

IMPLEMENTASI ALGORITMA *K-MEANS CLUSTERING* DATA PENJUALAN PADA WARUNG SEMBAKO ISAN MENGGUNAKAN *RAPIDMINER*

¹Daviqia Fadel, ²Fajri Maulana Azzam Harahap, ³Irsad Fauzan, ⁴Muhammad Azriel, ⁵Muhammad Fadlan Jabbar, ⁶Maulana Fansyuri

¹Sistem Informasi, Ilmu Komputer, Universitas Pamulang, Kota Tangerang Selatan, Indonesia

²Sistem Informasi, Ilmu Komputer, Universitas Pamulang, Kota Tangerang Selatan, Indonesia

³Sistem Informasi, Ilmu Komputer, Universitas Pamulang, Kota Tangerang Selatan, Indonesia

⁴Sistem Informasi, Ilmu Komputer, Universitas Pamulang, Kota Tangerang Selatan, Indonesia

⁵Sistem Informasi, Ilmu Komputer, Universitas Pamulang, Kota Tangerang Selatan, Indonesia

⁶Sistem Informasi, Ilmu Komputer, Universitas Pamulang, Kota Tangerang Selatan, Indonesia

¹daviqiafadel01@gmail.com, ²fajriazzam90@gmail.com, ³irsadfzn@gmail.com,
⁴azrielm826@gmail.com, ⁵mohamadfadlan1204@gmail.com, ⁶dosen02359@unpam.ac.id

Abstract

This study aims to apply the K-Means Clustering algorithm with the help of RapidMiner software on sales data at Warung Sembako Isan. In managing small businesses such as grocery stores, processing sales data manually often faces various challenges, such as errors in recording and difficulties in identifying sales trends. Therefore, data mining techniques, especially Clustering methods, are used to categorize products based on their sales capabilities. This process is carried out using RapidMiner, which allows analysis without the need for programming through a visual interface. The data were analyzed using the K-Means algorithm with parameter $k = 3$, which produces three categories: products with high potential, medium potential, and low potential. The results of this Clustering make it easier for shop owners to understand product performance, develop storage strategies, and plan more efficient promotions. This study shows that the use of simple technology can improve operational efficiency and assist MSMEs in data-based decision making.

Keywords: *Data Mining, Clustering, K-Means, RapidMiner, Grocery Store..*

Abstrak

Studi ini bertujuan untuk menerapkan algoritma K-Means Clustering dengan bantuan perangkat lunak *RapidMiner* pada data penjualan di Warung Sembako Isan. Dalam pengelolaan usaha kecil seperti toko sembako, melakukan pengolahan data penjualan dengan cara manual sering kali menghadapi berbagai tantangan, seperti kesalahan dalam pencatatan dan kesulitan dalam mengidentifikasi tren penjualan. Oleh karena itu, teknik penambangan data, khususnya metode pengelompokan, digunakan untuk mengkategorikan produk berdasarkan kemampuan penjualannya. Proses ini dilakukan dengan menggunakan *RapidMiner*, yang memungkinkan analisis tanpa keharusan untuk melakukan pemrograman melalui antarmuka yang bersifat visual. Data dianalisis dengan menggunakan algoritma *K-Means* dengan parameter $k = 3$, yang menghasilkan tiga kategori: produk dengan potensi tinggi, potensi sedang, dan potensi rendah. Hasil dari pengelompokan ini memudahkan pemilik toko untuk memahami kinerja produk, menyusun strategi penyimpanan, serta merencanakan promosi yang lebih efisien. Penelitian ini menunjukkan bahwa penggunaan teknologi sederhana dapat meningkatkan efisiensi operasional dan membantu UMKM dalam pengambilan keputusan yang berbasis data.

Kata Kunci: *Data Mining, Clustering, K-Means, RapidMiner, Toko Sembako.*

A. PENDAHULUAN

Toko sembako adalah jenis usaha kecil yang memiliki peranan penting dalam memenuhi kebutuhan dasar warga sehari-hari. Di tengah persaingan bisnis dan perubahan pasar yang terus berlangsung, warung atau toko sembako masih menjadi pilihan utama bagi masyarakat untuk memenuhi kebutuhan pokok seperti beras, minyak goreng, gula, dan barang-barang rumah tangga lainnya. Contohnya adalah Toko Sembako Isan, yang telah melayani pelanggan di sekitar tempatnya selama beberapa tahun. Dengan meningkatnya jumlah transaksi dan variasi barang yang ditawarkan, pengelolaan data penjualan dan persediaan menjadi lebih rumit dan memerlukan metode yang lebih terorganisir.

Selama ini, Toko Sembako Isan masih menggunakan cara manual dalam mengatur data penjualan dan inventaris, yang berisiko terkena kesalahan dalam pencatatan, kesulitan dalam mengenali pola penjualan, dan pengambilan keputusan yang kurang efektif mengenai pembelian stok. Dalam konteks usaha kecil, penerapan teknologi informasi yang sederhana dan relevan bisa sangat membantu dalam meningkatkan efisiensi dan efektivitas operasional. Salah satu solusi yang dapat diterapkan adalah analisis data dengan menggunakan metode data mining, khususnya teknik pengelompokan.

Pengelompokan dijelaskan sebagai metode penting dalam penambangan data yang menyusun data tanpa label menjadi bentuk informasi yang bermanfaat. Juga ditekankan ada tantangan seperti: volume data yang besar, kemampuan untuk meningkatkan skala, dan pemeriksaan keabsahan hasil Alasali dan Ortakci (2024).

Dalam konteks Toko Sembako Isan, penggunaan *K-Means Clustering* memungkinkan pemilik untuk mengelompokkan produk berdasarkan ciri-ciri tertentu seperti frekuensi penjualan, margin keuntungan, atau perputaran stok. Contohnya, barang-barang dengan penjualan yang tinggi dapat dikelompokkan menjadi kategori “produk utama”, sementara barang yang jarang terjual bisa dimasukkan ke dalam kategori “produk dengan penjualan lambat”. Pengelompokan ini menjadi dasar untuk merancang strategi persediaan barang, promosi, dan juga membantu dalam pengambilan keputusan pembelian dengan lebih terarah.

Untuk mempermudah penggunaan algoritma tersebut, perangkat lunak *RapidMiner* digunakan, yaitu platform analisis data yang menawarkan antarmuka visual sehingga pengguna yang tidak memiliki keterampilan pemrograman dapat membuat dan menjalankan model analisis data. *RapidMiner* mendukung berbagai metode pembelajaran mesin, termasuk *K-Means*, dan menyediakan fitur lengkap untuk persiapan data, validasi model, serta visualisasi hasil analisis (*RapidMiner Documentation*, 2023).

Berdasarkan latar belakang ini, penelitian ini bertujuan untuk menerapkan algoritma *K-Means Clustering* pada data penjualan Toko Sembako Isan dengan menggunakan *RapidMiner*. Melalui pendekatan ini, diharapkan toko dapat mendapatkan pemahaman yang lebih mendalam tentang karakteristik penjualannya dan meningkatkan efisiensi operasional dalam pengelolaan stok serta strategi pemasaran. Selain itu, penelitian ini juga berupaya mendorong penggunaan teknologi sederhana di kalangan pelaku UMKM agar dapat bersaing di era digital.

B. TEORI DAN METODELOGI

RapidMiner merupakan sebuah aplikasi perangkat lunak untuk ilmu data yang dibuat agar analisis data dan penambangan data bisa dilakukan tanpa memerlukan penulisan kode secara langsung. Aplikasi ini menyediakan antarmuka visual yang dapat digunakan dengan cara seret dan lepas, yang memungkinkan pengguna untuk menyusun alur kerja analisis dengan cara yang mudah dimengerti. Dengan kemampuan untuk mengintegrasikan data dari banyak sumber seperti *Excel*, *basis data*, *CSV*, dan lainnya, *RapidMiner* sangat sesuai untuk para analis data, ilmuwan, maupun perusahaan yang ingin memanfaatkan kecerdasan buatan meskipun tidak memiliki latar belakang di bidang pemrograman.

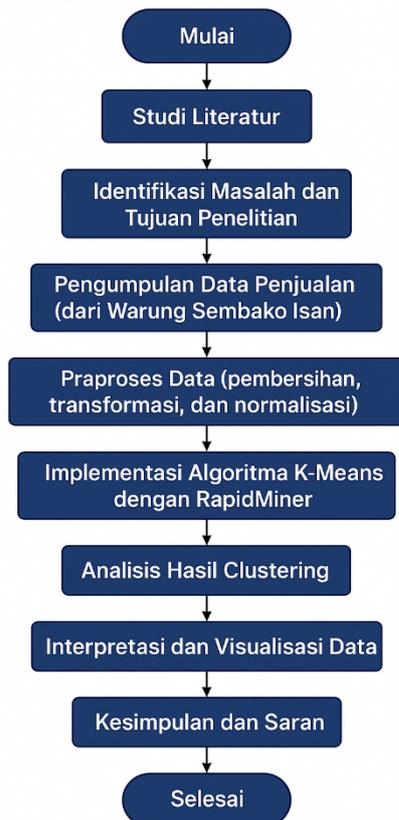
Clustering merupakan sebuah metode dalam data mining dan pembelajaran mesin yang berfungsi untuk mengelompokkan data ke dalam beberapa grup berdasarkan kesamaan atau karakteristik yang serupa. Tujuan utama dari clustering adalah agar data yang berada dalam satu grup memiliki tingkat kesamaan yang tinggi di antara mereka. Salah satu algoritma clustering yang sangat terkenal adalah *K-Means*, di mana pengguna terlebih dahulu menentukan jumlah kelompok yang diinginkan, kemudian algoritma akan berupaya untuk membagi data ke dalam kelompok tersebut dengan cara mengurangi jarak antara data dan pusat kelompok (*centroid*).

K-Means merupakan metode pembelajaran tanpa pengawasan yang digunakan untuk mengelompokkan data ke dalam beberapa kelompok berdasarkan kesamaan. Prosesnya dimulai dengan menetapkan jumlah kelompok (*K*), kemudian memilih *centroid* secara acak, yang diambil dari rata-rata setiap kelompok. Proses ini diulang sampai posisi *centroid* tidak berubah. *K-Means* sangat cocok untuk data dalam skala besar dan mudah diterapkan, namun memiliki kekurangan seperti harus menentukan jumlah kelompok sebelumnya dan kepekaan terhadap variasi data, pengelompokan dokumen, serta kompresi gambar.

Data penjualan yang dianalisis dalam studi ini berasal dari Warung Sembako Isan, yang dikelola oleh Ibu Warnisih. Pengumpulan data dilakukan antara bulan April hingga akhir Mei 2025. Dalam studi ini, peneliti menerapkan metode kuantitatif dengan menggunakan algoritma *K-Means*. Proses penelitian dilaksanakan secara sistematis

melalui langkah-langkah yang tertata, dimulai dari tahap pengumpulan data, pengolahan, hingga mencapai output dalam bentuk hasil pengelompokan.

Di dalam penelitian ini, langkah-langkah prosedural diilustrasikan melalui diagram alur yang menunjukkan perjalanan kegiatan penelitian dari awal sampai selesai. Diagram ini berfungsi untuk memudahkan pemahaman terhadap setiap tahapan yang dilakukan selama penelitian, sehingga lebih jelas.



Gambar 1 Alur Metodologi Penelitian

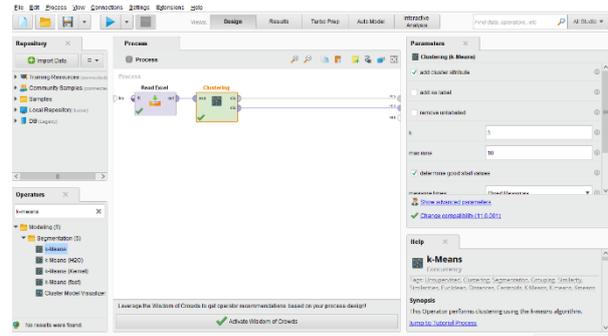
C. HASIL DAN PEMBAHASAN

Implementasi dari studi ini menggunakan metode perhitungan Algoritma *K-means* yang dilakukan dengan bantuan alat *RapidMiners*, yang dirancang untuk mempermudah pengguna dalam melakukan perhitungan melalui operator. Fungsi dari operator ini adalah untuk mengubah data, dan kemudian dihubungkan dengan berbagai node Operator. Sementara itu, peran pengguna adalah menyambungkan node tersebut agar dapat melihat hasil dalam bentuk tampilan grafik dari perhitungan.

1. Implementasi *K-means*

Setelah memasukkan dan memeriksa data, langkah selanjutnya adalah memodelkan dataset Warung Sembako Isan. Algoritma yang digunakan adalah *K-means* dengan

dukungan perangkat lunak bernama RapidMiner. Proses dimulai dengan memasukkan data ke dalam RapidMiner, kemudian data tersebut diproses menggunakan metode *K-means* dengan parameter $k = 3$ dan jenis pengukuran = campuran. Hal ini dilakukan karena dalam Warung Sembako Isan terdapat data yang berupa kode dan angka serta menggunakan Jarak Euclidean sebagai standar untuk mengukur jarak titik kluster.



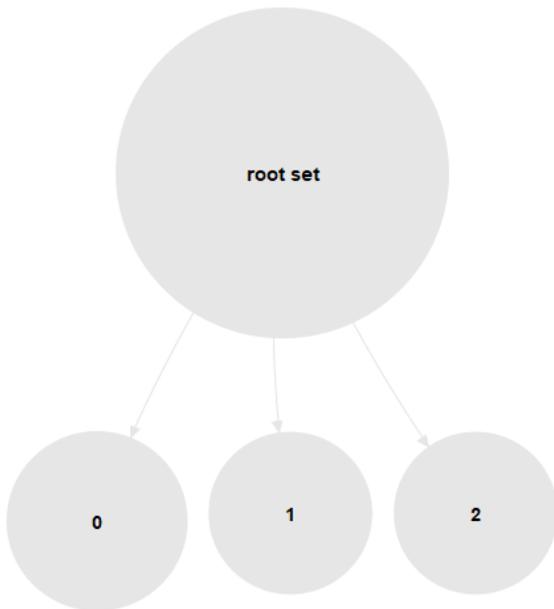
Gambar 2 Implementasi K-means

Selanjutnya, informasi dapat diproses menggunakan RapidMiner dan menghasilkan tiga model kluster yang terdiri dari kluster 0 yang memiliki 36 item, kluster 1 dengan 32 item, dan kluster 2 juga dengan 32 item. Penomoran dimulai dari 0 karena dalam bahasa pemrograman, angka 0 adalah angka pertama dalam urutan penomoran

Tabel 1 Tabel Cluster

Cluster 0	36 Items
Cluster 1	32 Items
Cluster 2	32 Items
Total	100 Items

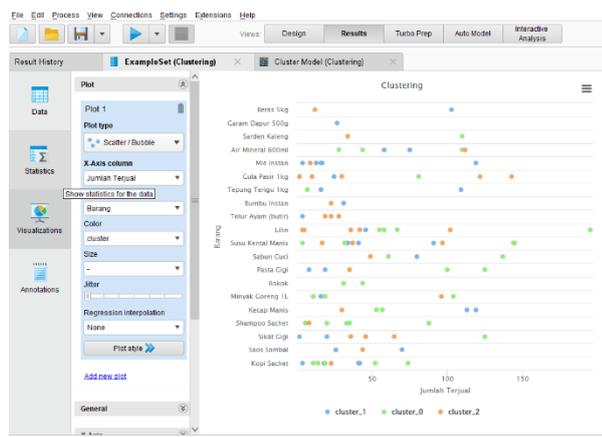
Adapun hasil yang berupa pohon dalam pilihan yang diberikan oleh *RapidMiners* memberikan gambaran mengenai seberapa besar perbedaan antara cluster dengan seluruh data. Dengan opsi ini, kita dapat melihat bagian dari data utama atau *root set*, yang terbagi menjadi beberapa cabang yaitu sebuah cluster.



Gambar 3 Hasil Diagram Tree(Pohon)

2. Visualizations Clustering

Visualisasi pengelompokan data penjualan menggunakan algoritma *k-means* melalui *rapidminer*, Warna biru terang melambangkan cluster 1 (Potensial) karena dihitung berdasarkan total penjualan, sementara warna hijau terang adalah cluster 0 (cukup potensial), dan warna oranye menunjukkan cluster 2 (kurang potensial).



Gambar 4 Visualisasi Clustering

3. Tabel Centroid

Dari tiga kelompok yang telah ditetapkan, pelanggan kemudian digolongkan berdasarkan hasil perhitungan *k means*. Hal ini menunjukkan bahwa kelompok 0 (produk yang sangat laku) memiliki jumlah sebesar 63. 861, kelompok 1 memperlihatkan (produk yang cukup laku) sebesar 43. 969, dan kelompok 2 menunjukkan (produk yang kurang laku) sebanyak 41. 906

Attribute	cluster_0	cluster_1	cluster_2
Jumlah Barang	109.333	93.938	106.062
Jumlah Terjual	63.861	43.969	41.906
Stok Barang	45.472	49.969	84.156
Harga Barang (Rp)	33624.139	12437	58802.531

Gambar 4 Tabel Centroid

D. PENUTUP

Kesimpulan

Penelitian ini berhasil menerapkan algoritma K-Means Clustering melalui RapidMiner pada data penjualan Warung Sembako Isan untuk mengelompokkan produk sesuai dengan potensi penjualannya. Hasil dari pengelompokan menunjukkan bahwa:

- Cluster 0 merepresentasikan produk yang sangat potensial dengan tingkat penjualan tertinggi (63. 861),
- Cluster 1 menggambarkan produk yang cukup potensial (43. 969),
- Cluster 2 menunjukkan produk yang kurang potensial (41. 906).

Dengan total sebanyak 100 produk yang dikelompokkan ke dalam tiga cluster, analisis ini memberikan pemahaman lebih mendalam tentang kinerja masing-masing produk. Visualisasi dari hasil clustering juga mempermudah pengertian dengan penggunaan warna yang mencerminkan setiap kategori produk.

Penerapan ini menunjukkan bahwa penggunaan teknologi sederhana seperti RapidMiner dapat mendukung pelaku UMKM untuk meningkatkan efisiensi dalam pengelolaan stok dan strategi penjualan. Di samping itu, pendekatan ini bisa menjadi langkah awal bagi pelaku usaha kecil untuk memanfaatkan teknologi data mining dalam pengambilan keputusan yang lebih akurat dan berbasis data.

Saran

Sebagai langkah selanjutnya, disarankan agar penelitian ini diperluas dengan cara membandingkan hasil dari algoritma K-Means dengan teknik clustering lainnya seperti DBSCAN atau Hierarchical Clustering. Selain itu, pengembangan sistem yang dapat memproses data secara otomatis dan dalam waktu nyata akan membuat analisis menjadi lebih praktis dan efektif. Penting juga untuk melakukan validasi terhadap hasil clustering dengan menggunakan metode evaluasi tambahan, sehingga

clustering data benar-benar mencerminkan kondisi sebenarnya di lapangan. Secara keseluruhan, studi ini menunjukkan bahwa penerapan metode data mining yang sederhana dapat memberikan bantuan bagi usaha kecil dalam memahami data penjualan serta membuat keputusan yang lebih tepat berdasarkan informasi yang terstruktur.

Ucapan Terima Kasih

Penulis mengucapkan terima kasih kepada semua pihak yang turut serta dalam proses pembuatan dan penyelesaian penelitian ini, terutama kepada platform Kaggle yang memberikan akses ke dataset terbuka untuk tujuan pembelajaran dan penelitian.

E. DAFTAR PUSTAKA

- Amin, F. A. (2023). Penerapan metode K-Means dalam penjualan produk Souq.com. *Applied Information System and Management (AISM)*, 5(1), 12–19.
- Andini, L. R., & Darmawan, B. (2023). Pengaruh visualisasi hasil clustering terhadap keputusan manajerial UMKM. *Jurnal Teknologi Bisnis*, 5(2), 66–74.
- Dahil, I. (2022). Penerapan data mining dalam menentukan potensi data penjualan tim telemarketing di Bank ABC dengan metode clustering dan algoritma K-Means menggunakan software RapidMiner. *ISMETEK: Jurnal Sistem Informasi, Manajemen dan Teknologi Informasi*, 13(2), 45–52.
- Hidayat, M. R., & Maulana, T. (2021). Perbandingan performa algoritma K-Means dan K-Medoids dalam segmentasi pelanggan. *Jurnal Ilmu Komputer Terapan*, 7(4), 88–96.
- Kartika, A. F. (2020). Clustering data penjualan dengan algoritma K-Means dan DBSCAN. *Jurnal Informatika dan Komputer*, 5(1), 13–20.
- Kurniawan, E. (2021). Evaluasi metode K-Means dan Hierarchical Clustering pada data penjualan ritel. *Jurnal Penelitian Komputer*, 10(1), 25–33.
- Kusumawardhani, R. (2020). Strategi berbasis data untuk pengembangan UMKM digital. *Jurnal Ekonomi dan Teknologi Informasi*, 6(2), 102–109.
- Nugroho, A. S., & Cahyadi, R. (2020). Penerapan K-Means untuk analisis pola pembelian pelanggan toko ritel. *Jurnal Teknologi Informasi dan Komputer*, 6(1), 41–50.
- Pratama, Y. D. (2022). Penggunaan RapidMiner dalam pengelompokan data pelanggan e-commerce. *Jurnal Sistem dan Teknologi Informasi*, 4(1), 44–52.
- Putri, R. M., & Santoso, A. (2021). Pengaruh clustering K-Means terhadap efektivitas strategi pemasaran. *Jurnal Sistem Informasi Bisnis*, 4(2), 55–62.
- Rahmawati, R. P. (2024). Optimasi stok dengan clustering data transaksi penjualan menggunakan algoritma K-Means di Konter Agung Cell. *Jurnal Informatika dan Teknik Elektro Terapan*, 4(2), 37–45.
- RapidMiner Documentation. (2023). RapidMiner official documentation. Retrieved from <https://docs.rapidminer.com>
- Setiawan, B., & Lestari, D. (2020). Pemanfaatan RapidMiner untuk segmentasi pelanggan dalam bisnis UMKM. *Jurnal Ekonomi Digital*, 2(3), 78–86.
- Sudrajat, W. C. (2022). Penerapan algoritma K-Means dalam segmentasi pelanggan pada toko sembako menggunakan RapidMiner. *Jurnal Ilmiah Binary*, 3(2), 59–66.
- Susanto, I., & Rahayu, S. (2022). Implementasi data mining dalam peningkatan layanan pelanggan pada UMKM. *Jurnal Ilmu Komputer dan Informatika*, 11(1), 67–74.
- Utami, N., & Wijaya, T. (2023). Penerapan teknik clustering untuk klasifikasi produk unggulan UMKM. *Jurnal Informatika UMKM*, 3(1), 15–22.
- Wahyuni, A. S. (2021). Penerapan data mining untuk menentukan strategi penjualan pada Toko Raja Komputer menggunakan metode clustering. *JURIKOM (Jurnal Riset Komputer)*, 8(2), 33–39.
- Widad, R., & Sahal, A. (2021). Penerapan algoritma K-Means clustering pada penjualan produk skincare. *Jurnal Mahasiswa Teknik Informatika Universitas Ibrahimy*, 5(1), 20–28.
- Yuliana, D. (2020). Analisis data pelanggan menggunakan teknik data mining berbasis clustering. *Jurnal Teknologi dan Sistem Informasi*, 8(2), 21–29.
- Zainuddin, M. F. (2021). Optimalisasi stok barang menggunakan algoritma K-Means. *Jurnal Teknik Informatika dan Sistem Informasi*, 9(3), 31–38.