



DOI: <https://doi.org/10.38035/dit.v4i2>
<https://creativecommons.org/licenses/by/4.0/>

Received: 01 October 2026 | Revised: 06 October 2026 | Accepted: 13 October 2026

Klasterisasi Masyarakat Kurang Mampu dalam Penyaluran Bansos Menggunakan Algoritma OPTICS di Desa Lubuk Rotan

Haliza Silviya¹, Guidio Leonarde Ginting²

¹Sekolah Tinggi Ilmu Manajemen Sukma, Kota Medan, Indonesia, halizasilviya@gmail.com

²Sekolah Tinggi Ilmu Manajemen Sukma, Kota Medan, Indonesia

Corresponding Author: halizasilviya@gmail.com¹

Abstract: *The distribution of social assistance requires information that can objectively describe the socioeconomic conditions of the community. This study aims to cluster low-income households in Lubuk Rotan Village based on socioeconomic characteristics using the Ordering Points to Identify the Clustering Structure (OPTICS) algorithm and to generate supporting information for the distribution of social assistance. The dataset consists of 639 community records with attributes including income, number of dependents, asset ownership, and home ownership status. Data processing was performed using the Knowledge Discovery in Databases (KDD) framework, which includes selection, preprocessing, transformation, data mining, and evaluation. The min_samples test was conducted at values of 20, 25, 30, 35, and 40. The test results showed that min_samples = 25 produced the highest Silhouette Index of 0.833517, which is classified as “very good.” With these parameters, OPTICS generated 7 main clusters and 1 noise group consisting of 55 data points. The clustering results reveal differences in socioeconomic characteristics among the groups. These findings can serve as supporting information for the Lubuk Rotan village government in understanding the characteristics of the community and determining the distribution of social assistance based on socioeconomic conditions.*

Keywords: *Clustering, Low-Income Communities, Social Assistance, OPTICS, Silhouette Index.*

Abstrak: Penyaluran bantuan sosial memerlukan informasi yang dapat menggambarkan kondisi sosial ekonomi masyarakat secara objektif. Penelitian ini bertujuan mengelompokkan masyarakat kurang mampu di Desa Lubuk Rotan berdasarkan karakteristik sosial ekonomi menggunakan algoritma *Ordering Points to Identify the Clustering Structure* (OPTICS) serta menghasilkan informasi pendukung dalam penyaluran bantuan sosial. Data yang digunakan berjumlah 639 data masyarakat dengan atribut pendapatan, jumlah tanggungan, kepemilikan aset, dan status kepemilikan rumah. Pengolahan data dilakukan menggunakan tahapan *Knowledge Discovery in Database* (KDD), meliputi *selection*, *preprocessing*, *transformation*, *data mining*, dan *evaluation*. Pengujian min_samples dilakukan pada nilai 20, 25, 30, 35, dan 40. Hasil pengujian menunjukkan min_samples=25 menghasilkan *Silhouette Index* tertinggi sebesar 0,833517 dengan kriteria sangat baik. Dengan parameter tersebut, OPTICS

menghasilkan 7 *cluster* utama dan 1 kelompok *noise* sebanyak 55 data. Hasil clustering menunjukkan perbedaan karakteristik sosial ekonomi antar kelompok. Hasil pengelompokan dapat menjadi informasi pendukung bagi pemerintah Desa Lubuk Rotan dalam memahami karakteristik masyarakat dan mempertimbangkan penyaluran bantuan sosial berdasarkan kondisi sosial ekonomi.

Kata Kunci: Klasterisasi, Masyarakat kurang mampu, Bantuan Sosial, OPTICS, *Silhouette Index*.

PENDAHULUAN

Desa Lubuk Rotan merupakan salah satu desa di Kecamatan Perbaungan, Kabupaten Serdang Bedagai, Sumatera Utara. Desa ini memiliki luas wilayah sekitar 390 Ha, dengan jumlah penduduk yang terdiri atas 858 kepala keluarga dan 2.673 jiwa yang tersebar di 5 dusun. Sebagian besar masyarakat Desa Lubuk Rotan bermata pencaharian pada sektor pertanian, perkebunan, perdagangan kecil, dan pekerjaan informal lainnya. Keberagaman mata pencaharian tersebut mencerminkan adanya perbedaan kondisi sosial ekonomi antar keluarga yang menjadi salah satu karakteristik keberagaman masyarakat yang menunjukkan bahwa tingkat kesejahteraan setiap keluarga tidak sepenuhnya sama. Perbedaan kondisi tersebut berkaitan dengan pelaksanaan berbagai program kesejahteraan sosial oleh pemerintah, salah satunya melalui program bantuan sosial.

Bantuan sosial (bansos) merupakan salah satu program pemerintah yang diberikan dalam berbagai bentuk bantuan berupa uang, barang, maupun kebutuhan pokok yang disalurkan melalui beberapa skema sesuai dengan kebijakan pemerintah. Dalam pelaksanaannya, pemerintah memiliki berbagai program bantuan sosial, di antaranya Program Keluarga Harapan (PKH), Bantuan Pangan Non Tunai (BPNT), Bantuan Langsung Tunai (BLT), serta beberapa program bantuan lainnya yang disesuaikan dengan kondisi dan kebutuhan masyarakat (Kementerian Sosial Republik Indonesia, 2026). Seiring dengan perkembangan kebijakan pemerintah, program bantuan sosial terus mengalami penyesuaian baik dari segi mekanisme maupun bentuk bantuan yang diberikan.

Proses penentuan penerima bantuan sosial dapat didukung dengan metode *clustering*, seperti algoritma K-Means, untuk mengelompokkan masyarakat berdasarkan layak atau tidak layak dalam menerima bantuan. (Musthafa et al., 2025). Hal tersebut menunjukkan bahwa metode *clustering* dapat dimanfaatkan sebagai pendekatan dalam mengolah data masyarakat dan memberikan informasi pendukung dalam proses seleksi penerima bantuan sosial.

Masyarakat yang menjadi calon penerima bantuan memiliki karakteristik yang berbeda-beda, seperti tingkat pendapatan, jenis pekerjaan, jumlah tanggungan keluarga, kondisi tempat tinggal, maupun kepemilikan aset. Keberagaman karakteristik tersebut menyebabkan data masyarakat menjadi lebih bervariasi, sehingga proses pengelompokan masyarakat yang memiliki kondisi sosial ekonomi serupa menjadi lebih kompleks. Akibatnya, kelompok masyarakat yang memiliki karakteristik serupa menjadi lebih sulit diidentifikasi, terutama ketika data memiliki tingkat kepadatan yang berbeda. Kondisi ini memerlukan metode klasterisasi yang mampu mengelompokkan data berdasarkan tingkat kemiripan karakteristik sekaligus menangani variasi kepadatan data yang terbentuk. Salah satu metode klasterisasi yang memiliki kemampuan tersebut adalah algoritma *Ordering Points To Identify the Clustering Structure* (OPTICS).

Algoritma *Ordering Points To Identify the Clustering Structure* (OPTICS) adalah salah satu metode klasterisasi berbasis kepadatan yang mampu mengidentifikasi struktur kelompok data dengan tingkat kepadatan yang bervariasi tanpa harus menentukan jumlah *cluster* di awal (Mustikasari & Salman, 2023). Selain itu, OPTICS juga mampu mengenali data yang tidak

termasuk ke dalam kelompok tertentu (*noise*), sehingga sesuai diterapkan pada data masyarakat yang memiliki karakteristik sosial ekonomi yang beragam (Ankerst et al., 1999).

Penelitian terdahulu menunjukkan bahwa OPTICS telah diterapkan pada berbagai permasalahan *clustering*. Salah satunya pengelompokan kemiskinan multidimensi di Indonesia dan menunjukkan bahwa algoritma tersebut dapat menghasilkan kualitas *cluster* yang baik, meskipun terdapat data yang dikategorikan sebagai *noise* (Supardi et al., 2026). Penelitian lain juga menunjukkan bahwa OPTICS mampu membentuk beberapa *cluster* sekaligus mengidentifikasi data *noise* pada data dengan karakteristik kepadatan yang berbeda (Nurhalizah et al., 2025). Selain itu, penerapan OPTICS pada data *log server web* memperoleh hasil yang menunjukkan kemampuan algoritma dalam mengidentifikasi pola dan aktivitas anomali (Buana et al., 2026).

Berdasarkan penelitian terdahulu, algoritma OPTICS telah banyak digunakan pada berbagai bidang, namun penerapannya dalam pengelompokan masyarakat kurang mampu pada tingkat desa masih jarang dilakukan, khususnya dalam mendukung penyaluran bantuan sosial berbasis kondisi sosial ekonomi masyarakat. Oleh karena itu, penelitian ini dilakukan dengan tujuan membantu pihak desa dalam mengelompokkan masyarakat berdasarkan kondisi sosial ekonomi sebagai pendukung penentuan penerima bantuan sosial yang lebih tepat sasaran.

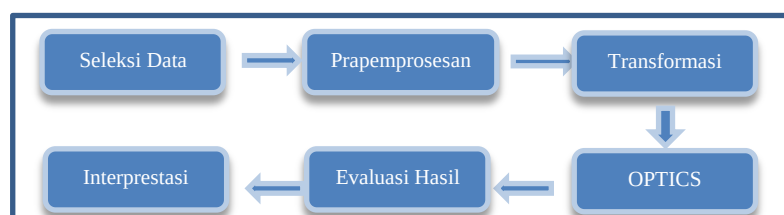
METODE

Penelitian ini menggunakan pendekatan kuantitatif dengan metode data mining untuk mengelompokkan masyarakat berdasarkan karakteristik sosial ekonomi. *Data mining* merupakan proses pengolahan dan analisis data secara otomatis untuk memperoleh informasi tersembunyi serta menemukan pola baru yang bermanfaat dari sejumlah besar data sehingga dapat menghasilkan pengetahuan yang berguna (Zhou, 2022).

Objek penelitian adalah masyarakat di Desa Lubuk Rotan, Kecamatan Perbaungan, Kabupaten Serdang Bedagai, Sumatera Utara. Data penelitian diperoleh dari Pemerintah Desa Lubuk Rotan dalam bentuk Data Tunggal Sosial Ekonomi Nasional (DTSEN). Penelitian dilakukan pada April sampai Agustus 2026 di Desa Lubuk Rotan. Pengumpulan data dilakukan melalui dokumentasi dan wawancara dengan perangkat desa. Dokumentasi digunakan untuk memperoleh data sosial ekonomi masyarakat, sedangkan wawancara digunakan untuk memperoleh informasi pendukung yang berkaitan dengan pendataan dan kondisi sosial ekonomi masyarakat.

Data awal yang tersedia pada administrasi desa berjumlah 858 Kepala Keluarga (KK), sedangkan data DTSEN yang diperoleh dan dapat digunakan dalam penelitian berjumlah 639 KK. Seluruh 639 data tersebut digunakan sebagai populasi sekaligus sampel penelitian dengan teknik total sampling. Atribut yang digunakan dalam proses clustering terdiri atas pendapatan keluarga, jumlah tanggungan, status kepemilikan rumah, dan kepemilikan aset. Pendapatan keluarga dan jumlah tanggungan merupakan atribut yang berkaitan dengan kondisi ekonomi rumah tangga, sedangkan status kepemilikan rumah dan kepemilikan aset digunakan untuk menggambarkan kondisi kepemilikan tempat tinggal dan aset masyarakat.

Tahapan pengolahan data mengacu pada proses *Knowledge Discovery in Database* (KDD) yang meliputi *selection*, *preprocessing*, *transformation*, *data mining*, dan *evaluation* (Arhami & Nasir, 2020).



Sumber: Diolah Peneliti(2026)

Gambar 1. Diagram Alur Penelitian

Pada tahap *selection* dilakukan pemilihan data dan atribut yang sesuai dengan tujuan penelitian. Tahap *preprocessing* dilakukan dengan memeriksa dan membersihkan data agar memiliki format yang sesuai untuk proses analisis (Alghifari & Juardi, 2021). Tahap *prapemrosesan* dalam penelitian ini meliputi pemeriksaan data yang kosong (*missing value*), data yang tidak konsisten, serta data duplikat. Apabila ditemukan data yang tidak sesuai, dilakukan perbaikan atau penyesuaian agar tidak memengaruhi hasil proses klasterisasi (Dista & Abdulloh, 2022). Selanjutnya, tahap *transformation* dilakukan dengan mengubah data kategorikal menjadi bentuk numerik berdasarkan kategori yang telah ditetapkan dalam penelitian, sebagaimana ditunjukkan pada tabel berikut.

Tabel 1. Transformasi Kategori Pendapatan

Kategori Pendapatan	Nilai
$\leq$ Rp1.500.000	1
$>$ Rp1.500.000 – Rp2.500.000	2
$>$ Rp2.500.000 – Rp3.500.000	3
$>$ Rp3.500.000	4

Sumber : Diolah Peneliti (2026)

Kategori pendapatan dikelompokkan berdasarkan rentang jumlah pendapatan bulanan, kemudian setiap kategori diberikan nilai 1 sampai 4. Nilai yang lebih kecil menunjukkan kategori pendapatan yang lebih rendah, sedangkan nilai yang lebih besar menunjukkan kategori pendapatan yang lebih tinggi. Transformasi ini dilakukan agar variabel pendapatan dapat digunakan dalam proses pengolahan data menggunakan algoritma OPTICS.

Tabel 2. Transformasi Kategori Jumlah Tanggungan

Jumlah Tanggungan	Nilai
1–2 orang	1
3–4 orang	2
≥ 5 orang	3

Sumber : Diolah Peneliti (2026)

Jumlah tanggungan dikelompokkan menjadi tiga kategori, yaitu 1–2 orang, 3–4 orang, dan ≥ 5 orang, yang masing-masing diberi nilai 1, 2, dan 3. Pemberian nilai tersebut digunakan untuk mengubah data kategorikal menjadi data numerik sehingga dapat diproses pada tahap *clustering* menggunakan algoritma OPTICS.

Tabel 3. Tabel Transformasi Data status kepemilikan rumah

Status Kepemilikan Rumah	Nilai
Menumpang	0
Milik sendiri	1

Sumber : Diolah Peneliti (2026)

Pada Kategori Status Kepemilikan Rumah, Status rumah *menumpang* diberi nilai 0, sedangkan status *milik sendiri* diberi nilai 1. Transformasi ini dilakukan untuk mengubah data kategorikal menjadi data numerik dalam proses *clustering*.

Tabel 4. Tabel Transformasi Data kepemilikan aset

Kepemilikan Aset	Nilai
Tidak Memiliki	1
Memiliki Aset Dasar	2
Memiliki Aset Lengkap	3

Sumber : Diolah Peneliti (2026)

Kepemilikan aset dikelompokkan menjadi tiga kategori, yaitu tidak memiliki, memiliki aset dasar, dan memiliki aset lengkap, yang masing-masing diberi nilai 1, 2, dan 3. Nilai tersebut digunakan untuk merepresentasikan tingkat kepemilikan aset dalam bentuk numerik sehingga dapat digunakan dalam proses pengolahan data menggunakan algoritma OPTICS.

Setelah proses transformasi, data dinormalisasi menggunakan metode *Min-Max Scaling* untuk menyamakan rentang nilai antaratribut. Normalisasi dilakukan menggunakan fungsi *MinMaxScaler* pada *Python*.

Data yang telah melalui proses transformasi dan normalisasi selanjutnya dianalisis menggunakan algoritma *Ordering Points To Identify the Clustering Structure* (OPTICS). OPTICS digunakan untuk mengelompokkan data berdasarkan struktur kepadatan dan kedekatan antarobjek. Pada penelitian ini, perhitungan jarak antar data menggunakan *Euclidean Distance*. Hasil perhitungan jarak tersebut digunakan sebagai dasar dalam menentukan *core distance* dan *reachability distance* pada proses pembentukan struktur *cluster* (Hastuti et al., 2024).

Perhitungan *core distance* dilakukan menggunakan persamaan berikut:

$$CoreDistance(p) = distance(p, MinPts)$$

Keterangan:

p = titik data

$MinPts$ = jumlah minimum tetangga

$distance$ = jarak antar data.

Nilai *core distance* digunakan untuk menentukan apakah suatu objek memenuhi syarat sebagai *core point* berdasarkan jumlah tetangga minimum yang telah ditentukan. Sementara itu, perhitungan *Reachability distance* dilakukan menggunakan persamaan berikut:

$$ReachabilityDistance(p, o) = \max(CoreDistance(o), d(o, p))$$

Keterangan:

o = objek tetangga

p = objek data

$d(o, p)$ = jarak antar objek

Nilai *reachability distance* digunakan untuk menentukan tingkat keterjangkauan suatu objek terhadap *core point*. Berdasarkan nilai *core distance* dan *reachability distance*, algoritma OPTICS melakukan proses *ordering* untuk menyusun urutan data berdasarkan tingkat kepadatannya.

Tahap terakhir adalah evaluasi hasil *clustering* menggunakan *Silhouette Index*. Evaluasi ini dilakukan untuk mengetahui kualitas kelompok yang terbentuk berdasarkan tingkat kedekatan data dengan anggota dalam *cluster* yang sama dan tingkat keterpisahannya dari *cluster* lain (Hastuti et al., 2023). Nilai *Silhouette* untuk setiap objek dihitung menggunakan persamaan:

$$(i) SC(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))}$$

Keterangan:

$SC(i)$ = nilai *Silhouette Coefficient* objek ke- i

$a(i)$ = rata-rata jarak objek terhadap anggota dalam *cluster* yang sama

$b(i)$ = rata-rata jarak objek terhadap *cluster* terdekat lainnya

Nilai *Silhouette Index* berada pada rentang -1 hingga 1. Semakin mendekati nilai 1, menunjukkan *cluster* yang dihasilkan semakin baik, sedangkan nilai yang mendekati 0 atau bernilai negatif menunjukkan kualitas *cluster* yang kurang baik. Dalam perkembangannya, *Silhouette Index* menjadi salah satu metode validasi *cluster* yang banyak digunakan karena

mampu memberikan ukuran kualitas *cluster* yang mudah dipahami dan memiliki tingkat interpretasi yang baik dalam mengevaluasi hasil pengelompokan data (Yolandari et al., 2025).

HASIL DAN PEMBAHASAN

Hasil

Tahap Seleksi Data

Tahap *selection* dilakukan untuk memilih atribut yang digunakan dalam proses klusterisasi. Data DTSEN yang diperoleh dari Pemerintah Desa Lubuk Rotan memiliki beberapa atribut, namun tidak seluruhnya digunakan dalam penelitian. Atribut yang bersifat identitas, seperti NIK, nama kepala keluarga, dan alamat, tidak digunakan karena tidak memengaruhi proses klusterisasi. Atribut yang dipilih pada terdiri atas pendapatan keluarga, jumlah tanggungan keluarga, status kepemilikan rumah, dan kepemilikan aset. Keempat atribut tersebut dipilih karena sesuai dengan variabel penelitian dan dapat menggambarkan kondisi sosial ekonomi masyarakat. Hasil seleksi atribut ditunjukkan pada Tabel 5:

Tabel 5 Hasil Seleksi Atribut Data Masyarakat

No	Atribut	Status	Alasan
1.	NIK	Tidak Digunakan	Atribut Identitas
2	Nama Kepala Keluarga	Tidak Digunakan	Atribut Identitas
3	Alamat	Tidak Digunakan	Tidak Digunakan dalam Proses Klusterisasi
4	Pendapatan Keluarga	Digunakan	Menggambarkan Kondisi Ekonomi
5	Jumlah Tanggungan	Digunakan	Menggambarkan Beban Tanggungan
6	Status Kepemilikan Rumah	Digunakan	Menggambarkan Kondisi Tempat Tinggal
7	Kepemilikan Aset	Digunakan	Menggambarkan Tingkat kepemilikan Aset

Sumber : Diolah Peneliti (2026)

Berdasarkan hasil *selection*, diperoleh empat atribut yang digunakan sebagai data masukan pada tahap *preprocessing* sebelum dilakukan proses transformasi dan klusterisasi menggunakan algoritma OPTICS. Data setelah proses *selection* ditunjukkan pada Gambar 2

	A	B	C	D	E
	ID	Pendapatan/bh	Jumlah Tanggungan	Aset	Status Kepemilikan Rumah
1	KK001	2.400.000	4	SEPEDA MOTOR + SMARTPHONE + KULKA	Mlha Sendri
2	KK002	2.800.000	3	SEPEDA MOTOR + SMARTPHONE	Mlha Sendri
3	KK003	0	4	SEPEDA MOTOR + SMARTPHONE	Mlha Sendri
4	KK004	1.000.000	5	SEPEDA MOTOR + SMARTPHONE	Mlha Sendri
5	KK005	1.000.000	4	SEPEDA MOTOR + SMARTPHONE	Mlha Sendri
6	KK006	4.000.000	2	SEPEDA MOTOR + SMARTPHONE + KULKA	Mlha Sendri
7	KK007	1.000.000	2	SEPEDA MOTOR + SMARTPHONE	Mlha Sendri
8	KK008	2.800.000	3	SEPEDA MOTOR + SMARTPHONE	Mlha Sendri
9	KK009	0	3	SEPEDA MOTOR + SMARTPHONE	Mlha Sendri
10	KK010	2.700.000	2	SEPEDA MOTOR + SMARTPHONE	Mlha Sendri
11	KK011	1.000.000	3	SEPEDA MOTOR + SMARTPHONE + KULKA	Mlha Sendri
12	KK012	1.000.000	3	SEPEDA MOTOR + SMARTPHONE + KULKA	Mlha Sendri
13	KK013	1.500.000	3	SEPEDA MOTOR + SMARTPHONE + KULKA	Mlha Sendri
14	KK014	2.800.000	2	SEPEDA MOTOR + SMARTPHONE + TV	Mlha Sendri
15	KK015	2.500.000	1	SEPEDA MOTOR + SMARTPHONE	Mlha Sendri
16	KK016	2.500.000	2	SEPEDA MOTOR + SMARTPHONE	Mlha Sendri
17	KK017	1.000.000	3	SEPEDA MOTOR + SMARTPHONE + KULKA	Mlha Sendri
18	KK018	1.000.000	4	SEPEDA MOTOR + SMARTPHONE + KULKA	Mlha Sendri
19	KK019	2.800.000	3	SEPEDA MOTOR + SMARTPHONE	Mlha Sendri
20	KK020	2.800.000	1	SEPEDA MOTOR + SMARTPHONE	Mlha Sendri
21	KK021	4.000.000	2	SEPEDA MOTOR + SMARTPHONE + KULKA	Mlha Sendri
22	KK022	1.000.000	1	SEPEDA MOTOR + SMARTPHONE	Mlha Sendri

Gambar 2 Data Setelah Proses Seleksi

Sumber : DTSEN Pemerintah Desa Lubuk Rotan, Diolah Peneliti(2026)

Data hasil *selection* tersebut kemudian digunakan pada tahap *preprocessing* untuk memastikan kualitas data sebelum dilakukan proses transformasi dan implementasi algoritma OPTICS.

Tahap Preprocessing Data

Tahap *pra-pemrosesan* dilakukan untuk memastikan kualitas data sebelum memasuki tahap transformasi dan proses *pengelompokan* menggunakan algoritma OPTICS. Proses ini bertujuan untuk mengidentifikasi data yang hilang (*missing value*), data duplikat, dan data

penulisan yang tidak konsisten yang dapat memengaruhi hasil analisis. Tahap *pra-pemrosesan* dilakukan menggunakan bahasa pemrograman *Python*.

Pemeriksaan nilai yang hilang dilakukan menggunakan fungsi *isnull().sum()* yang ada di dalam pustaka, pustaka *Pandas*. Pemeriksaan ini bertujuan untuk mengetahui apakah ada data kosong pada setiap atribut yang digunakan dalam penelitian. Hasil pemeriksaan nilai yang hilang ditunjukkan pada gambar 3:

```
ID      0
Pendapatan/ bln      0
Jumlah Tanggungan      0
Aset      0
Status Kepemilikan Rumah      0
dtype: int64
```

Gambar 3. Hasil Pemeriksaan Data yang Hilang

Sumber: Hasil Pengolahan Data (Python, 2026)

Pengecekan data duplikat dilakukan menggunakan fungsi *duplicated().sum()* untuk memastikan tidak ada data yang tercantum lebih dari satu kali. Hasil pemeriksaan data duplikat ditunjukkan pada gambar 4:

```
ID      0
Pendapatan/ bln      0
Jumlah Tanggungan      0
Aset      0
Status Kepemilikan Rumah      0
dtype: int64
0
```

Gambar 4. Hasil Pemeriksaan Data Duplikat

Sumber : Hasil Pengolahan Data (Python, 2026)

Berdasarkan Gambar 4, hasilnya adalah 0, yang menunjukkan bahwa tidak ada data duplikat dalam dataset. Oleh karena itu, proses penghapusan data tidak dilakukan, dan semua data masih digunakan dalam proses penelitian.

Berdasarkan hasil *praproses*, dapat disimpulkan bahwa dataset penelitian tidak memiliki *nilai yang hilang* atau data duplikat. Selain itu, konsistensi penulisan pada atribut Aset telah diperbaiki melalui proses pembersihan data sehingga seluruh kategori menjadi konsisten. Oleh karena itu, dataset telah memenuhi kualitas yang dibutuhkan dan siap digunakan pada tahap transformasi data sebelum proses *clustering* dilakukan menggunakan algoritma OPTICS

Tahap Transformasi Data

Setelah proses *preprocessing* selesai, data ditransformasikan ke dalam bentuk numerik sesuai dengan kategori yang telah ditentukan. *Transformasi* dilakukan terhadap atribut pendapatan keluarga, jumlah tanggungan, kepemilikan aset, dan status kepemilikan rumah agar seluruh atribut dapat diproses oleh algoritma OPTICS. Hasil *transformasi* setiap atribut dijelaskan sebagai berikut:

Tabel 6. Transformasi Kategori Pendapatan

Kategori	Jumlah data
1	99
2	228
3	269
4	43
Total	639

Sumber : Hasil Pengolahan Data (Python, 2026)

Berdasarkan hasil *transformasi*, kategori pendapatan terbanyak berada pada kategori 3 sebanyak 269 kepala keluarga, diikuti kategori 2 sebanyak 228 kepala keluarga. Sementara itu, kategori 1 yang merupakan kelompok pendapatan terendah berjumlah 99 kepala keluarga dan kategori 4 sebanyak 43 kepala keluarga.

Tabel 7. Transformasi Kategori Jumlah Tanggungan

Kategori	Jumlah data
1	311
2	272
3	56
Total	639

Sumber : Hasil Pengolahan Data (Python, 2026)

Hasil transformasi jumlah tanggungan menunjukkan bahwa kategori jumlah tanggungan didominasi oleh kategori 1 sebanyak 311 kepala keluarga, diikuti kategori 2 sebanyak 272 kepala keluarga, sedangkan kategori 3 berjumlah 56 kepala keluarga.

Tabel 8. Hasil Transformasi Aset

Kategori	Jumlah data
1	13
2	624
3	2
Total	639

Sumber : Hasil Pengolahan Data (Python, 2026)

Hasil transformasi aset menunjukkan bahwa mayoritas kepala keluarga berada pada kategori aset 2 sebanyak 624 data. Kategori 1 yang tidak memiliki aset berjumlah 13 data, sedangkan kategori 3 yang memiliki mobil hanya berjumlah 2 data.

Tabel 9. Hasil Transformasi Status Kepemilikan Rumah

Status	Nilai	Jumlah Data
Milik Sendiri	0	634
Menumpang	1	5

Sumber : Hasil Pengolahan Data (Python, 2026)

Berdasarkan hasil transformasi menggunakan *Label Encoding*, sebanyak 634 kepala keluarga memiliki status rumah milik sendiri, sedangkan 5 kepala keluarga masih berstatus menumpang.

Berdasarkan hasil transformasi, seluruh atribut penelitian telah berhasil diubah ke dalam bentuk numerik sehingga dapat digunakan sebagai data masukan pada tahap normalisasi. Proses transformasi menghasilkan empat atribut numerik, yaitu Kategori Pendapatan, Tanggungan, Aset, dan Status Kepemilikan Rumah. Selanjutnya, dilakukan normalisasi data setelah seluruh atribut berhasil ditransformasikan ke dalam bentuk numerik. Tujuannya adalah menyamakan rentang nilai pada setiap atribut agar tidak terjadi perbedaan skala yang dapat memengaruhi proses perhitungan jarak pada algoritma OPTICS. Normalisasi dilakukan menggunakan metode *Min-Max Scaling* yang diimplementasikan melalui library *Scikit-learn* pada bahasa pemrograman *Python*. Metode ini mengubah nilai setiap atribut ke dalam rentang 0 sampai 1 tanpa mengubah karakteristik maupun hubungan antar data (Setiawan et al., 2025).

Proses normalisasi juga diterapkan pada empat atribut yang digunakan dalam penelitian. Dengan penyamaan rentang nilai tersebut, setiap atribut memiliki kontribusi yang seimbang dalam proses perhitungan *Euclidean Distance* pada algoritma OPTICS sehingga tidak ada atribut yang mendominasi proses pembentukan *cluster*. Hasil normalisasi data menggunakan Python ditunjukkan pada Gambar 5

```
===== HASIL NORMALISASI =====
```

	Kategori Pendapatan	Kategori Tanggungan	Kategori Aset	Status Kepemilikan Rumah
0	0.333333	0.5	0.5	1.0
1	0.666667	0.5	0.5	1.0
2	0.000000	0.5	0.5	1.0
3	0.666667	1.0	0.5	1.0
4	0.666667	0.5	0.5	1.0

Sumber : Hasil Pengolahan Data (Python, 2026)

Gambar 5. Output Hasil Normalisasi Menggunakan Python

Berdasarkan Gambar 5, seluruh atribut telah berhasil dinormalisasi ke dalam rentang nilai 0 hingga 1. Nilai 0 menunjukkan nilai minimum, sedangkan nilai 1 menunjukkan nilai maksimum pada masing-masing atribut. Hasil tersebut menunjukkan bahwa proses normalisasi menggunakan metode *Min-Max Scaling* telah berhasil menyamakan skala seluruh atribut tanpa mengubah karakteristik data. Dengan demikian, proses perhitungan jarak pada algoritma OPTICS dapat dilakukan secara lebih proporsional karena setiap atribut memiliki pengaruh yang seimbang dalam proses pembentukan *cluster*. Untuk memastikan bahwa proses normalisasi telah berjalan dengan baik, dilakukan pemeriksaan statistik terhadap data hasil normalisasi. Hasil pemeriksaan menunjukkan bahwa seluruh atribut memiliki nilai minimum sebesar 0 dan nilai maksimum sebesar 1, sebagaimana disajikan pada tabel

Tabel 10. Statistik Hasil Normalisasi

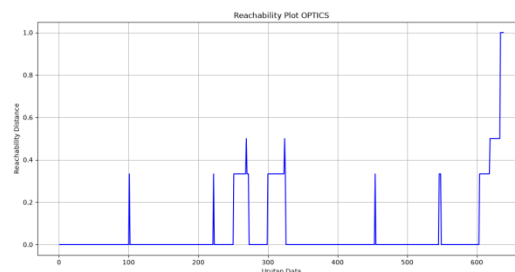
Atribut	Nilai Minimum	Nilai Maksimum
Kategori Pendapatan	0,000	1,000
Kategori Tanggungan	0,000	1,000
Kategori Aset	0,000	1,000
Status Kepemilikan Rumah	0,000	1,000

Sumber: Hasil Pengolahan Data Menggunakan Python (2026)

Berdasarkan hasil normalisasi tersebut, seluruh data telah berada pada skala yang sama sehingga memenuhi kebutuhan proses *clustering* menggunakan algoritma OPTICS. Dataset hasil normalisasi selanjutnya digunakan sebagai data masukan pada tahap implementasi algoritma OPTICS untuk mengelompokkan masyarakat berdasarkan karakteristik kondisi sosial ekonomi.

Implementasi Algoritma OPTICS

Proses pengelompokan data masyarakat kurang mampu dilakukan menggunakan algoritma *Ordering Points To Identify the Clustering Structure* (OPTICS) Implementasi algoritma dilakukan menggunakan bahasa pemrograman *Python* dengan memanfaatkan library *Scikit-Learn* terhadap 639 data Kepala Keluarga (KK) Desa Lubuk Rotan. Variabel yang digunakan dalam proses klasterisasi meliputi pendapatan keluarga, jumlah tanggungan keluarga, status kepemilikan rumah, dan kepemilikan aset. dengan parameter *min_samples=25*. Algoritma OPTICS mengelompokkan data berdasarkan tingkat kepadatan menggunakan perhitungan *Euclidean Distance*, kemudian menentukan *Core Distance* dan *Reachability Distance* sehingga mampu membentuk *cluster* dan mengidentifikasi data yang termasuk *noise*. Untuk melihat struktur kepadatan data yang terbentuk, digunakan *Reachability Plot* yang ditunjukkan pada Gambar 6

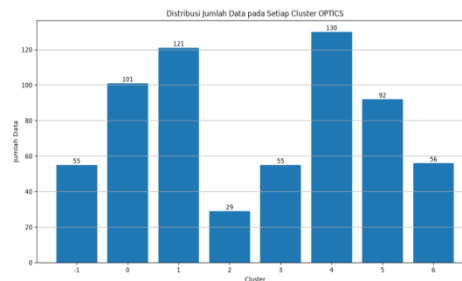


Sumber: Hasil Pengolahan Data Menggunakan Python (2026).

Gambar 6. Reachability Plot Hasil Clustering Menggunakan Algoritma OPTICS

Berdasarkan Gambar 6, terlihat bahwa nilai *Reachability Distance* setiap data bervariasi sesuai dengan tingkat kepadatan data yang dihitung oleh algoritma OPTICS. Sebagian besar data memiliki nilai *Reachability Distance* yang mendekati 0, yang menunjukkan bahwa data berada pada wilayah dengan kepadatan tinggi sehingga membentuk suatu *cluster*.

Namun, pada beberapa urutan data terlihat peningkatan nilai *Reachability Distance*. Peningkatan tersebut menunjukkan adanya batas antar *cluster* atau data yang memiliki karakteristik berbeda dibandingkan data di sekitarnya. Semakin tinggi nilai *Reachability Distance*, semakin rendah tingkat kemiripan suatu data dengan kelompoknya sehingga data tersebut berpotensi menjadi *noise* atau data pencilan. Secara keseluruhan, grafik menunjukkan bahwa algoritma OPTICS berhasil mengidentifikasi struktur kepadatan data masyarakat Desa Lubuk Rotan dan membentuk beberapa *cluster* berdasarkan tingkat kemiripan karakteristik sosial ekonomi tanpa harus menentukan jumlah *cluster* sejak awal. Grafik distribusi jumlah data pada setiap *cluster* OPTICS di tampilkan pada gambar 7

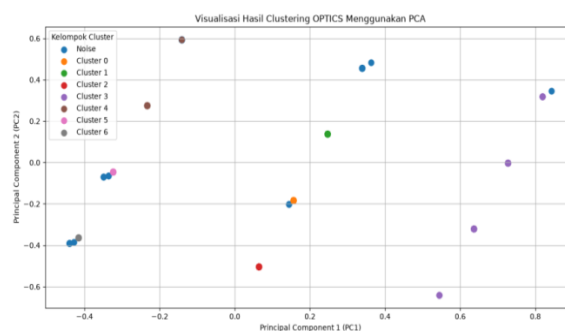


Sumber: Hasil Pengolahan Data Menggunakan Python (2026).

Gambar 7. Grafik Distribusi Jumlah Data pada Setiap Custer OPTICS

Berdasarkan Gambar 7, hasil penerapan algoritma OPTICS terhadap 639 data masyarakat menghasilkan 7 *cluster* dan 1 kelompok *noise*. Setiap *cluster* memiliki jumlah anggota yang berbeda, menunjukkan bahwa kepadatan data pada masing-masing kelompok tidak sama. *Cluster 4* merupakan *cluster* dengan jumlah anggota terbanyak, yaitu 130 data (20,34%), diikuti oleh *Cluster 1* sebanyak 121 data (18,94%), *Cluster 0* sebanyak 101 data (15,81%), dan *Cluster 5* sebanyak 92 data (14,40%). Sementara itu, *Cluster 6* terdiri atas 56 data (8,76%), *Cluster 3* sebanyak 55 data (8,61%), dan *Cluster 2* merupakan *cluster* dengan jumlah anggota paling sedikit, yaitu 29 data (4,54%).

Selain itu, algoritma OPTICS mengidentifikasi 55 data (8,61%) sebagai *noise* atau data pencilan. Data tersebut tidak dimasukkan ke dalam *cluster* tertentu karena memiliki tingkat kepadatan yang berbeda dibandingkan kelompok lainnya. Hal ini menunjukkan bahwa algoritma OPTICS mampu membedakan data yang tidak memiliki tingkat kemiripan yang cukup dengan anggota *cluster* lain sehingga hasil pengelompokan menjadi lebih representatif terhadap kondisi data yang sebenarnya. Secara keseluruhan, hasil *clustering* menunjukkan bahwa karakteristik sosial ekonomi masyarakat Desa Lubuk Rotan bersifat beragam. Perbedaan karakteristik tersebut menyebabkan terbentuknya beberapa *cluster* dengan jumlah anggota yang berbeda sesuai tingkat kemiripan antar data. Untuk mempermudah interpretasi hasil *clustering*, dilakukan visualisasi menggunakan metode *Principal Component Analysis* (PCA). Visualisasi ini bertujuan mereduksi empat variabel penelitian menjadi dua komponen utama sehingga pola penyebaran *cluster* dapat diamati dengan lebih jelas (Sari et al., 2026). Hasil visualisasi ditunjukkan pada Gambar 8:

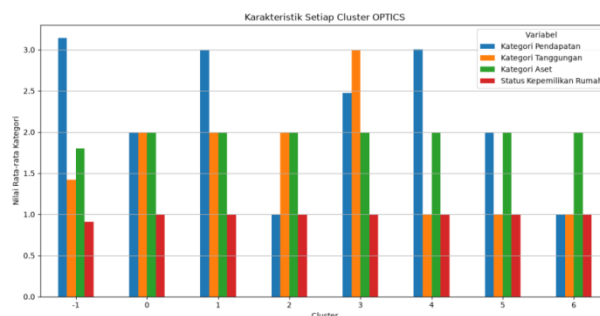


Sumber: Hasil Pengolahan Data Menggunakan Python (2026).

Gambar 8. Visualisasi hasil clustering menggunakan (PCA)

Berdasarkan Gambar 8, visualisasi hasil *clustering* menggunakan *Principal Component Analysis* (PCA) menunjukkan bahwa data berhasil diproyeksikan ke dalam dua komponen utama, yaitu *Principal Component 1* (PC1) dan *Principal Component 2* (PC2). Visualisasi ini memudahkan pengamatan terhadap pola penyebaran data hasil *clustering* yang diperoleh dari algoritma OPTICS. Setiap *cluster* ditampilkan dengan warna yang berbeda sehingga penyebaran antar*cluster* dapat terlihat dengan jelas. Sebagian besar *cluster* menempati posisi yang berbeda pada bidang PCA, menunjukkan bahwa data dalam setiap *cluster* memiliki tingkat kemiripan yang lebih tinggi dibandingkan dengan data pada *cluster* lainnya. Selain itu, terdapat beberapa data yang dikategorikan sebagai *noise*, yang menunjukkan bahwa data tersebut tidak memenuhi tingkat kepadatan yang cukup untuk bergabung ke dalam *cluster* tertentu.

Berdasarkan hasil visualisasi *clustering* menggunakan algoritma OPTICS, diketahui bahwa data berhasil dikelompokkan ke dalam beberapa *cluster* dengan karakteristik yang berbeda. Untuk mengetahui karakteristik masing-masing *cluster* secara lebih rinci, dilakukan analisis terhadap nilai rata-rata setiap variabel, Hasil analisis karakteristik setiap *cluster* ditunjukkan pada Gambar 9



Sumber: Hasil Pengolahan Data Menggunakan Python (2026).
Gambar 9. Hasil Analisis Karakteristik Setiap Kluster OPTICS

Berdasarkan Gambar 9. terlihat bahwa setiap *cluster* memiliki karakteristik yang berbeda berdasarkan nilai rata-rata kategori pendapatan, jumlah tanggungan, aset, dan status kepemilikan rumah. Perbedaan karakteristik tersebut menunjukkan bahwa algoritma OPTICS mampu mengelompokkan data berdasarkan tingkat kemiripan masing-masing kelompok.

Hasil Evaluasi Clustering Menggunakan *Silhouette Index*

Setelah proses *clustering* menggunakan algoritma OPTICS selesai dilakukan, tahap selanjutnya adalah mengevaluasi kualitas hasil *clustering* menggunakan metode *Silhouette Index*. Sebelumnya dilakukan pengujian terhadap parameter *min_samples* dengan beberapa jumlah minimum sampel, yaitu 20, 25, 30, 35, dan 40 sampel. Pengujian tersebut dilakukan menggunakan *metric = euclidean* dan *cluster_method = XI*. Berdasarkan hasil pengujian, nilai *min_samples = 25* dipilih sebagai parameter yang digunakan pada proses clustering akhir karena memberikan hasil clustering yang paling sesuai berdasarkan hasil pengujian. Evaluasi ini bertujuan untuk mengukur tingkat kemiripan data dalam *cluster* yang sama (*cohesion*) serta tingkat keterpisahan antar *cluster* (*separation*). Nilai *Silhouette Index* berada pada rentang -1 hingga 1, di mana nilai yang semakin mendekati 1 menunjukkan kualitas *cluster* yang semakin baik. Hasil evaluasi *Silhouette Index* ditampilkan pada Tabel 11:

Tabel 11. Hasil Evaluasi Clustering Menggunakan *Silhouette Index*

Metode Evaluasi	Nilai	Kriteria
<i>Silhouette Index</i>	0,833517	Sangat Baik

Sumber: Hasil Pengolahan Data Menggunakan Python (2026)

Berdasarkan Tabel 11, diperoleh nilai *Silhouette Index* sebesar 0,833517. Nilai tersebut menunjukkan bahwa hasil *clustering* menggunakan algoritma OPTICS memiliki kualitas yang

sangat baik, karena berada pada rentang nilai 0,70–1,00. Hal ini menunjukkan bahwa data dalam setiap *cluster* memiliki tingkat kemiripan yang tinggi dengan anggota pada *cluster* yang sama serta memiliki pemisahan yang jelas dengan *cluster* lainnya. Dengan demikian, hasil *clustering* yang diperoleh dinilai mampu membentuk kelompok masyarakat dengan karakteristik yang jelas sehingga dapat digunakan sebagai dasar dalam analisis karakteristik masyarakat.

Pembahasan

Analisis Karakteristik Setiap Cluster

Berdasarkan hasil *clustering* menggunakan algoritma OPTICS, diperoleh tujuh *cluster* dan satu kelompok *noise* (*Cluster -1*). Karakteristik setiap *cluster* dianalisis berdasarkan nilai rata-rata kategori pendapatan, kategori jumlah tanggungan, kategori kepemilikan aset, dan status kepemilikan rumah. Hasil analisis karakteristik setiap *cluster* adalah sebagai berikut:

Cluster 0

Cluster 0 memiliki rata-rata kategori pendapatan sebesar 2,00, yang menunjukkan bahwa sebagian besar masyarakat memiliki pendapatan pada rentang lebih dari Rp1.500.000 hingga Rp2.500.000 per bulan. Rata-rata kategori jumlah tanggungan sebesar 2,00 menunjukkan bahwa mayoritas masyarakat memiliki 3-4 orang tanggungan. Selain itu, rata-rata kategori aset sebesar 2,00 menunjukkan bahwa masyarakat umumnya memiliki aset dasar, sedangkan status kepemilikan rumah sebesar 1,00 menunjukkan bahwa mayoritas masyarakat memiliki rumah sendiri. Karakteristik tersebut menunjukkan kondisi sosial ekonomi yang relatif sedang.

Cluster 1

Cluster 1 memiliki rata-rata kategori pendapatan sebesar 3,00, yang menunjukkan bahwa sebagian besar masyarakat memiliki pendapatan pada rentang lebih dari Rp2.500.000 hingga Rp3.500.000 per bulan. Rata-rata kategori jumlah tanggungan sebesar 2,00 menunjukkan bahwa mayoritas masyarakat memiliki 3–4 orang tanggungan. Selain itu, masyarakat pada *cluster* ini umumnya memiliki aset dasar dan rumah milik sendiri. Karakteristik tersebut menunjukkan kondisi ekonomi yang relatif lebih baik dibandingkan *Cluster 0* dan *Cluster 2*.

Cluster 2

Cluster 2 memiliki rata-rata kategori pendapatan sebesar 1,00, yang menunjukkan bahwa sebagian besar masyarakat memiliki pendapatan $\leq$ Rp1.500.000 per bulan. Rata-rata kategori jumlah tanggungan sebesar 2,00 menunjukkan bahwa mayoritas masyarakat memiliki 3–4 orang tanggungan. Selain itu, masyarakat pada *cluster* ini didominasi oleh kepemilikan aset dasar dan rumah milik sendiri. Karakteristik tersebut menunjukkan bahwa *Cluster 2* memiliki tingkat pendapatan paling rendah dibandingkan sebagian besar *cluster* lainnya.

Cluster 3

Cluster 3 memiliki rata-rata kategori pendapatan sebesar 2,47, yang menunjukkan bahwa karakteristik pendapatan anggotanya cenderung berada pada kategori pendapatan menengah dengan sebagian anggota berada pada kategori yang lebih tinggi. Rata-rata kategori jumlah tanggungan sebesar 3,00 menunjukkan bahwa sebagian besar masyarakat memiliki lima orang atau lebih tanggungan, yang merupakan jumlah tanggungan tertinggi dibandingkan *cluster* lainnya. Selain itu, masyarakat pada *cluster* ini umumnya memiliki aset dasar dan rumah milik sendiri. Kondisi tersebut menunjukkan bahwa beban tanggungan keluarga pada *Cluster 3* relatif lebih tinggi dibandingkan *cluster* lainnya.

Cluster 4

Cluster 4 memiliki rata-rata kategori pendapatan sebesar 3,01, yang menunjukkan bahwa sebagian besar masyarakat memiliki pendapatan pada rentang lebih dari Rp2.500.000 hingga Rp3.500.000 per bulan. Rata-rata kategori jumlah tanggungan sebesar 1,00 menunjukkan bahwa mayoritas masyarakat memiliki 1–2 orang tanggungan. Selain itu, masyarakat pada *cluster* ini didominasi oleh kepemilikan aset dasar dan rumah milik sendiri. Karakteristik tersebut menunjukkan kondisi ekonomi yang relatif lebih baik dengan jumlah tanggungan yang rendah.

Cluster 5

Cluster 5 memiliki rata-rata kategori pendapatan sebesar 2,00, yang menunjukkan bahwa sebagian besar masyarakat memiliki pendapatan pada rentang lebih dari Rp1.500.000 hingga Rp2.500.000 per bulan. Rata-rata kategori jumlah tanggungan sebesar 1,00 menunjukkan bahwa mayoritas masyarakat memiliki 1–2 orang tanggungan. Selain itu, masyarakat pada *cluster* ini umumnya memiliki aset dasar dan rumah milik sendiri. Karakteristik tersebut menunjukkan kondisi ekonomi yang relatif stabil dengan jumlah tanggungan yang lebih rendah.

Cluster 6

Cluster 6 memiliki rata-rata kategori pendapatan sebesar 1,00, yang menunjukkan bahwa sebagian besar masyarakat memiliki pendapatan $\leq$ Rp1.500.000 per bulan. Rata-rata kategori jumlah tanggungan sebesar 1,00 menunjukkan bahwa mayoritas masyarakat memiliki 1–2 orang tanggungan. Selain itu, masyarakat pada *cluster* ini didominasi oleh kepemilikan aset dasar dan rumah milik sendiri. Karakteristik tersebut menunjukkan bahwa *cluster* ini memiliki tingkat pendapatan yang rendah meskipun jumlah tanggungannya relatif sedikit.

Cluster -1 (Noise)

Cluster -1 (noise) memiliki rata-rata kategori pendapatan sebesar 3,15, rata-rata kategori jumlah tanggungan sebesar 1,42, rata-rata kategori aset sebesar 1,80, dan rata-rata status kepemilikan rumah sebesar 0,91. *Cluster* ini merupakan kelompok data yang tidak memenuhi tingkat kepadatan sehingga tidak tergabung ke dalam *cluster* tertentu. Oleh karena itu, data pada *cluster* ini tidak digunakan sebagai dasar dalam analisis karakteristik kelompok utama.

Implikasi Hasil Klasterisasi terhadap Penyaluran Bantuan Sosial.

Hasil klasterisasi OPTICS menunjukkan adanya perbedaan karakteristik sosial ekonomi masyarakat Desa Lubuk Rotan berdasarkan pendapatan, jumlah tanggungan, kepemilikan aset, dan status kepemilikan rumah. *Cluster 2* memiliki rata-rata kategori pendapatan terendah, yaitu 1,00 atau $\leq$ Rp1.500.000 per bulan, dengan jumlah tanggungan 3–4 orang, sehingga perlu mendapat perhatian dalam proses verifikasi calon penerima bantuan sosial. *Cluster 3* memiliki rata-rata jumlah tanggungan tertinggi, yaitu 3,00 atau 5 orang atau lebih, dengan rata-rata pendapatan 2,47. Kondisi tersebut menunjukkan bahwa beban tanggungan juga menjadi faktor yang perlu diperhatikan meskipun pendapatannya tidak berada pada kategori terendah.

Sementara itu, *Cluster 1* dan *Cluster 4* memiliki pendapatan relatif lebih tinggi, dengan *Cluster 4* memiliki jumlah tanggungan 1–2 orang. *Cluster 0*, *Cluster 5*, dan *Cluster 6* menunjukkan karakteristik yang berada di antara kelompok lainnya, dengan perbedaan pada tingkat pendapatan dan jumlah tanggungan. Hasil tersebut menunjukkan bahwa setiap *cluster* memiliki karakteristik dan tingkat kebutuhan yang berbeda. Oleh karena itu, hasil klasterisasi dapat digunakan sebagai informasi pendukung bagi pemerintah desa untuk menentukan

kelompok masyarakat yang perlu diprioritaskan dalam proses verifikasi bantuan sosial, bukan sebagai penentu tunggal penerima bantuan.

KESIMPULAN DAN SARAN

Kesimpulan

Berdasarkan hasil penelitian, penerapan algoritma *Ordering Points To Identify the Clustering Structure* (OPTICS) mampu mengelompokkan masyarakat Desa Lubuk Rotan berdasarkan karakteristik sosial ekonomi melalui tahapan *Knowledge Discovery in Database* (KDD). Proses tersebut menghasilkan data yang siap dianalisis sehingga pengelompokan dapat dilakukan secara lebih objektif berdasarkan variabel pendapatan keluarga, jumlah tanggungan, status kepemilikan rumah, dan kepemilikan aset.

Hasil implementasi OPTICS terhadap 639 data Kepala Keluarga (KK) menghasilkan 7 *cluster* dan 1 kelompok *noise*, yang menunjukkan bahwa kondisi sosial ekonomi masyarakat memiliki karakteristik yang beragam. Keberadaan data *noise* membuktikan bahwa algoritma OPTICS mampu mengidentifikasi data yang tidak memiliki tingkat kemiripan yang cukup dengan kelompok lain, sehingga hasil pengelompokan menjadi lebih representatif terhadap kondisi data yang sebenarnya.

Berdasarkan evaluasi menggunakan *Silhouette Index*, diperoleh nilai 0,833517 yang termasuk dalam kategori sangat baik. Nilai tersebut menunjukkan bahwa setiap *cluster* memiliki tingkat kemiripan yang tinggi di dalam kelompoknya serta memiliki pemisahan yang jelas dengan *cluster* lain. Dengan demikian, algoritma OPTICS terbukti efektif dalam mengelompokkan masyarakat berdasarkan kondisi sosial ekonomi dan dapat dimanfaatkan sebagai pendukung pengambilan keputusan dalam penyaluran bantuan sosial agar lebih tepat sasaran. Meskipun demikian, hasil *clustering* tetap perlu didukung dengan proses verifikasi lapangan sehingga keputusan yang dihasilkan sesuai dengan kondisi masyarakat yang sebenarnya.

Keterbatasan

Penelitian ini memiliki keterbatasan pada variabel yang digunakan dalam proses *clustering*, yang masih berfokus pada data sosial ekonomi yang tersedia sehingga belum mencakup seluruh faktor yang dapat menggambarkan kondisi masyarakat secara menyeluruh. Selain itu, penelitian hanya menggunakan algoritma OPTICS sehingga belum dilakukan perbandingan dengan algoritma *clustering* lainnya. Hasil *clustering* juga bergantung pada kualitas dan kondisi data yang digunakan, sehingga perubahan atau pembaruan data dapat memengaruhi hasil pengelompokan.

Saran

Data sosial ekonomi masyarakat yang digunakan dalam *clustering* sebaiknya diperbarui secara berkala agar hasil pengelompokan tetap sesuai dengan kondisi terkini dan mendukung ketepatan sasaran bantuan sosial. Penelitian selanjutnya dapat menambahkan variabel sosial ekonomi yang relevan, seperti pendidikan, pekerjaan, kesehatan, dan pengeluaran rumah tangga, serta membandingkan algoritma OPTICS dengan algoritma *clustering* lainnya menggunakan dataset yang sama untuk mengevaluasi kualitas hasil pengelompokan.

REFERENSI

- Alghifari, F., & Juardi, D. (2021). Penerapan Data Mining Pada Penjualan Makanan Dan Minuman Menggunakan Metode Algoritma Naïve Bayes. *Jurnal Ilmiah Informatika*, 9(02), 75–81. <https://doi.org/10.33884/jif.v9i02.3755>
- Ankerst, M., Breunig, M. M., & Kriegel, H. (1999). *OPTICS: Ordering Points To Identify the Clustering Structure*. 49–60.
- Arhami, M., & Nasir, M. (2020). *Data Mining: Algoritma dan Implementasi* (R. I. Utami

- (ed.); 1st ed.). ANDI.
- Buana, L. A. C., Data, M., & Trisnawan, P. H. (2026). *Pengembangan Alat Bantu Investigasi Serangan Siber Berdasarkan Log Server Web Menggunakan Algoritma Optics*. 10(2), 1–9.
- Dista, T. M., & Abdulloh, F. F. (2022). Clustering Pengunjung Mall Menggunakan Metode K-Means dan Particle Swarm Optimization. *Jurnal Media Informatika Budidarma*, 6(3), 1339. <https://doi.org/10.30865/mib.v6i3.4172>
- Hastuti, S. H., Nurmayanti, W. P., & Saputri, A. A. (2023). Penerapan Metode Clustering Self Organizing Maps (Som) Dan K-Affinity Propagation (K-Ap) Dalam Mengelompokkan Nilai Tukar Petani Di Indonesia 2022. *VARIANCE: Journal of Statistics and Its Applications*, 5(1), 79–88. <https://doi.org/10.30598/variancevol5iss1page79-88>
- 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>
- Kementerian Sosial Republik Indonesia. (2026). *Program Bantuan Sosial*. <https://kemensos.go.id/>
- Musthafa, A. Z., Basukl, B. M., & Habibi, A. (2025). *Implementasi Sistem Klasterisasi Menggunakan Algoritma K-Means Untuk Seleksi Penerima Bantuan Sosial*. 19, 54–60.
- Mustikasari, & Salman, N. (2023). *ANALISIS KLASSTER BERBASIS KEPADATAN DENGAN DBSCAN DAN OPTICS*. 8, 1–8. <https://doi.org/https://doi.org/10.24252/insypro.v8i1.36347>
- Nurhalizah, C. D., Damaliana, A. T., & Prasetya, D. A. (2025). *OPTICS-Based Clustering of East Java Regencies and Cities by Divorce Factors*. 4(3), 113–123. <https://doi.org/https://doi.org/10.55537/jistr.v4i3.1227>
- Sari, I. D. P., Pratiwi, S. M., & Al Rosyid, H. (2026). Clustering Wilayah Pulau Jawa Berdasarkan Indikator Sosial Ekonomi Menggunakan Metode K-Means. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 10(1), 1032–1036. <https://doi.org/10.36040/jati.v10i1.16934>
- Setiawan, M. A., Insani, F., Vitriani, Y., & Yusra. (2025). Analisis Perbandingan Metode DBSCAN dan Meanshift dalam Klasterisasi Data Gempa Bumi di Indonesia. *Bulletin of Computer Science Research*, 5(4), 554–563. <https://doi.org/10.47065/bulletincsr.v5i4.605>
- Supardi, N. E., Handhayani, T., & Lewenusa, I. (2026). *CLUSTERING WILAYAH KEMISKINAN MULTIDIMENSI DI INDONESIA MENGGUNAKAN ALGORITMA FUZZY C-MEANS DAN OPTICS*. 20(2), 275–283. <https://doi.org/https://doi.org/10.33480/inti.v20i2.7814>
- Yolandari, N. A., Butarbutar, L. E., Rajagukguk, G. C. H., Zulfi, M. F., & Ramadhani, F. (2025). *ANALISIS PERBANDINGAN K-MEANS DAN DBSCAN DALAM PENGELOMPOKAN DATA TRAVEL REVIEW RATINGS MENGGUNAKAN EVALUASI SILHOUETTE INDEX DAN DAVIES-BOULDIN INDEX*. *Jurnal Informatika Dan Teknik Terapan*, 13(3). <https://doi.org/https://doi.org/10.23960/jitet.v13i3.6884>
- Zhou, G. (2022). *Data Mining for Co-location Patterns : principles and aplications*. CRC Press. <https://doi.org/https://doi.org/10.1201/9781003139416>