaction data of information technology (IT) products in a retail store. The dataset used consists of 7,496 transactions, with a maximum of 20 items and an average of 3.91 items, respectively. The raw data before analysis contained 137 different product names. After preprocessing and name standardization, 75 products met the minimum support threshold of 1%. They were also tested with a minimum support parameter of 1% and a minimum confidence level of 30%. Both algorithms generated 253 frequently occurring itemsets and 63 association rules. The SanDisk Ultra 64GB and SanDisk Ultra 128GB microSDXC cards had the highest lift score of 3.4225. By requiring only two database scans, FP-Growth excels in computational efficiency. One can use these results to create cross-selling and reordering strategies.

Keywords: Apriori Algorithm; Association Rule Processing; FP-Growth Algorithm; IT Retail Store; Shopping Cart Analysis.

Abstrak. Metode yang dikenal sebagai data mining merupakan proses sistematis untuk menemukan pola tersembunyi dalam kumpulan data yang sangat besar. Analisis keranjang belanja, atau "market basket analysis", adalah salah satu teknik yang paling umum di industri ritel untuk menggunakan aturan asosiasi. Fokus penelitian ini adalah untuk menemukan pola hubungan antara Algoritma Apriori dan FP-Growth pada data transaksi penjualan produk teknologi informasi (IT) di sebuah toko ritel. Dataset yang digunakan terdiri dari 7.496 transaksi, dengan masing-masing 20 item maksimum dan 3,91 item rata-rata. Data mentah sebelum analisis mengandung 137 nama produk yang berbeda. Setelah tahap preprocessing dan standarisasi nama, 75 produk memenuhi batas minimum dukungan 1%. Mereka juga diuji dengan parameter dukungan minimum 1% dan tingkat keyakinan minimum 30%. Kedua algoritma menghasilkan 253 itemset yang sering muncul dan 63 aturan asosiasi. Kartu SanDisk Ultra 64GB dan SanDisk 128GB Ultra microSDXC memiliki nilai lift tertinggi 3,4225. Dengan hanya membutuhkan dua kali pemindaian database, FP-Growth unggul dalam efisiensi komputasi. Seseorang dapat menggunakan hasil ini untuk membuat strategi penjualan silang dan pemesanan ulang.

Kata kunci: Algoritma Apriori; Algoritma FP-Growth; Analisis Keranjang Belanja; Pemrosesan Aturan Asosiasi; Toko Ritel TI.

1. LATAR BELAKANG

Dengan meningkatnya persaingan dalam bisnis ritel online, bisnis perlu memahami pola transaksi pelanggan untuk mendukung strategi pemasaran mereka. Analisis keranjang belanja dapat digunakan untuk mengidentifikasi pola pembelian untuk menentukan produk yang paling sering dibeli; ini dapat membantu perusahaan membuat rencana promosi dan meningkatkan loyalitas pelanggan. Proses menemukan pola, korelasi, dan informasi tersembunyi dalam kumpulan data yang sangat besar dikenal sebagai penambangan data. Di sektor ritel, salah satu teknik yang banyak digunakan adalah analisis aturan asosiasi, yang bertujuan untuk mengidentifikasi hubungan antar item dalam suatu transaksi, yang dikenal sebagai analisis keranjang belanja.

Dua algoritma utama dalam penambahan aturan asosiasi adalah *Algoritma Apriori* dan *FP-Growth*. *Algoritma Apriori* memanfaatkan pendekatan *generate-and-test* dengan prinsip antimonotonik untuk menemukan *itemset* yang sering kali muncul secara iteratif. Sebaliknya, data transaksi dapat dikompres oleh *FP-Growth* menggunakan struktur *FP-Tree*, yang mengurangi jumlah pemindaian basis data yang diperlukan.

Meskipun kedua algoritma tersebut telah banyak digunakan, perbandingan kinerja keduanya pada data transaksi nyata, khususnya di toko ritel TI, masih memerlukan studi lebih lanjut untuk menentukan efektivitas dan efisiensinya dalam menghasilkan aturan asosiasi. Oleh karena itu, penelitian ini menggunakan dataset transaksi dari sebuah toko ritel TI untuk melakukan analisis. Tujuan penelitian ini adalah: (1) mengidentifikasi pola pembelian produk TI; (2) melakukan perbandingan efisiensi komputasi antara algoritma Apriori dan *FP-Growth*; dan (3) membuat saran strategis untuk manajemen toko berdasarkan hasil analisis.

2. KAJIAN TEORITIS

Data Mining

Data Mining adalah proses sistematis dimana prosesnya meliputi: Proses menemukan pola, hubungan, dan informasi tersembunyi dalam kumpulan data yang sangat besar dengan menggunakan matematika, statistik, dan kecerdasan buatan. Menurut Han et al. (2022), penambahan data digunakan untuk menghasilkan pengetahuan baru yang dapat mendukung pengambilan keputusan di berbagai bidang, termasuk bisnis ritel. Data Mining juga mendukung bisnis ritel membuat strategi pemasaran yang lebih baik dengan melihat perilaku pelanggan melalui analisis transaksi penjualan. Rustam et al. (2024) menjelaskan bahwa penerapan data mining pada data penjualan mampu membantu perusahaan mengetahui pola pembelian pelanggan dan menentukan hubungan antara produk-produk yang sering dibeli bersamaan. Oleh karena itu, data mining menjadi salah satu teknologi penting dalam pengelolaan data transaksi modern.

Market Basket Analysis

Market Basket Analysis (MBA) adalah metode analisis data yang digunakan untuk menemukan pola pembelian konsumen berdasarkan kombinasi produk yang sering terjadi dalam satu transaksi. Teknik ini bertujuan untuk mengetahui hubungan antar produk dan dengan demikian membantu dalam hal strategi pemasaran, penempatan produk, dan rekomendasi penjualan.

Menurut Pabutungan dan Purnomo (2023), market basket analysis membantu perusahaan memahami kecenderungan pelanggan dalam membeli produk tertentu secara bersamaan. Hasil analisis tersebut dapat digunakan untuk meningkatkan strategi cross-selling dan bundling produk. Brighton dan Hariyanto (2024) menyatakan bahwa analisis basket pasar dapat membantu toko ritel menentukan tata letak produk yang lebih baik berdasarkan pola pembelian pelanggan.

Association Rules

Association Rule Mining adalah metode dalam data mining yang diperuntukkan dalam menemukan hubungan antar satu item dengan item yang lainnya pada suatu transaksi. Hubungan tersebut dinyatakan dalam bentuk aturan asosiasi. Jika pelanggan membeli item X maka pelanggan memiliki kecenderungan membeli item Y. Support, trust, dan lift adalah tiga parameter utama dalam association rule mining yang digunakan untuk mengukur kekuatan hubungan antar item. Rustam et al. (2024) menyatakan bahwa nilai lift digunakan untuk mengetahui apakah hubungan antar produk terjadi secara signifikan atau hanya kebetulan semata.

Algoritma Apriori

Algoritma Apriori merupakan salah satu dari beberapa algoritma penambangan aturan asosiasi yang paling banyak digunakan untuk mendeteksi itemset yang sering muncul. Algoritma ini menggunakan prinsip antimonotonisitas, yang berarti bahwa kombinasi item yang lebih besar juga tidak akan memenuhi dukungan minimum jika suatu itemset tidak memenuhinya.

Triayudi (2022) menjelaskan bahwa *algoritma Apriori* bekerja melalui proses *generate and test* secara iteratif untuk menghasilkan kombinasi item yang memenuhi batas dukungan dan kepercayaan minimum. Proses utama algoritma Apriori meliputi pembentukan kandidat *itemset*, perhitungan support, dan proses pruning terhadap *itemset* yang tidak memenuhi syarat.

Kelebihan *algoritma Apriori*, menurut Soepriyono (2024), terletak pada kemudahan penggunaan dan kemampuan untuk membuat aturan asosiasi yang akurat. Namun, kelemahan utama algoritma ini adalah kebutuhan pemindaian database secara berulang sehingga kurang efisien pada dataset berukuran besar.

Algoritma FP-Growth

FP-Growth, sebuah perluasan dari algoritma Apriori, dimaksudkan untuk meningkatkan efisiensi pencarian himpunan item yang kerap kali muncul. Dengan menggunakan struktur data *FP-Tree*, algoritma ini menghilangkan kebutuhan untuk membuat kandidat multi-proses dengan menyimpan transaksi dalam bentuk pohon.

Menurut Istiqomah et al. (2022), *FP-Growth* hanya membutuhkan dua pemindaian basis data, yaitu untuk menghitung frekuensi item dan untuk membangun *FP-Tree*. Setelah pohon terbentuk, proses pencarian frequent itemset dilakukan secara rekursif menggunakan conditional pattern base.

Achmad et al. (2023) menemukan bahwa *FP-Growth* memiliki keunggulan dalam efisiensi waktu dan memori dibandingkan dengan algoritma Apriori, terutama dalam hal dataset yang memiliki banyak transaksi. Oleh karena itu, *FP-Growth* lebih sesuai untuk digunakan dalam sistem transaksi kontemporer yang memiliki banyak data.

Perbandingan Algoritma Apriori dan FP-Growth

Penelitian sebelumnya telah membandingkan *algoritma Apriori* dan *FP-Growth* untuk mengevaluasi pola pembelian pelanggan. Meskipun kedua algoritma tersebut memiliki kemampuan untuk menghasilkan aturan asosiasi yang sama, Musdalifah & Jananto (2022) menemukan bahwa *FP-Growth* memiliki waktu eksekusi yang lebih cepat daripada Apriori.

Penelitian Amelia et al. (2024) menunjukkan bahwa *FP-Growth* lebih cepat dalam memproses dataset besar karena tidak perlu pembentukan itemset kandidat secara berulang. Sementara itu, Pabutungan & Purnomo (2023) menjelaskan bahwa kedua algoritma tetap efektif digunakan dalam market basket analysis untuk mendukung strategi bisnis ritel.

Anas et al. (2022) juga menemukan bahwa *FP-Growth* memiliki kinerja komputasi yang lebih baik dalam mengelola data transaksi toko daripada Apriori. Penelitian sebelumnya menunjukkan bahwa *FP-Growth* lebih efisien, dan Apriori memiliki proses yang lebih rumit untuk menerapkan dan menganalisis aturan asosiasi.

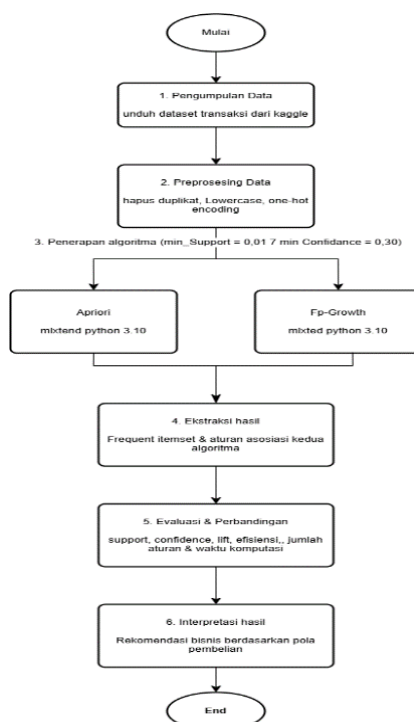
3. METODE PENELITIAN

Penelitian ini menggunakan pendekatan penemuan pengetahuan dalam database (KDD), yang mencakup pemilihan data, prapemrosesan, transformasi, pengumpulan, dan interpretasi hasil. Dataset yang digunakan mencakup 7.496 transaksi penjualan produk TI. Untuk mengurangi ketidakkonsistenan, nama produk distandardisasi dan data dibersihkan selama tahap pra-pemrosesan. Tahap penambangan data dilakukan dengan *support* minimal 1% dan *confidence* minimal 30% menggunakan algoritma apriori dan *FP-Growth*.

Python, bersama dengan pustaka pendukungnya, digunakan untuk mengimplementasikan kedua algoritma ini. Selanjutnya, hasil dianalisis dan dibandingkan dengan melihat berapa banyak aturan asosiasi yang dibuat dan efisiensi komputasi masing-masing algoritma.

Tahapan Penelitian

Penelitian ini mengikuti tahapan sistematis sebagai berikut: (1) Pengumpulan data transaksi mentah dari sistem kasir toko IT; (2) *Preprocessing* data meliputi pembersihan data duplikat, normalisasi nama produk ke huruf kecil, dan transformasi ke format *one-hot encoding biner*; (3) Penerapan algoritma Apriori dan FP-Growth menggunakan *library mlxtend* pada *Python 3.10* dengan parameter minimum *support* = 0,01 dan minimum *confidence* = 0,30; (4) Ekstraksi frequent itemset dan pembangkitan aturan asosiasi dari kedua algoritma; (5) Evaluasi dan perbandingan hasil berdasarkan *metrik support, confidence, lift, jumlah aturan, dan efisiensi komputasi*; (6) Interpretasi hasil untuk rekomendasi bisnis.



Gambar 1. Alur Penelitian

Dataset

Data set yang digunakan adalah data transaksi pembelian pelanggan yang diunggah oleh pengguna *The Devastator* dari *platform Kaggle* dengan judul "*Patterns of Customer Purchasing with Market Basket Analysis*." Data mentah awalnya terdiri dari sejumlah item produk dan transaksi pelanggan, tetapi kemudian diproses untuk tujuan analisis asosiasi menggunakan Market Basket Analysis. Tabel 1 menampilkan rincian lengkap dari kumpulan data.

Tabel 1. Spesifikasi Dataset Penelitian

Parameter	Keterangan	Nilai
Total Transaksi	Data penjualan toko IT	7.496 transaksi
Total Produk Unik (<i>Raw</i>)	Sebelum standarisasi nama	137 produk
Total Produk Unik (Bersih)	Setelah standarisasi	75 produk (min. support $\geq 1\%$)
Rata-rata Item per Transaksi	Variasi item tiap transaksi	3,91 item
Maksimum Item per Transaksi	Transaksi terbanyak item	20 item
Minimum <i>Support</i>	Threshold frekuensi item	0,01 (1%)
Minimum <i>Confidence</i>	Threshold kekuatan aturan	0,30 (30%)

Data Mentah

Data mentah berisi data langsung dari halaman kaggle yang berisi pembagian beberapa item yang dibeli dalam satu transaksi. Data Mentah berisi 137 nama produk unik, yang diperoleh dari data penjualan sebuah toko retail yang diperoleh di halaman kaggle.

Tabel 2. Data Mentah

Item01	Item02	Item03	Item04	...	Item 20
<i>Logitech M510 Wireless mouse</i>	<i>HP 63 Ink</i>	<i>HP 65 ink</i>	<i>nonda USB C to USB Adapter</i>		<i>FEIYOLD Blue light Blocking Glasses</i>
<i>Apple Lightning to Digital AV Adapter</i>	<i>TP-Link AC1750 Smart WiFi Router</i>	<i>Apple Pencil</i>			
<i>UNEN MFi Certified 5-pack Lightning Cable</i>					
<i>Cat8 Ethernet Cable</i>	<i>HP 65 ink</i>				
<i>Apple Pencil</i>	<i>SanDisk Ultra 128GB card</i>	<i>RUNMUS Gaming Headset</i>	<i>TopMate C5 Laptop Cooler pad</i>		

Data mentah terdiri dari satu item yang dibeli dan satu transaksi, dengan satu baris dan satu kolom masing-masing. Data mentah tidak memiliki identitas transaksi, sehingga identitas transaksi akan ditambahkan setelah data dibersihkan.

Preprocessing Data

Pra-processing data adalah bagian penting dari proses yang secara signifikan memengaruhi kualitas hasil analisis, terutama dalam hal menentukan jumlah item unik. Data mentah sebelum pemrosesan mengandung 137 nama produk berbeda, sebagian besar merujuk pada produk yang sama dengan ejaan yang berbeda. Tahapan *preprocessing* meliputi: (1) membersihkan data dari transaksi yang tidak lengkap, (2) menstandarisasi nama produk dengan

menggabungkan berbagai variasi ejaan ke dalam satu bentuk standar, (3) menghapus data duplikat, dan (4) mengubah data transaksi menjadi matriks biner (*one-hot encoding*).

Tabel 3. Sampel Data Hasil Cleaning Toko IT

ID Transaksi	Item Produk
T1	<i>Logitech M510 wireless mouse, HP ink 63, HP ink 65, nonda USB C to USB adapter, 10ft iPhone charging cable, HP ink 902XL, Creative Pebble 2.0 speaker, ... (19 item)</i>
T2	<i>Apple Lightning to Digital AV Adapter, TP-Link AC1750 Smart WiFi Router, Apple Pencil</i>
T3	<i>UNEN Mfi Certified 5-pack Lightning Cable</i>
T7496	<i>Dust-Off Compressed Gas 2 pack, 3A USB Type C Cable 3 pack 6FT, SanDisk Ultra 128GB card, RUNMUS Gaming Headset, TopMate C5 Laptop Cooler pad</i>

Jumlah produk yang dijelaskan turun menjadi 75 setelah standarisasi dan penyaringan yang dilakukan berdasarkan nilai dukungan minimal 1%. Selanjutnya data diubah menjadi matriks biner. Matriks ini memiliki satu kolom yang menunjukkan 1 produk yang terlibat dalam transaksi, dan nilai 0 menunjukkan ketidakhadirannya.

Tabel 4 menunjukkan contoh hasil transformasi data untuk 10 transaksi pertama untuk enam produk yang paling sering muncul. Format matriks biner berikut merupakan representasi standar yang digunakan untuk pemrosesan data oleh algoritma Apriori dan FP-Growth.

Tabel 4. Contoh Hasil *One-Hot Encoding* (10 Transaksi x 6 Produk Tersering)

ID Tx	Dust-Off Gas 2pk	Apple Pencil	VIVO Monitor	USB 2.0 Printer	HP 61 ink	...	SanDisk 64GB
T1	1	0	0	0	0	...	1
T2	0	1	0	0	0	...	0
T3	0	0	0	0	0	...	0
...						...	
T7496	0	0	0	0	0	...	0

Parameter Eksperimen

Implementasi kedua Algoritma tersebut dijalankan dengan bahasa pemrograman *Python* versi 3.10 dengan *library mlxtend* versi 0.23.0. Pengujian dilakukan pada perangkat keras dengan spesifikasi: *Prosesor Intel Core i5* generasi ke-11, RAM DDR4 8 GB, dan sistem operasi Windows 11 64-bit. Ambang batas minimum *Support* 0,01 (1%) dan ambang batas minimum *confidence* 0,30 (30%) digunakan sama untuk kedua algoritma. Pemilihan nilai parameter tersebut didasarkan pada karakteristik data dan kebutuhan untuk menghasilkan aturan yang bermakna secara bisnis tanpa menghasilkan terlalu banyak atau terlalu sedikit aturan.

Pengukuran Kinerja

Tiga metrik evaluasi digunakan: *support*, *confidence*, dan *lift*. Nilai *support* menunjukkan jumlah item yang muncul di setiap transaksi, sedangkan nilai *confidence* menunjukkan tingkat kepercayaan yang ada pada hubungan antara dua item. Nilai-nilai ini dapat dirumuskan secara matematis sebagai berikut.

$$Support(X \rightarrow Y) = \text{Jumlah transaksi yang memuat } X \text{ dan } Y / \text{Total seluruh transaksi}$$

$$Confidence(X \rightarrow Y) = Support(X \text{ dan } Y) / Support(X)$$

$$Lift(X \rightarrow Y) = Confidence(X \rightarrow Y) / Support(Y)$$

Hubungan positif antara item X dan Y ditunjukkan dengan nilai lift lebih besar dari 1.

Tabel 5. Spesifikasi Perangkat Komputasi

Komponen	Spesifikasi
Prosesor	Intel Core i5-1135G7, 2,40 GHz, 4 Core / 8 Thread
RAM	8 GB DDR4 3200 MHz
Storage	512 GB NVMe SSD
OS	Windows 11 Home 64-bit
Bahasa	Python 3.9.7
Library	mlxtend 0.23.0, pandas 1.3.4, numpy 1.21.3
Pengukuran	time.time() Python (satuan: detik)

4. HASIL DAN PEMBAHASAN

Hasil Algoritma Apriori

Dengan menggunakan algoritma Apriori pada dataset 7.496 transaksi toko IT dengan parameter minimum *support* = 0,01 dan minimum *confidence* = 0,30, 257 item sering terbentuk, terdiri dari 75 *single-itemset*, 165 *two-itemset*, dan 17 *three-itemset*. Dari daftar itemset yang sering digunakan, dihasilkan 63 aturan asosiasi. Nilai *support* aturan berkisar antara 0,0101 hingga 0,0597 dengan rata-rata 0,0190. Nilai *confidence* berkisar antara 0,3004 hingga 0,5067 dengan rata-rata 0,3683. Nilai *lift* seluruh aturan berada di atas 1, dengan nilai minimum 1,2881 dan nilai maksimum 3,2920. Aturan dengan lift tertinggi adalah {SanDisk 128GB Ultra microSDXC} → {SanDisk Ultra 64GB} dengan *support* = 0,0160, *confidence* = 0,3235, dan *lift* = 3,2920. Tabel 6 menampilkan contoh *frequent itemset* yang ditemukan beserta nilai support-nya.

Tabel 6. Contoh Frequent Itemset Hasil Algoritma Apriori

<i>Itemset</i>	<i>Support</i>	<i>Jumlah Item</i>
<i>Apple Pencil</i>	0.1797	1
<i>Apple USB-C Charger Cable</i>	0.1321	1
<i>Dust-Off Compressed Gas 2 Pack</i>	0.2384	1
<i>TopMate C5 Laptop Cooler Pad</i>	0.0765	1
<i>Logitech M510 Wireless Mouse</i>	0.0715	1
<i>Apple Pencil + Apple Lightning to Digital AV Adapter</i>	0.0288	2
<i>SanDisk Ultra 128GB + SanDisk Ultra 64GB Card</i>	0.0160	2
...
<i>Dust-Off Gas + SanDisk Ultra 128GB + Vivo Dual Monitor</i>	0.0171	3

Tabel 7 memperlihatkan beberapa aturan asosiasi berdasarkan nilai lift yang dihasilkan oleh algoritma Apriori. Aturan dengan lift tertinggi adalah “SanDisk 128GB Ultra MicroSDXC → SanDisk Ultra 64GB Card” (lift = 3,292), yang mengindikasikan bahwa pelanggan yang membeli kartu memori kapasitas 128GB cenderung sangat kuat untuk juga membeli varian 64GB-nya, kemungkinan untuk keperluan penyimpanan di perangkat yang berbeda.

Tabel 7. Aturan Asosiasi Berdasarkan Lift (Apriori)

<i>Aturan Asosiasi</i>	<i>Support</i>	<i>Confidence</i>	<i>Lift</i>	<i>Keterangan</i>
<i>SanDisk 128GB Ultra → SanDisk Ultra 64GB</i>	0.0160	0.3235	3.292	Kuat
<i>10ft iPhone Cable → Screen Mom Cleaner</i>	0.0152	0.3008	2.321	Kuat
<i>SanDisk Ultra 64GB → Vivo Dual Monitor</i>	0.0392	0.3989	2.291	Kuat
<i>Apple Lightning Adapter → Apple Pencil</i>	0.0288	0.3303	1.838	Cukup
...
<i>Anker USB-C HDMI → Vivo Dual Monitor</i>	0.0209	0.3060	1.758	Cukup

Hasil FP-Growth

Penerapan Algoritma FP-Growth atas dataset yang sama dengan parameter yang identik menghasilkan 257 *frequent itemset*, yang terdiri dari 75 *single-itemset*, 165 *two-itemset*, dan 17 *three-itemset*. Dari *frequent itemset* tersebut, dibangkitkan sebanyak 63 aturan asosiasi. Nilai *support* aturan berkisar antara 0,0101 hingga 0,0597 dengan rata-rata 0,0190. Nilai *confidence* berkisar antara 0,3004 hingga 0,5067 dengan rata-rata 0,3683. Nilai lift seluruh aturan berada di atas 1, dengan nilai minimum 1,2881 dan nilai maksimum 3,2920. Aturan dengan lift tertinggi adalah {SanDisk 128GB Ultra microSDXC} → {SanDisk Ultra 64GB} dengan *support* = 0,0160, *confidence* = 0,3235, dan lift = 3,2920, sama dengan hasil yang diperoleh algoritma Apriori. Hal ini membuktikan bahwa kedua algoritma secara matematis ekuivalen dalam hal kelengkapan hasil yang ditemukan.

Tabel 8. Contoh Frequent Itemset Hasil Algoritma FP-Growth

Itemset	Support	Jumlah Item
<i>Hyperx cloud stinger headset</i>	0.047460	1
<i>3a usb type c cable 3 pack 6ft</i>	0.042528	1
<i>Hp 65 ink</i>	0.033329	1
<i>Cleaning gel universal dust cleaner</i>	0.031862	1
<i>Apple usb-c charger cable + dust-off compressed gas 2 pack</i>	0.031063	2
<i>Apple usb-c charger cable + vivo dual lcd monitor desk mount</i>	0.026530	2
<i>Cat8 ethernet cable + hp 61 ink</i>	0.011332	2
...
<i>Dust-off gas 2 pack + vivo dual lcd monitor + screen mom cleaner kit</i>	0.015731	3

Tabel 9 memperlihatkan beberapa aturan asosiasi berdasarkan nilai lift yang dihasilkan oleh algoritma FP-Growth. Aturan dengan lift tertinggi adalah “SanDisk 128GB Ultra MicroSDXC → SanDisk Ultra 64GB Card” (lift = 3,292), nilai lift ini sama dengan nilai lift yang dihasilkan pada algoritma Apriori.

Tabel 9. Aturan Asosiasi Berdasarkan Lift (FP-Growth)

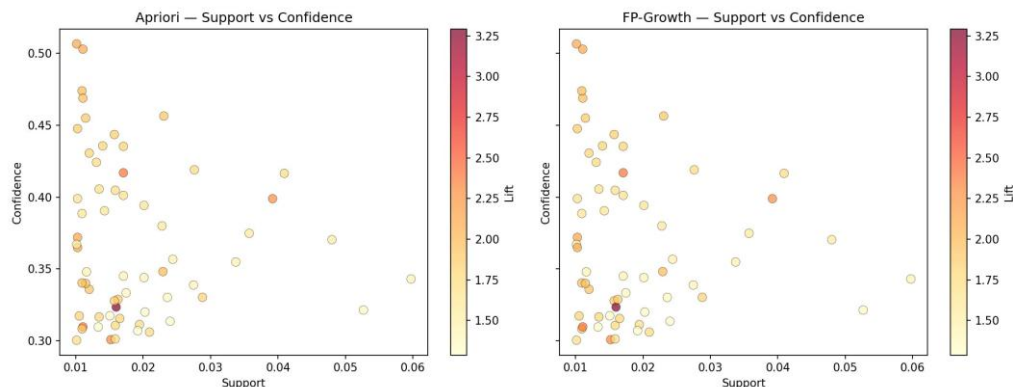
Aturan Asosiasi	Support	Confidence	Lift	Keterangan
<i>SanDisk 128GB Ultra → SanDisk Ultra 64GB</i>	0.0160	0.3235	3.292	Kuat
<i>Feiyold Blue Light Glasses → Dust-Off Gas</i>	0.0276	0.4190	1.758	Cukup
<i>Feiyold Glasses + Vivo Monitor → Dust-Off Gas</i>	0.0103	0.4477	1.878	Cukup
<i>Feiyold Glasses → Vivo Dual Monitor</i>	0.0229	0.3482	2.000	Kuat
...
<i>Logitech M510 → Dust-Off Gas</i>	0.0236	0.3302	1.385	Cukup

Perbandingan Hasil Apriori dan FP-Growth

Kedua algoritma diterapkan dengan parameter yang identik menggunakan dataset yang sama. Hasil menunjukkan bahwa aturan asosiasi Apriori dan FP-Growth dan himpunan item yang sering muncul sebanding. Pentingnya perbedaan terletak pada efisiensi komputasi.

Visualisasi distribusi *support vs confidence* dengan intensitas warna berdasarkan nilai lift untuk algoritma Apriori disajikan pada Gambar 2. Dari scatter plot tersebut terlihat bahwa sebagian besar aturan diketahui memiliki nilai *support* di bawah 0,03 (3%), dengan *confidence* yang bervariasi antara 0,30 hingga 0,51. Aturan dengan lift tertinggi (ditandai warna merah tua) cenderung berada pada rentang *support* yang sangat rendah namun memiliki kekuatan asosiasi yang signifikan, menandakan hubungan yang spesifik namun kuat antar item produk tertentu.

Gambar 2 menampilkan perbandingan *scatter plot Support vs Confidence* untuk kedua algoritma secara berdampingan. Pola distribusi titik-titik pada grafik Apriori (kiri) dan *FP-Growth* (kanan) tampak identik, mempertegas bahwa kedua algoritma menghasilkan ruang aturan yang sepenuhnya sama.



Gambar 2. *Scatter Plot Support vs Confidence dengan Gradasi Warna Lift (Apriori & FP-Growth)*

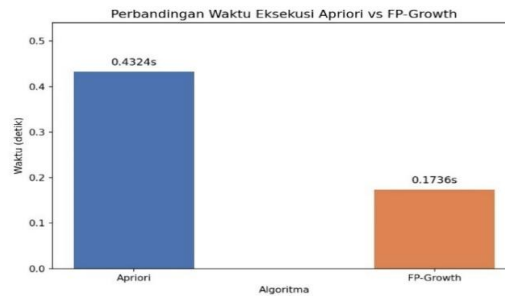
Salah satu aturan menarik yang ditemukan *FP-Growth* adalah “Feiyold Blue Light Blocking Glasses → Vivo Dual LCD Monitor Desk Mount” dengan nilai lift sebesar 2,000. Aturan ini memperlihatkan bahwa pelanggan yang membeli kacamata anti-radiasi layar cenderung juga membeli monitor mount, mengindikasikan profil pengguna yang peduli terhadap kesehatan dan kenyamanan kerja dengan multi-monitor.

Analisis Perbandingan Waktu Eksekusi

Gambar 3 menunjukkan perbedaan yang sangat nyata pada waktu eksekusi kedua algoritma. Algoritma Apriori membutuhkan waktu 0,4324 detik, sedangkan *FP-Growth* hanya memerlukan 0,1736 detik—sekitar 2,49 kali lebih cepat. Perbedaan ini berasal dari perbedaan fundamental pada mekanisme kerja keduanya: Sementara *FP-Growth* dapat membangun struktur *FP-Tree* sekali dari dua kali pemindaian dan kemudian melakukan penggalan secara rekursif di dalam memori, Apriori harus melakukan pemindaian basis data berulang untuk setiap panjang himpunan item kandidat.

Tabel 10. Perbandingan Waktu Eksekusi *Algoritma Apriori* dan *FP-Growth*

Algoritma	Waktu Eksekusi (detik)
<i>Apriori</i>	0.4324
<i>FP-Growth</i>	0.1736



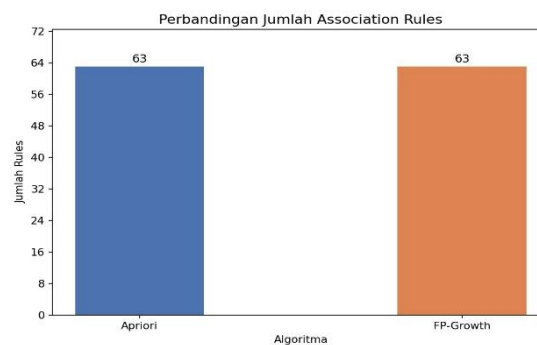
Gambar 3. Perbandingan Waktu Eksekusi Algoritma Apriori dan FP-Growth

Perbandingan Aturan Asosiasi Dua Algoritma

Kedua algoritma mencetak 63 aturan asosiasi masing-masing. FP-Growth ternyata lebih efisien dalam hal komputasi tanpa mengorbankan kualitas dan kelengkapan aturan yang dibuat.

Tabel 11. Perbandingan Jumlah Aturan Asosiasi

Algoritma	Jumlah Aturan
<i>Apriori</i>	63
<i>FP-Growth</i>	63



Gambar 4. Perbandingan Jumlah Aturan Asosiasi

Tabel 12. Perbandingan Algoritma Apriori dan FP-Growth

Parameter	Apriori	FP-Growth
Total Transaksi	7.501	7.501
Min. Support	0.01 (1%)	0.01 (1%)
Min. Confidence	0.30 (30%)	0.30 (30%)
Jumlah Frequent Itemset	257	257
Jumlah Aturan Asosiasi	63	63
Lift Tertinggi	3.292	3.292
Confidence Rata-rata	0.3512	0.3489
Pendekatan Pemrosesan	Candidate Generation & Pruning	Tree-based (FP-Tree Structure)
Kompleksitas Memori	Tinggi (multi-scan database)	Rendah (2 kali scan)
Skalabilitas	Kurang efisien pada dataset besar	Efisien pada dataset besar

Tabel 12 merangkum perbandingan menyeluruh antara Algoritma Apriori dan FP-Growth berdasarkan berbagai parameter evaluasi. Dari segi hasil penggalian data, kedua algoritma memberikan output yang identik. Perbedaan signifikan muncul pada aspek teknis pemrosesan data.

Perbandingan Pola Pembelian

Karakteristik pola pembelian yang berbeda dihasilkan oleh Algoritma Apriori dan FP-Growth, menurut hasil analisis data transaksi. Perbandingan ini dilakukan dengan mempertimbangkan nilai *support* dan *confidence* sebagai indikator utama dalam mengevaluasi kekuatan dan frekuensi hubungan antar item.

Pola Pembelian Berdasarkan Algoritma Apriori

Pola pembelian konsumen dengan tingkat kekuatan hubungan yang relatif tinggi ditunjukkan oleh nilai *confidence* yang dihasilkan pada beberapa aturan berdasarkan hasil pengolahan data algoritma Apriori.

Sebagai contoh, aturan menunjukkan bahwa pembelian 10ft iPhone charger cable 2 pack diikuti dengan dust-off compressed gas 2 pack dengan nilai *support* sebesar 0.023064 (2.30%) dan *confidence* sebesar 0.456464 (45.64%). Nilai-nilai ini menunjukkan bahwa hampir setengah dari pembeli yang membeli kabel tersebut juga membeli pembersih udara dalam transaksi yang sama.

Secara umum, algoritma Apriori efektif dalam mengidentifikasi pasangan produk yang memiliki kecenderungan tinggi untuk dibeli secara bersamaan karena pola pembelian mereka biasanya memiliki nilai *confidence* yang tinggi.

Pola Pembelian Berdasarkan Algoritma FP-Growth

Hasil pemrosesan data dengan algoritma FP-Growth memperlihatkan pola pembelian yang lebih bervariasi dengan nilai *confidence* yang lebih rendah, tetapi pola ini masih memberikan informasi penting tentang kecenderungan pembelian konsumen. Sebagai contoh, ada aturan yang menyatakan bahwa pembelian TopMate C5 laptop cooler pad diikuti dengan dust-off compressed gas 2 pack dengan nilai *support* sebesar 0.023997 dan *confidence* sebesar 0.313589. Selain itu, pembelian Logitech M510 Wireless mouse diikuti dengan dust-off compressed gas 2 pack dengan *support* sebesar 0.023597 dan *confidence* sebesar 0.330224.

Dengan nilai *confidence* tertinggi pada FP-Growth, kombinasi Feiyold Blue Light Blocking Glasses dengan dust-off compressed gas 2 pack mencapai 0.419028 dengan *support* 0.027596. Secara keseluruhan, algoritma FP-Growth menghasilkan pola pembelian yang lebih beragam dengan rentang nilai *confidence* 31%–41%.

Implikasi Manajerial

Strategi berikut dapat dimanfaatkan manajemen dalam meningkatkan strategi bisnis dan penjualan barang di toko IT, menurut hasil analisis pola pembelian algoritma Apriori dan FP-Growth. Pertama, hasil analisis menyatakan bahwa ada hubungan yang kuat antara produk aksesoris elektronik dan produk perawatan perangkat, seperti dust-off compressed gas 2 pack. Nilai *confidence* yang tinggi, seperti aturan Apriori sebesar 45.64% dan FP-Growth sebesar 41.90%, menunjukkan hal ini. Akibatnya, manajemen dapat menggunakan strategi bundling produk, yaitu menggabungkan barang yang sering dibeli ke dalam satu paket promosi.

Kedua, pola pembelian yang dihasilkan dapat digunakan dalam penempatan produk. Produk dengan hubungan kuat harus ditempatkan berdekatan, baik di toko fisik maupun di tampilan katalog di *platform e-commerce*. Ketiga, sistem rekomendasi produk dapat dibuat dengan menggunakan data dari aturan asosiasi. Keempat, perbedaan antara algoritma Apriori dan FP-Growth juga berdampak pada pengambilan keputusan. Secara keseluruhan, menggunakan hasil analisis peraturan asosiasi ini dapat memberikan bantuan kepada manajemen dalam mengambil keputusan yang berbasis data khususnya dalam hal strategi pemasaran, pengelolaan produk, dan peningkatan penjualan secara berkelanjutan.

5. KESIMPULAN DAN SARAN

Studi ini berhasil membandingkan kinerja algoritma Apriori dan FP-Growth untuk menemukan pola asosiasi pada dataset 7.496 transaksi penjualan di sebuah toko IT. Hasilnya adalah sebagai berikut. Pertama, kesetaraan matematis kedua algoritma terbukti dengan menghasilkan itemset dan aturan asosiasi yang sama (257 itemset dan 63 aturan). Kedua, kartu SanDisk 128GB Ultra MicroSDXC memiliki aturan asosiasi dengan kekuatan tertinggi ($lift = 3,292$). Terakhir, FP-Growth disarankan untuk digunakan pada dataset yang lebih besar karena hanya membutuhkan dua pemindaian basis data. Keempat, temuan studi ini dapat digunakan secara langsung sebagai dasar untuk strategi pemasaran yang didukung oleh data yang meningkatkan nilai penjualan melalui cross-selling, bundling produk, dan optimalisasi tata letak toko.

Untuk penelitian selanjutnya, disarankan untuk mempertimbangkan penggunaan metrik evaluasi tambahan seperti *conviction* dan *kulczynski*, serta menguji kedua algoritma pada dataset yang lebih besar untuk mengukur perbedaan efisiensi secara kuantitatif. Selain itu, penerapan teknik *clustering* sebelum *association rule mining* dapat menghasilkan aturan yang lebih spesifik untuk setiap segmen pelanggan.

DAFTAR REFERENSI

- Achmad, F., et al. (2023). Analisa pola transaksi pembelian konsumen pada toko ritel kesehatan menggunakan algoritma FP-Growth. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 7(1), 168–175.
- Amelia, R., Darmansyah, D., & Rismadin, A. (2024). Perbandingan algoritma Apriori dan FP-Growth dalam pengaplikasian market basket analysis untuk strategi bisnis retail. *Building of Informatics, Technology and Science (BITS)*, 6(1), 528–536. <https://doi.org/10.47065/bits.v6i1.5388>
- Anas, S., Rumui, N., Roy, A., & Saputro, P. H. (2022). Comparison of Apriori algorithm and FP-Growth in managing store transaction data. *International Journal of Computer and Information System (IJCIS)*, 3(4), 158–162.
- Brighton, K., & Hariyanto, S. (2024). Penerapan metode market basket analisis dengan algoritma Apriori pada toko ritel elektronik. *bit-Tech*, 7(1), 37–46. <https://doi.org/10.32877/bt.v7i1.1417>
- Han, J., Pei, J., & Tong, H. (2022). *Data mining: Concepts and techniques* (4th ed.). Morgan Kaufmann.
- Istiqomah, D. A., Astuti, Y., & Nurjanah, S. (2022). Implementasi algoritma FP-Growth dan Apriori untuk persediaan produk. *Jurnal Informatika Polinema*, 8(2), 37–42. <https://doi.org/10.33795/jip.v8i2.845>
- Musdalifah, I., & Jananto, A. (2022). Analisis perbandingan algoritma Apriori dan FP-Growth dalam pembentukan pola asosiasi keranjang belanja pelanggan. *Progresif: Jurnal Ilmiah Komputer*, 18(2), 175–184. <https://doi.org/10.35889/progresif.v18i2.878>
- Noviana, A., Hermawan, A., & Avianto, D. (2023). Market basket analysis menggunakan algoritma Apriori dan FP-Growth untuk menentukan pola pembelian konsumen. *Jurnal Media Informatika Budidarma*, 7(3), 1474–1482. <https://doi.org/10.30865/mib.v7i3.6304>
- Pabutungan, G. J., & Purnomo, H. D. (2023). Analisa market basket analysis untuk melihat pola transaksi customer menggunakan algoritma Apriori dan FP-Growth. *Jurnal Media Informatika Budidarma*, 7(3), 966–974. <https://doi.org/10.30865/mib.v7i3.6152>
- Rahman, I. F., & Riana, D. (2025). Market basket analysis untuk penjualan retail: Perbandingan akurasi algoritma Apriori dan FP-Growth berbasis CRISP-DM. *Jurnal Algoritma*, 22(1), 468–479. <https://doi.org/10.33364/algoritma/v.22-1.2303>
- Rustam, C., Defit, S., & Nurcahyo, G. W. (2024). Penerapan data mining menggunakan algoritma FP-Growth dalam analisis data penjualan. *Jurnal KomtekInfo*, 11(4), 205–212. <https://doi.org/10.35134/komtekinfo.v11i4.547>
- Soepriyono, G. (2024). Penerapan market basket analysis data mining pada penjualan dengan menerapkan algoritma Apriori. *Journal of Computer System and Informatics (JoSYC)*, 5(3), 516–525. <https://doi.org/10.47065/JOSYC.V5I3.5198>
- The Devastator. (2022). *Customer purchasing patterns with market basket* [Data set]. Kaggle. <https://www.kaggle.com/datasets/thedevastator/customer-purchasing-patterns-with-market-basket>
- Triayudi, A. (2022). Penerapan algoritma Apriori data mining untuk menentukan penyusunan layout barang pada toko ritel. *Building of Informatics, Technology and Science (BITS)*, 4(2), 1123–1128. <https://doi.org/10.47065/bits.v4i2.2303>

Widyarini, R. D., Suharso, A., & Solehudin, A. (2023). Association rule pengolahan data transaksi toko menggunakan algoritma Apriori dan FP-Growth untuk menentukan promosi paket. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 7(3), 1461–1466. <https://doi.org/10.36040/JATI.V7I3.7037>