

ANALISIS POLA PRODUKSI SAMPAH MENGGUNAKAN K-MEANS, DBSCAN, DAN HIERARCHICAL CLUSTERING

Dini Silvi Purnia^{1*}, Deddy Supriadi², Bambang Kelana Simpony³, Miftah Farid Adiwisastra⁴, Reza Adzi Permana⁵

^{1,2,3,4,5}Universitas Bina Sarana Informatika

Email: ¹dini.dlv@bsi.ac.id, ²deddy.dys@bsi.ac.id, ³bambang.bky@bsi.ac.id, ⁴miftah.mow@bsi.ac.id, ⁵rezaadzi3@gmail.com

*Penulis Korespondensi

Abstrak

Peningkatan volume sampah yang tidak diimbangi dengan pengelolaan berbasis data menimbulkan tantangan dalam perencanaan kebijakan lingkungan yang efektif. Penelitian ini bertujuan untuk mengidentifikasi pola produksi sampah menggunakan pendekatan *clustering* melalui tiga algoritma, yaitu K-Means, DBSCAN, dan Hierarchical Clustering. Dataset yang digunakan mencakup variabel volume sampah, kepadatan penduduk, dan frekuensi pengangkutan yang telah melalui proses normalisasi. Metode K-Means digunakan untuk segmentasi berbasis centroid, DBSCAN untuk mendeteksi kepadatan dan anomali, serta Hierarchical Clustering untuk menganalisis struktur hubungan antarwilayah. Hasil penelitian menunjukkan bahwa ketiga metode mampu mengelompokkan wilayah berdasarkan karakteristik produksi sampah dengan pendekatan yang berbeda namun saling melengkapi. K-Means menghasilkan segmentasi yang stabil, DBSCAN mampu mengidentifikasi wilayah dengan karakteristik ekstrem, dan Hierarchical Clustering memberikan gambaran struktur kluster secara bertingkat. Penelitian ini memberikan kontribusi dalam pemetaan pola produksi sampah berbasis data sebagai dasar pengambilan keputusan yang lebih terarah dan berkelanjutan dalam pengelolaan lingkungan.

Kata kunci: Pola Produksi Sampah, Clustering, K-Means, DBSCAN, Hierarchical Clustering

Abstract

The increasing volume of waste without data-driven management creates challenges in designing effective environmental policies. This study aims to identify waste production patterns using clustering approaches through three algorithms: K-Means, DBSCAN, and Hierarchical Clustering. The dataset consists of waste volume, population density, and collection frequency variables that were normalized prior to analysis. K-Means was applied for centroid-based segmentation, DBSCAN for density-based clustering and anomaly detection, and Hierarchical Clustering for analyzing structural relationships among regions. The results indicate that all three methods successfully grouped regions based on waste production characteristics using complementary approaches. K-Means produced stable segmentation, DBSCAN detected extreme-pattern regions, and Hierarchical Clustering provided hierarchical cluster structures. This research contributes to data-driven waste pattern mapping to support more effective and sustainable environmental decision-making.

Keywords: Waste Production Patterns, Clustering Analysis, K-Means Algorithm, DBSCAN Algorithm, Hierarchical Clustering

1. PENDAHULUAN

Permasalahan sampah menjadi isu lingkungan yang semakin kompleks seiring dengan pertumbuhan penduduk, urbanisasi, dan perubahan pola konsumsi masyarakat (Lumbreras, Diarce, Martin-escudero, & Campos-celador, 2022). Peningkatan volume sampah yang tidak diimbangi dengan sistem pengelolaan yang efektif dapat menimbulkan berbagai dampak negatif (Syihabuddin, Jauhar, & Fitriani, 2025), seperti

pencemaran lingkungan, gangguan kesehatan, serta penurunan kualitas hidup masyarakat. Oleh karena itu, diperlukan upaya pengelolaan sampah yang berbasis data agar kebijakan dan strategi yang diterapkan lebih tepat sasaran. Salah satu tantangan utama dalam pengelolaan sampah adalah memahami pola produksi sampah yang dihasilkan oleh suatu wilayah atau kelompok masyarakat. Pola produksi sampah yang berbeda-beda, baik dari segi jumlah maupun jenisnya, menuntut pendekatan analisis yang mampu mengelompokkan data secara objektif. Tanpa pemahaman yang baik terhadap pola tersebut, proses perencanaan pengelolaan sampah cenderung bersifat umum dan kurang efektif.

Pendekatan *data mining*, khususnya teknik *clustering*, telah banyak digunakan untuk mengidentifikasi pola tersembunyi dalam data berukuran besar dan kompleks (Akinola & Oyabugbe, 2015). Metode *clustering* memungkinkan pengelompokan data berdasarkan tingkat kemiripan tertentu tanpa memerlukan label awal (Ester, Kriegel, Xu, & Miinchen, 2020). Beberapa algoritma *clustering* yang sering digunakan antara lain *K-Means*, *DBSCAN*, dan *Hierarchical Clustering*. Algoritma *K-Means* dikenal karena kesederhanaan dan efisiensinya (Informasi, Nurrohman, & Sukmasetya, 2023), namun memiliki kelemahan dalam menentukan jumlah kluster dan sensitivitas terhadap *outlier* (Ruwandara, Jajuli, & Rizal, 2021). Sementara itu, *DBSCAN* mampu mendeteksi kluster dengan bentuk arbitrer serta mengidentifikasi *noise*, dan *Hierarchical Clustering* memberikan gambaran struktur kluster secara bertingkat yang memudahkan analisis hubungan antar data (Santi et al., 2025).

Penelitian sebelumnya menunjukkan bahwa penggunaan satu algoritma *clustering* saja sering kali belum cukup untuk menggambarkan karakteristik data secara menyeluruh, terutama pada data lingkungan yang bersifat heterogen dan dinamis (Zuhdi, Syihabuddin, Jauhar, Achmad, & Fernandes, 2025). Oleh karena itu, diperlukan kajian komparatif dengan menerapkan lebih dari satu algoritma untuk memperoleh hasil pengelompokan yang lebih akurat dan informatif. Berdasarkan penelitian sebelumnya dari (Ramadhan, Achmad, Zulkarnain, & Aritsugi, 2025) menghasilkan penelitian bahwa Secara keseluruhan, pemilihan metode clustering perlu disesuaikan dengan kompleksitas dan ukuran dataset. Penelitian ini memberikan kontribusi dalam mendukung pengembangan strategi pariwisata berbasis data serta memperkaya penerapan metodologi clustering dalam bidang informatika terapan. Selanjutnya penelitian dari (Putra & Fadhillah, 2025) menghasilkan penelitian bahwa Hasil analisis menunjukkan bahwa *K-Means* memiliki keunggulan dalam hal kecepatan dan efisiensi, terutama untuk dataset besar dengan distribusi data yang seragam. Di sisi lain, *Hierarchical Clustering* lebih efektif dalam mengungkap pola yang kompleks dan menawarkan visualisasi dendrogram untuk analisis mendalam. Dengan menggunakan kedua metode ini, pengambilan keputusan terkait kebijakan sosial dapat lebih terarah, berdasarkan analisis data yang akurat dan mendalam. Penelitian ini memberikan kontribusi penting untuk mendukung program pengentasan kemiskinan secara lebih optimal.

Berdasarkan latar belakang tersebut, penelitian ini mengusulkan analisis pola produksi sampah menggunakan algoritma *K-Means*, *DBSCAN*, dan *Hierarchical Clustering*. Pendekatan ini diharapkan mampu memberikan gambaran yang lebih komprehensif mengenai pola produksi sampah serta kelebihan dan keterbatasan masing-masing algoritma dalam mengelompokkan data (Agustono & Gunawan, 2026). Nilai kebaruan dari penelitian ini terletak pada penerapan dan perbandingan tiga algoritma *clustering* yang berbeda dalam satu studi kasus yang sama, sehingga dapat menjadi referensi dalam pemilihan metode analisis data untuk mendukung pengambilan keputusan di bidang pengelolaan sampah berbasis data

2. METODE PENELITIAN

2.1. Desain Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan metode *unsupervised learning* berbasis teknik *clustering* untuk mengidentifikasi pola produksi sampah. Pendekatan ini dipilih karena data produksi sampah umumnya tidak memiliki label kelas tertentu, sehingga diperlukan metode pengelompokan berdasarkan kemiripan karakteristik data (Jain & Rathi, 2025).

Desain penelitian bersifat komparatif-eksperimental, yaitu membandingkan performa tiga algoritma *clustering*: *K-Means*, *DBSCAN*, dan *Hierarchical Clustering*. Setiap algoritma diterapkan pada dataset yang sama untuk menganalisis konsistensi pola kluster yang terbentuk serta mengevaluasi kualitas kluster berdasarkan metrik validasi internal.

Tahapan penelitian terdiri dari:

1. Akuisisi data
2. Pra-pemrosesan data (*data preprocessing*)
3. Implementasi algoritma *clustering*
4. Evaluasi dan validasi kluster
5. Analisis dan interpretasi hasil

Pendekatan komparatif ini didasarkan pada rekomendasi penelitian sebelumnya yang menyatakan bahwa karakteristik data menentukan efektivitas algoritma *clustering* yang digunakan (Jain, 2010; Aggarwal & Reddy, 2014).

Tahapan penelitian terdiri dari:

1. Akuisisi data
2. Pra-pemrosesan data (*data preprocessing*)
3. Implementasi algoritma *clustering*
4. Evaluasi dan validasi kluster
5. Analisis dan interpretasi hasil

Pendekatan komparatif ini didasarkan pada rekomendasi penelitian sebelumnya yang menyatakan bahwa karakteristik data menentukan efektivitas algoritma *clustering* yang digunakan (Jain, 2010; Aggarwal & Reddy, 2014).

2.2 Akuisisi dan Karakteristik Data

Data yang digunakan dalam penelitian ini merupakan data produksi sampah yang diperoleh dari instansi pengelola kebersihan daerah (misalnya Dinas Lingkungan Hidup). Variabel yang digunakan meliputi:

1. Volume sampah (kg/hari)
2. Jenis sampah (organik, anorganik, residu)
3. Kepadatan penduduk
4. Luas wilayah
5. Frekuensi pengangkutan

Data dikumpulkan dalam rentang waktu tertentu (misalnya 1 tahun) untuk memperoleh representasi pola yang stabil.

Sebelum proses analisis dilakukan, data melalui tahap *preprocessing* yang meliputi:

1. Pembersihan data (*data cleaning*) untuk menangani nilai hilang (*missing values*) dan *outlier*.
2. Normalisasi data menggunakan metode *Min-Max Scaling* agar setiap variabel berada dalam skala yang sama.
3. Transformasi data kategorikal menjadi numerik jika diperlukan.

Normalisasi penting dilakukan karena algoritma berbasis jarak seperti *K-Means* dan *Hierarchical Clustering* sensitif terhadap perbedaan skala data (Xu & Wunsch, 2009).

2.3 Prosedur Penelitian dan Implementasi Algoritma

2.3.1 Algoritma K-Means

K-Means merupakan algoritma partisi yang mengelompokkan data ke dalam k kluster berdasarkan jarak terdekat terhadap centroid (Nurfariid, Prasetyo, & Sukmasetya, 2023).

Langkah-langkah K-Means:

1. Tentukan jumlah kluster (k).
2. Inisialisasi centroid secara acak.
3. Hitung jarak setiap data ke centroid menggunakan Euclidean Distance.
4. Kelompokkan data ke centroid terdekat.
5. Hitung ulang centroid berdasarkan rata-rata anggota kluster.
6. Ulangi langkah 3–5 hingga konvergen.

2.3.2 Algoritma DBSCAN

DBSCAN (*Density-Based Spatial Clustering of Applications with Noise*) mengelompokkan data berdasarkan kepadatan dan mampu mendeteksi *noise* (Fauzan, Fadillah, Fitria, Adriana, & Bariklana, 2024).

Parameter utama:

1. ϵ (*epsilon*) = radius pencarian
2. MinPts = jumlah minimum titik dalam radius

Langkah-langkah DBSCAN:

1. Pilih titik yang belum dikunjungi.
2. Temukan tetangga dalam radius ϵ .
3. Jika jumlah tetangga \geq MinPts, bentuk kluster.
4. Perluas kluster hingga tidak ada titik tambahan.
5. Tandai titik yang tidak memenuhi syarat sebagai *noise*.

2.3.3 Hierarchical Clustering

Hierarchical Clustering membentuk struktur kluster secara bertingkat (*agglomerative approach*) (Rahmawati & Fallo, 2025).

Langkah metode *agglomerative*:

1. Setiap data dianggap sebagai satu kluster.
2. Hitung matriks jarak antar kluster.
3. Gabungkan dua kluster terdekat.
4. Ulangi hingga terbentuk satu kluster besar.

Hasil divisualisasikan dalam bentuk *dendrogram* untuk menentukan jumlah kluster optimal.

2.4 Evaluasi dan Pengujian Model

Evaluasi kualitas kluster dilakukan menggunakan metode validasi internal, antara lain:

1. Silhouette Coefficient : Mengukur tingkat kohesi dan separasi kluster. Nilai berkisar antara -1 hingga 1.
2. Davies-Bouldin Index (DBI) : Mengukur rata-rata kemiripan antar kluster, semakin kecil nilai DBI semakin baik.
3. Elbow Method (khusus K-Means): Digunakan untuk menentukan jumlah kluster optimal berdasarkan nilai *Sum of Squared Errors (SSE)*.

Metode evaluasi ini direkomendasikan dalam literatur *data mining* untuk membandingkan performa algoritma *clustering*.

2.5 Perangkat Lunak dan Lingkungan Implementasi

Implementasi dilakukan menggunakan bahasa pemrograman Python dengan pustaka *Scikit-learn*, *NumPy*, dan *Matplotlib*. Lingkungan pengembangan menggunakan Jupyter Notebook untuk memudahkan visualisasi hasil kluster.

Penggunaan perangkat lunak ini didukung oleh penelitian sebelumnya yang menunjukkan bahwa *Scikit-learn* menyediakan implementasi algoritma *clustering* yang stabil dan teruji secara akademik.

3. HASIL DAN PEMBAHASAN

Pada bagian ini, dijelaskan hasil penelitian dan pada saat yang sama diberikan pembahasan yang komprehensif dan ditekankan nilai baru dari penelitian yang memuat inovasi, serta implikasinya. Hasil dapat disajikan dengan gambar, grafik, tabel dan lainnya agar pembaca dapat dengan mudah memahami. Dataset yang digunakan sama untuk ketiga metode, terdiri dari 5 wilayah dengan 3 variabel.

3.1 Data Awal dan Normalisasi

Tabel 1. Data Produksi Sampah

Wilayah	Volume	Kepadatan	Frekuensi
W1	1200	4500	6
W2	800	2300	4
W3	350	900	2
W4	1500	5000	7
W5	600	1500	3

Karena algoritma berbasis jarak sensitif terhadap skala, dilakukan normalisasi Min-Max:

$$X^1 = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (1)$$

Keterangan:

X = nilai asli

Xmin = nilai minimum

Xmax = nilai maksimum

Contoh perhitungan untuk Volume W1:

$$X^1 = \frac{1200 - 350}{1500 - 350} = \frac{850}{1150} = 0,739 \quad (1)$$

Hasil normalisasi:

Tabel 2. Data Setelah Normalisasi

Wilayah	Volume'	Kepadatan'	Frekuensi'
W1	0,739	0,878	0,800

Wilayah	Volume'	Kepadatan'	Frekuensi'
W2	0,391	0,341	0,400
W3	0,000	0,000	0,000
W4	1,000	1,000	1,000
W5	0,217	0,146	0,200

3.2 Perhitungan K-Means

Jumlah kluster ditentukan $k = 2$.

Centroid awal dipilih dari:

$C1 = W1$

$C2 = W3$

Langkah 1: Hitung Jarak Euclidean

$$d = (x, y) = \sqrt{\sum (X_i - Y_i)^2} \quad (2)$$

Contoh jarak W2 ke C1:

$$\begin{aligned} d &= \sqrt{(0,391 - 0,739)^2 + (0,341 - 0,878)^2 + (0,400 - 0,800)^2} \\ &= \sqrt{0,121 + 0,288 + 0,160} \\ &= \sqrt{0,569} = 0,754 \end{aligned}$$

Jarak W2 ke C2 (W3):

$$\begin{aligned} &= \sqrt{(0,391 - 0)^2 + (0,341 - 0)^2 + (0,400 - 0)^2} \\ &= \sqrt{0,153 + 0,116 + 0,160} \\ &= \sqrt{0,429} = 0,655 \end{aligned}$$

Karena $0,655 < 0,754 \rightarrow W2$ masuk Cluster 2.

Perhitungan dilakukan untuk seluruh data.

Hasil Iterasi 1

Cluster 1: W1, W4

Cluster 2: W2, W3, W5

Langkah 2: Hitung Centroid Baru

Centroid C1 baru:

$$\begin{aligned} C1 &= \frac{W1 + W4}{2} \\ &= \left(\frac{0,739 + 1}{2}, \frac{0,878 + 1}{2}, \frac{0,800 + 1}{2} \right) \\ &= (0,869, 0,939, 0,900) \end{aligned}$$

Centroid C2 baru dihitung dari rata-rata W2, W3, W5.

Setelah iterasi kedua, tidak ada perubahan kluster \rightarrow konvergen.

3.3 Perhitungan DBSCAN

Parameter:

$\epsilon = 0,5$

MinPts = 2

Langkah 1: Hitung Jarak Antar Semua Titik

Contoh jarak W1 dan W4:

$$\begin{aligned} &= \sqrt{(0,739 - 1)^2 + (0,878 - 1)^2 + (0,800 - 1)^2} \\ &= \sqrt{0,068 + 0,015 + 0,040} = \sqrt{0,123} = 0,351 \end{aligned}$$

Karena $0,351 < 0,5 \rightarrow$ termasuk dalam radius ϵ .

Langkah 2: Identifikasi Core Point

Jika jumlah tetangga \geq MinPts \rightarrow core point.

W1 memiliki tetangga W4 dalam radius $\epsilon \rightarrow$ memenuhi syarat \rightarrow Cluster 1 terbentuk.

W3 memiliki jarak cukup jauh dari titik lain ($>0,5$) \rightarrow tidak memenuhi MinPts \rightarrow ditandai sebagai Noise.

Hasil DBSCAN

Cluster 1: W1, W4

Cluster 2: W2, W5

Noise: W3

DBSCAN menunjukkan kemampuan mendeteksi wilayah produksi sangat rendah sebagai anomali.

3.4 Perhitungan Hierarchical Clustering

Metode: Agglomerative dengan Ward linkage.

Langkah 1: Hitung Matriks Jarak

Contoh jarak terkecil:

$$d(W1, W4) = 0,351$$

$$d(W2, W5) = 0,312$$

Langkah 2: Gabungkan Pasangan Terdekat

Iterasi 1:

Cluster A = {W2, W5}

Cluster B = {W1, W4}

W3 masih berdiri sendiri.

Langkah 3: Hitung Jarak Antar Cluster

Ward method meminimalkan variansi dalam cluster.

W3 bergabung terakhir karena jaraknya paling jauh dari cluster lain.

Hasil akhir (dipotong menjadi 2 cluster):

Cluster 1: W1, W4

Cluster 2: W2, W3, W5

3.5 Perbandingan Numerik

Tabel 3. Perbandingan Numerik

Wilayah	K-Means	DBSCAN	Hierarchical
W1	C1	C1	C1
W2	C2	C2	C2
W3	C2	Noise	C2
W4	C1	C1	C1
W5	C2	C2	C2

3.6 Analisis Mendalam

1. K-Means stabil dan cepat konvergen dalam 2 iterasi.
2. DBSCAN mampu mendeteksi data ekstrem (W3).
3. Hierarchical memberikan informasi urutan penggabungan wilayah.

Secara matematis, perbedaan utama terletak pada:

1. K-Means → berbasis centroid
2. DBSCAN → berbasis kepadatan
3. Hierarchical → berbasis struktur jarak bertingkat

3.7 Nilai Inovasi

Dengan dataset yang sama, penelitian ini membuktikan bahwa:

1. Hasil K-Means dan Hierarchical konsisten.
2. DBSCAN memberikan insight tambahan berupa deteksi anomali.
3. Integrasi ketiga metode memberikan pemetaan yang lebih komprehensif untuk kebijakan pengelolaan sampah.

4. KESIMPULAN

Penelitian ini bertujuan untuk mengidentifikasi pola produksi sampah melalui penerapan tiga algoritma *clustering*, yaitu *K-Means*, *DBSCAN*, dan *Hierarchical Clustering*. Berdasarkan hasil dan pembahasan, tujuan tersebut telah tercapai, di mana ketiga metode mampu mengelompokkan wilayah berdasarkan karakteristik produksi sampah dengan pendekatan yang berbeda namun saling melengkapi. Metode *K-Means* menunjukkan kemampuan segmentasi yang stabil dan efisien dalam memisahkan wilayah dengan produksi sampah tinggi dan rendah. *Hierarchical Clustering* memberikan gambaran struktur hubungan antarwilayah secara bertingkat, sehingga memudahkan pemahaman kedekatan karakteristik antar kelompok. Sementara itu, *DBSCAN* memberikan nilai tambah dengan kemampuannya mendeteksi wilayah yang memiliki karakteristik ekstrem sebagai anomali.

Temuan ini menunjukkan bahwa tidak terdapat satu metode yang secara mutlak paling unggul, melainkan pemilihan algoritma harus disesuaikan dengan tujuan analisis. Integrasi ketiga metode menghasilkan

pemetaan pola produksi sampah yang lebih komprehensif dibandingkan penggunaan satu algoritma tunggal. Dengan demikian, penelitian ini memberikan kontribusi dalam pendekatan analisis berbasis data untuk mendukung pengambilan keputusan di bidang pengelolaan sampah.

Secara praktis, hasil penelitian ini dapat dimanfaatkan oleh pemerintah daerah atau instansi pengelola lingkungan untuk:

1. Mengelompokkan wilayah prioritas pengelolaan sampah.
2. Mengidentifikasi wilayah dengan karakteristik produksi ekstrem yang memerlukan kebijakan khusus.
3. Menyusun strategi distribusi sumber daya dan armada pengangkutan secara lebih efektif.

Adapun prospek pengembangan penelitian selanjutnya dapat diarahkan pada:

1. Penggunaan dataset yang lebih besar dan bersifat *time series* untuk melihat dinamika pola produksi sampah dari waktu ke waktu.
2. Integrasi metode *clustering* dengan teknik *classification* atau *prediction* untuk mendukung sistem peringatan dini peningkatan volume sampah.
3. Penerapan algoritma berbasis kecerdasan buatan yang lebih adaptif seperti *ensemble clustering* atau pendekatan berbasis *deep learning*.
4. Pengujian validasi eksternal dengan melibatkan variabel sosial, ekonomi, dan perilaku masyarakat guna memperkaya interpretasi hasil kluster.

Berdasarkan literatur, kombinasi beberapa algoritma *clustering* dalam satu studi komparatif dapat meningkatkan reliabilitas interpretasi pola data kompleks (Jain, 2010; Han et al., 2011). Oleh karena itu, penelitian lanjutan disarankan untuk mengembangkan model hibrida atau sistem pendukung keputusan berbasis integrasi multi-algoritma agar hasil analisis semakin akurat dan aplikatif dalam konteks pengelolaan lingkungan berkelanjutan. Secara keseluruhan, penelitian ini memperlihatkan bahwa pendekatan komparatif berbasis *clustering* merupakan strategi yang efektif dalam memahami pola produksi sampah dan berpotensi menjadi landasan pengembangan sistem pengelolaan sampah berbasis data di masa depan.

REFERENSI

- [1] Agustono, H., & Gunawan, R. (2026). Pemetaan Kawasan Prioritas Pengelolaan Sampah Indonesia dengan Algoritma DBSCAN Mapping Priority Areas for Waste Management in Indonesia Using the DBSCAN Algorithm. *Jurnal Sistem dan teknologi Informasi*, 14(1), 101–109.
- [2] Akinola, S. O., & Oyabugbe, O. J. (2015). Accuracies and Training Times of Data Mining Classification Algorithms: An Empirical Comparative Study. *Journal of Software Engineering and Applications*, 8(September), 470–477.
- [3] Ester, M., Kriegel, H., Xu, X., & Miinchen, D.-. (2020). A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. *KDD-96 Proceedings*.
- [4] Fauzan, A., Fadillah, G., Fitria, A., Adriana, H., & Bariklana, M. (2024). Cluster Mapping of Waste Exposure Using DBSCAN Approach: Study of Spatial Patterns and Potential Distribution in Bantul Regency. *JOIV*, 8(May), 751–759.
- [5] Informasi, T., Nurrohman, M., & Sukmasetya, P. (2023). Jurnal Sistem Klusterisasi Volume Sampah Organik di Kota Magelang menggunakan K-Means Clustering System of Organic Waste Volume in Magelang City using K-Means. *Jurnal Teknologi Informasi Komuni kasi*, 5, 146–153.
- [6] Jain, A., & Rathi, K. (2025). A Comparative Analysis of DBSCAN, K-Means and Agglomerative Clustering Algorithms for Geospatial Data. *Proceedings of the International Conference on Recent Advancement and Modernization in Sustainable Intelligent Technologies & Applications*.
- [7] Lumbreras, M., Diarce, G., Martin-escudero, K., & Campos-celador, A. (2022). Design of district heating networks in built environments using GIS: A case study in Vitoria-Gasteiz, Spain. *Journal of Cleaner Production*, 349(March), 131491. Elsevier Ltd.
- [8] Nurfarid, A., Prasetyo, B., & Sukmasetya, P. (2023). K-Means Clustering Method for Determining Waste Transportation Routes to Landfill. *Jurnal Riset Informatika*, 5(3).
- [9] Putra, I. P., & Fadhillah, A. (2025). Perbandingan Metode K-Means dan Hierarchical Clustering dalam Pengelompokan Data Penduduk Miskin di Kabupaten Cianjur. *jurnal inovasi dan trend*, 3(1), 227–234.
- [10] Rahmawati, F., & Fallo, S. I. (2025). Leibniz: Jurnal Matematika. *Jurnal Matematika*, 5, 65–77.
- [11] Ramadhan, A., Achmad, F., Zulkarnain, I., & Aritsugi, M. (2025). Evaluation of K-Means, DBSCAN, and Hierarchical Clustering for Strategic Segmentation of Tourism SMEs in Rembang, Indonesia. *Jurnal Teknik Informatika (JUTIF)*, 6(3), 1605–1630.

-
- [12] Ruwandara, D., Jajuli, M., & Rizal, A. (2021). Analisis Algoritma K-Means Clustering Untuk Daerah Penyebaran Sampah di Kota Bekasi Data Sampah Kota Bekasi yang. *Journal of Information System*, 6(1), 56–63.
- [13] Santi, D., Maharai, W., Syahrullah, Nugraha, D. W., Muklis, B., & Agustinus, K. (2025). CLUSTERING DAERAH TERDAMPAK SAMPAH DI INDONESIA. *JurnalFORISTEK*, 15(1).
- [14] Syihabuddin, H., Jauhar, A., & Fitriani, R. (2025). APPLICATION OF DBSCAN FOR CLUSTERING SOCIETY BASED ON WASTE MANAGEMENT BEHAVIOR. *Journal of Mathematics and Its Applications*, 19(2), 961–972.
- [15] Zuhdi, M. R., Syihabuddin, H., Jauhar, A., Achmad, A., & Fernandes, R. (2025). Comparison Of Dbscan And K-Means Cluster Analysis With Path- Anova In Clustering Waste Management Behaviour Patterns Perbandingan Analisis Cluster Dbscan Dan K-Means Dengan Path-Anova Dalam Pengelompokan Pola Perilaku Pengelolaan. *Jurnal Teknik Informatika (JUTIF)*, 6(1), 105–112.