

PENGELOMPOKAN DATA PENJUALAN ALAT PERABOTAN MENGUNAKAN K MEANS DI TOKO EKA PERABOT KOTA TANGERANG

Azis Suhendra

¹Universitas Pamulang, Jl. Raya Puspitek, Buaran, Kec. Pamulang, Kota Tangerang Selatan, Banten, Indonesia, 15310

e-mail: azissuhendra05@gmail.com

Abstract

This study uses the K-Means clustering algorithm to examine furniture sales patterns at Eka Perabot Store in Tangerang City. Understanding sales patterns can help the store develop better inventory management and marketing strategies. This study analyzed sales data to identify clustering patterns based on product price and quantity. Sales data collection, preprocessing, and the application of the K-Means algorithm using RapidMiner software were all part of the research process. The cleaned data was then grouped into clusters based on similar characteristics, resulting in product groups with specific characteristics, such as low-priced and high-selling products. The results showed that organizing with K-Means successfully divided products into categories appropriate to the store. The K-Means clustering method proved effective in helping Eka Perabot Store understand customer preferences and develop better sales strategies because each cluster has unique characteristics that can be used as a basis for business decision-making.

Keywords: K-Means Clustering, Sales, Sales Patterns, Machine Learning, RapidMiner, Furniture Store

Abstrak

Studi ini menggunakan algoritma clustering K-Means untuk mempelajari pola penjualan alat perabotan di Toko Eka Perabot Kota Tangerang. Toko dapat membuat strategi pengelolaan persediaan dan pemasaran yang lebih baik dengan mengetahui pola penjualan. Penelitian ini menganalisis data penjualan untuk menemukan pola pengelompokan berdasarkan harga dan jumlah produk. Pengumpulan data penjualan, preprocessing, dan penerapan algoritma K-Means melalui perangkat lunak RapidMiner adalah semua bagian dari proses penelitian. Data yang telah bersih kemudian dikelompokkan ke dalam beberapa cluster berdasarkan kesamaan karakteristik, yang menghasilkan kelompok produk dengan karakteristik tertentu, seperti produk dengan harga rendah dan penjualan tinggi. Hasil penelitian menunjukkan bahwa pengorganisasian dengan K-Means berhasil membagi produk ke dalam beberapa kategori yang sesuai dengan toko. Metode K-Means Clustering terbukti efektif dalam membantu Toko Eka Perabot memahami preferensi pelanggan dan membuat strategi penjualan yang lebih baik karena setiap cluster memiliki karakteristik unik yang dapat digunakan sebagai dasar pengambilan keputusan bisnis.

Kata Kunci: Clustering K-Means, Penjualan, Pola Penjualan, Machine Learning, RapidMiner, Toko Perabotan

1. PENDAHULUAN

Seiring berkembangnya teknologi informasi, pengelolaan data dalam dunia bisnis pun mengalami perubahan yang signifikan. Data tidak lagi sekadar catatan transaksi, tetapi data telah berkembang menjadi aset strategis yang dapat menambah nilai jika dianalisis dengan benar. Di industri retail, data penjualan merupakan sumber daya penting yang dapat diproses untuk memberikan wawasan tentang perilaku konsumen, tren pasar, dan kinerja produk. Namun tantangan bagi banyak perusahaan, termasuk Toko Eka Perabot Kota Tangerang, adalah menggunakan data yang semakin besar dan kompleks untuk mendukung keputusan bisnis yang lebih baik.

Toko Eka Perabot merupakan toko yang menjual berbagai macam furniture rumah mulai dari wajan, panci, baskom hingga peralatan dapur lainnya. Volume penjualan toko ini cukup tinggi, dan data penjualan yang dihasilkan setiap harinya terus meningkat. Namun, data penjualan tidak diproses secara sistematis untuk mengidentifikasi pola yang dapat membantu toko mengembangkan strategi yang lebih efektif, seperti manajemen inventaris, penetapan harga, dan strategi pemasaran.

Selama ini Toko Eka Perabot mengelola data penjualan hanya dengan mencatat setiap transaksi tanpa melakukan analisis data secara mendalam.

Akibatnya, toko kesulitan memahami produk mana yang paling laris, kapan waktu terbaik untuk menjalankan promosi, dan bagaimana mengelola stok dengan lebih efisien. Misalnya, produk dengan permintaan tinggi sering kali kehabisan stok, sedangkan produk dengan permintaan rendah tetap berada di rak untuk jangka waktu yang lama. Situasi ini menunjukkan bahwa Toko Eka Perabot membutuhkan suatu sistem yang dapat mengelompokkan data penjualannya dengan lebih baik sehingga informasi yang diberikan dapat digunakan untuk mengambil keputusan bisnis yang lebih akurat. [1]

2. TINJAUAN PUSTAKA

2.1. Pengelompokan

Secara umum, Pengelompokan adalah proses atau tindakan membagi objek, data, atau informasi ke dalam kelompok-kelompok berdasarkan fitur tertentu. Dalam arti yang lebih luas, pengelompokan dapat digunakan dalam berbagai domain, seperti data, bahasa, dan analisis statistik.

Pengertian pengelompokan menurut (Arifatul Ulya et al., 2023) adalah untuk membagi kumpulan data menjadi berbagai kelompok berdasarkan kesamaan variabel atau fitur. Di mana data dikelompokkan berdasarkan centroid, atau pusat cluster, yang paling dekat dengannya.

Pengelompokan adalah salah satu teknik yang paling signifikan dalam data mining yang mengeksplorasi kumpulan data. Dalam beberapa decade terakhir, beberapa pendekatan pengelompokan dengan kinerja yang lebih baik telah diterapkan untuk berbagai aplikasi, seperti bidang psikologi dan ilmu sosial lainnya, biologi, statistik, pengenalan pola, pencarian informasi, machine learning, dan data mining. (Zaki et al., 2022). [2]

2.2. Data

Data merupakan salah satu komponen yang penting di era digital, seluruh aktivitas data akan terekam dan disimpan dalam media penyimpanan data atau database, sehingga data ini dapat diproses untuk mendapatkan informasi yang berguna.

Data menurut Suharsimi Arikunto dalam jurnal Rike Limia Budiarti dan Gina Cendana, data adalah segala fakta dan angka yang dapat dijadikan bahan untuk menyusun suatu informasi. Sehingga dapat di artikan bahwa informasi adalah data yang sudah diolah untuk keperluan tertentu. (Limia Budiarti & Cendana, 2022).

2.3. Perabotan

Perabotan adalah berbagai barang atau benda yang digunakan untuk melengkapi dan memfungsikan suatu ruang, baik di rumah, kantor, atau area komersial. Perabotan biasanya berfungsi sebagai tempat duduk, tempat kerja, atau tempat istirahat, dan termasuk kursi, meja, lemari, tempat tidur, rak, dan perangkat lainnya yang membantu membuat ruang menjadi nyaman dan fungsional.

Perabotan dapat dibuat dari berbagai bahan, seperti kayu, logam, plastik, atau kaca, dan sering dirancang untuk memberikan fungsionalitas dan menambah nilai estetika ruangan. Tujuan pemilihan dan penataan perabotan biasanya disesuaikan dengan tujuan dan gaya dekorasi ruang tersebut, sehingga perabotan tersebut tidak hanya memenuhi kebutuhan praktis tetapi juga mempercantik tampilan keseluruhan. [3]

2.4. Machine Learning

Machine learning adalah cabang kecerdasan buatan yang memungkinkan algoritme untuk menemukan pola tersembunyi dalam kumpulan data, yang memungkinkan mereka untuk membuat prediksi pada data baru yang serupa tanpa melakukan pemrograman khusus untuk setiap tugas. *Machine Learning* tradisional menggunakan data dan alat statistik untuk memprediksi keluaran, menghasilkan wawasan yang dapat ditindaklanjuti. Aplikasi teknologi ini mencakup pemrosesan bahasa alami, pengenalan ucapan dan gambar, sistem rekomendasi, deteksi penipuan, pengoptimalan portofolio, dan pengotomatisan tugas.

Menurut Faggella dalam jurnal (Bella Pratiwi, 2024) *Machine Learning* adalah ilmu yang memungkinkan komputer bertindak seperti manusia, dimana mesin dapat secara otomatis belajar memahami lebih banyak seiring berjalannya waktu atau melalui pengalaman.

Adapun metode *machine learning* yang perlu diketahui, antara lain :

a. *Supervised Learning*

Supervised learning merupakan salah satu metode *machine learning* dimana pada metode ini data yang digunakan adalah data yang memiliki kelas atau label di dalamnya. Data terbagi menjadi dua yaitu data latih dan data uji. Data latih *Reinforcement Learning* digunakan untuk membentuk model, lalu model yang telah di bentuk itu di terapkan pada data uji, salah satu tujuannya adalah untuk mengetahui hasil atau *output* dan mengetahui tangka akurasi. Beberapa algoritma *supervised learning* yang dapat digunakan antara lain seperti *Decision Tree*, *Naive Bayes Classifier*, *K-NN*, *Artificial Neural Network*, dan *Support Vector Machine*.

b. *Unsupervised Learning*

Metode *unsupervised learning* tidak perlu menggunakan data untuk "berlatih", akan tetapi model akan dibentuk dari cara mengidentifikasi pola, mengenali karakteristik data, dll. Biasanya, metode ini digunakan pada prosedur yang lebih rumit.

c. *Reinforcement Learning*

Reinforcement learning merupakan metode *machine learning* yang berfungsi untuk mengoptimalkan sistem. Terdapat dua tipe *reinforcement learning* yaitu *positive reinforcement learning* dan *negative reinforcement learning*. Beberapa manfaat yang didapatkan jika menggunakan metode ini adalah menyelesaikan permasalahan perusahaan dan menghemat waktu dan tenaga.[4]

2.5. Data mining

Data mining menurut David Hand adalah analisis terhadap data (biasanya data yang berjumlah/berukuran besar) untuk menemukan hubungan yang jelas serta menyimpulkannya yang belum diketahui sebelumnya dengan cara terbaru yang dipahami dan berguna bagi pemilik data tersebut.

Data mining adalah proses untuk menemukan pola atau informasi yang berharga dan terpendam dari sejumlah besar data dengan menerapkan teknik ataupun metode tertentu. salah satu teknik data mining adalah teknik clustering. (Murdiaty et al., 2020)

Data mining adalah cara untuk menemukan informasi dalam suatu basis data. Ini juga dapat disebut sebagai proses mengekstraksi data dari kumpulan data atau informasi yang sangat besar. Pada dasarnya, data mining adalah bagian dari Knowledge Discovery in Database (KDD) (Yoliadi, 2023)

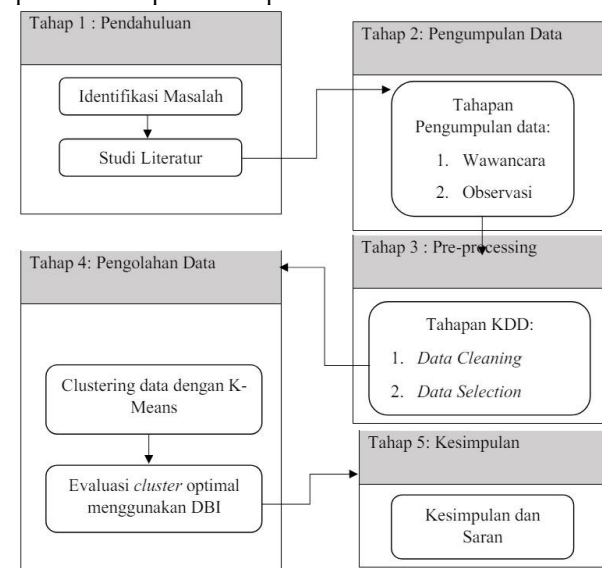
Berdasarkan pengertian Data Mining menurut para ahli penelitian diatas maka dapat disimpulkan data mining adalah proses menganalisis sejumlah besar data untuk menemukan pola atau informasi yang tersembunyi yang dapat bermanfaat bagi pemilik data. Ini dilakukan dengan menggunakan teknik tertentu seperti clustering dan merupakan komponen penting dari Knowledge Discovery in Database (KDD). Data mining juga memungkinkan identifikasi dan pemanfaatan hubungan antar data yang sebelumnya tidak terlihat.

2.6 Knowledge Discovery in Database (KDD)

Menurut (Santoso, 2017) dalam jurnal (Limia Budiarti & Cendana, 2022) KDD adalah kegiatan yang meliputi pengumpulan, pemakaian data, historis untuk menemukan keteraturan, pola atau hubungan dalam set data berukuran besar. [5]

3. METODE PENELITIAN

Tahapan dalam proses penyelesaian permasalahan bertujuan agar penelitian tersebut menghasilkan data yang benar dan tepat. Tahapan penelitian dapat dilihat pada Gambar di bawah ini.



Gambar 1. Tahapan Penelitian

Tahapan yang ada dalam penelitian ini mencakup kegiatan yang akan dilakukan diagram ini menggambarkan proses penelitian yang akan ditempuh dan juga menggambarkan penelitian secara

keseluruhan. Tahapan yang akan dilakukan sebagai berikut:

a. Pendahuluan

Pada tahap ini penting untuk mengidentifikasi masalah dengan jelas dan terdefinisi dengan baik.

b. Pengumpulan Data

Pengumpulan data adalah bertujuan untuk mengumpulkan data yang diperlukan untuk penelitian dengan cara wawancara dengan bagian keuangan dan observasi.

c. Preprocessing

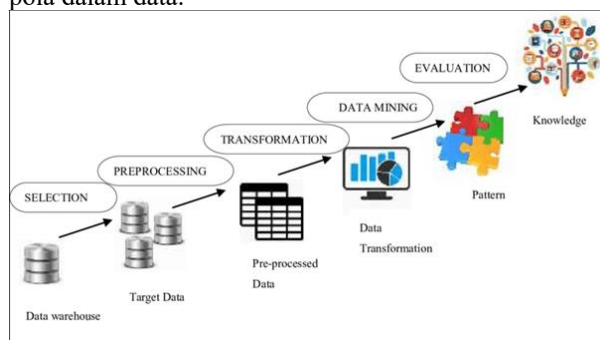
Preprocessing adalah penting untuk memastikan bahwa data tersebut bersih dan valid. Hal ini dapat dilakukan dengan melakukan beberapa langkah yaitu:

- *Missing values* adalah data yang hilang atau tidak tersedia. Missing values dapat dihapus, diimputasi, atau diabaikan, tergantung pada jenis data dan masalah yang ingin dipecahkan.
- *Outliers* adalah data yang jauh berbeda dari nilai-nilai lain dalam dataset. Outliers dapat dihapus atau diimputasi, tergantung pada jenis data dan masalah yang ingin dipecahkan.
- *Normalisasi data* adalah proses mengubah data menjadi skala yang sama. Normalisasi data dapat dilakukan dengan berbagai cara, seperti *min-max scaling* dan *z-score normalization*.

d. Pengolahan Data

Pada tahap ini data yang dikumpulkan diolah sesuai dengan tahapan *Knowledge Discovery in Database* (KDD) dengan menggunakan algoritma *K-Means*. Selanjutnya akan dilakukan pemrosesan data yang sebelumnya dari pengumpulan data dengan aplikasi *RapidMiner*.

Clustering data dengan *K-Means* adalah teknik pengolahan data yang digunakan untuk mengelompokkan data menjadi beberapa kelompok. Setiap kelompok memiliki karakteristik yang sama. Teknik ini dapat digunakan untuk mengidentifikasi pola dalam data.



Gambar 2. Proses KDD

e. Kesimpulan dan Saran

Membuat kesimpulan dari hasil penelitian yang dilakukan dan juga memberikan saran untuk pihak toko agar hasil dari penelitian ini menjadi lebih bermanfaat. proses KDD dimulai. Hasil dari data seleksi ini akan digunakan dalam proses data mining, disimpan di suatu berkas dan terpisah dari data operasional pengujian.

Table 1. Data Selection

No	Nama Produk	Kategori	Jumlah
1	Kuali Super 10	Peralatan Masak	27
2	Parutan Serbaguna 4 in 1 Warna	Peralatan dapur	110
3	Bak 28 Karet	Perlengkapan Rumah Tangga	11
4	Talenan Kayu Kecil	Peralatan Dapur	60
5	Keset Anvam	Alat Kebersihan	290
6	Pel Dragon S 15 Sumbu	Alat Kebersihan	33
7	Sapu Ijuk Dragon	Alat Kebersihan	98
8	Burner Kompor Gas Rinnai	Peralatan Masak	46
9	Piring Ramen Kotak Plastik	Peralatan Minum	180
10	Wiper Dragon	Alat Kebersihan	32
11	Wiper Boom Boom	Alat Kebersihan	33
12	Teflon Valentino Maxim 18	Peralatan Masak	4
13	Toples Sunrise 5 Liter	Wadah Penyimpanan	156
14	Baskom Plastik Olivia 14	Peralatan Dapur	51
15	Basko Jongsok Solido Shinpo	Perlengkapan Rumah Tangga	10
16	Teko 1.4 Liter Greenline	Peralatan Minum	30
17	Sendok 18 Gram	Peralatan Masak	115
18	Baskom Plastik Kudaku 12	Peralatan Dapur	44
19	Basko Bakso Smile	Perlengkapan Rumah Tangga	72
20	Gelas Mika Bening	Peralatan Minum	42
21	Keset Mie Tebal	Alat Kebersihan	41
22	Folding Hanger Lion Star 20 Stik	Perlengkapan Rumah Tangga	8
23	Jepitan Pengoreng Stenlis Kecil	Peralatan Masak	42
24	Teko Greenline 4.1 Liter	Peralatan Minum	33
25	Panci Susu Stenlis 20 CM	Peralatan Masak	12
26	Baskom Plastik Kudaku 20	Peralatan Dapur	39
27	Sikat Dorong Lantai Gazang Dragon	Alat Kebersihan	29
28	Botol Kecap Plastik 300 ML	Wadah Penyimpanan	144
29	Keranjang Sampah Kudaku Mini	Alat Kebersihan	96
30	Gayung Mandi Kudaku	Perlengkapan Mandi	306
31	---	---	---
32	Parutan Keju	Peralatan Dapur	48

f. Data Cleaning

Tahapan ini merupakan tahapan awal proses Knowledge Discovery in Database (KDD). Tahap ini, semua data mining value, berlebihan, dan tidak penting akan dibuang. Jika atribut dataset tidak ada atau kosong, data dianggap tidak memiliki nilai. Sebaliknya, jika atribut dataset lebih dari record berisikan nilai data yang sama, data dianggap redundant. Setelah data dibersihkan, data akan disimpan ke Microsoft Excel 2010 untuk langkah berikutnya. Berikut merupakan data yang sudah melalui tahap Cleaning.

g. Transformation Data

Tahap transformasi adalah tahap di mana data yang dipilih diubah sehingga sesuai untuk proses data mining. Proses transformasi dalam KDD sangat kreatif dan bergantung pada jenis atau pola informasi yang akan dicari dalam basis data.

Pada tahap ini, seluruh data operasional menghasilkan data pengelompokan atribut yang akan digunakan dalam proses data mining, yaitu atribut data jumlah dan kategori yang akan digunakan dalam proses data mining.

Pada atribut kategori data yang sebelumnya akan dilakukan perubahan data non-numeric ke bentuk numeric agar dapat diolah menggunakan algoritma K-Means.

h. Proses Clustering

Setelah data penjualan di Toko Eka Perabot pada bulan Januari hingga Februari 2025 berhasil dikumpulkan, kemudian dilakukan pengelompokan menggunakan metode K-Means Clustering. Algoritma K-Means digunakan untuk mengelompokkan 100 data penjualan dari toko tersebut. Dalam penelitian ini, dibentuk 3 kluster secara acak dengan 3 titik pusat awal, yaitu Centroid pertama, kedua, dan ketiga. Setelah titik pusat tersebut ditentukan, dihitung jarak antara setiap data penjualan ke masing-masing titik pusat dengan menggunakan metode Euclidean Distance. Selanjutnya, data dikelompokkan berdasarkan jarak terdekat hingga diperoleh titik pusat yang tidak berubah dan kluster yang sudah konvergen. Jika kluster belum konvergen, rumus rata-rata digunakan untuk mencari titik pusat baru selama proses iterasi. Proses iterasi berhenti ketika kluster sudah konvergen dan titik pusat tidak berubah lagi.

Pada penelitian ini, pengelompokan data penjualan Toko Eka Perabot menggunakan algoritma K-Means Clustering dilakukan secara komputerisasi dengan menggunakan software RapidMiner. Diagram alur algoritma K-Means Clustering pada pengelompokan data penjualan juga dibuat.

4. HASIL DAN PEMBAHASAN

4.1. Transformation

Tahap transformasi adalah tahap di mana data yang dipilih diubah sehingga sesuai untuk proses data mining. Proses transformasi dalam KDD sangat kreatif dan bergantung pada jenis atau pola informasi yang akan dicari dalam basis data.

Pada tahap ini, seluruh data operasional menghasilkan data pengelompokan atribut yang akan digunakan dalam proses data mining, yaitu atribut data jumlah dan kategori yang akan digunakan dalam proses data mining.

Pada atribut kategori data yang sebelumnya akan dilakukan perubahan data non-numeric ke bentuk numeric agar dapat diolah menggunakan algoritma K-Means.

No	Nama Produk	Kategori	Jumlah
1	Kuali Super 10	3	27
2	Parutan Serbaguna 4 in 1 Warna	2	110
3	Bak 28 Karet	5	11
4	Talenan Kayu Kecil	2	60
5	Keset Anyam	1	290
6	Pel Dragon 5 15 Sumbu	1	33
7	Sapu Ijuk Dragon	1	98
8	Burner Kompor Gas Rinnai	3	46
9	Piring Ramen Kotak Plastik	6	180
10	Wiper Dragon	1	32
11	Wiper Boom Boom	1	33
12	Teflon Valentino Maxim 18	3	4
13	Toples Sunrise 5 Liter	4	156
14	Baskom Plastik Olivia 14	2	51
15	Bangku Jongkok Solido Shinpo	5	10
16	Teko 1,4 Liter Greenline	7	30
17	Sendok 18 Gram	6	115
18	Baskom Plastik Kudaku 12	2	44
19	Bangku Bakso Smile	5	72
20	Gelas Mika Bening	7	42
21	Keset Mlie Tebal	1	41
22	Folding Hanger Lion Star 20 Stik	5	8
23	Jepitan Penggoreng Stenlis Kecil	3	42
24	Teko Greenline 4,1 Liter	7	33
25	Panci Susu Stenlis 20 CM	3	12
26	Baskom Plastik Kudaku 20	2	39
27	Sikat Dorong Lantai Gagang Dragon	1	29
28	Botol Kecap Plastik 300 ML	4	144
29	Keranjang Sampah Kudaku Mini	1	96
30	Gayung Mandi Kudaku	8	306
...		...	
222	Parutan Keju	2	48

4.2. Metode K-Means Clustering

Data siap untuk diolah dengan metode K-Means Clustering setelah melalui tahap pre-processing dan tranformasi. Perhitungan manual akan dilakukan menggunakan rumus jarak Euclidean dan data akan dipilih secara acak sebagai pusat cluster awal yaitu sebagai berikut:

a. Menentukan Jumlah Cluster

Untuk menentukan data penjualan yang banyak diminati, diminati, dan kurang diminati. Penelitian ini membutuhkan tiga cluster, banyak cluster tersebut telah diuji dengan metode Elbow yang menunjukkan bahwa K=3 adalah cluster yang tepat untuk penelitian ini.

b. Menentukan titik pusat awal

Penelitian ini menggunakan kategori harga murah, menengah, dan mahal untuk menentukan titik pusat awal.

Table 4. centroid awal

Titik Pusat	Nama	Kode Produk	jumlah
C0	Kuali Super 10	3	27
C1	Keset Anyam	1	290
C2	Toples Sunrise 5 Liter	4	156

Tabel di atas merupakan penentuan centroid awal yang digunakan untuk perhitungan data mining.

a. Menghitung jarak object dengan centroidTable

225

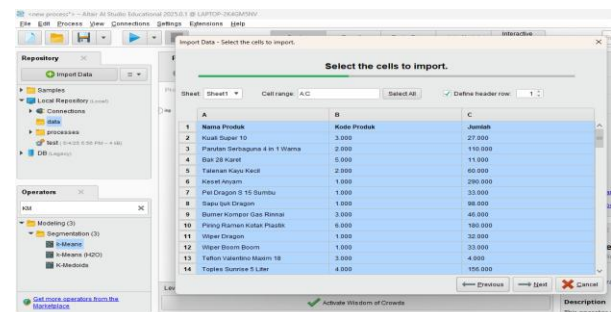
Nama Produk	C0	C1	C2	Terdekat	Cluster
Kuali Super 10	6	263	129	6	C0
Parutan Serbaguna 4 in 1 Warna	77	180	46	46	C2
Bak 28 Karet	22	279	145	22	C0
Talenan Kayu Kecil	27	230	96	27	C0
Keset Anyam	257	0	134	0	C1
Pel Dragon S 15 Sumbu	2	257	123	2	C0
Sapu Ijuk Dragon	65	192	58	58	C2
Burner Kompor Gas Rinnai	13	244	110	13	C0
Piring Ramen Kotak Plastik	147	110	24	24	C2
Wiper Dragon	3	258	124	3	C0
Wiper Boom Boom	2	257	123	2	C0
Teflon Valentino Maxim 18	29	286	152	29	C0
Toples Sunrise 5 Liter	123	134	0	0	C2
Baskom Plastik Olivia 14	18	239	105	18	C0
Bangku Jongkok Solido Shimpo	23	280	146	23	C0
Teko 1,4 Liter Greenline	5	260	126	5	C0
Sendok 18 Gram	82	175	41	41	C2
Baskom Plastik Kudaku 12	11	246	112	11	C0
Bangku Bakso Smile	39	218	84	39	C0
Gelas Mika Bening	9	248	114	9	C0
Keset Mie Tebal	8	249	115	8	C0
Folding Hanger Lion Star 20 Stik	25	282	148	25	C0
Jepitan Penggoreng Stenlis Kecil	9	248	114	9	C0
Teko Greenline 4,1 Liter	4	257	123	4	C0
Panci Susu Stenlis 20 CM	21	278	144	21	C0
Baskom Plastik Kudaku 20	6	251	117	6	C0
Sikat Dorong Lantai Gagang Dragon	5	261	127	5	C0
Botol Kecap Plastik 300 ML	111	146	12	12	C2
Keranjang Sampah Kudaku Mini	63	194	60	60	C2
Gayung Mandi Kudaku	273	17	150	17	C1
....					
Bak Air USA 24 Warna Yutaka	3	258	124	3	C0

Lalu ulangi langkah perhitungan untuk menentukan cluster pada setiap iterasi sampai iterasi ke-n dengan posisi data cluster tidak berubah.

Hasil akhir perhitungan dari jumlah 100 data menghasilkan C0 = 69 data, C1 = 7 data, dan C2 = 24 data yang dapat dilihat pada table berikut:

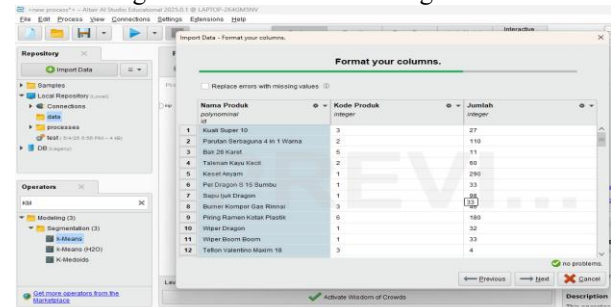
4.3. Penerapan Algoritma K Means dengan Rapidminer

Pada tahapan ini diawal dengan pemilihan data yang ingin di uji, dapat dilihat pada Gambar berikut:



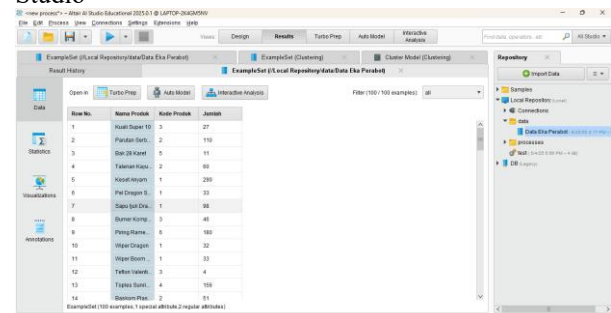
Gambar 4. Data yg ingin di uji

Kemudian Sistem akan menampilkan Format Columns. Format tersebut dapat diubah sesuai type data. Dalam penelitian data train yang akan digunakan semua colom berisi data teks dan data angka maka format yang dipilih adalah polynominal dan integer. Format Colomns untuk bulan adalah polynominal(label) dan menu adalah polynominal (id). Serta Kode Menu adalah Integer dan Jumlah adalah Integer.



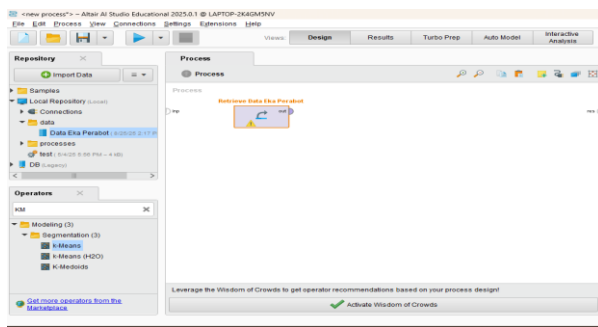
Gambar 5. Format Data

Data sudah berhasil masuk kedalam Rapid Miner Studio



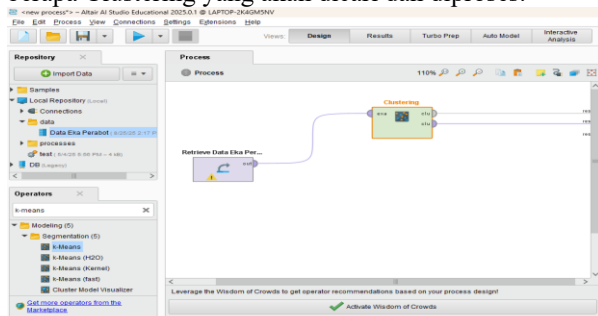
Gambar 6. Hasil Import Data

Selanjutnya pengujian data > pilih menu Design > Drag and drop file Uji Data kedalam halaman process pada RapidMiner.



Gambar 7. Uji Data

sambungkan retrieve Uji Data ke Clustering lalu pada pojok kanan parameters K diubah untuk menentukan berapa Clustering yang akan dicari dan diproses.



Gambar 8. Proses K Means

Dibawah ini menampilkan hasil Clustering dimana data-data penduduk tersebut sudah tertera Clusternya.

Row No.	Nama Produk	cluster	Kode Produk	Jumlah
1	Kopi Super 10	cluster_0	3	27
2	Pastilan Seris...	cluster_2	2	110
3	Ban 20 Karat	cluster_0	5	11
4	Takonan Kars...	cluster_0	2	60
5	Kopi Kaya	cluster_1	1	290
6	Pet Dragon S...	cluster_0	1	33
7	Barang Eka...	cluster_2	1	98
8	Bumer Kamp...	cluster_0	3	46
9	Poteng Rama...	cluster_2	6	180
10	Wiper Dragon	cluster_0	1	32
11	Wiper Boom...	cluster_0	1	33
12	Teknis Vatel...	cluster_0	3	4
13	Teknis Sani...	cluster_2	4	105
14	Baskom Pias...	cluster_0	2	51

Gambar 9. Hasil data

Dengan hasil jumlah Cluster 0 terdapat 7 items, Cluster 1 terdapat 15 items, dan Cluster 2 terdapat 8 item

4.3.1 Hasil Pengujian

a. Proses ke 2

Avg. within centroid distance

Avg. within centroid distance: -1577.384

Cluster Model

Cluster 0: 17 items
 Cluster 1: 83 items
 Total number of items: 100

b. Proses ke 3

Avg. within centroid distance

Avg. within centroid distance: -430.290

Cluster Model

Cluster 0: 24 items
 Cluster 1: 7 items
 Cluster 2: 69 items
 Total number of items: 100

c. Proses ke 4

Avg. within centroid distance

Avg. within centroid distance: -217.751

Cluster Model

Cluster 0: 66 items
 Cluster 1: 7 items
 Cluster 2: 17 items
 Cluster 3: 10 items
 Total number of items: 100

d. Proses ke 5

Avg. within centroid distance

Avg. within centroid distance: -110.173

Cluster Model

Cluster 0: 50 items
 Cluster 1: 7 items
 Cluster 2: 10 items
 Cluster 3: 17 items
 Cluster 4: 16 items
 Total number of items: 100

e. Proses ke 6

Avg. within centroid distance

Avg. within centroid distance: -64.839

Cluster Model

Cluster 0: 43 items
Cluster 1: 7 items
Cluster 2: 10 items
Cluster 3: 16 items
Cluster 4: 14 items
Cluster 5: 10 items
Total number of items: 100

f. Proses ke 7

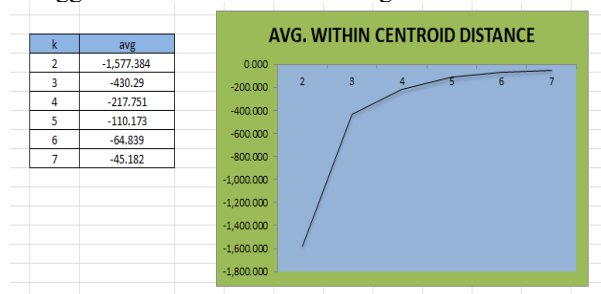
Avg. within centroid distance

Avg. within centroid distance: -45.182

Cluster Model

Cluster 0: 16 items
Cluster 1: 6 items
Cluster 2: 14 items
Cluster 3: 7 items
Cluster 4: 7 items
Cluster 5: 46 items
Cluster 6: 4 items
Total number of items: 100

Dari hasil analisa Cluster yang didapatkan data yang diukur untuk mencari jumlah Cluster yang paling baik untuk data penduduk Desa Rengasbandung dengan menggunakan Elbow Method sebagai berikut:



Gambar 10. Metode Siku (Elbow Method)

Gambar diatas menunjukkan bahwa Cluster yang paling baik digunakan adalah Cluster 3 karena pada Cluster 3 menunjukkan sudut yang paling siku dengan hasil avgnya -430,29. Maka dari itu Cluster 3 merupakan Cluster yang paling baik untuk penelitian ini.

5. KESIMPULAN

5.1 Berdasarkan kegiatan penelitian yang telah dilakukan, hasil implementasi dan pengujian menunjukkan bahwa

- Dari hasil pengujian data mining menggunakan analisis data Knowledge in Database dengan penerapan metode K-Means Clustering pada produk alat perabotan di Toko Eka Perabot, ditemukan 69 produk yang masuk ke dalam cluster 0 yang paling kurang diminat, 7 produk yang masuk ke dalam cluster 1 yang sangat diminati, dan 24 produk yang masuk ke dalam cluster 2 yang diminati. Dengan hasil tersebut, Toko Eka Perabot dapat mengetahui klasifikasi produk dan meningkatkan strategi penjualan berdasarkan data yang telah diperoleh.
- Dari hasil pengujian performa dalam penelitian ini menggunakan RapidMiner dengan alat avg. within centroid distance hingga proses ke-7, kemudian diukur menggunakan Elbow Method diperoleh nilai $k-2 = -1577,384$, $k-3 = -430,290$, $k-4 = -217,751$, $k-5 = -110,173$, $k-6 = -64,839$, $k-7 = -45,182$. Dari hasil tersebut, cluster yang paling baik adalah cluster 3 yang menunjukkan sudut paling siku dengan nilai avg. -430,290. Dengan hasil performa ini, Toko Eka Perabot dapat menerapkan metode ini untuk mengetahui klasifikasi terbaik dari data penjualan di toko tersebut.

DAFTAR PUSTAKA

- [1] Arifatul Ulya, F., Abdullah, A. N., Aisya Hanan, T., & Manfaati Nur, I. (2023). Pengelompokan Tingkat Pengangguran Terbuka Di Jawa Tengah Menggunakan Metode K – Means Clustering. *Journal Of Data Insights*, 1(2), 71–80. <https://doi.org/10.26714/jodi.v1i2.148>
- [2] Bella Pratiwi. (2024). Penerapan Model Machine Learning dalam Pengembangan Web App “CRACKSAFE” untuk Deteksi Keretakan pada Dinding Bangunan. *Inisiatif: Jurnal Dedikasi Pengabdian Masyarakat*, 3(1), 59–68. <https://doi.org/10.61227/inisiatif.v3i1.148>
- [3] Dari, W., Aliyah, S., & Setiyawati, R. I. (2023). Penerapan Aplikasi Microsoft Excel Dalam Pengelolaan Data Nilai Siswa Pada TK Kartika I-1 Medan Helvetia. *Jurnal Pengabdian Masyarakat Sains Dan Teknologi*, 2(2), 198–205. <https://doi.org/10.58169/JPMsAINTEK.V2I2.203>
- [4] Haryanti, M. F., Fauzi, A., Jelita, A. A., Setiyowati, A., Octarina, A., Edina, E. P., Aulia, R. Z., & Fitriana, S. (2024). Pengaruh Data Mining, Strategi Perusahaan, Terhadap Laporan Kinerja Perusahaan. *Jurnal Portofolio : Jurnal*

- Manajemen Dan Bisnis, 3(1), 71–90.
<https://doi.org/10.70704/jpjm.v3i1.285>
- [5] Kurniadi, D., & Sugiyono, A. (2020). Pengelompokan Data Akademik Menggunakan Algoritma K-Means Pada Data Akademik Unissula. *Jurnal Transformatika*, 18(1), 93–101.
<https://doi.org/10.26623/transformatika.v18i1.2277>
- [6] Limia Budiarti, R., & Cendana, G. (2022). KLASIFIKASI DATA NASABAH KREDIT PINJAMAN MENGGUNAKAN DATA MINING DENGAN METODE K-MEANS PADA MEGA CENTRAL FINANCE. *JURNAL AKADEMIKA*, 14(2), 88–94.
<https://doi.org/10.53564/akademika.v14i2.866>
- [7] Murdiaty, M., Angela, A., & Sylvia, C. (2020). Pengelompokan Data Bencana Alam Berdasarkan Wilayah, Waktu, Jumlah Korban dan Kerusakan Fasilitas Dengan Algoritma K-Means. *JURNAL MEDIA INFORMATIKA BUDIDARMA*, 4(3), 744.
<https://doi.org/10.30865/mib.v4i3.2213>
- [8] Nasari, F., Yos, J. K. L., Km, S. 6, & Medan, T. M. (2021). Pengelompokan Daerah Produksi Kelapa Sawit Menggunakan Algoritma K-Means Clustering: Grouping Oil Palm Production Areas Using K-Means Clustering Algorithm. *Jurnal Sains Dan Ilmu Terapan*, 4(2), 34–38.
<https://doi.org/10.59061/JSIT.V4I2.108>
- [9] Putri, D. A., & Z, G. (2025). PERBANDINGAN ALGORITMA K-MEANS DAN K-MEDOIDS DALAM PENGELOMPOKKAN BANK UMUM SYARIAH DI INDONESIA DENGAN DAVIES BOULDIN INDEKS, CALINSKI HARABASZ INDEKS & KOEFISIEN SILHOUTTE. *Inovasi Pembangunan : Jurnal Kelitbangan*, 13(1).
<https://doi.org/10.35450/JIP.V13I01.623>
- [10] Sahri, A. eka. (2023). Implementasi K-Means Clustering Dalam Pengelompokan Produksi Daging Ayam Menurut Provinsi Di Indonesia. *Jurnal PROCESSOR*, 18(2).
<https://doi.org/10.33998/processor.2023.18.2.853>
- [11] Sholeh, M., & Aeni, K. (2023). Perbandingan Evaluasi Metode Davies Bouldin, Elbow dan Silhouette pada Model Clustering dengan Menggunakan Algoritma K-Means. *STRING (Satuan Tulisan Riset Dan Inovasi Teknologi)*, 8(1), 56–65.
<https://doi.org/10.30998/STRING.V8I1.16388>
- [12] Yoliadi, D. N. (2023). Data Mining Dalam Analisis Tingkat Penjualan Barang Elektronik Menggunakan Algoritma K-Means. *Insearch: Information System Research Journal*, 3(01).
<https://doi.org/10.15548/isrj.v3i01.5829>
- [13] aki, A., Irwan, & Sembe, I. A. (2022). Penerapan K-Means Clustering dalam Pengelompokan Data (Studi Kasus Profil Mahasiswa Matematika FMIPA UNM). *Journal of Mathematics, Computations and Statistics*, 5(2), 163–176.
<http://journal.unm.ac.id/index.php/JMATHCOS/article/view/3932>