

PENGELOMPOKAN JENIS SAMPAH MENGGUNAKAN ALGORITMA K-MEANS PADA BANK SAMPAH BUNGA RAYA

Rizky Ramadhan¹, Anita Diana², Yudi Wiharto^{3*}

^{1,2} Sistem Informasi, Fakultas Teknologi Informasi, Universitas Budi Luhur, Jakarta, Indonesia

³ Teknik Informatika, Fakultas Teknologi Informasi, Universitas Budi Luhur, Jakarta, Indonesia

Email: ¹kiki.adonk@gmail.com, ²anita.diana@budiluhur.ac.id, ^{3*}yudi.wiharto@budiluhur.ac.id

(* : corresponding author)

Abstrak- Penelitian ini bertujuan untuk menerapkan algoritma *K-Means Clustering* dalam menganalisis pola penyetoran sampah di Bank Sampah Bunga Raya RW 07, Perumahan Griya Sangiang Mas. Permasalahan utama yang diangkat adalah belum optimalnya pemanfaatan data historis penyetoran sampah dalam mendukung pengambilan keputusan operasional. Padahal, data tersebut memiliki potensi besar untuk dievaluasi dan dimanfaatkan dalam merancang strategi pengelolaan yang lebih efektif dan berbasis data. Data yang digunakan mencakup informasi mengenai jenis sampah, berat (kg), serta nilai ekonomis dari setiap transaksi penyetoran dalam periode Januari 2024 hingga Januari 2025. Metode analisis dalam penelitian ini mengacu pada pendekatan CRISP-DM yang terdiri dari tahapan pemahaman bisnis dan data, pembersihan dan transformasi data, normalisasi menggunakan metode *Min-Max Scaling*, serta proses klasterisasi dengan algoritma *K-Means*. Penentuan jumlah klaster optimal dilakukan menggunakan metode Elbow, kemudian divalidasi menggunakan nilai *Davies-Bouldin Index* (DBI). Hasil evaluasi menunjukkan bahwa DBI terbaik sebesar 0,537 diperoleh saat jumlah klaster adalah dua, yang menandakan kualitas pemisahan klaster yang cukup baik. Berdasarkan hasil klasterisasi, data penyetoran sampah terbagi ke dalam dua klaster utama. Klaster pertama (Cluster 0) mencakup jenis-jenis sampah yang disetor dalam jumlah besar, frekuensinya tinggi, serta memiliki kontribusi besar terhadap nilai total transaksi, seperti Bodong Kotor, Asoy, dan Botol Beling. Klaster kedua (Cluster 1) terdiri dari jenis-jenis sampah yang lebih spesifik, jumlahnya lebih kecil, namun memiliki nilai ekonomis per satuan yang tinggi, seperti Minyak Jelantah, UBC, dan Putihah. Hasil klasterisasi ini memberikan wawasan yang bermanfaat bagi pengelola Bank Sampah untuk menyusun strategi pengelolaan sampah yang lebih efisien, edukatif, dan berbasis data.

Kata Kunci: *Data Mining, K-Means Clustering, Sampah, Klasterisasi, Bank Sampah*

WASTE TYPE CLUSTERING USING K-MEANS ALGORITHM AT BUNGA RAYA WASTE BANK

Abstract- *This study aims to apply the K-Means clustering algorithm to analyze the pattern of waste deposits at Bunga Raya Waste Bank, RW 07, Griya Sangiang Mas Housing. The main problem addressed is the suboptimal utilization of historical waste deposit data to support operational decision-making. The data used in this study consists of information on waste types, weight (kg), and economic value for each deposit transaction from January 2024 to January 2025. The research methodology follows the CRISP-DM (Cross-Industry Standard Process for Data Mining) approach, including business and data understanding, data cleaning and transformation, normalization using Min-Max Scaling, and clustering using the K-Means algorithm. The optimal number of clusters was determined using the Elbow method and validated with the Davies-Bouldin Index (DBI). The evaluation results showed the best DBI value of 0.537 was obtained when the number of clusters was two, indicating a good separation quality. Based on the clustering results, the waste deposit data was grouped into two main clusters: Cluster 0 includes waste types deposited in large quantities with high frequency and significant contribution to the total transaction value (such as Bodong Kotor, Asoy, and Glass Bottles), while Cluster 1 contains more specific waste types deposited in smaller quantities but with a higher economic value per unit (such as Used Cooking Oil, UBC, and Putihah). These results provide valuable insights for waste bank managers to develop more efficient, educational, and data-driven waste management strategies.*

Keywords: *Data Mining, K-Means Clustering, Waste, Clustering, Waste Bank*

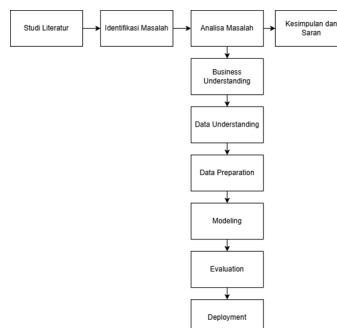
1. PENDAHULUAN

Dalam konteks *data mining*, *clustering* merupakan teknik analisis data tanpa label (*unsupervised learning*) yang digunakan untuk mengelompokkan data berdasarkan kesamaan karakteristik. Salah satu algoritma yang populer adalah *K-Means Clustering*. Metode ini dapat membantu mengidentifikasi pola penyetoran sampah

berdasarkan karakteristik tertentu, sehingga strategi pengelolaan yang diterapkan dapat lebih efektif dan sesuai dengan kebutuhan lapangan [1]. Dengan dukungan metode *Elbow* untuk menentukan jumlah kluster optimal serta *Davies-Bouldin Index (DBI)* untuk mengevaluasi kualitas hasil klusterisasi. Selain itu, metode *Min-Max Scaling* digunakan agar setiap atribut berada pada skala yang sama sehingga tidak mendominasi proses perhitungan. Beberapa penelitian sebelumnya telah menerapkan teknik clustering pada data pengelolaan sampah. Misalnya, Salsabila dkk. (2024) mengelompokkan kecamatan berdasarkan volume sampah di Karawang, sedangkan Nurrohman dkk. (2023) fokus pada klasifikasi volume sampah organik di Kota Magelang. Penelitian lain oleh Suryanti dkk. (2022) mengembangkan sistem informasi pemetaan volume sampah di Surakarta. Namun, sebagian besar penelitian masih menggunakan data agregat (wilayah/kota), bukan data transaksi riil di tingkat bank sampah. Hal ini menunjukkan adanya gap penelitian yang coba diisi melalui penelitian ini, yaitu dengan memanfaatkan data historis transaksi di Bank Sampah Bunga Raya secara mandiri. Meski kontribusinya positif, tantangan tetap ada, terutama dalam pengelolaan volume sampah yang terus meningkat dan belum diklasifikasikan berdasarkan jenis secara optimal [2]. Selain itu setiap daerah dapat dikelompokkan berdasarkan karakteristik pengelolaan sampahnya, sehingga dapat membantu dalam perumusan kebijakan yang lebih tepat dan berbasis data [3]. Dengan penerapan teknik data mining, evaluasi pola yang ditemukan, serta penyajian informasi akhir yang dapat dimanfaatkan dalam pengambilan keputusan bertujuan untuk membantu pengambilan keputusan yang lebih efisien dan berbasis data [4].

2. METODE PENELITIAN

Penelitian ini menggunakan pendekatan *CRISP-DM (Cross Industry Standard Process for Data Mining)* yang terdiri dari enam tahap. Tahap pertama adalah pemahaman bisnis, yaitu mengidentifikasi kebutuhan Bank Sampah Bunga Raya dalam mengelompokkan jenis sampah untuk mendukung strategi pengelolaan yang efisien. Tahap kedua adalah pemahaman data, dengan menggunakan *dataset* transaksi penyetoran sampah periode Januari 2024 hingga Januari 2025 yang berjumlah 1.166 entri. *Dataset* ini memuat atribut seperti tanggal setor, nama penyetor, nomor rekening, jenis sampah, berat, harga satuan, dan nilai ekonomis. Tahap ketiga adalah persiapan data. Pada tahap ini dilakukan pembersihan data dari duplikasi, pengecekan nilai kosong, serta pemberian kode numerik pada variabel kategorik menggunakan teknik label encoding. Normalisasi atribut numerik seperti berat dan nilai transaksi dilakukan dengan metode *Min-Max Scaling* agar setiap atribut berada pada skala yang sama. Proses persiapan data dilakukan dengan bantuan Microsoft Excel untuk pengolahan awal *dataset* dan *RapidMiner* untuk tahapan *preprocessing*. Tahap keempat adalah pemodelan, yaitu penerapan algoritma K-Means untuk melakukan proses klusterisasi. Penentuan jumlah kluster optimal ditentukan menggunakan metode *Elbow* dengan parameter *Sum of Squared Errors (SSE)*. Pada tahap ini digunakan *Google Colab* dengan bahasa pemrograman *Python* untuk menjalankan algoritma dan membandingkan hasil pemodelan dengan *RapidMiner*. Tahap kelima adalah evaluasi, di mana kualitas hasil kluster diukur menggunakan *Davies-Bouldin Index (DBI)*. Nilai *DBI* yang rendah menunjukkan kualitas kluster yang baik. Evaluasi dilakukan baik melalui *RapidMiner* maupun *Google Colab* untuk memastikan konsistensi hasil. Tahap terakhir adalah *deployment*, yaitu penerapan hasil analisis ke dalam rekomendasi strategi pengelolaan Bank Sampah. Hasil klusterisasi dimanfaatkan untuk memberikan saran terkait penataan ruang penyimpanan, penjadwalan pengangkutan, serta prioritas program edukasi kepada masyarakat. *Cross Industry Standard Process for Data Mining (CRISP-DM)* Metode penelitian ini menggunakan pendekatan *Cross Industry Standard Process for Data Mining (CRISP-DM)* yang terdiri dari enam tahapan utama [5]. Penggunaan metode *CRISP-DM* guna menjamin setiap tahap analisis data dilakukan secara terorganisir dan sistematis [6]



Gambar 1. Tahapan Penelitian

2.1. Business Understanding

Mengidentifikasi tujuan penelitian, yaitu mengelompokkan jenis sampah di Bank Sampah Bunga Raya untuk mendukung strategi pengelolaan yang lebih efisien.

2.2. Data Understanding

Data yang digunakan adalah catatan penyetoran sampah periode Januari 2024–Januari 2025 yang berjumlah 1.166 transaksi. Atribut yang dianalisis meliputi nama sampah, berat (kg), harga, dan nilai ekonomis.

2.3. Data Preparation

Tahapan ini mencakup pembersihan data dari duplikasi, pengecekan nilai kosong (tidak ditemukan atau *missing value*), pemberian kode numerik pada variabel kategorik (label encoding untuk kolom NAMA_SAMPAH), serta normalisasi atribut numerik menggunakan metode *Min-Max Scaling* pada kolom JUMLAH(KG) dan NILAI.

2.4. Modeling

Penerapan algoritma *K-Means* untuk mengelompokkan jenis sampah. Penentuan jumlah kluster optimal dilakukan menggunakan Elbow Method dengan parameter *Sum of Squared Errors(SSE)*.

2.5. Evaluation

Kualitas hasil kluster diukur menggunakan *Davies-Bouldin Index(DBI)*, di mana nilai terendah menunjukkan pemisahan kluster yang lebih baik.

2.6. Deployment

Hasil analisis digunakan sebagai dasar rekomendasi strategi pengelolaan, termasuk penataan ruang simpan, penjadwalan pengangkutan, dan prioritas program edukasi kepada masyarakat.

2.7. K-Means

Algoritma *K-Means* merupakan salah satu metode *clustering* non-hierarki yang banyak digunakan dalam analisis data. Salah satu keunggulan utama dari algoritma ini adalah fleksibilitasnya, di mana pengguna dapat menentukan sendiri jumlah kluster (K) yang diinginkan sesuai dengan kebutuhan analisis. *K-Means* bekerja dengan mempartisi data ke dalam sejumlah kelompok, di mana data yang memiliki karakteristik serupa akan ditempatkan dalam satu kluster yang sama, sedangkan data yang memiliki perbedaan karakteristik akan dikelompokkan ke dalam kluster yang berbeda. Proses ini memungkinkan terbentuknya kelompok-kelompok data yang homogen di dalam kluster dan heterogen antar kluster, dengan tingkat variasi internal yang relatif kecil [7]. Setiap titik data kemudian dikelompokkan ke dalam kluster yang memiliki *centroid* terdekat berdasarkan jarak *euclidean* atau metrik tertentu. Selanjutnya, *centroid* tiap kluster diperbarui dengan menghitung rata-rata dari seluruh titik data yang tergabung dalam kluster tersebut. Langkah ini terus diulang secara iteratif hingga posisi *centroid* stabil atau tidak terjadi lagi perubahan signifikan dalam pengelompokan data, atau hingga jumlah iterasi maksimum tercapai[8] Selanjutnya dengan metode Elbow menetapkan nilai *cluster* yang berbentuk siku [9]. Selain itu berbeda dengan klasifikasi, klusterisasi termasuk dalam pendekatan unsupervised learning karena proses pengelompokannya dilakukan tanpa label awal [10]. Langkah-langkah melakukan *clustering* dengan metode *K-Means* adalah sebagai berikut:

- Tentukan nilai k sebagai jumlah kluster yang ingin dibentuk.
- Inisialisasi k pusat *cluster* ini bisa dilakukan dengan berbagai cara, namun yang paling sering dilakukan adalah dengan cara random yang diambil dari data yang ada.
- Menghitung jarak setiap data input terhadap masing – masing *centroid* menggunakan rumus jarak *Euclidean* (*Euclidean Distance*) hingga ditemukan jarak yang paling dekat dari setiap data dengan centroid. Berikut adalah persamaan *Euclidian Distance* :

$$De = \sqrt{\sum (x_i - u_j)^2} \quad (1)$$

Keterangan:

De adalah *Euclidean Distance*

x_i : data kriteria.

u_j : centroid pada cluster ke- j .

- d. Mengklasifikasikan setiap data berdasarkan kedekatannya dengan centroid (jarak terkecil).

$$C_k = \frac{1}{n_k} \sum D_i \quad (2)$$

Keterangan:

n_k : Jumlah data dalam cluster k .

D_i : Jumlah nilai jarak yang masuk dalam masing-masing *cluster*.

- e. Memperbaharui nilai centroid. *Nilai Centroid* baru di peroleh dari rata-rata cluster yang bersangkutan.

3. HASIL DAN PEMBAHASAN

3.1. Business Understanding

Pada tahap ini, pemahaman bisnis difokuskan pada kebutuhan Bank Sampah Bunga Raya dalam mengelola beragam jenis sampah yang disetorkan masyarakat. Selama ini, data transaksi penyetoran sampah hanya digunakan untuk pencatatan administrasi, belum dimanfaatkan lebih lanjut untuk mendukung pengambilan keputusan strategis. Kondisi tersebut menimbulkan beberapa kebutuhan utama, yaitu:

- Perlunya sistem pengelompokan jenis sampah berdasarkan karakteristik seperti volume setoran dan nilai ekonomis.
- Perlunya identifikasi pola setoran agar kapasitas penyimpanan dapat diatur secara lebih efisien.
- Perlunya dasar analisis berbasis data untuk mendukung strategi operasional, penjadwalan pengangkutan, dan edukasi masyarakat.

Dengan demikian, klusterisasi data jenis sampah menjadi sangat penting agar pengelolaan bank sampah dapat dilakukan secara lebih terstruktur, tepat sasaran, dan berkelanjutan.

3.2. Data Understanding

Penelitian ini menggunakan data publik mengenai aktivitas penyetoran sampah yang dilakukan oleh nasabah Bank Sampah Bunga Raya di Kota Tangerang. Data diperoleh dari dokumen internal yang telah dikumpulkan dalam bentuk digital dan memuat total 1.166 baris data. Rentang waktu data mencakup periode Januari 2024 hingga Januari 2025. *Dataset* ini berisi informasi harian mengenai jenis sampah yang disetorkan, nama penyetor (biasanya institusi seperti sekolah), serta berat dan nilai ekonomis dari masing-masing jenis sampah. Pengelolaan Bank Sampah berbasis partisipasi masyarakat, namun belum menerapkan metode pengelompokan atau analisis berbasis data [11]. Informasi ini berguna untuk melakukan analisis pola penyetoran sampah berdasarkan jenis sampah Adapun atribut-atribut dalam *dataset* tersebut adalah sebagai berikut:

- TANGGAL SETOR: Tanggal saat transaksi penyetoran sampah dilakukan
- NAMA: Nama institusi atau individu penyetor sampah.
- NO REK: Nomor rekening nasabah pada sistem Bank Sampah.
- NAMA SAMPAH: Jenis sampah yang disetorkan, seperti kardus, botol beling, kertas, dan lain-lain.
- JUMLAH(KG): Berat sampah yang disetorkan dalam satuan kilogram.
- HARGA (BS INDUK): Harga satuan sampah per kilogram yang ditetapkan oleh Bank Sampah Induk.
- NILAI: Total nilai ekonomis dari hasil penyetoran (hasil perkalian antara berat dan harga satuan).

3.3. Data Preparation

Pada tahap ini dilakukan proses pembentukan *dataset* sebagai dasar analisis. *Dataset* berasal dari dokumen pencatatan transaksi penyetoran sampah di Bank Sampah Bunga Raya periode Januari 2024 hingga Januari 2025 yang telah dikompilasi dalam bentuk digital. Data awal diolah menggunakan *Microsoft Excel* untuk membersihkan duplikasi, menghapus nilai kosong (missing value), serta menyatukan atribut yang relevan, yaitu TANGGAL SETOR, NAMA, NO REK, NAMA SAMPAH, JUMLAH (KG), HARGA (BS INDUK), dan NILAI.

Selanjutnya, dilakukan proses verifikasi agar setiap transaksi tercatat dengan benar sesuai identitas penyetor, jenis sampah, berat setoran, harga satuan, dan nilai total transaksi. *RapidMiner* digunakan untuk membantu proses *preprocessing* lanjutan, seperti integrasi atribut dan pengecekan konsistensi data. Hasil akhir dari tahap ini adalah *dataset* yang sudah bersih dan siap digunakan untuk analisis klusterisasi dengan algoritma *K-Means*.

Contoh susunan *dataset* hasil persiapan data ditunjukkan pada Tabel 1, yang merupakan sampel data transaksi riil Bank Sampah Bunga Raya setelah melalui tahap pembersihan dan penyatuan atribut.

3.4. Dataset Asli

Tabel 1. *Dataset Asli (sample data).*

TANGGAL SETOR	NAMA	NO REK	NAMA SAMPAH	JUMLAH (KG)	HARGA (BS INDUK)	NILAI
1/6/2024	SD Penerus Bangsa	096	Putihan Kertas	24.9	1500	37350
1/6/2024	SD Penerus Bangsa	096	Bodong Kotor	13.8	2300	31740
1/6/2024	SD Penerus Bangsa	096	Emberan	2.5	1500	3750
1/6/2024	SD Penerus Bangsa	096	Kardus	3.2	1800	5760
1/7/2024	Nur Latif	090	Kardus	24.3	1800	43740

3.5. Dataset Transformasi

Tahap transformasi data dilakukan untuk menyiapkan dataset agar dapat digunakan dalam proses klusterisasi. Salah satu langkah utama adalah menambahkan atribut ID sebagai identitas unik pada setiap *record* atau baris data. Pemberian ID ini dilakukan secara *auto-incremental*, dimulai dari angka 1 hingga jumlah total baris data, sehingga memudahkan proses pelacakan dan pengolahan. Proses transformasi dilakukan dengan bantuan *Microsoft Excel* untuk penambahan kolom ID secara otomatis, serta *RapidMiner* untuk memastikan struktur data sudah konsisten. Selain itu, *Google Colab* juga digunakan untuk mengecek kembali hasil transformasi melalui skrip *Python* sederhana, sehingga data siap digunakan pada tahap normalisasi. Hasil dari tahap ini berupa *dataset* yang lebih terstruktur dengan adanya ID pada setiap baris data. Contoh hasil transformasi ditunjukkan pada Tabel 2, yang diperoleh dari data transaksi penyetoran sampah Bank Sampah Bunga Raya setelah melalui proses penambahan atribut ID.

Tabel 2. *Dataset Transformasi*

ID	TANGGAL SETOR	NAMA	NO REK	NAMA SAMPAH	JUMLAH (KG)	HARGA (BS INDUK)	NILAI
1	1/6/2024	SD Penerus Bangsa	096	Putihan Kertas	1500	37350	37350
2	1/6/2024	SD Penerus Bangsa	096	Bodong Kotor	2300	31740	31740
3	1/6/2024	SD Penerus Bangsa	096	Emberan	1500	3750	3750
4	1/6/2024	SD Penerus Bangsa	096	Kardus	1800	5760	5760
5	1/6/2024	SD Penerus Bangsa	096	Botol Beling	1800	43740	840

3.6. Dataset Normalisasi

Dalam *data mining*, normalisasi merupakan proses penskalaan nilai atribut numerik agar berada pada rentang tertentu, umumnya 0 hingga 1. Tujuan utamanya adalah menyetarakan skala antar atribut sehingga tidak ada satu atribut yang mendominasi proses perhitungan hanya karena memiliki nilai jauh lebih besar dibanding atribut lain. Proses normalisasi pada penelitian ini dilakukan menggunakan metode *Min-Max Scaling* terhadap atribut JUMLAH (KG) dan NILAI transaksi. *Microsoft Excel* digunakan untuk perhitungan awal, kemudian *RapidMiner* dan *Google Colab (Python)* digunakan untuk memastikan proses normalisasi berjalan konsisten. Hasilnya, setiap atribut numerik memiliki nilai yang telah ditransformasi ke dalam rentang 0–1, sehingga siap digunakan dalam tahap pemodelan dengan algoritma *K-Means*. Tahap pemodelan dilakukan dengan menerapkan algoritma *K-Means Clustering*. Penentuan jumlah kluster optimal dilakukan menggunakan metode *Elbow*, sedangkan evaluasi kualitas kluster menggunakan *Davies-Bouldin Index (DBI)*. Hasil pengolahan menunjukkan nilai DBI terbaik sebesar 0,537 ketika jumlah kluster adalah dua, yang menandakan kualitas pemisahan kluster cukup baik. Berdasarkan hasil klusterisasi, diperoleh dua kelompok utama yaitu Kluster 0 terdiri dari jenis sampah dengan volume setoran tinggi, frekuensi besar, dan kontribusi signifikan terhadap nilai transaksi, seperti kardus, botol beling, dan besi. Kluster 1 mencakup jenis sampah dengan volume lebih kecil namun memiliki nilai ekonomis tinggi per satuan, seperti minyak jelantah, UBC, dan putihan. Hasil ini menunjukkan bahwa terdapat perbedaan pola penyetoran antara sampah bernilai tinggi secara frekuensi dengan sampah bernilai tinggi secara ekonomi. Temuan penelitian ini memberikan wawasan yang dapat dimanfaatkan pengelola Bank Sampah Bunga Raya. Beberapa rekomendasi yang dapat diterapkan antara lain:

- Menyusun strategi penyimpanan yang lebih efisien dengan memprioritaskan jenis sampah dari kluster ber-

- volume tinggi.
- Menentukan jadwal pengangkutan lebih teratur untuk menghindari penumpukan klaster volume besar.
 - Mengoptimalkan program edukasi kepada masyarakat dengan mendorong peningkatan setoran pada jenis sampah bernilai ekonomis tinggi.
 - Menggunakan hasil klasterisasi sebagai dasar untuk perencanaan jangka panjang dalam pengelolaan bank sampah berbasis data.

Tabel 3 menampilkan hasil normalisasi dari data transaksi penyetoran sampah Bank Sampah Bunga Raya. Data pada tabel ini diperoleh dari hasil *preprocessing* dengan *RapidMiner* yang kemudian divalidasi ulang menggunakan *Google Colab*.

Tabel 3. Dataset Normalisasi

Jumlah (Kg)	Harga (Bs Induk)	Nilai	Nama_Sampah_Encoded
0.336499	0.009634	0.337249	0.852941
0.185889	0.0158	0.28656	0.176471
0.032564	0.009634	0.033657	0.441176
0.042062	0.011946	0.051818	0.647059
0.036635	0.000385	0.007364	0.235294
0.328358	0.011946	0.394985	0.647059
0.592944	0.009634	0.593404	0.441176

Hasil Komparasi model Penelitian ini membandingkan algoritma *K-Means* dan *K-Medoids* untuk menentukan metode klasterisasi terbaik dalam mengelompokkan data penyetoran sampah. Evaluasi dilakukan menggunakan nilai *Davies-Bouldin Index (DBI)*, di mana semakin kecil nilai DBI menunjukkan kualitas klaster yang lebih baik.

Tabel 4. Komparasi *K-means*

Cluster	Nilai DBI
Cluster 1	0.537
Cluster 2	0.731
Cluster 3	0.676

Berdasarkan Tabel 4 algoritma *K-Means* dengan jumlah 2 *cluster* merupakan *cluster* terbaik karena paling mendekati nilai 0 sebesar 0,537. Berdasarkan nilai DBI pada algoritma *K-Medoids*, jumlah *cluster* yang paling baik adalah 2 *cluster* sebesar 0,596 dibandingkan nilai DBI pada 3 *cluster* dan 4 *cluster*. Dari hasil perbandingan kedua algoritma, *K-Means* dengan 2 *cluster* menghasilkan nilai DBI paling rendah (0.537) dibandingkan algoritma *K-Medoids*. Oleh karena itu, algoritma *K-Means* dipilih sebagai metode klasterisasi terbaik untuk penelitian ini, dengan jumlah *cluster* optimal sebanyak 2 *cluster*.

Tabel 5 Komparasi *Kmedoids*

Cluster	Nilai DBI
Cluster 2	0.596
Cluster 3	0.589
Cluster 4	0.630

4. KESIMPULAN

Penelitian ini berhasil menerapkan algoritma *K-Means Clustering* untuk mengelompokkan data penyetoran sampah di Bank Sampah Bunga Raya. Hasil klasterisasi menunjukkan adanya dua kelompok utama, yaitu sampah dengan volume tinggi dan sampah dengan nilai ekonomis tinggi per satuan. Temuan ini dapat membantu pengelola dalam merancang strategi penataan ruang, penjadwalan pengangkutan, serta program edukasi kepada masyarakat. Evaluasi model menggunakan *Davies-Bouldin Index* menghasilkan nilai 0,537, yang menandakan kualitas klaster cukup baik. Penelitian selanjutnya dapat mengembangkan sistem informasi berbasis data mining agar hasil analisis dapat diimplementasikan secara langsung dalam operasional bank sam Penelitian ini berhasil menerapkan algoritma *K-Means Clustering* dalam menganalisis data penyetoran sampah di Bank Sampah Bunga Raya periode Januari 2024–Januari 2025. Hasil klasterisasi menunjukkan terbentuknya dua kelompok utama. Klaster pertama terdiri dari jenis sampah dengan volume setoran tinggi dan frekuensi besar, seperti kardus, botol beling, dan besi,

yang memberikan kontribusi signifikan terhadap total volume transaksi. Klaster kedua terdiri dari jenis sampah dengan volume relatif kecil tetapi memiliki nilai ekonomis tinggi per satuan, seperti minyak jelantah, UBC, dan putihan. Evaluasi model menggunakan Davies-Bouldin Index menghasilkan nilai 0,537 pada jumlah klaster dua, yang menandakan pemisahan klaster cukup baik. Temuan ini memberikan gambaran nyata mengenai pola penyetoran sampah dan dapat dijadikan dasar strategi pengelolaan berbasis data. Berdasarkan hasil penelitian, beberapa rekomendasi bagi pengelola Bank Sampah antara lain:

- a. Mengoptimalkan penataan ruang penyimpanan dengan memperhatikan jenis sampah ber-volume tinggi.
- b. Menyusun jadwal pengangkutan lebih teratur untuk mengurangi risiko penumpukan.
- c. Mengembangkan program edukasi kepada masyarakat untuk meningkatkan setoran jenis sampah bernilai ekonomis tinggi
- d. Memanfaatkan hasil klasterisasi sebagai acuan dalam perencanaan jangka panjang yang lebih efisien dan berkelanjutan.

Untuk penelitian selanjutnya, disarankan pengembangan sistem informasi berbasis data mining sehingga hasil analisis dapat diintegrasikan langsung ke dalam operasional bank sampah, serta penggunaan metode *clustering* lain sebagai perbandingan untuk meningkatkan kualitas model.

UCAPAN TERIMA KASIH

Penulis mengucapkan syukur ke hadirat Allah SWT atas segala rahmat dan karunia-Nya sehingga penelitian ini dapat terselesaikan dengan baik. Ucapan terima kasih yang sebesar-besarnya penulis sampaikan kepada kedua orang tua dan keluarga atas doa, dukungan, dan motivasi yang tiada henti. Tidak lupa penulis berterima kasih kepada pengelola dan nasabah Bank Sampah Bunga Raya yang telah memberikan data dan informasi sehingga penelitian ini dapat terlaksana. Terima kasih juga kepada rekan-rekan yang senantiasa memberikan dukungan, semangat, dan kebersamaan selama proses penelitian ini berlangsung. Semoga segala bantuan dan dukungan yang diberikan mendapatkan balasan yang berlipat ganda dari Allah SWT.

DAFTAR PUSTAKA

- [1] M. Nurrohman, M. Maimunah, and P. Sukmasetya, "Sistem Klasterisasi Volume Sampah Organik di Kota Magelang menggunakan K-Means," *TEMATIK*, vol. 10, no. 1, pp. 146–153, Jun. 2023, doi: 10.38204/tematik.v10i1.1338.
- [2] F. Salsabila, T. Ridwan, and H. H., "ANALISA VOLUME PENYEBARAN SAMPAH DI KARAWANG MENGGUNAKAN ALGORITMA K-MEANS CLUSTERING," *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 12, no. 2, Apr. 2024, doi: 10.23960/jitet.v12i2.4226.
- [3] A. Sulistiyawati and E. Supriyanto, "Implementasi Algoritma K-means Clustering dalam Penentuan Siswa Kelas Unggulan," vol. 15, no. 2.
- [4] Dodi Nofri Yoliadi, "Data mining Dalam Analisis Tingkat Penjualan Barang Elektronik Menggunakan Algoritma K-means," 2023.
- [5] S. Isnanto and S. Widodo, "PENERAPAN DATA MINING PADA PENERIMAAN MAHASISWA BARU DENGAN ALGORITMA K-MEANS CLUSTERING," *Jurnal Teknik Informasi dan Komputer (Tekinkom)*, vol. 4, no. 2, p. 158, Dec. 2021, doi: 10.37600/tekinkom.v4i2.367.
- [6] W. Usino, "KLASTERISASI TINGKAT KELAYAKAN PROVINSI DALAM PEMBANGUNAN KAWASAN INDUSTRI MENGGUNAKAN ALGORITMA K-MEANS," 2024.
- [7] A. Sulistiyawati and E. Supriyanto, "Implementasi Algoritma K-means Clustering dalam Penentuan Siswa Kelas Unggulan," *J. Tekno Kompak*, vol. 15, no. 2, p. 25, 2021.
- [8] D. Syaripudin, A. Rachmat Raharja, S. Informasi, F. Kesehatan dan Teknik, and U. Bandung, "PENERAPAN ALGORITMA K-MEANS CLUSTERING PADA DATA NILAI SISWA UNTUK MENENTUKAN KELOMPOK PENERIMA BEASISWA".
- [9] R. Yuliana Sari, H. Oktavianto, and H. Wahyu Sulisty, "ALGORITMA K-MEANS DENGAN METODE ELBOW UNTUK MENGELOMPOKKAN KABUPATEN/KOTA DI JAWA TENGAH BERDASARKAN KOMPONEN PEMBENTUK INDEKS PEMBANGUNAN MANUSIA K-MEANS ALGORITHM WITH ELBOW METHOD TO GROUPING DISTRICT/CITY IN CENTRAL JAVA BASED ON COMPONENTS OF HUMAN DEVELOPMENT INDEX," 2022. [Online]. Available: <http://jurnal.unmuhjember.ac.id/index.php/JST>

- [10] S. K. Dirjen *et al.*, “Terakreditasi SINTA Peringkat 4 Klasterisasi Karakter Konsumen Terhadap Kecenderungan Pemilihan Produk Menggunakan K-Means,” 2018.
- [11] O. A. Halid, K. Yulianto, and M. Saleh, “STRATEGI PENGELOLAAN BANK SAMPAH di NTB (Studi Kasus Bank Sampah Bintang Sejahtera),” *Januari*, no. 8, p. 763, 2022.