



Sistem Klusterisasi Volume Sampah Organik di Kota Magelang menggunakan K-Means

Clustering System of Organic Waste Volume in Magelang City using K-Means

Muhamad Nurrohman¹, Maimunah², Pristi Sukmasetya³

^{1,2,3}Prodi Teknik Informatika, Fakultas Teknik, Universitas Muhammadiyah Magelang
¹mnurrohman2020@gmail.com, ²maimunah@unimma.ac.id*, ³pristi.sukmasetya@ummgl.ac.id

Abstract

Waste is one of the most serious problems faced by many regions. In Magelang City alone, 70 tons of waste is generated per day, and 60% of it is organic waste. Business sectors such as restaurants, culinary tours, markets, and hotels are high producers of organic waste. Organic waste can be utilized and sorted to be useful for maggot farmers, compost and biogas makers, but the lack of information about the source and availability of this waste is an obstacle for organic waste users and the Environmental Service. The purpose of this research is to assist the Environmental Agency in planning waste transportation, as well as making it easier for organic waste users to get information about the waste. In this study, the data used is waste volume data and data on the number of visitors and traders in hotels, markets, and street vendors in Magelang City, the data is then processed using the K-Means clustering algorithm. The data that has been processed produces the optimal number of clusters is 2, cluster 1 is a low waste volume producing category, while cluster 2 is a high waste volume producing category. After obtaining the clustering results using the K-Means algorithm, an evaluation of the results was carried out using the silhouette score method which resulted in a score value of 0.66, from the evaluation results it can be concluded that the application of the K-Means algorithm in clustering the volume of organic waste in Magelang City is quite good. With these results, it is hoped that it can help the Magelang City government, especially the Magelang City Environmental Agency, maggot cultivators, and other organic waste users to more easily obtain information about the availability of organic waste, which is expected to help reduce the volume of organic waste.

Keywords: organic waste; k-means; clustering; data mining

Abstrak

Permasalahan sampah merupakan salah satu masalah serius yang dihadapi berbagai daerah. Di Kota Magelang sendiri sampah yang dihasilkan per-hari mencapai 70 ton, dan 60% di antara itu merupakan sampah organik. Sektor usaha seperti restoran, wisata kuliner, pasar, dan hotel merupakan penghasil sampah organik dengan kapasitas tinggi. Sampah organik dapat dimanfaatkan dan dipilah agar menjadi bermanfaat bagi peternak maggot, pembuat pupuk kompos dan biogas, akan tetapi kurangnya informasi mengenai sumber dan ketersediaan sampah tersebut menjadi kendala tersendiri bagi pemanfaat sampah organik, serta Dinas Lingkungan Hidup. Tujuan dari penelitian ini adalah untuk membantu Dinas Lingkungan Hidup dalam perencanaan pengangkutan sampah, serta memudahkan pemanfaat sampah organik untuk mendapatkan informasi mengenai sampah tersebut. Dalam penelitian ini, data yang digunakan adalah data volume sampah dan data jumlah pengunjung serta pedagang di Hotel, Pasar, maupun Pedagang Kaki Lima yang ada di Kota Magelang, data kemudian diolah menggunakan algoritma klusterisasi K-Means. Data yang sudah diolah menghasilkan jumlah kluster optimal sebanyak 2, kluster 1 adalah kategori penghasil volume sampah rendah, sedangkan kluster 2 adalah kategori penghasil volume sampah tinggi. Setelah didapatkan hasil klusterisasi menggunakan algoritma K-Means, dilakukan evaluasi terhadap hasil tersebut menggunakan metode silhouette score yang menghasilkan nilai score sebesar 0.66, dari hasil evaluasi dapat disimpulkan bahwa penerapan algoritma K-Means dalam klusterisasi volume sampah organik di Kota Magelang sudah cukup baik. Dengan hasil tersebut diharapkan dapat membantu pemerintah Kota Magelang, khususnya Dinas Lingkungan Hidup Kota Magelang, pembudidaya maggot, serta pemanfaat sampah organik lainnya supaya lebih mudah mendapatkan informasi mengenai sumber dan ketersediaan sampah organik, yang diharapkan dapat membantu mengurangi volume sampah organik.

Kata kunci: sampah organik; k-means; clustering; data mining

1. Pendahuluan

Permasalahan sampah menjadi salah satu masalah serius yang dihadapi di berbagai daerah. Salah satu penyebabnya adalah kurangnya kesadaran masyarakat mengenai pentingnya mengelola sampah dengan baik, sehingga lingkungan dapat terjaga kebersihannya dan tetap sehat. Banyak masyarakat masih beranggapan bahwa menjaga lingkungan yang bersih, sehat, dan nyaman merupakan tanggung jawab dari pemerintah saja, padahal masyarakat seharusnya juga ikut aktif dalam menjaga dan mengelola sampah supaya tercipta lingkungan yang sehat. Lingkungan adalah tempat hidup semua makhluk yang ada di bumi, khususnya manusia.

Kota Magelang merupakan salah satu kota yang ada di provinsi Jawa Tengah dengan luas 18,12 km² yang terdiri dari 3 kecamatan, 17 kelurahan, dan jumlah penduduk yang terus mengalami peningkatan setiap tahunnya, sehingga pada tahun 2020 jumlah penduduknya mencapai 122.539 jiwa [1]. Dengan bertambahnya jumlah penduduk maka jumlah sampah yang dihasilkan juga akan semakin bertambah. Rata-rata setiap penduduk menghasilkan 0,7 kg sampah di setiap hari, sehingga perlu dilakukan penanganan khusus untuk mengatasi masalah sampah di Kota Magelang.

Di Kota Magelang sendiri, sampah yang dapat dihasilkan per-harinya sebanyak 70 ton dan 60% di antara itu merupakan sampah organik. Sampah organik sendiri memiliki dampak buruk bagi lingkungan seperti pencemaran udara dan dapat mendatangkan wabah penyakit [2]. Sampah organik dapat menghasilkan gas metana yang bisa menyebabkan menipisnya lapisan ozon. Volume sampah organik yang kian menumpuk dapat diatasi dengan cara memilah dan memanfaatkannya menjadi pakan maggot, pupuk kompos dan biogas [3]. Sampah organik sendiri banyak ditemukan di beberapa sektor usaha seperti Rumah Makan, Wisata Kuliner, Pasar dan Hotel/Resto [4], akan tetapi akses untuk mendapatkan informasi mengenai sampah organik di tempat-tempat tersebut masih minim, sehingga dikhawatirkan terjadinya penumpukan sampah organik di tempat-tempat tersebut.

Salah satu solusi untuk menangani minimnya informasi tentang sampah organik yaitu dengan cara membuat sistem klasterisasi terhadap volume sampah. Sistem klasterisasi sampah organik merupakan suatu langkah awal dalam membuat strategi pengelolaan sampah organik agar dapat dimanfaatkan oleh peternak maggot sebagai sumber pakan, pembuat pupuk kompos dan biogas serta membantu proses perencanaan angkutan sampah yang dilakukan oleh pihak Dinas Lingkungan Hidup Kota Magelang.

Dalam menemukan penyebaran volume sampah pada sektor usaha seperti Rumah Makan, Wisata Kuliner, Pasar dan Hotel/Resto di Kota Magelang ini dapat menggunakan metode Data Mining. Data Mining adalah proses menganalisis dan mengekstraksi pengetahuan secara otomatis menggunakan satu atau lebih teknik pembelajaran komputer. Salah satu fungsi dari data mining adalah *clustering*. *Clustering* adalah pengelompokan item data ke dalam kelompok-kelompok yang lebih kecil sehingga setiap kelompok memiliki kemiripan yang signifikan [5]. Algoritma yang dapat digunakan dalam *clustering* yaitu *K-Means*, *K-Medoids*, *DBSCAN*, dan *Fuzzy CMeans*. Pada penelitian ini algoritma yang digunakan adalah dengan metode *K-Means*.

Beberapa penelitian pada bidang informatika sebelumnya sudah dilaksanakan untuk menyelesaikan permasalahan sampah, seperti pada penelitian [2], dengan membuat game edukasi yang interaktif dan menarik masyarakat, khususnya anak usia dini, dapat memberikan materi bahaya dan praktik membuang sampah kepada anak-anak menggunakan model game berbasis *Multimedia Development Live Cycle* (MDLC). Pada penelitian yang dilakukan oleh [6], data mining digunakan untuk menyelesaikan permasalahan sampah dengan cara memprediksi volume sampah yang masuk ke Tempat Pengelolaan Sampah Akhir (TPSA) setiap harinya, proses prediksi volume sampah dilakukan menggunakan metode *backpropagation Artificial Neural Network* (ANN), tujuan dari penelitian tersebut adalah untuk memberikan informasi bagi pihak Pemerintah Kota melalui Dinas Lingkungan Hidup (DLH) dalam rangka mengatasi permasalahan pengelolaan sampah, seperti persiapan penyediaan lahan TPSA, penyediaan Tempat Pembuangan Sementara (TPS), kendaraan pengangkut, dan petugas pengangkutan sampah. Penelitian berikutnya dilakukan oleh [5], yang menerapkan salah satu algoritma data mining yaitu *K-Means*, untuk mengelompokkan distribusi sampah dengan kategori sampah tinggi, sedang dan rendah di Kota Bekasi, agar dapat diketahui daerah mana saja yang memiliki timbulan sampah tinggi dan penanganan yang tepat sehingga permasalahan sampah di Kota Bekasi dapat teratasi.

Algoritma klasterisasi *K-Means* merupakan salah satu dari algoritma data *mining* yang populer di kalangan peneliti [7]. Jika dibandingkan dengan algoritma klasterisasi lainnya, *K-Means* juga masih cukup tangguh di berbagai jenis data [8]. *K-Means* adalah teknik pengelompokan non-hierarki yang bertujuan untuk memisahkan data yang ada ke dalam satu atau lebih klaster atau kelompok, dengan data yang mirip ditempatkan ke dalam klaster yang sama dan data yang berbeda dibagi ke dalam kelompok yang berbeda [5].

Beberapa penelitian memanfaatkan algoritma klasterisasi *K-Means* untuk berbagai keperluan, antara

lain untuk melakukan Pengelompokan atau Klasterisasi Kasus Covid-19 di DKI Jakarta yang dilakukan oleh [9] pada penelitian tersebut menyebutkan bahwa algoritma K-Means dapat bekerja dengan baik untuk menyelesaikan permasalahan kasus Covid-19 di tingkat wilayah Provinsi DKI Jakarta dan memiliki kontribusi praktis dalam membantu penyajian data kasus Covid-19 secara lebih baik. K-Means juga dapat digunakan secara efektif untuk mengklasifikasikan daerah-daerah di Indonesia yang menyumbangkan sampah berdasarkan sampah masuk, sampah organik yang diolah, sampah anorganik yang diolah, jumlah pemulung, dan jumlah penduduk [10].

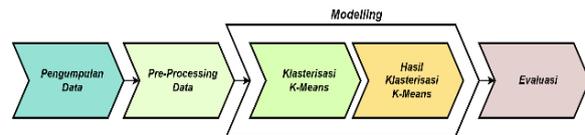
Pada penelitian yang dilakukan oleh [11] K-Means dapat digunakan sebagai data model dalam aplikasi yang dapat membantu dalam pengolahan data untuk identifikasi segmen dalam pasar, agar dapat mengetahui segmen dalam data calon pelanggan serta dataset lain sehingga mampu menganalisa atribut yang membentuk segmen tersebut. Kemudian penelitian oleh [12] K-Means digunakan untuk menganalisis data kasus Covid-19 agar dapat mengetahui pengelompokan pada masalah kasus Covid-19 di Provinsi Lampung.

Metode klasterisasi K-Means dapat dioptimalkan menggunakan metode *local outlier factor* seperti pada penelitian yang dilakukan oleh [13], dengan menggunakan metode *local outlier factor* dapat memberikan hasil yang lebih optimal pada pengelompokan data dengan metode K-Means baik dari jumlah iterasi, waktu pemrosesan dan kualitas klaster. Pada penelitian [14] dilakukan perbandingan antara algoritma K-Means dan K-Medoids untuk pengelompokan data obat dengan *silhouette coefficient*, sehingga didapatkan hasil *score* untuk K-Means sebesar 0.627 dan K-Medoids sebesar 0.536, dari hasil tersebut dapat disimpulkan hasil klaster pada K-Means lebih tinggi daripada K-Medoids dan masuk ke dalam kategori *medium structure* atau dapat dikatakan cukup baik.

Berdasarkan uraian diatas, maka pada penelitian ini akan dilakukan klasterisasi atau pengelompokan terhadap data volume sampah organik di Kota Magelang, khususnya pada sektor perdagangan dan perhotelan yang memiliki jumlah kepadatan interaksi manusia tinggi, yaitu antara pengunjung dan pedagang. Dengan data volume sampah organik yang sudah dikelompokkan menggunakan Algoritma *K-Means Clustering* ini diharapkan dapat membantu pengambilan keputusan yang tepat dalam menangani persebaran sampah organik dan meminimalisir jumlah sampah organik di Kota Magelang.

2. Metode Penelitian

Tahapan yang dilakukan dalam penelitian ini disajikan dalam bentuk alur metode penelitian seperti pada Gambar 1.



Gambar 1. Alur Metode Penelitian

Pada gambar 1 terdapat 5 tahapan dalam penelitian terhadap klasterisasi volume sampah organik di Kota Magelang terutama pada sektor perdagangan dan perhotelan yang meliputi pengumpulan data, *pre-processing* data, klasterisasi K-Means, hasil klasterisasi K-Means, dan Evaluasi. Penjelasan tiap alur yang dilakukan dalam penelitian ini adalah sebagai berikut.

2.1. Pengumpulan Data

Pengumpulan data adalah tahapan yang penting sebelum melakukan sebuah penelitian [15]. Dengan data yang memenuhi kriteria dan berkualitas akan membuat hasil penelitian yang maksimal, sehingga hasil penelitian dapat membawa informasi yang bermanfaat.

Penelitian Klasterisasi Volume Sampah Organik di Kota Magelang dilakukan menggunakan dataset jumlah pengunjung Hotel Resto, jumlah pedagang Pasar dan jumlah Pedagang Kaki Lima yang berada di Wisata Kuliner serta volume sampah pada Hotel, Resto, Pasar dan Wisata Kuliner Tahun 2022 Semester Satu di Kota Magelang. Data ini diperoleh dari situs Datago Magelang Kota yang dikelola oleh Pemerintah Kota Magelang dan data dari TPSA Banyuwirip.

2.2. Pre-Processing Data

Pre-processing data adalah proses membersihkan data mentah dan mengubahnya menjadi bentuk yang dapat diproses oleh algoritma pembelajaran mesin atau analisis data. Proses *pre-processing* data meliputi beberapa tahapan seperti penghapusan nilai kosong (*missing value*), normalisasi data, penyatuan data dari beberapa sumber, dan transformasi variable seperti pengujian *outlier* [16]. Tujuan dari *pre-processing* data adalah untuk memastikan data yang digunakan dalam analisis atau model *machine learning* adalah data yang akurat, konsisten, dan terstruktur dengan baik [17].

Pengujian *outlier* dilakukan untuk melakukan filter pada data-data yang memiliki karakteristik yang berbeda dari data lainnya. Normalisasi bertujuan untuk menskalakan nilai dari fitur yang digunakan pada penelitian ini memiliki rentang yang sama dari 0-1 [13].

2.3. Klasterisasi K-Means

Clustering merupakan aktivitas yang bertujuan mengelompokkan data yang memiliki kemiripan antara satu data dengan data lainnya ke dalam klaster atau kelompok sehingga data dalam satu klaster memiliki tingkat kemiripan yang maksimum dan data antar klaster memiliki kemiripan yang minimum [18]. K-

Means merupakan algoritma heuristik yang memisah kumpulan data ke dalam kluster k dengan meminimalkan jumlah jarak kuadrat di setiap kluster [19]. Algoritma K-Means adalah salah satu algoritma *clustering* yang masuk ke dalam kelompok *unsupervised learning* dan digunakan untuk mengelompokkan data ke dalam beberapa kelompok dengan sistem partisi [5]. *Unsupervised learning* adalah teknik tanpa fase pembelajaran dan tanpa memberi label pada setiap kelompok atau kluster [20].

Secara umum terdapat langkah-langkah dari *K-Means clustering*, yaitu menentukan nilai k atau jumlah kluster pada data set, menentukan nilai titik pusat (*centroid*), menghitung jarak terhadap masing-masing *centroid* menggunakan rumus jarak *euclidean distance* sehingga ditemukan jarak yang paling dekat dari setiap data dengan *centroid*, melakukan klasifikasi data berdasarkan kedekatannya dengan *centroid* (jarak terkecil), memperbaharui nilai *centroid*, dan kemudian melakukan perulangan dari langkah penentuan *centroid* hingga didapatkan nilai *centroid* baru sampai anggota tiap kluster tidak ada yang berubah. Jika langkah perulangan terpenuhi, nilai pusat kluster digunakan sebagai parameter iterasi terakhir untuk menentukan klasifikasi data [10]. Berikut persamaan rumus 1 *euclidean distance*.

$$d(x, y) = \sqrt{\sum_{i=0}^n (x_i - y_i)^2} \quad (1)$$

Dimana jarak antara data pada titik x dan y adalah $d(x, y)$, x adalah titik data objek, y adalah titik data *centroid* dan i adalah jumlah atribut data. Nilai *centroid* baru diperoleh dari rata-rata kluster yang bersangkutan dengan menggunakan persamaan rumus 2 berikut.

$$C_k = \left(\frac{1}{n_k}\right) \sum d_i \quad (2)$$

Dimana C_k adalah *centroid* kluster, n_k adalah jumlah data dalam *cluster* k dan d_i adalah jumlah dari nilai jarak yang masuk dalam masing masing *cluster*.

2.4. Hasil Klasterisasi K-Means

Pada penelitian ini dilakukan pengujian terhadap hasil dari klasterisasi volume sampah organik menggunakan Metode Elbow. Metode Elbow merupakan salah satu metode yang dapat digunakan untuk menguji efektifitas jumlah kluster yang digunakan pada proses klasterisasi. Metode ini bekerja dengan cara melihat persentase hasil perbandingan antara jumlah kluster yang akan membentuk siku pada suatu titik [21].

Teknik data mining dengan metode *clustering* menggunakan algoritma K-Means dikombinasikan dengan metode Elbow agar jumlah kluster mendapatkan hasil yang optimal. Jika nilai kluster pertama dengan nilai kluster kedua nilainya mengalami penurunan paling besar maka nilai kluster tersebut yang terbaik.

2.5. Evaluasi

Ada beberapa metode untuk mengevaluasi hasil pengelompokan, seperti *indeks rand*, *indeks rand* yang disesuaikan, skor distorsi, dan *silhouette score*. Meskipun sebagian besar metode evaluasi kinerja membutuhkan set pelatihan, *silhouette score* tidak membutuhkan set pelatihan untuk mengevaluasi hasil pengelompokan. Hal ini membuat *silhouette score* lebih sesuai untuk tugas pengelompokan atau klasterisasi. Berikut rumus 3 *silhouette score* dengan lebar *silhouette* $s(x_i)$ [22].

$$s(x_i) = \frac{b(x_i) - a(x_i)}{\max\{b(x_i), a(x_i)\}} \quad (3)$$

Dimana x_i adalah elemen dalam *cluster* k, $a(x_i)$ adalah jarak rata-rata dari x_i ke semua elemen lain dalam *cluster* k (dalam ketidaksamaan), dan $b(x_i) = \min\{d_l(x_i)\}$ diantara semua *cluster* $l \neq k$. Dimana $d_l(x_i)$ adalah jarak rata-rata dari x_i ke semua titik dalam *cluster* l untuk $l \neq k$ (di antara ketidaksamaan) [22].

Silhouette score dinilai berdasarkan *score* yang didapatkan, dari rentang nilai -1 sampai 1. Jika nilai rata-rata mendekati nilai 1, maka *clustering* dianggap semakin baik. Jika nilai rata-rata mendekati -1, maka *clustering* dianggap tidak baik [11]. Tabel 1 adalah kriteria *silhouette coefficient* [14].

Tabel 1. Tabel Kriteria *Silhouette Coefficient*

<i>Silhouette Coefficient (SC)</i>	Kriteria Penilaian
$0.7 < SC \leq 1.0$	<i>Strong Structure</i>
$0.5 < SC \leq 0.7$	<i>Medium Structure</i>
$0.25 < SC \leq 0.5$	<i>Weak Structure</i>
$SC \leq 0.25$	<i>No Structure</i>

3. Hasil dan Pembahasan

3.1. Pengumpulan Data

Sumber data yang digunakan sebagai variabel untuk membangun metode *K-Means Clustering* adalah data dari situs Datago Magelang Kota Tahun 2022 Semester Satu yang dikombinasikan dengan data dari TPSA Banyuwirup dengan format excel.

Data yang digunakan adalah data volume sampah organik pada pasar, hotel dan resto, serta pedagang kaki lima (PKL) per-hari. Data yang akan digunakan dalam penelitian ini disajikan pada tabel 2.

Tabel 2. Tabel Data Volume Sampah Organik

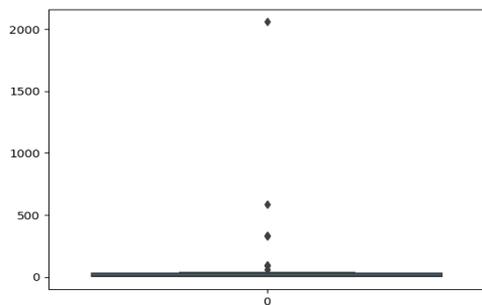
No	Nama Usaha	Volume Sampah (kg)	Jumlah
1	Pasar Kebon Polo	326	467
2	Pasar Cacaban	91	130
3	Pasar Gotong Royong	583	833
4	Pasar Rejowinangun	2057	2939
43	Hotel Ahava	6	9
44	Hotel Wisma Sejahtera	5	8
45	Hotel Educitra	1	1
46	Villa Grand Artos	3	4

3.2. Pre-Processing Data

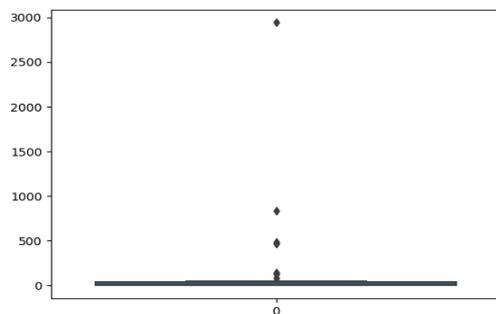
Pre-processing adalah langkah berikutnya setelah melakukan pengumpulan data. Setelah data diperoleh kemudian diseleksi berdasarkan volume sampah pada sektor usaha tersebut dan jumlah pengunjung hotel serta resto maupun jumlah pedagang yang ada di pasar, dan wisata kuliner.

Langkah selanjutnya setelah melakukan seleksi data adalah dengan pengujian *outlier* dan normalisasi data. Pengujian *outlier* bertujuan untuk memisahkan data dengan karakter yang berbeda. Gambar 2 dan 3 berikut adalah visualisasi dari data *outlier*.

Kemudian setelah menghapus data *outlier* langkah selanjutnya adalah melakukan normalisasi data menggunakan *min-max normalization*, normalisasi data ditujukan untuk menskalakan nilai variabel yang memiliki rentang fitur 0-1 sehingga dapat menghasilkan kluster yang lebih optimal.



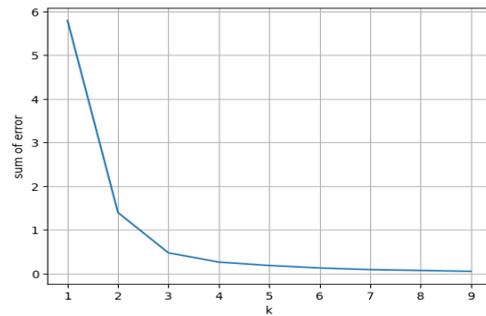
Gambar 2. Data *outlier* dari variabel volume sampah (kg)



Gambar 3. Data *outlier* dari variabel jumlah

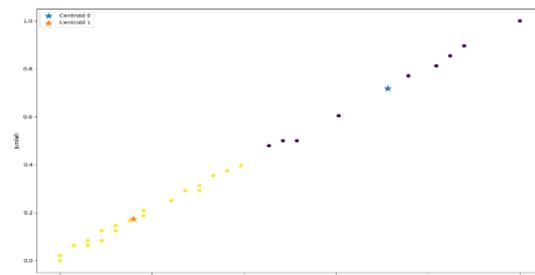
3.3. Proses dan Hasil Klasterisasi K-Means (*Modelling*)

Setelah tahap *pre-processing* langkah selanjutnya adalah menerapkan metode untuk menentukan jumlah kluster terbaik menggunakan *elbow method*. Hasil dari penggunaan metode elbow pada penelitian ini menghasilkan jumlah kluster terbaik sebanyak 2 kluster. Pada gambar 4 berikut dapat dilihat visualisasi hasil dari metode elbow dimana menunjukkan nilai k sebagai kluster mengalami penurunan yang signifikan pada angka 2, sehingga menjadikan 2 kluster sebagai hasil yang paling optimal.



Gambar 4. Visualisasi *Elbow Method*

Selanjutnya setelah tahap *pre-processing* dan mencari nilai kluster terbaik menggunakan *elbow method*, kemudian dilanjutkan pada tahap berikutnya yaitu *modelling* atau penerapan algoritma *K-Means Clustering*. Pada tahap *modelling* langkah pertama adalah memasukkan jumlah kluster yang sudah didapatkan dari penerapan *elbow method* yaitu sebanyak 2 kluster agar dapat menghasilkan 2 pengelompokan atau penyebaran kluster volume sampah organik. Gambar 5 adalah visualisasi dari hasil kluster volume sampah organik.



Gambar 5. Visualisasi Penyebaran Kluster Volume Sampah

Berdasarkan visualisasi gambar 5, titik warna kuning adalah kluster 1 yang dapat diartikan sebagai penghasil sampah organik kategori rendah, kemudian warna ungu adalah kluster 2 yang menunjukkan sebagai kategori penghasil sampah organik tinggi, masing-masing kluster ditandai titik tengah (*centroid*), *centroid* adalah nilai rata-rata kluster atau titik tengah dari kluster yang bersangkutan, dengan simbol bintang berwarna oranye adalah *centroid* dari kluster 1, bintang berwarna biru adalah *centroid* dari kluster 2.

Tabel 3. Tabel Jumlah nama usaha dan rata-rata volume sampah setiap kluster

Kluster	Jumlah Nama Usaha	Jumlah Pengunjung/Pedagang	Rata-Rata Volume Sampah (Kg)
Kluster 1	27	252	6.25
Kluster 2	10	355	24.5

Pada tabel 3 dijelaskan bahwa kluster 1 dengan kategori penghasil sampah rendah memiliki rata-rata volume sampah yang lebih rendah apabila dibandingkan dengan rata-rata dari kluster 2 meskipun pada kluster 1 memiliki jumlah data yang lebih banyak dari kluster 2, akan tetapi

jumlah pengunjung/pedagang pada kluster 2 lebih dominan. Hal ini menjelaskan bahwa data jumlah pengunjung/pedagang memiliki pengaruh yang sangat penting terhadap volume sampah yang dihasilkan. Tabel 4 dan 5 adalah data dari jumlah data dan jumlah seluruh volume sampah berdasarkan kategori tempat penyebarannya pada tiap kluster.

Tabel 4. Tabel Data jumlah dan rata-rata volume sampah kluster 1

Kluster 1	Jumlah Nama Usaha	Jumlah Pengunjung/Pedagang	Jumlah Volume Sampah (Kg)
Pasar/PKL	9	124	83
Hotel/Resto	18	128	86

Tabel 5. Tabel Data jumlah dan rata-rata volume sampah kluster 2

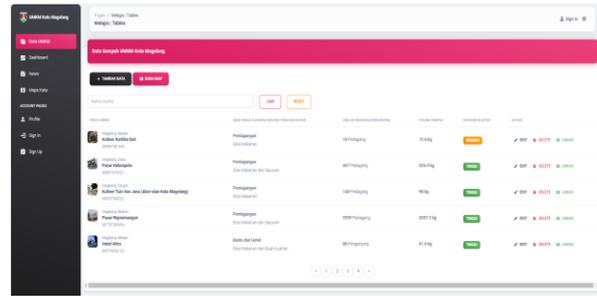
Kluster 2	Jumlah Nama Usaha	Jumlah Pengunjung/Pedagang	Jumlah Volume Sampah (Kg)
Pasar/PKL	8	281	193
Hotel/Resto	2	74	52

Hasil dari klusterisasi ini dapat digunakan oleh pencari sampah organik yang memanfaatkan sampah tersebut sebagai ladang penghasilan sehari-hari, seperti peternak maggot, pembuat pupuk kompos, dan biogas. Dari hasil tersebut para pemanfaat sampah organik dapat dengan mudah mencari informasi dan lokasi mengenai ketersediaan sampah organik di Kota Magelang. Data tersebut juga dapat membantu Dinas Lingkungan Hidup Kota Magelang dalam merencanakan proses pengangkutan sampah di tempat-tempat yang ada pada hasil kluster tersebut. Pada tabel 6 ditampilkan data nama usaha pada setiap kluster.

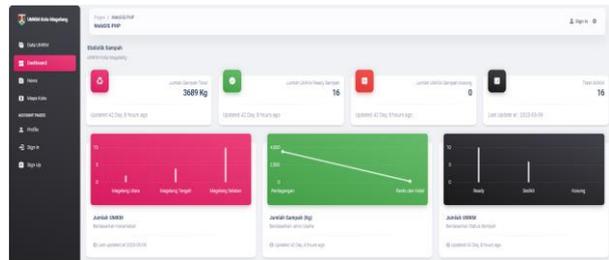
Tabel 6. Tabel Data Jumlah nama usaha setiap kluster

Kluster	Jumlah Nama Usaha
Kluster 1	PKL Sari Boga Kencana, PKL Sejuta Bunga, PKL Daha, PKL Jendralan, PKL Jl. Pajajaran, PKL Rejomulyo, PKL Kauman, PKL Kalingga, PKL S. Parman, Hotel Puri Asri, Hotel Borobudur Indah, Hotel Sriti, Hotel Oxalis, Hotel Wisata, Hotel Safari, Hotel Bharata, Hotel Pringgading, Hotel Sumber Waras, Hotel Lokasari, Hotel Mutiara, Hotel Ardiva, Hotel Wijaya, Hotel Safira, Hotel Ahava, Hotel Wisma Sejahtera, Hotel Educitra, Villa Grand Artos.
Kluster 2	PKL Sigaluh, PKL Armada Estate, PKL Kartika Sari, PKL Jenggolo, PKL Badaan, PKL Jl. Sriwijaya, PKL Jl. Alibasah, PKL Lembah Tidar, Hotel Trio, Hotel City Hub.

Berdasarkan persebaran dari hasil kluster tersebut, dikembangkan sebuah *website*. *Website* tersebut digunakan untuk menampilkan data hasil kluster supaya dapat dimanfaatkan oleh pemanfaat sampah organik, agar lebih mudah lagi dalam mencari informasi mengenai ketersediaan sampah tersebut di tempat-tempat tertentu, dengan hanya bermodalkan ponsel dan internet. Gambar 6 adalah tampilan dari *website* informasi klusterisasi sampah organik.



Gambar 6. Tampilan *website* informasi klusterisasi sampah organik



Gambar 7. Tampilan menu grafik pada *website*

Pada *website* tersebut juga ditampilkan grafik dari data sampah organik yang telah dimasukkan. Gambar 7 adalah tampilan dari menu grafik.

3.4. Evaluasi

Pada tahap terakhir, dilakukan evaluasi hasil dari klusterisasi volume sampah organik menggunakan metode *silhouette score*. Metode ini adalah salah satu cara untuk mengukur seberapa efektif algoritma K-Means bekerja dalam menghasilkan pengelompokan atau klusterisasi.

Pada penelitian sebelumnya oleh [14], dilakukan perbandingan klusterisasi menggunakan algoritma K-Means dan K-Medoids dalam kasus pengelompokan obat di Puskesmas dengan menggunakan metode *silhouette score*, didapatkan hasil *score* sebesar 0.627 pada algoritma K-Means, sedangkan pada algoritma K-Medoids menghasilkan *score* sebesar 0.536. Berdasarkan hasil nilai *score* tersebut dapat dinyatakan bahwa hasil dari algoritma K-Means lebih baik dibandingkan K-Medoids. Pada penelitian [10], dalam pengelompokan atau klusterisasi daerah penyumbang sampah, algoritma K-Means dapat bekerja lebih cepat dan efisien dalam mengelompokkan data dibandingkan dengan algoritma Fuzzy C-Means.

Berdasarkan ulasan tersebut, penulis memilih algoritma K-Means dalam kasus klusterisasi volume sampah organik. Pada penelitian ini hasil dari klusterisasi mendapatkan nilai *score* sebesar 0.66, dan dapat disimpulkan bahwa penerapan algoritma K-Means dalam pengelompokan atau klusterisasi pada penelitian ini sudah cukup baik (*Medium Structure*), karena nilai *silhouette score* lebih mendekati angka 1 dibandingkan dengan -1.

4. Kesimpulan

Berdasarkan hasil penerapan algoritma K-Means dalam klusterisasi volume sampah organik di Kota Magelang, dapat disimpulkan bahwa metode K-Means bekerja dengan cukup baik. Hasil klusterisasi dengan nilai elbow, sehingga hasil tersebut merupakan nilai kluster optimal atau terbaik. Kluster 1 adalah kategori penghasil volume sampah rendah dengan jumlah anggota 27 dan rata-rata volume sampah per-hari sebesar 6.25 kg, sedangkan kluster 2 yaitu kategori penghasil sampah tinggi memiliki 10 anggota dengan rata-rata volume sampah per-hari sebesar 24.5 kg.

Dari analisis hasil data klusterisasi dapat diketahui jumlah anggota pada kluster 1 kategori Pasar/PKL tidak lebih banyak dari Hotel/Resto, akan tetapi Pasar/PKL memiliki jumlah pengunjung/pedagang yang lebih banyak, hal itu membuat volume sampah per-hari menjadi tinggi, pada kluster 2 volume sampah juga tinggi apabila total pengunjung/pedagang per-harinya tinggi.

Hasil tersebut akan membantu pengambilan keputusan yang tepat dalam menangani permasalahan sampah organik, khususnya bagi Dinas Lingkungan Hidup Kota Magelang, pembudidaya maggot di Kota Magelang dan sekitarnya, serta pemanfaatan sampah organik lainnya supaya mendapatkan informasi mengenai lokasi ketersediaan sampah organik, yang diharapkan dapat membantu mengurangi volume sampah organik.

Pada penelitian selanjutnya diharapkan bisa menambahkan data penghasil sampah khususnya organik dari sektor lainnya yang memungkinkan untuk dilakukan klusterisasi terhadap volume sampah organik dan dengan menerapkan lebih dari satu algoritma, agar dapat membandingkan hasil penerapan K-Means dengan penerapan algoritma *clustering* yang lain.

Ucapan Terimakasih

Terima kasih kepada Dinas Lingkungan Hidup Kota Magelang yang telah memberikan pendampingan serta masukan dalam penelitian ini, dan juga kepada pihak PHRI (Persatuan Hotel dan Restoran Indonesia) Magelang yang telah mendampingi dalam proses pengumpulan data.

Daftar Rujukan

- [1] N. F. Muto'in and A. Utami, "Analisis Tingkat Kecelakaan Lalu Lintas Menggunakan Metode Accident Rate Dan Equivalent Accident Number (EAN) Di Kota Magelang," *J. Rekayasa Sipil JRS-Unand*, vol. 18, no. 1, p. 60, May 2022, doi: 10.25077/jrs.18.1.60-67.2022.
- [2] R. I. Borman and Y. Purwanto, "Implementasi Multimedia Development Live Cycle pada Pengembangan Game Edukasi Pengenalan Bahaya Sampah pada Anak," vol. 5, no. 2, 2019, doi: <http://dx.doi.org/10.26418/jp.v5i2.25997>.
- [3] R. Mulyani, D. I. Anwar, and N. Nurbaeti, "Pemanfaatan Sampah Organik untuk Pupuk Kompos dan Budidaya Maggot

- Sebagai Pakan Ternak," *JPM J. Pemberdaya. Masy.*, vol. 6, no. 1, pp. 568–573, May 2021, doi: 10.21067/jpm.v6i1.4911.
- [4] N. Salman, E. Nofiyanti, and T. Nurfadhilah, "Pengaruh dan Efektivitas Maggot Sebagai Proses Alternatif Penguraian Sampah Organik Kota di Indonesia," *J. Serambi Eng.*, vol. 5, no. 1, Dec. 2019, doi: 10.32672/jse.v5i1.1655.
- [5] D. Ruwandara, M. Jajuli, and A. Rizal, "Analisis Algoritma K-Means Clustering Untuk Daerah Penyebaran Sampah di Kota Bekasi," *JOINS J. Inf. Syst.*, vol. 6, no. 1, pp. 56–63, May 2021, doi: 10.33633/joins.v6i1.4085.
- [6] W. Santoso and P. Sukmasetya, "Prediksi Volume Sampah di TPSA Banyuwirip Menggunakan Metode Backpropagation Neural Network," vol. 7, 2023, doi: <http://dx.doi.org/10.30865/mib.v7i1.5499>.
- [7] M. Ahmed, R. Seraj, and S. M. S. Islam, "The k-means Algorithm: A Comprehensive Survey and Performance Evaluation," *Electronics*, vol. 9, no. 8, p. 1295, Aug. 2020, doi: 10.3390/electronics9081295.
- [8] M. Z. Rodriguez *et al.*, "Clustering algorithms: A comparative approach," *PLOS ONE*, vol. 14, no. 1, p. e0210236, Jan. 2019, doi: 10.1371/journal.pone.0210236.
- [9] A. Solichin and K. Khairunnisa, "Klusterisasi Persebaran Virus Corona (Covid-19) Di DKI Jakarta Menggunakan Metode K-Means," *Fountain Inform. J.*, vol. 5, no. 2, p. 52, Oct. 2020, doi: 10.21111/fij.v5i2.4905.
- [10] B. N. S. Desi Kristina Sitinjak, "Clustering Daerah Penyumbang Sampah Berdasarkan Provinsi di Indonesia Menggunakan Algoritma K-Means," Sep. 2022, doi: 10.5281/ZENODO.7059032.
- [11] Y. Christian and K. O. Y. R. Qi, "Penerapan K-Means pada Segmentasi Pasar untuk Riset Pemasaran pada Startup Early Stage dengan Menggunakan CRISP-DM," *JURIKOM J. Ris. Komput.*, vol. 9, no. 4, p. 966, Aug. 2022, doi: 10.30865/jurikom.v9i4.4486.
- [12] Z. Nabila, A. R. Isnain, and Z. Abidin, "Analisis Data Mining Untuk Clustering Kasus Covid-19 di Povinsi Lampung Dengan Algoritma K-Means," *J. Teknol. Dan Sist. Inf.*, vol. 2, no. 2, doi: <https://doi.org/10.33365/jtsi.v2i2.868>.
- [13] P. A. Ariawan, "Optimasi Pengelompokan Data Pada Metode K-means dengan Analisis Outlier," *J. Nas. Teknol. Dan Sist. Inf.*, vol. 5, no. 2, pp. 88–95, Sep. 2019, doi: 10.25077/TEKNOSI.v5i2.2019.88-95.
- [14] R. A. Farissa, R. Mayasari, and Y. Umaidah, "Perbandingan Algoritma K-Means dan K-Medoids Untuk Pengelompokan Data Obat dengan Silhouette Coefficient," vol. 5, no. 2, doi: <https://doi.org/10.30871/jaic.v5i1.3237>.
- [15] A. Handayanto, K. Latifa, N. D. Saputro, and R. R. Waliansyah, "Analisis dan Penerapan Algoritma Support Vector Machine (SVM) dalam Data Mining untuk Menunjang Strategi Promosi," *JUITA J. Inform.*, vol. 7, no. 2, p. 71, Nov. 2019, doi: 10.30595/juita.v7i2.4378.
- [16] L. Karlina and O. Nurdiawan, "Penerapan K-Medoids Dalam Klasifikasi Persebaran Lahan Kritis di Jawa Barat Berdasarkan Kabupaten/Kota," vol. 7, no. 1, 2023, doi: <https://doi.org/10.36040/jati.v7i1.6348>.
- [17] P. Mishra, A. Biancolillo, J. M. Roger, F. Marini, and D. N. Rutledge, "New data preprocessing trends based on ensemble of multiple preprocessing techniques," *TrAC Trends Anal. Chem.*, vol. 132, p. 116045, Nov. 2020, doi: 10.1016/j.trac.2020.116045.
- [18] N. T. S. Saptadi and V. P. Taga, "Penerapan Algoritma K-Means Dalam Mengatasi Permasalahan Sampah (Studi Kasus: Pola Perilaku Masyarakat Dalam Membuang Sampah)," vol. 10, no. 1, 2022, doi: <https://doi.org/10.56963/tematika.v10i1.134>.
- [19] M. Nishom, "Perbandingan Akurasi Euclidean Distance, Minkowski Distance, dan Manhattan Distance pada Algoritma K-Means Clustering berbasis Chi-Square," *J. Inform. J. Pengemb. IT*, vol. 4, no. 1, pp. 20–24, Jan. 2019, doi: 10.30591/jpit.v4i1.1253.
- [20] H. Priyatman, F. Sajid, and D. Haldivany, "Klusterisasi Menggunakan Algoritma K-Means Clustering untuk

- Memprediksi Waktu Kelulusan Mahasiswa,” *J. Edukasi Dan Penelit. Inform. JEPIN*, vol. 5, no. 1, p. 62, Apr. 2019, doi: 10.26418/jp.v5i1.29611.
- [21] V. A. Ekasetya and A. Jananto, “Klusterisasi Optimal Dengan Elbow Method Untuk Pengelompokan Data Kecelakaan Lalu Lintas di Kota Semarang,” *J. Din. Inform.*, vol. 12, no. 1, pp. 20–28, Aug. 2020, doi: 10.35315/informatika.v12i1.8159.
- [22] M. Shutaywi and N. N. Kachouie, “Silhouette Analysis for Performance Evaluation in Machine Learning with Applications to Clustering,” *Entropy*, vol. 23, no. 6, p. 759, Jun. 2021, doi: 10.3390/e23060759.