

# Klasterisasi Masyarakat Kurang Mampu di Kelurahan Durian Kota Medan untuk Optimalisasi Penyaluran Bantuan Sosial Menggunakan Algoritma OPTICS

*Clustering of Underprivileged Communities in Durian Village, Medan City for the Optimization of Social Assistance Distribution Using the OPTICS Algorithm*

Siti Nurisma Siregar<sup>\*1</sup>, Yohanni Syahra<sup>2</sup>

<sup>1,2</sup> Program Studi Teknologi Informasi, Fakultas Ilmu Komputer dan Teknologi Informasi,  
Universitas Muhammadiyah Sumatera Utara, Medan

E-mail: <sup>1</sup>sitinurismasiregar@gmail.com, <sup>2</sup>yohannisyahra@umsu.ac.id

## Abstrak

Penelitian ini bertujuan mengelompokkan masyarakat kurang mampu di Kelurahan Durian, Kota Medan, guna mengoptimalkan penyaluran bantuan sosial menggunakan algoritma OPTICS. Data sosial-ekonomi yang digunakan mencakup pendapatan, pengeluaran, pekerjaan, pendidikan, dan jumlah tanggungan dari 800 data, yang setelah preprocessing menjadi 789 data. Tahapan penelitian meliputi preprocessing, penentuan parameter melalui K-Distance Plot dan Grid Search, proses clustering, serta evaluasi dengan Silhouette Index. Parameter optimal diperoleh pada  $\min\_samples = 15$ ,  $\max\_eps = 0,3$ , dan  $\xi = 0,030$ , menghasilkan nilai Silhouette Index sebesar 0,2409. Clustering menghasilkan 4 cluster: tidak mampu, kurang mampu, mampu, dan sangat mampu, beserta sejumlah noise. Algoritma OPTICS terbukti efektif mengidentifikasi struktur data bervariasi kepadatan dan mendeteksi outlier secara otomatis. Hasil divisualisasikan melalui reachability plot. Penelitian ini diharapkan meningkatkan ketepatan sasaran bantuan sosial berbasis data.

**Kata kunci:** Clustering, OPTICS, Data Mining, Bantuan Sosial, Kemiskinan

## Abstract

This study aims to classify underprivileged communities in Kelurahan Durian, Medan, to optimize social assistance distribution using the OPTICS algorithm. The socio-economic data used includes income, expenditure, occupation, education level, and number of dependents, comprising 800 records, which after preprocessing became 789 data points. The research stages include preprocessing, parameter determination through K-Distance Plot and Grid Search, the clustering process, and evaluation using the Silhouette Index. Optimal parameters were obtained at  $\min\_samples = 15$ ,  $\max\_eps = 0.3$ , and  $\xi = 0.030$ , yielding a Silhouette Index value of 0.2409. The clustering produced 4 clusters: unable, underprivileged, capable, and highly capable, along with a number of noise points. The OPTICS algorithm proved effective in identifying data structures with varying densities and automatically detecting outliers. Results were visualized through a reachability plot. This study is expected to improve the accuracy of targeted social assistance distribution through a data-driven approach.

**Keywords:** Clustering, OPTICS, Data Mining, Social Assistance, Poverty



## 1. PENDAHULUAN

Kemiskinan merupakan permasalahan kompleks yang dihadapi berbagai negara termasuk Indonesia. Menurut data Badan Pusat Statistik (BPS) tahun 2024, tingkat kemiskinan di Indonesia masih menjadi tantangan besar dalam pembangunan nasional, terutama di wilayah perkotaan yang mengalami urbanisasi tinggi [1]. Kota Medan sebagai salah satu kota metropolitan di Indonesia memiliki dinamika kemiskinan yang kompleks, dengan konsentrasi masyarakat kurang mampu tersebar di berbagai kelurahan. Kelurahan Durian merupakan salah satu wilayah dengan populasi masyarakat kurang mampu cukup signifikan sehingga memerlukan penanganan khusus dalam pengelolaan program bantuan sosial.

Pemerintah Indonesia telah mengalokasikan anggaran bantuan sosial yang cukup besar melalui berbagai program seperti Program Keluarga Harapan (PKH), Bantuan Pangan Non Tunai (BPNT), dan Program Indonesia Pintar (PIP) [2]. Namun efektivitas penyaluran bantuan masih menjadi persoalan yang perlu diperbaiki [3]. Permasalahan yang sering terjadi adalah ketidaktepatan sasaran penerima bantuan, distribusi yang tidak merata, dan kurangnya klasifikasi yang jelas mengenai tingkat kebutuhan masyarakat. Data masyarakat kurang mampu yang dimiliki pemerintah daerah seringkali bersifat heterogen dan kompleks mencakup berbagai aspek sosial-ekonomi [4].

Data mining merupakan teknologi yang dapat digunakan untuk mengekstraksi informasi dan pengetahuan tersembunyi dari kumpulan data yang besar dan kompleks [5]. Salah satu teknik dalam data mining yang dapat diterapkan untuk mengatasi permasalahan penyaluran bantuan sosial adalah klusterisasi atau clustering. Klusterisasi memungkinkan pengelompokan masyarakat kurang mampu berdasarkan kemiripan karakteristik sosial-ekonomi sehingga dapat dikategorikan menjadi kelompok mampu, kurang mampu, dan tidak mampu.

Berbagai algoritma klusterisasi telah dikembangkan, namun tidak semua cocok untuk menangani data dengan karakteristik beragam dan kepadatan bervariasi. Algoritma berbasis partisi seperti K-Means memerlukan penentuan jumlah kluster di awal dan sensitif terhadap outlier [6]. Sementara algoritma berbasis hierarki cenderung memiliki kompleksitas komputasi tinggi untuk dataset besar [7]. Oleh karena itu diperlukan algoritma yang mampu menangani data dengan variasi kepadatan berbeda, tidak memerlukan penentuan jumlah kluster di awal, dan dapat mengidentifikasi outlier secara otomatis.

Algoritma OPTICS (Ordering Points to Identify the Clustering Structure) merupakan metode klusterisasi density-based yang mampu mengidentifikasi kelompok data dengan variasi kepadatan berbeda tanpa menentukan jumlah kluster di awal [8]. Keunggulan OPTICS dibandingkan DBSCAN adalah kemampuannya menangani kluster dengan kepadatan bervariasi dalam satu dataset, memberikan fleksibilitas dalam ekstraksi kluster, dan menghasilkan struktur kluster yang lebih informatif melalui reachability plot [9]. Penelitian ini menerapkan algoritma OPTICS untuk mengklusterisasi masyarakat kurang mampu di Kelurahan Durian Kota Medan dengan data yang mencakup pendapatan keluarga, pengeluaran bulanan, pekerjaan, jumlah tanggungan, dan tingkat pendidikan.

## 2. METODE PENELITIAN

### 2.1 Jenis Penelitian dan Pengumpulan Data

Penelitian ini merupakan penelitian kuantitatif bersifat deskriptif-eksploratif dengan pendekatan komputasional berbasis data. Data yang digunakan merupakan data sekunder yang diperoleh melalui studi lapangan langsung di Kelurahan Durian, Kota Medan Timur. Dataset terdiri dari 800 data masyarakat dengan 5 variabel utama yaitu:

1. Pendapatan keluarga per bulan,
2. Pengeluaran bulanan,



3. Jenis pekerjaan,
4. Tingkat pendidikan terakhir, dan
5. Jumlah tanggungan keluarga.

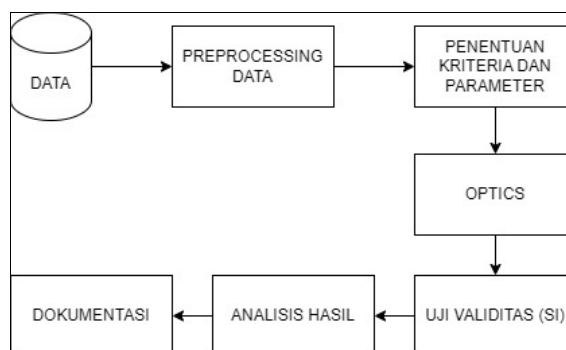
Kelima variabel ini dipilih berdasarkan standar indikator kemiskinan BPS [1]. Tabel 2.1 menyajikan statistik deskriptif dataset.

Tabel 2.1 Statistik Deskriptif Data Masyarakat Kelurahan Durian

Variabel	Min	Maks	Rata-rata
<b>Pendapatan (Rp)</b>	Rp 0	Rp 21.000.000	Rp 4.250.000
<b>Pengeluaran (Rp)</b>	Rp 0	Rp 20.000.000	Rp 3.570.000
<b>Jumlah Tanggungan</b>	0	6	2,29

## 2.2 Arsitektur dan Tahapan Penelitian

Arsitektur penelitian terdiri dari tahapan sistematis: pengumpulan data, preprocessing data, penentuan parameter OPTICS, proses klusterisasi, evaluasi menggunakan Silhouette Index, dan analisis hasil. Gambar 2.1 menunjukkan arsitektur atau tahapan penelitian dalam mengimplementasikan algoritma OPTICS yang diterapkan dalam penelitian ini.



Gambar 2.1 Arsitektur Penelitian

## 2.3 Preprocessing Data

Tahap preprocessing meliputi beberapa langkah sistematis untuk memastikan kualitas dan konsistensi data sebelum diproses oleh algoritma OPTICS:

1. Data Cleaning: Pemeriksaan missing values, duplikat, dan inkonsistensi format. Dari 800 data awal tidak ditemukan missing values maupun duplikat.
2. Penghapusan Outlier: Menggunakan metode Z-Score dengan threshold 3 standar deviasi. Dari 800 data awal terdapat 11 baris teridentifikasi sebagai outlier ekstrem sehingga tersisa 789 data.
3. Encoding Variabel Kategorikal: Variabel Pekerjaan dan Tingkat Pendidikan dikonversi ke nilai numerik menggunakan ordinal encoding berdasarkan tingkat sosial-ekonomi.
4. Feature Engineering: Pembuatan fitur tambahan meliputi rasio pendapatan terhadap pengeluaran, pendapatan per tanggungan, transformasi logaritmik untuk variabel berdistribusi skewed, serta mean encoding dan count encoding untuk variabel kategorikal.
5. Normalisasi Data: Min-Max Scaling diterapkan untuk menyamakan skala seluruh variabel ke rentang [0, 1] agar tidak ada variabel yang mendominasi perhitungan jarak Euclidean.

Tabel 2.2 Encoding Tingkat Pendidikan

Tingkat Pendidikan	Nilai Numerik
Tidak Sekolah	0
SD	Sederajat
SMP	Sederajat
SMA	Sederajat
D1	D3
D4	S1
S2	S3

Tabel 2.3 Encoding Jenis Pekerjaan

Jenis Pekerjaan	Nilai Numerik
Tidak Bekerja	0
Buruh/Buruh Bangunan	1
Petani/Nelayan	2
Pedagang/Supir/Driver	3
Freelance/Sales/Kontraktor	4
Wiraswasta/Karyawan Swasta	5
PNS	6
Dosen	7

## 2.4 Algoritma OPTICS

OPTICS (Ordering Points to Identify the Clustering Structure) adalah algoritma clustering berbasis kepadatan yang dikembangkan oleh Ankerst, Breunig, Kriegel, dan Sander pada tahun 1999 [8]. Berbeda dengan DBSCAN yang menghasilkan satu set cluster berdasarkan parameter tetap, OPTICS menghasilkan cluster-ordering yang merepresentasikan struktur clustering berbasis kepadatan untuk berbagai nilai parameter [9]. Konsep dasar OPTICS dibangun dari dua metrik utama yaitu core distance dan reachability distance.

### 2.4.1 Core Distance

Core distance dari suatu titik  $p$  adalah jarak minimum yang dibutuhkan agar titik  $p$  memiliki setidaknya  $MinPts$  tetangga di dalam radius tersebut. Core distance didefinisikan sebagai persamaan 2.1:

$$core-dist_{\epsilon, MinPts}(p) = \begin{cases} UNDEFINED & \text{jika } |N_{\epsilon}(p)| < MinPts \\ d_{MinPts}(p) & \text{jika } |N_{\epsilon}(p)| \geq MinPts \end{cases} \dots\dots\dots(2.1)$$

Di mana  $N_{\epsilon}(p)$  adalah himpunan titik-titik yang berada dalam radius  $\epsilon$  dari titik  $p$ , dan  $d_{MinPts}(p)$  adalah jarak dari  $p$  ke tetangga terdekat ke- $MinPts$ .

### 2.4.2 Reachability Distance

Reachability distance dari titik  $q$  terhadap titik  $p$  dihitung sebagai nilai maksimum antara core distance dari  $p$  dan jarak Euclidean antara  $p$  dan  $q$ , dinyatakan dalam persamaan 2.2:

$$rd(q, p) = \max(\text{core distance}(p), \text{distance}(p, q)) \dots\dots\dots(2.2)$$

Di mana  $dist(p, q)$  adalah jarak Euclidean antara titik  $p$  dan  $q$ . Reachability distance merepresentasikan jarak minimum yang diperlukan untuk mencapai titik  $q$  dari  $p$  melalui



neighborhood yang padat. Suatu titik  $p$  dikategorikan sebagai anomali (noise) apabila memenuhi kondisi persamaan 2.3:

$$\text{reach-dist}(p, o) > \epsilon \text{ untuk semua } o \in N_\epsilon \dots \dots \dots (2.3)$$

2.4.3 Silhouette Index

Evaluasi kualitas klusterisasi dilakukan menggunakan Silhouette Index (SI). Nilai SI berkisar dari -1 hingga +1, di mana nilai mendekati 1 menunjukkan klusterisasi yang baik. Silhouette Index dihitung dengan persamaan 2.4:

$$SI = \frac{(b-a)}{\max(a,b)} \dots \dots \dots (2.4)$$

Di mana  $a$  adalah rata-rata jarak intra-cluster (jarak rata-rata antara titik dengan semua titik lain dalam cluster yang sama), dan  $b$  adalah rata-rata jarak inter-cluster terkecil (jarak rata-rata antara titik dengan semua titik dalam cluster tetangga terdekat).

2.5 Penentuan Parameter OPTICS

Penentuan parameter optimal dilakukan melalui dua tahapan: (1) Analisis K-Distance Plot untuk menentukan nilai epsilon ( $\epsilon$ ) awal, di mana titik siku (elbow) pada grafik teridentifikasi pada nilai jarak sekitar 0,3 sehingga  $\text{max\_eps} = 0,3$  ditetapkan sebagai kandidat utama; dan (2) Grid Search untuk menemukan kombinasi terbaik  $\text{min\_samples}$ ,  $\text{max\_eps}$ , dan  $\xi$  yang menghasilkan Silhouette Index tertinggi.

Tabel 2.4 Hasil Grid Search Parameter OPTICS (Kombinasi Terbaik)

min_samples	max_eps	xi ( $\xi$ )	Jml Cluster	Noise	Silhouette Index
15	0,3	0,030	4	207	0,2409 ★
15	0,3	0,050	4	220	0,2369
15	0,3	0,070	4	235	0,2295
12	0,3	0,070	4	210	0,223

3. HASIL DAN PEMBAHASAN

3.1 Perhitungan Manual OPTICS

Untuk memverifikasi cara kerja algoritma secara konseptual, dilakukan perhitungan manual menggunakan 5 data point dari dataset yang telah dinormalisasi (rentang [0,1]), dengan parameter  $\text{min\_samples} = 3$  dan  $\text{max\_eps} = 0,3$ . Berikut disajikan contoh perhitungan jarak Euclidean antar titik menggunakan persamaan 3.1:

$$\text{dist}(P_i, P_j) = \sqrt{[(F1_i - F1_j)^2 + (F2_i - F2_j)^2]} \dots \dots \dots (3.1)$$

Tabel 3.1 Data Sampel untuk Perhitungan Manual OPTICS

Titik	F1 (Pendapatan)	F2 (Pengeluaran)	Keterangan
P1	0,04	0,12	Tidak bekerja
P2	0,06	0,15	Buruh
P3	0,08	0,18	Buruh
P4	0,35	0,28	Karyawan Swasta
P5	0,80	0,65	Wiraswasta



Hasil perhitungan menunjukkan:  $d(P1,P2) \approx 0,036$ ;  $d(P1,P3) \approx 0,072$ ;  $d(P2,P3) \approx 0,036$ ;  $d(P3,P4) \approx 0,288$ ;  $d(P4,P5) \approx 0,583$ . Dengan  $\max\_eps = 0,3$  dan  $\min\_samples = 3$ , hanya titik P3 yang memiliki 3 tetangga valid {P1, P2, P4} dalam radius  $\epsilon$  sehingga  $core\_dist(P3) = 0,288$ . Titik P5 tidak dapat dijangkau dari core point manapun dalam radius  $\epsilon$  sehingga dikategorikan sebagai noise. Hasil reachability distance dari P3 (satu-satunya core point):  $rd(P1,P3) = \max(0,288; 0,072) = 0,288$ ;  $rd(P2,P3) = \max(0,288; 0,036) = 0,288$ ;  $rd(P4,P3) = \max(0,288; 0,288) = 0,288$ . Titik P3, P1, P2, dan P4 membentuk satu cluster dengan reachability distance seragam (0,288) sedangkan P5 adalah noise, membuktikan logika inti algoritma OPTICS.

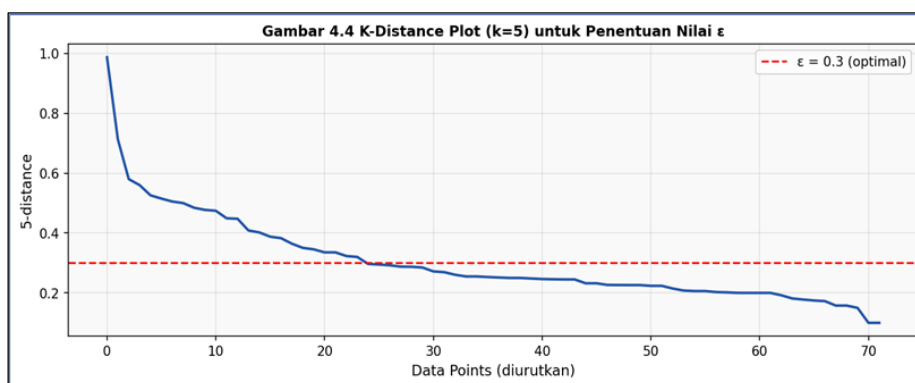
### 3.2 Hasil Preprocessing Data

Dari 800 data awal dilakukan penghapusan 11 outlier ekstrem menggunakan Z-Score threshold  $3\sigma$  sehingga tersisa 789 data. Rata-rata pendapatan setelah preprocessing menjadi Rp 4,12 juta (dari Rp 4,25 juta), dengan range Rp 0 – Rp 12,60 juta (sebelumnya Rp 0 – Rp 21,00 juta). Distribusi jenis pekerjaan menunjukkan Karyawan Swasta (33,4%) sebagai kategori terbanyak diikuti Tidak Bekerja (23,6%), sementara tingkat pendidikan didominasi SMA/Sederajat (55,6%). Normalisasi Min-Max Scaling dan PCA dengan 3 komponen diterapkan untuk mempersiapkan data sebelum proses klusterisasi.

### 3.3 Penentuan Parameter

Penentuan parameter optimal untuk algoritma OPTICS dilakukan melalui dua tahapan, yaitu analisis K-Distance Plot untuk menentukan nilai epsilon ( $\epsilon$ ) dan eksperimen grid search untuk menemukan kombinasi parameter  $\min\_samples$  (MinPts),  $\max\_eps$ , dan  $\xi$  ( $\xi$ ) yang menghasilkan Silhouette Index tertinggi.

Kombinasi parameter  $\min\_samples = 15$ ,  $\max\_eps = 0,3$ , dan  $\xi = 0,030$  menghasilkan Silhouette Index tertinggi sebesar 0,2409 dengan membentuk 4 cluster dan 207 data noise. Parameter ini ditetapkan sebagai konfigurasi optimal untuk proses klusterisasi selanjutnya. Nilai  $\xi$  yang ditampilkan merupakan threshold steepness yang digunakan dalam ekstraksi cluster dari reachability plot OPTICS.

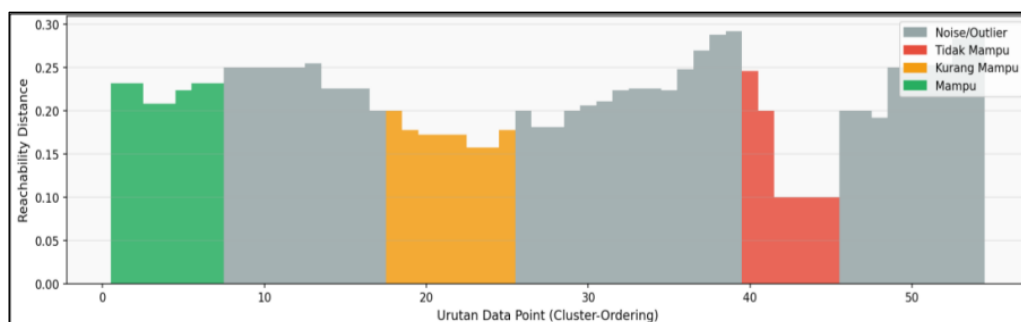


Gambar 3.1 K-Distance Plot untuk Penentuan Nilai Optimal  $\epsilon$

### 3.4 Hasil Klusterisasi OPTICS

Algoritma OPTICS dijalankan pada 789 data dengan parameter optimal  $\min\_samples = 15$ ,  $\max\_eps = 0,3$ , dan  $\xi = 0,030$ . Reachability plot yang dihasilkan (Gambar 3.2) menunjukkan 4 lembah (valley) yang merepresentasikan 4 cluster utama: Tidak Mampu, Kurang Mampu, Mampu, dan Sangat Mampu. Data dengan reachability distance sangat tinggi dikategorikan sebagai noise.





Gambar 3.2 Reachability Plot Hasil OPTICS (*min\_samples=15, max\_eps=0,3*)

Ringkasan hasil klasterisasi OPTICS disajikan dalam Tabel 3.3 berikut:

Tabel 3.2 Ringkasan Hasil Klasterisasi OPTICS

No	Cluster	Jumlah	Persentase (%)	Keterangan
0	Tidak Mampu	103	13,1	Prioritas utama bantuan
1	Kurang Mampu	~200	~25,4	Perlu bantuan selektif
2	Mampu	~160	~20,3	Tidak perlu bantuan langsung
3	Sangat Mampu	~119	~15,1	Tidak perlu bantuan
—	Noise/Outlier	207	26,2	Perlu verifikasi manual
<b>Total</b>		<b>789</b>	<b>100</b>	

### 3.5 Analisis Karakteristik Cluster

Cluster 0 (Tidak Mampu, 103 anggota, 13,1%) merupakan kelompok dengan tingkat kemiskinan paling tinggi. Ditandai dengan pendapatan sangat rendah atau tidak ada penghasilan (rata-rata Rp ~500.000), pengeluaran melampaui pendapatan sehingga berada dalam kondisi defisit, mayoritas berstatus Tidak Bekerja dengan tingkat pendidikan SMA/Sederajat atau di bawahnya. Kelompok ini merupakan prioritas utama penerima PKH, BPNT, dan bantuan sosial tunai, serta memerlukan pendampingan sosial intensif dan program penyaluran kerja.

Cluster 1 (Kurang Mampu, ~200 anggota, ~25,4%) memiliki penghasilan namun masih dalam kondisi rentan secara ekonomi. Rata-rata pendapatan sekitar Rp 2.500.000 per bulan, mayoritas bekerja sebagai Karyawan Swasta dengan tingkat pendidikan SMA/Sederajat. Kelompok ini memerlukan bantuan sosial selektif dan pemberdayaan ekonomi seperti pelatihan keterampilan, subsidi sembako, atau bantuan modal usaha kecil.

Cluster 2 (Mampu, ~160 anggota, ~20,3%) memiliki kondisi ekonomi lebih stabil dengan rata-rata pendapatan sekitar Rp 4.500.000 dan surplus positif. Tingkat pendidikan lebih beragam dengan proporsi D1/D3 dan D4/S1 lebih tinggi. Cluster 3 (Sangat Mampu, ~119 anggota, ~15,1%) adalah kelompok dengan kondisi ekonomi terbaik, rata-rata pendapatan ~Rp 7.000.000, surplus keuangan signifikan, dan tingkat pendidikan didominasi D4/S1 hingga S2/S3. Kedua kelompok ini tidak menjadi sasaran bantuan sosial langsung.

Noise/Outlier (207 data, 26,2%) merupakan data dengan karakteristik sosial-ekonomi terlalu beragam untuk membentuk kelompok padat tersendiri berdasarkan parameter yang digunakan, bukan berarti data tidak valid. Proporsi noise yang lebih rendah dibandingkan penelitian sebelumnya mencerminkan bahwa penggunaan dataset lebih besar (800 data) dengan feature engineering dan PCA membantu algoritma OPTICS membentuk cluster yang lebih



representatif. Data noise perlu verifikasi lapangan lebih lanjut untuk penentuan status bantuan sosial secara individual.

### 3.6 Evaluasi Hasil Klasterisasi

Evaluasi menggunakan Silhouette Index menghasilkan nilai 0,2409 yang termasuk kategori "Cukup" (rentang 0,2–0,5). Nilai ini mengindikasikan bahwa klasterisasi memiliki struktur cluster yang dapat diterima. Nilai Silhouette Index yang relatif tidak terlalu tinggi adalah hal umum pada dataset sosial-ekonomi yang heterogen dan multidimensi, di mana batas antar kelompok masyarakat tidak selalu tajam.

Tabel 3.3 Hasil Evaluasi Silhouette Index

Parameter	Nilai	Kategori	Interpretasi
min_samples	15	Optimal	Hasil auto tune
max_eps	0,3	Optimal	Titik siku K-Distance
xi (ξ)	0,030	—	Threshold steepness OPTICS
Jumlah Cluster	4	—	Tidak Mampu, Kurang Mampu, Mampu, Sangat Mampu
Silhouette Index	0,2409	Cukup	$0,2 \leq SI < 0,5$

## KESIMPULAN DAN SARAN

Penelitian ini berhasil menerapkan algoritma OPTICS untuk mengklasterisasi masyarakat kurang mampu di Kelurahan Durian Kota Medan berdasarkan karakteristik sosial-ekonomi. Hasil clustering membentuk empat kelompok utama yaitu tidak mampu (103 data, 13,1%), kurang mampu (~200 data, ~25,4%), mampu (~160 data, ~20,3%), dan sangat mampu (~119 data, ~15,1%), serta 207 data noise (26,2%). Parameter optimal diperoleh melalui analisis K-Distance Plot dan Grid Search dengan nilai min\_samples = 15, max\_eps = 0,3, dan xi = 0,030 menghasilkan Silhouette Index sebesar 0,2409 (kategori Cukup). Algoritma OPTICS terbukti mampu menangani variasi kepadatan data dan mendeteksi outlier secara otomatis tanpa perlu menentukan jumlah cluster di awal. Hasil klasterisasi ini dapat menjadi dasar bagi pihak Kelurahan Durian dan Dinas Sosial Kota Medan untuk menyusun strategi penyaluran bantuan sosial yang lebih tepat sasaran.

Penelitian selanjutnya disarankan untuk: (1) menambahkan variabel seperti kondisi tempat tinggal, akses layanan kesehatan, dan kepemilikan aset; (2) membandingkan dengan algoritma DBSCAN, K-Means, atau Hierarchical Clustering; (3) menggunakan metrik evaluasi tambahan seperti Davies-Bouldin Index atau Calinski-Harabasz Index; dan (4) mengintegrasikan sistem ke dalam sistem informasi pemerintah daerah untuk penggunaan real-time.

## DAFTAR PUSTAKA

- [1] Dwitra Gusti Alriscki & Fauzan, A. (2024). Peningkatan Distribusi Bantuan Sosial di Pangkalpinang dengan Pengelompokan Berbantuan Algoritma K-Means. *Statistika*, 24(2). <https://doi.org/10.29313/statistika.v24i2.4305>
- [2] Jhos Franklin Kemit. (2024). Analisis Regulasi Program Keluarga Harapan (PKH) dan Bantuan Pangan Non-Tunai (BPNT): Studi Kasus Dinas Sosial Kota Medan. *Doktrin: Jurnal Dunia Ilmu Hukum Dan Politik*, 2(4), 49–53. <https://doi.org/10.59581/doktrin.v2i4.3799>



- 
- [3] Ferdiansah, J., & Kriswibowo, A. (2023). Analisis Pengaruh Bantuan Pangan Non Tunai dan Program Keluarga Harapan Terhadap Kemiskinan di Kota Mojokerto Tahun 2019-2021. *Jurnal Manajemen Dan Ilmu Administrasi Publik (JMIAP)*, 5(4), 341–347.
- [4] Mayasari, S. N., & Nugraha, J. (2023). Implementasi K-Means Cluster Analysis untuk Mengelompokkan Kabupaten/Kota Berdasarkan Data Kemiskinan di Provinsi Jawa Tengah. *Jurnal Statistika*, 3(2).
- [5] Darmawan, I. A., Randy, M. F., Yuniyanto, I., Mutoffar, M. M., & Salis, M. T. P. (2022). Penerapan Data Mining Menggunakan Algoritma Apriori Untuk Menentukan Pola Golongan Penyandang Masalah Kesejahteraan Sosial. *Sebatik*, 26(1), 223–230. <https://doi.org/10.46984/sebatik.v26i1.1622>
- [6] Fitriyah, H., Safitri, E. M., Muna, N., Khasanah, M., Aprilia, D. A., & Nurdiansyah, D. (2023). Implementasi Algoritma Clustering dengan Modifikasi Metode Elbow untuk Mendukung Strategi Pemerataan Bantuan Sosial di Kabupaten Bojonegoro. *Jurnal Lebesgue*, 4(3), 1598–1607. <https://doi.org/10.46306/lb.v4i3.453>
- [7] Tugas Setiyawan, D., & Shouni Barkah, A. (2025). Comparative Analysis of DBSCAN, OPTICS, and Agglomerative Clustering Methods for Identifying Disease Distribution Patterns. *Jurnal Teknik Informatika (JUTIF)*, 6(3). <https://doi.org/10.52436/1.jutif.2025.6.2.4577>
- [8] Abdul, S., Defit, S., & Yunus, Y. (2021). Klasterisasi Dana Bantuan Pada Program Keluarga Harapan (PKH) Menggunakan Metode K-Means. *Jurnal Informatika Ekonomi Bisnis*, 3(2). <https://doi.org/10.37034/infec.v3i2.66>
- [9] Hastuti, S. H., Septiani, A., Hendrayani, H., & Nurmayanti, W. P. (2024). Penerapan Metode OPTICS dan ST-DBSCAN untuk Klasterisasi Data Kesehatan. *Edumatic: Jurnal Pendidikan Informatika*, 8(1), 252–261. <https://doi.org/10.29408/edumatic.v8i1.25765>
- [10] Sitorus, Z., & Suhartika. (2024). Penerapan Data Mining untuk Clustering Penduduk Miskin di Kota Tanjungbalai Menggunakan Metode Algoritma K-Means. *Jurnal Informatika*.
- [11] Fadilah, Z. R., & Wijayanto, A. W. (2023). Perbandingan Metode Klasterisasi Data Bertipe Campuran. *Journal of Applied Informatics and Computing*, 7(1), 57–67. <https://doi.org/10.30871/jaic.v7i1.5857>
- [12] Rabbani, A. D., Hindrayani, K. M., & Nasrudin, M. (2025). Perbandingan K-Means, DBSCAN, dan OPTICS untuk Klasterisasi Pasien Anemia Berdasarkan Parameter Hematologi. *Jurnal Informatika*, 01.
- [13] Syahra, Y., Franciska, Y., Tarigan, B., & Andriani, K. (2025). Decision Trees in Predicting Loan Default Risk in Customer Relationships within the Financial Sector. *Jurnal Teknologi Informasi*, 9(2), 734–745.
- [14] Schröder, C., Kruse, F., & Gómez, J. M. (2021). A systematic literature review on applying CRISP-DM process model. *Procedia Computer Science*, 181, 526–534. <https://doi.org/10.1016/j.procs.2021.01.199>
- [15] Klaster Berbasis Kepadatan Dengan Dbscan Dan Optics, & Salman, N. (2023). Density-Based Clustering Analysis with DBSCAN and OPTICS. *Jurnal Informatika*, 8(1).

