



Analisis Metode K-Means Clustering Dalam Pengelompokan Penjualan Sembako

Rahma Fikria*, Sriani

Sains dan Teknologi, Ilmu Komputer, Universitas Islam Negeri Sumatera Utara, Deli Serdang
Jl. Lap. Golf No.120, Kp. Tengah, Kec. Pancur Batu, Kabupaten Deli Serdang, Sumatera Utara, Indonesia

Email: ^{1,*}rahmafikria2002@gmail.com, ²sriani@uinsu.ac.id

Email Penulis Korespondensi: rahmafikria2002@gmail.com

Submitted: 27/07/2024; Accepted: 31/07/2024; Published: 31/07/2024

Abstrak—Penelitian ini bertujuan untuk menganalisis pengelompokan penjualan sembako di PT. Sinarmas Distribusi Nusantara menggunakan metode K-Means Clustering. Permasalahan utama yang dihadapi perusahaan adalah ketidakjelasan dalam mengidentifikasi produk dengan performa penjualan tinggi dan produk yang memerlukan perhatian lebih lanjut. Hal ini berdampak negatif pada efektivitas strategi pemasaran dan distribusi perusahaan. Metode K-Means Clustering digunakan untuk membagi data penjualan produk sembako ke dalam beberapa cluster berdasarkan kesamaan karakteristik penjualannya. Data penjualan dikumpulkan dan dianalisis untuk mengelompokkan produk berdasarkan tingkat penjualan. Hasil penelitian menunjukkan bahwa dari keseluruhan produk yang diteliti, 5 produk termasuk dalam kategori “Laris”, 2 produk dalam kategori “Kurang Laris” dan 51 produk dalam kategori “Tidak Laris”. Evaluasi hasil pengelompokan menggunakan indeks Davies-Bouldin menghasilkan nilai 0.8911, yang menunjukkan kualitas pengelompokan yang cukup baik. Kesimpulannya, metode K-Means Clustering efektif dalam mengidentifikasi pola penjualan produk sembako, sehingga dapat digunakan sebagai dasar untuk pengambilan keputusan strategis dalam manajemen penjualan.

Kata Kunci: K-Means Clustering; Pengelompokan; Penjualan Sembako; Strategi Pemasaran; Distribusi; Indeks Davies-Bouldin

Abstract—This research aims to analyze the clustering of staple food sales at PT. Sinarmas Distribusi Nusantara using the K-Means Clustering method. The main problem faced by the company is the lack of clarity in identifying which products have high sales performance and which products require more attention. This issue negatively impacts the effectiveness of the company's marketing and distribution strategies. The K-Means Clustering method is used to divide the sales data of staple food products into several clusters based on the similarity of their characteristics. Sales data is collected and analyzed to group products based on their sales levels. The research results show that out of all the products studied, 5 products fall into the "Fast-Selling" category, 2 products into the "Slow-Selling" category, and 51 products into the "Non-Selling" category. Evaluation of the clustering results using the Davies-Bouldin index yielded a value of 0.8911, indicating a reasonably good clustering quality. In conclusion, the K-Means Clustering method is effective in identifying sales patterns of staple food products, thus providing a basis for strategic decision-making in sales management.

Keywords: K-Means Clustering; Clustering; Staple Food Sales; Marketing Strategy; Distribution; Davies-Bouldin Index

1. PENDAHULUAN

PT. Sinarmas Distribusi Nusantara merupakan bagian dari Grup Sinarmas yang bergerak di bidang jasa penjualan dan distribusi. Perusahaan adalah sebuah perusahaan di bawah Grup Sinarmas yang mendistribusikan produk melalui berbagai saluran yang mencakup perdagangan umum, perdagangan modern, e-commerce, dan industri jasa makanan [1]. Dimana salah satu produknya adalah sembako. Persediaan stok produk merupakan salah satu hal penting untuk diperhatikan pada sebuah perusahaan distribusi, tidak terkecuali PT. Sinarmas Distribusi Nusantara. Salah satu permasalahan utama di PT. Sinarmas Distribusi Nusantara Medan adalah terjadinya penumpukan produk sembako di gudang dan terkadang mengalami kekurangan stok produk yang dibutuhkan pelanggan. Penumpukan produk di gudang PT. Sinarmas Distribusi Nusantara Medan dapat menyebabkan biaya penyimpanan yang tinggi, risiko kerusakan produk dan kesulitan dalam menemukan produk saat dibutuhkan [2]. Kualitas produk pun menurun akibat penyimpanan yang tidak optimal. Permasalahan ini berdampak pada kerugian finansial dan efisiensi operasional yang menurun. Akar permasalahan terletak pada perkiraan permintaan produk yang tidak akurat [3]. Untuk mengoptimalkan persediaan agar tidak terjadi masalah pada persediaan produk, dibutuhkan pengetahuan terhadap pola pembelian dari pelanggan. Dengan menganalisis pola pembelian pelanggan, efisiensi dari persediaan produk akan meningkat sehingga dapat meminimalisir terjadinya masalah kekurangan atau kelebihan persediaan produk yang dapat mengganggu proses distribusi [4].

Sembako merupakan kebutuhan pokok masyarakat yang harus dipenuhi setiap hari. Penjualan sembako menjadi salah satu sektor usaha yang sangat penting dan strategis [5]. Dalam menjalankan usaha penjualan sembako, perusahaan perlu memahami pola penjualan sembako agar dapat mengelola pendistribusiannya dengan baik. Salah satu cara untuk memahami pola penjualan sembako adalah dengan melakukan pengelompokan penjualan sembako dimana salah satunya adalah dengan memanfaatkan teknik data mining dengan algoritma K-means clustering. Teknik data mining ini dapat membantu perusahaan untuk mengidentifikasi produk yang paling laris dan yang tidak begitu laris [6]. Hal ini dapat menjadi acuan bagi perusahaan untuk melakukan pengambilan keputusan re-stok produk sehingga produk yang paling laris selalu tersedia di gudang dan produk yang tidak begitu

laris tidak menumpuk. Penelitian ini akan menghasilkan informasi yang dapat membantu perusahaan untuk meningkatkan efisiensi dan profitabilitas usaha melalui optimasi stok produk.

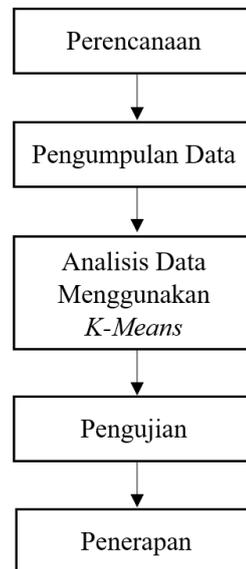
Data Mining adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan dan machine learning untuk mengekstraksi dan mengidentifikasi informasi yang terkait dari berbagai database besar [7]. Untuk memenuhi mengoptimalkan persediaan barang, banyak cara yang dapat ditempuh. Salah satunya dengan pemanfaatan data perusahaan. Data mining merupakan teknik yang menggabungkan teknik analisis data dan menemukan pola-pola yang penting pada data. Data mining merupakan proses informasi dengan mencari pola dan hubungan yang tersembunyi pada tumpukan data [8]. Data mining disebut juga sebagai Knowledge In Database (KDD) yaitu kegiatan Pengumpulan pemakaian data lampau untuk menentukan pola pembelian konsumen, salah satunya K-Means Clustering [9].

K-Means merupakan salah satu algoritma dalam fungsi clustering atau pengelompokan, dengan algoritma K-Means mengacu pada pengelompokan data, observasi atau kasus yang berdasarkan kemiripan suatu objek yang diteliti [10]. Algoritma K-Means cukup mudah dalam implementasi, relatif cepat, mudah disesuaikan dan banyak digunakan, dalam penggunaan algoritma K-Means untuk pengelompokan dimulai dengan pembentukan partisi/pusat cluster ini diperbaiki sehingga tidak terjadi perubahan yang signifikan pada partisi/pusat dari cluster [11]. Algoritma K-Means menggunakan pendekatan “top-down” yang semula dari awal sudah ditetapkan jumlah cluster kemudian dilakukan pengamatan. Dalam mengetahui jumlah dari kelompok cluster terbaik digunakan sebuah metode validasi cluster untuk memaksimalkan antara jarak inter cluster serta meminimalkan jarak intra cluster [12]. Sebelumnya pada penelitian dijelaskan bahwa metode K-Means dapat menentukan penjualan barang dengan penerapan data mining untuk menentukan kelompok barang yang lebih diminati berdasarkan penjualan. Dari data yang dianalisa dengan Algoritma K-Means Clustering akan menghasil satu tujuan yaitu mendapatkan kelompok produk. Dengan mengelompokkan produk berdasarkan tingkat kelarisan, bisnis dapat memastikan bahwa mereka memiliki persediaan yang cukup untuk barang yang paling laris dan tidak kelebihan persediaan untuk barang yang kurang laris. Hal ini dapat membantu meningkatkan efisiensi dan mengurangi biaya penyimpanan [13].

2. METODOLOGI PENELITIAN

2.1 Kerangka Penelitian

Pada penelitian ini dilakukan analisis data menggunakan teknik data mining algoritma K-Means untuk mengelompokkan beberapa produk ke dalam 3 kategori berdasarkan tingkat kelarisan. Adapun tahapan penelitian yang dilakukan adalah sebagai berikut:



Gambar 1. Kerangka Penelitian

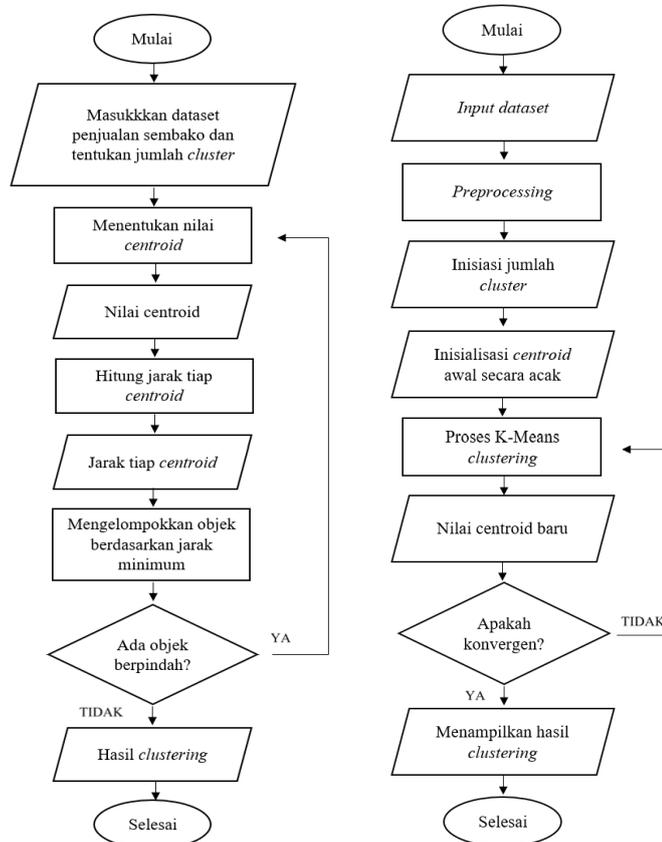
Gambar 1 merupakan kerangka penelitian. Pada tahapan penelitian yang dilakukan adalah perencanaan, pengumpulan data, analisis data menggunakan K-Means, pengujian dan penerapan. Berdasarkan sumber data yang digunakan pada penelitian ini, maka metode pengumpulan data yang digunakan adalah sebagai berikut.

1. Studi Literatur

Teknik ini dilakukan untuk mengungkap berbagai teori – teori yang relevan dengan permasalahan yang sedang diteliti sebagai bahan rujukan dalam pembahasan hasil penelitian [14]. Adapun kegiatan studi literatur yang dilakukan peneliti dengan mengumpulkan data pustaka, membaca dan memcatat, serta mengolah bahan penelitian.

2. Wawancara

Wawancara dilakukan untuk melengkapi data dan informasi terkait dengan penerapan metode K-Means Clustering dalam pengelompokan penjualan sembako di PT. Sinarmas Distribusi Nusantara [15]. Tujuan wawancara ini adalah meminta izin untuk menggunakan data penjualan sembako di PT. Sinarmas Distribusi Nusantara dalam penelitian yang akan dilakukan. Dari hasil wawancara dengan Bapak Alexander Ginting S.E. sebagai Branch Manager cabang Medan, diperoleh data penjualan produk sembako selama tahun 2023.



Gambar 2. (a) Flowchart Algoritma K-Means Clustering, (b) Flowchart Sistem

3. HASIL DAN PEMBAHASAN

3.1 Pengumpulan Data

Penelitian yang dilakukan melibatkan 58 jenis produk yang akan dikelompokkan ke dalam 3 kategori. Dataset memiliki 4 atribut yang akan memengaruhi proses pengelompokan yaitu penjualan semester pertama, penjualan semester kedua, rata-rata penjualan per semester dan margin stock. Berikut adalah 5 sampel jenis produk pada data yang akan digunakan.

Tabel 1. Representase Data

No	Produk	Semester I	Semester II	Rata-rata	Margin Stok
1	Cooking Oil CP	1426	0	713	1070
2	Filma 1L Pch	0	22	11	17
3	Filma 2L Pch	0	6323	3161.5	3162
4	Filma 18L Bib	201	0	100.5	101
5	Filma 1L Btl	625	394	509.5	127

Tabel 1. merupakan representase data 5 sampel jenis produk pada data yang akan digunakan.

3.2 Proses Clustering

Untuk proses clustering data dilakukan menggunakan Google Colaboratory (Colab). File dataset yang berformat .xlsx di upload ke lingkungan colab untuk selanjutnya dilakukan pre-processing data.

3.2.1 Pre-Processing Data

Terdapat beberapa tahapan yang perlu dilakukan untuk mempersiapkan data penjualan sembako sebelum diproses menggunakan metode K-Means Clustering. Adapun tahapan yang dimaksud adalah sebagai berikut.

- a. Pembersihan data
Tahap pembersihan data bertujuan untuk memastikan data yang digunakan bebas dari kesalahan dan siap untuk dianalisis. Langkah-langkah yang dilakukan meliputi pengecekan data duplikat dan penghapusan data yang tidak memiliki nilai atribut [16].
- b. Transformasi data
Transformasi data diperlukan untuk mengubah data ke dalam format yang sesuai untuk analisis K-Means Clustering untuk mengoptimalkan proses pengelompokan. Teknik transformasi yang digunakan pada penelitian ini adalah z-score [17].

Tabel 2. Transformasi Data

No	Produk	Semester I	Semester II	Rata-rata	Margin Stok
1	Cooking Oil CP	1.812	-0.541	-0.160	0.146
2	Filma 1L Pch	-0.836	-0.532	-0.765	-0.734
3	Filma 2L Pch	-0.836	1.996	1.950	1.894
4	Filma 18L Bib	-0.463	-0.541	-0.688	-0.664
5	Filma 1L Btl	0.324	-0.383	-0.336	-0.642

Tabel 2. merupakan hasil transformasi data pada sampel data pada tabel sebelumnya.

3.2.2 Pengelompokan K-Means Clustering

Data yang sudah melalui tahap pre-processing akan diteruskan ke tahap pengelompokan. Adapun tahapan dari proses pengelompokan adalah sebagai berikut.

- a. Tentukan jumlah cluster (c)
Proses penentuan jumlah cluster umumnya dilakukan dengan menggunakan berbagai metode seperti elbow atau silhouette. Namun pada penelitian ini proses penentuan jumlah cluster tidak menggunakan metode apapun dikarenakan tujuan awal dari penelitian yang dilakukan adalah mengelompokkan produk ke dalam kategori “Laris”, “Kurang Laris” dan “Tidak Laris”. Sehingga jumlah cluster yang diterapkan pada penelitian ini adalah sebanyak 3 cluster [18].
- b. Inisialisasi nilai centroid
Pada tahap awal, kita memilih k titik sebagai centroid awal. Inisialisasi nilai centroid awal dilakukan secara acak. Misalkan dataset awal adalah data pada tabel 2. Kemudian data akan dibagi ke dalam 3 cluster (k=3). Maka diinisialisasi centroid awal secara acak sebagai:
Centroid 1 : (0.2255, 0, 0.1128, 0.1692)
Centroid 2 : (0, 0.0035, 0.0017, 0.0027)
Centroid 3 : (0, 1, 0.5, 0.5001)
- c. Menghitung jarak data
Pada tahap ini, kita menghitung jarak Euclidean antara setiap titik data dengan setiap centroid. Rumus jarak Euclidean adalah:
$$\text{Jarak} = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \tag{1}$$

Contoh perhitungan data sampel :
Untuk data pertama pada tabel 2, jarak ke setiap centroid dihitung sebagai berikut:
Data 1 ke centroid 1 = $\sqrt{(1.812 - 0.2255)^2 + (-0.541 - 0)^2 + (-0.160 - 0.1128)^2 + (0.146 - 0.1692)^2} = \mathbf{1.69}$
Data 1 ke centroid 2 = $\sqrt{(1.812 - 0)^2 + (-0.541 - 0.0035)^2 + (-0.160 - 0.0017)^2 + (0.146 - 0.0027)^2} = \mathbf{1.9}$
Data 1 ke centroid 3 = $\sqrt{(1.812 - 0)^2 + (-0.541 - 1)^2 + (-0.160 - 0.5)^2 + (0.146 - 0.5001)^2} = \mathbf{2.49}$
- d. Penugasan cluster
Tabel 3. merupakan penugasan cluster. Setiap titik data ditugaskan ke cluster yang memiliki centroid dengan jarak terdekat. Misal dari perhitungan jarak di atas, data pertama ditugaskan ke cluster 1 karena jaraknya ke centroid 1 adalah yang terpendek yaitu 0.

Tabel 3. Penugasan Cluster

No	Produk	Cluster
1	Cooking Oil CP	1
2	Filma 1L Pch	2
3	Filma 2L Pch	1
4	Filma 18L Bib	2
5	Filma 1L Btl	2



e. Perhitungan nilai centroid baru

Setelah penugasan cluster, centroid baru dihitung sebagai rata-rata dari semua titik data dalam cluster tersebut. Dari proses perhitungan jarak euclidean pertama, kita memperoleh hasil bahwa terdapat 2 data pada cluster 1 dan 3 data pada cluster 2. Dari hasil tersebut, maka dicari nilai centroid baru dari rata-rata masing-masing data pada cluster. Misalkan untuk centroid 1, nilai dari data 1 dan data 3 dijumlahkan kemudian dibagi 2 untuk menjadi nilai centroid baru.

Centroid 1 = ((1.812 + -0.836)/2), (-0.541 + 1.996)/2, (-0.160 + 1.950)/2, (0.146 + 1.894)/2) = (0.488, 0.7275, 0.895, 1.02)

Centroid 2 = ((-0.836 + -0.463 + 0.324)/3, (-0.532 + -0.541 + -0.383)/3, (-0.765 + -0.688 + -0.336)/3, (-0.734 + -0.664 + -0.642)/3) = (-0.325, -0.485, -0.596, -0.68)

Centroid 3 = (0,1, 0,5, 0,5001)

f. Iterasi

Setelah diperoleh nilai centroid baru, langkah perhitungan jarak data dilakukan kembali, penugasa cluster dilakukan dan kemudian nilai centroid baru dicari kembali. Ketiga proses tersebut dilakukan secara terus menerus sampai nilai centroid tidak mengalami perubahan lagi (konvergen) atau perubahan nilai centroid sangat kecil. Apabila nilai centroid sudah mencapai konvergen, artinya titik data sudah dikelompokkan dengan stabil.

Tabel 4. Hasil Perhitungan Pada Iterasi 1

Table with 6 columns: No., Nama Produk, C1, C2, C3, Cluster. It lists various products like Cooking Oil Cp, Filma CO SNI, and Palmvita SHT DFF with their corresponding cluster assignments and distance values.

Tabel 4. merupakan hasil perhitungan jarak euclidean seluruh data terhadap centroid pada iterasi pertama. Setelah melakukan perhitungan jarak euclidean iterasi 1, maka akan diperoleh nilai centroid baru. Dengan menggunakan cara yang sama pada iterasi pertama, maka dilakukan perhitungan jarak euclidean iterasi kedua menggunakan centroid terbaru [19].

Tabel 5. Hasil Perhitungan Pada Iterasi 2

Table with 6 columns: No., Nama Produk, C1, C2, C3, Cluster. It lists the same products as Table 4 but with updated distance values for the second iteration.



No.	Nama Produk	C1	C2	C3	Cluster
11	Filma MGRN TF Salted (0420) 60X200G SCH	4.491926	2.364898	1.671861	2
12	Filma MGRN TF Salted (0520) 24X250G Tub	0.075118	2.753222	5.964514	0
13	Filma MGRN TF Salted (0819) 60x200G Sch	0.107469	2.793164	6.003148	0
14	Filma MGRN TF Salted 60x200G Sch	0.110579	2.795594	6.005978	0
15	Filma MGRNKrim TF Salted(0622)24X250GTub	0.103246	2.79083	5.999011	0
...	...				
58	Palmvita SHT DFF (0620) 1x15Kg Ctn	0.098279	2.781083	5.993165	0

Tabel 5. merupakan hasil perhitungan jarak euclidean terhadap centroid terbaru pada iterasi kedua dapat dilihat pada tabel diatas.

Dapat dilihat bahwa pada iterasi kedua, hasil jarak data terhadap centroid dari seluruh data tidak mengalami perubahan yang signifikan. Setelah melanjutkan ke iterasi yang ketiga pada sistem ternyata nilai cluster juga tidak mengalami perubahan. Oleh karena itu dapat disimpulkan bahwa iterasi pada penelitian ini berjumlah 2 iterasi. Dengan menggunakan nilai centroid pada iterasi kedua, maka diperoleh hasil pengelompokan pada data sebenarnya sebagaimana pada tabel 5 [20].

Pada tabel 5 telah diketahui masing-masing cluster dari keseluruhan data. Dimana setiap data termasuk ke dalam cluster 1, 2 atau 3. Untuk mencapai tujuan pada penelitian ini maka dibutuhkan pendefinisian dari setiap cluster. Dimana dengan pencocokan jumlah penjualan terhadap hasil pengelompokan maka diperoleh kesimpulan bahwa cluster 0 didefinisikan sebagai kategori “Tidak Laris”, cluster 1 sebagai “Kurang Laris” dan cluster 2 sebagai “Laris”. Sehingga dapat disimpulkan bahwa dari keseluruhan data produk yang diteliti menggunakan algoritma K-Means Clustering terdapat 5 produk yang termasuk pada kategori “Laris”, 2 produk pada kategori “Kurang Laris” dan 51 produk termasuk pada kategori “Tidak Laris”.

Tabel 6. Hasil Akhir Clustering

No	Nama Produk	Cluster
1	Cooking Oil Cp	Tidak Laris
2	Filma CO SNI (0223) 12x1L Pch	Tidak Laris
3	Filma CO SNI (0223) 6x2L Pch	Tidak Laris
4	Filma Co SNI (0920) 1x18L BIB	Tidak Laris
5	Filma CO SNI 12x1L Btl	Tidak Laris
6	FILMA CO SNI 12X1L PCH	Tidak Laris
7	Filma CO SNI 4x5L Jrc	Tidak Laris
8	Filma CO SNI 6x2L Btl	Tidak Laris
9	Filma CO SNI 6x2L Pch	Laris
10	Filma CO SNI NP 6x2L Pch	Tidak Laris
11	FILMA MGRN TF SALTED (0420) 60X200G SCH	Laris
12	Filma MGRN TF Salted (0520) 24X250G Tub	Tidak Laris
13	Filma MGRN TF Salted (0819) 60x200G Sch	Tidak Laris
14	Filma MGRN TF Salted 60x200G Sch	Tidak Laris
15	Filma MGRNKrim TF Salted(0622)24X250GTub	Tidak Laris
16	Filma Professional CO SNI(0420)1x18L BIB	Tidak Laris
17	Filma Professional CO Vit(0420)1x18L Jrc	Tidak Laris
18	Kunci Mas CO 24x450mL PCH(NOT USED)	Tidak Laris
19	Kunci Mas CO SNI (0323) 4x5L Jrc	Tidak Laris
20	Kunci Mas CO SNI (0323) 6x2L Pch	Tidak Laris
21	Kunci Mas CO SNI (0520) 12x900MI Pch	Kurang Laris
22	Kunci Mas CO SNI (0520) 24x450MI Pch	Tidak Laris
23	Kunci Mas CO SNI (0520) 48x200MI Pch	Tidak Laris
24	Kunci Mas CO SNI (0520) 4x5L Jrc	Tidak Laris
25	Kunci Mas CO SNI (0920) 1x18L Jrc	Tidak Laris
26	Kunci Mas CO SNI 12x900ML PCH	Tidak Laris
27	Kunci Mas CO SNI NP 12x900MI Pch	Tidak Laris
28	Kunci Mas CO SNI NP 6x2L Pch	Tidak Laris
29	Kunci Mas Cooking Oil SNI (0520)6x2L Pch	Kurang Laris



No	Nama Produk	Cluster
30	Kunci Mas Cooking Oil SNI 6x2L Pch	Tidak Laris
31	Masku CO SNI (0423) 12x1L Pch	Tidak Laris
32	Masku CO SNI (0423) 6x2L Pch	Tidak Laris
33	Masku CO SNI (0520)12x1L PCH	Tidak Laris
34	Masku Cooking Oil SNI (0520) 4x5L Jrc	Tidak Laris
35	Masku Cooking Oil SNI (0520) 6x2L PCH	Tidak Laris
36	MASKU COOKING OIL SNI 12X1L PCH	Tidak Laris
37	Menara MGRN (0720) 1x15Kg Ctn	Tidak Laris
38	Menara MGRN Krim (0720) 1x15Kg Ctn	Tidak Laris
39	Menara MGRN Krim (0922) 15Kg Ctn	Tidak Laris
40	Menara SHT BKF Classic (0620)1x15Kg Ctn	Tidak Laris
41	Menara SHT BKF Classic (0822) 15Kg Ctn	Tidak Laris
42	Menara SHT BKF Spc MRD (0420) 1x15Kg Ctn	Tidak Laris
43	Menara Spc SHT BKF (0322) 15Kg Ctn	Tidak Laris
44	Minyakita CO SNI (1022) 12x1L Pch	Laris
45	Mitra CO SNI (0323) 12x1L Pch	Tidak Laris
46	Mitra CO SNI (0323) 6x2L Pch	Laris
47	Mitra CO SNI (1121) 12x1L Pch	Tidak Laris
48	Mitra CO SNI (1121) 6x2L Pch	Laris
49	Mitra CO SNI NP 12x1L Pch	Tidak Laris
50	Mitra CO SNI NP 6x2L Pch	Tidak Laris
51	Mitra Cooking Oil SNI (0520) 12x1L Pch	Tidak Laris
52	Palmboom Margarine TF 24X250G Tub	Tidak Laris
53	Palmboom Margarine TF 24X500G Tub	Tidak Laris
54	Palmboom MGRN TF (0122) 24X500G Tub	Tidak Laris
55	Palmboom MGRN TF (1220) 24X250G Tub	Tidak Laris
56	Palmboom MGRN TF (1220) 24X500G Tub	Tidak Laris
57	Palmvita SHT BOS (0720) 1x18Kg Pal	Tidak Laris
58	Palmvita SHT DFF (0620) 1x15Kg Ctn	Tidak Laris

Tabel 6. merupakan hasil akhir clustering. Setelah proses pengelompokan selesai, file hasil pengelompokan yang telah dilakukan diunduh melalui colab. Data diatas adalah tampilan data hasil pengelompokan yang diunduh dari colab.

4. KESIMPULAN

Berdasarkan hasil penelitian, proses pengelompokan penjualan sembako di PT. Sinarmas Distribusi Nusantara dapat dilakukan menggunakan teknik data mining yang dengan tahapan proses pre-processing data, penentuan jumlah cluster, proses algoritma K-Means Clustering dan evaluasi hasil clustering. Dengan mengelompokkan data penjualan PT. Sinarmas Distribusi Nusantara, perusahaan dapat mengidentifikasi produk-produk yang paling diminati oleh setiap cluster konsumen. Dengan demikian, mereka dapat mengalokasikan stok secara proporsional untuk memenuhi permintaan dari masing-masing segmen pasar. Proses pengelompokan data memperoleh hasil bahwa dari keseluruhan produk yang diteliti, 5 produk termasuk pada kategori “Laris”, 2 produk pada kategori “Kurang Laris” dan 51 produk pada kategori “Tidak Laris”. Hasil dari proses pengelompokan memperoleh nilai davies-bouldin index sebesar 0.8911.

REFERENCES

- [1] K. A. Ningtyas, “Strategi Peningkatan Pembelian di PT. Haruman Multi Artha melalui Bauran Pemasaran.” Universitas Islam Indonesia, 2023.
- [2] K. Y. Saputra, “Analisis Komunikasi Organisasi Pada Pt. Duta Palma Nusantara Pks Sci. Kuko Kuantan Singingi.” Universitas Islam Riau, 2019.
- [3] A. Syamil et al., *Manajemen Rantai Pasok*. PT. Sonpedia Publishing Indonesia, 2023.
- [4] M. B. Soeltanong and C. Sasongko, “Perencanaan produksi dan pengendalian persediaan pada perusahaan manufaktur,” *J. Ris. Akunt. Perpajak.*, vol. 8, no. 01, pp. 14–27, 2021.
- [5] A. A. B. Sahbannur, “Analisis Pengendalian Internal atas Persediaan Barang Dagang pada CV. Aisyah,” vol. 1, p. 50, 2019.
- [6] D. S. Dewi, “Dampak Keberadaan Pasar Modern Terhadap Pendapatan Para Pedagang Pasar Tradisional (Studi Kasus di Desa Tanggul Angin Kecamatan Punggur).” IAIN Metro, 2020.
- [7] D. Ramdhan, G. Dwilestari, R. D. Dana, and A. Ajiz, “Clustering Data Persediaan Barang Dengan Menggunakan Metode K-Means,” *MEANS (Media Inf. Anal. dan Sist.*, pp. 1–9, 2022.
- [8] D. H. Depari, Y. Widiastiw, and M. M. Santoni, “Perbandingan Model Decision Tree, Naive Bayes dan Random Forest untuk Prediksi Klasifikasi Penyakit Jantung,” *Inform. J. Ilmu Komput.*, vol. 18, no. 3, p. 239, 2022, doi:



10.52958/iftk.v18i3.4694.

- [9] N. F. Adani, A. F. Boy, and R. Syahputra, “Implementasi Data Mining Untuk Pengelompokan Data Penjualan Berdasarkan Pola Pembelian Menggunakan Algoritma K-Means Clustering Pada Toko Syihan,” *J. Cyber Tech*, vol. 2, no. 5, 2019.
- [10] D. Dona, “Penerapan Metode K-Means Clustering Untuk Menentukan Status Gizi Baik Dan Gizi Buruk Pada Balita (Studi Kasus Kabupaten Rokan Hulu),” *Rabit J. Teknol. dan Sist. Inf. Univrab*, vol. 7, no. 2, pp. 179–191, 2022.
- [11] Y. Prayoga, A. Mahmudi, and H. Z. Zahro, “Penerapan Metode K Means Pada Sistem Informasi Akademik Untuk Pengelompokan Siswa Berprestasi Di Upt Sma Negeri 3 Kota Pasuruan Berbasis Web,” *Jati (Jurnal Mhs. Tek. Inform., vol. 5, no. 2, pp. 822–828, 2021.*
- [12] N. S. Aji, F. Natsir, and S. Istianah, “Penentuan Penjualan Barang Berdasarkan Pengelompokan Produk dengan K-Means Clustering Metode CRISP-DM Pada CV Sembako Dina,” *J. Zetroem*, vol. 05, no. 02, pp. 119–126, 2023.
- [13] A. Nugraha, O. Nurdiawan, and G. Dwilestari, “Penerapan Data Mining Metode K-Means Clustering Untuk Analisa Penjualan Pada Toko Yana Sport,” *JATI (Jurnal Mhs. Tek. Inform., vol. 6, no. 2, pp. 849–855, 2022.*
- [14] S. Sriani, M. S. Hasibuan, and R. Ananda, “Classification of Batu Bara Songket Using Gray-Level Co-Occurrence Matrix and Support Vector Machine,” *J. Ris. Inform., vol. 5, no. 1, pp. 481–490, 2022, doi: 10.34288/jri.v5i1.469.*
- [15] Suhardi, A. H. Lubis, A. Aprilia, and I. A. Ningrum, “Penerapan Metode Simple Multi Attribute Rating Technique pada Pemilihan Cafe Terfavorit,” *Sist. Pendukung Keputusan dengan Apl., vol. 2, no. 1, pp. 1–11, 2023, doi: 10.55537/spk.v2i1.114.*
- [16] P. W. Rahayu et al., *Buku Ajar Data Mining*. PT. Sonpedia Publishing Indonesia, 2024.
- [17] E. Akbar, “Perbandingan Algoritma Dbscan-K Means dan K Means Untuk Pengelompokan Madrasah Aliyah Provinsi Jawa Timur.” *Fakultas Sains dan Teknologi UIN Syarif Hidayatullah Jakarta, 2023.*
- [18] F. A. Dewa and M. T. Jatipaningrum, “SEGMENTASI E-COMMERCE DENGAN CLUSTER K-MEANS DAN FUZZY C-MEANS: Studi Kasus: Media Sosial di Indonesia yang diunduh di Play Store,” *J. Stat. Ind. dan Komputasi, vol. 4, no. 01, pp. 53–67, 2019.*
- [19] R. Mauliadi, “Data Mining Menggunakan Algoritma K-Means Clustering dalam Analisis Tingkat Potongan Harga Terhadap Harga Jual Sepeda Motor Honda,” *J. Inform. Ekon. Bisnis, pp. 124–129, 2022.*
- [20] M. R. Q. Baihaqie, “Analisis Perilaku Konsumen pada Usaha Ritel dengan menggunakan Metode Association Rule-Market Basket Analysis dan Clustering sebagai Usulan Strategi Peningkatan Penjualan (Studi Kasus: Intimart Gedongan).” *Universitas Islam Indonesia, 2023.*