



Computer Based Information System Journal

ISSN (Print): 2337-8794 | E- ISSN : 2621-5292
web jurnal : <http://ejournal.upbatam.ac.id/index.php/cbis>



PENERAPAN ALGORITMA APRIORI DAN K-MEANS UNTUK ANALISIS POLA PEMBELIAN PADA TOKO A8 ELECTRONIC & FURNITURE PIAYU

Tri Mey Wendelina Ambarita¹, Rika Harman²

Universitas Putera Batam, Indonesia

INFORMASI ARTIKEL

Diterima Redaksi: Januari 2025
Diterbitkan Online: Maret 2025

KATA KUNCI

Algoritma Apriori, K-Means, Pola Pembelian, Segmentasi Pelanggan

KORESPONDENSI

E-mail:
pb211510040@upbatam.ac.id
rika@puterabatam.ac.id

A B S T R A C T

A8 Electronic & Furniture Piayu Store faces challenges in identifying frequently purchased product combinations and arranging products for customer convenience. This research aims to analyze consumer purchasing patterns and group consumers based on their habits using the Apriori algorithm and K-Means. The Apriori algorithm identifies frequent itemsets, such as a 55% likelihood of purchasing an AC after buying a fan and a tendency to purchase a mic with a speaker. K-Means clustering, with an optimal configuration of 4 clusters (Davis-Bouldin score: 0.754), categorizes 54 items in cluster 0, 18 in cluster 1, 20 in cluster 2, and 28 in cluster 3. These insights are recommended for optimizing stock management, tailoring promotions, and improving customer service. The study demonstrates the potential of transaction data analysis to support strategic decision-making and business growth.

I. Latar Belakang

Peran teknologi informasi telah memberikan dampak signifikan terhadap banyak aspek kehidupan, termasuk industri perdagangan[1]. Semua bisnis, berapa pun ukurannya, harus memiliki kemampuan untuk mengambil keputusan bisnis yang tepat guna meningkatkan kepuasan pelanggan dan efektivitas operasional. Pemahaman menyeluruh tentang minat dan pola perilaku pelanggan sangat penting ketika membuat penilaian. Munculnya beberapa *platform e-commerce* telah mengubah cara konsumen berbelanja. Pelanggan kini

memiliki lebih banyak pilihan dan akses cepat terhadap informasi. Organisasi ritel harus memiliki pendekatan yang tepat untuk memahami dan memenuhi harapan pelanggan jika ingin berkembang[2].

Bisnis dapat mempelajari banyak hal tentang perilaku konsumen, tren pasar, dan preferensi produk dengan memeriksa data penjualan tersebut[3]. Salah satu pelaku pasar yang menawarkan beragam barang elektronik dan peralatan rumah tangga ini adalah Toko A8 Electronic & Furniture Piayu. Volume data transaksi yang dibuat oleh toko semakin

bertambah seiring dengan kemajuan ini. Setiap transaksi berisi data yang sangat berharga tentang kebiasaan membeli pelanggan. Namun demikian, banyak dari data ini yang sering kali tidak dianalisis secara menyeluruh, sehingga menggunakannya untuk menginformasikan keputusan bisnis mungkin tidak memberikan manfaat yang maksimal. Toko juga tidak memiliki metode yang efisien untuk menentukan produk mana yang sering dibeli bersama oleh pelanggan. Oleh karena itu, rencana pemasaran yang diterapkan terlalu luas dan tidak fokus, sehingga sulit untuk meningkatkan penjualan secara signifikan. Masalah lainnya adalah pengelolaan stok produk yang tidak efektif. Produk yang sering dibeli tidak selalu ditata di tempat yang berdekatan, sehingga menyulitkan pelanggan untuk berbelanja dengan nyaman. Hal ini mengurangi kemungkinan penjualan produk yang serupa dan menurunkan pengalaman berbelanja.

Cara yang lebih efektif dan efisien untuk mengatasi kelemahan analisis manual adalah melalui *data mining*. *Data mining* dapat memproses data transaksi dalam skala besar dan mengungkap pola tersembunyi yang sulit ditemukan secara manual dengan memanfaatkan algoritma seperti Apriori dan *K-Means*. Penerapan teknik *data mining* merupakan salah satu cara untuk memahami perilaku pembelian konsumen[2]. Toko dapat menemukan cara untuk meningkatkan penjualan, mempelajari lebih lanjut tentang produk yang biasa dibeli bersama oleh pelanggan, dan meningkatkan pemahaman mereka tentang perilaku pelanggan dengan menggunakan analisis data transaksi. Hasilnya, penggunaan *data mining* secara efektif dapat membantu proses pengambilan keputusan bisnis berbasis data yang lebih efisien.

Teknik *data mining* yang disebut algoritma Apriori digunakan untuk mencari pola korelasi antar produk dalam data transaksi. Algoritma ini dapat mengidentifikasi kumpulan item yang sering digunakan, atau kelompok produk yang sering dibeli bersama, dan memberikan aturan asosiasi yang membantu dalam memilih taktik pemasaran dan jajaran produk[4]. Sebaliknya, segmen konsumen dibuat menggunakan algoritma *K-Means* sesuai dengan pola pembeliannya. Pelanggan dapat dibagi menjadi beberapa subkelompok yang lebih homogen menggunakan pengelompokan. Pengecer dapat

menciptakan strategi pemasaran yang lebih terfokus, misalnya dengan menawarkan penawaran eksklusif kepada konsumen yang sering membeli dari kategori produk tertentu atau diskon eksklusif kepada segmen klien yang tidak melakukan pembelian secara rutin.

Analisis pola pembelian yang lebih menyeluruh dan mendalam dapat dilakukan dengan menggabungkan metode Apriori dan *K-Means*. Berdasarkan perilaku pembelian konsumen, dibagi menjadi beberapa kategori dengan algoritma *K-Means*. Algoritma Apriori kemudian digunakan untuk memeriksa tren pembelian di masing-masing segmen tersebut. Dengan menggabungkan kedua algoritma ini, Toko A8 Electronic & Furniture Piayu dapat memperoleh pemahaman yang lebih baik tentang perilaku konsumen, sehingga akan meningkatkan efektivitas operasional dan membantu mereka menciptakan strategi pemasaran yang lebih sukses.

II. Kajian Literatur

2.1. Clustering

Clustering adalah proses membagi suatu kumpulan data menjadi beberapa subset sehingga data pada setiap subset mempunyai arti yang bermakna, atau lebih spesifik lagi, mengelompokkan objek-objek yang sebanding ke dalam kelompok-kelompok yang berbeda. Mengorganisasikan data ke dalam beberapa grup atau *cluster* di mana *item* dalam satu grup memiliki sifat yang sama dengan objek lain dalam grup tersebut adalah tujuan utama *clustering* [5].

2.2. Association Rules

Terdapat dua metrik, yaitu *support* dan *confident* yang digunakan untuk menilai aturan asosiasi [6].

1. *Support* : Ini menampilkan frekuensi terjadinya *item* atau kelompok hal tertentu dalam setiap transaksi. Kesesuaian suatu *item* atau sekelompok item sebagai dasar perhitungan ditentukan oleh ukurannya *confidence*.
2. *Confidence* : Sejauh mana 2 *item* mempunyai hubungan bersyarat satu sama lain ditunjukkan oleh *confidence*. Ini mengukur seberapa sering *item* atau *itemset* B muncul ketika *itemset* A juga muncul.

Tahapan utama pada metode analisis asosiasi terdiri dari 2 tahap [7] yaitu:

- a. Menganalisis Pola Frekuensi Tinggi

Rumus untuk menentukan nilai support dari 1 item:

$$\text{Support (A)} = \frac{\sum \text{Transaksi Item A}}{\sum \text{Transaksi}}$$

Sementara itu, rumus untuk mencari nilai support 2 item adalah:

$$\text{Support (A} \rightarrow \text{B)} = \frac{\sum \text{Transaksi Item A dan B}}{\sum \text{Transaksi}}$$

- b. Pembentukan Aturan Sosiasi

Proses ini akan menghitung keyakinan (confidence) dengan menggunakan rumus:

$$\text{Confidence (B / A)} = \frac{\sum \text{Transaksi Item A dan B}}{\sum \text{Transaksi Item A}}$$

2.3. K-Means

Langkah-langkah pada algoritma *k-means* adalah sebagai berikut[8]:

1. Menentukan *cluster* (k) yang akan dibentuk dalam dataset
2. Mengatur nilai pusat (*Centroid*)
Nilai *centroid* awalnya ditentukan secara acak dan kemudian diiterasi menggunakan rumus berikut:

$$V_{ij} = \frac{1}{N_i} \sum_{k=0}^{N_i} X_{kj}$$

Keterangan:

- a. Untuk variabel ke-j, V_{ij} adalah rata-rata *centroid* dari *cluster* ke-I.
 - b. Indeks *cluster* adalah i dan k.
 - c. N_i adalah jumlah anggota *cluster* ke-i.
 - d. Indeks variabel dilambangkan dengan j.
 - e. Untuk *cluster* tersebut, X_{kj} adalah nilai data ke-k dari variabel ke-j.
3. Untuk setiap *record*, jarak terdekat ke *centroid* dihitung dengan menggunakan rumus Jarak *Euclidean* (*Euclidean Distance*) sebagai berikut:

$$De = \sqrt{(x_i - s_i)^2 + (y_i - t_i)^2}$$

Keterangan:

- a. De melambangkan jarak *Euclidean*.
- b. I melambangkan jumlah benda atau objek
- c. Koordinat benda adalah x, y.
- d. Koordinat pusat (*centroid*) adalah s, t.

4. Mengelompokkan objek berdasarkan seberapa dekat benda-benda tersebut dengan *centroid*.
5. Hingga nilai ideal tercapai, iterasi dari langkah 3 hingga 4 dilakukan berulang kali.

2.4. Algoritma Apriori

Metode sederhana untuk mengidentifikasi kumpulan item umum pada aturan asosiasi *Boolean* adalah algoritma apriori[9]. Konsep utama dari algoritma apriori adalah:

1. Mencari database transaksi untuk *frequent itemset*, yang merupakan kumpulan *item-item* dengan dukungan (*support*) minimal.
2. Menghapus *itemset* frekuensi rendah sesuai dengan tingkat dukungan minimal yang telah ditetapkan sebelumnya.

Membuat aturan untuk asosiasi antar *itemset* yang memenuhi keyakinan nilai *minimal database* adalah langkah berikutnya. Algoritma apriori menggunakan dua langkah utama untuk membuat kandidat *itemset*, yaitu sebagai berikut:

1. Langkah penggabungan (*JoinStep*)
Setiap *item* digabungkan dengan *item* lain selama prosedur hingga tidak ada lagi kombinasi yang dapat dibuat.
2. Pemangkasan (*PruneStep*)
Tahapan ini melibatkan penggabungan temuan *item* dan kemudian memangkasnya sesuai dengan aturan yang disebutkan sebelumnya.

2.5. Pola Pembelian

Faktor-faktor yang mempengaruhi perilaku pelanggan yaitu ada 5 fase perilaku konsumen yang terjadi selama proses pembelian yakni identifikasi kebutuhan, pencarian informasi, evaluasi informasi, pembelian, dan pasca pembelian. Budaya adalah faktor paling mendasar yang mempengaruhi perilaku dan keinginan, dan budaya mempunyai dampak yang mendalam dan luas terhadap perilaku konsumen[10].

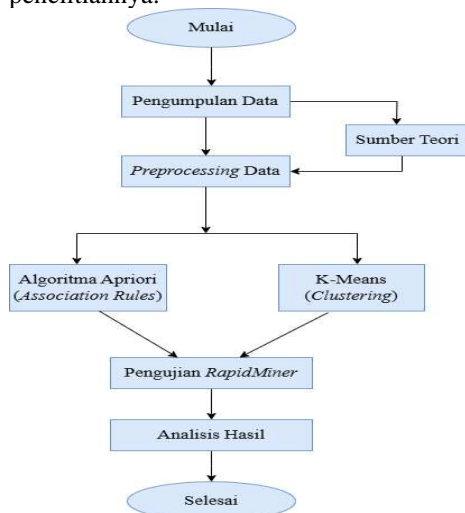
2.6. RapidMiner

RapidMiner adalah *platform* perangkat lunak yang tangguh untuk pembelajaran mesin dan penelitian data. *RapidMiner* menawarkan serangkaian alat untuk memodelkan, mengevaluasi, menerapkan, dan menyiapkan data[11]. Pada *RapidMiner* menggunakan sejumlah metode, termasuk metode deskriptif dan prediktif. *RapidMiner* adalah *open source*

dan dapat diakses oleh siapa saja yang menggunakan bahasa pemrograman Java [12].

III. Metodologi

Dalam mendukung penelitian ini diperlukan sebuah desain penelitian yang terstruktur dan sistematis untuk penyelesaian permasalahan pada penelitian. Berikut langkah-langkah yang dilakukan peneliti dalam proses penelitiannya:



Gambar 1 Desain Penelitian

Berdasarkan tahapan penelitian di atas, maka proses penelitian akan dijelaskan sebagai berikut:

1. Pengumpulan Data

Observasi langsung di Toko A8 Electronic & Furniture Piayu digunakan untuk mengumpulkan data yang diperlukan untuk penelitian ini, sehingga dapat dilakukan identifikasi langsung terhadap permasalahan yang sedang terjadi. Selanjutnya melakukan wawancara terarah untuk mendapatkan data dan informasi yang dibutuhkan peneliti.

2. Sumber Teori

Setelah pengumpulan data, teori-teori tersebut berkontribusi pada analisis data dengan menawarkan konteks yang relevan, memungkinkan peneliti untuk mendapatkan pemahaman dan interpretasi yang lebih dalam atas temuan penelitian. Sumber teoretis membantu peneliti dalam merumuskan argumen metodis dan logis selama proses pembuatan kerangka kerja, sehingga meningkatkan koherensi alur penelitian.

3. Preprocessing Data

<http://ejournal.upbatam.ac.id/index.php/cbis>

Preprocessing data digunakan untuk mengubah kumpulan data mentah menjadi kumpulan data yang siap diuji menggunakan algoritma. Pembersihan data, transformasi data, reduksi data, dan integrasi data adalah beberapa langkah yang membentuk *preprocessing*. Pembersihan data yang diterima pada tahap pra-pemrosesan data merupakan langkah yang dilakukan setelah pra-pemrosesan data. Data yang tidak diperlukan disaring selama proses pembersihan, meninggalkan data untuk digunakan dalam perhitungan selanjutnya.

4. Algoritma Apriori (Association Rule)

Algoritma Apriori bertujuan untuk mengidentifikasi korelasi antara produk yang sering dibeli dalam setiap kelompok pelanggan.

Adapun tahapan pada *association rule* adalah sebagai berikut:

a. Menganalisis Pola Frekuensi Tinggi

b. Pembentukan Aturan Asosiasi

5. Clustering (K-Means)

Setelah proses pembersihan, dilakukan seleksi pada data. Hasil pemilihan data merupakan data yang akan digunakan pada tahap *clustering* metode *K-Means* untuk perhitungan. Tujuan dari teknik analisis data *clustering* adalah untuk menyusun objek data ke dalam kelompok-kelompok yang sebanding menurut karakteristik atau atribut yang dimiliki. Menemukan struktur tersembunyi dalam data dan mengenali pola atau hubungan antar objek data adalah tujuan utama *clustering*. Tahapan dalam algoritma *k-means* yaitu:

a. Menentukan *cluster* (k) yang akan dibentuk dalam dataset

b. Mengatur nilai pusat (*Centroid*)

c. Jarak terdekat ke *centroid* dihitung dengan menggunakan rumus Jarak *Euclidean* (*Euclidean Distance*)

d. Mengelompokkan objek berdasarkan seberapa dekat benda-benda tersebut dengan *centroid*.

e. Hingga nilai ideal tercapai, iterasi dari langkah 3 hingga 4 dilakukan berulang kali.

6. Pengujian RapidMiner

Perangkat lunak *RapidMiner* adalah *software* yang digunakan peneliti untuk menguji hasil perhitungan yang didapatkan pada tahapan di atas.

7. Analisis Hasil

Proses analisis hasil pengujian dilakukan untuk memahami pola dan cluster yang dihasilkan oleh algoritma, menilai efektivitas metodologi yang digunakan, memastikan hasil selaras dengan tujuan penelitian, dan menyajikannya dengan cara yang mudah dipahami.

IV. Pembahasan

4.1. Hasil Penelitian

Pada bagian sub bab ini, peneliti akan menjelaskan tentang hasil penelitian yang sudah dilakukan. Adapun hasil penelitian yang didapat dari penelitian ini adalah sebagai berikut:

1. Jika membeli Kipas maka 55% akan membeli Kulkas. Hasil ini diperoleh berdasarkan perhitungan itemset-2 yang telah memenuhi nilai *support* 10% dan nilai *confidence* 55%.
2. Jika membeli Speaker maka 55% akan membeli Mic.
3. Berdasarkan hasil *Cluster Distance Performance*, diperoleh bahwa jarak terdekat data ke pusat cluster adalah dengan mengelompokkan data menjadi 4 *cluster* dengan Davis Bouldin = 0,754.
4. Berdasarkan perhitungan dengan menggunakan 4 kali iterasi dihasilkan *cluster* 0 terdiri dari 54 items, *cluster* 1 terdapat 18 items, *cluster* 2 terdapat 20 item, dan *cluster* 3 terdapat 28 items.

4.2. Pembahasan

Menginterpretasikan hasil penerapan algoritma Apriori dan *K-Means* pada data transaksi toko A8 Elektronik dan *Furniture* Piayu menjadi tujuan utama analisis. Penjelasan pola asosiasi yang ditemukan oleh algoritma Apriori memulai proses pembahasan. Temuan analisis ini menjelaskan produk-produk yang sering dibeli konsumen secara bersamaan, seperti beberapa barang furnitur menarik dan perangkat listrik tertentu. Peneliti menggunakan nilai *support minimum* yaitu 10% dan nilai *confidence* adalah 55%. Selain itu, berdasarkan kebiasaan pembelian konsumen, pelanggan dikelompokkan ke dalam beberapa segmen menggunakan hasil algoritma *K-Means*. Setiap *cluster* menunjukkan ciri-ciri yang berbeda, misalnya pelanggan yang minat utamanya adalah furnitur, elektronik, atau keduanya.

4.3. Pembahasan Algoritma Apriori

4.3.1. Data Transaksi Pembelian

Berikut adalah data transaksi pada Toko A8 Electronic & Furniture Piayu yang digunakan peneliti dalam penelitian ini:

Tabel 1 Data Transaksi

Transaksi	Item Pembelian
1	AC,TV, Magic Com, Setrika
2	AC, Kompom
3	Mixer, AC, Kulkas, Air Coler, Magic Com
4	Speaker, Kipas, Magic Com, Kulkas
5	Kasur, AC, Kipas, Magic Com, Kulkas
6	AC, Magic Com, Mixer
7	Magic Com, Kipas, SmartTV, Braket, Kulkas, Blender
8	Kipas, AC
9	Kulkas, Setrika, AC
10	Mesin Cuci, Selang, Magic Com, AC
...	
...	
119	Magic Com, Mesin Cuci, Rak Piring, Showchase
120	Mesin Cuci, AC, Dispenser, Kompom, Setrika, Magic Com

4.3.2. Tabulasi Penjualan

Bentuk tabulasi penjualan yang terbentuk adalah sebagai berikut:

Tabel 2 Tabulasi Data

Transaksi	AC	TV	Magic Com	Setrika	Kompom	Mixer			Bantal	Rak Piring
1	1	1	1	1	0	0	...	...	0	0
2	1	0	0	0	1	0	...	...	0	0
3	1	0	1	0	0	1	...	...	0	0
4	0	0	0	0	0	0	...	...	0	0
5	1	0	1	0	0	0	...	...	0	0
6	1	0	1	0	0	1	...	...	0	0
7	0	0	1	0	0	0	...	...	0	0
8	1	0	0	0	0	0	...	...	0	0
9	1	0	0	1	0	0	...	...	0	0
10	1	0	1	0	0	0	...	...	0	0
...	...	...	...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...	...	...	...

Transaksi	AC	TV	Magic Com	Setrika	Kompom	Mixer			Bantal	Rak Piring
119	0	0	1	0	0	0	...	...	0	1
120	1	0	1	1	1	0	...	...	0	0

4.3.3. Pemilihan Data Seleksi

Data yang diambil peneliti adalah berdasarkan hasil *random sampling* yaitu sebanyak 120 data dari data penjualan selama 3 bulan.

4.3.3.1. Menentukan Frequent Itemset

Menemukan objek atau komponen yang paling sering muncul dalam kumpulan data dilakukan melalui penggunaan analisis frekuensi tinggi. Penelitian ini sering digunakan untuk mengidentifikasi korelasi, hubungan, atau pola hubungan antar item data transaksi. Analisis frekuensi tinggi dilakukan untuk mendapatkan algoritma yang efisien dalam menentukan itemset yang pertama, kedua, dan ketiga. Peneliti terlebih dahulu mengubah data transaksi ke dalam tabel tabulasi, yang kemudian digunakan sebagai bahan pengujian dengan *software RapidMiner* untuk memudahkan mendapatkan nilai dari *itemset*.

4.3.3.2. Analisis Pola Frekuensi Tinggi

Tahapan ini peneliti menentukan kombinasi *item* yang memenuhi syarat *minimum* dari nilai *support*. Nilai *support* ditentukan dengan menggunakan rumus berikut:

$$\text{Support (A)} = \frac{\sum \text{Transaksi Item A}}{\sum \text{Transaksi}}$$

1. Pembentukan Itemset 1

1. *Support* (AC) = $\frac{48}{120} \times 100\% = 40\%$
2. *Support* (TV) = $\frac{25}{120} \times 100\% = 20,8\%$
3. *Support* (Magic Com) = $\frac{31}{120} \times 100\% = 25,8\%$
4. *Support* (Setrika) = $\frac{16}{120} \times 100\% = 13,3\%$
5. *Support* (Kompom) = $\frac{24}{120} \times 100\% = 20\%$
6. *Support* (Mixer) = $\frac{2}{120} \times 100\% = 1,6\%$
7. *Support* (Kulkas) = $\frac{52}{120} \times 100\% = 43,3\%$

$$8. \text{Support (Air Coler)} = \frac{4}{120} \times 100\% = 3,3\%$$

$$9. \text{Support (Speaker)} = \frac{27}{120} \times 100\% = 22,5\%$$

$$10. \text{Support (Kipas)} = \frac{42}{120} \times 100\% = 35\%$$

$$11. \text{Support (Ampia)} = \frac{1}{120} \times 100\% = 0,8\%$$

$$12. \text{Support (Blender)} = \frac{16}{120} \times 100\% = 13,3\%$$

$$13. \text{Support (Kasur)} = \frac{14}{120} \times 100\% = 11,6\%$$

$$14. \text{Support (SmartTV)} = \frac{9}{120} \times 100\% = 7,5\%$$

$$15. \text{Support (Braket)} = \frac{13}{120} \times 100\% = 10,8\%$$

$$16. \text{Support (Mesin Cuci)} = \frac{55}{120} \times 100\% = 45,8\%$$

$$17. \text{Support (Selang)} = \frac{14}{120} \times 100\% = 11,6\%$$

$$18. \text{Support (Meja)} = \frac{2}{120} \times 100\% = 1,6\%$$

$$19. \text{Support (Mic)} = \frac{28}{120} \times 100\% = 23,3\%$$

$$20. \text{Support (Showcase)} = \frac{6}{120} \times 100\% = 5\%$$

$$21. \text{Support (HP)} = \frac{1}{120} \times 100\% = 0,8\%$$

$$22. \text{Support (Subwoofer)} = \frac{2}{120} \times 100\% = 1,6\%$$

$$23. \text{Support (Lemari)} = \frac{7}{120} \times 100\% = 5,8\%$$

$$24. \text{Support (Dispenser)} = \frac{10}{120} \times 100\% = 8,3\%$$

$$25. \text{Support (Antena)} = \frac{6}{120} \times 100\% = 5\%$$

$$26. \text{Support (Divan)} = \frac{2}{120} \times 100\% = 1,6\%$$

$$27. \text{Support (Rak Sepatu)} = \frac{1}{120} \times 100\% = 0,8\%$$

$$28. \text{Support (Freezer)} = \frac{4}{120} \times 100\% = 3,3\%$$

$$29. \text{Support (Oven)} = \frac{1}{120} \times 100\% = 0,8\%$$

$$30. \text{Support (Rice Box)} = \frac{1}{120} \times 100\% = 0,8\%$$

$$31. \text{Support (Juicer)} = \frac{1}{120} \times 100\% = 0,8\%$$

$$32. \text{Support (Pipa)} = \frac{1}{120} \times 100\% = 0,8\%$$

$$33. \text{Support (Rak Kompom)} = \frac{1}{120} \times 100\% = 0,8\%$$

$$34. \text{Support (Bantal)} = \frac{2}{120} \times 100\% = 1,6\%$$

$$35. \text{Support (Rak Piring)} = \frac{1}{120} \times 100\% = 0,8\%$$

Berdasarkan hasil yang didapatkan dengan menentukan *support minimum* 10%, maka diperoleh item yang mencapai nilai *support* 10% yang dibuat dalam tabel berikut:

Tabel 3 Hasil Itemset 1

No	Itemset	Support
1	AC	48/120 x 100% = 40%
2	TV	25/120 x 100% = 20,8%
3	Magic Com	31/120 x 100% = 25,8%
4	Setrika	16/120 x 100% = 13,3%
5	Kompore	24/120 x 100% = 20%
6	Kulkas	52/120 x 100% = 43,3%
7	Speaker	27/120 x 100% = 22,5%
8	Kipas	42/120 x 100% = 35%
9	Blender	16/120 x 100% = 13,3%
10	Kasur	14/120 x 100% = 11,6%
11	Braket	13/120 x 100% = 10,8%
12	Mesin Cuci	55/120 x 100% = 45,8%
13	Selang	14/120 x 100% = 11,6%
14	Mic	28/120 x 100% = 23,3%

2. Pembentukan Itemset 2

Pembentukan *itemset* 2 sering disebut dengan C2. Rumus yang digunakan untuk menentukan *itemset* 2 adalah sebagai berikut:

$$\text{Support (A} \rightarrow \text{B)} = \frac{\sum \text{Transaksi Item A dan B}}{\sum \text{Transaksi}}$$

Berdasarkan hasil perhitungan menggunakan rumus di atas, maka hasil yang telah memenuhi syarat nilai *support minimum* 10% adalah:

Tabel 4 Hasil Itemset 2

No	Itemset 2		Fk	Support
1	AC	Magic Com	13	13/120 x 100% = 10,8%
2	AC	Kulkas	17	17/120 x 100% = 14,1%
3	AC	Kipas	14	14/120 x 100% = 11,6%
4	AC	Mesin Cuci	20	20/120 x 100% = 16,6%
5	Magic Com	Kulkas	15	15/120 x 100% = 12,5%
6	Magic Com	Mesin Cuci	13	13/120 x 100% = 10,8%
7	Kulkas	Kipas	24	24/120 x 100% = 20%
8	Kulkas	Mesin Cuci	19	19/120 x 100% = 15,8%
9	Speaker	Mic	15	15/120 x 100% = 12,5%

3. Pembentukan Itemset 3

Setelah hasil *itemset* 2 telah didapatkan, maka selanjutnya adalah mencari *itemset* 3 hingga nilai *support* yang sudah ditetapkan tidak ditemukan lagi. Pembentukan *itemset* 3 adalah sebagai berikut:

Tabel 5 Hasil Itemset 3

No	Itemset 3			Fk	Support
1	AC	Magic Com	Kulkas	6	6/120 x 100% = 5%
2	AC	Magic Com	Kipas	5	5/120 x 100% = 4,1%

No	Itemset 3			Fk	Support
3	AC	Magic Com	Mesin Cuci	6	$6/120 \times 100\% = 5\%$
4	AC	Kulkas	Kipas	6	$6/120 \times 100\% = 5\%$
5	AC	Kulkas	Mesin Cuci	8	$8/120 \times 100\% = 6,6\%$
6	AC	Kipas	Mesin Cuci	2	$2/120 \times 100\% = 1,6\%$
7	Magi c Com	Kulkas	Mesin Cuci	4	$4/120 \times 100\% = 3,3\%$
8	Magi c Com	Kulkas	Kipas	6	$6/120 \times 100\% = 5\%$
9	Kulkas	Mesin Cuci	Kipas	6	$6/120 \times 100\% = 5\%$
10	Magi c Com	Mesin Cuci	Kipas	2	$2/120 \times 100\% = 1,6\%$
11	Speaker	Mic	Kulkas	8	$8/120 \times 100\% = 6,6\%$

Berdasarkan perhitungan diatas pencarian nilai *support* berakhir karena hasil *minimum item support* yang ditentukan tidak ditemukan. Hal ini menunjukkan bahwa pencarian *item support* hanya sampai pada *itemset 2*.

4. Pembentukan Aturan Asosiasi

Langkah selanjutnya adalah menentukan aturan asosiasi yang memenuhi syarat minimum untuk *confidence* dengan menghitung nilai *confidence* aturan asosiasi $A \rightarrow B$ setelah diperoleh semua pola frekuensi tinggi dengan minimal *confidence* = 55%. Nilai *confidence* didapatkan dengan menggunakan rumus berikut:

$$Confidence (B | A) = \frac{\sum \text{Transaksi Item A dan B}}{\sum \text{Transaksi Item A}}$$

Dengan menggunakan rumus di atas dapat diperoleh nilai *confidence* dari 2 *itemset* seperti tabel berikut:

Tabel 6 Confidence Itemset 2

No	Item	Confidence	
1	AC → Magic Com	$13/48 \times 100\%$	27,0 %
2	Magic Com → AC	$13/31 \times 100\%$	41,9 %
3	AC → Kulkas	$17/48 \times 100\%$	35,4 %
4	Kulkas → AC	$17/52 \times 100\%$	32,6 %
5	AC → Kipas	$14/48 \times 100\%$	29,1 %
6	Kipas → AC	$14/42 \times 100\%$	33,3 %
7	AC → Mesin Cuci	$20/48 \times 100\%$	41,6 %
8	Mesin Cuci → AC	$20/55 \times 100\%$	36,3 %
9	Magic Com → Kulkas	$15/31 \times 100\%$	48,3 %
10	Kulkas → Magic Com	$15/52 \times 100\%$	28,8 %
11	Magic Com → Mesin Cuci	$13/31 \times 100\%$	41,9 %
12	Mesin Cuci → Magic Com	$13/55 \times 100\%$	23,6 %
13	Kulkas → Kipas	$24/52 \times 100\%$	46,1 %
14	Kipas → Kulkas	$24/42 \times 100\%$	57,1 %
15	Kulkas →	$19/52 \times 100\%$	36,5 %

No	Item	Confidence	
	Mesin Cuci		%
16	Mesin Cuci → Kulkas	19/55 x 100%	34,5 %
17	Speaker → Mic	15/27 x 100%	55,6 %
18	Mic → Speaker	15/28 x 100%	53,5 %

Berdasarkan tabel di atas, maka hasil penentuan *confidence* yang memenuhi *confidence* 55% didapatkan seperti pada tabel berikut:

Tabel 7 Hasil Confidence

No	Item	Confidence	
1	Kipas → Kulkas	24/42 x 100%	57,1%
2	Speaker → Mic	15/27 x 100%	55,6%

Jika semua nilai telah diperoleh, maka didapatkan hasil *rules* dari pola kombinasi *item* seperti tabel berikut:

Tabel 8 Hasil Kombinasi Item

No	Item	Support	Confidence
1	Jika membeli Kipas maka akan membeli Kulkas	20%	57,1%
2	Jika membeli Speaker maka akan membeli Mic	12,5%	55,6%

4.4. Pembagian K-Means

4.4.1. Data Transaksi Pembelian

Data yang digunakan peneliti adalah sampel data yaitu 120 transaksi pembelian pada Toko A8 Electronic & Furniture Piayu. Pada data ini peneliti membagi data menjadi 2 bagian, yaitu elektronik dan furniture.

Tabel 9 Data Pembelian

Transaksi	Elektronik(x)	Furniture(y)
1	4	0

<http://ejournal.upbatam.ac.id/index.php/cbis>

2	1	1
3	5	0
4	4	1
5	4	1
6	6	0
7	5	1
8	2	0
9	3	0
...	...	...
...	...	...
119	3	1
120	5	1

4.4.2. Pembentukan Cluster

Dalam melakukan analisis *cluster* algoritma *K-Means* terlebih dahulu membuat nilai *centroid*. Tujuan dari analisis *cluster* adalah untuk mendapatkan sebuah aturan *cluster* yang memenuhi syarat minimum *centroid*. Perhitungan nilai *centroid* menggunakan rumus *Euclidean Distance*, seperti berikut:

$$De = \sqrt{(xi-si)^2 + (yi-ti)^2}$$

Penjelasan secara rinci dari setiap langkah-langkah pembentukan iterasi menggunakan algoritma *k-means* seperti berikut.

1. Iterasi Pertama

Proses iterasi pertama dilakukan dengan cara menginisialisasi posisi *centroid* secara acak dan menghitung jarak antara setiap titik data dengan *centroid*. Adapun titik pusat awal *cluster* yang diambil peneliti secara acak adalah:

Tabel 10 Cluster Awal

data ke	centroid	x	y
4	0	4	1
42	1	3	3
105	2	5	1
90	3	1	1

Selanjutnya melakukan perhitungan jarak setiap data dengan pusat *cluster* menggunakan rumus *Euclidean Distance*.

Tabel 11 Perhitungan Euclidean Distance

Data ke	C0	C1	C2	C3	Mini mum	Cl uster
1	1	3.16 2278	1.41 4214	3.16 2278	1	0
2	3	2.82 8427	4	0	0	3

Data ke	C0	C1	C2	C3	Minimum	Cluster
3	1.4 142 14	3.60 5551	1	4.12 3106	1	2
4	0	2.23 6068	1	3	0	0
5	0	2.23 6068	1	3	0	0
6	2.2 360 68	4.24 2641	1.41 4214	5.09 902	1.41 4214	2
7	1	2.82 8427	0	4	0	2
8	2.2 360 68	3.16 2278	3.16 2278	1.41 4214	1.41 4214	0
9	1.4 142 14	3	2.23 6068	2.23 6068	1.41 4214	0
10	1	2	2	2	1	0
...	...	...	...	...	...	...
120	1	2.82 8427	0	4	0	2

Karena posisi titik pusat masih jauh dari posisi ideal, maka pengelompokan pada iterasi ini mungkin tampak acak dan tidak mencerminkan struktur *cluster* secara akurat.

2. Iterasi Kedua

Proses iterasi kedua dilakukan untuk mengetahui apakah ada perpindahan *cluster* antara iterasi pertama dan iterasi kedua dengan menggunakan *centroid* baru. Berikut adalah nilai *centroid* yang baru:

Tabel 12 Centroid Baru

Penentuan cluster baru	x	y
centroid baru yg ke-0	3.403509	0.333333
centroid baru yg ke-1	2.888889	2.555556
centroid baru yg ke-2	5.35	0.7
centroid baru yg ke-3	1.823529	0.823529

Selanjutnya algoritma menghitung ulang jarak antara seluruh data dan *centroid* baru menggunakan posisi *centroid* yang baru,

<http://ejournal.upbatam.ac.id/index.php/cbis>

kemudian mengatur ulang data berdasarkan *centroid* terdekat. Beberapa data sering kali berpindah *cluster* sebagai akibat dari proses ini. Setelah itu, data baru dari setiap *cluster* digunakan untuk memperbarui *centroid* lagi.

Tabel 13 Perhitungan Jarak Centroid

Data ke	C0	C1	C2	C3	Min	Cluster
1	0.68 331	2.78 665 2	1.52 069 1	2.32 7063 68	0.6 833 1	0
2	2.49 425 3	2.44 696 8	4.36 033 3	0.84 2224 77	0.8 422 25	3
3	1.63 091 9	3.31 476 3	0.78 262 4	3.28 1488 4	0.7 826 24	2
4	0.89 456 5	1.91 162 8	1.38 293 2	2.18 3613 08	0.8 945 65	0
5	0.89 456 5	1.91 162 8	1.38 293 2	2.18 3613 08	0.8 945 65	0
6	2.61 78	4.02 614 9	0.95 524 9	4.25 6889 39	0.9 552 49	2
7	1.73 009 5	2.62 231 6	0.46 097 7	3.18 1368 77	0.4 609 77	2
8	1.44 254 9	2.70 573 2	3.42 235 3	0.84 2224 77	0.8 422 25	3
9	0.52 338 4	2.55 797	2.45 204	1.43 6065 37	0.5 233 84	0
10	0.77 927 1	1.55 951 9	2.36 907 2	1.18 9632 26	0.7 792 71	0
...	...	...	...	...	...	...
120	1.73 009 5	2.62 231 6	0.46 097 7	3.18 1368 77	0.4 609 77	2

Pada iterasi kedua terlihat bahwa beberapa data mengalami perpindahan *cluster*, maka selanjutnya adalah menentukan nilai *centroid* yang baru untuk proses iterasi selanjutnya.

3. Iterasi Ketiga

Pada proses iterasi ketiga proses penetapan dan penghitungan jarak dilakukan kembali

dengan menggunakan posisi *centroid* yang ditentukan dari iterasi 2.

Tabel 14 *Centroid* dari Iterasi 2

cluster baru	x	y
centroid 0	3.345455	0.272727
centroid 1	2.823529	2.294118
centroid 2	5.35	0.7
centroid 3	1.785714	0.857143

Berikut hasil perhitungan dan penetapan nilai cluster pada iterasi ketiga:

Tabel 15 Perhitungan *Cluster* Iterasi 3

Data ke	C0	C1	C2	C3	Min	Cluster
1	0.70 9091	2.57 818 9	1.52 0691	2.37 4396	0.7 090 91	0
2	2.45 5623	2.23 606 8	4.36 0333	0.79 8596	0.7 985 96	3
3	1.67 6872	3.16 227 8	0.78 2624	3.32 6609	0.7 826 24	2
4	0.97 8445	1.74 894 9	1.38 2932	2.21 8889	0.9 784 45	0
5	0.97 8445	1.74 894 9	1.38 2932	2.21 8889	0.9 784 45	0
6	2.66 8519	3.91 828 3	0.95 5249	4.30 0569	0.9 552 49	2
7	1.80 7331	2.53 214 6	0.46 0977	3.21 7459	0.4 609 77	2
8	1.37 2818	2.43 745 3	3.42 2353	0.88 3523	0.8 835 23	3
9	0.44 0135	2.30 089 5	2.45 204	1.48 6332	0.4 401 35	0
10	0.80 5149	1.30 609 4	2.36 9072	1.22 266	0.8 051 49	0
...	...	...	...	...	...	...
120	1.80 7331	2.53 214 6	0.46 0977	3.21 7459	0.4 609 77	2

4. Iterasi Keempat

Pada iterasi keempat *cluster* dihitung kembali dengan menggunakan *cluster* terbaru.

Tabel 16 *Cluster* Terbaru

Centroid baru	x	y
Centroid 0	3.37037	0.240741
Centroid 1	2.777778	2.77778
Centroid 2	5.35	0.7
Centroid 3	1.785714	0.571429

Selanjutnya, peneliti melakukan iterasi keempat untuk mendapatkan nilai *cluster*.

Tabel 17 Perhitungan *Cluster* Iterasi 4

Data ke	C0	C1	C2	C3	Min	Cluster
1	0.67 408 4	2.5 849 76	1.52 0691	2.28 683	0.76 4084	0
2	2.48 900 2	2.1 893 4	4.36 0333	0.89 4997	0.89 4997	3
3	1.64 731 6	3.1 822 23	0.78 2624	3.26 4684	0.78 2624	2
4	0.98 636 1	1.7 682 03	1.38 2932	2.25 5379	0.98 6361	0
5	0.98 636 1	1.7 682 03	1.38 2932	2.25 5379	0.98 6361	0
6	2.64 062 6	3.9 460 09	0.95 5249	4.25 285	0.95 5249	2
7	1.79 782 3	2.5 633 94	0.46 0977	3.24 2731	0.46 0977	2
8	1.39 135 6	2.4 069 09	3.42 2353	0.61 0286	0.61 0286	3
9	0.44 173 6	2.2 885 92	2.45 204	1.34 2021	0.44 1736	0
10	0.84 477 7	1.2 969 58	2.36 9072	1.28 7697	0.84 4777	0
...	...	...	...	...	...	...
120	1.79 782 3	2.5 633 94	0.46 0977	3.24 2731	0.46 0977	2

Berdasarkan proses pada iterasi ketiga dan keempat, tidak ada lagi data *cluster* yang mengalami perpindahan sehingga proses iterasi dihentikan pada proses iterasi keempat. Adapun jumlah masing-masing *cluster* yang didapat yaitu pada *cluster* 0 terdapat 54 data, *cluster* 1 terdapat 18 data, *cluster* 2 terdapat 20 data, dan *cluster* 3 terdapat 28 data.

4.5. Pengujian

4.5.1. Pengujian Algoritma Apriori

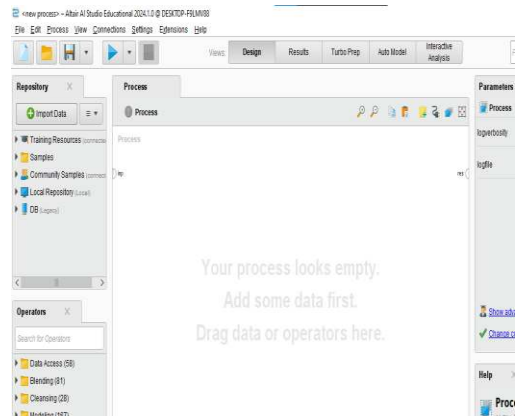
Pada penelitian ini peneliti menggunakan *software RapidMiner* untuk menguji kebenaran dari hasil data yang diperoleh menggunakan perhitungan manual. Berikut adalah langkah-langkah pengujian menggunakan *software RapidMiner*:

- Langkah pertama yang dilakukan peneliti adalah mengubah data transaksi menjadi tabel tabulasi pada Microsoft Excel seperti gambar berikut:

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
AC	TV	Magic ComSetrika	Kompom	Mixer	Kulkas	Air Coler	Speaker	Kipas	Amplo	Blender	Kasur	SmartTV	Braket	Mesin Cuci	Selang	Meja	Mic	Showcase		
1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
1	0	1	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	
0	0	0	0	0	0	1	0	1	1	1	1	0	0	0	0	0	0	0	0	
1	0	1	0	0	0	1	0	0	1	0	0	1	0	0	0	0	0	0	0	
1	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
0	0	1	0	0	0	1	0	0	1	0	1	0	1	1	0	0	0	0	0	
1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	
1	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	
1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	
0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	1	0	0	
0	1	0	0	0	0	1	0	1	1	0	0	0	0	1	1	0	0	1	0	
1	0	1	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	
1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	
0	0	0	0	0	0	1	0	0	1	0	1	0	0	0	1	0	0	0	0	
1	0	0	0	1	0	1	0	1	0	0	0	0	0	0	0	1	0	0	1	
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	
1	0	0	0	0	0	1	0	1	1	0	0	1	1	1	0	0	0	1	0	
0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	1	
0	0	0	0	0	0	1	0	0	1	0	0	1	1	0	0	Activate Windows				

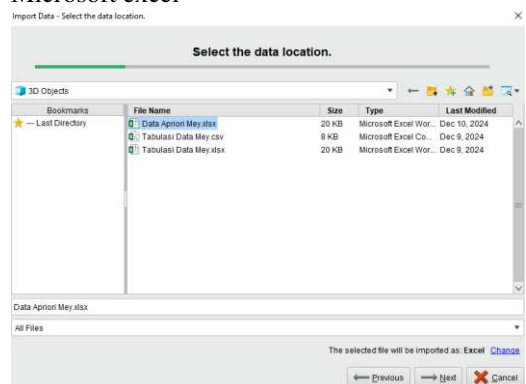
Gambar 2 Tabulasi Data

- Selanjutnya adalah membuka *software RapidMiner* dan akan menampilkan seperti gambar di bawah:



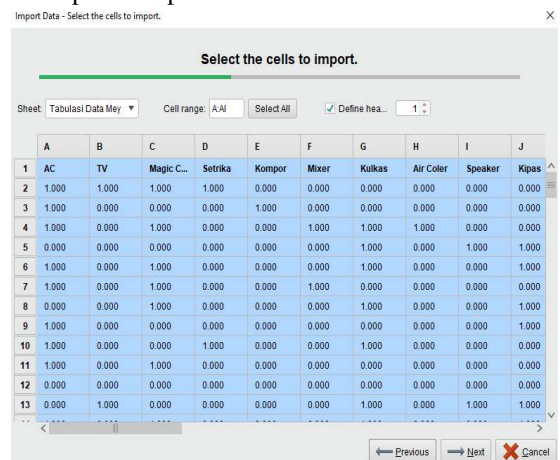
Gambar 3 Tampilan awal RapidMiner

- Pada bagian *Repository* klik *import data*, pilih data tabel tabulasi yang telah dibuat di Microsoft excel



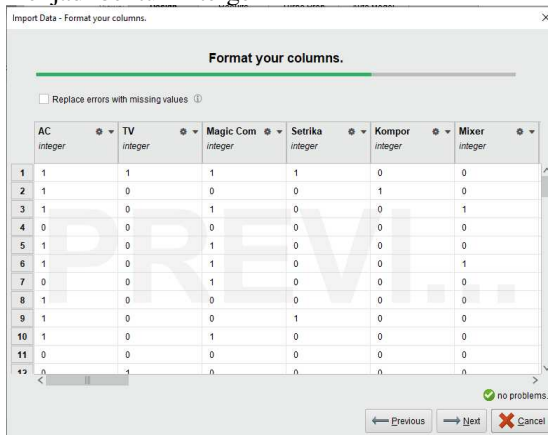
Gambar 4 Import Data

- Kemudian klik *Next*, dan data akan ditampilkan seperti berikut:



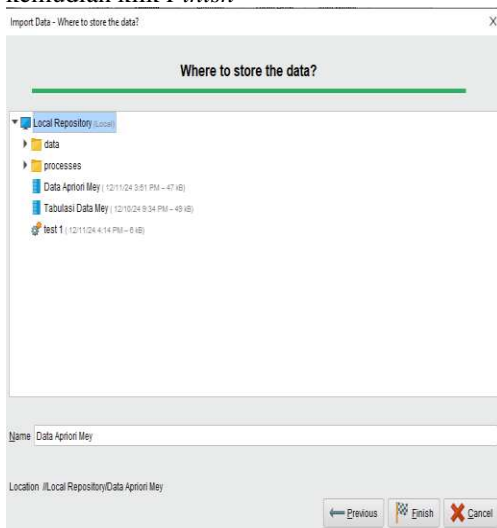
Gambar 5 Select Data

5. Klik *Next*, kemudian data akan diubah menjadi bentuk integer



Gambar 6 Format Kolom

6. Pilih *Next* dan simpan data dalam *repository*, kemudian klik *Finish*



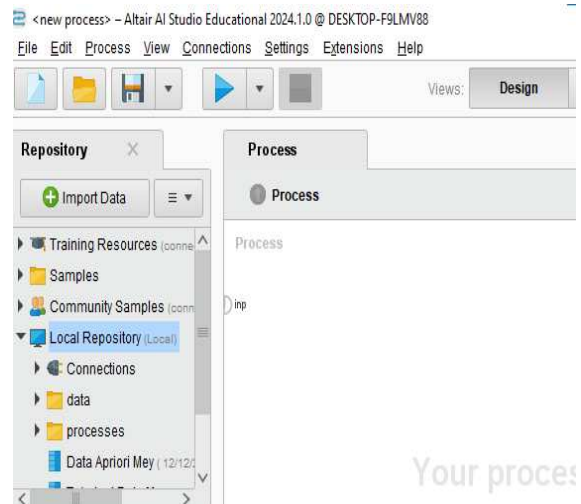
Gambar 7 Penyimpanan Data

7. Setelah penyimpanan data selesai, maka akan ditampilkan seperti gambar di bawah ini

Row No.	AC	TV	Magic Com	Setrika	Kompor	Mixer	Kukus	Air
1	1	1	1	1	0	0	0	0
2	1	0	0	0	1	0	0	0
3	1	0	1	0	0	1	1	1
4	0	0	0	0	0	0	1	0
5	1	0	1	0	0	0	1	0
6	1	0	1	0	0	1	0	0
7	0	0	1	0	0	0	1	0
8	1	0	0	0	0	0	0	0
9	1	0	0	1	0	0	0	0
10	1	0	1	0	0	0	0	0
11	0	0	0	0	0	0	0	0
12	0	1	0	0	0	0	1	0
13	1	0	1	0	0	0	1	0
14	1	0	0	0	0	0	1	0

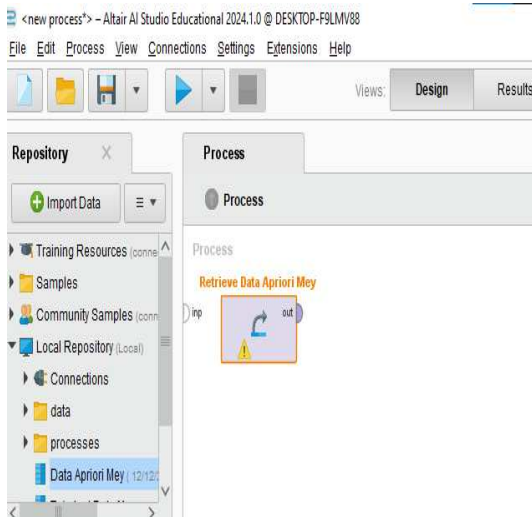
Gambar 8 Tampilan Data Tersimpan

8. Selanjutnya pilih *Design* dan pada bagian *Local repository*, data tabulasi yang telah diinput akan muncul



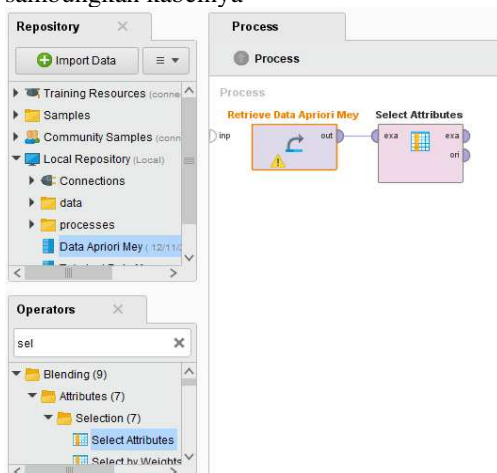
Gambar 9 Data tabulasi yang Diinput

9. Langkah selanjutnya adalah *drag and drop* data tabulasi yang ada di bagian *repository* ke bagian *Process*



Gambar 10 Drag Data ke Proses

10. Untuk proses perhitungan, pada bagian *Operators* pilih *Select Attributes* kemudian *drag and drop* ke dalam *Process* dan sambungkan kabelnya

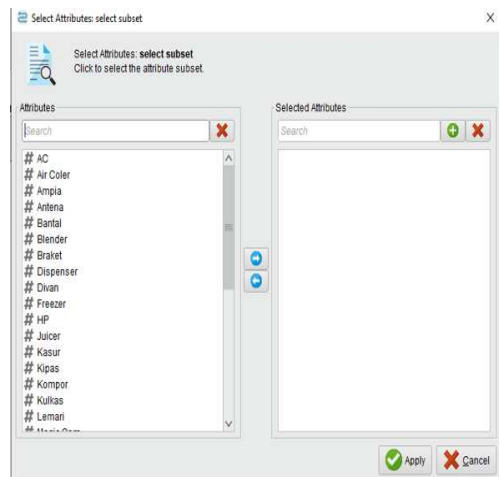


Gambar 11 Select Attributes

11. Klik *Select Attributes*, pada bagian *Parameters attribute filter type* pilih *a subset*, kemudian klik *select attributes* dan akan menampilkan atribut yang sudah di *upload* pada excel

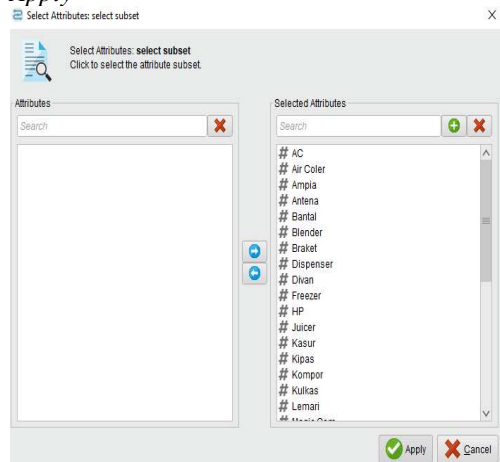


Gambar 12 Parameter Atribut



Gambar 13 Pemilihan Atribut

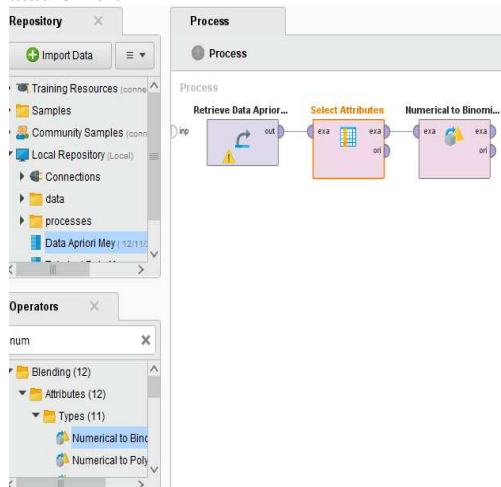
12. Langkah berikutnya yaitu pilih semua atribut dengan cara double klik item tersebut, atribut yang telah dipilih akan berpindah ke kolom *Selected attributes*. Kemudian pilih *Apply*



Gambar 14 Atribut yang Dipilih

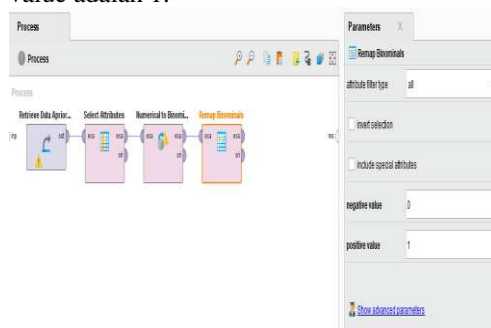
13. Selanjutnya pilih operator *Numerical to Binominal* kemudian *drag and drop* ke bagian *Process* dan sambungkan kabelnya.

Operator ini berfungsi untuk mengkonversi atribut numerik menjadi atribut binominal atau biner.



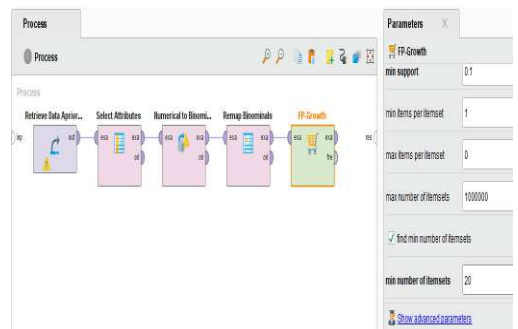
Gambar 15 Proses Numerical to Binominal

14. Langkah berikutnya pilih operator *Remap Binominal*, kemudian *drag and drop* ke bagian *Process* dan sambungkan kabelnya. Operator ini bertujuan untuk mengubah nilai atribut binominal yang sudah ada, seperti mengubah nilai 1 dan 0 menjadi positif dan negatif. Pada data ini peneliti membuat nilai negative value adalah 0 dan nilai positive value adalah 1.



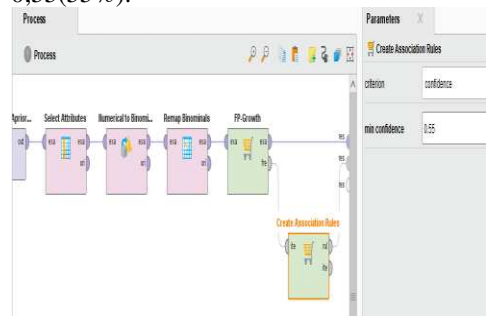
Gambar 16 Mengubah nilai Binominal

15. Selanjutnya pilih operator *FP-Growth*, kemudian *drag and drop* ke bagian *Process* dan sambungkan. Operator ini digunakan untuk menemukan pola atau asosiasi yang sering muncul dalam dataset. Pada bagian *FP-Growth* masukkan nilai minimal support sebesar 0,1(10%) yang sudah ditentukan peneliti sebelumnya dan nilai *minimal number of itemset* sebanyak 20.



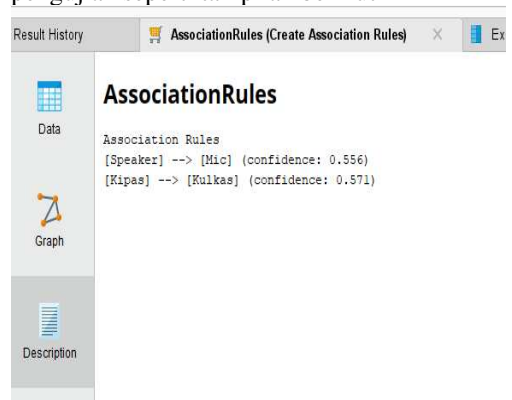
Gambar 17 Nilai Support

16. Langkah selanjutnya adalah pilih operator *Create Association Rules*, kemudian *drag and drop* ke bagian *Process* dan sambungkan. Hasil dari operator ini berupa daftar aturan asosiasi dengan matriks *support* dan *confidence*. Pada data ini peneliti membuat nilai *confidence* sebesar 0,55(55%).



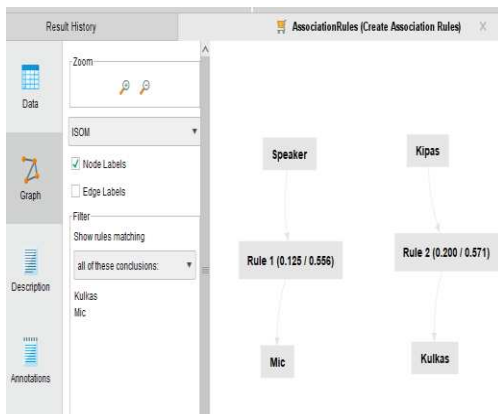
Gambar 18 Nilai Confidence

17. Setelah semua process selesai selanjutnya klik *Connections* untuk menampilkan hasil pengujian seperti tampilan berikut



Gambar 19 Association Rules

18. Bentuk grafik dari hasil *Association rules* adalah seperti di bawah ini



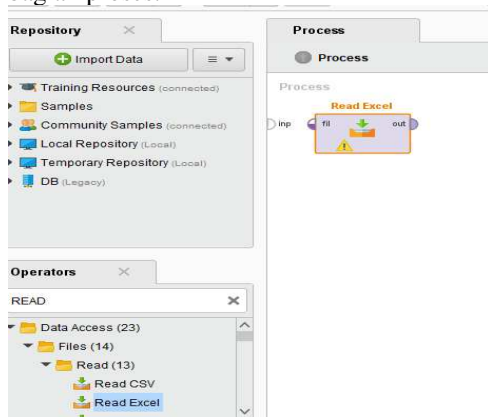
Gambar 20 Grafik Association Rules

Berdasarkan hasil yang didapatkan menggunakan *software RapidMiner* menunjukkan bahwa perhitungan *association rule* dengan algoritma apriori secara manual menampilkan hasil yang sama, yaitu mencakup 2 aturan asosiasi atau kombinasi item yang sering dibeli secara bersamaan seperti yang dijelaskan pada sub bab sebelumnya.

4.5.2. Pengujian K-Means Clustering

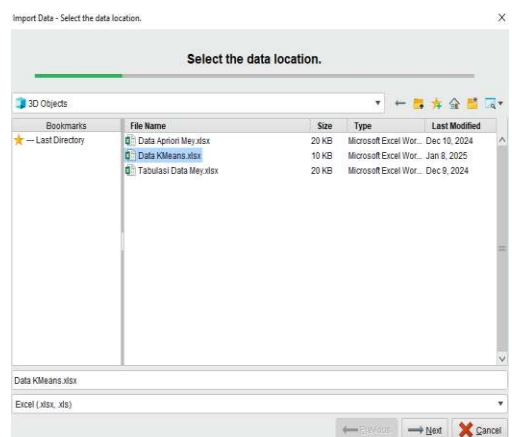
Berikut adalah langkah-langkah pengujian data menggunakan *software RapidMiner*:

1. Pada bagian operator cari *Read Excel*. Bagian ini berfungsi untuk membaca dataset dari file excel, kemudian *drag* ke bagian proses.



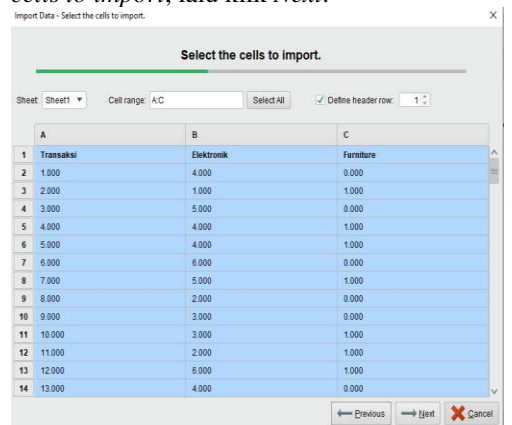
Gambar 21 Operator Read Excel

2. Kemudian klik dua kali pada proses *read excel* untuk memasukkan data yang akan diolah. Lalu klik *Next*.



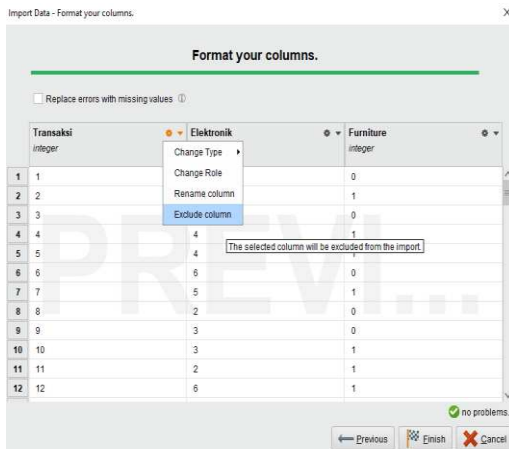
Gambar 22 Input Data

3. Berikutnya akan muncul tampilan *select the cells to import*, lalu klik *Next*.



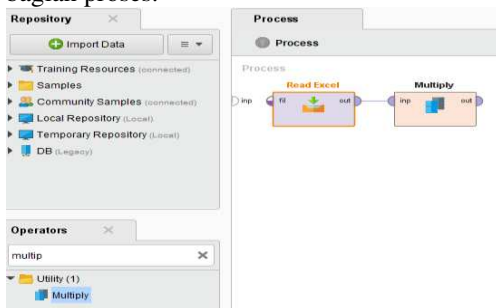
Gambar 23 Tampilan Select Data

4. Pada tampilan *Format your columns*, bagian Transaksi pilih *Exclude column* karena tabel tersebut tidak digunakan. Kemudian klik *Finish*.



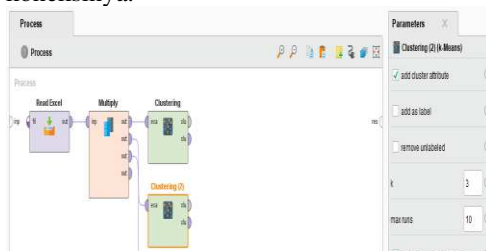
Gambar 24 Format Kolom

5. Selanjutnya pilih operator *Multiply*. Operator ini berfungsi agar peneliti bisa melakukan pengolahan data menggunakan beberapa *kmeans* dengan nilai *k* (jumlah cluster) yang berbeda. Kemudian *drag* ke bagian proses.

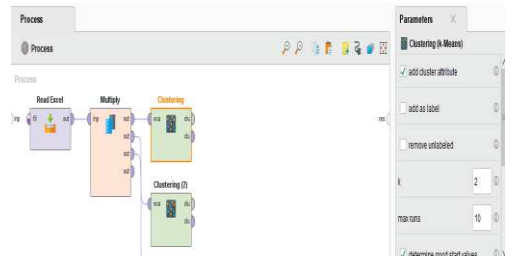


Gambar 25 Operator Multiply

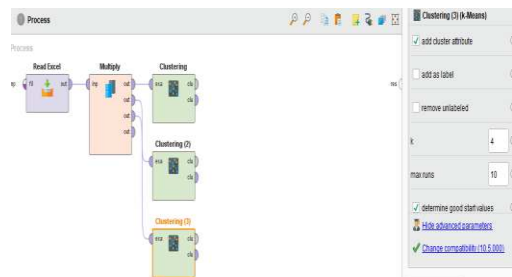
6. Kemudian pilih operator *Kmeans*, lalu *drag* ke bagian proses. Pada bagian ini peneliti membagi *kmeans* menjadi 3 bagian dengan nilai *cluster* yang berbeda. *Clustering 1* menggunakan *k* adalah 2 (2 cluster), *clustering 2* menggunakan 3 cluster, dan *clustering 3* menggunakan 4 cluster. Selanjutnya hubungkan masing-masing koneksinya.



Gambar 26 Clustering 1

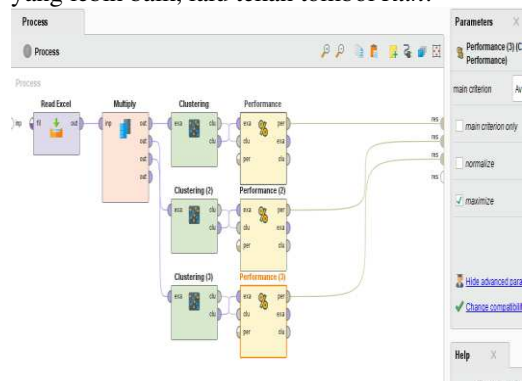


Gambar 27 Clustering 2



Gambar 28 Clustering 3

7. Langkah selanjutnya pilih operator *Cluster Distance Performance*. Tujuan operator ini adalah untuk membandingkan performa setiap *clustering* dalam menentukan konfigurasi yang terbaik. Operator ini juga dapat memastikan seberapa baik data dikelompokkan ke dalam *cluster*. Kemudian *drag* ke bagian proses dan koneksikan setiap *clustering* dengan cara silang, dan pada bagian parameter centang *maximize*. Selanjutnya, hubungkan setiap *performance* ke *result* untuk mengetahui kelompok mana yang lebih baik, lalu tekan tombol *Run*.

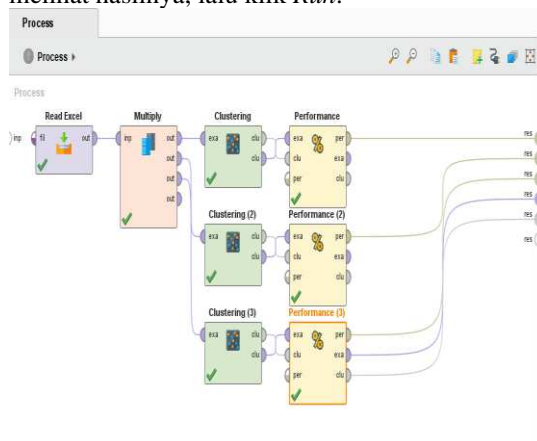


Gambar 29 Cluster Distance Performance

8. Pada bagian *Result* dihasilkan performa masing-masing *clustering* dan nilai *Davies Bouldin*. Kualitas hasil *clustering* dinilai menggunakan *Davis-Bouldin Index (DBI)*. Nilai DBI membantu dalam menilai seberapa

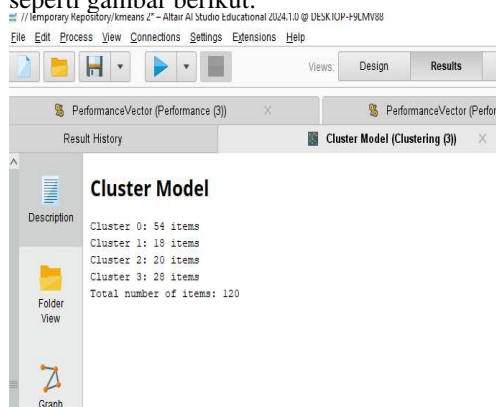
sukses teknik pengelompokan, seperti *K-Means*, dalam mengelompokkan data ke dalam kelompok. Pada tampilan ini dihasilkan:

- Clustering 1 (2 cluster) : Davies Bouldin = 0.922*
 - Clustering 2 (3 cluster) : Davies Bouldin = 0.855*
 - Clustering 3 (4 cluster) : Davies Bouldin = 0.754*
9. Berdasarkan nilai *Davies Bouldin*, maka *Clustering 3 (4 cluster)* adalah yang pilihan terbaik karena nilai DBI nya paling kecil yaitu 0.754. Maka selanjutnya hubungkan *Performance(3)* ke bagian *result* untuk melihat hasilnya, lalu klik *Run*.



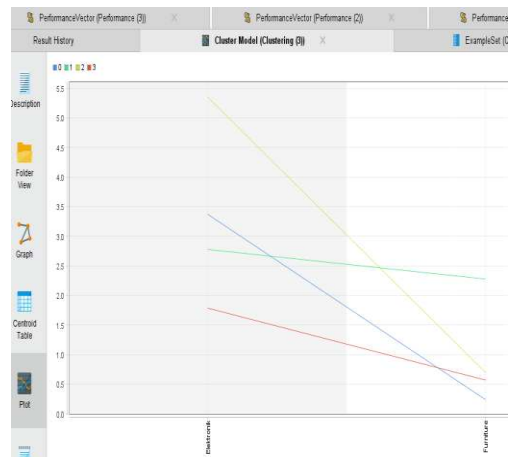
Gambar 30 Koneksi Performance ke Result

10. Hasil untuk tampilan *cluster model* adalah seperti gambar berikut:



Gambar 31 Hasil Cluster Model

11. Klik *Plot* untuk melihat grafik yang terbentuk pada *cluster* seperti tampilan berikut:



Gambar 32 Tampilan Plot Cluster

V. Kesimpulan

Berdasarkan penelitian yang telah dilakukan dengan menggunakan algoritma Apriori untuk analisis pola pembelian dan algoritma *K-Means* untuk klasifikasi data, maka dapat diambil kesimpulan sebagai berikut:

- Penelitian ini berhasil menentukan bahwa pada Toko Piayu Elektronik & Furnitur A8, terdapat kemungkinan 55% konsumen yang membeli kipas angin juga akan membeli AC. Penelitian ini menunjukkan hubungan antar item yang dapat digunakan untuk meningkatkan strategi penjualan.
- Hasil penelitian juga menunjukkan bahwa dengan tingkat kepercayaan 55%, terdapat korelasi besar antara pembelian speaker dan mikrofon. Hasil ini dapat menjadi landasan untuk menciptakan promosi yang lebih sukses.
- Berdasarkan kebiasaan pembelian konsumen, data konsumen secara efektif dibagi menjadi 4 *cluster* menggunakan analisis *cluster*. Nilai *Davis Bouldin* sebesar 0.754 yang menunjukkan kualitas *cluster* yang dapat diterima, digunakan untuk menghitung jarak terdekat ke pusat *cluster*.
- Berdasarkan hasil pengelompokan diperoleh bahwa terdapat 54 item di *cluster* 0, 18 item di *cluster* 1, 20 di *cluster* 2, dan 28 di *cluster* 3. Berdasarkan persebaran data pada *Visualization* di *RapidMiner* dapat dilihat bahwa barang elektronik adalah produk yang paling populer, sementara barang furniture memerlukan lebih banyak fokus dalam strategi pemasaran.

Ucapan Terima Kasih

Penulis mengucapkan terima kasih kepada Tuhan Yang Maha Esa atas berkat dan penyertaan-Nya penulis dapat menyelesaikan penelitian ini. Penulis juga berterima kasih kepada Bapak Rika Harman selaku Dosen pembimbing dan semua pihak yang telah memberikan kontribusi positif dalam penyelesaian penelitian ini.

Daftar Pustaka

- [1] N. Oktaviani, "Implementasi Algoritma Apriori Untuk Analisis Pola Pembelian Konsumen Pada Toko Serba," *JATI (Jurnal Mhs. Tek. Inform.,* vol. 8, no. 3, pp. 3706–3711, 2024, doi: 10.36040/jati.v8i3.9624.
- [2] N. F. Adani *et al.*, "Implementasi Data Mining Untuk Pengelompokan Data Penjualan Berdasarkan Pola Pembelian Menggunakan Algoritma K-Means Clustering Pada Toko Syihan," *J. Cyber Tech*, vol. x. No.x, no. x, pp. 1–11, 2019, [Online]. Available: <https://ojs.trigunadharma.ac.id/index.php/jct/article/view/4648%0Ahttps://ojs.trigunadharma.ac.id/index.php/jct/article/download/4648/791>
- [3] M. Arif Saifudin, H. Endah Wahanani, and A. Junaidi, "Implementasi Algoritma Asosiasi Fp-Growth Dan Klasifikasi K-Means Terhadap Pola Pembelian Konsumen Di Marketplace Shopee," *JATI (Jurnal Mhs. Tek. Inform.,* vol. 8, no. 1, pp. 764–771, 2024, doi: 10.36040/jati.v8i1.8848.
- [4] L. N. Nainggolan and E. Elisa, "Analisis Pola Pembelian Produk Menggunakan Algoritma Apriori Pada Toko Surya Elektronik," *J. Comasie*, vol. 07, no. 06, pp. 26–31, 2022.
- [5] R. A. Indraputra and R. Fitriana, "K-Means Clustering Data COVID-19," *J. Tek. Ind.,* vol. 10, no. 3, pp. 275–282, 2020, doi: 10.25105/jti.v10i3.8428.
- [6] A. N. Fuadi, H. Bhakti, and A. Premana, "RITEL DMART MENGGUNAKAN ALGORITMA APRIORI," vol. 12, no. 3, pp. 3269–3277, 2024.
- [7] D. Setiawan, E. Sedyono, and I. Sembiring, "Pemanfaatan Metode Association Rules dan Holt-Winter Multiplicative untuk Meningkatkan Peluang Penjualan Obat Pertanian pada PT . ABC," vol. 01, pp. 46–55, 2020.
- [8] E. M. Fitri, R. R. Suryono, and A. Wantoro, "Klasterisasi Data Penjualan Berdasarkan Wilayah Menggunakan Metode K-Means Pada Pt Xyz," *J. Komputasi*, vol. 11, no. 2, pp. 157–168, 2023, doi: 10.23960/komputasi.v11i2.12582.
- [9] S. Saefudin and D. Fernando, "Penerapan Data Mining Rekomendasi Buku Menggunakan Algoritma Apriori," *JSiI (Jurnal Sist. Informasi)*, vol. 7, no. 1, p. 50, 2020, doi: 10.30656/jsii.v7i1.1899.
- [10] S. Adha, "Faktor Revolusi Perilaku Konsumen Era Digital : Sebuah Tinjauan Literatur," *Jipis*, vol. 31, no. 2, pp. 134–148, 2022, doi: 10.33592/jipis.v31i2.3286.
- [11] M. R. Nahjan, N. Heryana, A. Voutama, F. I. Komputer, U. S. Karawang, and R. Miner, "IMPLEMENTASI RAPIDMINER DENGAN METODE CLUSTERING K-MEANS UNTUK ANALISA PENJUALAN PADA TOKO OJ CELL," vol. 7, no. 1, pp. 101–104, 2023.
- [12] Y. R. Sari, A. Sudewa, D. A. Lestari, and T. I. Jaya, "Penerapan Algoritma K-Means Untuk Clustering Data Kemiskinan Provinsi Banten Menggunakan Rapidminer," *CESS (Journal Comput. Eng. Syst. Sci.,* vol. 5, no. 2, p. 192, 2020, doi: 10.24114/cess.v5i2.18519.