



Penerapan Metode K-Means Dalam Menentukan Klasifikasi Produk Pembelian Pelanggan

Kevin Noel Nababan

Fakultas Ilmu Komputer dan Teknologi Informasi, Program Studi Teknik Informatika, Universitas Budi Darma, Medan, Indonesia
Email: Kevinnoelnababa@gmail.com

INFORMASI ARTIKEL

Histori Artikel:

Submitted : Jan 10, 2022

Accepted : Nov 25, 2022

Published : Nov 31, 2022

KORESPONDENSI

Email: Kevinnoelnababa@gmail.com

A B S T R A K

Menurut survey ada terdapat kira-kira 3.500 jenis produk makanan dan non-makanan tersedia dengan harga bersaing, memenuhi hampir semua kebutuhan konsumen sehari-hari. Dengan jumlah transaksi yang cukup besar, perusahaan membutuhkan alat bantu analisis dan klasifikasi untuk memberikan informasi yang bermanfaat bagi perusahaan dalam penentuan tata letak barang, barang apa yang paling diminati oleh konsumen dan lain-lain. Penentuan tata letak produk makanan dan minuman dilakukan untuk mempermudah konsumen dalam mencari produk makanan dan minuman sehingga tidak mengecewakan para konsumen dalam mencari letak barang produk mana yang cocok digabungkan dengan produk lain yang sering diminati konsumen sehingga para konsumen bisa menghemat waktu. Berdasarkan uraian di atas, maka dibutuhkan alat bantu analisis data mining dengan menggunakan Association Rules untuk diolah dengan metode K-Means. Saat ini pemanfaatan data-data yang dimiliki belum maksimal, baru sebatas untuk pembuatan laporan.

Kata Kunci: Normalisasi; K-Means; Data Penjualan; Klasifikasi; Pola

A B S T R A C T

According to the survey, there are approximately 3,500 types of food and non-food products available at competitive prices, meeting almost all the daily needs of consumers. With a fairly large number of transactions, companies need analysis and classification tools to provide useful information for companies in determining the layout of goods, what goods are most in demand by consumers and others. Determination of the layout of food and beverage products is carried out to make it easier for consumers to find food and beverage products so as not to disappoint consumers in finding the location of which products are suitable to be combined with other products that are often in demand by consumers so that consumers can save time. Based on the description above, a data mining analysis tool is needed using Association Rules to be processed using the K-Means method. Currently, the utilization of the available data has not been maximized, it is only limited to making reports.

Keywords: Normalization; K-Means; Sales Data; Classification; Pattern.

1. PENDAHULUAN

Supermarket adalah salah satu swayalan yang sedang berkembang memiliki jaringan minimarket yang menyediakan bahan pokok sehari-hari. Lebih dari 3.500 jenis produk makanan dan non-makanan tersedia dengan harga bersaing, memenuhi hampir semua kebutuhan konsumen sehari-hari. Dengan jumlah transaksi yang cukup besar, perusahaan membutuhkan alat bantu analisis untuk memberikan informasi yang bermanfaat bagi perusahaan dalam penentuan tata letak barang, barang apa yang paling diminati oleh konsumen dan lain-lain. Penentuan tata letak produk makanan dan minuman dilakukan untuk mempermudah konsumen dalam mencari produk makanan dan minuman sehingga tidak mengecewakan para konsumen dalam mencari letak barang produk mana yang cocok digabungkan dengan produk lain yang sering diminati konsumen sehingga para konsumen bisa menghemat waktu. [1]

Berdasarkan permasalahan yang dihadapi perusahaan, maka dibutuhkan alat bantu analisis data mining. Saat ini pemanfaatan data-data yang dimiliki belum maksimal, baru sebatas untuk pembuatan laporan. Data transaksi penjualan yang terkumpul dan tersimpan dapat memberikan pengetahuan yang bermanfaat bagi pengambil keputusan dalam melakukan usaha-usaha yang terkait dengan peningkatan penjualan dengan melakukan promosi yang tepat dan mengetahui kebiasaan berbelanja para *customer*. Oleh karena itu yang menjadi permasalahan dalam penelitian ini adalah menentukan pola pembelian *customer* untuk penentuan tata letak barang dengan memanfaatkan data mining. [2]

Menurut Peneliti terdahulu H. Sy, Rismayani, and A. Syam Dengan memanfaatkan data mining yang mengoptimasi proses pencarian informasi dalam basis data yang besar, serta menemukan pola-pola yang tidak

diketahui sebelumnya. Oleh sebab itu dalam mencari informasi yang ingin dilakukan terlebih dahulu menyipkan data yang siap diolah salah satunya dengan menormalisasikan data penjualan supermarket dengan cara mengelompokkan data jenis prodak yang dijual dengan menggunakan metode K-Means. Metode K-Means adalah suatu metode penganalisa data atau metode Data Mining yang melakukan proses pemodelan tanpa supervisi (*unsupervised*) dan merupakan salah satu metode yang melakukan pengelompokan data dengan sistem partisi.[3] Menurut Peneliti terdahulu E. Elisa Sehingga data yang telah diolah dengan metode K-means dijadikan data dalam memperoleh informasi pola penjualan untuk penentuan tata letak barang dengan menggunakan klasifikasi .klasifikasi adalah suatu metode untuk mencari pola hubungan antar satu atau lebih item dalam suatu dataset. klasifikasi banyak digunakan pada data transaksi atau biasa disebut market basket, misalnya sebuah swalayan memiliki market basket, dengan adanya klasifikasi, pemilik swalayan dapat mengetahui pola pembelian seorang konsumen, jika seorang konsumen membeli item A , B, punya kemungkinan 50% dia akan membeli item C, pola ini sangat signifikan dengan adanya data transaksi selama ini. Menurut peneliti terdahulu yang dilakuakn oleh H. Annur tujuan dari penelitian ini untuk mengelompokkan data penjualan kedalam sebuah cluster dengan metode Data Mining Algoritma K-Means Clustering. Data Penjualan nantinya akan dikelompokkan berdasarkan kemiripan data tersebut sehingga data dengan karakteristik yang sama akan berada dalam satu cluster. Atribut yang digunakan adalah brand dan penjualan[4]

Serta peneliti yang dilakuakn oleh R. Rachman, untuk dapat melakukan hal tersebut toko luxor variasi mobil membutuhkan strategi-strategi penjualan untuk dapat menarik minat pembeli dan meningkatkan laba atau pendapatan perusahaan. Pada penelitian ini analisa data mining dilakukan dengan Teknik Clustering menggunakan metode K-Means[5].

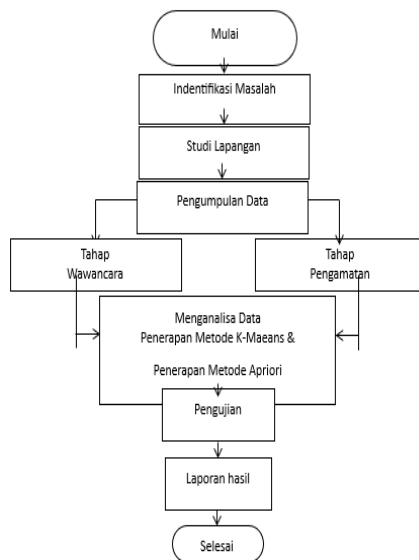
. Serta peneliti yang dilakuakn oleh A. A. Tri Susilo, penyusunan pola pembelian barang yang biasa digunakan dalam data mining dengan algoritme association rule, algoritme yang digunakan menggunakan klasifikasi dan FPGrowth. klasifikasi menghasilkan 5 pola aturan asosiasi minimum support sebesar 10% dan minimum confidence sebesar 50% menghasilkan nilai confidence tertinggi yaitu 100% , sedangkan algoritme FP- Growth menghasilkan 5 pola aturan asosiasi dengan minimum support count 2 menghasilkan nilai support count tertinggi yaitu 8[6].

Serta peneliti yang dilakukan oleh Tri Susilo A. A.Sunardi L.Waruwu Y. menjelaskan analisis transaksi penjualan bertujuan untuk merancang strategi penjualan atau pemasaran yang efektif dengan memanfaatkan data transaksi penjualan yang telah tersedia.Selain itu penggunaan teknik analisis ini juga dapat menemukan pola berupa produk-produk yang sering dibeli bersamaan dalam sebuah transaksi. Pengolahan data untuk memperoleh informasi mengenai pola-pola dapat dilakukan dengan datamining, yaitu Association Rules.klasifikasi yang merupakan salah satu metode Association Rules cocok untuk diterapkan bila terdapat beberapa hubungan item yang ingin dianalisis[7].

2. METODOLOGI PENELITIAN

2.1 Kerangka Kerja Penelitian

Metodologi penelitian merupakan suatu strategis yang disusun sedemikian rupa, untuk mencari tahu suatu kebenaran atau kenyataan terhadap yang di kaji berdasarkan masalah yang terjadi. Pada metode penelitian ini disusun sangat terstruktur guna untuk mendapatkan suatu kebenaran atau fakta dari beberapa sumber masalah dengan kajian yang ada untuk menjangkau hasil akhir dan memberikan solusi yang ideal terhadap waktu penelitian yang di lakukan berdasarkan sumber data yang ada dan kondisi suatu tempat, sehingga dengan strategis yang disusun tersebut dapat diolah dengan baik



Gambar 1. Kerangka Kerja Penelitian

2.2 Data Mining

Data mining adalah proses yang menggunakan berbagai perangkat analisis data untuk menemukan pola dan hubungan dalam data yang mungkin dapat digunakan untuk membuat prediksi yang valid [8], [9]. Langkah pertama dan paling sederhana dalam data mining yaitu menggambarkan data – menyimpulkan atribut statistik (seperti rata-rata dan standar deviasi), mereview secara visual menggunakan diagram dan grafik, serta mencari relasi berarti yang potensial antar variabel (misalnya nilai yang sering muncul bersamaan). Mengumpulkan, mengeksplor, dan memilih data yang tepat adalah sangat penting.

2.3 Metode K-Means

K-means clustering merupakan salah satu metode data clustering non-hirarki yang mengelompokkan data dalam bentuk satu atau lebih cluster/kelompok [10]–[12]. Data-data yang memiliki karakteristik yang sama dikelompokkan dalam satu cluster/kelompok dan data yang memiliki karakteristik yang berbeda dikelompokkan dengan cluster/kelompok yang lain sehingga data yang berada dalam satu cluster/kelompok memiliki tingkat variasi yang kecil. Langkah-langkah melakukan clustering dengan metode K-Means adalah sebagai berikut:

1. Pilih jumlah cluster k.
2. Inisialisasi k pusat cluster ini bisa dilakukan dengan berbagai cara. Namun yang paling sering dilakukan adalah dengan cara random. Pusat-pusat cluster diberiduberi nilai awal dengan angka-angka random.
3. Alokasikan semua data/ objek ke cluster terdekat. Kedekatan dua objek ditentukan berdasarkan jarak kedua objek tersebut. Demikian juga kedekatan suatu data ke cluster tertentu ditentukan jarak antara data dengan pusat cluster. Dalam tahap ini perlu dihitung jarak tiap data ke tiap pusat cluster. Jarak paling antara satu data dengan satu cluster tertentu akan menentukan suatu data masuk dalam cluster mana. Untuk menghitung jarak semua data ke setiap titik pusat cluster dapat menggunakan teori jarak Euclidean yang dirumuskan sebagai berikut:

$$D(i,j) = \sqrt{(X_{1i} - Y_{1j})^2 + (X_{2i} - Y_{2j})^2 + \dots + (X_{ki} - Y_{kj})^2} \quad (1)$$

Dimana

$D(i,j)$ = Jarak data ke i ke pusat cluster j

X_{ki} = Data ke i pada atribut data ke k

X_{kj} = Titik pusat ke j pada atribut ke k

4. Hitung kembali pusat cluster dengan keanggotaan cluster yang sekarang. Pusat cluster adalah rata – rata dari semua data/ objek dalam cluster tertentu. Jika dikehendaki bisa juga menggunakan median dari cluster tersebut. Jadi rata – rata (mean) bukan satu – satunya ukuran yang bisa di pakai.
5. Tugaskan lagi setiap objek memakai cluster yang bery. Jika pusat cluster tidak berubah lagi maka proses clustering selesai. Atau, kembali ke langkah nomor 3 sampai pusat cluster tidak berubah lagi.

3. HASIL DAN PEMBAHASAN

3.1 Analisa Penerapan Metode

Setiap lembaga mempunyai sistem operasional yang setiap transaksi kegiatan operasinya selalu dicatat dan didokumentasikan. Pendokumentasian setiap transaksi sangat berguna bagi lembaga tersebut untuk segala keperluan. Data-data tersebut tersimpan dalam sebuah basis data berkepastian besar. Bagi perusahaan, data data yang tersimpan di basis data dan dapat dimanfaatkan untuk membuat laporan penjualan, kontrol inventaris, dan sebagainya, yang pada akhirnya dapat digunakan untuk mengetahui kondisi keuangan perusahaan. Selain kebutuhan pada hal tersebut, perusahaan juga perlu mengetahui pola grafik jumlah produksi berdasarkan data-data produksi yang sudah diketahui sebelumnya, perkiraan nilai penjualan dan layout swalayan , serta pola pembelian konsumen, analisis kebutuhan pasar yang sangat signifikan dalam sumbangan laba perusahaan, dan sebagainya. Semua hal tersebut bisa didapatkan dengan memanfaatkan data-data yang sudah ada di dalam basis data.

Dalam statistik dan mesin pembelajaran. Pengelompokan K-Means merupakan metode analisis kelompok yang mengarah pada pemartisian N objek pengamatan ke dalam K kelompok (cluster) di mana setiap objek pengamatan dimiliki oleh sebuah kelompok dengan mean (rata-rata) terdekat, mirip dengan algoritma Expectation Maximization untuk Gaussian Mixture di mana keduanya mencoba untuk menemukan pusat dari kelompok dalam data sebanyak iterasi perbaikan yang dilakukan oleh kedua algoritma.

K-Means merupakan salah satu metode pengelompokkan dan non hierarki (sekatan) yang berusaha mempartisi data yang ada ke dalam bentuk dua atau lebih kelompok. Metode ini mempartisi data yang ada ke dalam bentuk dua atau lebih kelompok. Metode ini mempartisi data ke dalam kelompok sehingga data berkarakteristik sama dimasukkan ke dalam satu kelompok yang sama dan data yang berkarakteristik berbea dikelompokkan ke dalam kelompok yang lain. Adapaun tujuan pengelompokkan data ini adalah untuk meminimalkan fungsi objektif yang diset dalam proses pengelompokkan.

3.1.1 Penerapan Metode K-Means

Pengelompokan produk merupakan tahap awal dari pengolahan data. Pada tahap ini, produk-produk yang berada di dalam swalayan dikelompokkan berdasarkan beberapa pertimbangan. Pengelompokan produk dengan memperlihatkan pertimbangan current layout dan jenis produk. Penelitian ini tidak membahas persediaan, sehingga meskipun “layout” diubah diharapkan tidak akan mempengaruhi persediaan. Dengan memperlihatkan current layout maka akan diketahui banyaknya persediaan produk yang biasa diletakkan di Swalayan. Dari sini dapat dibuat pembagian kelompok produk

berdasarkan current layout. Dalam current layout, penyajian produk dalam toko dibagi menjadi empat macam yaitu penyajian dalam rak, etalase kaca, pendingin dan bebas. Rak dibagi menjadi dua macam yaitu rak utama yang terdapat di tengah-tengah ruangan. Rak utama dibagi menjadi dua bagian yaitu rak besar dan kecil. Selain rak utama dan rak samping yang berdiri melekat pada dinding. Etalase kaca terdapat pada bagian depan toko. Etalase ini terlihat seperti stand yang dijaga oleh satu atau dua karyawan. Proses pembelian produknya pun berbeda dengan penyajian rak.

Pembagian berdasarkan jenis produk dilakukan dengan memperlihatkan pertimbangan sebelumnya. Produk yang memilikisifat sama atau sejenis dikelompokkan menjadi satu. Meskipun secara pertimbangan penyajian rak produk tersebut berada pada rak yang berbeda namun jika memiliki tipe dan jenis yang sama, maka tetap dikelompokkan menjadi satu kelompok produk. Contoh produk yang berbeda dalam rak namun berbeda tipe dan jenis sama yaitu minuman botol, snack, dan sabun cuci. Sebaliknya, produk yang berbeda dalam satu rak namun berbeda jenis dan memungkinkan untuk dipisahkan meskipun dalam current layout disatukan. Kemungkinan produk ini dipisahkan karena memiliki persediaan yang banyak, contoh produk yang mengalami penguraian ini adalah minuman botol dan minuman kotak yang dijadikan kelompok sendiri. Hasil pengelompokan produk dapat dilihat pada Tabel 1.

Tabel 1. Data Penjualan Mingguan Swalayan Prima Smart

| No | Struk ke- | Merek Produk | Jumlah |
|----|-----------|---|--------|
| 1 | 00021 | Berbeard,Sari roti,Kratindeng, Canned food, Kecap | 5 |
| 2 | 00022 | Blueband, Beras sunco, Apel,panadol | 4 |
| 3 | 00023 | Dancow, Kue bohong,Anggur,Indocafe | 4 |
| 4 | 00024 | Marboro,Fanta,kapal api,The Melati | 4 |
| 5 | 00025 | Bolpoint,Clasmild,Tisu Paseo | 3 |
| 6 | 00026 | Palmia,semangka,Promag,Kopi Sidikalang | 4 |
| 7 | 00027 | Mentos,Antangin,Teh Melati,Kit Mobil | 4 |
| 8 | 00028 | M150,Ajinamoto,Matrai,Camel | 4 |
| 9 | 00029 | Pisang,Teh Melati,Bodrex,Segita biru,Lays | 5 |
| 10 | 00030 | Beras Sunco,Sarden,Dancow,M150 | 4 |
| 11 | 00041 | Sajiku,Kertas A4, Kapal Api,Tisu paseo,Teh Wangi | 5 |
| 12 | 00042 | Coco cola,Bimoli,kratindeng,Royco,Kopi sidikalang | 5 |
| 13 | 00043 | Lux,Kapal api,Surya 16,Pisang | 4 |
| 14 | 00044 | Royco,Chitatos,Dancow,Nabati | 4 |
| 15 | 00045 | Kratindeng,Kologen,Teh Melati,Marboro | 4 |
| 16 | 00046 | Somperna,Sprint,Bolpoint | 3 |
| 17 | 00047 | Promaag,Nescafe, Anggur,Chitatos | 4 |
| 18 | 00048 | Marboro,Kecap Bango,Biore | 3 |
| 19 | 00049 | Panadol,Kopi Sidikalang,Kertas A4,Teh Sari | 4 |
| 20 | 00050 | Lifeboy,Sari Roti,Marboro,Anggur | 4 |

Tabel 2. Data Jenis Prodak Swalayan Prima Smart

| No | Merek Produk Makanan | Merek produk Minuman | Kebutuhan Rumah Tangga |
|----|--|----------------------|------------------------|
| 1 | Sari roti | Berbeard | Antangin |
| 2 | Canned food | Kopi Sidikalang | Beras sunco |
| 3 | Semangka | Fanta | Tisu Paseo |
| 4 | Lays | Teh Sari | Kecap |
| 5 | Ikan sadern | M150 | sajiku |
| 6 | Chitatos | Coco cola | Tepung segitiga biru |
| 7 | Nabati | Teh Melat | ajinamoto |
| 8 | Roti bohong | Nescafe | Sampo Life boy |
| 9 | Anggur | Susu Dancow | Kecap Asin |
| 10 | Instant Cup Noodle Ayam Pedas 90G | Yougort | Garam |
| 11 | Ufo Mi Goreng Ala Indonesia 88G | Lee mineral | Minyak Palmia |
| 12 | SarimiMie Goreng Puass Fried Chicken 100G. | Kopi good day | Obat kologen |
| 13 | Migelas Mie Instant Paket Keluarga Soto Ayam 6X28g | Teh kotak | Royco |
| 14 | Migelas Mie Instant Ayam Bawang 6X28g | Pocary sweet | Kit mobil |
| 15 | Potato | karatindeng | Matrai |
| 16 | Popcron | Ber bintang | Kertas A4 |
| 17 | Soyjoy | Cap kaki tiga | Minyak bimoli |
| 18 | Sari roma | Kopi luak | Bolpoint |
| 19 | Kacang Almond | Kopi golda | Show klin |
| 20 | Mie instan | Fresh tea | Popok bayi |

Tabel 3. Data Penentuan Jenis Produk Swalayan Prima Smart

| No | Struk Ke- | Makanan Ringan | Minuman | Kebuthan Rumah Tangga |
|----|-----------|----------------|---------|-----------------------|
| 1 | 00021 | 2 | 2 | 2 |
| 2 | 00022 | 1 | 0 | 3 |

| No | Struk Ke- | Makanan Ringan | Minuman | Kebutuhan Rumah Tangga |
|----|-----------|----------------|---------|------------------------|
| 3 | 00023 | 1 | 1 | 1 |
| 4 | 00024 | 0 | 2 | 0 |
| 5 | 00025 | 0 | 1 | 2 |
| 6 | 00026 | 1 | 1 | 2 |
| 7 | 00027 | 0 | 1 | 2 |
| 8 | 00028 | 1 | 1 | 2 |
| 9 | 00029 | 2 | 1 | 2 |
| 10 | 00030 | 1 | 1 | 2 |
| 11 | 00041 | 0 | 1 | 3 |
| 12 | 00042 | 0 | 3 | 2 |
| 13 | 00043 | 1 | 1 | 1 |
| 14 | 00044 | 2 | 0 | 1 |
| 15 | 00045 | 0 | 2 | 1 |
| 16 | 00046 | 0 | 1 | 1 |
| 17 | 00047 | 2 | 1 | 1 |
| 18 | 00048 | 0 | 0 | 2 |
| 19 | 00049 | 0 | 2 | 2 |
| 20 | 00050 | 1 | 0 | 2 |

Dari tabel diatas dapat dilihat dataset berjumlah 16 record struk belanja dengan 9 jenis produk. Dataset yang dimasukkan berupa merk produk apa saja yang dibeli dan berapa jumlah merk produk yang dibeli oleh konsumen didalam satu struk kemudian dikelompokkan berdasarkan jenis produknya.

K-Means adalah salah satu metode data clustering non hirarki yang berusaha mempartisi data yang ada ke dalam bentuk satu atau lebih cluster. Metode ini mempartisi data ke dalam cluster sehingga data yang memiliki karakteristik yang sama dikelompokkan ke dalam satu cluster yang sama dan data yang mempunyai katakteristik yang berbeda dikelompokkan ke dalam cluster yang lain.

Tahapan clustering dengan menggunakan K-Means dimulai dengan pembentukan cluster dimana penulis di sini membagi cluster maksimal sesuai dengan jumlah dataset, karena nantinya setiap cluster yang diinputkan dapat memilih data sheet untuk menentukan nilai centroid akan tetapi dengan syarat nilai centroid tidak dapat muncul dengan nilai yang sama. Kemudian dilakukan proses perhitungan jarak dari centroid dengan menggunakan Euclidean Distance, dimana centroid diperoleh dari objek dataset yang telah diinputkan.

Setelah diketahui nilai centroid pada masing-masing cluster, tahap selanjutnya adalah melakukan perhitungan untuk menentukan jarak setiap data dengan centroid yang telah dibentuk dengan menggunakan rumus euclidiance distance. Hasil dari perhitungan jarak dengan rumus euclidiance distance ini akan berpengaruh pada penempatan setiap merk produk ke cluster tertentu. Setelah hasil dari jarak setiap data diperoleh tahapan selanjutnya adalah mengelompokkan jarak data minuman atau yang terkecil dari masing-masing cluster berdasarkan parameter jenis produknya.

Menentukan titik pusat awal cluster pada metode Data Mining algoritma K-Means Clustering adalah dipilih secara random. Adapun titik pusat awal cluster yang dipilih oleh penulis secara random dapat dilihat pada tabel 4. berikut:

Tabel 4. Data dengan titik awal cluster

| No | Struk Ke- | Makanan Ringan | Minuman | Kebutuhan Rumah Tangga |
|----|-----------|----------------|---------|------------------------|
| 1 | 00021 | 2 | 2 | 2 |
| 2 | 00022 | 1 | 0 | 3 |
| 3 | 00023 | 1 | 1 | 1 |
| 4 | 00024 | 0 | 2 | 0 |
| 5 | 00025 | 0 | 1 | 2 |
| 6 | 00026 | 1 | 1 | 2 |
| 7 | 00027 | 0 | 1 | 2 |
| 8 | 00028 | 1 | 1 | 2 |
| 9 | 00029 | 2 | 1 | 2 |
| 10 | 00030 | 1 | 1 | 2 |
| 11 | 00041 | 0 | 1 | 3 |
| 12 | 00042 | 0 | 3 | 2 |
| 13 | 00043 | 1 | 1 | 1 |
| 14 | 00044 | 2 | 0 | 1 |
| 15 | 00045 | 0 | 2 | 1 |
| 16 | 00046 | 0 | 1 | 1 |
| 17 | 00047 | 2 | 1 | 1 |
| 18 | 00048 | 0 | 0 | 2 |
| 19 | 00049 | 0 | 2 | 2 |
| 20 | 00050 | 1 | 0 | 2 |

Dari tabel 4. tersebut, dapat diambil kesimpulan masing-masing titik awal Cluster berdasarkan urutan yang akan dilakukan penulis yakni daerah rawan tindak kriminal, sedang dan tidak rawan tindak kriminal adalah sebagai berikut:

1. Pusat Cluster I: Data ke-6 Struk Pembelian dengan Makanan Ringan=1 dan Minuman=1 dan kebutuhan rumah tangga =2
2. Pusat Cluster II: Data ke-17 Struk Pembelian dengan Makanan Ringan=2 dan Minuman=1 dan kebutuhan rumah tangga =1

Pada algoritma K-Means Clustering, menghitung jarak setiap data ke pusat cluster adalah bertujuan untuk menentukan titik pusat cluster yang baru menggunakan rumus Euclidian Distance. Pada tahapan ini, perhitungan jarak setiap data ke titik pusat cluster dilakukan secara perulangan. Dan apabila nilai pusat cluster pada iterasi tersebut sama dengan nilai pusat cluster pada iterasi sebelumnya, perulangan dihentikan karena hasil telah ditemukan. Adapun perhitungan jarak data ke pusat cluster pada masing-masing iterasi adalah sebagai berikut:

1. Iterasi I

| | |
|----------------------------|---|
| Data-1 ke Pusat Cluster I | $= \sqrt{(1-2)^2 + (1-2)^2 + (2-2)^2}$ $= \sqrt{1+1+0}=1,414$ |
| Data-2 ke Pusat Cluster I | $= \sqrt{(1-1)^2 + (1-0)^2 + (2-3)^2}$ $= \sqrt{0+1+1}=1,414$ |
| Data-3 ke Pusat Cluster I | $= \sqrt{(1-1)^2 + (1-1)^2 + (2-1)^2}$ $= \sqrt{0+0+1}=1$ |
| Data-4 ke Pusat Cluster I | $= \sqrt{(1-0)^2 + (1-2)^2 + (2-0)^2}$ $= \sqrt{1+1+4}=2,449$ |
| Data-5 ke Pusat Cluster I | $= \sqrt{(1-0)^2 + (1-1)^2 + (2-2)^2}$ $= \sqrt{1+0+0}=0$ |
| Data-6 ke Pusat Cluster I | $= \sqrt{(1-1)^2 + (1-1)^2 + (2-2)^2}$ $= \sqrt{4+0+0}=2$ |
| Data-7 ke Pusat Cluster I | $= \sqrt{(1-0)^2 + (1-1)^2 + (2-2)^2}$ $= \sqrt{1+0+0}=1$ |
| Data-8 ke Pusat Cluster I | $= \sqrt{(1-1)^2 + (1-1)^2 + (2-2)^2}$ $= \sqrt{4+4+0}=2,828$ |
| Data-9 ke Pusat Cluster I | $= \sqrt{(1-2)^2 + (1-1)^2 + (2-2)^2}$ $= \sqrt{1+0+0}=1$ |
| Data-10 ke Pusat Cluster I | $= \sqrt{(1-1)^2 + (1-1)^2 + (2-2)^2}$ $= \sqrt{0+0+0}=0$ |
| Data-11 ke Pusat Cluster I | $= \sqrt{(1-0)^2 + (1-1)^2 + (2-3)^2}$ $= \sqrt{0+0+1}=1$ |
| Data-12 ke Pusat Cluster I | $= \sqrt{(1-0)^2 + (1-1)^2 + (2-3)^2}$ $= \sqrt{1+0+1}=1,414$ |
| Data-13 ke Pusat Cluster I | $= \sqrt{(1-1)^2 + (1-1)^2 + (2-1)^2}$ $= \sqrt{0+0+1}=1$ |
| Data-14 ke Pusat Cluster I | $= \sqrt{(1-2)^2 + (1-0)^2 + (2-1)^2}$ $= \sqrt{1+1+1}=1,732$ |
| Data-15 ke Pusat Cluster I | $= \sqrt{(1-0)^2 + (1-2)^2 + (2-1)^2}$ $= \sqrt{0+1+1}=1,41$ |
| Data-16 ke Pusat Cluster I | $= \sqrt{(1-0)^2 + (1-1)^2 + (2-1)^2}$ $= \sqrt{1+0+1}=1,414$ |
| Data-17 ke Pusat Cluster I | $= \sqrt{(1-2)^2 + (1-1)^2 + (2-1)^2}$ $= \sqrt{1+0+1}=1,414$ |
| Data-18 ke Pusat Cluster I | $= \sqrt{(1-0)^2 + (1-0)^2 + (2-2)^2}$ $= \sqrt{1+1+0}=1,732$ |
| Data-19 ke Pusat Cluster I | $= \sqrt{(1-0)^2 + (1-2)^2 + (2-2)^2}$ $= \sqrt{1+1+0}=1,732$ |
| Data-20 ke Pusat Cluster I | $= \sqrt{(1-1)^2 + (1-0)^2 + (2-2)^2}$ $= \sqrt{0+1+0}=1$ |
| Data-1 ke Pusat Cluster II | $= \sqrt{(2-2)^2 + (1-2)^2 + (1-2)^2}$ $= \sqrt{0+1+1}=1,414$ |

| | |
|-----------------------------|--|
| Data-2 ke Pusat Cluster II | $= \sqrt{(2-1)^2 + (1-0)^2 + (1-3)^2}$ $= \sqrt{1+1+3}=2,236$ |
| Data-3 ke Pusat Cluster II | $= \sqrt{(2-1)^2 + (1-1)^2 + (1-1)^2}$ $= \sqrt{1+0+0}=1$ |
| Data-4 ke Pusat Cluster II | $= \sqrt{(2-0)^2 + (1-2)^2 + (1-0)^2}$ $= \sqrt{4+1+1}=2,449$ |
| Data-5 ke Pusat Cluster II | $= \sqrt{(2-0)^2 + (1-1)^2 + (1-2)^2}$ $= \sqrt{4+0+1}=3,471$ |
| Data-6 ke Pusat Cluster II | $= \sqrt{(2-1)^2 + (1-1)^2 + (1-2)^2}$ $= \sqrt{1+1+3}=2,236$ |
| Data-7 ke Pusat Cluster II | $= \sqrt{(2-0)^2 + (1-1)^2 + (1-2)^2}$ $= \sqrt{4+0+1}=2,236$ |
| Data-8 ke Pusat Cluster II | $= \sqrt{(2-1)^2 + (1-1)^2 + (1-2)^2}$ $= \sqrt{1+0+1}=1,414$ |
| Data-9 ke Pusat Cluster II | $= \sqrt{(2-2)^2 + (1-1)^2 + (1-2)^2}$ $= \sqrt{1+0+1}=1,414$ |
| Data-10 ke Pusat Cluster II | $= \sqrt{(2-1)^2 + (1-1)^2 + (1-2)^2}$ $= \sqrt{1+0+1}=1,414$ |
| Data-11 ke Pusat Cluster II | $= \sqrt{(2-0)^2 + (1-1)^2 + (1-3)^2}$ $= \sqrt{4+0+4}=2,828$ |
| Data-12 ke Pusat Cluster II | $= \sqrt{(2-0)^2 + (1-3)^2 + (1-2)^2}$ $= \sqrt{4+4+1}=3$ |
| Data-13 ke Pusat Cluster II | $= \sqrt{(2-1)^2 + (1-1)^2 + (1-1)^2}$ $= \sqrt{1+0+0}=1,414$ |
| Data-14 ke Pusat Cluster II | $= \sqrt{(2-2)^2 + (1-0)^2 + (1-1)^2}$ $= \sqrt{0+1+0}=1$ |
| Data-15 ke Pusat Cluster II | $= \sqrt{(2-0)^2 + (1-2)^2 + (1-1)^2}$ $= \sqrt{4+1+0}=2,236$ |
| Data-16 ke Pusat Cluster II | $= \sqrt{(2-0)^2 + (1-1)^2 + (1-1)^2}$ $= \sqrt{4+0+0}=2$ |
| Data-17 ke Pusat Cluster II | $= \sqrt{(2-2)^2 + (1-1)^2 + (1-1)^2}$ $= \sqrt{0+0+0}=0$ |
| Data-18 ke Pusat Cluster II | $= \sqrt{(2-0)^2 + (1-0)^2 + (1-2)^2}$ $= \sqrt{4+1+1}=2,449$ |
| Data-19 ke Pusat Cluster II | $= \sqrt{(2-0)^2 + (1-2)^2 + (1-2)^2}$ $= \sqrt{4+1+1}=2,449$ |
| Data-20 ke Pusat Cluster II | $= \sqrt{(2-1)^2 + (1-0)^2 + (1-2)^2}$ $= \sqrt{1+1+1}=1,732$ |

Hasil perhitungan menggunakan Euclidian Distance pada iterasi I tersebut diatas dapat dilihat pada tabel 4.5 berikut:

Tabel 5. Hasil Iterasi I

| Daka ke- | Struk Ke- | Jarak Ke Centroid | |
|----------|-----------|-------------------|-------|
| | | C1 | C2 |
| 1 | 00021 | 1,414 | 1,414 |
| 2 | 00022 | 1,414 | 2,236 |
| 3 | 00023 | 1 | 1 |
| 4 | 00024 | 2,449 | 2,449 |
| 5 | 00025 | 0 | 3,471 |
| 6 | 00026 | 2 | 2,236 |
| 7 | 00027 | 1 | 2,236 |
| 8 | 00028 | 2,828 | 1,414 |
| 9 | 00029 | 1 | 1 |
| 10 | 00030 | 0 | 1,414 |
| 11 | 00041 | 1 | 2,828 |
| 12 | 00042 | 1,414 | 3 |

| Daka ke- | Struk Ke- | Jarak Ke Centroid | |
|----------|-----------|-------------------|-------|
| | | C1 | C2 |
| 13 | 00043 | 1 | 1,414 |
| 14 | 00044 | 1,732 | 1 |
| 15 | 00045 | 1,41 | 2,236 |
| 16 | 00046 | 1,414 | 2 |
| 17 | 00047 | 1,414 | 0 |
| 18 | 00048 | 1,732 | 2,449 |
| 19 | 00049 | 1,732 | 2,449 |
| 20 | 00050 | 1 | 1,732 |

Tahap selanjutnya adalah menentukan jarak terpendek data terhadap pusat cluster awal yang dipilih secara random sebelumnya. Menentukan jarak terpendek dengan pusat cluster data dapat menggunakan rumus MIN (Cij : Ckj). Adapun hasil penentuan jarak terpendek terhadap pusat cluster yang telah ditabulasikan dapat dilihat pada tabel 4.6 berikut:

Tabel 6. Tabulasi Binner Jarak Terdekat Data terhadap Cluster

| Daka ke- | Struk Ke- | Jarak Ke Centroid | | Cluster |
|----------|-----------|-------------------|----|---------|
| | | C1 | C2 | |
| 1 | 00021 | 1 | 1 | C1, C2 |
| 2 | 00022 | 1 | 0 | C1 |
| 3 | 00023 | 1 | 1 | C1, C2 |
| 4 | 00024 | 1 | 1 | C1, C2 |
| 5 | 00025 | 1 | 0 | C1 |
| 6 | 00026 | 1 | 0 | C1 |
| 7 | 00027 | 1 | 0 | C1 |
| 8 | 00028 | 0 | 1 | C2 |
| 9 | 00029 | 1 | 1 | C1, C2 |
| 10 | 00030 | 1 | 0 | C1 |
| 11 | 00041 | 1 | 0 | C1 |
| 12 | 00042 | 1 | 0 | C1 |
| 13 | 00043 | 1 | 0 | C1 |
| 14 | 00044 | 0 | 1 | C2 |
| 15 | 00045 | 1 | 0 | C1 |
| 16 | 00046 | 1 | 0 | C1 |
| 17 | 00047 | 0 | 1 | C2 |
| 18 | 00048 | 1 | 0 | C1 |
| 19 | 00049 | 1 | 0 | C1 |
| 20 | 00050 | 1 | 0 | C1 |

Berdasarkan nilai minimum yang telah ditentukan menggunakan rumus dan hasil terlihat pada tabel 6. tersebut di atas, maka diperoleh hasil pengelompokan seperti terlihat pada tabel 7. berikut

Tabel 7. Hasil Pengelompokan

| Kelompok (Cluster) | Anggota Cluster | Jumlah Anggota |
|--------------------|--|----------------|
| 1 | {1,2,3,4,5,6,7,9,10,11,12,13,15,16,18,19,20} | 17 |
| 2 | {1,3,4,8,9,14,17} | 7 |

Tahapan selanjutnya adalah menentukan titik pusat cluster baru yakni dengan mencari rata-rata semua anggota cluster dengan cara menjumlahkan seluruh value anggota cluster dibagi jumlah anggota. Adapun titik pusat cluster baru beserta value setiap variable adalah sebagai berikut:

$$C_1(X1) = \frac{\text{Value Makanan } R (2 + 1 + 1 + 1 + 0 + 1 + 0 + 2 + 1 + 0 + 0 + 1 + 0 + 0 + 0 + 0 + 1)}{\text{jlh anggota}(17)}$$

$$= \frac{11}{17} = 0,647$$

$$C_1(X2) = \frac{\text{Value Minuman } (2 + 0 + 1 + 2 + 1 + 1 + 1 + 1 + 1 + 1 + 3 + 1 + 2 + 1 + 0 + 2 + 0)}{\text{jlh anggota } (17)}$$

$$= \frac{21}{17} = 1,235$$

$$C_1(X2) = \frac{\text{ValueK RT } (2 + 3 + 1 + 0 + 2 + 2 + 2 + 2 + 2 + 2 + 3 + 2 + 1 + 1 + 1 + 1 + 2 + 2 + 2 + 2)}{\text{jlh anggota } (17)}$$

$$= \frac{34}{17} = 2$$

$$C_2(X1) = \frac{\text{Value Makanan R } (2 + 1 + 0 + 1 + 2 + 2 + 2)}{\text{jlh anggota } (7)}$$

$$= \frac{10}{7} = 1,428$$

$$C_2(X2) = \frac{\text{Value Minuman } (2 + 1 + 2 + 1 + 1 + 0 + 1)}{\text{jlh anggota } (7)}$$

$$= \frac{(8)}{7} = 1,142$$

$$C_2(X2) = \frac{\text{Value Kebutuhan RT } (2 + 1 + 0 + 2 + 2 + 1 + 1 + 1)}{\text{jlh anggota } (7)}$$

$$= \frac{9}{7} = 1,285$$

Berdasarkan proses perhitungan di atas, maka diperoleh centroid baru sebagai tolak ukur pada iterasi selanjutnya dengan nilai sebagai berikut:

$$C_1 = [0,647, 1,235, 2]$$

$$C_2 = [1,428, 1,142, 1,285]$$

2. Interasi 2

$$\text{Data-1 ke Pusat Cluster I} = \sqrt{(0,647 - 2)^2 + (1,235 + 2)^2 + (2 + 2)^2}$$

$$= \sqrt{1,834 + 10,465 + 16} = 5,319$$

$$\text{Data-2 ke Pusat Cluster I} = \sqrt{(0,647 - 1)^2 + (1,235 - 0)^2 + (2 - 3)^2}$$

$$= \sqrt{0,124 + 1,525 + 1} = 1,627$$

$$\text{Data-3 ke Pusat Cluster I} = \sqrt{(0,647 - 1)^2 + (1,235 - 1)^2 + (2 - 1)^2}$$

$$= \sqrt{0,124 + 0,055 + 1} = 1,085$$

$$\text{Data-4 ke Pusat Cluster I} = \sqrt{(0,647 - 0)^2 + (1,235 - 2)^2 + (2 - 0)^2}$$

$$= \sqrt{0,418 + 0,005 + 4} = 2,103$$

$$\text{Data-5 ke Pusat Cluster I} = \sqrt{(0,647 - 0)^2 + (1,235 - 1)^2 + (2 - 2)^2}$$

$$= \sqrt{0,418 + 0,055 + 0} = 0,687$$

$$\text{Data-6 ke Pusat Cluster I} = \sqrt{(0,647 - 1)^2 + (1,235 - 1)^2 + (2 - 2)^2}$$

$$= \sqrt{0,124 + 0,055 + 0} = 0,423$$

$$\text{Data-7 ke Pusat Cluster I} = \sqrt{(0,647 - 0)^2 + (1,235 - 1)^2 + (2 - 2)^2}$$

$$= \sqrt{0,418 + 0,055 + 0} = 0,687$$

$$\text{Data-8 ke Pusat Cluster I} = \sqrt{(0,647 - 1)^2 + (1,235 - 1)^2 + (2 - 2)^2}$$

$$= \sqrt{0,124 + 0,055 + 0} = 0,423$$

$$\text{Data-9 ke Pusat Cluster I} = \sqrt{(0,647 - 2)^2 + (1,235 - 1)^2 + (2 - 2)^2}$$

$$= \sqrt{1,830 + 0,055 + 0} = 1,372$$

$$\text{Data-10 ke Pusat Cluster I} = \sqrt{(0,647 - 1)^2 + (1,235 - 1)^2 + (2 - 2)^2}$$

$$= \sqrt{0,124 + 0,055 + 0} = 0,423$$

$$\text{Data-11 ke Pusat Cluster I} = \sqrt{(0,647 - 0)^2 + (1,235 - 1)^2 + (2 - 3)^2}$$

$$= \sqrt{0,418 + 0,055 + 1} = 1,213$$

$$\text{Data-12 ke Pusat Cluster I} = \sqrt{(0,647 - 0)^2 + (1,235 - 3)^2 + (2 - 2)^2}$$

$$= \sqrt{0,418 + 3,11 + 0} = 1,878$$

$$\text{Data-13 ke Pusat Cluster I} = \sqrt{(0,647 - 1)^2 + (1,235 - 1)^2 + (2 - 2)^2}$$

$$= \sqrt{0,124 + 0,055 + 0} = 0,423$$

$$\text{Data-14 ke Pusat Cluster I} = \sqrt{(0,647 - 2)^2 + (1,235 - 0)^2 + (2 - 1)^2}$$

$$= \sqrt{1,830 + 1,525 + 1} = 2,086$$

$$\text{Data-15 ke Pusat Cluster I} = \sqrt{(0,647 - 0)^2 + (1,235 - 2)^2 + (2 - 1)^2}$$

$$= \sqrt{0,418 + 0,585 + 1} = 1,415$$

$$\text{Data-16 ke Pusat Cluster I} = \sqrt{(0,647 - 0)^2 + 1}$$

$$\begin{aligned}
 &= \sqrt{0,418 + 0,055 + 1} = 1,213 \\
 \text{Data-17 ke Pusat Cluster I} &= \sqrt{(0,647 - 2)^2 + (1,235 - 2)^2 + (2 - 1)^2} \\
 &= \sqrt{1,830 + 0,585 + 1} = 1,847 \\
 \text{Data-18 ke Pusat Cluster I} &= \sqrt{(0,647 - 0)^2 + (1,235 - 0)^2 + (2 - 2)^2} \\
 &= \sqrt{0,418 + 1,525 + 0} = 1,393 \\
 \text{Data-19 ke Pusat Cluster I} &= \sqrt{(0,647 - 0)^2 + (1,235 - 2)^2 + (2 - 2)^2} \\
 &= \sqrt{0,418 + 0,585 + 0} = 1,001 \\
 \text{Data-20 ke Pusat Cluster I} &= \sqrt{(0,647 - 1)^2 + (1,235 - 0)^2 + (2 - 2)^2} \\
 &= \sqrt{0,124 + 1,525 + 0} = 1,284 \\
 \text{Data-1 ke Pusat Cluster II} &= \sqrt{(1,4281 - 2)^2 + (1,142 - 2)^2 + (1,285 - 2)^2} \\
 &= \sqrt{0,327 + 0,736 + 0,511} = 1,254 \\
 \text{Data-2 ke Pusat Cluster II} &= \sqrt{(1,4281 - 1)^2 + (1,142 - 0)^2 + (1,285 - 3)^2} \\
 &= \sqrt{0,183 + 1,304 + 2,941} = 2,104 \\
 \text{Data-3 ke Pusat Cluster II} &= \sqrt{(1,4281 - 1)^2 + (1,142 - 1)^2 + (1,285 - 1)^2} \\
 &= \sqrt{0,183 + 0,020 + 0,081} = 0,532 \\
 \text{Data-4 ke Pusat Cluster II} &= \sqrt{(1,4281 - 0)^2 + (1,142 - 2)^2 + (1,285 - 0)^2} \\
 &= \sqrt{2,039 + 0,736 + 1,651} = 2,103 \\
 \text{Data-5 ke Pusat Cluster II} &= \sqrt{(1,4281 - 0)^2 + (1,142 - 1)^2 + (1,285 - 2)^2} \\
 &= \sqrt{2,039 + 0,020 + 0,511} = 1,603 \\
 \text{Data-6 ke Pusat Cluster II} &= \sqrt{(1,4281 - 1)^2 + (1,142 - 1)^2 + (1,285 - 2)^2} \\
 &= \sqrt{0,183 + 0,020 + 0,511} = 0,844 \\
 \text{Data-7 ke Pusat Cluster II} &= \sqrt{(1,4281 - 0)^2 + (1,142 - 1)^2 + (1,285 - 2)^2} \\
 &= \sqrt{2,039 + 0,020 + 0,511} = 1,603 \\
 \text{Data-8 ke Pusat Cluster II} &= \sqrt{(1,4281 - 1)^2 + (1,142 - 1)^2 + (1,285 - 2)^2} \\
 &= \sqrt{0,183 + 0,020 + 0,511} = 0,844 \\
 \text{Data-9 ke Pusat Cluster II} &= \sqrt{(1,4281 - 2)^2 + (1,142 - 1)^2 + (1,285 - 2)^2} \\
 &= \sqrt{0,327 + 0,020 + 0,511} = 0,926 \\
 \text{Data-10 ke Pusat Cluster II} &= \sqrt{(11,4281 - 1)^2 + (1,142 - 1)^2 + (1,285 - 2)^2} \\
 &= \sqrt{0,183 + 0,020 + 0,511} = 0,844 \\
 \text{Data-11 ke Pusat Cluster II} &= \sqrt{(1,4281 - 0)^2 + (1,142 - 1)^2 + (1,285 - 3)^2} \\
 &= \sqrt{2,039 + 0,020 + 2,941} = 2,236 \\
 \text{Data-12 ke Pusat Cluster II} &= \sqrt{(1,4281 - 0)^2 + (1,142 - 3)^2 + (1,285 - 2)^2} \\
 &= \sqrt{2,039 + 3,452 + 0,511} = 2,449 \\
 \text{Data-13 ke Pusat Cluster II} &= \sqrt{(1,4281 - 1)^2 + (1,142 - 1)^2 + (1,285 - 1)^2} \\
 &= \sqrt{0,183 + 0,020 + 0,081} = 0,532 \\
 \text{Data-14 ke Pusat Cluster II} &= \sqrt{(1,4281 - 2)^2 + (1,142 - 0)^2 + (1,285 - 1)^2} \\
 &= \sqrt{0,327 + 0,020 + 0,081} = 0,654 \\
 \text{Data-15 ke Pusat Cluster II} &= \sqrt{(1,4281 - 0)^2 + (1,142 - 2)^2 + (1,285 - 1)^2} \\
 &= \sqrt{2,039 + 0,736 + 0,081} = 1,689 \\
 \text{Data-16 ke Pusat Cluster II} &= \sqrt{(1,4281 - 0)^2 + (1,142 - 1)^2 + (1,285 - 1)^2} \\
 &= \sqrt{2,039 + 0,020 + 0,081} = 1,462 \\
 \text{Data-17 ke Pusat Cluster II} &= \sqrt{(1,4281 - 1)^2 + (1,142 - 1)^2 + (1,285 - 1)^2} \\
 &= \sqrt{0,183 + 0,020 + 0,081} = 0,532 \\
 \text{Data-18 ke Pusat Cluster II} &= \sqrt{(1,4281 - 0)^2 + (1,142 - 0)^2 + (1,285 - 2)^2} \\
 &= \sqrt{2,196 + 1,304 + 0,511} = 2,002 \\
 \text{Data-19 ke Pusat Cluster II} &= \sqrt{(1,4281 - 0)^2 + (1,142 - 2)^2 + (1,285 - 2)^2} \\
 &= \sqrt{2,039 + 0,736 + 0,511} = 1,812 \\
 \text{Data-20 ke Pusat Cluster II} &= \sqrt{(1,4281 - 1)^2 + (1,142 - 0)^2 + (1,285 - 2)^2} \\
 &= \sqrt{0,183 + 1,304 + 0,511} = 1,413
 \end{aligned}$$

Hasil perhitungan menggunakan Euclidian Distance pada iterasi I tersebut diatas dapat dilihat pada Tabel 8. berikut:

Tabel 8. Hasil Iterasi II

| Daka ke- | Struk Ke- | Jarak Ke Centroid | |
|----------|-----------|-------------------|-------|
| | | C1 | C2 |
| 1 | 00021 | 5,319 | 1,254 |
| 2 | 00022 | 1,085 | 2,104 |
| 3 | 00023 | 1,085 | 0,532 |
| 4 | 00024 | 2,103 | 2,103 |
| 5 | 00025 | 0,687 | 1,603 |
| 6 | 00026 | 0,423 | 0,844 |
| 7 | 00027 | 0,687 | 1,603 |
| 8 | 00028 | 0,423 | 0,844 |
| 9 | 00029 | 1,372 | 0,926 |
| 10 | 00030 | 0,423 | 0,844 |
| 11 | 00041 | 1,213 | 2,236 |
| 12 | 00042 | 1,878 | 2,449 |
| 13 | 00043 | 0,423 | 0,532 |
| 14 | 00044 | 2,086 | 0,654 |
| 15 | 00045 | 1,415 | 1,689 |
| 16 | 00046 | 1,213 | 1,462 |
| 17 | 00047 | 1,847 | 0,532 |
| 18 | 00048 | 1,393 | 2,002 |
| 19 | 00049 | 1,001 | 1,812 |
| 20 | 00050 | 1,284 | 1,413 |

Adapun hasil penentuan jarak terpendek terhadap pusat cluster yang telah ditabulasikan ke dalam biner dapat dilihat pada tabel 9. berikut:

Tabel 9. Tabulasi Binner Iterasi II

| Daka ke- | Struk Ke- | Jarak Ke Centroid | | Cluster |
|----------|-----------|-------------------|----|---------|
| | | C1 | C2 | |
| 1 | 00021 | 0 | 1 | C2 |
| 2 | 00022 | 1 | 0 | C1 |
| 3 | 00023 | 0 | 1 | C2 |
| 4 | 00024 | 1 | 1 | C1C2 |
| 5 | 00025 | 1 | 0 | C1 |
| 6 | 00026 | 1 | 0 | C1 |
| 7 | 00027 | 1 | 0 | C1 |
| 8 | 00028 | 1 | 0 | C1 |
| 9 | 00029 | 0 | 1 | C2 |
| 10 | 00030 | 1 | 0 | C1 |
| 11 | 00041 | 1 | 0 | C1 |
| 12 | 00042 | 1 | 0 | C1 |
| 13 | 00043 | 1 | 0 | C1 |
| 14 | 00044 | 0 | 1 | C2 |
| 15 | 00045 | 1 | 0 | C1 |
| 16 | 00046 | 1 | 0 | C1 |
| 17 | 00047 | 0 | 1 | C2 |
| 18 | 00048 | 1 | 0 | C1 |
| 19 | 00049 | 1 | 0 | C1 |
| 20 | 00050 | 1 | 0 | C1 |

Berdasarkan nilai minimum yang telah ditentukan menggunakan rumus dan hasil terlihat pada tabel 9. tersebut di atas, maka diperoleh hasil pengelompokan seperti terlihat pada tabel 10. Berikut

Tabel 10. Hasil Pengelompokan

| Kelompok (Cluster) | Anggota Cluster | Jumlah Anggota |
|--------------------|--|----------------|
| 1 | {2,4,5,6,7,8,10,11,12,13,15,16,18,19,20} | 15 |
| 2 | {1,3,9,14,17} | 5 |

Tahapan selanjutnya adalah menentukan titik pusat cluster baru yakni dengan mencari rata-rata semua anggota cluster dengan cara menjumlahkan seluruh value anggota cluster dibagi jumlah anggota. Adapun titik pusat cluster baru beserta value setiap variable adalah sebagai berikut:

$$C_1(X1) = \frac{C_1(X1) \text{ Value Makanan R } (2 + 1 + 1 + 1 + 0 + 1 + 0 + 2 + 1 + 0 + 0 + 1 + 0 + 0 + 0 + 0 + 1)}{j\text{lh anggota}(15)} = \frac{11}{15} = 0,733$$

$$C_1(X2) = \frac{C_1(X2) \text{ Value Minuman } (2 + 0 + 1 + 2 + 1 + 1 + 1 + 1 + 1 + 1 + 3 + 1 + 2 + 1 + 0 + 2 + 0)}{j\text{lh anggota } (17)} = \frac{21}{15} = 1,4$$

$$C_1(X2) = \frac{C_1(X2) \text{ Value K RT } (2 + 3 + 1 + 0 + 2 + 2 + 2 + 2 + 2 + 2 + 3 + 2 + 1 + 1 + 1 + 2 + 2 + 2 + 2)}{j\text{lh anggota } (17)} = \frac{34}{15} = 2,266$$

$$C_2(X1) = \frac{C_2(X1) \text{ Value Makanan R } (2 + 1 + 0 + 1 + 2 + 2 + 2)}{j\text{lh anggota } (7)} = \frac{10}{6} = 1,666$$

$$C_2(X2) = \frac{C_2(X2) \text{ Value Minuman } (2 + 1 + 2 + 1 + 1 + 0 + 1)}{j\text{lh anggota } (7)} = \frac{8}{6} = 1,333$$

$$C_2(X2) = \frac{C_2(X2) \text{ Value Kebutuhan RT } (2 + 1 + 0 + 2 + 2 + 1 + 1 + 1)}{j\text{lh anggota } (7)} = \frac{9}{6} = 1,5$$

Berdasarkan proses perhitungan di atas, maka diperoleh centroid baru sebagai tolak ukur pada iterasi selanjutnya dengan nilai sebagai berikut:

$$C_1 = [0,733, 1,4, 2,266]$$

$$C_2 = [1,666, 1,333, 1,5]$$

3. Iterasi 3

$$\text{Data-1 ke Pusat Cluster I} = \sqrt{(0,733 - 2)^2 + (1,4 - 2)^2 + (2,266 - 2)^2} = \sqrt{1,605 + 0,36 + 0,070} = 1,426$$

$$\text{Data-2 ke Pusat Cluster I} = \sqrt{(0,733 - 1)^2 + (1,4 - 0)^2 + (2,266 - 3)^2} = \sqrt{0,071 + 1,96 + 0,538} = 1,602$$

$$\text{Data-3 ke Pusat Cluster I} = \sqrt{(0,733 - 1)^2 + (1,4 - 1)^2 + (2,266 - 1)^2} = \sqrt{0,863 + 0,16 + 1,602} = 1,620$$

$$\text{Data-4 ke Pusat Cluster I} = \sqrt{(0,733 - 0)^2 + (1,4 - 2)^2 + (2,266 - 0)^2} = \sqrt{0,537 + 0,36 + 5,134} = 1,858$$

$$\text{Data-5 ke Pusat Cluster I} = \sqrt{(0,733 - 0)^2 + (1,4 - 1)^2 + (2,266 - 2)^2} = \sqrt{1,605 + 0,16 + 0,070} = 1,354$$

$$\text{Data-6 ke Pusat Cluster I} = \sqrt{(0,733 - 1)^2 + (1,4 - 1)^2 + (2,266 - 2)^2} = \sqrt{0,071 + 0,16 + 0,070} = 0,548$$

$$\text{Data-7 ke Pusat Cluster I} = \sqrt{(0,733 - 0)^2 + (1,4 - 1)^2 + (2,266 - 2)^2} = \sqrt{0,537 + 0,36 + 0,070} = 0,983$$

$$\text{Data-8 ke Pusat Cluster I} = \sqrt{(0,733 - 1)^2 + (1,4 - 1)^2 + (2,266 - 2)^2} = \sqrt{0,071 + 0,36 + 0,070} = 0,707$$

$$\text{Data-9 ke Pusat Cluster I} = \sqrt{(0,733 - 2)^2 + (1,4 - 1)^2 + (2,266 - 2)^2} = \sqrt{1,605 + 0,36 + 0,070} = 1,426$$

$$\text{Data-10 ke Pusat Cluster I} = \sqrt{(0,733 - 1)^2 + (1,4 - 1)^2 + (2,266 - 2)^2} = \sqrt{0,071 + 0,36 + 0,070} = 0,707$$

$$\text{Data-11 ke Pusat Cluster I} = \sqrt{(0,733 - 0)^2 + (1,4 - 1)^2 + (2,266 - 3)^2} = \sqrt{0,537 + 0,36 + 0,538} = 1,197$$

$$\text{Data-12 ke Pusat Cluster I} = \sqrt{(0,733 - 0)^2 + (1,4 - 3)^2 + (2,266 - 2)^2}$$

$$= \sqrt{0,017 + 2,56 + 0,070} = 1,626$$

$$\text{Data-13 ke Pusat Cluster I} = \sqrt{(0,733 - 1)^2 + (1,4 - 1)^2 + (2,266 - 1)^2}$$

$$= \sqrt{0,071 + 0,16 + 1,602} = 0,891 \text{ 1,353}$$

$$\text{Data-14 ke Pusat Cluster I} = \sqrt{(0,733 - 2)^2 + (1,4 - 0)^2 + (2,266 - 1)^2}$$

$$= \sqrt{1,605 + 1,96 + 1,602} = 2,273$$

$$\text{Data-15 ke Pusat Cluster I} = \sqrt{(0,733 - 0)^2 + (1,4 - 2)^2 + (2,266 - 1)^2}$$

$$= \sqrt{0,537 + 0,36 + 1,602} = 1,580$$

$$\text{Data-16 ke Pusat Cluster I} = \sqrt{(0,733 - 0)^2 + (1,4 - 1)^2 + (2,266 - 1)^2}$$

$$= \sqrt{0,537 + 0,16 + 1,602} = 1,522$$

$$\text{Data-17 ke Pusat Cluster I} = \sqrt{(0,733 - 2)^2 + (1,4 - 1)^2 + (2,266 - 1)^2}$$

$$= \sqrt{1,605 + 0,16 + 1,602} = 1,834$$

$$\text{Data-18 ke Pusat Cluster I} = \sqrt{(0,733 - 0)^2 + (1,4 - 0)^2 + (2,266 - 2)^2}$$

$$= \sqrt{0,53 + 1,96 + 0,070} = 1,6$$

$$\text{Data-19 ke Pusat Cluster I} = \sqrt{(0,733 - 0)^2 + (1,4 - 2)^2 + (2,266 - 2)^2}$$

$$= \sqrt{0,733 + 0,36 + 0,070} = 1,078$$

$$\text{Data-20 ke Pusat Cluster I} = \sqrt{(0,733 - 1)^2 + (1,4 - 0)^2 + (2,266 - 2)^2}$$

$$= \sqrt{0,071 + 1,96 + 0,070} = 1,449$$

$$\text{Data-1 ke Pusat Cluster II} = \sqrt{(1,666 - 2)^2 + (1,333 - 2)^2 + (1,5 - 2)^2}$$

$$= \sqrt{0,111 + 0,448 + 0,25} = 0,899$$

$$\text{Data-2 ke Pusat Cluster II} = \sqrt{(1,666 - 1)^2 + (1,333 - 0)^2 + (1,5 - 3)^2}$$

$$= \sqrt{0,435 + 1,768 + 2,25} = 2,110$$

$$\text{Data-3 ke Pusat Cluster II} = \sqrt{(1,666 - 1)^2 + (1,333 - 1)^2 + (1,5 - 1)^2}$$

$$= \sqrt{0,435 + 0,108 + 0,25} = 0,890$$

$$\text{Data-4 ke Pusat Cluster II} = \sqrt{(1,666 - 0)^2 + (1,333 - 2)^2 + (1,5 - 0)^2}$$

$$= \sqrt{2,755 + 0,448 + 0,25} = 1,858$$

$$\text{Data-5 ke Pusat Cluster II} = \sqrt{(1,666 - 0)^2 + (1,333 - 1)^2 + (1,5 - 2)^2}$$

$$= \sqrt{2,755 + 1,08 + 0,25} = 2,021$$

$$\text{Data-6 ke Pusat Cluster II} = \sqrt{(1,666 - 1)^2 + (1,333 - 1)^2 + (1,5 - 2)^2}$$

$$= \sqrt{0,435 + 0,110 + 0,25} = 0,638$$

$$\text{Data-7 ke Pusat Cluster II} = \sqrt{(1,666 - 0)^2 + (1,333 - 1)^2 + (1,5 - 2)^2}$$

$$= \sqrt{2,755 + 0,108 + 0,25} = 1,764$$

$$\text{Data-8 ke Pusat Cluster II} = \sqrt{(1,666 - 1)^2 + (1,333 - 1)^2 + (1,5 - 2)^2}$$

$$= \sqrt{0,443 + 0,108 + 0,25} = 0,934$$

$$\text{Data-9 ke Pusat Cluster II} = \sqrt{(1,666 - 2)^2 + (1,333 - 1)^2 + (1,5 - 2)^2}$$

$$= \sqrt{1,11 + 0,108 + 0,25} = 1,211$$

$$\text{Data-10 ke Pusat Cluster II} = \sqrt{(1,666 - 1)^2 + (1,333 - 1)^2 + (1,5 - 2)^2}$$

$$= \sqrt{0,435 + 0,110 + 0,25} = 0,891$$

$$\text{Data-11 ke Pusat Cluster II} = \sqrt{(1,666 - 0)^2 + (1,333 - 1)^2 + (1,5 - 3)^2}$$

$$= \sqrt{2,755 + 0,110 + 0,25} = 1,764$$

$$\text{Data-12 ke Pusat Cluster II} = \sqrt{(1,666 - 0)^2 + (1,333 - 3)^2 + (1,5 - 2)^2}$$

$$= \sqrt{2,755 + 2,778 + 0,25} = 2,404$$

$$\text{Data-13 ke Pusat Cluster II} = \sqrt{(1,666 - 1)^2 + (1,333 - 1)^2 + (1,5 - 1)^2}$$

$$= \sqrt{0,435 + 0,110 + 0,25} = 1,353$$

$$\text{Data-14 ke Pusat Cluster II} = \sqrt{(1,666 - 2)^2 + (1,333 - 0)^2 + (1,5 - 1)^2}$$

$$= \sqrt{2,755 + 1,768 + 0,25} = 1,459$$

$$\text{Data-15 ke Pusat Cluster II} = \sqrt{(1,666 - 0)^2 + (1,333 - 2)^2 + (1,5 - 1)^2}$$

$$= \sqrt{0,711 + 0,844 + 0,25} = 1,897$$

$$\text{Data-16 ke Pusat Cluster II} = \sqrt{(1,666 - 0)^2 + (1,333 - 1)^2 + (1,5 - 1)^2}$$

$$= \sqrt{0,711 + 0,710 + 0,25} = 1,686$$

$$\text{Data-17 ke Pusat Cluster II} = \sqrt{(1,666 - 2)^2 + (1,333 - 1)^2 + (1,5 - 1)^2}$$

$$= \sqrt{0,111 + 0,110 + 0,25} = 0,689$$

$$\begin{aligned} \text{Data-18 ke Pusat Cluster II} &= \sqrt{(1,666 - 0)^2 + (1,333 - 0)^2 + (1,5 - 2)^2} \\ &= \sqrt{2,775 + 1,776 + 0,25} = 2,193 \end{aligned}$$

$$\begin{aligned} \text{Data-19 ke Pusat Cluster II} &= \sqrt{(1,666 - 0)^2 + (1,333 - 2)^2 + (1,5 - 2)^2} \\ &= \sqrt{2,775 + 0,444 + 0,25} = 1,862 \end{aligned}$$

$$\begin{aligned} \text{Data-20 ke Pusat Cluster II} &= \sqrt{(1,666 - 1)^2 + (1,333 - 0)^2 + (1,5 - 2)^2} \\ &= \sqrt{0,443 + 1,776 + 0,25} = 1,571 \end{aligned}$$

Hasil perhitungan menggunakan Euclidian Distance pada iterasi I tersebut diatas dapat dilihat pada tabel 11. berikut:

Tabel 11. Hasil Iterasi III

| Daka ke- | Struk Ke- | Jarak Ke Centroid | |
|----------|-----------|-------------------|-------|
| | | C1 | C2 |
| 1 | 00021 | 1,426 | 0,899 |
| 2 | 00022 | 1,602 | 2,110 |
| 3 | 00023 | 1,620 | 0,890 |
| 4 | 00024 | 1,858 | 1,858 |
| 5 | 00025 | 1,354 | 2,021 |
| 6 | 00026 | 0,548 | 0,638 |
| 7 | 00027 | 0,983 | 1,764 |
| 8 | 00028 | 0,707 | 0,934 |
| 9 | 00029 | 1,426 | 1,211 |
| 10 | 00030 | 0,707 | 0,891 |
| 11 | 00041 | 1,197 | 1,764 |
| 12 | 00042 | 1,626 | 2,404 |
| 13 | 00043 | 0,891 | 1,353 |
| 14 | 00044 | 2,273 | 1,459 |
| 15 | 00045 | 1,580 | 1,897 |
| 16 | 00046 | 1,522 | 1,686 |
| 17 | 00047 | 1,834 | 0,689 |
| 18 | 00048 | 1,6 | 2,193 |
| 19 | 00049 | 1,078 | 1,862 |
| 20 | 00050 | 1,449 | 1,571 |

Adapun hasil penentuan jarak terpendek terhadap pusat cluster yang telah ditabulasikan ke dalam biner dapat dilihat pada tabel 12. berikut:

Tabel 12. Tabulasi Binner Iterasi II

| Daka ke- | Struk Ke- | Jarak Ke Centroid | | Cluster |
|----------|-----------|-------------------|----|---------|
| | | C1 | C2 | |
| 1 | 00021 | 0 | 1 | C2 |
| 2 | 00022 | 1 | 0 | C1 |
| 3 | 00023 | 0 | 1 | C2 |
| 4 | 00024 | 1 | 1 | C1C2 |
| 5 | 00025 | 1 | 0 | C1 |
| 6 | 00026 | 1 | 0 | C1 |
| 7 | 00027 | 1 | 0 | C1 |
| 8 | 00028 | 1 | 0 | C1 |
| 9 | 00029 | 0 | 1 | C2 |
| 10 | 00030 | 1 | 0 | C1 |
| 11 | 00041 | 1 | 0 | C1 |
| 12 | 00042 | 1 | 0 | C1 |
| 13 | 00043 | 1 | 0 | C1 |
| 14 | 00044 | 0 | 1 | C2 |
| 15 | 00045 | 1 | 0 | C1 |
| 16 | 00046 | 1 | 0 | C1 |
| 17 | 00047 | 0 | 1 | C2 |
| 18 | 00048 | 1 | 0 | C1 |
| 19 | 00049 | 1 | 0 | C1 |
| 20 | 00050 | 1 | 0 | C1 |

Tabel 13. Hasil Pengelompokan

| Kelompok (Cluster) | Anggota Cluster | Jumlah Anggota |
|--------------------|--|----------------|
| 1 | {2,4,5,6,7,8,10,11,12,13,15,16,18,19,20} | 15 |

| Kelompok (Cluster) | Anggota Cluster | Jumlah Anggota |
|-----------------------|-----------------|-------------------|
| 2 | {1,3,9,14,17} | 5 |

Berdasarkan hasil iterasi ke 3 memiliki hasil pengelompokan yang sama dengan iterasi ke 2 maka proses pencarian berhenti dan di dapatka untuk masing-masing kelopak berdasar dari pembagian data kelopak transaksi pembelian maka yang digunakan untuk penentuan pola pembelian pelanggan dalam menentukan pola tatak letak barang disuper market menggunakan anggota kelompok 1 seperti tabel dibawah ini:

Tabel 14. Hasil Pengelompokan Cluster 1

| No | Struk ke- | Merek Produk | Jumlah |
|----|-----------|---|--------|
| 1 | 00022 | Blueband, Beras sunco, Apel, panadol | 4 |
| 2 | 00024 | Marboro, Fanta, kapal api, The Melati | 4 |
| 3 | 00025 | Bolpoint, Clasmild, Tisu Paseo | 3 |
| 4 | 00026 | Palmia, semangka, Promag, Kopi Sidikalang | 4 |
| 5 | 00027 | Mentos, Antangin, Teh Melati, Kit Mobil | 4 |
| 6 | 00028 | M150, Ajinamoto, Matrai, Camel | 4 |
| 7 | 00030 | Beras Sunco, Sarden, Dancow, M150 | 4 |
| 8 | 00041 | Sajiku, Kertas A4, Kapal Api, Tisu paseo, Teh Wangi | 5 |
| 9 | 00042 | Coco cola, Bimoli, kratindeng, Royco, Kopi sidikalang | 5 |
| 10 | 00043 | Lux, Kapal api, Surya 16, Pisang | 4 |
| 11 | 00045 | Kratindeng, Kologen, Teh Melati, Marboro | 4 |
| 12 | 00046 | Somperna, Sprint, Bolpoint | 3 |
| 13 | 00048 | Marboro, Kecap Bango, Biore | 3 |
| 14 | 00049 | Panadol, Kopi Sidikalang, Kertas A4, Teh Sari | 4 |
| 15 | 00050 | Lifeboy, Sari Roti, Marboro, Anggur | 4 |

Tahap dalam klasifikasi pembelian barang dimulai dengan menyeleksi data-data yang akan dianalisa, kemudian dicari semua jenis item yang ada di dalam list transaksi penjualan, selanjutnya dicari jumlah item yang ada pada semua transaksi penjualan. Pembentukan pola kombinasi didasarkan pada nilai support minimal, jika nilai support minimal terpenuhi dan pola kombinasi itemsets yang ada lebih dari satu pola maka pola kombinasi itemsets yang berikutnya bisa dibentuk. Langkah yang selanjutnya adalah pembentukan aturan asosiasi yang diperoleh dari pola kombinasi itemsets. Adapun data yang digunakan adalah hasil normalisasi pengelompokan dengan menggunakan metode K-Means dimana data yang digunakan caluster 1 seperti tabel dibawah ini:

Tabel 15. Hasil Pengelompokan Cluster

| No | Struk Ke- | Merek Produk | Jumlah |
|----|-----------|-----------------|--------|
| 1 | 00022 | Blueband | 4 |
| | | Beras Sunco | |
| | | Apel | |
| | | Panadol | |
| 2 | 00024 | Marboro | 4 |
| | | Fanta, | |
| | | Kapal Api | |
| | | The Melati | |
| 3 | 00025 | Bolpoint, | 3 |
| | | Clasmild | |
| | | Tisu Paseo | |
| 4 | 00026 | Palmia, | 4 |
| | | Semangka, | |
| | | Promag | |
| | | Kopi Sidikalang | |
| 5 | 00027 | Mentos, | 4 |
| | | Antangin, | |
| | | Teh Melati, | |
| | | Kit Mobil | |
| 6 | 00028 | M150, | 4 |
| | | Ajinamoto, | |
| | | Matrai, | |
| | | Camel | |
| 7 | 00030 | Beras Sunco, | 4 |
| | | Sarden, | |
| | | Dancow, | |
| | | M150 | |
| 8 | 00041 | Sajiku, | 5 |

| No | Struk Ke- | Merek Produk | Jumlah |
|----|-----------|---|--------|
| 9 | 00042 | Kertas A4, Kapal Api, Tisu Paseo, Teh Wangi Coco Cola, Bimoli, Kratindeng, Royco, Kopi Sidikalang Lux, | 5 |
| 10 | 00043 | Kapal Api, Surya 16, Pisang Kratindeng, | 4 |
| 11 | 00045 | Kologen, Teh Melati, Marboro | 4 |
| 12 | 00046 | Somperna, Sprint, Bolpoint | 3 |
| 13 | 00048 | Marboro, Kecap Bango, Biore | 3 |
| 14 | 00049 | Panadol, Kopi Sidikalang, Kertas A4, Teh Sari | 4 |
| 15 | 00050 | Lifeboy, Sari Roti, Marboro, Anggur | 4 |

4. KESIMPULAN

Berdasarkan hasil penelitian mengenai normalisasi data penjualan supermarket menggunakan metode k-means untuk menentukan pola pembelian menerapkan metode apriori penulis dapat mengambil kesimpulan bahwa Dengan cara menormalisasikan data penjualan supermarket akan memudahkan memberikan informasi tata letak barang apa yang paling diamati oleh konsumen. Dengan demikian pola pembelian mempermudah konsumen dalam mencari produk makanan dan minuman. Karena telah diklasifikasikan data pada tempat yang di tentukan. Dengan menerapkan metode k-means dan metode apriori berhasil dibangun dan bertujuan dengan baik. Serta metode dapat di gunakan sebagai alat bantu untuk melakukan pengklasifikasian data dalam jumlah banyak. Sehingga tingkat kemudahan mencari produk yang diinginkan tepat dan akurat sesuai dengan data penjualan yang dijadikan sebagai cuan. Pada data mining menggunakan metode k-means untuk menentukan pola pembelian menerapkan metode apriori dalam normalisasi data penjualan disupermarket tentukannya dapat mengetahui kelemahan dan kelebihan prosedur sytem penjualan yang telah berjalan pada supermarket serta aplikasi yang dibangun akan menormalisasikan data secara berurutan. Sehingga dengan cara tersebut akan mempermudah mengecek produk yang diaminati yang laris di pasaran.

REFERENCES

- [1] I. A. K. Dalam, M. Produk, T. Dan, K. Laris, P. Toko, and A. Cikarang, "Pelita teknologi," vol. 16, no. 1, pp. 79–87, 2021.
- [2] S. M. Hutabarat and A. Sindar, "Data Mining Penjualan Suku Cadang Sepeda Motor Menggunakan Algoritma K-Means," *J. Nas. Komputasi dan Teknol. Inf.*, vol. 2, no. 2, p. 126, 2019.
- [3] B. M. Metisen and H. L. Sari, "Analisis clustering menggunakan metode K-Means dalam pengelompokan penjualan produk pada Swalayan Fadhila," *J. Media Infotama*, vol. 11, no. 2, pp. 110–118, 2015.
- [4] F. Ferlanda, "Implementasi Data Mining Untuk Menentukan Persediaan Stok Obat Di Apotek Enok Menggunakan Metode K-Means Clustering," *JATISI (Jurnal Tek. Inform. dan Sist. Informasi)*, vol. 8, no. 3, pp. 1294–1306, 2021.
- [5] I. H. T. U. Ida Fitriani, "Sistem Informasi Manajemen Stok Menggunakan Metode K-Means," *Univ. Teknol. Yogyakarta*, 2019.
- [6] F. Pratiwi, Fince Tinus Waruwu, D. P. Utomo, and R. Syahputra, "Penerapan Metode Aras Dalam Pemilihan Asisten Perkebunan Terbaik Pada PTPN V," in *Seminar Nasional Teknologi Komputer & Sains (SAINTEKS)*, 2019, pp. 651–662.
- [7] R. Kharisman Ndruru, "Penerapan Metode Additive Ratio Assessment (ARAS) dan Rank Order Centroid (ROC) Dalam Pemilihan Jaksa Terbaik Pada Kejaksaan Negeri Medan," *Semin. Nas. Teknol. Komput. Sains*, pp. 367–372, 2020.
- [8] X. Wu, X. Zhu, G. Q. Wu, and W. Ding, "Data mining with big data," *IEEE Trans. Knowl. Data Eng.*, vol. 26, no. 1, 2014.
- [9] E. Buulolo, *Data Mining Untuk Perguruan Tinggi*. Deepublish, 2020.
- [10] Y. D. Darmi and A. Setiawan, "PENERAPAN METODE CLUSTERING K-MEANS DALAM PENGELOMPOKAN PENJUALAN PRODUK," *J. MEDIA INFOTAMA*, vol. 12, no. 2, 2017.

- [11] F. Tambunan, "Implementation of Data Mining using the Clustering Method (Case: Region of the Actors of Theft Crime by Province)," *IJISTECH (International J. Inf. Syst. Technol.*, vol. 2, no. 2, p. 75, 2019.
- [12] I. Parlina, A. P. Windarto, A. Wanto, and M. R. Lubis, "MEMANFAATKAN ALGORITMA K-MEANS DALAM MENENTUKAN PEGAWAI YANG LAYAK MENGIKUTI ASESSMENT CENTER UNTUK CLUSTERING PROGRAM SDP," *CESS (Journal Comput. Eng. Syst. Sci.*, vol. 3, no. 1, pp. 87–93, 2018.