

PENERAPAN *K-MEDOIDS* UNTUK MENGELOMPOKKAN PRODUK PADA TOKO *ONLINE* ALDAF STORE DEPOK

Nabila Khatami

Universitas Indraprasta PGRI
Jalan Raya Tengah No.80, RT.6/RW.1, Gedong, Kec. Pasar Rebo, Jakarta Timur
nabilakhatami01@gmail.com

ABSTRAK

Penelitian ini bertujuan mengelompokkan produk di Aldaf Store berdasarkan harga, jumlah penjualan dan kategori menggunakan algoritma *K-Medoids*. Algoritma ini dipilih karena efektif dan tahan terhadap *outlier*. Data dinormalisasi agar skala seimbang sebelum proses *clustering*. Tahapan dimulai dari penentuan jumlah *cluster* optimal, perhitungan jarak *Euclidean*, hingga pengelompokan produk berdasarkan kedekatan ke *medoid*. Hasil *clustering* mengidentifikasi beberapa pola, seperti produk dengan penjualan tinggi dan harga sedang, serta produk dengan penjualan rendah dan harga tinggi. Evaluasi dengan *Silhouette Score* menghasilkan rata-rata 0.683, menunjukkan kualitas *cluster* yang cukup baik. Untuk memudahkan analisis, hasil ditampilkan melalui sistem web interaktif dalam bentuk grafik dan tabel. Penelitian ini diharapkan dapat mendukung manajemen Aldaf Store dalam pengambilan keputusan strategis, terutama terkait promosi, pengelolaan stok dan penentuan produk unggulan.

Kata Kunci: *K-Medoids*, *Clustering*, Penjualan Pakaian

ABSTRACT

This study aims to group products at Aldaf Store based on price, sales volume, and category using the K-Medoids algorithm. This algorithm was chosen because it is effective and resistant to outliers. The data were normalized to ensure a balanced scale before clustering. The stages began with determining the optimal number of clusters, calculating the Euclidean distance, and grouping products based on their proximity to the medoid. The clustering results identified several patterns, such as products with high sales and moderate prices, as well as products with low sales and high prices. Evaluation using the Silhouette Score yielded an average of 0.683, indicating fairly good cluster quality. To facilitate analysis, the results are displayed in an interactive web system as graphs and tables. This research is expected to support Aldaf Store management in strategic decision-making, particularly regarding promotions, inventory management, and flagship product selection.

Keywords: *K-Medoids*, *Clustering*, *Apparel Sales*

PENDAHULUAN

Aldaf Store, toko fashion muslim di Depok, menghadapi kesulitan mengelompokkan produk karena banyaknya variasi. Penelitian ini menerapkan *K-Medoids* untuk mengelompokkan produk berdasar harga penjualan, dan kategori guna membantu mengenali pola penjualan dan Menyusun strategi bisnis yang lebih efektif. *Data Mining* adalah proses menemukan pola penting dari data jumlah besar dengan teknik tertentu dan merupakan bagian *Knowledge Discovery in Database* (KDD) (Fayyad et al., 1996) teknik-teknik data mining meliputi:

1. Klasifikasi: Mengelompokkan data berdasarkan kategori yang telah ada.
2. Prediksi: Memperkirakan nilai atau kejadian di masa depan dari pola historis.

3. Clustering: Mengelompokkan data tanpa tabel berdasarkan kemiripan karakteristik.
4. Estimasi: Menentukan nilai numerik dari data baru.
5. Asosiasi: Mencari hubungan antar atribut yang sering muncul bersama.

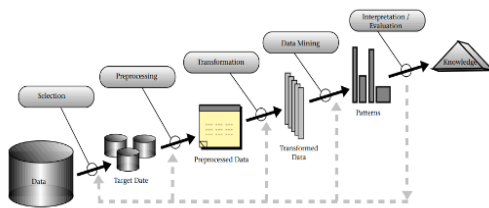
Penelitian ini menggunakan metode *clustering* karena mampu menemukan kelompok tersembunyi dalam data secara objektif (Triayudi et al., 2024). *Custering* yang baik memiliki kemiripan tinggi dalam satu cluster dan perbedaan jelas antar *cluster* (Silitonga & Morina, 2017). Sebagai bagian dari *Unsupervised Learning*, *Clustering* bertujuan mengungkap pola atau struktur tersembunyi pada data yang tidak berlabel (Madhulatha, 2011).

Penelitian ini menggunakan algoritma *K-Medoids* yang memiliki objek nyata sebagai medoid sehingga *cluster* lebih stabil dan representatif (Kamalia & Nawangsih, 2025). Implementasi dilakukan menggunakan *Python* dengan dukungan library seperti *pandas*, *NumPy*, dan *Matplotlib* untuk analisis dan visualisasi data (Anggraini et al., 2025). Selain itu, RapidMiner digunakan sebagai alat bantu visual untuk mengeksekusi algoritma data mining secara efisien melalui antarmuka yang mudah dipahami (Aprilla Donny Aji Baskoro Lia Ambarwati I Wayan Simri Wicaksana Editor & Sanjaya, 2013).

METODE PENELITIAN

1. Knowledge Discovery in Database

Penelitian ini menggunakan pendekatan Knowledge Discovery in Database (KDD) untuk menggali informasi penting melalui teknik data mining (Narasati et al., 2024). Tahapan KDD terdiri dari:



Gambar 1. Tahapan KDD
 Sumber: (Fayyad et al., 1996)

- Data Selection*: Identifikasi masalah dan pengumpulan data yang akan digunakan dalam algoritma *K-Medoids*.
- Preprocessing*: pembersihan data, termasuk penanganan *missing value* untuk meningkatkan akurasi.
- Transformation*: Konversi data, seperti nama produk menjadi kode numerik agar siap dianalisis.
- Data Mining*: Penerapan algoritma *K-Medoids* untuk mengelompokkan data berdasarkan kemiripan.
- Interpretation/Evaluation*: Evaluasi hasil *cluster* menggunakan metrik untuk menilai performa model.

2. Algoritma K-Medoids

K-Medoids adalah metode *clustering* yang menggunakan objek nyata (*medoid*) sebagai pusat *cluster*. Prosesnya meliputi:

- Inisialisasi *Medoid*: Pemilihan *k* objek secara acak.
- Perhitungan Jarak: Menggunakan *Euclidean Distance* untuk mengukur kedekatan antar objek.
- Pengelompokan: Setiap data dialokasikan ke medoid terdekat.
- Evaluasi dan Pertukaran: Medoid diganti bila ditemukan konfigurasi dengan total jarak yang lebih kecil.
- Iterasi: Proses diulang hingga tidak ada perubahan medoid atau total *cost* minimum tercapai.

Perhitungan jarak antar objek menggunakan rumus Eulidean Distance (Fitriyani et al., 2017).

$$d_{ij} = \sqrt{\sum_{k=1}^p (x_{ik} - x_{jk})^2}$$

dengan:

d_{ij} = Jarak antar objek *i* dan *j*.

x_{ik}, x_{jk} = Nilai variabel ke-*k* pada objek *i* dan *j*.
P = Jumlah variabel.

HASIL DAN PEMBAHASAN

Penelitian ini menghasilkan pengelompokan produk Aldaf Store dengan algoritma *K-Medoids*, mencakup pengujian sistem, analisis *cluster* dan evaluasi efektivitasnya dalam mendukung strategi bisnis. Data penjualan Aldaf Store selama 1 tahun ini berjumlah 1.122. Sebelum algoritma diterapkan, data melalui tahap *preprocessing* berupa pembersihan, konversi data kategorikal ke numerik, dan agregasi. Setelah itu dilakukan normalisasi dengan Min-Max Scaling agar skala tiap variabel seimbang, sehingga pembentukan *cluster* menjadi lebih akurat dan representatif (Flomina et al., 2025).

Rumus Min-Max Scaling:

$$X_{Scaled} = \frac{X - X_{Min}}{X_{Max} - X_{Min}}$$

dengan:

X = Nilai asli yang dinormalisasi.

X_{Min} = Nilai minimum dari seluruh kolom tersebut.

X_{Max} = Nilai maksimul dari seluruh kolom tersebut.

Tabel 1. Tabel Hasil Preprocessing Data

ID	Produk	Pesanan	Transaksi	Pendapatan	Diskon
031	Kurta FL – BTAK 1 XL	0	0	0	0
153	Sirwal Formal 1 AL - M	0	0	0	0
239	Sirwal Kouzi – Hitam XXL	1	0	1	0,867
23B	Sirwal Casual – Dongker M	0	0	0,278	1
27M	Ankle Kid – Mocca XXL	0	0	0,111	0,4
1JX	Sirwal Semi – Dark Milo M	0	0	0,278	1

Setelah dinormalisasi, tahap selanjutnya adalah:

1. Menentukan k (Jumlah Cluster)

Tabel 2. Tabel Medoid Awal

ID	Produk	Pesanan	Transaksi	Pendapatan	Diskon
C0	Sirwal Kouzi – Hitam XXL	1	0	1	0,86666667
C1	Sirwal Casual – Dongker M	0	0	0,27777778	1
C2	Ankle Kid – Mocca XXL	0	0	0,11111111	0,4

2. Menghitung Euclidean Distance

$$d_{ij} = \sqrt{\sum_{k=1}^p (x_{ik} - x_{jk})^2}$$

$d_{Bunayya Sirwal Kouzi-Hitam,XXL,C0}$
 $= \sqrt{(1-0)^2 + (0-0)^2 + (1-0)^2 + (0,867-0)^2} = 1,658647$

$d_{Bunayya Sirwal Casual-Dongker,M,C1}$
 $= \sqrt{(0-0)^2 + (0-0)^2 + (0,278-0)^2 + (1-0)^2} = 1,037863$

$d_{Bunayya Angkel Pants Kid-Mocca,XXL,C2}$
 $= \sqrt{(0-0)^2 + (0-0)^2 + (0,111-0)^2 + (0,4-0)^2} = 0,415145$

Tabel 3. Hasil Perhitungan K-Medoids Iterasi ke-1

No	Produk	Jarak K-Medoids			Jarak	Cluster
		Cost 1	Cost 2	Cost 3		
1	Kurta FL – BTAK 1 XL	1,658647	1,037863	0,415145	0,415	3
2	Sirwal Formal AL - M	1,658647	1,037863	0,415145	0,415	3
3	Sirwal Kouzi – Hitam XXL	0	1,240719	1,417008	0	1
4	Sirwal Casual – Dongker M	1,240719	0	0,622718	0	2
5	Ankle Kid – Mocca XXL	1,417008	0,622718	0	0	3
13	Sirwal Semi – Dark Milo M	1,240719	0	0,622718	0	2

Setelah diperoleh nilai *cost* pada iterasi pertama, proses dilanjutkan ke iterasi kedua dengan mencoba calon medoid baru yang berbeda dengan medoid sebelumnya.

Tabel 4. Tabel Medoid Baru

ID	Produk	Pesan	Transaksi	Pendapatan	Diskon
D1	Kurta Orizont – Army XL	0	0	0,203703704	0,733333333
D2	Kurta Orizont Kid – Army 12	0	0	0,203703704	0,733333333
D3	Sirwal Semi – Dark Milo M	0	0	0,277777778	1

Tabel 5. Hasil Perhitungan K-Medoids Iterasi ke-2

No	Produk	Jarak K-Medoids			Jarak	Cluster
		Cost 1	Cost 2	Cost 3		
1	Kurta FL – BT AK1 XL	0,7611	0,7611	1,033	0,761099847	1
2	Sirwal Formal AL - M	0,7611	0,7611	1,033	0,761099847	1
3	Sirwal Kouzi – Hitam XX L	1,285249	1,28524	1,24071	1,240718629	3
4	Sirwal Casual – Donker M	0,276764	0,276764	0	0	3

5	Ankle Kid – Moca XXL	0,345954	0,345954	0,622718	0,345954	1
...
13	Sirwal Semi – Dark Milo M	0,276764	0,276764	0	0	3

3. Hitung Total Simpangan (S)

Setelah iterasi ke-1 dan ke-2, total simpangan (S) dihitung dengan mengurangkan total *cost* baru dengan *cost* sebelumnya (Andini & Arifin, 2020). Berikut persamannya:

$$S = \text{Total Cost Baru} - \text{Total Cost Lama} \\ = 3,93916354 - 2,14491775 \\ = 1,794246$$

Karena nilai $S > 0$, maka *medoid* tidak diganti.

4. Silhouette Score

Setelah menghitung total simpangan, langkah berikutnya menghitung *Silhouette Score* untuk menilai kualitas *clustering* (Rousseeuw, 1987). Berikut persamaannya:

$$s(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))}$$

Berdasarkan (Kaufman et al., 2005). Interpretasi nilai *silhouette* sebagai berikut:

Tabel 6. Interpretasi Nilai Silhouette Coefficient

Silhouette Coefficient	Akurasi
$\geq 0.7-1$	Struktur Kuat
0.5-0.7	Struktur Sedang
$0.25 \leq 0.5$	Struktur Lemah

Tabel 7. Hasil Silhouette Score

Produk Kombinasi	Cluster	Silhouette Score
Sirwal Kouzi – Hitam,xxl	1	0.000
Sirwal Casual – Dongker,M	2	0.821
Sirwal Semi – Abu2,Xl	2	0.841
Sirwal Casual – Hitam,Xl	2	0.841
Kemko Fadl – Maroon,M	2	0.739
...
Ankle Pants Kid - Charcoal,XXL	3	0.700

Rata-rata *Silhouette Score* sebesar 0.683 menunjukkan struktur *clustering* yang cukup baik. Mayoritas produk di *cluster* 2 dan 3 memiliki nilai diatas 0.7 (struktur kuat).

5. Pelabelan Otomatis *Top, Medium, Low Sales*

Tabel 8. Hasil Perhitungan Label

Rat a Pes ana n	Rata Tra nsak si	Rata Pend apata n	Rata Disko n	Comb ined Score	R an k	La bel
1	0	1	0,866 66666 7	0,716 66666 7	1	Top Sales
0	0	0,240 74074 1	0,866 66666 7	0,276 85185 2	2	Me diu m Sal es
0	0	0,041 66666 7	0,15	0,047 91666 7	3	Lo w Sal es

6. Pengujian Data Menggunakan *RapidMiner*

Cluster Model

Cluster 0: 1 items
 Cluster 1: 4 items
 Cluster 2: 8 items
 Total number of items: 13

Gambar 2. Centroid K-Medoids Data Uji

c. Tampilan Hasil *Clustering*

Laporan Analisis Produk

Nama Produk	Nama Variasi	Jumlah Pesana	Total Pendapatan	Jumlah Transak	Rata2 Harga Awal	Cluster Label	Jarak ke Medoid
Bunayya Ankle Pants Kid	Charcoal,XXL	1	140,000	1	140,000	Low Sales	0.0
Bunayya Ankle Pants Kid	Mocca,XXL	1	140,000	1	140,000	Low Sales	0.0
Bunayya Kemko Fadl	Maroon,M	2	410,000	1	235,000	Top Sales	0.0
Bunayya Kurta Orizont	Army,XL	1	215,000	1	215,000	Medium Sales	0.1116
Bunayya Kurta Orizont Kids	Army,12	1	170,000	1	170,000	Low Sales	0.3348
Bunayya Kurta Orizont Kids	Maroon,2	1	155,000	1	155,000	Low Sales	0.1674
Bunayya Sirwal Casual	Dongker,M	1	205,000	1	205,000	Medium Sales	0.0
Bunayya Sirwal Casual	Hitam,XL	1	205,000	1	205,000	Medium Sales	0.0
Bunayya Sirwal Formal	Abu Lumut,M	1	195,000	1	195,000	Medium Sales	0.1116
Bunayya Sirwal Kouzi	Hitam,xxl	1	215,000	1	215,000	Medium Sales	0.1116
Bunayya Sirwal Semiformal	Abu2,XL	1	195,000	1	195,000	Medium Sales	0.1116
Bunayya Sirwal Semiformal	Dark milo,M	1	195,000	1	195,000	Medium Sales	0.1116
Kurta Indonesia Facelift	BTAK 1,XL	1	215,000	1	215,000	Medium Sales	0.1116

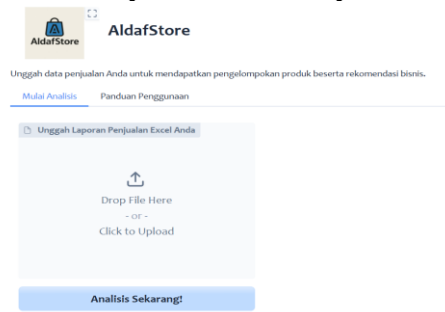
Depok, 27 Juli 2025


khaerani ulfah

Gambar 5. Tampilan Hasil Clustering

7. Tampilan Layar

a. Tampilan Beranda Aplikasi



Gambar 3. Tampilan Beranda

b. Tampilan Setelah Analisis



Gambar 4. Tampilan Setelah Analisis Data Uji

d. Tampilan Visual Data Penjualan Aldaf Store dalam 1 Tahun



Gambar 6. Tampilan Visual Data Penjualan Aldaf Store dalam 1 Tahun

SIMPULAN

Berdasarkan hasil penelitian, K-Medoids dapat mengelompokkan produk Aldaf Store secara efisien dan terstruktur, membantu melihat pola penjualan, menentukan kategori *Top*, *Medium*, dan *Low Sales*, serta mendukung keputusan bisnis berbasis data.

SARAN

Sistem disarankan terintegrasi ke *platform* toko, menambah variabel seperti *rating* dan dievaluasi rutin agar tetap relevan dengan tren pasar.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada dosen pembimbing atas bimbingan dan arahnya selama proses penelitian. Ucapan terima kasih juga ditujukan kepada diri sendiri atas usaha dan ketekunan yang telah diberikan hingga penelitian ini selesai.

DAFTAR PUSTAKA

Andini, A. D., & Arifin, T. (2020). Implementasi Algoritma K-Medoids Untuk Klasterisasi Data Penyakit Pasien di RSUD Kota Bandung. *Jurnal Responsif*, 2(2), 128–138. <http://ejurnal.ars.ac.id/index.php/jti>

Anggraini, D., Ayu Natalia Hutabarat, R., De Luisa Sitorus, S., Niarina Nabilah, W., Andri Hutapea, T., William Iskandar Ps, J. V, Baru, K., Percut Sei Tuan, K., Deli Serdang, K., & Utara, S. (2025). PT. Media Akademik Publisher Visualisasi Supremum dan Infimum Menggunakan Pemrograman Python.

JMA), 3(4), 3031–5220. <https://doi.org/10.62281>

Aprilla Donny Aji Baskoro Lia Ambarwati I Wayan Simri Wicaksana Editor, D. C., & Sanjaya, R. (2013). *Identitas Belajar Data Mining dengan RapidMiner Hak Cipta © pada Penulis Hak Guna mengikuti Open Content model Desain sampul: Dennis Aprilla C.*

Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996). *From Data Mining to Knowledge Discovery in Databases* (© AAAI) (Vol. 17). www.ffly.com/

Fitriyani, Fitriani, R., & Rosmawanti, N. (2017). Penerapan Algoritma Euclidean Distance Untuk Pemilihan Paket Internet Berdasarkan Wilayah.

Flomina, K., Rosman, E., Nasution, M. I., Febrina, Y. K., Hasibuan, R. S., & Padang, P. N. (2025). Penerapan Data Mining untuk Pemetaan Kinerja Akademik Mahasiswa dengan Metode K-Means. In *Journal of Science and Social Research: Vol. VIII* (Issue 2). <http://jurnal.goretanpena.com/index.php/JSSR>

Kamalia, A. Z., & Nawangsih, I. (2025). Identifikasi Pola Tingkat Kesenjangan Ketuntasan Pendidikan Di Indonesia Dengan Menggunakan Metode K-Medoids Clustering. *Jurnal Teknologi Informasi Dan Ilmu Komputer*, 12(2), 321–330. <https://doi.org/10.25126/jtiik.2025129219>

Kaufman, Leonard, Rousseeuw, & Peter J. (2005). *Finding Groups in Data: An Introduction to Cluster Analysis.*

Madhulatha, T. S. (2011). Comparison between K-Means and K-Medoids Clustering Algorithms. *International Journal of Advanced Computing (IJAC)*.

Narasati, R., Lestari, R., Herdiana, R., Hamonangan, R., & Anawar, S. (2024). Penerapan Algoritma K-Medoid Pada Penjualan Kerudung E-

- Commerce Shopee: Fifau Hijab. In *Jurnal Mahasiswa Teknik Informatika* (Vol. 8, Issue 2).
- Rousseeuw, P. J. (1987). Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. In *Journal of Computational and Applied Mathematics* (Vol. 20).
- Silitonga, P. D. P., & Morina, I. S. (2017). *Klusterisasi Pola Penyebaran Penyakit Pasien Berdasarkan Usia Pasien Dengan Menggunakan K-Means Clustering* (Issue 2).

Triayudi, A., Pinastawa, I. W. R., Pratama, J. D., Ningsih, S., Ilham, Wibowo, A., Judijanto, L., Sabililrasyad, I., Matondang, N., Hasnita, Desnelita, Y., & Gustientedina. (2024). *Clustering. Pertama*, 1–8.

Biografi Penulis



Nabila Khatami lahir di Jakarta, 17 Desember 2000.