

CLUSTERING PENDUDUK MISKIN UNTUK PENERIMA BANTUAN SOSIAL MENGGUNAKAN ALGORITMA K-MEANS (STUDI KASUS: DESA GREGED)

Muhamad Subur¹, Martanto², Umi Hayati³

^{1,3}Teknik Informatika, ²Manajemen Informatika STMIK IKMI Cirebon
Jalan Perjuangan No.10 B Majasem, Kota Cirebon, Indonesia
Email: muhamadsubur211@gmail.com

ABSTRAK

Penyaluran bantuan sosial merupakan strategi vital yang diterapkan oleh pemerintah dan organisasi kemanusiaan untuk memberikan dukungan finansial atau materi kepada individu atau kelompok yang memerlukan. Salah satu kesulitan yang terkadang dihadapi Pemerintah dalam proses penanganan kemiskinan adalah proses pembagian bantuan sosial yang tidak merata dan tidak tepat sasaran. Ini disebabkan karena validasi data sering diabaikan sehingga sering menimbulkan data yang tidak akurat. Tujuan dari penelitian ini penulis mencoba untuk menerapkan Teknik clustering dengan menggunakan algoritma K-Means sebagai metode penyelesaian dalam mengelompokkan data penduduk di Desa Greged yang memang dikatakan tergolong penduduk yang berhak menerima bantuan sosial. . Evaluasi dilakukan di software rapidminer dengan operator cluster distance performance dengan hasil 10 percobaan dengan menggunakan algoritma K-Means Clustering menunjukkan bahwa nilai K yang Paling optimal adalah K=10. Nilai DBI untuk K=10 adalah 0,049, yang menunjukan bahwa cluster-cluster penerima bantuan sosial memiliki tingkat kekompakan yang tinggi dan tingkat kesenjangan antar cluster yang rendah. Hal ini mengindikasikan bahwa penerima bantuan sosial di Desa Greged dapat di kelompokkan menjadi 10 cluster yang berbeda berdasarkan karakteristik demografis, ekonomi, dan sosial mereka.

Kata kunci : Data mining, Clusteirng, Penduduk Miskin, Algoritma K-Means, Knowledge Discovery in Database (KDD)

1. PENDAHULUAN

Penyaluran bantuan sosial merupakan strategi vital yang diterapkan oleh pemerintah dan organisasi kemanusiaan untuk memberikan dukungan finansial atau materi kepada individu atau kelompok yang memerlukan[1]. Tujuan utama dari pemberian bantuan sosial adalah untuk meringankan beban ekonomi dan meningkatkan kesejahteraan penerima, khususnya mereka yang mengalami kerentanan ekonomi atau menghadapi situasi krisis. Jenis bantuan sosial meliputi bantuan tunai, bantuan pangan, pemberian perumahan, dukungan pendidikan, atau layanan kesehatan, dan sering ditujukan kepada kelompok seperti keluarga miskin, penyandang disabilitas, atau korban bencana alam[2]. Proses penyaluran bantuan sosial memerlukan perencanaan yang teliti, identifikasi penerima yang akurat, serta evaluasi berkelanjutan untuk memastikan keefektifan dalam memberikan dampak positif bagi masyarakat yang membutuhkan bantuan[3].

Pelaksanaan bantuan sosial melibatkan serangkaian langkah kompleks, mulai dari perumusan kebijakan, pengidentifikasian penerima, hingga distribusi yang adil. Meski Pemerintah telah meluncurkan berbagai program untuk mengurangi kemiskinan seperti PKH, Raskin, dan BIDIKMISI, tetapi tantangan kemiskinan tetap ada. Salah satu hambatan utama adalah distribusi bantuan yang tidak efisien dan tidak sesuai sasaran akibat data yang kurang valid[4]. Oleh karena itu, studi ini bertujuan untuk mengeksplorasi solusi untuk masalah kemiskinan di Desa Greged. Mengadaptasi diri dengan

perubahan sosial dan ekonomi, penting bagi pemerintah untuk mengembangkan strategi yang inovatif, termasuk memanfaatkan teknologi informasi dan analisis data. Penelitian tentang klasterisasi penerima bantuan dengan algoritma K-Means di Desa Greged, Kabupaten Cirebon, menjadi penting karena potensinya dalam meningkatkan akurasi penentuan penerima manfaat dan efisiensi penggunaan dana. Dengan pendekatan ini, diharapkan tercipta sistem bantuan yang lebih adaptif, efisien, dan mampu memenuhi kebutuhan masyarakat dengan lebih efektif[5].

Tujuan dari penelitian ini penulis mencoba untuk menerapkan Teknik clustering dengan menggunakan algoritma K-Means sebagai metode penyelesaian dalam mengelompokkan data penduduk di Desa Greged yang memang dikatakan tergolong penduduk yang berhak menerima bantuan sosial. Algoritma K-Means Clustering atau metode analisis kluster yang bertujuan untuk mengelompokkan data menjadi kluster-kluster yang serupa berdasarkan kesamaan atribut atau karakteristik tertentu[6]. Dalam penelitian ini, Algoritma K-Means akan digunakan untuk mengelompokkan penduduk Desa Greged ke dalam kelompok yang memiliki kesamaan dalam hal kondisi ekonomi dan sosial. Banyak metode clustering yang telah diusulkan oleh para ahli, salah satunya adalah K-Means. Metode K-Means merupakan Algoritma klasterisasi yang paling banyak digunakan dalam berbagai aplikasi kecil hingga menengah karena kemudahan mengaplikasikannya[7].

Dalam penelitian ini, digunakan metode penelitian yang bersumber dari data penduduk Desa Greged tahun 2020. Teknik pengumpulan data yang dilakukan dalam penelitian ini yaitu observasi dan wawancara. Dalam metode ini penulis mengumpulkan data pada bulan Oktober 2023 dengan melakukan observasi / pengamatan langsung ke Kantor Desa Greged dan melakukan tanya jawab/wawancara yang dilakukan dengan staf Kantor Desa Greged untuk mengetahui gambaran tentang kondisi perekonomian Penduduk. Teknik analisis data dalam penelitian ini menggunakan metode Knowledge Discovery in Database (KDD)[8]. Data yang dari observasi dan 14 wawancara kemudian diolah menggunakan metode KDD untuk mendapatkan informasi yang lebih mendalam tentang kondisi perekonomian penduduk[9].

Penelitian ini diharapkan dapat menghasilkan informasi yang lebih terperinci dan berharga dalam upaya mengatasi masalah kemiskinan di wilayah Desa Greged. Sehingga dapat membantu dan mendukung Kantor Desa Greged dalam pengambilan keputusan yang lebih efisien dalam penyaluran bantuan sosial dan perencanaan program-program pemberdayaan ekonomi yang lebih efektif untuk penduduk miskin. Selain itu, penelitian ini diharapkan dapat menyajikan temuan dalam bentuk laporan penelitian yang dapat digunakan oleh pemerintah setempat atau instansi terkait dan dapat memberikan wawasan yang lebih dalam tentang penduduk miskin di Desa Greged. Penelitian ini juga diharapkan dapat menjadi pendorong bagi Masyarakat Desa Greged untuk mewujudkan desa yang sejahtera. Namun, masih ada permasalahan yang sering terjadi pada saat pembagian bantuan sosial adalah kurang tepat sasaran. Sehingga dapat memunculkan kecurigaan masyarakat terhadap Lembaga yang mengurus tentang bantuan sosial[10]. Diharapkan penelitian ini dapat memberikan informasi yang berguna bagi pihak yang terkait dalam penyaluran dana dan juga penentuan Masyarakat yang menerima manfaat bantuan sosial di Desa Greged.

2. TINJAUAN PUSTAKA

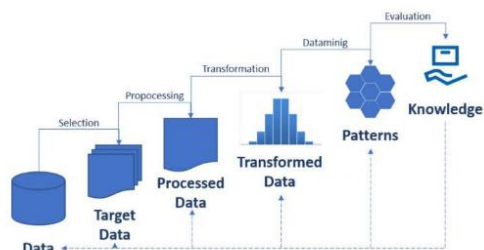
Penelitian ini memanfaatkan pendekatan review literature dengan metode systematic reviews untuk mengkaji artikel-artikel yang relevan dengan tema penelitian. Penelitian yang dilakukan dalam lima tahun terakhir oleh para peneliti sebelumnya memiliki relevansi signifikan dengan artikel yang sedang diteliti saat ini. Kajian-kajian terdahulu tersebut memberikan kontribusi penting dalam pemahaman yang lebih mendalam terkait isu yang sedang dibahas dalam penelitian ini. Seperti penelitian yang dilakukan oleh Suhartini dan Ria Yuliani[11] berjudul "Penerapan Data Mining untuk Mengcluster Data Penduduk Miskin Menggunakan Algoritma K-Means di Dusun Bagik Endep Sukamulia Timur" menggunakan data mining dengan algoritma K-means untuk mengelompokkan penduduk miskin di Sukamulia Timur ke dalam cluster. Mereka

mengumpulkan data tahun 2019 sebanyak 200 entri data yang menampilkan sembilan atribut seperti nama penduduk, pekerjaan, pendapatan per bulan, jumlah anak bersekolah, jumlah anggota keluarga, dan masih banyak lagi. Mereka mengamati tiga klaster: Klaster 1 (18 warga), populasi ekonomi tinggi; Klaster 2 (72 warga), penduduk ekonomi menengah; Klaster 3 (110 jiwa), penduduk ekonomi rendah. Pandangan distributif ini membantu pemerintah desa dalam memberikan bantuan secara efisien dan akurat kepada mereka yang membutuhkan. Para penulis menyimpulkan efektivitas cluster terkait dengan penentuan centroid awal, dengan pemilihan centroid yang berbeda menghasilkan hasil yang beragam. Tujuan dari strategi ini adalah untuk memberikan alat kepada pemerintah Desa Sukamulia Timur untuk membantu pengambilan keputusan dan memecahkan masalah terkait pencatatan penduduk miskin secara akurat agar distribusi bantuan menjadi efisien. Pada penelitian yang dilakukan oleh Baskoro[1] dengan judul "Klasifikasi Penduduk Kurang Mampu Dengan Metode K-Means untuk Optimalisasi Program Bantuan Sosial", fokus pada optimalisasi penyaluran bantuan sosial melalui penggunaan metode K-Means untuk mendklasifikasikan populasi yang kurang mampu. Penelitian ini dilakukan dalam konteks Indonesia, dimana tingkat kemiskinan cenderung meningkat seiring dengan pandemi COVID-19. Tujuan dari penelitian ini adalah untuk memastikan efektivitas dan keakuratan penyaluran program bantuan sosial yang dirasakan kurang terarah saat ini. Metode yang digunakan adalah algoritma K-Means, yang membantu mengelompokkan data menjadi tiga kelompok berdasarkan empat variabel: Pendidikan, Pekerjaan, Tanggungan, dan Pengeluaran perbulan. Hal ini diharapkan dapat membantu mengungkap kemiskinan yang lebih akurat, serta mendistribusikan bantuan sosial yang lebih tepat sasaran dan efisien. Proses K-Means menghasilkan tiga kluster dari penduduk: yang sangat layak mendapatkan bantuan, yang layak mendapatkan bantuan, dan yang kurang layak mendapatkan bantuan. Pada penelitian berjudul "Implementasi Data Mining Dalam Mengelompokkan Jumlah Penduduk Miskin Berdasarkan Provinsi Menggunakan Algoritma K-Means" oleh Yuni Radana Sembiring, S Saifullah, dan Riki Winanjaya[12]. Penelitian ini fokus pada penerapan teknik data mining untuk mengelompokkan jumlah penduduk miskin di Indonesia berdasarkan provinsi. Metode yang digunakan dalam penelitian ini adalah metode K-Means Clustering. Para peneliti menggunakan data jumlah penduduk miskin dari tahun 2007-2019 yang diperoleh dari Badan Pusat Statistik. Dalam penentuan cluster, peneliti menetapkan tiga kelompok, yaitu cluster tingkat tinggi, sedang, dan rendah. Proses pengelompokan penduduk miskin dilakukan dalam beberapa langkah. Pertama, tentukan jumlah data yang akan dikumpulkan dan jumlah cluster (k) yang diinginkan. Kedua, menentukan nilai pusat (centroid) awal yang ditentukan secara acak. Ketiga, menghitung

jarak setiap data jumlah penduduk miskin terhadap pusat cluster. Hasil dari penelitian ini menunjukkan bahwa dengan pendekatan data mining dan algoritma K-Means, peneliti berhasil mengelompokkan jumlah penduduk miskin di tiap provinsi Indonesia. Umumnya, metode K-Means digunakan dalam berbagai aplikasi kecil hingga menengah karena kemudahan dalam pengaplikasiannya. Dari hasil penelitian ini, peneliti berharap mampu memberikan informasi atau masukan bagi pemerintah dalam mengambil langkah yang tepat sebagai upaya untuk menurunkan jumlah kemiskinan di setiap provinsi yang ada di Indonesia.

3. METODE PENELITIAN

Pada penelitian ini, metode yang diambil adalah *Knowledge in Database (KDD)* yaitu proses yang bertujuan untuk menggali dan menganalisis data yang sangat besar menjadi informasi yang berguna untuk pengetahuan. Algoritma yang digunakan dalam melakukan clustering penduduk miskin untuk penerimaan bantuan sosial menggunakan metode algoritma *K-Means*[2].



Gambar 1 Tahapan Metode KDD

Tahapan *Knowledge Discovery in Database (KDD)* secara garis besar dapat dijelaskan sebagai berikut.

1. **Selection:** Pilih sub kumpulan data yang relevan untuk dianalisis.
2. **Preprocessing:** Bersihkan dan ubah data agar siap untuk dianalisis.
3. **Transformation:** Mengubah data menjadi formal yang cocok untuk penambahan data.
4. **Data Mining:** Terapkan teknik dan Algoritma penambahan data ke data untuk mengekstrak informasi dan wawasan yang bermanfaat.
5. **Evaluation:** Evaluasi hasil untuk memastikan bahwa pengetahuan yang diekstraksi bermanfaat, akurat, dan bermakna.

4. HASIL DAN PEMBAHASAN

4.1. Selection

Pada tahap awal, peneliti melakukan seleksi data dari sumber data sekunder yang disimpan oleh kantor Desa Greded. Data ini mencakup 4996 catatan dengan 17 variabel atau atribut, termasuk RT, RW, Dusun, Kode Keluarga, Nama Kepala Keluarga, Jumlah

Tanggungan, NIK, Nama Anggota Keluarga, Tanggal Lahir, Status, Agama, Golongan Darah, Kewarganegaraan, Etnis, Pendidikan, Pekerjaan, dan Penghasilan.

Gambar 2 Data sebelum diseleksi

Dalam studi ini, peneliti memilih untuk fokus pada data per kepala keluarga. Dari seleksi tersebut, ditemukan 1404 contoh data yang kemudian disimpan dalam sebuah file baru.

Gambar 3 Data setelah diseleksi

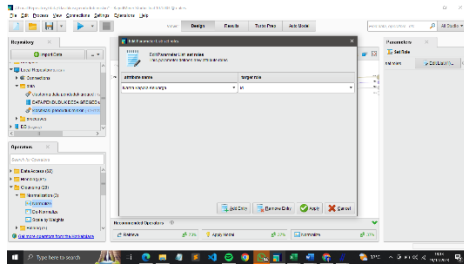
4.2. Preprocessing

Data yang digunakan dalam penelitian ini adalah data yang sudah lengkap. Setelah dilakukan pengecekan pada hasil statistik data ini tidak terdapat missing value.

Gambar 4 Proses pengecekan missing values

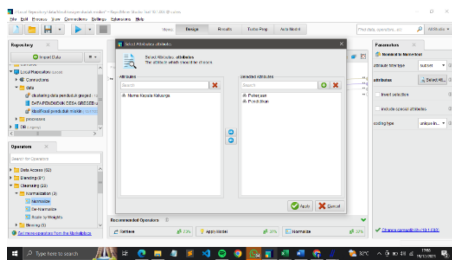
4.3. Transformation

Dalam tahap transformasi data, diperlukan pembentukan set role untuk mengubah nama kepala keluarga menjadi ID. Tujuan dari pembentukan set role ini adalah untuk memberikan identifikasi yang khas bagi setiap kepala keluarga, memfasilitasi proses analisis, dan menjamin kestabilan dalam pengurusan data.



Gambar 5 Operator set role

Dalam langkah ini, digunakan operator "Nominal to Numerical" untuk mengkonversi data dari format nominal menjadi bentuk numerik, memungkinkan rapidminer untuk memproses dan menganalisis data dengan lebih efisien. Atribut yang diubah menjadi format numerik meliputi Pendidikan dan Pekerjaan.

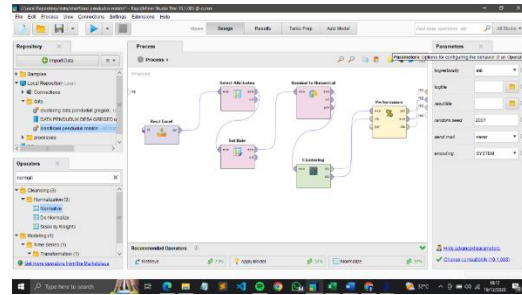


Gambar 6 Proses Nominal to Numerical

Dalam kategori pendidikan, nilai 0 mengindikasikan lulusan SLTA atau setara, nilai 1 untuk lulusan SD atau setara, 2 untuk lulusan SLTP atau setara, 3 untuk lulusan D-3 atau setara, 4 untuk individu yang tidak menyelesaikan pendidikan, 5 untuk lulusan S-1 atau setara, 6 untuk lulusan D-2 atau setara, 7 untuk mereka yang tidak pernah bersekolah, 8 untuk individu yang tidak bisa membaca dan menulis, dan 9 untuk lulusan D-1 atau setara. Sementara untuk kategori pekerjaan, nilai 0 mencerminkan status sebagai Pegawai Negeri Sipil, nilai 1 untuk Buruh Harian Lepas, 2 untuk Wiraswasta, 3 untuk Pedagang Keliling, 4 untuk Pedagang Barang, 5 untuk Buruh Tani, 6 untuk Peternak, 7 untuk Karyawan Perusahaan, 8 untuk Karyawan Honorer, 9 untuk Petani, 10 untuk Pemilik Warung, 11 untuk Perangkat Desa, 12 untuk individu tanpa pekerjaan tetap, 13 untuk Guru Swasta, 14 untuk Wartawan, 15 untuk Sopir, 16 untuk Tukang Jahit, 17 untuk Tukang Listrik, 18 untuk Satpam, 19 untuk Purnawirawan atau pensiunan, dan 20 untuk Pengusaha Perdagangan.

4.4. Data Mining

Dalam tahapan ini, penulis menerapkan teknik data mining menggunakan algoritma K-Means. Algoritma K-Means adalah pendekatan untuk mengelompokkan data berdasarkan jarak antar data, yang mengorganisir data ke dalam sejumlah cluster yang telah ditentukan. Penulis melakukan eksperimen mulai dari K2 hingga K10 dengan batasan iterasi maksimal sebanyak 10.



Gambar 7 Proses Clustering

Nilai cluster (K) yang paling optimal adalah yang nilai Davies Bouldin Index nya mendekati 0, dan sudah dilakukan dengan maksimal 10 kali iterasi untuk mengetahui hasil Davies Bouldien Index yang mendekati 0.

Algoritma k-means clustering adalah metode pengelompokan data non-hirarkis yang bekerja dengan mengidentifikasi objek yang akan dikelompokkan. Algoritma ini relatif sederhana dan bisa digunakan baik untuk data berjumlah kecil maupun besar. Pada tahap awal, pusat setiap cluster ditentukan secara acak. Kemudian, jarak antara setiap data dengan pusat cluster dihitung. Proses pengelompokan ini berlangsung berulang kali hingga cluster mencapai stabilitas.

- Tentukan jumlah *cluster*
- Hitung jarak dari setiap data ke dalam cluster dengan menggunakan persamaan ukuran jarak yang paling dekat, yaitu Euclidean distance.

$$D_i(X_1, X_2) = \|x_2 - x_1\| = \sqrt{\sum_{j=1}^p \{x_{2j} - x_{1j}\}^2} \quad (1)$$

Dimana $D_i(X_1, X_2)$ adalah jarak diantara data ke-I dan data ke-j

x_{2j} adalah jarak kordinat data x_2 pada dimensi j
 x_{1j} adalah kordinat data x_1 pada dimensi j dan P dimensi data

- Kelompokkan data ke dalam *cluster* dengan jarak yang paling pendek dengan menggunakan rumus pada persamaan :

$$V_{ij} = \frac{\sum_{j=1}^{N_i} x_{kj}}{N_i} \quad (2)$$

V_{ij} adalah data cluster ke - i kolom j

x_{kj} adalah data ke -k kolom ke - j

N_i banyaknya anggota cluster ke - j

- Hitung pada pusat *cluster* yang baru menggunakan persamaan
- Kelompokkan objek berdasarkan jarak ke centroid terdekat. Ulangi langkah ke-2, lakukan iterasi hingga centroid bernilai optimal.

Davies Bouldin

Davies Bouldin: 0.131

Gambar 8 Hasil DBI K=2

Davies Bouldin

Davies Bouldin: 0.074

Gambar 9 Hasil DBI K=3

Davies Bouldin

Davies Bouldin: 0.076

Gambar 10 Hasil DBI K=4

Davies Bouldin

Davies Bouldin: 0.103

Gambar 11 Hasil DBI K=5

Davies Bouldin

Davies Bouldin: 0.075

Gambar 12 Hasil K=6

Davies Bouldin

Davies Bouldin: 0.066

Gambar 13 Hasil K=6

Davies Bouldin

Davies Bouldin: 0.063

Gambar 14 Hasil DBI K=8

Davies Bouldin

Davies Bouldin: 0.059

Gambar 15 Hasil DBI K=9

Davies Bouldin

Davies Bouldin: 0.049

Gambar 16 Hasil DBI K=10

Setelah eksplorasi dari K=2 hingga K=10, ditemukan bahwa DBI paling mendekati nol dengan nilai 0,049 tercapai saat K=10. Dari hasil ini, kesimpulannya adalah K=10 merupakan pilihan terbaik untuk nilai K.

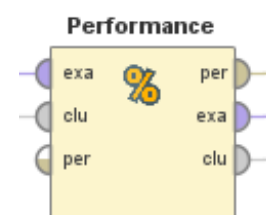
Cluster Model

```
Cluster 0: 431 items
Cluster 1: 51 items
Cluster 2: 210 items
Cluster 3: 1 items
Cluster 4: 216 items
Cluster 5: 28 items
Cluster 6: 1 items
Cluster 7: 412 items
Cluster 8: 18 items
Cluster 9: 36 items
Total number of items: 1404
```

Gambar 17 Hasil Cluster K=10

Berdasarkan standar yang diterapkan oleh Desa Greded, keluarga dengan pendapatan antara Rp. 500.000 hingga Rp. 1.000.000 dikategorikan sebagai keluarga dengan status ekonomi menengah ke bawah. Dalam cluster ini, mereka dianggap sangat membutuhkan bantuan sosial. Sementara keluarga dengan pendapatan antara Rp. 1.500.000 hingga Rp. 2.000.000 dianggap sebagai keluarga dengan status ekonomi menengah, dan di dalam cluster ini, mereka adalah kandidat yang layak untuk menerima bantuan sosial. Di sisi lain, keluarga dengan pendapatan antara Rp. 2.500.000 hingga Rp. 5.000.000 dikategorikan sebagai keluarga dengan status ekonomi menengah ke atas, sehingga dianggap kurang memerlukan bantuan sosial.

4.5. Evaluation



Gambar 18 Operator Performance

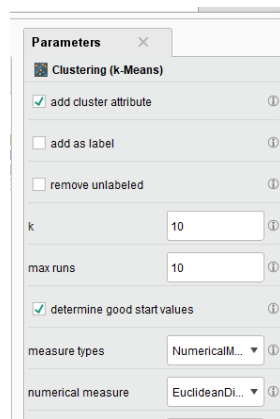
Setelah dilakukan proses evaluasi didapatkan tingkat performance yang paling baik terdapat pada

$K(10) = 0.049$, sehingga dipilih K yang akan digunakan yaitu $K=10$ dengan performance vector sebagai berikut.

Tabel 1Tabel jumlah nilai DBI K2-K10

Percobaan	Jumlah Cluster	Nilai DBI
1	K2	0,131
2	K3	0,074
3	K4	0,076
4	K5	0,103
5	K6	0,075
6	K7	0,066
7	K8	0,063
8	K9	0,059
9	K10	0,049

Proses clustering dilakukan dengan 10 kali iterasi dan pada parameter measure types diterapkan Numerical Measures dan pada Numerical Measure diterapkan Euclidean Distance.



Gambar 19 Operator Parameters

Berikutnya, dilakukan penentuan pusat cluster/centroid. Dalam penelitian ini, didapat centroid pada masing-masing kluster sebagai berikut.

Attribute	cluster_0	cluster_1	cluster_2	cluster_3	cluster_4	cluster_5	cluster_6	cluster_7	cluster_8	cluster_9
Pendidikan	1.460	2.833	1.317	1.386	1.225	1.429	2.539	1.451	1.474	1.417
Pekerjaan	5.180	4	1.357	2.588	1.241	8.179	0	3.480	1.782	2.250
Jumlah Tawar...	3.580	3.022	3.358	3.500	3.185	3.571	4.839	2.824	3.119	3.889
Penghasilan	1530000	3000000	500000	999074.074	7167192.297	1993857.143	4500000	2500000	600000	1200000

Gambar 20 Pusat Centroid

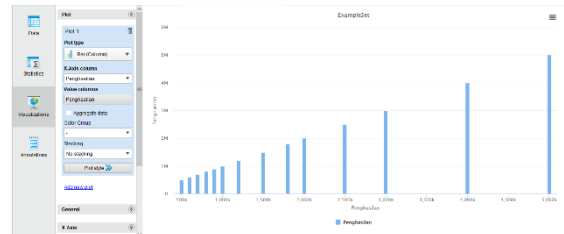
PerformanceVector

```
PerformanceVector:
Avg. within centroid distance: 126860146.045
Avg. within centroid distance_cluster_0: 3.889
Avg. within centroid distance_cluster_1: 10.883
Avg. within centroid distance_cluster_2: 1.577
Avg. within centroid distance_cluster_3: 22933815.103
Avg. within centroid distance_cluster_4: 89362138.649
Avg. within centroid distance_cluster_5: 344387760.440
Avg. within centroid distance_cluster_6: 62500000002.125
Avg. within centroid distance_cluster_7: 2.808
Avg. within centroid distance_cluster_8: 1.834
Avg. within centroid distance_cluster_9: 0.980
Davies Bouldin: 0.049
```

Gambar 21 Performance Vector

Setelah menentukan pusat cluster, langkah selanjutnya adalah menentukan jarak antar cluster. Operator yang digunakan untuk menentukan jarak antar cluster adalah Cluster Distance Performance, dan hasilnya dapat dilihat pada Performance vector pada gambar 21.

Selanjutnya, data dari pengklasteran dianalisis dalam bentuk diagram batang (Column Chart) untuk memperjelas interpretasi data tersebut.



Gambar 22 (Bar Column)

Dengan menerapkan metode Knowledge Discovery in Database (KDD) dan teknik data mining untuk mengklasifikasikan penerima bantuan sosial menggunakan algoritma K-Means, ditemukan bahwa jumlah K optimal berdasarkan Davies Bouldin Index adalah $K=10$.

5. KESIMPULAN DAN SARAN

Dari analisis menggunakan algoritma K-Means Clustering pada data penerima bantuan sosial Desa Greged, beberapa kesimpulan dapat ditarik sebagai berikut: Studi ini memanfaatkan pendekatan data mining melalui metode Knowledge Discovery in Database dan algoritma K-Means Clustering untuk mengkategorikan warga berdasarkan status ekonominya, mulai dari kategori ekonomi menengah kebawah, menengah, hingga menengah ke atas. Dari sepuluh eksperimen yang dilakukan dengan K-Means Clustering, ditemukan bahwa $K=10$ adalah pilihan terbaik. Indeks Davies Bouldin untuk $K=10$ mencapai 0,049, menandakan bahwa cluster-cluster penerima bantuan sosial memiliki kohesi yang kuat dan perbedaan antara cluster minim. Ini menunjukkan bahwa warga penerima bantuan di Desa Greged dapat dikelompokkan ke dalam 10 kategori berbeda berdasarkan berbagai aspek seperti demografi, ekonomi, dan faktor sosial.

DAFTAR PUSTAKA

- [1] B. Baskoro, A. Gunaryati, and A. Rubhasy, "Klasifikasi Penduduk Kurang Mampu Dengan Metode K-Means untuk Optimalisasi Program Bantuan Sosial," *Infomatek*, vol. 25, no. 1, pp. 41–48, 2023, doi: 10.23969/infomatek.v25i1.7271.
- [2] A. Surahman and U. Hayati, "MPLEMENTASI ALGORITMA NAİVE BAYES UNTUK PREDIKSI PENERIMA BANTUAN SOSIAL (STUDI KASUS: DESA JATIPANCUR)," pp. 31–41, 2023.

- [3] F. Azzahra, A. Zakir, and D. Irwan, "PEMANFAATAN ALGORITMA K MEANS UNTUK MENENTUKAN PENERIMAAN BANTUAN LANGSUNG TUNAI PADA KECAMATAN MEDAN AREA," *Djtechno J. Inf. Technol. Res.*, vol. 3, no. 2, pp. 260–267, 2022.
- [4] A. E. Putra Haryanto and M. Ulfa Yanuar, "Metode K-Means Clustering untuk Pengelompokan Kabupaten/Kota dalam Upaya Pengendalian Tingkat Inflasi di Pulau Jawa dan Sumatera," *Gov. Stat.*, vol. 1, no. 1, pp. 29–42, 2022, doi: 10.21787/govstat.1.2022.29-42.
- [5] M. Orisa, "Optimasi Cluster pada Algoritma K-Means," *Pros. SENIATI*, vol. 6, no. 2, pp. 430–437, 2022, doi: 10.36040/seniati.v6i2.5034.
- [6] Y. Kusnadi and M. S. Putri, "Clustering Menggunakan Metode K-Means Untuk Menentukan Prioritas Penerima Bantuan Bedah Rumah (Studi Kasus : Desa Ciomas Bogor)," *J. Teknol. Inform. dan Komput.*, vol. 7, no. 1, pp. 17–24, 2021, doi: 10.37012/jtik.v7i1.498.
- [7] M. Khandava Mulyadien and U. Enri, "Algoritma K-Means Untuk Pengelompokan Bantuan Langsung Tunai (BLT)," *J. Ilm. Wahana Pendidik.*, vol. 8, no. 12, pp. 198–210, 2022.
- [8] Y. Filki, "Algoritma K-Means Clustering dalam Memprediksi Penerima Bantuan Langsung Tunai (BLT) Dana Desa," *J. Inform. Ekon. Bisnis*, vol. 4, pp. 166–171, 2022, doi: 10.37034/infec.v4i4.166.
- [9] T. A. Munandar, "Penerapan Algoritma Clustering Untuk Pengelompokan Tingkat Kemiskinan Provinsi Banten," *JSiI (Jurnal Sist. Informasi)*, vol. 9, no. 2, pp. 109–114, 2022, doi: 10.30656/jsii.v9i2.5099.
- [10] D. Agustina, Y. I. Mukti, and S. Muntari, "Integrasi Particle Swam Optimization Menggunakan K-Means untuk Klasterisasi Pengangguran di Kota Pagar Alam," vol. 3, no. 1, pp. 34–41, 2023.
- [11] S. Suhartini and R. Yuliani, "Penerapan Data Mining untuk Mengcluster Data Penduduk Miskin Menggunakan Algoritma K-Means di Dusun Bagik Endep Sukamulia Timur," *Infotek J. Inform. dan Teknol.*, vol. 4, no. 1, pp. 39–50, 2021, doi: 10.29408/jit.v4i1.2986.
- [12] Yuni Radana Sembiring, Saifullah, and Riki Winanjaya, "Implementasi Data Mining Dalam Mengelompokkan Jumlah Penduduk Miskin Berdasarkan Provinsi Menggunakan Algoritma," *KESATRIA J. Penerapan Sist. Inf. (Komputer Manajemen) Vol. 2, No. 2*, vol. 2, no. 2, pp. 125–132, 2021.