

KLASTERISASI KERENTANAN SOSIAL DAN PENDIDIKAN PENERIMA PROGRAM KELUARGA HARAPAN MENGGUNAKAN METODE K-MEANS DI KECAMATAN RENGASDENGKLOK

Ainun Nadzifah, Shofa Shofiah Hilabi, Baenil Huda

Sistem Informasi, Universitas Buana Perjuangan Karawang
Jalan HS. Ronggo Waluyo, Telukjambe Timur, Puseurjaya,
Telukjambe Timur, Kabupaten Karawang, Jawa Barat 41361, Indonesia
si22.ainunnadzifah@mhs.ubpkarawang.ac.id

ABSTRAK

Program Keluarga Harapan (PKH) merupakan instrumen strategis penanggulangan kemiskinan, namun data administratif penerima di Kecamatan Rengasdengklok belum termanfaatkan optimal untuk memetakan kerentanan sosial dan pendidikan secara terukur. Penelitian ini bertujuan mengelompokkan 1.1511 Kepala Keluarga penerima PKH tahun 2025 ke dalam tiga kategori kerentanan memanfaatkan metode klasterisasi K-Means. Metodologi penelitian mencakup tahapan pra-pemrosesan data yang meliputi pembersihan, konversi variabel kategorik ke numerik, serta normalisasi Min-Max Scaling, dilanjutkan penentuan jumlah klaster optimal melalui Metode Elbow, eksekusi algoritma K-Means, dan evaluasi kualitas klaster menggunakan Silhouette Coefficient. Hasil komputasi menunjukkan jumlah klaster optimal $K=3$ dengan nilai inertia sebesar 19,34 serta konvergensi tercapai pada iterasi kedelapan. Skor Silhouette Coefficient sebesar 0,2281 mengonfirmasi pemisahan klaster yang wajar bagi data bersifat multidimensi. Terbentuk tiga profil karakteristik distinct: Klaster 0 merepresentasikan kerentanan sedang dengan sebaran pendidikan dasar hingga menengah merata, Klaster 1 mencerminkan kerentanan tinggi yang didominasi penyandang disabilitas dan tingkat pendidikan rendah, sedangkan Klaster 2 menunjukkan kerentanan rendah dengan proporsi pendidikan SMA paling tinggi. Studi ini membuktikan efektivitas K-Means dalam memetakan kerentanan penerima PKH secara objektif, sehingga output klasterisasi dapat menjadi acuan strategis penetapan prioritas pendampingan serta intervensi bantuan yang lebih presisi dan tepat sasaran.

Kata kunci : K-Means, Klasterisasi, PKH, Kerentanan Sosial, Kerentanan Pendidikan

1. PENDAHULUAN

Kemiskinan merupakan rintangan global yang dialami oleh banyak wilayah, tidak terkecuali pada negara yang berkembang seperti halnya negara Indonesia saat ini. Untuk meningkatkan penanggulangan kemiskinan dengan lebih cepat, sejak tahun 2007 pemerintah kita sudah mengoperasikan sistem Program Keluarga Harapan (PKH) [1]. PKH ini juga dikenal di mancanegara sebagai *Conditional Cash Transfers* (CCT) dan telah teruji efektif dalam menanggulangi persoalan kemiskinan akut di sejumlah negara-negara di dunia [2].

PKH telah diterapkan di Kabupaten Karawang sejak 2008 untuk membantu keluarga prasejahtera memenuhi kebutuhan dasar, biaya sekolah, dan layanan Kesehatan [3].

Penelitian ini difokuskan pada Kecamatan Rengasdengklok pada tahun 2025 yang mencatat 1.151 Kartu Keluarga terdaftar, dengan 772 anggota yang berhak menerima komponen bantuan. Program ini telah menjangkau banyak keluarga, kenyataan di lapangan menunjukkan bahwa beban yang ditanggung tiap rumah tangga berbeda-beda ada yang berat di biaya pendidikan, ada pula yang terkendala urusan dalam hal beban perawatan anggota keluarga. Teknik pengumpulan data dalam penelitian ini memanfaatkan data sekunder melalui perolehan pendamping PKH di Kecamatan Rengasdengklok. Namun, data yang

tersedia masih bersifat inventarisasi administratif dan belum diolah menjadi indikator kerentanan yang terukur. Akibatnya, bantuan dan pendampingan yang diberikan masih bersifat umum. Kondisi tersebut menunjukkan perlunya proses pengolahan data lebih lanjut menggunakan pendekatan data mining, khususnya teknik klasterisasi. Atas dasar tersebut, penelitian ini bertujuan untuk mengelompokkan keluarga penerima PKH berdasarkan kerentanan yang dipahami sebagai total beban sosial dan pendidikan tanggungan keluarga penerima PKH, yang dilihat dari jumlah serta jenis komponen bantuan dalam setiap Kartu Keluarga.

Untuk mengatasi kerentanan ini, pendekatan klasterisasi menggunakan metode K-Means dapat menjadi alternatif yang tepat. Metode ini secara konsisten terbukti efektif dalam memetakan data kesejahteraan sosial dan Pendidikan, sebagaimana diadaptasi dalam sejumlah studi terdahulu.

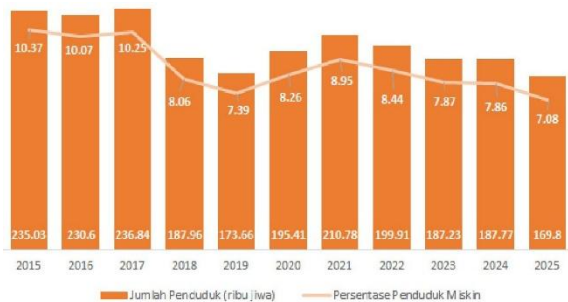
2. TINJAUAN PUSTAKA

2.1. Program Keluarga Harapan (PKH)

Program Keluarga Harapan (PKH) merupakan agenda penyaluran santunan dana tunai berprasyarat berdasarkan aturan resmi pemerintah bagi keluarga prasejahtera (KM) [4].

Program ini menyasar KM yang tercatat pada Data Terpadu Kesejahteraan Sosial (DTKS) serta

sudah melengkapi kriteria sesuai regulasi yang tengah berjalan [5].



Gambar 1. Diagram Batang Jumlah dan Persentase Keluarga Miskin

Sumber : (Badan Pusat Statistik Kabupaten Karawang, 2025)

Menurut Badan Pusat Statistik (BPS), total penduduk miskin pada daerah Karawang tercatat merangkak naik saat memasuki tahun 2021 akibat pandemi Covid-19, namun kembali menurun hingga mencapai 169,80 ribu jiwa (7,08%) pada Maret 2025. Program PKH bertujuan membantu memenuhi kebutuhan jangka pendek sekaligus mengakhiri siklus kemiskinan jangka panjang. Objek penerima manfaat PKH yaitu keluarga miskin yang mempunyai unsur pendidikan dan kesehatan seperti ibu hamil, anak kecil, lansia, hingga difabel guna menjadi ukuran hambatan sosial [6].

2.2. Kerentanan Sosial

Kerentanan sosial adalah kondisi ketika individu atau keluarga memiliki risiko tinggi terhadap tekanan sosial dan ekonomi yang dapat menyebabkan kemiskinan [7].

Faktor yang memengaruhi kerentanan antara lain kondisi ekonomi lemah, pendidikan rendah, banyaknya tanggungan keluarga, serta pekerjaan yang tidak tetap [8].

Kerentanan diukur dari jumlah anggota keluarga yang termasuk kelompok rentan seperti lansia, penderita disabilitas, ibu hamil, perempuan menyusui, dan anak usia sekolah. Semakin banyak anggota keluarga dalam kategori tersebut, semakin tinggi tingkat kerentanannya [9].

Penilaian ini digunakan untuk menentukan prioritas penerima bantuan melalui data DTKS. Program PKH bertujuan mengurangi kerentanan keluarga miskin melalui dukungan pendidikan, kesehatan, dan social [10].

2.3. Kerentanan Pendidikan

Kerentanan pendidikan adalah kondisi ketika individu atau keluarga mengalami kesulitan mengakses pendidikan sehingga peluang meningkatkan kualitas hidup menjadi terbatas [10].

Dalam Program Keluarga Harapan (PKH), kerentanan ini dilihat dari partisipasi anak dalam pendidikan mulai dari SD hingga SMA, di mana

bantuan bersyarat diberikan agar anak tetap bersekolah [11].

Sektor pendidikan memegang peranan vital untuk memajukan taraf hidup warga juga memangkas alur kemiskinan [12].

Karena hal tersebut, penelitian ini memanfaatkan metode K-Means untuk proses pengklasifikasian penerima PKH berdasarkan kondisi sosial dan pendidikan guna mengetahui tingkat kerentanan pendidikan sebagai dasar penentuan prioritas pendampingan.

2.4. K-means Clustering

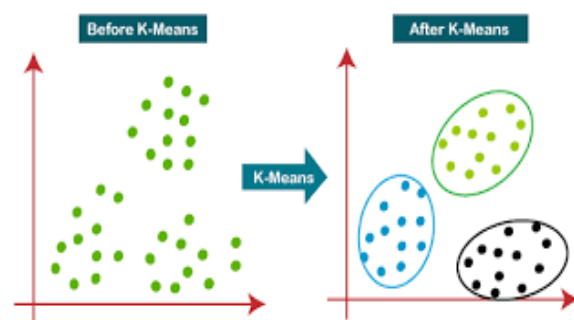
K-Means merupakan metode unsupervised learning dalam data mining untuk mengklasifikasikan sekumpulan data pada beberapa kluster menurut kedekatan cirinya [13].

Metode kerja ini memecah data ke total k kluster pilihan, di mana tiap-tiap bagian mempunyai pusat massa yang dinamakan centroid sebagai hasil hitung rata-rata dari seluruh elemen di dalam kluster itu.

Alur K-Means diawali dengan menetapkan total kelompok dan menentukan centroid awal secara random, lalu mengukur kedekatan tiap data ke centroid memakai rumus Euclidean Distance.

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

di mana x dan y adalah titik data dan i menunjukkan atribut. Setelah perhitungan jarak dilakukan tiap data dimasukkan ke dalam kelompok dengan rentang paling pendek. Selanjutnya, titik centroid direvisi melalui penghitungan angka rerata dari semua elemen kelompok. Mekanisme tersebut dilakukan berulang sampai letak centroid bersifat tetap (konvergen) atau menyentuh ambang batas iterasi [14].



Gambar 2. Sebelum dan Sesudah di Klustering

Gambar tersebut mengilustrasikan proses K-Means mengelompokkan data acak menjadi kluster teratur berdasarkan jarak. Dalam penelitian ini, algoritma diterapkan pada data penerima PKH melalui tahapan: penentuan kluster optimal (Metode Elbow), inisialisasi centroid, pengelompokan, dan evaluasi [15].

Hasilnya memberikan pemetaan tingkat kerentanan sosial dan pendidikan yang terstruktur, sebagai landasan rekomendasi kebijakan intervensi yang tepat sasaran.

2.5. Penelitian Terkait

Beberapa penelitian terdahulu telah membuktikan efektivitas metode K-Means dalam pemetaan data sosial. Abdul et al. (2021) mengelompokkan penyaluran dana PKH tingkat kabupaten ke dalam tiga kluster ekonomi, namun fokusnya masih pada distribusi anggaran bukan pada kerentanan sosial dan pendidikan di tingkat kecamatan [16].

Maharani (2023) juga memanfaatkan K-Means untuk mengelompokkan calon penerima bantuan berdasarkan profil ekonomi, meskipun hasil kluster masih memerlukan penyempurnaan [17].

Dalam konteks ketepatan sasaran, Rady Putra dan Anggrawan (2021) menunjukkan bahwa K-Means mampu membedakan penerima bantuan tepat dan tidak tepat sasaran berdasarkan pekerjaan dan penghasilan [18]. Temuan serupa ditunjukkan Risaychi (2024) yang membagi penerima bantuan ke dalam tiga kategori berdasarkan jenis bantuan [19], serta Febriansyah dan Muntari (2023) yang mengonfirmasi peran signifikan variabel pendidikan dalam membentuk kluster rumah tangga rentan [20].

Meskipun berbagai studi telah membuktikan kegunaan K-Means, sebagian besar masih berfokus pada variabel ekonomi umum atau ketepatan penyaluran. Integrasi antara indikator kerentanan sosial dan pendidikan khususnya pada penerima PKH di tingkat kecamatan masih jarang dilakukan. Oleh karena itu, penelitian ini mengisi celah tersebut dengan menggabungkan kedua aspek ke dalam satu model klusterisasi, sehingga menghasilkan pengelompokan yang lebih terukur dan dapat dijadikan dasar penentuan prioritas pendampingan.

3. METODE PENELITIAN

Penelitian ini memanfaatkan algoritma K-Means untuk mengelompokkan data penerima PKH berdasarkan pola kerentanan. Hasil pengelompokan akan disajikan secara visual untuk memudahkan analisis pola sebaran data. Berikut adalah tahapan metode penelitian yang digunakan.

3.1. Pengumpulan Data

Tahap pengumpulan data menggunakan data sekunder dari pendamping PKH Kecamatan Rengasdengklok. Variabel penelitian mencakup kondisi sosial dan pendidikan, meliputi jumlah anggota keluarga, anak prasekolah, murid sekolah (SD/SMP/SMA), lansia, kaum disabilitas, serta ibu hamil dan menyusui. Peneliti memastikan validitas dan relevansi data sebelum dilanjutkan ke tahap pra-pemrosesan. Hasil tahap ini adalah dataset mentah yang siap untuk proses klusterisasi.

3.2. Pra Pemrosesan Data

Tahap pra-pemrosesan dijalankan secara sistematis demi menyusun data agar kompatibel dengan algoritma K-Means. Pertama, dilakukan filterisasi data dengan mengatasi kolom kosong sekaligus menghilangkan data duplikat. Kedua, variabel kategorik dialihkan menjadi bentuk angka agar dapat dieksekusi secara komputasional. Terakhir, normalisasi data dikerjakan melalui Min-Max Scaling untuk mengatur skala semua variabel ke nilai 0–1. Langkah tersebut krusial sebab K-Means berbasis perhitungan jarak, sehingga variabel berskala besar tidak mendominasi hasil kluster. Hasil akhir tahap ini adalah dataset bersih dan terstandarisasi yang siap untuk proses klusterisasi.

3.3. Penentuan Jumlah Kluster

Sebelum proses klusterisasi dilakukan, terlebih dahulu ditentukan jumlah kluster (K) yang optimal menggunakan metode Elbow. Penentuan ini dilakukan dengan menghitung nilai Within-Cluster Sum of Squares (WCSS) pada beberapa nilai K, kemudian dipilih nilai K yang menunjukkan penurunan paling signifikan atau membentuk titik siku pada grafik. Nilai K tersebut selanjutnya dimanfaatkan dalam langkah klusterisasi melalui algoritma K-Means.

3.4. Proses Klusterisasi K-Means

Ketika jumlah kluster terbaik diputuskan melalui Metode Elbow, skema klusterisasi dikerjakan dengan siklus rutin: (1) memulai centroid pertama secara acak; (2) menghitung selisih tiap data ke centroid menggunakan Euclidean Distance; (3) memasukkan data ke kluster berdasarkan nilai kedekatan jarak terkecil; dan (4) pembaruan posisi centroid berdasarkan rata-rata nilai atribut anggota kluster. Tahapan ini dijalankan berulang sampai konvergen, yakni saat letak centroid tetap atau pergeseran yang ada tidak berarti. Hasil akhir tahap ini adalah pengelompokan data penerima PKH ke dalam tiga kluster tingkat kerentanan sosial dan pendidikan: rendah, sedang, dan tinggi.

3.5. Analisis dan Visualisasi Hasil

Setelah proses klusterisasi selesai, Dilaksanakan pengujian kualitas kluster melalui *Silhouette Coefficient* untuk menghitung level kohesi (kedekatan data di dalam kluster) juga separasi (perbedaan antar kluster). Skor metrik ini berkisar dari -1 hingga 1, di mana skor yang mendekati 1 mengindikasikan bahwa objek telah ditempatkan pada kluster yang tepat dengan batas pemisah yang jelas. Temuan analisis ini dimanfaatkan untuk membuktikan bahwa kluster yang tercipta memiliki standar yang baik dan pengelompokan data dinilai optimal. Selanjutnya, hasil klusterisasi divisualisasikan dalam bentuk tabel distribusi dan grafik (seperti *scatter plot* atau diagram batang) untuk memudahkan interpretasi pola kerentanan di setiap kelompok. Tampilan ini menyajikan potret edukatif mengenai karakteristik

setiap kelompok sebagai dasar analisis lanjutan dan rekomendasi kebijakan.

4. HASIL DAN PEMBAHASAN

4.1. Deskripsi Dataset

Fokus penelitian ini adalah Kecamatan Rengasdengklok tahun 2025, dengan populasi data sekunder sebesar 1.151 KK penerima PKH dan 772 jiwa yang memenuhi kriteria sebagai penerima komponen bantuan. Dataset mencakup 7 atribut kondisi sosial dan pendidikan, meliputi jumlah anak usia dini, jenjang pendidikan sekolah, keberadaan lansia, disabilitas, dan ibu hamil. Data mentah ini menjadi dasar proses klasterisasi untuk memetakan tingkat kerentanan sosial dan pendidikan penerima bantuan.

Table 1. Data Set Awal PKH

AK	AUD	SD	SMP	SMA	Disabilitas	Lansia	Hamil
4	0	0	1	0	0	0	0
6	0	1	0	0	0	0	0
5	0	0	0	0	0	2	0
4	0	0	1	0	0	0	0
1	0	0	0	0	0	1	0
6	0	2	0	0	0	0	0
4	0	1	0	1	0	0	0
1	0	0	0	0	0	1	0
4	0	0	1	0	0	0	0

4.2. Variabel Penelitian

Variabel penelitian yang dimanfaatkan dalam skema klasterisasi diseleksi menurut kriteria dasar yang merepresentasikan kondisi sosial dan pendidikan keluarga penerima Program Keluarga Harapan (PKH). Variabel-variabel ini mencerminkan tingkat tanggungan keluarga serta komponen bantuan yang diterima, yang berhubungan langsung dengan tingkat kerentanan sosial dan pendidikan. beban tanggungan (anggota keluarga & bantuan), tanggung jawab edukasi (anak per jenjang), serta faktor rentan (lansia, disabilitas, ibu hamil). Variabel-variabel ini memungkinkan K-Means mengelompokkan penerima PKH secara objektif terbagi tiga urutan kategori kerentanan: rendah, sedang, dan tinggi.

Table 2. Variabel Penelitian

No	Variabel Penelitian	Keterangan
1	Jumlah Anak Usia Dini (AUD)	Jumlah anak usia 0–6 tahun dalam keluarga
2	Jumlah Anak SD	Jumlah anak yang sedang menempuh studi tingkat SD
3	Jumlah Anak SMP	Jumlah populasi anak pada jenjang pendidikan tingkat SMP
4	Jumlah Anak SMA	Jumlah individu dalam kategori fase pendidikan tingkat SMA
5	Jumlah Lansia	Jumlah anggota keluarga berusia ≥ 60 tahun
6	Jumlah Penyandang Disabilitas	Jumlah anggota keluarga dengan kondisi disabilitas
7	Jumlah Ibu Hamil	Jumlah ibu hamil dalam keluarga

4.3. Pemeriksaan Data Kosong (Missing Value)

Pemeriksaan missing value menggunakan fungsi `isnull()` pada 1151 data menunjukkan bahwa seluruh variabel penelitian memiliki kelengkapan data 100% (nilai kosong = 0). Dengan demikian, tidak diperlukan penanganan khusus untuk missing value, dan dataset dapat langsung diproses ke tahapan pra-pemrosesan selanjutnya.

4.4. Konversi Variabel Numerik

Tahap konversi dilakukan untuk mengubah variabel bertipe kategorik menjadi numerik menggunakan `LabelEncoder` dari library `scikit-learn`. Hasil konversi menunjukkan bahwa seluruh kolom dalam dataset berhasil diubah dari tipe data object menjadi `int64`. Proses ini penting karena algoritma K-Means bekerja berdasarkan perhitungan matematis jarak Euclidean yang mensyaratkan seluruh input data berbentuk numerik. Setelah konversi, dataset terdiri dari 9 variabel numerik yang siap untuk dinormalisasi.

4.5. Normalisasi Data dengan Min-Max Scaling

Tahap normalisasi dilakukan dengan metode Min-Max Scaling untuk mentransformasi seluruh nilai variabel ke dalam rentang 0 hingga 1. Berdasarkan hasil verifikasi normalisasi data, diketahui bahwa seluruh variabel komponen PKH (AUD, SD, SMP, SMA, disabilitas, dan lansia) menghasilkan nilai minimum 0,0 dan nilai maksimum 1,0, yang mengindikasikan bahwa proses normalisasi telah berhasil dilakukan pada rentang yang diharapkan. Namun, pada variabel Hamil, nilai maksimum tercatat 0,0 bukan 1,0. Ini mengindikasikan bahwa dalam dataset yang digunakan, tidak terdapat komponen ibu hamil.

Table 3. Normalisasi Data dengan Min-Max Scaling

Komponen PKH	Nilai Minimum	Nilai Maksimum	Keterangan
Aud	0,0	1,0	Terdistribusi normal
Sd	0,0	1,0	Terdistribusi normal
Smp	0,0	1,0	Terdistribusi normal
Sma	0,0	1,0	Terdistribusi normal
Disabilitas	0,0	1,0	Terdistribusi normal
Lansia	0,0	1,0	Terdistribusi normal
Hamil	0,0	0,0	Konstan (tidak ada variasi)

4.6. Hasil Penentuan Jumlah Klaster

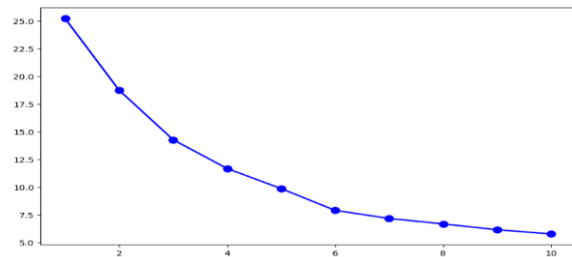
Mengacu luaran komputasi metode Elbow, didapati bahwa skor WCSS (Within-Cluster Sum of Squares) senantiasa merosot seiring dengan peningkatan kuantitas klaster (K). Nilai WCSS tertinggi berada pada K=1 yaitu sebesar 25,21, kemudian terus turun hingga mencapai nilai terendah

pada K=10 yaitu sebesar 5,78. Penurunan nilai WCSS yang paling signifikan terjadi dari K=9 ke K=10, sehingga titik siku (elbow) terdeteksi pada K=10. Hal ini menunjukkan bahwa jumlah kluster optimal yang direkomendasikan untuk proses klusterisasi K-Means adalah sebanyak 10 kluster.

Table 4. Hasil Perhitungan Wcss (Elbow Method)

Jumlah Kluster (K)	WCSS (Inertia)
1	25.207483
2	18.743386
3	14.277614
4	11.682254
5	9.869120
6	7.905853
7	7.176687
8	6.682870
9	6.158598
10	5.779663

Nilai WCSS untuk setiap jumlah kluster (K=1 hingga K=10) telah dihitung menggunakan metode Elbow. Hasil perhitungan tersebut selanjutnya divisualisasikan dalam bentuk grafik pada gambar berikut ini.



Gambar 3. Hasil Elbow Method

Pada grafik tersebut, tampak bahwa skor WCSS kian merosot sejalan peningkatan K. Reduksi paling drastis muncul dari K=1 ke K=3 serta K=3 menuju K=4, lalu mulai stabil sesudah K=4. Maka, titik siku berada pada K=4, yang mengindikasikan bahwa total kelompok terbaik adalah 4.

4.7. Hasil Proses Klusterisasi K-Means

Proses klusterisasi K-Means diawali dengan inisialisasi centroid awal secara acak. Setelah melalui proses iterasi hingga konvergen, diperoleh posisi centroid akhir untuk masing-masing kluster seperti disajikan pada tabel .

Table 5. Hasil Centroid Akhir

Kluster	AUD	SD	SMP	SMA	Disabilitas	Lansia	Hamil
0	0.1522	0.2464	0.0000	0.2391	0,0000	0.2464	0,0000
1	0.0789	0.1053	0.6316	0,2368	0,0000	0, 1754	0,0000
2	0,0000	0.0952	0.0714	0.1429	1,0000	0, 1905	0,0000

Hasil centroid akhir dengan K=3 menunjukkan bahwa kluster 0 memiliki nilai cukup merata pada AUD, SD, SMA, dan Lansia. kluster 1 didominasi oleh komponen SMP. Sedangkan kluster 2 didominasi oleh komponen Disabilitas.

Setelah memperoleh nilai centroid akhir dari proses iterasi, tahap selanjutnya adalah melakukan penugasan setiap objek ke kelompok tertentu berdasarkan pada selisih centroid. Berdasarkan hasil penugasan data ke kluster menggunakan centroid akhir, diperoleh distribusi anggota kluster yang bervariasi. Kluster 0 merupakan kluster terbesar

dengan 23 Kepala Keluarga (46,9% dari total data) dan memiliki rata-rata jarak ke centroid sebesar 0,5469. Kluster 1 menempati posisi kedua dengan 19 KK (38,8%) dan rata-rata jarak ke centroid 0,5147. Adapun Kluster 2 sebagai kluster terkecil terdiri dari 7 KK (14,3%) dengan nilai rata-rata jarak ke centroid terendah yaitu 0,4227. Rendahnya nilai rata-rata jarak pada Kluster 2 mengindikasikan tingkat homogenitas yang lebih tinggi dibandingkan kluster lainnya, meskipun memiliki jumlah anggota paling sedikit. Hasil penugasan data ke dalam kluster berdasarkan centroid akhir disajikan pada Tabel 6.

Table 6. Hasil Penugasan Data per Kluster

Kluster	Jumlah Data (KK)	Persentase (%)	Rata-rata Jarak ke Centroid
0	23	46,9	0,5469
1	19	38,8	0,5147
2	7	14,3	0,4227
Total	49	100	-

Setelah diperoleh hasil centroid akhir pada setiap kluster, proses lanjutan merupakan penghitungan selisih tiap data menuju centroid melalui metode *Euclidean Distance*. Hasil perhitungan jarak tersebut

kemudian dianalisis lebih lanjut dengan menghitung statistik jarak minimum dan maksimum untuk mengukur tingkat penyebaran data dalam setiap kluster.

Table 7. Tabel Jarak Data terhadap Centroid Menggunakan Euclidean Distance

Kluster	Jumlah Data	Rata-rata Jarak	Jarak Minimum	Jarak Maksimum	Rentang
0	23	0,5469	0,3090	0,9473	0,6383
1	19	0,5147	0,3401	0,8049	0,4648
2	7	0,4227	0,2345	0,6232	0,3887

Analisis lebih lanjut menunjukkan bahwa Klaster 2 memiliki tingkat kekompakan tertinggi dengan rentang jarak 0,2345–0,6232 dan rata-rata 0,4227. Sebaliknya, Klaster 0 menunjukkan penyebaran terlebar dengan jarak maksimum mencapai 0,9473, meskipun jarak minimumnya relatif dekat (0,3090). Klaster 1 berada di posisi tengah dengan rentang jarak 0,3401–0,8049. Variasi rentang jarak ini mencerminkan tingkat homogenitas yang berbeda di antara ketiga klaster.

Proses pembaruan centroid dilakukan secara iteratif hingga mencapai konvergensi. Berdasarkan hasil algoritma K-Means, proses ini berlangsung selama 8 iterasi. Di setiap fase, titik centroid diperbaiki melalui kalkulasi rerata fitur dari anggota klaster, kemudian setiap data ditugaskan ke klaster terdekat. Proses berhenti ketika perubahan centroid sudah tidak signifikan, yang menandakan klasterisasi telah stabil.

Table 8. Tabel Nilai Centroid Baru Berdasarkan Rata-rata Atribut Klaster

Klaster	0	1	2
AUD	0.1600	0.0000	0.0588
SD	0.2133	0.0952	0.1373
SMP	0.3200	0.0714	0.2353
SMA	0.0000	0.1429	0.5882

Table 9. Nilai rata-rata silhouette keseluruhan

Klaster	0	1	2	Total
Kategori	Sedang	Tinggi	Rendah	-
Jumlah Anggota	9	6	6	21
Rata-rata Silhouette	0,1900	0,2712	0,2231	0,2281
Std Deviasi	0,1861	0,1449	0,1137	-
Rentang Nilai	-0,0309 – 0,4546	-0,0175 – 0,4308	0,0816 – 0,4105	-

Berdasarkan Tabel 9, Klaster 1 (Kerentanan Tinggi) memiliki nilai silhouette tertinggi sebesar 0,2712 dengan standar deviasi terendah (0,1449), menunjukkan bahwa anggota klaster ini memiliki karakteristik yang paling homogen dan terpisah dengan baik dari klaster lain. Klaster 2 (Kerentanan Rendah) menunjukkan stabilitas terbaik dengan seluruh nilai silhouette bernilai positif (0,0816–0,4105), yang berarti tidak ada anggota yang salah penempatan. Sementara itu, Klaster 0 memiliki nilai minimum negatif (-0,0309) yang mengindikasikan adanya beberapa data yang berada di batas antar klaster. Secara keseluruhan, nilai silhouette rata-rata sebesar 0,2281 menunjukkan struktur klaster yang wajar untuk data kerentanan sosial yang bersifat kompleks dan multidimensi.

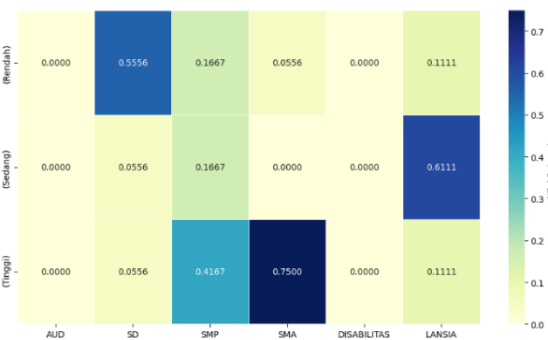
Untuk memudahkan interpretasi karakteristik masing-masing klaster, hasil klasterisasi divisualisasikan dalam bentuk heatmap yang menampilkan nilai rata-rata setiap atribut per klaster. Visualisasi ini memberikan gambaran yang jelas mengenai profil kerentanan sosial dan pendidikan pada setiap kelompok, sehingga perbedaan karakteristik antar klaster dapat teridentifikasi secara visual. Heatmap karakteristik rata-rata per klaster ditampilkan pada Gambar 4.

Klaster	0	1	2
Disabilitas	0.0000	1.0000	0.0000
Lansia	0.2267	0.1905	0.1961
Kategori Kerentanan	Sedang	Tinggi	Rendah

Berdasarkan Tabel 1, Klaster 1 dikategorikan sebagai kerentanan tinggi karena memiliki nilai DISABILITAS = 1,0000 (seluruh anggota klaster menyandang disabilitas) dan proporsi pendidikan dasar yang dominan. Sebaliknya, Klaster 2 menunjukkan kerentanan rendah dengan nilai SMA tertinggi (0,5882), sedangkan Klaster 0 berada pada tingkat kerentanan sedang.

4.8. Hasil Analisis dan Visualisasi Hasil

Setelah proses klasterisasi selesai, dilakukan evaluasi kualitas klaster menggunakan Silhouette Coefficient untuk mengevaluasi tingkat kecocokan setiap data dimasukkan ke klaster yang benar. Parameter silhouette terentang dari -1 hingga 1, di mana hasil yang lebih besar menandakan jarak antar klaster yang kian jelas. Hasil perhitungan silhouette score untuk setiap klaster ditampilkan pada Tabel 9 berikut



Gambar 4. Heatmap Karakteristik Rata-rata per Klaster

Berdasarkan Gambar 4, setiap klaster memiliki ciri khas tersendiri. Klaster 0 (Rendah) didominasi oleh pendidikan SD, Klaster 1 (Sedang) oleh variabel Lansia, dan Klaster 2 (Tinggi) oleh pendidikan SMA. Variabel AUD dan Disabilitas bernilai nol di semua klaster. Hasil ini menegaskan bahwa ketiga kelompok memiliki profil kerentanan yang berbeda dan terpisah dengan jelas.

5. KESIMPULAN DAN SARAN

Penelitian ini sukses mengaplikasikan metode K-Means untuk mengelompokkan penerima PKH di Kecamatan Rengasdengklok ke dalam tiga kluster tingkat kerentanan sosial dan pendidikan. Jumlah kluster optimal $K=3$ didapatkan melalui Metode Elbow dengan perolehan nilai inersia 19,34, dan proses iterasi mencapai konvergensi pada iterasi ke-8. Evaluasi melalui Silhouette Coefficient mencapai skor rata-rata 0,2281 yang menunjukkan kualitas kluster yang wajar untuk data kerentanan multidimensi. Ketiga kluster yang terbentuk memiliki karakteristik berbeda: kerentanan sedang (pendidikan dasar-menengah merata), kerentanan tinggi (dominasi disabilitas dan pendidikan rendah), serta kerentanan rendah (proporsi SMA tertinggi). Hasil tersebut membuktikan bahwasanya metode K-Means efektif memetakan kerentanan penerima PKH secara objektif sebagai dasar prioritas pendampingan.

Dinas Sosial disarankan memanfaatkan hasil klusterisasi untuk memprioritaskan pendampingan intensif pada keluarga berkerentanan tinggi, sementara kluster lain difokuskan pada pemberdayaan ekonomi dan akses pendidikan. Penelitian selanjutnya direkomendasikan mengintegrasikan indikator ekonomi, menguji model pada seluruh populasi penerima, serta membandingkannya dengan algoritma lain seperti DBSCAN atau Hierarchical Clustering guna meningkatkan akurasi dan validitas pemetaan kerentanan.

DAFTAR PUSTAKA

- [1] Mahizah Rani, "Pengelompokan Data Calon Penerima Bantuan PKH (Program Keluarga Harapan) Menggunakan Metode K-Means Clustering," Jakarta, Aug. 2023.
- [2] J. Ekonomi Pertanian Dan Agribisnis *Et Al.*, "Karakteristik Rumah tangga Petani Usaha Padi Penerima PKH Di Kabupaten Karawang, Jawa Barat Rice Smallholder Characteristics Of PKH Recipient In Kabupaten Karawang, West Java," Vol. 4, No. 1, Pp. 123–132, 2020, Doi: 10.21776/Ub.Jepa.2020.004.01.12.
- [3] A. Rivai, B. M. Sinaga, H. Siregar, And H. Harianto, "Karakteristik Rumah tangga Petani Usaha Padi Penerima PKH Di Kabupaten Karawang, Jawa Barat," *Jurnal Ekonomi Pertanian Dan Agribisnis*, Vol. 4, No. 1, Pp. 123–132, Jan. 2020, Doi: 10.21776/Ub.Jepa.2020.004.01.12.
- [4] M. Heru Widodo, A. Rohmad Jatmiko, M. Fathur Rohman, J. Ekonomi Pembangunan, And F. Ekonomi, "Analisis Program Keluarga Harapan (PKH) Untuk Meningkatkan Kesejahteraan Keluarga Penerima Manfaat (PKM) Di Kabupaten Jombang," *Jurnal Eba*, Vol. 9, No. 2, 2022.
- [5] M. Heru Widodo, A. Rohmad Jatmiko, M. Fathur Rohman, J. Ekonomi Pembangunan, And F. Ekonomi, "Analisis Program Keluarga Harapan (PKH) Untuk Meningkatkan Kesejahteraan Keluarga Penerima Manfaat (PKM) Di Kabupaten Jombang," *Jurnal Eba*, Vol. 9, No. 2, 2022.
- [6] Badan Pusat Statistik Kabupaten Karawang, "Profil Kemiskinan Di Kabupaten Karawang Maret 2025," Karawang, Sep. 2025.
- [7] L. G. Rady Putra And A. Anggrawan, "Pengelompokan Penerima Bantuan Sosial Masyarakat Dengan Metode K-Means," *Matrik : Jurnal Manajemen, Teknik Informatika Dan Rekayasa Komputer*, Vol. 21, No. 1, Pp. 205–214, Nov. 2021, Doi: 10.30812/Matrik.V21i1.1554.
- [8] A. Z. Musthafa, B. Minto Basukl, And A. Habibi, "Implementasi Sistem Klusterisasi Menggunakan Algoritma K-Means Untuk Seleksi Penerima Bantuan Sosial," *Science Electro*, Vol. 19, No. 1, Sept 2025.
- [9] A. Sukron, Y. Cikusin, And T. Rahman Ilyas, "Implementasi Program Keluarga Harapan (PKH) Dalam Menanggulangi Masyarakat Miskin Di Kota Batu (Studi Pada Dinas Sosial Kota Batu)," *Jurnal Respon Publik* Vol. 18, No. 3, Tahun 2024, Hal: 90-97.
- [10] B. Huda And B. Priyatna, "Penggunaan Aplikasi Content Management System (Cms) Untuk Pengembangan Bisnis Berbasis E-Commerce," *Systematics*, Vol. 1, No. 2, P. 81, Dec. 2019, Doi: 10.35706/Sys.V1i2.2076.
- [11] Vianti & Fauziah, "Analisis Pengaruh Bantuan Sosial Bersyarat Program Keluarga Harapan Terhadap Partisipasi Pendidikan Anak Usia Sekolah Di Indonesia," *Jurnal Kebijakan Ekonomi*, Vol. 18, No. 2, Oct. 2023, Doi: 10.7454/Jke.V18i2.1135.
- [12] Djumura Et Al, "Program Keluarga Harapan Dalam Peningkatan Kesejahteraan Masyarakat Desa Bubeya," *Jurnal Kajian Ekonomi Dan Bisnis*, Vol. 15, No 1, Juni. 2022
- [13] S. Shofiah Hilabi, Y. Aprianti, A. Lia Hananto, S. Informasi, And U. Buana Perjuangan Karawang, "Klasifikasi Sentimen Komentar Pengguna Pada Aplikasi Ruangguru Menggunakan Algoritma Naive Bayes," Vol. 9, P. 2025, Doi: 10.47002/Metik.V9i1.1023.
- [14] S. S. Hilabi, D. D. Susilo, B. Priyatna, And E. Novalia, "Implementasi Data Mining Dalam Pengelompokan Data Pembelian Menggunakan Algoritma K-Means Pada Pt.Otomotif 1," *Jutisi : Jurnal Ilmiah Teknik Informatika Dan Sistem Informasi*, Vol. 13, No. 1, P. 476, Apr. 2024, Doi: 10.35889/Jutisi.V13i1.1836.
- [15] S. S. Hilabi, "Rancang Bangun Situs Responsif Di Universitas Buana Perjuangan Karawang Dengan Menggunakan Metode Perpaduan Grid System Dan Css Media Query," *Techno Xplore : Jurnal Ilmu Komputer Dan Teknologi Informasi*, Vol. 2, No. 1, Apr. 2017, Doi: 10.36805/Technoxplore.V2i1.220.

- [16] S. Abdul, A. 1*, S. Defit, And Y. Yunus, “Jurnal Informatika Ekonomi Bisnis Klasterisasi Dana Bantuan Pada Program Keluarga Harapan (PKH) Menggunakan Metode K-Means,” *Jurnal Informatika Ekonomi Bisnis*, Vol. 3 No. 2 (2021) 53-59. Doi: 10.37034/InfEb.V3i2.66.
- [17] Maharani Et Al, “Pengaruh Kerentanan Sosial Terhadap Ketangguhan Masyarakat Dalam Menghadapi Bencana Erupsi Gunung Merapi Di Kabupaten Sleman,” *Jurnal Dialog Penanggulangan Bencana*, Vol. 11, No. 1, 2020.
- [18] L. G. Rady Putra And A. Anggrawan, “Pengelompokan Penerima Bantuan Sosial Masyarakat Dengan Metode K-Means,” *Matrik : Jurnal Manajemen, Teknik Informatika Dan Rekayasa Komputer*, Vol. 21, No. 1, Pp. 205–214, Nov. 2021, Doi: 10.30812/Matrik.V21i1.1554.
- [19] A. B. Risaychi, “Clustering Bantuan Sosial Sesuai Tingkat Kemiskinan Di Kelurahan Girikerto Menggunakan Algoritma K-Means,” *Seminar Nasional Mahasiswa Fakultas Teknologi Informasi*, Vol 3, No 2, September, 2024. [Online]. Available: <https://Senafiti.Budiluhur.Ac.Id>
- [20] F. Febriansyah And S. Muntari, “Penerapan Algoritma K-Means Untuk Klasterisasi Penduduk Miskin Pada Kota Pagar Alam,” *Jurnal Informatika Sunan Kalijaga*, Vol. 8, No. 1, Januari, 2023, Pp. 66 – 77.