

CLUSTERING PENERIMA BANTUAN PANGAN BERBASIS ALGORITMA K-MEANS UNTUK MENINGKATKAN EFEKTIVITAS PROGRAM SOSIAL DI KOTA/KABUPATEN CIREBON

Naufal Arif Hidayatullah^{1*}, Willy Prihartono², Fathurrohman³

^{1,2,3}Sekolah Tinggi Manajemen Informatika dan Komputer (STMIK) IKMI Cirebon; Jl. Perjuangan No.10B, Karyamulya, Kec. Kesambi, Kota Cirebon, Jawa Barat 45135; Telp. (0231) 490480

Received: 11 Desember 2024

Accepted: 14 Januari 2025

Published: 20 Januari 2025

Keywords:

Algoritma K-Means, Data Mining, Bantuan Sosial Pangan, Clustering, Distribusi Bantuan.

Correspondent Email:

naufalhidayatullah1604@gmail.com

Abstrak. Distribusi bantuan sosial pangan sering menghadapi tantangan berupa ketidakmerataan akibat pengelompokan penerima yang belum optimal. Penelitian ini bertujuan untuk mengelompokkan penerima bantuan sosial di Kota dan Kabupaten Cirebon menggunakan *Algoritma K-Means*. Variabel utama yang dianalisis meliputi kelurahan, kecamatan, jarak ke lokasi distribusi, dan usia penerima. *Algoritma K-Means* dipilih karena kemampuannya memproses data besar secara efisien dan menghasilkan kluster yang terpisah dengan baik. Optimalisasi jumlah kluster dilakukan menggunakan *Davies-Bouldin Index (DBI)* pada rentang $K=2$ hingga $K=6$. Hasil menunjukkan nilai optimal $K=3$ dengan DBI sebesar 0,505, menghasilkan tiga kluster. *Cluster 0* memiliki rata-rata jumlah kelurahan tertinggi (188,434) dan jarak distribusi terpendek (412,873 meter), cocok untuk wilayah padat penduduk. *Cluster 1* memiliki jarak distribusi terjauh (2979,891 meter) dan usia penerima rata-rata 49,217 tahun, mencakup area dengan tantangan distribusi yang lebih kompleks. *Cluster 2* memiliki jarak distribusi sedang (1422,595 meter) dan usia rata-rata 51,564 tahun, mencerminkan wilayah yang luas tetapi tetap mudah dijangkau. Penelitian ini menunjukkan bahwa *Algoritma K-Means* mampu meningkatkan efektivitas distribusi bantuan sosial dengan memastikan alokasi yang lebih tepat berdasarkan karakteristik demografis dan geografis penerima.

Abstract. *The distribution of social assistance for food often encounters challenges related to uneven targeting, primarily due to suboptimal grouping of recipients. This study aims to classify food aid recipients in Cirebon City and Regency using the K-Means algorithm. The key variables analyzed include subdistrict, district, distance to distribution centers, and recipient age. K-Means was chosen for its efficiency in handling large datasets and its ability to create well-separated Cluster s. The optimal number of Cluster s was determined using the Davies-Bouldin Index (DBI) with K values ranging from 2 to 6. The results indicate that the optimal K is 3, with a DBI score of 0.505, resulting in three Cluster s. Cluster 0 features the highest average number of subdistricts (188.434) and the shortest average distance to distribution points (412.873 meters), suitable for densely populated areas. Cluster 1 is characterized by the longest average distance (2979.891 meters) and an average recipient age of 49.217 years, covering areas with more complex distribution challenges. Cluster 2 has a moderate average distance (1422.595 meters) and an average age of 51.564 years, representing wider but still accessible regions. This research demonstrates that the K-Means algorithm is effective in improving the accuracy and efficiency of food aid distribution by*

ensuring better targeting based on recipients' demographic and geographic characteristics.

1. PENDAHULUAN

Program bantuan pangan merupakan salah satu upaya pemerintah untuk mengurangi kesenjangan sosial dan memenuhi kebutuhan dasar masyarakat rentan. Agar bantuan ini dapat tepat sasaran dan memaksimalkan manfaatnya, diperlukan efisiensi tinggi dalam proses distribusinya. Dalam era transformasi digital, penggunaan teknologi seperti machine learning menjadi semakin relevan untuk mendukung pengambilan keputusan berbasis data yang lebih baik [1].

Bantuan sosial memungkinkan distribusi yang lebih terukur dengan mengelompokkan penerima berdasarkan atribut tertentu. Salah satu algoritma yang sering digunakan adalah *K-Means* karena kemampuannya dalam mengelompokkan data berdasarkan berbagai atribut secara efektif [2]. Kota/Kabupaten Cirebon, dengan populasi yang beragam dan tantangan sosial yang kompleks, membutuhkan analisis berbasis data untuk meningkatkan efisiensi program bantuan pangan [3].

Meskipun program bantuan pangan telah rutin dilakukan, masih terdapat kendala seperti ketidakakuratan dalam menentukan penerima, sehingga bantuan tidak selalu sampai kepada pihak yang paling membutuhkan. Selain itu, ketidaksesuaian antara kebutuhan nyata di lapangan dan jumlah penerima terdaftar menyebabkan distribusi menjadi tidak efektif [4].

Karakteristik penerima seperti usia, wilayah, dan jarak belum dimanfaatkan secara maksimal dalam proses segmentasi. Oleh karena itu, diperlukan metode berbasis data yang lebih objektif untuk mengelompokkan penerima dan menentukan prioritas berdasarkan kebutuhan yang sebenarnya. *Algoritma K-Means* dinilai relevan dalam konteks ini karena mampu menemukan pola klaster secara lebih akurat [5].

Penelitian sebelumnya menunjukkan bahwa *K-Means* telah digunakan dalam berbagai industri, termasuk untuk mengidentifikasi kelompok rentan dalam program bantuan sosial. Misalnya, penelitian oleh [6] menunjukkan bahwa *K-Means* dapat mempercepat proses penentuan prioritas penerima bantuan di

wilayah tertentu. Namun, penelitian tersebut terbatas pada analisis atribut dasar tanpa mempertimbangkan elemen seperti seperti usia, wilayah, dan jarak.

Konteks penelitian ini adalah mengintegrasikan variabel penting seperti usia, wilayah, dan jarak ke dalam model klasterisasi untuk meningkatkan efektivitas distribusi bantuan pangan. Data dari PT. Pos Indonesia akan dianalisis menggunakan perangkat lunak *RapidMiner*, yang memungkinkan pengolahan data besar secara efisien tanpa kebutuhan pengkodean khusus. Model yang dihasilkan akan membantu pemerintah daerah menentukan penerima bantuan berdasarkan klaster secara lebih efisien.

Penelitian ini diharapkan memberikan kontribusi praktis dalam perencanaan distribusi bantuan pangan yang lebih tepat sasaran di Kota/Kabupaten Cirebon. Selain itu, temuan ini juga dapat digunakan sebagai referensi untuk pengembangan aplikasi kebijakan sosial berbasis data yang lebih transparan dan akuntabel. Dengan demikian, penelitian ini tidak hanya mendukung kebijakan lokal tetapi juga menambah literatur ilmiah di bidang analitik data dan kebijakan sosial.

2. TINJAUAN PUSTAKA

2.1. *Algoritma K-Means*

Algoritma K-Means Clustering merupakan teknik pengelompokan yang mengorganisasi data ke dalam cluster berdasarkan tingkat kemiripan, dengan tujuan agar data dalam satu *cluster* memiliki karakteristik yang serupa. Langkah-langkahnya meliputi menentukan jumlah *cluster* (K), menginisialisasi titik pusat (*centroid*), mengelompokkan data berdasarkan jarak ke *centroid*, serta memperbarui posisi *centroid* hingga mencapai kestabilan. Keunggulan algoritma ini adalah mudah dipahami dan efisien, sementara kekurangannya mencakup perlunya menetapkan jumlah *cluster* di awal serta kerentanannya terhadap pemilihan titik pusat awal [7].

2.2. Davies-Bouldin Index (DBI)

Davies-Bouldin Index (DBI) adalah metrik yang digunakan untuk menilai kualitas hasil pengelompokan data dengan mengukur sejauh mana kluster yang dihasilkan saling terpisah dan seberapa rapat kluster tersebut. DBI dihitung dengan membandingkan jarak antara *centroid* kluster dengan tingkat kerapatan dalam kluster itu sendiri. Nilai DBI yang lebih rendah menunjukkan kualitas kluster yang lebih baik, karena mencerminkan jarak antar kluster yang lebih besar dibandingkan jarak dalam kluster. Dengan kata lain, semakin kecil nilai DBI, semakin baik tingkat pemisahan dan kekompakan kluster yang terbentuk [8].

2.3. Clustering

Clustering adalah metode untuk mengelompokkan data ke dalam beberapa kelompok berdasarkan kemiripan, sehingga data dalam satu kelompok memiliki kesamaan yang lebih besar dibandingkan dengan data di kelompok lainnya. Tujuan *Clustering* adalah mengidentifikasi pola atau struktur dalam data yang tidak memiliki label, dan metode ini dapat diterapkan dalam berbagai bidang, termasuk analisis kemiskinan [9].

2.4. Data Mining

Data Mining adalah serangkaian langkah yang digunakan untuk mengekstrak informasi berharga yang sebelumnya tidak diketahui dari pengolahan data secara manual dalam sebuah basis data. Konsep *Data Mining* mulai dikenal pada tahun 1990-an sebagai metode yang efektif untuk menemukan pola dan informasi dalam data, serta mengidentifikasi hubungan antar data. Proses ini memungkinkan pengelompokan data ke dalam satu atau lebih *Cluster*, di mana objek dalam satu *Cluster* memiliki kesamaan yang signifikan. *Data Mining* merupakan bagian dari proses penemuan pengetahuan dalam Knowledge Discovery in Databases (KDD), yang bertujuan untuk memahami informasi yang terkandung dalam basis data [10].

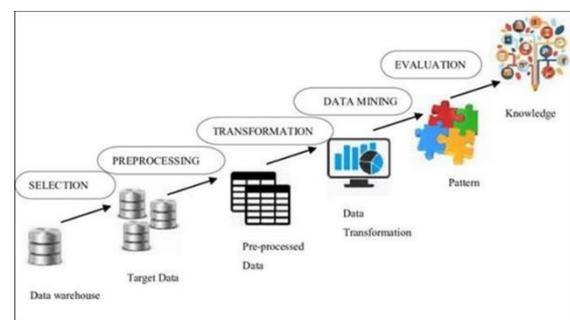
2.5. Bantuan Pangan Non Tunai

Bantuan Pangan Non Tunai (BPNT) adalah program yang diperkenalkan oleh pemerintah Indonesia pada tahun 2017 untuk mendukung keluarga miskin dalam memenuhi kebutuhan

pangan mereka. Tujuan utama BPNT adalah untuk memperkuat ketahanan pangan masyarakat serta meringankan beban ekonomi keluarga kurang mampu [11].

3. METODE PENELITIAN

Knowledge Discovery in Database (KDD) merupakan pendekatan sistematis yang dirancang untuk mengidentifikasi pola atau informasi penting dari data mentah. Metode ini melibatkan serangkaian tahapan yang saling berkesinambungan, mulai dari pengumpulan data, proses transformasi, analisis mendalam, hingga interpretasi hasil [12]. KDD telah banyak diterapkan dalam berbagai bidang penelitian karena kemampuannya dalam menangani dataset berskala besar serta menghasilkan wawasan yang dapat mendukung pengambilan keputusan. Pada penelitian ini, KDD dimanfaatkan untuk melakukan klusterisasi terhadap penerima bantuan sosial pangan di Kota dan Kabupaten Cirebon menggunakan *Algoritma K-Means*. tahap berikutnya adalah menentukan proses *Data Mining*. Pada Gambar 3.1 ditunjukkan tahapan KDD sebagai berikut:



Gambar 3.1 Tahapan KDD

3.1. Data Selection

Proses seleksi data dilakukan dengan mengidentifikasi dan memilih data relevan dari sumber yang tersedia, yaitu data penerima bantuan dari PT. Pos Indonesia di wilayah Kota dan Kabupaten Cirebon. Informasi penting yang dipilih meliputi usia penerima, wilayah domisili, dan jarak dari pusat distribusi. Seleksi data ini bertujuan memastikan bahwa hanya data yang signifikan bagi penelitian digunakan, sehingga proses analisis menjadi lebih fokus dan efektif.

3.2. Data Preprocessing

Tahap ini mencakup pembersihan data dari inkonsistensi, data duplikat, atau nilai kosong. Proses ini penting untuk mengurangi bias dalam analisis dan memastikan kualitas data. Data yang tidak lengkap akan diisi atau dihapus sesuai prosedur yang tepat, sementara variabel seperti usia dan jarak akan dinormalisasi guna memudahkan penerapan *Algoritma K-Means*.

3.3. Data Transformation

Transformasi data bertujuan memformat dan mengkonversi data agar sesuai dengan kebutuhan analisis *Clustering*. Pada tahap ini, data numerik seperti usia dan jarak akan dinormalisasi ke dalam rentang tertentu agar setiap variabel memiliki bobot yang setara dalam proses pembentukan kluster. Selain itu, kategori wilayah tempat tinggal akan dikelompokkan berdasarkan kode atau koordinat geografis untuk mendukung pemisahan kluster yang akurat.

3.4. Data Mining

Data Mining adalah bagian penting dari *Knowledge Discovery in Database (KDD)* yang bertujuan untuk mengekstraksi pola dan informasi bermakna dari dataset. Proses ini diawali dengan pengumpulan data penerima bantuan dari PT. Pos Indonesia, mencakup atribut usia, wilayah, dan jarak distribusi. Atribut signifikan dipilih untuk analisis, sementara data yang tidak lengkap disaring menggunakan *purposive sampling*, dan data duplikat dihapus untuk memastikan keunikan. Atribut nominal dikonversi menjadi numerik agar kompatibel dengan *Algoritma K-Means*, yang kemudian diterapkan untuk mengelompokkan data ke dalam kluster yang homogen berdasarkan atribut tersebut. Data juga diduplikasi untuk pengujian berbagai skenario parameter. Hasil klusterisasi dievaluasi menggunakan *Davies-Bouldin Index (DBI)*, di mana nilai DBI yang rendah menunjukkan kualitas kluster yang baik. Tahapan ini memastikan pengolahan data yang efektif untuk mendukung pengambilan keputusan berbasis data.

3.5. Evaluation

Evaluasi pola yang dihasilkan dari proses *Clustering* dilakukan untuk menilai seberapa

baik setiap kluster mewakili karakteristik penerima bantuan. Metrik DBI digunakan untuk memastikan bahwa kluster yang terbentuk memiliki pemisahan yang baik dan tingkat homogenitas tinggi. Semakin rendah nilai DBI, semakin optimal kualitas kluster yang terbentuk. Selain itu, hasil *Clustering* juga akan dibandingkan dengan informasi distribusi aktual untuk melihat kesesuaian dan efektivitas pengelompokan.

3.6. Knowledge

Hasil akhir dari proses KDD disajikan dalam bentuk tabel, grafik, atau peta tematik untuk memvisualisasikan setiap kluster penerima bantuan. Visualisasi ini memudahkan dalam memahami pola distribusi dan karakteristik penerima di berbagai wilayah. Temuan penelitian akan diinterpretasikan secara komprehensif untuk memberikan rekomendasi bagi pemangku kebijakan, seperti strategi distribusi yang lebih efisien dan tepat sasaran. Pengetahuan yang diperoleh melalui proses ini diharapkan dapat mendukung peningkatan efektivitas program bantuan sosial di Cirebon.

4. HASIL DAN PEMBAHASAN

Pada tahap ini, peneliti menerapkan metode *Clustering* dengan menggunakan *Algoritma K-Means* untuk mengelompokkan penerima bantuan pangan di Kota dan Kabupaten Cirebon. Proses ini bertujuan untuk meningkatkan efektivitas distribusi bantuan dengan membentuk kluster penerima yang relevan berdasarkan atribut tertentu seperti usia, wilayah, dan jarak dari pusat distribusi.

4.1. Data Selection

Pada tahap *Data Selection*, langkah pertama yang dilakukan adalah memilih data yang sesuai untuk analisis. Dalam penelitian ini, data yang digunakan adalah data penerima bantuan pangan yang diperoleh dari PT. Pos Indonesia, yang mencakup atribut seperti usia penerima, wilayah tempat tinggal, dan jarak dari pusat distribusi bantuan. Tujuan dari tahap ini adalah untuk memastikan bahwa hanya data yang relevan dan memiliki informasi yang dibutuhkan untuk analisis *K-Means* yang dipilih. Atribut yang tidak penting atau tidak lengkap akan dihapus, sehingga hanya data yang memenuhi kriteria yang akan digunakan.

Proses ini penting untuk memastikan bahwa analisis dilakukan dengan data yang tepat, mendukung tujuan penelitian dalam mengelompokkan penerima bantuan berdasarkan karakteristik yang relevan.

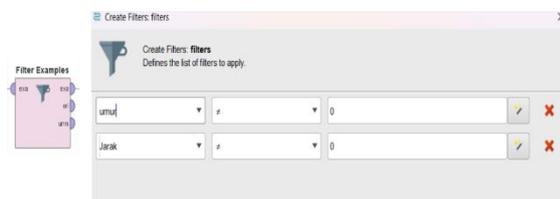
Row No.	nama	kelurahan	kecamatan	umur	Jarak
1	ANAH	ARJAWINAN...	ARJAWINAN...	53	1000
2	ANAH	ARJAWINAN...	ARJAWINAN...	53	1000
3	KHASANAH	ARJAWINAN...	ARJAWINAN...	52	1000
4	KHASANAH	ARJAWINAN...	ARJAWINAN...	69	1000
5	KHASANAH	ARJAWINAN...	ARJAWINAN...	52	1000
6	MASKANAH	ARJAWINAN...	ARJAWINAN...	58	1600
7	MASKANAH	ARJAWINAN...	ARJAWINAN...	45	1600
8	SANTI	ARJAWINAN...	ARJAWINAN...	45	850
9	SASI	ARJAWINAN...	ARJAWINAN...	54	850
10	SASI	ARJAWINAN...	ARJAWINAN...	36	850
11	SRI YATI	ARJAWINAN...	ARJAWINAN...	55	850
12	SRI YATI	ARJAWINAN...	ARJAWINAN...	51	850
13	EMI	BULAK	ARJAWINAN...	60	500
14	EMI	BULAK	ARJAWINAN...	60	500
15	ERNAWATI	BULAK	ARJAWINAN...	35	700
16	ERNAWATI	BULAK	ARJAWINAN...	48	700
17	MELANI	BULAK	ARJAWINAN...	42	500
18	MELANI	BULAK	ARJAWINAN...	42	500

ExampleSet (9,809 examples,0 special attributes,5 regular attributes)

Gambar 4. 1 Data Penerima Bantuan pangan

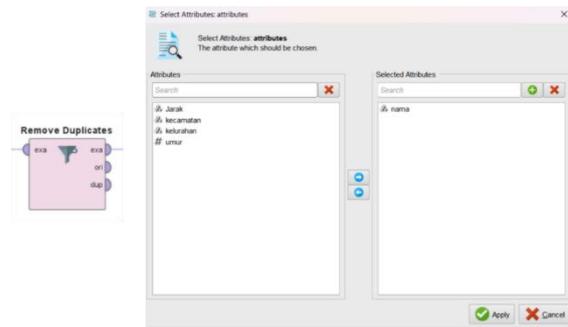
4.2. Data Preprocessing

Tahap *preprocessing* melibatkan pembersihan data dari nilai kosong (missing values) dan duplikasi. Dalam *RapidMiner*, peneliti menggunakan operator *Filter Examples* untuk mengidentifikasi dan menghapus data yang mengandung nilai kosong atau nol. Hal ini bertujuan agar hanya data yang valid digunakan dalam analisis.



Gambar 4. 2 Operator Filter Example

Setelah tahap pembersihan selesai, peneliti menggunakan operator *Remove Duplicates* untuk menghindari terjadinya duplikasi data yang dapat mempengaruhi hasil *Clustering*.



Gambar 4. 3 Operator Remove Duplicates

Jika sudah tersambung jalankan proses, maka data yang terdapat duplikat akan difilter seperti pada gambar 4.4

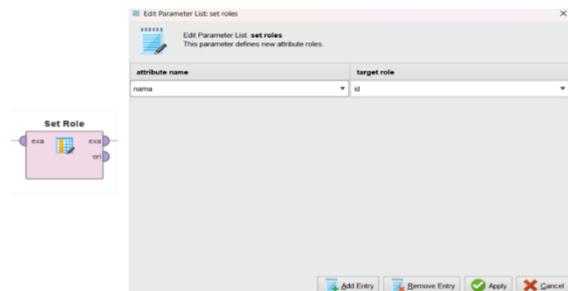
Row No.	nama	kelurahan	kecamatan	umur	Jarak
1	ANAH	ARJAWINAN...	ARJAWINAN...	53	1000
2	KHASANAH	ARJAWINAN...	ARJAWINAN...	52	1000
3	MASKANAH	ARJAWINAN...	ARJAWINAN...	58	1600
4	SANTI	ARJAWINAN...	ARJAWINAN...	45	850
5	SASI	ARJAWINAN...	ARJAWINAN...	54	850
6	SRI YATI	ARJAWINAN...	ARJAWINAN...	55	850
7	EMI	BULAK	ARJAWINAN...	60	500
8	ERNAWATI	BULAK	ARJAWINAN...	35	700
9	MELANI	BULAK	ARJAWINAN...	42	500
10	RASMINI	BULAK	ARJAWINAN...	55	500
11	RATINI	BULAK	ARJAWINAN...	63	700
12	SARIA	BULAK	ARJAWINAN...	52	700
13	SUMINI	BULAK	ARJAWINAN...	62	500
14	SUNARTI	BULAK	ARJAWINAN...	39	700
15	SUTARI	BULAK	ARJAWINAN...	58	700
16	KARNESI	GEYONGAN	ARJAWINAN...	44	900
17	MINA	GEYONGAN	ARJAWINAN...	75	400
18	RUMINI	GEYONGAN	ARJAWINAN...	40	400

ExampleSet (2,226 examples,0 special attributes,5 regular attributes)

Gambar 4. 4 Hasil Remove Duplicates

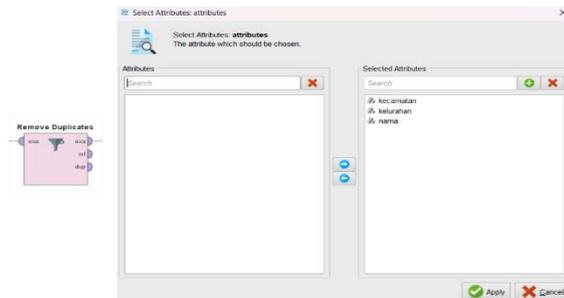
4.3. Data Transformation

Transformasi data dilakukan untuk memastikan data sesuai dengan kebutuhan algoritma. Dalam penelitian ini, operator *Set Role* digunakan untuk menetapkan peran atribut nama menjadi id.



Gambar 4. 5 Operator Set Role

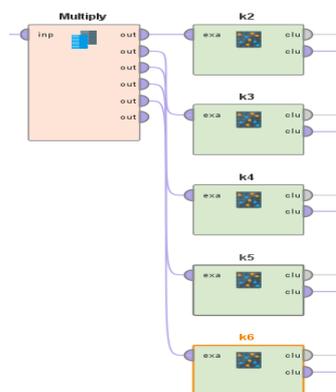
Pada tahapan berikutnya yaitu menggunakan operator *Nominal to Numeric* untuk mengonversi data yang berbentuk kategori (nominal) menjadi data numerik agar dapat diterima oleh *Algoritma K-Means*. Proses ini memastikan bahwa semua data telah berada dalam format yang tepat untuk dianalisis.



Gambar 4. 6 Operator Nominal To Numerical

4.4. Data Mining

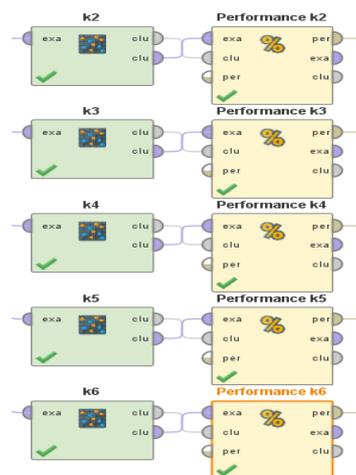
Pada tahap ini, diterapkan *Algoritma K-Means* menggunakan metode *Clustering* untuk menganalisis data yang telah diproses. Langkah pertama adalah menentukan jumlah k pada kluster, yang bertujuan untuk menentukan jumlah kluster yang paling efektif dalam proses *Clustering*. Untuk menguji akurasi kluster, dilakukan pengujian dengan lima kelompok berbeda, yaitu kelompok satu dengan jumlah k=2, kelompok dua dengan k=3, kelompok tiga dengan k=4, kelompok empat dengan k=5, dan kelompok lima dengan k=6. Untuk menambahkan lebih dari satu operator kluster pada *RapidMiner*, digunakan operator *multiply*, yang berfungsi untuk menduplikasi operator dan memungkinkan analisis dengan beberapa skenario parameter sekaligus.



Gambar 4. 7 Operator Multiply

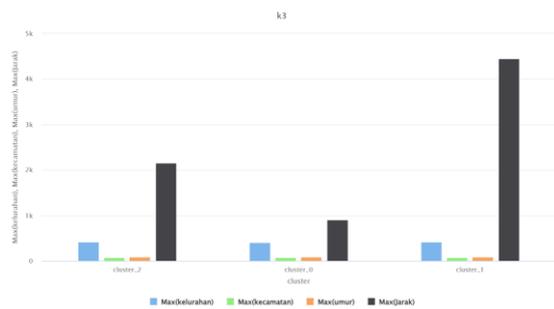
4.5. Evaluation

Pada tahap ini, dilakukan evaluasi hasil klusterisasi untuk menilai kinerja klusterisasi data penerima bantuan pangan dengan berbagai nilai k (k=2, k=3, k=4, k=5, k=6). Evaluasi bertujuan untuk menentukan jumlah kluster yang optimal, yang menghasilkan pemisahan yang baik antar-kluster dan kedekatan yang kuat dalam setiap kluster. Pengujian dilakukan dengan menggunakan operator *Cluster Distance Performance* di *RapidMiner* dan perhitungan *Davies-Bouldin Index (DBI)* untuk setiap nilai k. DBI mengukur kedekatan dalam kluster dan pemisahan antar-kluster, di mana nilai DBI yang lebih rendah menunjukkan kluster yang terbentuk dengan baik. Jumlah kluster dengan DBI terendah akan dipilih sebagai jumlah kluster optimal, yang mendukung segmentasi penerima bantuan yang lebih tepat dan meningkatkan ketepatan sasaran distribusi bantuan sosial.



Gambar 4. 8 Operator Cluster distance performance

Berdasarkan hasil pada masing-masing nilai *DBI (Davies Bouldin Index)* hasil yang paling rendah diantara masing-masing *Cluster* adalah dengan jumlah k=3 dengan nilai DBI sebesar 0.505, maka *Cluster* dengan jumlah k=3 lebih baik dibandingkan dengan *Cluster* jumlah k=2, k=4, k=5 dan k=6. Maka dari itu nilai k=3 akan dilakukan analisis untuk melihat *Cluster* dengan jumlah kelurahan, kecamatan, jarak, dan umur tertinggi dan terendah. Adapun hasil dari analisis *Cluster* didapatkan hasil pada



Gambar 4. 9 Grafik Maximum Cluster

Berdasarkan grafik tersebut, dapat disimpulkan bahwa ketiga *Cluster* menunjukkan perbedaan signifikan dalam distribusi kelurahan, kecamatan, jarak, dan umur tertinggi. *Cluster* 0 memiliki jarak terjauh 900 meter, umur tertua 93 tahun, dengan 71 kecamatan dan 414 kelurahan, menunjukkan wilayah yang lebih terpusat dan mudah diakses dengan populasi penerima bantuan yang lebih tua serta kepadatan area yang memungkinkan distribusi lebih efisien. *Cluster* 1, dengan jarak terjauh 4.450 meter, umur tertua 83 tahun, 71 kecamatan, dan 415 kelurahan, mencerminkan wilayah yang sangat luas dan terpencil, sehingga memerlukan perencanaan logistik lebih kompleks untuk menjangkau seluruh penerima bantuan, meskipun usia penerima relatif lebih muda dibandingkan dengan *Cluster* 0. *Cluster* 2 memiliki jarak terjauh 2.150 meter, umur tertua 88 tahun, 71 kecamatan, dan 415 kelurahan, mencerminkan area yang luas namun lebih mudah dijangkau dibandingkan *Cluster* 1, sehingga distribusi masih dapat dilakukan secara efisien meskipun tidak serumit *Cluster* 1. Secara keseluruhan, *Cluster* 1 memerlukan rencana distribusi paling kompleks karena jarak yang jauh, sementara *Cluster* 0 dan *Cluster* 2 memungkinkan distribusi yang lebih sederhana dan efisien. Pendekatan distribusi yang disesuaikan di setiap *Cluster* akan memastikan penyaluran bantuan yang tepat sasaran di seluruh wilayah penelitian.

4.6. Knowledge

Berdasarkan analisis Davies-Bouldin Index (DBI), pembagian data menjadi tiga *Cluster* ($k=3$) optimal dengan DBI 0,505, menunjukkan pemisahan antar-klaster yang baik. *Cluster* 0 (jarak maksimal 900 meter, usia rata-rata 93 tahun) membutuhkan distribusi efisien untuk

lansia. *Cluster* 1 (jarak 4.450 meter, usia rata-rata 83 tahun) memerlukan logistik kompleks untuk wilayah terpencil. *Cluster* 2 (jarak 2.150 meter, usia rata-rata 88 tahun) memungkinkan distribusi efisien dengan penyesuaian untuk area luas. Pendekatan distribusi spesifik diperlukan untuk memenuhi kebutuhan tiap *Cluster*.

5. KESIMPULAN

- Penelitian ini berhasil menentukan titik pusat (*centroid*) optimal dengan $K=3$ untuk mengelompokkan penerima bantuan sosial di Cirebon. Hasil klusterisasi menunjukkan bahwa *Cluster* 0 mencakup wilayah padat penduduk dekat lokasi distribusi, memudahkan distribusi. *Cluster* 1 meliputi area yang luas dengan jarak distribusi jauh, membutuhkan strategi lebih kompleks. *Cluster* 2 mencakup wilayah luas yang lebih mudah dijangkau, memungkinkan distribusi lebih efisien dibandingkan *Cluster* 1.
- Penelitian ini berhasil memastikan jarak antar-klaster yang signifikan dalam proses *K-Means*, menghasilkan tiga klaster dengan karakteristik berbeda berdasarkan kelurahan, kecamatan, usia, dan jarak distribusi. *Cluster* 0 mencakup penerima berusia tua yang tinggal dekat dengan lokasi distribusi, memudahkan penyaluran bantuan secara cepat. *Cluster* 1 mencakup penerima muda di area dengan jarak distribusi terjauh, membutuhkan strategi khusus. *Cluster* 2 mencakup penerima usia menengah dengan jarak distribusi sedang, memudahkan penyaluran dengan upaya yang moderat.
- Nilai K yang optimal pada pengujian performa DBI setelah dilakukan lima kali pengujian dengan menetapkan nilai $K=2$, $K=3$, $K=4$, $K=5$ dan $K=6$. Didapatkan hasil DBI untuk $K=2$ mendapatkan nilai 0.580, $K=3$ mendapatkan nilai

0.505, K=4 mendapatkan nilai 0.563, K=5 mendapatkan nilai 0.619 dan K=6 mendapatkan nilai 0.654. Dengan demikian didapatkan nilai K yang optimal yaitu K=3 dengan nilai 0.505 lebih rendah dengan nilai K yang lain.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada pihak-pihak terkait yang telah memberi dukungan terhadap penelitian ini.

DAFTAR PUSTAKA

- [1] B. Mahesh, "Machine Learning Algorithms - A Review," *Int. J. Sci. Res.*, vol. 9, no. 1, pp. 381–386, 2020, doi: 10.21275/art20203995.
- [2] A. Abdulhafedh, "Incorporating *K-Means*, Hierarchical *Clustering* and PCA in Customer Segmentation," *J. City Dev.*, vol. 3, no. 1, pp. 12–30, 2021, doi: 10.12691/jcd-3-1-3.
- [3] A. A. H. Tasyifa Nafsiah Muthmainnah, Apriade Voutama, "MENGUNAKAN ALGORITMA *K-MEANS*," vol. 8, no. 3, pp. 3916–3919, 2024.
- [4] B. Wang *et al.*, "Tortuous Pore Path Through the Glaucomatous Lamina Cribrosa," *Sci. Rep.*, vol. 8, no. 1, 2018, doi: 10.1038/s41598-018-25645-9.
- [5] D. Suhardinata, A. K. Ningsih, F. Kasyidi, U. Jenderal, and A. Yani, "Klasterisasi Data Penduduk Untuk Penerimaan Bantuan Pangan Non Tunai (BPNT) Menggunakan *K-Means* (Studi Kasus : Desa Tanimulya Bandung Barat)," *IJESPG (International J. Eng. Econ. Soc. Polit. Gov.)*, vol. 1, no. 3, pp. 221–228, 2023, [Online]. Available: <http://ijespgjournal.org/index.php/ijespg/article/view/55>
- [6] N. Nurahman and J. Susanto, "Klasterisasi Data Penerima Bantuan Langsung Tunai Menggunakan *Algoritma K-Means*," *JURIKOM (Jurnal Ris. Komputer)*, vol. 10, no. 2, p. 461, 2023, doi: 10.30865/jurikom.v10i2.5807.
- [7] I. D. Anjani and A. Bahtiar, "Penerapan *Algoritma K-Means Clustering* untuk Mengelompokkan Penerima Bantuan Sosial Tunai (BST) di Jawa Barat," *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 8, no. 3, pp. 2743–2747, 2024.
- [8] F. Febriansyah and S. Muntari, "Penerapan *Algoritma K-Means* untuk Klasterisasi Penduduk Miskin pada Kota Pagar Alam," *JISKA (Jurnal Inform. Sunan Kalijaga)*, vol. 8, no. 1, pp. 66–77, 2023, doi: 10.14421/jiska.2023.8.1.66-77.
- [9] Z. Sitorus, "Penerapan *Data Mining* Untuk *Clustering* Penduduk Miskin Di Kota Tanjungbalai Menggunakan Metode *Algoritma K-Means*," *J. Sci. Soc. Res.*, vol. 4307, no. 1, pp. 212–218, 2024, [Online]. Available: <http://jurnal.goretanpena.com/index.php/JSSR>
- [10] N. Siti Paridah and M. Martanto, "Klasterisasi Penerima Dana Bantuan Program Keluarga Harapan Menggunakan Metode *K-Means* Pada Desa Gereba," *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 8, no. 1, pp. 1036–1043, 2024, doi: 10.36040/jati.v8i1.8873.
- [11] F. Juliawati, R. Buaton, R. Saragih, and S. Kaputama, "Pengelompokan *Data Mining* Penerimaan Bantuan Pangan Non Tunai (BPNT) Menggunakan Metode *Clustering* (Studi Kasus : Kantor Desa Payabakung Hampan Perak)," *J. Comput. Sci. Inf. Technol. E-ISSN*, vol. 3, no. 2, p. 69, 2023.
- [12] D. Fitriyani, M. Jajuli, and G. Garno, "Implementasi *Algoritma K-Means* Untuk Klasterisasi Dalam Pengelolaan Persediaan Obat (Studi Kasus : Apotek Naza)," *J. Inform. dan Tek. Elektro Terap.*, vol. 12, no. 3, pp. 2841–2848, 2024, doi: 10.23960/jitet.v12i3.4921.