



Penerapan Algoritma FP-Growth dan K-Means Clustering dalam Analisis Pola Asosiasi Berdasarkan Segmentasi Pelanggan

Aldiansyah Pramudia Hasibuan, Fitri Insani*, Alwis Nazir, Iis Afrianty

Sains dan Teknologi, Teknik Informatika, Universitas Islam Negeri Sultan Syarif Kasim Riau, Pekanbaru
Panam, Jl. HR. Soebrantas No.Km. 15, RW.15, Simpang Baru, Kota Pekanbaru, Riau, Indonesia

Email: ¹11950115003@students.uin-suska.ac.id, ^{2,*}fitri.insani@uin-suska.ac.id,

³alwis.nazir@uin-suska.ac.id, ⁴iis.afrianty@uin-suska.ac.id

Email Penulis Korespondensi: fitri.insani@uin-suska.ac.id

Submitted: 11/03/2025; Accepted: 05/04/2025; Published: 12/04/2025

Abstrak—Industri farmasi mengalami pertumbuhan pesat yang mendorong perusahaan untuk memanfaatkan data penjualan secara efektif guna meningkatkan strategi pemasaran berbasis data. Namun, pemanfaatan data penjualan masih menjadi tantangan bagi perusahaan XYZ, distributor farmasi. Penelitian ini bertujuan untuk menganalisis pola pembelian pelanggan melalui penerapan algoritma FP-Growth dalam analisis asosiasi, yang dikombinasikan dengan segmentasi pelanggan menggunakan algoritma K-Means berbasis analisis RFM (Recency, Frequency, Monetary). Proses segmentasi menghasilkan empat cluster pelanggan, yaitu pelanggan aktif dan loyal (Cluster 1), pelanggan pasif (Cluster 2), pelanggan kurang aktif (Cluster 3), dan pelanggan baru (Cluster 4). Analisis FP-Growth pada setiap cluster menunjukkan bahwa Cluster 1 memiliki 10 aturan asosiasi signifikan dengan minimum support 0,01 dan confidence 0,7, sementara Cluster 2, 3, dan 4 masing-masing menghasilkan 2, 3, dan 4 aturan asosiasi dengan parameter yang disesuaikan. Seluruh aturan memiliki nilai lift > 1, menandakan hubungan antar produk yang positif. Hasil penelitian ini memberikan wawasan strategis bagi perusahaan dalam merancang pendekatan pemasaran berbasis data, seperti penawaran produk yang lebih tepat sasaran untuk pelanggan loyal atau strategi retensi bagi pelanggan pasif, sehingga dapat mengoptimalkan penjualan dan meningkatkan profitabilitas di setiap segmen pelanggan.

Kata Kunci: Analisis Asosiasi; Analisis RFM; FP-Growth; K-Means; Segmentasi Pelanggan; Strategi Pemasaran

Abstract—The pharmaceutical industry has experienced rapid growth, urging companies to leverage sales data effectively to enhance data-driven marketing strategies. However, utilizing sales data remains a challenge for XYZ company, a pharmaceutical distributor. This study aims to analyze customer purchasing patterns by applying the FP-Growth algorithm for association analysis, combined with customer segmentation using the K-Means algorithm based on RFM (Recency, Frequency, Monetary) analysis. The segmentation process resulted in four customer clusters: active and loyal customers (Cluster 1), passive customers (Cluster 2), less active customers (Cluster 3), and new customers (Cluster 4). FP-Growth analysis for each cluster revealed that Cluster 1 generated 10 significant association rules with a minimum support of 0.01 and confidence of 0.7, while Clusters 2, 3, and 4 produced 2, 3, and 4 association rules, respectively, with adjusted parameters. All rules showed a lift value > 1, indicating positive relationships between products. The findings of this study provide strategic insights for companies in designing data-driven marketing approaches, such as more targeted product offerings for loyal customers or retention strategies for passive customers, thereby optimizing sales and increasing profitability in each customer segment.

Keywords: Association Analysis; RFM Analysis; FP-Growth; K-Means; Customer Segmentation; Marketing Strategy

1. PENDAHULUAN

Dalam beberapa dekade terakhir, industri farmasi dan sektor kesehatan telah mengalami pertumbuhan pesat seiring dengan meningkatnya kesadaran masyarakat akan pentingnya kesehatan dan kebutuhan untuk menjaga kualitas hidup [1]. Peningkatan ini mendorong persaingan yang semakin ketat di antara para pemain industri farmasi, baik di tingkat produksi maupun distribusi [2]. Perusahaan tidak hanya dituntut untuk menyediakan produk yang lengkap dan berkualitas, tetapi juga harus memahami perilaku dan kebutuhan pelanggan untuk meningkatkan loyalitas serta kepuasan mereka [3]. Salah satu kunci keberhasilan di sektor ini adalah pemanfaatan data penjualan secara efektif untuk mendukung strategi pemasaran yang lebih terarah dan berbasis data [4].

Pemanfaatan data penjualan secara optimal masih menjadi tantangan bagi banyak perusahaan, termasuk perusahaan XYZ yang bergerak di bidang distribusi obat dan produk kesehatan. Berdasarkan wawancara dengan apoteker di perusahaan tersebut, data penjualan yang tercatat hanya disimpan tanpa diolah lebih lanjut. Padahal, data tersebut menyimpan informasi penting yang dapat digunakan untuk mengidentifikasi pola pembelian pelanggan, tren pasar, serta hubungan antar produk yang sering dibeli bersamaan [5]. Temuan ini dapat membantu perusahaan dalam merancang strategi pemasaran, seperti promosi produk bundling, mengoptimalkan stok, dan meningkatkan efisiensi distribusi [6]. Untuk mengekstraksi pola-pola seperti ini dari data penjualan yang besar, diperlukan teknik analisis yang tepat [7].

Association Rule Mining adalah salah satu metode yang dapat digunakan untuk menemukan pola atau keterkaitan antar item dalam data transaksi penjualan [8]. Salah satu metode yang efektif dalam melakukan analisis asosiasi adalah algoritma FP-Growth (Frequent Pattern Growth) [9], [10]. FP-Growth dirancang untuk mengidentifikasi pola frekuensi tinggi secara efisien dan terbukti unggul dibandingkan algoritma lain [11]. Dalam penelitian menggunakan framework CRISP-DM, FP-Growth menghasilkan 14 aturan asosiasi dengan kekuatan total 2,32, melampaui Apriori yang hanya menghasilkan 9 aturan dengan kekuatan 0,72 [12]. Penelitian lain juga menunjukkan FP-Growth lebih unggul dengan menghasilkan 156 aturan asosiasi dibandingkan 136 aturan dari

Apriori, meskipun keduanya mencapai support tertinggi 59% dan confidence 100% [13]. Dengan keunggulannya, FP-Growth dianggap ideal untuk menganalisis hubungan item yang sering dibeli bersama [11], [12], [13].

Pola asosiasi yang dihasilkan dari data yang luas sering kali kurang optimal jika tidak mempertimbangkan perbedaan karakteristik pelanggan [14]. Oleh karena itu, customer segmentation menjadi penting, yaitu proses pengelompokan pelanggan berdasarkan karakteristik atau perilaku mereka [15]. Segmentasi ini memungkinkan perusahaan menyesuaikan strategi pemasaran atau rekomendasi produk dengan kebutuhan spesifik setiap segmen [16], [17]. Dalam konteks ini, RFM Analysis (Recency, Frequency, Monetary) digunakan untuk mengelompokkan pelanggan berdasarkan seberapa baru, sering, dan besar pengeluaran mereka [18]. Kemudian, K-Means Clustering diterapkan untuk membentuk kelompok pelanggan dengan karakteristik RFM yang serupa, dengan cara mengelompokkan data berdasarkan kesamaan nilai RFM [19]. Jumlah cluster optimal dalam K-Means ditentukan menggunakan kombinasi Elbow Method dan Silhouette Score [20]. Setelah segmen pelanggan terbentuk, analisis asosiasi dilakukan pada setiap cluster menggunakan algoritma FP-Growth untuk menghasilkan pola yang lebih relevan dan spesifik bagi setiap segmen [21].

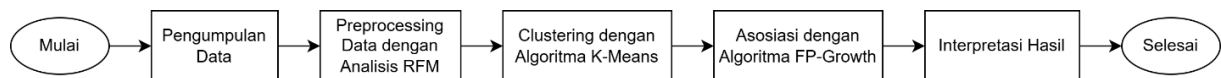
Penelitian terkait yang menggunakan teknik analisis asosiasi berbasis segmentasi pelanggan telah diterapkan di berbagai sektor bisnis. Salah satunya adalah penelitian yang mengimplementasikan metode clustering dengan algoritma K-Medoids di Aurel Mini Mart menggunakan model RFM. Penelitian ini berhasil membagi pelanggan menjadi dua kategori utama, yaitu Golden Customers (pelanggan dengan loyalitas tinggi) di Cluster 1, dan Dormant Customers (pelanggan yang jarang bertransaksi) di Cluster 2 hingga Cluster 5. Analisis asosiasi menggunakan FP-Growth dengan nilai minimum support 10% dan confidence 60% menghasilkan 10 aturan pada Cluster 1, 9 aturan pada Cluster 3, dan 3 aturan pada Cluster 4, sementara Cluster 2 dan Cluster 5 tidak menghasilkan aturan asosiasi yang memenuhi kriteria [22].

Penelitian lain di Swalayan 212 Mart memanfaatkan kombinasi analisis RFM dan algoritma DBSCAN untuk segmentasi pelanggan, yang berhasil mengidentifikasi 5 cluster pelanggan. Selanjutnya, algoritma FP-Growth diterapkan dengan nilai minimum support 30% dan confidence 70%, menghasilkan variasi aturan asosiasi, yaitu 7 aturan pada Cluster 1, 6 aturan pada Cluster 2, 10 aturan pada Cluster 3, serta masing-masing 2 dan 6 aturan pada Cluster 4 dan Cluster 5 [23]. Penelitian di Bazmart Pelalawan mengimplementasikan algoritma FP-Growth untuk menganalisis asosiasi produk yang sering dibeli bersama, serta algoritma Fuzzy C-Means untuk segmentasi pelanggan berdasarkan model LRFM. Hasil penelitian ini mengidentifikasi dua cluster pelanggan dengan kualitas baik (DBI = 0,628) dan menghasilkan 8 aturan asosiasi di Cluster 1 serta 17 aturan di Cluster 2 [24].

Melalui kombinasi metode asosiasi dan segmentasi pelanggan, penelitian ini diharapkan dapat memberikan wawasan yang lebih mendalam mengenai pola pembelian pada setiap segmen pelanggan. Dengan mengidentifikasi pola asosiasi dalam setiap cluster, perusahaan XYZ dapat merancang strategi pemasaran yang lebih spesifik, seperti promosi produk atau rekomendasi yang disesuaikan dengan kebutuhan masing-masing segmen pelanggan [23], [24].

2. METODOLOGI PENELITIAN

Penelitian ini menggunakan pendekatan data mining deskriptif untuk mengidentifikasi pola asosiasi antar produk berdasarkan segmentasi pelanggan. Adapun tahapan penelitian ini ada pada Gambar 1.



Gambar 1. Tahapan Penelitian

Penelitian ini dimulai dengan pengumpulan data transaksi penjualan obat di perusahaan XYZ selama periode Januari hingga Desember 2023. Setelah data diperoleh, dilakukan tahap preprocessing, yang mencakup seleksi data untuk memilih atribut yang relevan, diikuti dengan pembersihan data, seperti penanganan nilai yang hilang (missing values) dan penghapusan data duplikasi. Setelah data bersih, dilakukan transformasi dengan analisis RFM (Recency, Frequency, Monetary) untuk menghasilkan nilai yang digunakan sebagai dasar segmentasi pelanggan. Kemudian, hasil nilai RFM dinormalisasi ke rentang 0 hingga 1 menggunakan rumus min-max normalization (1), agar sesuai untuk algoritma K-Means.

$$X_{\text{norm}} = \frac{X - X_{\text{min}}}{X_{\text{max}} - X_{\text{min}}} \quad (1)$$

Langkah selanjutnya adalah melakukan clustering menggunakan algoritma K-Means, yang menghasilkan segmentasi pelanggan. Kemudian, data transaksi dikelompokkan berdasarkan masing-masing segmen. Setelah itu, data transaksi yang telah dikelompokkan dipersiapkan dalam format transaksi-item untuk analisis asosiasi menggunakan algoritma FP-Growth.

2.1 Model RFM

Model RFM adalah pendekatan analitis yang digunakan untuk mengevaluasi nilai pelanggan berdasarkan tiga dimensi utama [16]:



- a. Recency (R): Mengukur seberapa baru pelanggan melakukan transaksi terakhirnya. Pelanggan yang melakukan transaksi terakhir dalam waktu yang lebih dekat dianggap lebih bernilai karena menunjukkan tingkat keterlibatan yang aktif. Perhitungan Recency menggunakan persamaan (2).

$$R = T_{\text{acuan}} - T_{\text{terakhir transaksi}} \quad (2)$$

- b. Frequency (F): Menghitung jumlah transaksi yang dilakukan oleh pelanggan dalam periode tertentu. Semakin sering pelanggan bertransaksi, semakin tinggi tingkat loyalitas mereka terhadap perusahaan. Perhitungan Frequency menggunakan persamaan (3).

$$F = \text{Jumlah Transaksi} \quad (3)$$

- c. Monetary (M): Mengukur total nilai finansial yang dihasilkan oleh pelanggan selama periode tertentu. Pelanggan yang menghasilkan pendapatan lebih besar dianggap lebih bernilai karena memberikan kontribusi signifikan terhadap keuntungan perusahaan. Perhitungan Monetary menggunakan persamaan (4).

$$M = \sum_{i=1}^n \text{Nilai Transaksi ke } i \quad (4)$$

RFM banyak digunakan dalam pengelompokan pelanggan karena sifatnya yang sederhana namun efektif dalam memberikan wawasan berbasis data [18]. Model ini telah terbukti membantu bisnis dalam memahami pola belanja pelanggan, meningkatkan loyalitas, dan mengoptimalkan kampanye pemasaran [19].

2.2 Algoritma K-Means

K-Means adalah algoritma unsupervised learning yang digunakan untuk membagi data ke dalam k kelompok (cluster) berdasarkan kemiripan antar data [25]. Algoritma ini bekerja dengan meminimalkan variansi dalam setiap cluster, sehingga data dalam satu cluster memiliki karakteristik yang lebih mirip satu sama lain dibandingkan dengan data di cluster lain [16].

Dalam implementasinya, algoritma K-Means mengikuti langkah-langkah sistematis untuk menemukan pengelompokan optimal dengan meminimalkan jarak antara data dan pusat cluster (centroid). Langkah-langkah dasar dalam algoritma K-Means adalah sebagai berikut [15]:

- Menentukan jumlah cluster (k)
Untuk menentukan nilai k terbaik, dilakukan analisis dengan mengkombinasikan dua metode, yaitu Elbow Method dan Silhouette Score, dengan rentang percobaan k dari 2 hingga 10.
- Inisialisasi vektor centroid cluster secara acak.
- Untuk setiap vektor data, hitung jarak antara data dan masing-masing centroid cluster. Tentukan cluster untuk setiap data berdasarkan jarak terdekat dengan centroid yang dihitung menggunakan rumus (5).

$$d_{(x,y)} = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (5)$$

- Hitung pusat cluster yang baru berdasarkan keanggotaan cluster yang baru. Pusat cluster dihitung dengan cara menghitung rata-rata posisi objek dalam cluster tersebut.
- Hitung kembali jarak setiap objek dengan pusat cluster yang baru, hingga cluster tidak berubah. Jika tidak ada perubahan pada posisi centroid, maka proses pengklasteran selesai.

2.3 Elbow Method

Elbow Method mengevaluasi Within-Cluster Sum of Squares (WCSS) untuk berbagai nilai K dan menentukan titik optimal (elbow point), yaitu ketika penurunan WCSS mulai melambat [26]. WCSS dihitung menggunakan persamaan (6).

$$WCSS = \sum_{i=1}^k \sum_{x \in C_i} \|x - \mu_i\|^2 \quad (6)$$

Keterangan:

- x = Titik data dalam cluster C_i
 μ_i = Centroid cluster C_i
 $\|x - \mu_i\|^2$ = Jarak Euclidean antara titik dan centroid

2.4 Silhouette Score

Silhouette Score merupakan metrik evaluasi dalam clustering yang mengukur sejauh mana suatu titik data berada dalam cluster-nya dibandingkan dengan cluster lain. Nilainya berkisar antara -1 hingga 1, di mana nilai mendekati 1 menunjukkan bahwa titik data berada dalam cluster yang tepat, nilai mendekati 0 mengindikasikan bahwa titik berada di batas antara dua cluster, dan nilai negatif menunjukkan bahwa titik lebih dekat ke cluster lain dibandingkan dengan cluster-nya sendiri [27]. Silhouette Score untuk setiap titik data dihitung menggunakan rumus (7).

$$S(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))} \quad (7)$$



Keterangan:

a(i) = Rata-rata jarak antara titik i dan semua titik lain dalam klasternya sendiri (cohesion).

b(i) = Rata-rata jarak antara titik i dan titik-titik dalam kluster terdekat lainnya (separation).

2.5 Association Rules

Association rules adalah teknik data mining yang digunakan untuk menemukan hubungan atau pola tersembunyi antar item dalam kumpulan data [17]. Teknik ini sering digunakan pada data transaksi untuk mengidentifikasi pola pembelian barang yang sering muncul bersamaan. Hasilnya berupa aturan dalam format "jika-X-maka-Y," yang dievaluasi menggunakan metrik seperti [7]:

a. Support

Mengukur seberapa sering kombinasi item X dan Y muncul dalam dataset. Hal ini dihitung dengan rumus (8):

$$\text{Support}(X \rightarrow Y) = \frac{\text{Jumlah transaksi yang mengandung } (XY)}{\text{Jumlah total transaksi}} \quad (8)$$

b. Confidence

Mengukur probabilitas bahwa item Y muncul dalam transaksi yang sudah mengandung X. Hal ini dihitung dengan rumus (9):

$$\text{Confidence}(X \rightarrow Y) = \frac{\text{Jumlah transaksi yang mengandung } (XY)}{\text{Jumlah transaksi yang mengandung } X} \quad (9)$$

c. Lift

Mengukur kekuatan hubungan antara X dan Y, dibandingkan dengan kemungkinan keduanya muncul secara independen. Hal ini dihitung dengan rumus (10):

$$\text{Confidence}(X \rightarrow Y) = \frac{\text{Jumlah transaksi yang mengandung } (XY)}{\text{Jumlah transaksi yang mengandung } X} \quad (10)$$

2.6 Algoritma FP-Growth

FP-Growth (Frequent Pattern Growth) adalah algoritma data mining yang digunakan untuk menemukan pola frekuensi tinggi dalam dataset secara efisien [8]. Algoritma ini bekerja dengan membangun struktur data FP-Tree (Frequent Pattern Tree), yang merepresentasikan data transaksi dalam bentuk pohon, memungkinkan eksplorasi pola tanpa perlu menghasilkan kandidat kombinasi seperti pada algoritma Apriori. Algoritma FP-Growth dibagi menjadi tiga langkah utama, yaitu [10]:

a. Tahap pembangkitan conditional pattern base.

b. Tahap pembangkitan conditional FP-tree.

c. Tahap pencarian frequent itemset.

3. HASIL DAN PEMBAHASAN

3.1 Pengumpulan Data

Data penelitian diperoleh dari sistem penjualan internal perusahaan XYZ, berupa 90.046 transaksi penjualan obat dari Januari hingga Desember 2023. Data tersebut terdiri dari 23 atribut yang mencakup informasi transaksi, produk, dan pelanggan. Beberapa atribut utama yang terdapat dalam data antara lain KODE_JUAL (kode penjualan), NODOKJDI (nomor transaksi), KODE_PELANGGAN (kode pelanggan), KODE_KOTA (kode kota), NAMA_LANG (nama lengkap pelanggan), ALAMAT (alamat pengiriman), KODEPROD (kode produk), NAMAPROD (nama produk), UNIT (kuantitas produk), HARGA (harga satuan produk), TGLDOK (tanggal transaksi), Klaim Disc Prins (diskon harga), dan atribut lainnya. Data ini digunakan sebagai dasar analisis untuk menemukan pola asosiasi antar produk berdasarkan segmentasi pelanggan.

3.2 Preprocessing Data

Tahap preprocessing data dilakukan untuk mempersiapkan data agar siap digunakan dalam analisis lebih lanjut. Proses preprocessing ini terdiri dari beberapa tahap, yaitu seleksi data, pembersihan data, transformasi data, dan normalisasi data, yang bertujuan untuk memastikan bahwa data yang digunakan relevan, konsisten, dan dalam format yang sesuai.

3.2.1 Seleksi Data

Tahap preprocessing data diawali dengan melakukan seleksi data yang akan digunakan dalam penelitian. Atribut yang digunakan mencakup ID_Transaksi (nomor transaksi), ID_Pelanggan (kode pelanggan), Kode_Prod (kode produk), Tgl_Transaksi (tanggal transaksi), Jumlah (kuantitas produk), Harga (harga satuan produk), dan Diskon (diskon harga). Atribut Jumlah, Harga, dan Diskon digabungkan menjadi satu tabel bernama Total_Harga, yang merepresentasikan total nilai transaksi. Hasil dari proses seleksi data dapat dilihat pada Tabel 1.

Tabel 1. Hasil Seleksi Data

ID_Transaksi	ID_Pelanggan	Kode_Prod	Tgl_Transaksi	Jumlah	Harga	Diskon	Total_Harga
2302061000034	DMIAPJAYA	ERECTP	1/2/2023	70	15400	0	1078000
2302061000034	DMIAPJAYA	MS221	1/2/2023	25	15750	0	393750
2302061000034	DMIAPJAYA	TFCTMO	1/2/2023	34	5542	0	188428
2302061000034	DMIAPJAYA	TFTFMP	1/2/2023	94	22500	0	2115000
2302061000029	KARAPALYA	EPEPX1	1/2/2023	20	11628	0	232560
...
2302061028392	UTJAPHASAN	PP09-01420	12/30/2023	20	164273	887074	2398386

3.2.2 Pembersihan Data

Tahap pembersihan data dilakukan untuk mengatasi nilai yang hilang (missing values) dan menghapus data duplikasi. Setelah tahap pembersihan data, jumlah data transaksi berkurang dari 90.046 menjadi 88.463. Hasil dari proses pembersihan data dapat dilihat pada Tabel 2.

Tabel 2. Hasil Pembersihan Data

ID_Transaksi	ID_Pelanggan	Kode_Prod	Tgl_Transaksi	Total_Harga
2302061000034	DMIAPJAYA	ERECTP	1/2/2023	1078000
2302061000034	DMIAPJAYA	MS221	1/2/2023	393750
2302061000034	DMIAPJAYA	TFCTMO	1/2/2023	188428
2302061000034	DMIAPJAYA	TFTFMP	1/2/2023	2115000
2302061000029	KARAPALYA	EPEPX1	1/2/2023	232560
...
2302061028392	UTJAPHASAN	PP09-01420	12/30/2023	2398386

3.2.3 Transformasi Data

Pada tahap selanjutnya, dilakukan transformasi atribut yang ada menjadi nilai kriteria RFM. Hasil transformasi data RFM dapat dilihat pada Tabel 3.

Tabel 3. Hasil Transformasi Data

No.	ID_Pelanggan	Recency	Frequency	Monetary
1	ARMAPBAGUS	27	9	5233779
2	ARMAPFITRI	103	22	52413343
3	ARMAPGRACE	271	4	18460184
4	ARMAPLINDO	107	16	87455025
5	ARMAPMISKA	26	32	80927424
...
1052	UTJRSATHAY	12	12	48177278

3.2.4 Normalisasi Data RFM

Normalisasi data RFM dilakukan untuk menyelaraskan skala nilai dari tiga kriteria utama, yaitu Recency (R), Frequency (F), dan Monetary (M), sehingga dapat dianalisis secara setara tanpa bias terhadap salah satu kriteria. Setiap kriteria memiliki rentang nilai yang berbeda. Sebagai contoh, nilai Recency dalam hitungan hari dapat jauh lebih kecil atau besar dibandingkan nilai Frequency atau Monetary. Hasil normalisasi data RFM dapat dilihat pada Tabel 4.

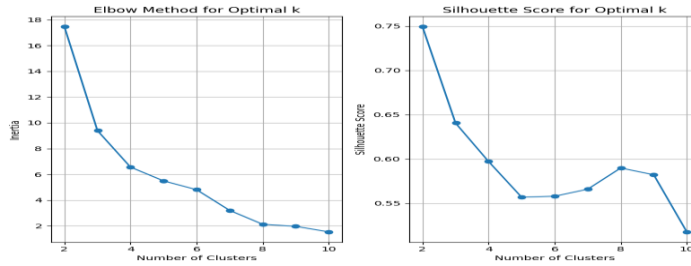
Tabel 4. Hasil Normalisasi Data RFM

No.	ID_Pelanggan	Recency	Frequency	Monetary
1	ARMAPBAGUS	0.071225	0.003210	0.180870
2	ARMAPFITRI	0.287749	0.008427	0.190581
3	ARMAPGRACE	0.766382	0.001204	0.183592
4	ARMAPLINDO	0.299145	0.006019	0.197794
5	ARMAPMISKA	0.068376	0.012440	0.196450
...
1052	UTJRSATHAY	0.162393	0.004414	0.189709

3.3 Clustering dengan Algoritma K-Means

Tahap clustering ini bertujuan untuk melakukan segmentasi pelanggan berdasarkan hasil transformasi data RFM, di mana setiap cluster merepresentasikan kelompok pelanggan dengan karakteristik yang serupa. Penelitian ini menggunakan algoritma K-Means untuk mengelompokkan data yang telah melalui tahap preprocessing.

Langkah awal dalam proses ini adalah menentukan jumlah cluster optimal dengan menggunakan kombinasi Elbow Method dan Silhouette Score. Hasil analisis jumlah cluster optimal disajikan pada Gambar 2.



Gambar 2. Grafik Cluster (k) Terbaik

Dapat dilihat pada Gambar 2, grafik Elbow (kiri) menunjukkan penurunan nilai inertia yang signifikan hingga k=4. Setelah k=4, penurunan melambat dan menjadi tidak signifikan, membentuk "sudut siku" (elbow point), yang menunjukkan jumlah cluster optimal. Pada grafik Silhouette (kanan), skor tertinggi (sekitar 0.60) juga tercapai pada k=4. Setelah k=4, nilai Silhouette Score menurun, menandakan kualitas klusterisasi berkurang. Berdasarkan hasil dari kedua grafik tersebut, jumlah cluster optimal yang dipilih adalah k=4, karena menghasilkan pembagian data yang paling efisien dan berkualitas.

Selanjutnya, segmentasi pelanggan dilakukan dengan menggunakan algoritma K-Means dengan nilai k=4. Kemudian, perhitungan nilai rata-rata RFM dilakukan untuk masing-masing cluster guna menentukan simbol peringkat (rank) berdasarkan nilai rata-rata yang diperoleh. Simbol (↑) digunakan untuk menunjukkan nilai rata-rata yang lebih tinggi, sedangkan simbol (↓) digunakan untuk nilai rata-rata yang lebih rendah. Hasil segmentasi pelanggan, beserta simbol peringkatnya, dapat dilihat pada Tabel 5.

Tabel 5. Hasil Segmentasi Pelanggan

Cluster	Anggota Cluster	Recency	Frequency	Monetary	Simbol Rank RFM
1	634	11.99	41.55	95549529.79	R↓F↑M↑
2	83	289.36	2.69	3752698.08	R↑F↓M↓
3	105	154.18	6.03	14843525.49	R↑F↓M↓
4	230	69.40	10.02	17130687.33	R↓F↓M↓

Berdasarkan Tabel 5, setiap cluster pelanggan dapat dikategorikan sebagai berikut:

- Cluster 1 adalah pelanggan aktif dan loyal, dengan Recency rendah (baru transaksi), serta Frequency dan Monetary tinggi (sering membeli dengan pengeluaran besar). Mereka merupakan pelanggan utama yang memberikan kontribusi signifikan terhadap pendapatan perusahaan.
- Cluster 2 adalah pelanggan pasif dengan Recency tinggi (jarang transaksi) dan Frequency serta Monetary rendah (jarang membeli dan pengeluaran kecil). Mereka kurang menguntungkan dan membutuhkan strategi reaktivasi.
- Cluster 3 adalah pelanggan kurang aktif dengan Recency tinggi serta Frequency dan Monetary rendah. Mereka berada di antara kategori pasif dan potensial, menunjukkan aktivitas yang sedikit lebih tinggi dibandingkan Cluster 2, namun memerlukan dorongan untuk lebih aktif bertransaksi.
- Cluster 4 adalah pelanggan baru yang belum aktif, dengan Recency rendah (baru transaksi), namun Frequency dan Monetary rendah. Mereka adalah pelanggan baru yang memerlukan pendekatan khusus untuk meningkatkan loyalitas dan aktivitas pembelian.

3.4 Asosiasi dengan Algoritma FP-Growth

Sebelum tahap asosiasi dilakukan menggunakan algoritma FP-Growth, data perlu melalui tahap preprocessing untuk diubah ke dalam format transaksi-item pada setiap cluster. Format ini disusun berdasarkan data dari masing-masing cluster. Selanjutnya, transaksi yang hanya mengandung satu itemset dihapus, karena salah satu syarat dalam penentuan pola association rules adalah transaksi harus memiliki setidaknya dua itemset. Selain itu, item yang duplikat dalam satu transaksi dihapus dengan menggabungkan item yang sama menjadi satu. Hasil dari proses ini, berupa data dalam format transaksi-item, ditampilkan pada Tabel 6.

Tabel 6. Hasil Preprocessing untuk Tiap Cluster

Cluster	ID Transaksi	Transaksi Obat
1	2302061000034	TFTFMP, ERECTP, MS221, TFCTMO
	2302061000029	ERELF6, ERGETM, EPEPX1, ERCCPS, EPPRN1

2	2302061028392	AP43003, PP09-01420, PP09-01020
	2302061000033	SDKTMS, CDMGT0, CDXTLDR, DPAVTC2, SDTNO1, ERECTP, ERGENO, CDMT02



Cluster	ID_Transaksi	Transaksi_Obat
	2302061000100	SDTCDC5, ERGENO, TPOTEM6

	2302061010171	ITGNS6, A0033
	2302061000022	SDPRCS3, SDBZCG
3	2302061000006	PBBONES, SLNFKS, PBECV1, SDTSC3, PBFTV1

	2302061018911	ERECTP, ERGENO
	2302061000028	ERIFS6, TPOBTM
4	2302061000039	STAXC5, TFPCT51

	2302061024861	PYPXMDS, SDMCC5, SDMDC5, SDMCC0

Setelah pengelompokan data transaksi pada setiap cluster, data transaksi ditransformasikan ke dalam bentuk boolean. Proses ini melibatkan pengisian atribut baru dengan nilai "TRUE" jika item atau produk yang sesuai dengan atribut tersebut terdapat dalam data transaksi, dan nilai "FALSE" jika tidak terdapat. Hasil transformasi data untuk cluster 1 hingga 5 disajikan pada Tabel 7, Tabel 8, Tabel 9, dan Tabel 10.

Tabel 7. Hasil Transformasi Data Transaksi Cluster 1

A0007	A0033	A0038	A0042	A0047	ALAEMFM	ALAEMFMG	...	TPUVXK
TRUE	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	...	FALSE
FALSE	TRUE	FALSE	FALSE	TRUE	FALSE	FALSE	...	FALSE
FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	...	FALSE
...
FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	...	FALSE

Tabel 8. Hasil Transformasi Data Transaksi Cluster 2

A0007	A0033	A0038	A0042	ALAEMFM	ALAHMFM	ALALCHS	...	TPOTM1B
FALSE	FALSE	FALSE	FALSE	TRUE	TRUE	FALSE	...	FALSE
FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	...	FALSE
FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	...	FALSE
...
FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	...	FALSE

Tabel 9. Hasil Transformasi Data Transaksi Cluster 3

A0007	A0033	A0038	A0042	A0047	ALAEMFM	ALAEMFMG	...	TPTRC3
FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	...	FALSE
FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	...	FALSE
TRUE	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	...	FALSE
...
FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	...	FALSE

Tabel 10. Hasil Transformasi Data Transaksi Cluster 4

A0007	A0033	A0038	A0042	A0047	ALAEMFM	ALAHMFM	...	TPTRZS6
FALSE	FALSE	TRUE	TRUE	FALSE	FALSE	FALSE	...	FALSE
FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	...	FALSE
FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	...	FALSE
...
FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	...	FALSE

Langkah terakhir dalam penelitian ini adalah melakukan analisis asosiasi pada data hasil segmentasi menggunakan algoritma FP-Growth. Analisis ini bertujuan untuk mengidentifikasi pola asosiasi antarproduk dalam masing-masing cluster. Parameter minimum support dan minimum confidence dipilih berdasarkan karakteristik jumlah transaksi dan atribut item dalam setiap cluster, sehingga menghasilkan pola asosiasi yang lebih bermakna dan tidak terlalu jarang. Berikut adalah hasil analisis asosiasi untuk masing-masing cluster.

3.4.1 Asosiasi Cluster 1

Pada Cluster 1, digunakan nilai minimum support sebesar 0.01 dan minimum confidence sebesar 0.7. Nilai minimum support ini dipilih berdasarkan distribusi transaksi dalam cluster, yang terdiri dari 13.376 transaksi dan 1.100 atribut item. Berdasarkan observasi, support tertinggi untuk satu item dalam cluster ini mencapai 0.06, sehingga nilai minimum support 0.01 dipilih agar tetap dapat menangkap pola asosiasi yang signifikan tanpa

kehilangan aturan yang memiliki frekuensi kemunculan cukup tinggi. Hasil analisis asosiasi pada cluster 1 dapat dilihat pada Tabel 11.

Tabel 11. Hasil Association Rules Cluster 1

No.	Antecedents	Consequents	Support	Confidence	Lift
1	ALGVSL	ALGVSM	0.011	0.754	30.004
2	TPOBPS6	TPOBPJ6	0.010	0.875	69.254
3	TPOBPJ6	TPOBPS6	0.010	0.828	69.254
4	ALAHMFM	ALAEMFM	0.010	0.820	47.709
5	A0038	A0007	0.031	0.829	21.785
6	A0007	A0038	0.031	0.809	21.785
7	A0042	A0007	0.014	0.778	20.451
8	A0042	A0038	0.014	0.766	20.619
9	A0038, A0042	A0007	0.011	0.774	20.332
10	A0007, A0042	A0038	0.011	0.762	20.499

Berdasarkan Tabel 11, terdapat 10 aturan asosiasi (rules) yang memenuhi kriteria confidence ≥ 0.7 . Hasil analisis menunjukkan bahwa seluruh aturan memiliki nilai lift > 1 , yang mengindikasikan hubungan yang positif antar item. Hal ini berarti bahwa kemunculan item dalam antecedents berpotensi meningkatkan kemungkinan kemunculan item dalam consequents, melebihi peluang kemunculan secara acak. Adapun penjelasan tabel sebagai berikut:

- Jika pelanggan Cluster 1 membeli ALTAMED GLOVE SIZE L (ALGVSL) maka akan membeli ALTAMED GLOVE SIZE M (ALGVSM) dengan support 0.011, confidence 0.754, dan lift 30.004.
- Jika pelanggan Cluster 1 membeli OBH TROP.PLUS ANAK STRAW 60 ML (TPOBPS6) maka akan membeli OBH TROP.PLUS ANAK JERUK 60 ML (TPOBPJ6) dengan support 0.010, confidence 0.875, dan lift 69.254.
- Jika pelanggan Cluster 1 membeli OBH TROP.PLUS ANAK JERUK 60 ML (TPOBPJ6) maka akan membeli OBH TROP.PLUS ANAK STRAW 60 ML (TPOBPS6) dengan support 0.010, confidence 0.828, dan lift 69.254.
- Jika pelanggan Cluster 1 membeli ALTAMED HEADLOOP MEDICAL FACE MASK (BLUE) (ALAHMFM) maka akan membeli ALTAMED EARLOOP MEDICAL FACE MASK (BLUE) (ALAEMFM) dengan support 0.010, confidence 0.820, dan lift 47.709.
- Jika pelanggan Cluster 1 membeli HUFAGRIPP BP - SYR (A0038) maka akan membeli HUFAGRIPP - SYR (A0007) dengan support 0.031, confidence 0.829, dan lift 21.785.
- Jika pelanggan Cluster 1 membeli HUFAGRIPP - SYR (A0007) maka akan membeli HUFAGRIPP BP - SYR (A0038) dengan support 0.031, confidence 0.809, dan lift 21.785.
- Jika pelanggan Cluster 1 membeli HUFAGRIPP PILEK - SYR (A0042) maka akan membeli HUFAGRIPP - SYR (A0007) dengan support 0.014, confidence 0.778, dan lift 20.451.
- Jika pelanggan Cluster 1 membeli HUFAGRIPP PILEK - SYR (A0042) maka akan membeli HUFAGRIPP BP - SYR (A0038) dengan support 0.014, confidence 0.766, dan lift 20.619.
- Jika pelanggan Cluster 1 membeli HUFAGRIPP BP - SYR (A0038) dan HUFAGRIPP PILEK - SYR (A0042) maka akan membeli HUFAGRIPP - SYR (A0007) dengan support 0.011, confidence 0.774, dan lift 20.332.
- Jika pelanggan Cluster 1 membeli HUFAGRIPP - SYR (A0007) dan HUFAGRIPP PILEK - SYR (A0042) maka akan membeli HUFAGRIPP BP - SYR (A0038) dengan support 0.011, confidence 0.762, dan lift 20.499.

3.4.2 Asosiasi Cluster 2

Pada Cluster 2, digunakan nilai minimum support sebesar 0.08 dan minimum confidence sebesar 0.7. Cluster ini memiliki jumlah transaksi yang jauh lebih sedikit dibanding Cluster 1, yaitu 128 transaksi dengan 243 atribut item. Support tertinggi untuk satu item dalam cluster ini mencapai 0.1, sehingga nilai minimum support 0.08 dipilih agar aturan asosiasi yang dihasilkan tetap relevan dan tidak terlalu banyak mengabaikan kombinasi item yang memiliki frekuensi kemunculan signifikan. Hasil analisis asosiasi pada cluster 2 dapat dilihat pada Tabel 12.

Tabel 12. Hasil Association Rules Cluster 2

No.	Antecedents	Consequents	Support	Confidence	Lift
1	A0038	A0007	0.086	0.917	9.778
2	A0007	A0038	0.086	0.917	9.778

Berdasarkan Tabel 12, terdapat 2 aturan asosiasi (rules) yang memenuhi kriteria confidence ≥ 0.7 . Hasil analisis menunjukkan bahwa seluruh aturan memiliki nilai lift > 1 , yang mengindikasikan hubungan yang positif antar item. Hal ini berarti bahwa kemunculan item dalam antecedents berpotensi meningkatkan kemungkinan kemunculan item dalam consequents, melebihi peluang kemunculan secara acak. Adapun penjelasan tabel sebagai berikut:



- a. Jika pelanggan Cluster 2 membeli HUFAGRIPP BP - SYR (A0038) maka mereka akan membeli HUFAGRIPP - SYR (A0007) dengan support 0.086, confidence 0.917, dan lift 9.778.
- b. Jika pelanggan Cluster 2 membeli HUFAGRIPP - SYR (A0007) maka mereka akan membeli HUFAGRIPP BP - SYR (A0038) dengan support 0.086, confidence 0.917, dan lift 9.778.

3.4.3 Asosiasi Cluster 3

Pada Cluster 3, digunakan nilai minimum support sebesar 0.03 dan minimum confidence sebesar 0.7. Cluster ini memiliki 347 transaksi dengan 495 atribut item. Support tertinggi untuk satu item dalam cluster ini mencapai 0.09, sehingga nilai minimum support 0.03 dipilih untuk memungkinkan deteksi pola asosiasi yang lebih luas tanpa kehilangan aturan penting. Hasil analisis asosiasi pada cluster 3 dapat dilihat pada Tabel 13.

Tabel 13. Hasil Association Rules Cluster 3

No.	Antecedents	Consequents	Support	Confidence	Lift
1	ECMPCD	ECMPC6	0.043	0.882	14.580
2	ECMPC6	ECMPCD	0.043	0.714	14.580
3	SLFLT1	SLCMS1	0.035	0.800	14.611

Berdasarkan Tabel 13, terdapat 3 aturan asosiasi (rules) yang memenuhi kriteria confidence ≥ 0.7 . Hasil analisis menunjukkan bahwa seluruh aturan memiliki nilai lift > 1 , yang mengindikasikan hubungan yang positif antar item. Hal ini berarti bahwa kemunculan item dalam antecedents berpotensi meningkatkan kemungkinan kemunculan item dalam consequents, melebihi peluang kemunculan secara acak. Adapun penjelasan tabel sebagai berikut:

- a. Jika pelanggan Cluster 3 membeli MAGIC POWER CRIMSON DESIRE 6 (ECMPCD) maka mereka akan membeli MAGIC POWER CLASSIC 6 (ECMPC6) dengan support 0.043, confidence 0.882, dan lift 14.580.
- b. Jika pelanggan Cluster 3 membeli MAGIC POWER CLASSIC 6 (ECMPC6) maka mereka akan membeli MAGIC POWER CRIMSON DESIRE 6 (ECMPCD) dengan support 0.043, confidence 0.714, dan lift 14.580.
- c. Jika pelanggan Cluster 3 membeli FOLATE KAP BOX/100 (SLFLT1) maka mereka akan membeli CAFOSOL (SLCMS1) dengan support 0.035, confidence 0.800, dan lift 14.611.

3.4.4 Asosiasi Cluster 4

Pada Cluster 4, digunakan nilai minimum support sebesar 0.03 dan minimum confidence sebesar 0.7. Cluster ini memiliki jumlah transaksi yang cukup besar, yaitu 1.170 transaksi dengan 753 atribut item. Support tertinggi untuk satu item dalam cluster ini mencapai 0.09, sehingga nilai minimum support 0.03 dipilih agar dapat menangkap pola asosiasi yang bermakna tanpa terlalu banyak aturan yang jarang terjadi. Hasil analisis asosiasi pada cluster 4 dapat dilihat pada Tabel 14.

Tabel 14. Hasil Association Rules Cluster 4

No.	Antecedents	Consequents	Support	Confidence	Lift
1	ECMPCD	ECMPC6	0.069	0.988	10.138
2	ECMPC6	ECMPCD	0.069	0.711	10.138
3	A0038	A0007	0.038	0.865	21.094
4	A0007	A0038	0.038	0.938	21.094

Berdasarkan Tabel 14, terdapat 4 aturan asosiasi (rules) yang memenuhi kriteria confidence ≥ 0.7 . Hasil analisis menunjukkan bahwa seluruh aturan memiliki nilai lift > 1 , yang mengindikasikan hubungan yang positif antar item. Hal ini berarti bahwa kemunculan item dalam antecedents berpotensi meningkatkan kemungkinan kemunculan item dalam consequents, melebihi peluang kemunculan secara acak. Adapun penjelasan tabel sebagai berikut:

- a. Jika pelanggan Cluster 4 membeli MAGIC POWER CRIMSON DESIRE 6 (ECMPCD) maka mereka akan membeli MAGIC POWER CLASSIC 6 (ECMPC6) dengan support 0.069, confidence 0.988, dan lift 10.138.
- b. Jika pelanggan Cluster 4 membeli MAGIC POWER CLASSIC 6 (ECMPC6) maka mereka akan membeli MAGIC POWER CRIMSON DESIRE 6 (ECMPCD) dengan support 0.069, confidence 0.711, dan lift 10.138.
- c. Jika pelanggan Cluster 4 membeli HUFAGRIPP BP - SYR (A0038) maka mereka akan membeli HUFAGRIPP - SYR (A0007) dengan support 0.038, confidence 0.865, dan lift 21.094.
- d. Jika pelanggan Cluster 4 membeli HUFAGRIPP - SYR (A0007) maka mereka akan membeli HUFAGRIPP BP - SYR (A0038) dengan support 0.038, confidence 0.938, dan lift 21.094.

4. KESIMPULAN

Penerapan algoritma K-Means berhasil mengidentifikasi empat cluster pelanggan dengan menggunakan kombinasi Elbow Method dan Silhouette Score, berdasarkan analisis RFM (Recency, Frequency, dan Monetary). Cluster 1 terdiri dari pelanggan aktif dan loyal dengan Recency rendah serta Frequency dan Monetary tinggi, yang



memberikan kontribusi besar terhadap pendapatan perusahaan. Cluster 2 mencakup pelanggan pasif dengan Recency tinggi serta Frequency dan Monetary rendah, yang saat ini kurang menguntungkan dan membutuhkan strategi reaktivasi. Cluster 3 adalah pelanggan kurang aktif dengan karakteristik serupa Cluster 2, yang menunjukkan sedikit lebih banyak aktivitas dibandingkan Cluster 2, namun masih memerlukan dorongan untuk lebih aktif bertransaksi. Sementara itu, Cluster 4 adalah pelanggan baru yang belum aktif, dengan Recency rendah namun Frequency dan Monetary juga rendah, yang membutuhkan upaya lebih lanjut untuk membangun loyalitas. Selanjutnya, analisis pola pembelian dilakukan pada setiap cluster menggunakan algoritma FP-Growth. Pada Cluster 1, ditemukan 10 aturan asosiasi signifikan dengan nilai minimum support 0,01 dan confidence 0,7, di mana seluruh aturan memiliki lift > 1, menunjukkan hubungan antar item yang kuat. Pada Cluster 2, dengan parameter minimum support 0,08 dan confidence 0,7, ditemukan 2 aturan asosiasi yang relevan. Analisis pada Cluster 3 menggunakan minimum support 0,03 dan confidence 0,7 menghasilkan 3 aturan signifikan, sedangkan pada Cluster 4 ditemukan 4 aturan asosiasi yang relevan dengan parameter yang sama. Hasil penelitian ini menunjukkan bahwa setiap cluster memiliki pola pembelian yang unik dan memerlukan strategi pemasaran yang berbeda. Untuk mempertahankan kontribusi pendapatan yang tinggi, strategi retensi dan loyalitas harus difokuskan pada Cluster 1. Sedangkan untuk Cluster 2, pendekatan reaktivasi yang agresif diperlukan, sementara Cluster 3 dan Cluster 4 memerlukan upaya untuk meningkatkan aktivitas dan loyalitas pembelian. Dengan memahami karakteristik dan pola pembelian setiap cluster, perusahaan dapat merancang strategi pemasaran yang lebih efektif untuk memaksimalkan potensi pendapatan dari setiap segmen pelanggan. Penelitian selanjutnya dapat mempertimbangkan penggunaan algoritma lain dalam proses clustering dan analisis asosiasi sebagai bahan perbandingan, guna mengevaluasi efektivitas yang lebih baik.

REFERENCES

- [1] World Health Organization, "Global Health Observatory Data Repository." Accessed: Jan. 01, 2025. [Online]. Available: <https://www.who.int/data/gho>
- [2] IQVIA, "The Global Use of Medicine in 2020." Accessed: Jan. 01, 2025. [Online]. Available: <https://www.iqvia.com/>
- [3] P. Kotler and K. L. Keller, *Marketing Management*, 15th ed. Harlow: Pearson Education, 2020.
- [4] E. Anggraiani, Satibi, and L. R. Fauzi, "Analisis Faktor-Faktor yang Mempengaruhi Rencana Kebutuhan Obat dan E-Procurement di Rumah Sakit di Indonesia," *Majalah Farmaseutik*, vol. 19, no. 2, pp. 214–220, 2023, doi: 10.22146/farmaseutik.v19i2.77944.
- [5] F. Sulianta and E. Prayogo, "Aturan Asosiasi Menggunakan Algoritma Apriori Untuk Menciptakan Strategi Pemasaran Pada Apotek," *Jurnal Teknik Elektro dan Informatika*, vol. 19, no. 1, pp. 83–89, 2024.
- [6] R. Yogasuwara, Ferdiansyah, Andri, and L. Atika, "Implementasi Algoritma Frequent Growth (FP-Growth) Menentukan Asosiasi Antar Produk," *Jurnal Sistem Komputer dan Informatika (JSON)*, vol. 4, no. 1, pp. 165–171, Sep. 2022, doi: 10.30865/json.v4i1.4894.
- [7] A. Wilrose, M. Afdal, S. Monalisa, and M. R. Munzir, "Penerapan Algoritma FP-Growth untuk Menentukan Strategi Promosi Berdasarkan Waktu dan Pembelian Produk," *Building of Informatics, Technology and Science (BITS)*, vol. 5, no. 1, pp. 104–113, Jun. 2023, doi: 10.47065/bits.v5i1.3577.
- [8] H. D. Ariyantini, D. Malita Puspita, and A. Triyono, "Implementasi Algoritma Fp-Growth Untuk Rekomendasi Produk Di Toko Lm Mart," *Jurnal Ilmu Komputer An Nur*, vol. 4, no. 1, pp. 1–12, Jan. 2024.
- [9] M. R. A. Utama, R. Umar, and A. Yudhana, "Penerapan Algoritma Fp-Growth Untuk Penentuan Pola Pembelian Transaksi Penjualan Pada Toko Kgs Rizky Motor," *Jurnal DINAMIK*, vol. 25, no. 1, pp. 20–28, Jan. 2020.
- [10] R. Fauzi, A. W. Aranski, Nopriadi, and E. Hutabri, "Implementasi Data Mining Pada Penjualan Pakaian dengan Algoritma FP-Growth," *JURIKOM (Jurnal Riset Komputer)*, vol. 10, no. 2, pp. 436–445, Apr. 2023, doi: 10.30865/jurikom.v10i2.5795.
- [11] Y. Husain, E. D. Oktaviyani, and S. Christina, "Analisis Perbandingan Algoritma Apriori, FP-Growth, dan Eclat dalam Menemukan Pola Pembelian Konsumen," *KONSTELASI: Konvergensi Teknologi dan Sistem Informasi*, vol. 3, no. 2, pp. 231–243, Dec. 2023.
- [12] I. Musdalifah and A. Jananto, "Analisis Perbandingan Algoritma Apriori Dan FP-Growth Dalam Pembentukan Pola Asosiasi Keranjang Belanja Pelanggan," *Jurnal Ilmiah Komputer*, vol. 18, no. 2, pp. 175–184, Aug. 2022.
- [13] D. Rachmawati, Y. Cahyana, E. E. Awal, and S. Faisal, "Perbandingan Algoritma Apriori dan Algoritma FP-Growth dalam Menentukan Pola Penjualan Pupuk," *JURNAL RESISTOR*, vol. 21, no. 1, pp. 21–31, Apr. 2024, [Online]. Available: <https://s.id/jurnalresistor>
- [14] A. T. Suseno, A. R. Naufal, and M. Al Amin, "Market Based Analysis Sebagai Peningkatan Penjualan Produk Menggunakan Algoritma K-Medoids Dan Fp-Growth," *Jurnal Teknik Informasi dan Komputer (Tekinkom)*, vol. 5, no. 2, pp. 301–308, Dec. 2022, doi: 10.37600/tekinkom.v5i2.646.
- [15] S. A. Perdana, S. F. Florentin, and A. Santoso, "Analisis Segmentasi Pelanggan Menggunakan K-Means Clustering Studi Kasus Aplikasi Alfagift," *Sebatik*, vol. 26, no. 2, pp. 446–457, Dec. 2022, doi: 10.46984/sebatik.v26i2.2134.
- [16] A. Pramudiansyah and H. Munte, "Segmentasi Pelanggan Menggunakan Algoritma K-Means Berdasarkan Model Recency Frequency Monetary," *Jurnal Ilmiah Ilmu Komputer*, vol. 7, no. 2, pp. 6–19, 2021, [Online]. Available: <http://ejournal.fikom-unasman.ac.id>
- [17] A. H. Fazri, A. Muhammad, and P. A. W. Purnama, "Algoritma K-Means Cluster Untuk Segmentasi Pelanggan," *CBIS JOURNAL*, vol. 11, no. 02, pp. 42–51, 2023, [Online]. Available: <http://ejournal.upbatam.ac.id/index.php/cbis>
- [18] W. A. Silamantha and K. Hadiono, "Analisis RFM dan K-Means Clustering untuk Segmentasi Pelanggan pada PT. Sanutama Bumi Arto," *KESATRIA: Jurnal Penerapan Sistem Informasi*, vol. 5, no. 3, pp. 1297–1305, Jul. 2024.



- [19] R. B. Ardi, F. Ely Nastiti, and S. Sumarlinda, “Algoritma K-Means Clustering Untuk Segmentasi Pelanggan (Studi Kasus: Fashion Viral Solo),” *INFOTECH journal*, vol. 9, no. 1, pp. 124–131, May 2023, doi: 10.31949/infotech.v9i1.5214.
- [20] D. Hartama, Wanayumini, and I. S. Damanik, “Pengelompokan Algoritma K-Means dan K-Medoid Berdasarkan Lokasi Daerah Rawan Bencana di Indonesia dengan Optimasi Elbow, DBI, dan Silhouette,” *Building of Informatics, Technology and Science (BITS)*, vol. 6, no. 2, pp. 1151–1158, 2024, doi: 10.47065/bits.v6i2.5851.
- [21] S. Genjang Setyorini, E. K. Sari, L. Rahma Elita, and S. A. Putri, “Analisis Keranjang Pasar Menggunakan Algoritma K-Means dan FP-Growth pada PT. Citra Mustika Pandawa,” *MALCOM*, vol. 1, no. 1, pp. 41–46, Apr. 2021.
- [22] T. Ayunita Pertiwi, M. Afdal, R. Novita, and Mustakim, “Penerapan Algoritma K-Medoids dan FP-Growth dalam Penentuan Pola Kombinasi Produk Berdasarkan Hasil Segmentasi Pelanggan,” *Technology and Science (BITS)*, vol. 6, no. 2, pp. 721–732, Sep. 2024, doi: 10.47065/bits.v6i2.5268.
- [23] D. Anjainah and S. Monalisa, “Analisis Rekomendasi Produk Berdasarkan Segmentasi Pelanggan Menggunakan Algoritma DBSCAN dan FP-Growth,” *Techno.COM*, vol. 21, no. 4, pp. 948–956, Nov. 2022.
- [24] A. Rahmah and M. Afdal, “Segmentasi Pelanggan Menggunakan Fuzzy C-Means dan FP-Growth Berdasarkan Model LRFM untuk Rekomendasi Produk,” *JURNAL MEDIA INFORMATIKA BUDIDARMA*, vol. 8, no. 3, pp. 1218–1227, Jul. 2024, doi: 10.30865/mib.v8i3.7737.
- [25] F. Pangestu, N. Yasin, R. C. Hasugian, and Yunita, “Penerapan Algoritma K-Means Untuk Mengklasifikasi Data Obat,” *Jurnal Sisfokom (Sistem Informasi dan Komputer)*, vol. 12, no. 1, pp. 53–62, Mar. 2023, doi: 10.32736/sisfokom.v12i1.1461.
- [26] A. Rajsya, Purnawansyah, and A. R. Rachman, “Rancang Bangun Penerapan Metode Elbow Pada K-Means Untuk Clustering Data Persediaan Barang,” *Literatur Informatika & Komputer*, vol. 1, no. 4, pp. 395–403, 2024, doi: 10.33096/linier.vxix.xxxx.
- [27] Y. Hasan, “Pengukuran Silhouette Score dan Davies-Bouldin Index pada Hasil Cluster K-Means dan Dbscan,” *Kumpulan Artikel Karya Ilmiah Fakultas Ilmu Komputer*, vol. 06, no. 01, pp. 60–74, 2024.