

Penerapan Algoritma K-Means Dalam Menentukan Penerima Bantuan Keluarga Pra Sejahtera

Leoni Sancai¹, Indra M Sarkis Simamora², Doli Hasibuan³
^{1,2,3}Fakultas Ilmu Komputer, Universitas Methodist Indonesia

Info Artikel

Histori Artikel:

Received, 10 Agustus 2026
Revised, 3 September 2026
Accepted, 15 September 2026

Keywords:

Pengelompokkan,
K-Means,
Bantuan Sosial,
Keluarga Prasejahtera

ABSTRAK

Program bantuan sosial dirancang untuk membantu masyarakat prasejahtera dalam memenuhi kebutuhan dasar. Namun, distribusi bantuan sering menghadapi kendala seperti kesalahan dalam penentuan sasaran dan proses yang memakan waktu akibat belum adanya metode klasifikasi yang efektif. Penelitian ini bertujuan untuk menerapkan algoritma K-Means dalam mengelompokkan data penduduk calon penerima bantuan di Desa Lama Pancur Batu berdasarkan kriteria seperti usia, pendapatan, pekerjaan dan tanggungan. Penentuan jumlah cluster optimal dilakukan menggunakan metode elbow yang menunjukkan 4 cluster sebagai konfigurasi terbaik. Metode K-Means dipilih karena kemampuannya mengelompokkan data berukuran besar secara cepat dan efisien. Hasil penelitian menunjukkan bahwa dari 983 data penduduk yang dianalisis, terbentuk empat kluster: Sangat Miskin (C1:222 jiwa), Miskin (C2:222 jiwa), Rentan Miskin (C3:206 jiwa), dan Tidak Miskin (C4:333 jiwa). Kluster Sangat Miskin (C1) ditetapkan sebagai prioritas utama dalam penyaluran bantuan. Untuk mendukung hasil klusterisasi, dilakukan visualisasi menggunakan Scatter Plot yang menggambarkan penyebaran data berdasarkan kedekatan karakteristik masing-masing kluster, serta Bar Chart yang menunjukkan jumlah anggota pada setiap kluster. Penerapan algoritma K-Means membantu mempercepat proses identifikasi dan memastikan bantuan tepat sasaran, sehingga meningkatkan efektivitas program bantuan sosial.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Penulis Koresponden:

Leoni Sancai
Fakultas Ilmu Komputer,
Universitas Methodist Indonesia, Medan,
Jl. Hang Tuah No.8, Medan - Sumatera Utara.
Email: leoni@gmail.com

1. PENDAHULUAN

Program bantuan sosial bertujuan untuk mendukung masyarakat yang berada di bawah garis kemiskinan agar dapat memenuhi kebutuhan dasar mereka. Program bantuan pra sejahtera diimplementasikan untuk mengurangi beban ekonomi masyarakat miskin. Namun, tantangan dalam distribusi bantuan sering kali muncul karena belum adanya metode yang digunakan dalam mengelompokkan calon penerima manfaat yang menyebabkan waktu yang banyak, rawan kesalahan dan tidak tepat sasaran. Oleh karena itu, diperlukan sebuah metode yang mampu menganalisis data secara efisien dan mengelompokkan penduduk berdasarkan kriteria tertentu agar penyaluran bantuan menjadi lebih efektif.

Algoritma *K-means* adalah metode *clustering* yang banyak digunakan dalam analisis data untuk mengelompokkan objek-objek yang memiliki karakteristik serupa ke dalam kelompok-

kelompok yang berbeda [1]. Algoritma ini mampu menangani data dalam jumlah besar dengan cepat, sehingga cocok diterapkan dalam analisis data penduduk penerima bantuan. Dengan menggunakan *K-means*, data penduduk dapat dikelompokkan berdasarkan faktor-faktor seperti usia, tingkat pendapatan, dan pekerjaan, yang membantu dalam menentukan tingkat prioritas penerima bantuan.

Pelaksanaan penyaluran bantuan sosial masih menghadapi berbagai tantangan, terutama terkait keakuratan data dan efisiensi distribusi. Penelitian oleh [2] di Indonesia menunjukkan bahwa distribusi bantuan sosial sering kali tidak tepat sasaran akibat dari proses validasi data yang kurang optimal. Untuk mengatasi hal tersebut, pendekatan berbasis teknologi seperti algoritma *K-Means Clustering* terbukti efektif dalam mengelompokkan calon penerima berdasarkan kondisi sosial ekonomi yang serupa. Dengan pendekatan berbasis data ini, pihak terkait dapat mengidentifikasi kelompok penduduk yang benar-benar membutuhkan bantuan, sehingga program yang diimplementasikan dapat berjalan dengan lebih tepat sasaran dan membawa dampak positif yang lebih signifikan bagi masyarakat.

2. METODE PENELITIAN

2.1 Data Mining

Data Mining adalah proses mengekstraksi pola atau informasi yang berguna dari kumpulan data yang besar [3]–[6]. Dalam konteks metode *K-Means*, data mining digunakan untuk mengelompokkan data berdasarkan karakteristik tertentu, sehingga pola dalam data dapat dianalisis lebih lanjut. Metode *K-Means Clustering* adalah salah satu teknik dalam data mining yang digunakan untuk mengelompokkan data ke dalam beberapa kluster berdasarkan kesamaan karakteristik. Algoritma ini bekerja dengan cara membagi data ke dalam K kelompok sehingga setiap data dalam satu kelompok memiliki kesamaan lebih besar dibandingkan dengan kelompok lain.

2.2 Bantuan Pra Sejahtera

Bantuan Pra Sejahtera adalah program yang bertujuan untuk membantu masyarakat yang berada dalam kondisi ekonomi sulit, dengan memberikan bantuan berupa uang tunai atau barang yang dapat meningkatkan kesejahteraan hidup mereka. Program ini penting karena dapat memberikan dukungan kepada individu atau keluarga yang berisiko terjatuh dalam kemiskinan atau kesulitan ekonomi. Secara umum, bantuan ini menyasar kelompok masyarakat yang memiliki pendapatan rendah, pekerjaan yang tidak stabil, atau keluarga yang banyak menanggung beban [7].

2.3 Algoritma K-Means

Algoritma *K-Means* adalah algoritma *clustering* yang digunakan untuk mengelompokkan data ke dalam k kelompok berdasarkan kemiripan fitur [8][9][10]. Algoritma ini bekerja dengan cara membagi dataset ke dalam k cluster dengan iterasi yang berulang sehingga mendapatkan hasil yang optimal [11]. Langkah-langkah algoritma kmeans [12][13]:

Langkah 1 : Tentukan jumlah *Cluster*.

Langkah 2 : Tentukan nilai yang menjadi lokasi pusat *Cluster*

Langkah 3 : Hitung pusat *Cluster* terdekat untuk setiap *record* dengan rumus

$$d(x_i, y_j) = \sqrt{\sum_j^n (x_i - y_j)^2} \dots\dots\dots (1)$$

Keterangan [14]:

d = jarak setiap objek

y_j = *centroid* pada *cluster* ke-j

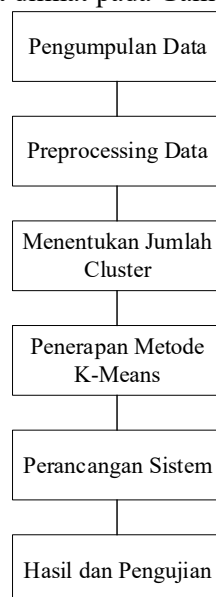
x_i = data kriteria

Langkah 4 : Untuk setiap *Cluster*, hitung *centroid Cluster* dan memperbarui lokasi setiap pusat *Cluster*.

Langkah 5 : Ulangi langkah 2 sampai 4 sampai konvergensi atau penghentian.

2.4 Framework Penelitian

Framework penelitian dalam penelitian ini bertujuan untuk mendeskripsikan seluruh tahapan kegiatan yang dilakukan selama penelitian agar sesuai dengan tujuan yang telah ditetapkan. Kerangka kerja pada penelitian ini dapat dilihat pada Gambar 1.



Gambar 1. Kerangka Kerja Penelitian

2.5 Pengumpulan Data

Jenis data yang digunakan dalam penelitian ini adalah data yang didapatkan langsung dari tempat penelitian merupakan data penerima bantuan pra sejahtera di Desa Lama Pancur Batu. Data yang digunakan dalam penelitian ini terdapat dengan sampel seperti Tabel 1

Tabel 1 Data Calon Penerima Bantuan

No	Nama	Umur	Pekerjaan	Gaji	Tanggungan
1	M.Nasrun	60	Buruh Harian	2400000	5
2	Syahrudin	60	Buruh Harian	2400000	4
3	Sri Martha	49	Buruh Harian	2400000	2
4	Salomo Sinuraya	73	Buruh Harian	2400000	5
5	Khairul Andri	50	Buruh Harian	2400000	2
6	Mail Ambutan	54	Tidak Bekerja	0	5
7	Hendri Bangun	46	Buruh Harian	2400000	6
8	Nurlina Br Pinem	63	Pembantu Rumah Tangga	1000000	2
9	Eli Anhar	47	Buruh Harian	2400000	4
10	Emmi Br Tarigan	54	Kader Posyandu	900000	2
...	...	...	...	...	...
983	Erwi Nuriana	57	Mengurus Rumah Tangga	0	2

Sumber : Kantor Desa Lama Pancur Batu

Data keseluruhan calon penerima bantuan pra sejahtera di Desa Lama Pancur Batu sebanyak 983 data. Adapun untuk seluruh data dapat diakses pada link google drive berikut: <https://drive.google.com/drive/folders/1HOJchfgqrda9SeU8em5KrBjNVK2gU7B>

2.6 Preprocessing Data

Data penduduk Desa Lama Pancur Batu yang terdiri dari atribut umur, pekerjaan, gaji, dan tanggungan dilakukan *preprocessing* untuk memastikan kompatibilitas dengan algoritma K-Means. Tahapanya dilakukan dengan transformasi data kategorikal seperti pekerjaan ke dalam bentuk numerik menggunakan teknik label encoding. Atribut numerik seperti gaji dan umur dinormalisasi menggunakan *Min-Max Scaling* agar memiliki skala yang seragam, sementara atribut tanggungan disesuaikan bobotnya agar tidak mendominasi perhitungan jarak. Hasil *preprocessing* ini menghasilkan dataset yang siap untuk proses *clustering* dengan *K-Means*, sehingga pengelompokan keluarga pra-sejahtera dapat dilakukan secara lebih akurat. Sebelum dilakukan proses hitungan *K-Means*, data penduduk terlebih dahulu di transformasi dengan ketentuan sebagai berikut:

1. Ketentuan Transformasi Umur

Tabel 2 Transformasi Umur

Nilai	Umur	Keterangan
0	20-30	Muda
1	31-59	Dewasa
2	> 59	Lansia

2. Ketentuan Transformasi Pekerjaan

Label encoding untuk pekerjaan dengan logika dimana semakin tinggi stabilitas/pendapatan pekerjaan, semakin besar nilainya

Tabel 3 Transformasi Pekerjaan

Nilai	Pekerjaan	Keterangan
0	Tidak Bekerja/Mengurus Rumah Tangga	Tidak memiliki penghasilan
1	Buruh Harian	Pekerja lepas tanpa jaminan tetap
2	Pedagang Kaki Lima	Usaha dengan ketidakpastian ekonomi
3	Pembantu Rumah Tangga	Pekerjaan informal, upah relatif rendah
4	Buruh Pabrik	Pekerja kasar dengan upah minimum
5	Peternak	Penghasilan tidak stabil
6	Petani	Bergantung pada musim dan lahan
7	Supir Angkot/Supir Sekolah	Pekerjaan transportasi informal
8	Ojek Online	Pekerja gig economy
9	Kader Posyandu	Relawan/pekerja sosial
10	Guru Honorer	Pendidikan tetapi tanpa status tetap
11	Staff Sekolah/Admin	Pekerjaan administratif tingkat dasar
12	Security	Pekerjaan jasa dengan risiko fisik
13	Wiraswasta	Usaha mandiri
14	Pegawai Swasta	Stabilitas lebih baik
15	Pegawai Negeri Sipil	Stabilitas tinggi dan tunjangan

3. Ketentuan Transformasi Pendapatan

Tabel 4 Transformasi Pendapatan

Nilai	Rentang Pendapatan	Keterangan
0	≤ Rp1.500.000	Sangat Rendah
1	Rp1.500.000 – Rp3.000.000	Menengah
2	> Rp3.000.000	Tinggi

4. Ketentuan Transformasi Tanggungan

Tabel 5 Transformasi Tanggungan

Nilai	Jumlah Tanggungan	Keterangan
0	0-2	Sedikit
1	3-4	Sedang
2	≥ 5	Banyak

2.7 Penerapan Metode K-Means

Pengelompokan menggunakan algoritma *K-Means* dengan penentuan jumlah *cluster* dan *centroid* yang ditentukan secara acak. Selanjutnya menghitung jarak data terhadap pusat *cluster* menggunakan *Euclidean Distance*. Setelah menghitung jarak data terhadap *cluster* maka dilanjutkan dengan mengelompokkan data ke dalam *cluster* dengan jarak paling dekat. Kemudian menghitung pusat *cluster* dengan mencari nilai rata rata pada kelompok data. Proses penentuan pusat *cluster* dan penempatan data dalam *cluster* diulangi sampai nilai *centroid* tidak berubah lagi. Metode *K-Means* diterapkan untuk mengelompokkan data menjadi sejumlah klaster yang sudah ditentukan sebelumnya. *K-Means* adalah algoritma pembelajaran mesin yang membagi data menjadi klaster-klaster berdasarkan kedekatannya dengan pusat klaster. Berdasarkan pada Tabel 1, jumlah data yang digunakan sebanyak 983 penduduk. Berdasarkan dari informasi ketentuan transformasi data yang mencakup kriteria seperti umur, pekerjaan, pendapatan, jumlah tanggungan, berikut adalah hasil transformasinya seperti pada Tabel 6.

Tabel 6 Hasil Transformasi Data Penduduk

No	Nama	Umur	Pekerjaan	Gaji	Tanggungan
1	M.Nasrun	2	1	1	2
2	Syahrudin	2	1	1	1
3	Sri Martha	2	1	1	0
4	Salomo Sinuraya	1	1	1	2
5	Khairul Andri	2	1	1	0
6	Mail Ambutan	1	0	0	2
7	Hendri Bangun	1	1	1	2
8	Nurlina Br Pinem	1	3	0	0
9	Eli Anhar	2	1	1	1
10	Emmi Br Tarigan	1	9	0	0
...	...	...	...	...	...
983	Erwi Nuriana	1	0	0	0

2.7.1 Penentuan Jumlah Cluster

Adapun jenis dan jumlah *cluster* yang digunakan pada penelitian ini didapat dari proses pencarian dengan metode *elbow* terdapat pada Tabel 7.

Tabel 7 Cluster Penerimaan Bantuan

No	Cluster	Jenis Pengelompokan
1	C1	Sangat Miskin
2	C2	Miskin
3	C3	Rentan Miskin
4	C4	Tidak Miskin

2.7.2 Menentukan Titik Pusat Cluster (Centroid)

Inisialisasi pusat *cluster* sebanyak dari data pendudukan pada Tabel 4.6. Untuk pemilihan setiap pusat cluster awal dilakukan secara random. Pada penelitian ini diambil data cendroid awal dari baris pertama hingga baris keempat pada data penduduk. Adapun pusat cluster awal *K-Means* seperti pada Tabel 8.

Tabel 8 Pusat *Clutser* Awal

Centroid	Usia	Pekerjaan	Gaji	Tanggungan
C1	2	1	1	2
C2	2	1	1	1
C3	2	1	1	0
C4	1	1	1	2

2.7.3 Menghitung Jarak Data Terhadap Centroid

Menghitung nilai jarak terdekat dengan persamaan *Euclidian Distance* yakni persamaan (2.1). Maka perhitungan untuk menghitung jarak setiap baris data penduduk dengan *Centroid* awal adalah sebagai berikut:

$$\begin{aligned}
 Data1, C1 &= \sqrt{(2-2)^2 + (1-1)^2 + (1-1)^2 + (2-2)^2} = 0 \\
 Data1, C2 &= \sqrt{(2-2)^2 + (1-1)^2 + (1-1)^2 + (2-1)^2} = 1 \\
 Data1, C3 &= \sqrt{(2-2)^2 + (1-1)^2 + (1-1)^2 + (2-0)^2} = 2 \\
 Data1, C4 &= \sqrt{(2-1)^2 + (1-1)^2 + (1-1)^2 + (2-2)^2} = 1 \\
 &\vdots \\
 Data983, C1 &= \sqrt{(1-2)^2 + (0-1)^2 + (0-1)^2 + (0-2)^2} = 2.645751 \\
 Data983, C2 &= \sqrt{(1-2)^2 + (0-1)^2 + (0-1)^2 + (0-1)^2} = 2 \\
 Data983, C3 &= \sqrt{(1-2)^2 + (0-1)^2 + (0-1)^2 + (0-0)^2} = 1.732051 \\
 Data983, C4 &= \sqrt{(1-1)^2 + (0-1)^2 + (0-1)^2 + (0-2)^2} = 2.44949
 \end{aligned}$$

Hasil dari keseluruhan dapat dilihat pada Tabel 9:

Tabel 9 Hasil Perhitungan Algoritma *K-Means* Iterasi ke-1

No	Jarak C1	Jarak C2	Jarak C3	Jarak C4	Terdekat
1	0	1	2	1	C1
2	1	0	1	1.414214	C2
3	2	1	0	2.236068	C3
4	1	1.414214	2.236068	0	C4
5	2	1	0	2.236068	C3
6	1.732051	2	2.645751	1.414214	C4
7	1	1.414214	2.236068	0	C4
8	3.162278	2.645751	2.44949	3	C3

No	Jarak C1	Jarak C2	Jarak C3	Jarak C4	Terdekat
9	1	0	1	1.414214	C2
10	8.3666	8.185353	8.124038	8.306624	C3
...	...	...	...	...	...
983	2.645751	2	1.732051	2.44949	C3

Berdasarkan dari hasil *clustering* data penduduk iterasi k-1 pada Tabel 9, didapati bahwa jumlah C1 sebanyak 79 *cluster*, C2 sebanyak 331 *cluster*, C3 sebanyak 323 *cluster* dan C4 sebanyak 250 *cluster*. Setelah didapatkan hasil jarak dari setiap objek pada iterasi ke-1 maka lanjut ke iterasi ke-2. Pembentukan *centroid* baru dilakukan dengan mencari jumlah rata-rata dari setiap nilai atribut pendudukan berdasarkan *clusternya*. Adapun hasil dari pembentukan *centroid* untuk iterasi k-2 apat dilihat pada Tabel 10.

Tabel 10 Pusat *Clutser* Baru Iterasi ke-2

Centroid	Usia	Pekerjaan	Gaji	Tanggungan
C1	2	6.341772	0.759494	2
C2	1.21148	7.003021	0.963746	1
C3	1.179567	7.092879	0.944272	0
C4	0.92	6.016	0.824	1.96

Menghitung nilai jarak terdekat dengan persamaan *Euclidian Distance* yakni persamaan (2.1). Maka perhitungan untuk menghitung jarak setiap baris data penduduk dengan *Centroid* iterasi k2-2 adalah sebagai berikut:

$$Data1, C1 = \sqrt{(2-2)^2 + (1-6.341772)^2 + (1-0.759494)^2 + (2-2)^2} = 5.347184$$

$$Data1, C2 = \sqrt{(2-1.21148)^2 + (1-7.003021)^2 + (1-0.963746)^2 + (2-1)^2} = 6.136721$$

$$Data1, C3 = \sqrt{(2-1.179567)^2 + (1-7.092879)^2 + (1-0.944272)^2 + (2-0)^2} = 6.465245$$

$$Data1, C4 = \sqrt{(2-0.92)^2 + (1-6.016)^2 + (1-0.824)^2 + (2-1.96)^2} = 5.134124$$

⋮

$$Data983, C1 = \sqrt{(1-2)^2 + (0-6.341772)^2 + (0-0.759494)^2 + (0-2)^2} = 6.767193$$

$$Data983, C2 = \sqrt{(1-1.21148)^2 + (0-7.003021)^2 + (0-0.963746)^2 + (0-1)^2} = 7.142537$$

$$Data983, C3 = \sqrt{(1-1.179567)^2 + (0-7.092879)^2 + (0-0.944272)^2 + (0-0)^2} = 7.157711$$

$$Data983, C4 = \sqrt{(1-0.92)^2 + (0-6.016)^2 + (0-0.824)^2 + (0-1.96)^2} = 6.381162$$

Hasil dari keseluruhan dapat dilihat pada Tabel 11:

Tabel 11 Hasil Perhitungan Algoritma *K-Means* Iterasi ke-2

No	Jarak C1	Jarak C2	Jarak C3	Jarak C4	Terdekat
1	5.347184	6.136721	6.465245	5.134124	C4
2	5.439887	6.054696	6.228916	5.222952	C4

No	Jarak C1	Jarak C2	Jarak C3	Jarak C4	Terdekat
3	5.708973	6.136721	6.148121	5.495383	C4
4	5.439887	6.089524	6.415491	5.019884	C4
5	5.708973	6.136721	6.148121	5.495383	C4
6	6.464898	7.142537	7.431879	6.072827	C4
7	5.439887	6.089524	6.415491	5.019884	C4
8	4.091977	4.242371	4.204231	3.690966	C4
9	5.439887	6.054696	6.228916	5.222952	C4
10	3.5557	2.441609	2.135651	3.664865	C3
...	...	...	...	...	...
983	6.767193	7.142537	7.157711	6.381162	C4

Berdasarkan dari hasil *clustering* data penduduk iterasi k-2 pada Tabel 11, didapati bahwa jumlah C1 sebanyak 14 *cluster*, C2 sebanyak 259 *cluster*, C3 sebanyak 254 *cluster* dan C4 sebanyak 456 *cluster*. Setelah didapatkan hasil jarak dari setiap objek pada iterasi ke-2 belum ditemukan persamaan jumlah *cluster* maka lanjut ke iterasi ke-3 Pembentukan *centroid* baru dilakukan dengan mencari jumlah rata-rata dari setiap nilai atribut pendudukan berdasarkan *clusternya*. Adapun hasil dari pembentukan *centroid* untuk iterasi k-3 apat dilihat pada Tabel 12.

Tabel 12 Pusat *Clutser* Baru Iterasi ke-3

Centroid	Usia	Pekerjaan	Gaji	Tanggungan
C1	2	5.857143	0.785714	1.7142
C2	1.220077	10.39382	1.200772	1.4594
C3	1.161417	11.11417	1.267717	0.2086
C4	1.164474	2.230263	0.539474	1.1491

Proses yang sama dilakukan hingga terjadi kovergensi.

2.7.4 Hasil Clustering

Pada iterasi ke-6 ditemukan kesamaan jumlah *cluster* pada iterasi sebelumnya, maka konvergensi tercapai. Adapun data akhir dari hasil *clustering* seperti pada Tabel 13.

Tabel 13 Hasil *Clustering* Metode *K-Means*.

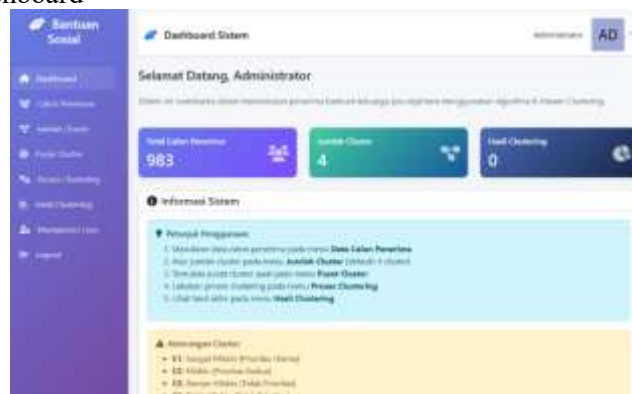
No	Nama	Usia	Pekerjaan	Gaji	Tanggungan	Hasil Cluster
1	M.Nasrun	2	1	1	2	C4 Tidak Miskin
2	Syahrudin	2	1	1	1	C4 Tidak Miskin
3	Sri Martha	2	1	1	0	C4 Tidak Miskin
4	Salomo Sinuraya	1	1	1	2	C4 Tidak Miskin
5	Khairul Andri	2	1	1	0	C4 Tidak Miskin
6	Mail Ambutan	1	0	0	2	C4 Tidak Miskin
7	Hendri Bangun	1	1	1	2	C4 Tidak Miskin
8	Nurlina Br Pinem	1	3	0	0	C4 Tidak Miskin
9	Eli Anhar	2	1	1	1	C4 Tidak Miskin
10	Emmi Br Tarigan	1	9	0	0	C2 Miskin
...	...	...	...	...	...	...
983	Erwi Nuriana	1	0	0	0	C4 Tidak Miskin

Berdasarkan pada Tabel 13, hasil *clustering* didapati bahwa pendudukan dengan *cluster* C1 sebanyak 222, C2 sebanyak 222, C3 sebanyak 206 dan C4 sebanyak 333.

3. HASIL DAN PEMBAHASAN

3.1 Hasil Pengujian Sistem

1. Tampilan Menu Dashboard



Gambar 2 Tampilan Menu Dashboard

2. Tampilan Menu Calon Penerima



Gambar 3 Tampilan Menu Calon Penerima

3. Tampilan Menu Pusat Cluster



Gambar 4 Tampilan Menu Hasil Simpan Pusat Cluster

4. Tampilan Menu Proses Clustering



Gambar 5 Tampilan Menu Proses Clustering

5. Tampilan Menu Hasil Clustering



Gambar 6 Hasil Clustering

3.2 Pembahasan

Adapun hasil akhir dari proses *clustering* didapati *cluster* setiap penduduk Penduduk Desa Lama Pancur Batu yang berhak menerima bantuan pra sejahtera seperti pada Gambar 7.

Detail Hasil Clustering

No	Nama Penerima	Umur	Pekerjaan	Gaji	Tanggungan	Cluster	Jarak Minimum
1	M. Nasrun	60	Buruh Harian	Rp 2.400.000	5	C1: Tidak Miskin	1.400824
2	Syahrudin	60	Buruh Harian	Rp 2.400.000	4	C1: Tidak Miskin	1.031730
3	Sri Mantha	60	Buruh Harian	Rp 2.400.000	2	C1: Tidak Miskin	1.471974
4	Sakomo Sinuraya	48	Buruh Harian	Rp 2.400.000	5	C1: Tidak Miskin	1.186047
5	Khairul Andri	73	Buruh Harian	Rp 2.400.000	2	C1: Tidak Miskin	1.471974
6	Mali Ambutan	50	Tidak Bekerja	Rp 0	5	C1: Tidak Miskin	1.560034
7	Hendri Bangun	54	Buruh Harian	Rp 2.400.000	6	C1: Tidak Miskin	1.186047
8	Nurlina Br Pinem	46	Pembantu Rumah Tangga	Rp 1.000.000	2	C1: Tidak Miskin	2.151490
9	Eli Anhar	63	Buruh Harian	Rp 2.400.000	4	C1: Tidak Miskin	1.031730
10	Emmi Br Tarigan	47	Kader Posyandu	Rp 900.000	2	C2: Miskin	1.587838
11	Yeni Indah Sari	54	Buruh Harian	Rp 2.400.000	2	C1: Tidak Miskin	1.269294
12	Mariono	52	Buruh Harian	Rp 2.400.000	5	C1: Tidak Miskin	1.186047
13	Mhd Cholli	60	Buruh Harian	Rp 2.400.000	2	C1: Tidak Miskin	1.471974
14	Mai Mawarhasibuan	48	Tidak Bekerja	Rp 0	6	C1: Tidak Miskin	1.560034
15	Sri Susilawati	63	Buruh Harian	Rp 2.400.000	2	C1: Tidak Miskin	1.471974
981	Ekel Pranata Tarigan	62	Mengurus Rumah Tangga	Rp 0	5	C1: Tidak Miskin	1.728961
982	Erina Br Sembiring	34	Pegawai Negeri Sipil	Rp 5.800.000	2	C3: Rentan Miskin	1.793603
983	Enai Nuriana	56	Mengurus Rumah Tangga	Rp 0	2	C1: Tidak Miskin	1.624225

Export Data

Export ke Excel

Gambar 7 Hasil Clustering Akhir

Berdasarkan dari hasil *clustering* pada Gambar 7, didapati bahwa jumlah pendudukan yang Sangat Miskin (C1) berjumlah sebanyak 222 jiwa, jumlah pendudukan yang Miskin (C2) berjumlah sebanyak 222 jiwa, jumlah pendudukan yang Rentan Miskin (C3) berjumlah sebanyak 206 jiwa dan jumlah pendudukan yang Tidak Miskin (C4) berjumlah sebanyak 333 jiwa.

4. KESIMPULAN

Proses menentukan kelompok keluarga calon penerima bantuan pra sejahtera yang ada di Desa Lama Pancur Batu berhasil dilakukan menggunakan metode K-Means dengan tahapan penentuan jumlah cluster optimal menggunakan metode *elbow* yang menghasilkan 4 cluster optimal, penentuan pusat *cluster*, proses iterasi *clustering* dan proses seleksi hasil *clustering*. Berdasarkan hasil *clustering* terhadap 983 data penduduk dengan variabel usia, pekerjaan, pendapatan dan tanggungan maka teridentifikasi 222 jiwa masuk kategori Sangat Miskin (C1), 222 jiwa Miskin (C2), 206 jiwa Rentan Miskin (C3), dan 333 jiwa Tidak Miskin (C4), di mana kelompok C1 sebanyak 222 jiwa menjadi prioritas utama penyaluran bantuan sosial. Data ini memungkinkan Kantor Desa Lama Pancur Batu untuk menargetkan penerima bantuan secara lebih akurat dan efisien, sehingga program bantuan pra-sejahtera dapat tepat sasaran dan berdampak optimal bagi warga yang paling membutuhkan.

REFERENSI

- [1] B. Harahap and A. Rambe, "Implementasi K-Means Clustering Terhadap Mahasiswa yang Menerima Beasiswa Yayasan Pendidikan Battuta di Universitas Battuta Tahun 2020/2021 Studi Kasus Prodi Informatika," *Informatika*, vol. 9, no. 3, pp. 90–97, 2021, doi: 10.36987/informatika.v9i3.2185.
- [2] N. Salsabila, N. Muna, V. H. Pradana, and W. F. Nurcahya, "Analisis Efektivitas Bantuan Sosial (Bansos) dalam mengatasi Kemiskinan di Indonesia," *J. Macroecon. Soc.*, vol. 1, no. 4, pp. 1–13, 2024, doi: 10.47134/jmsd.v1i4.317.
- [3] A. C. Sitorus, I. Kelana Jaya, and D. Hasibuhan, "Implementasi Data Mining Data Pencari Kerja Menggunakan Metode K-Means," *J. Ilm. Sist. Inf.*, vol. 3, no. 1, pp. 143–151, 2023.
- [4] T. R. Barudara, H. G. Manullang, and A. P. Silalahi, "ANALISA ALGORITMA K-MEDOIDS DALAM," vol. 4, no. 2, pp. 1–10, 2024.
- [5] S. Hartawan, T. Gulo, H. Rumapea, and H. G. Simanullang, "Penerapan Data Mining Dalam Pemilihan Laptop Mahasiswa Menggunakan Metode Random Forest (Studi Kasus : Universitas Methodist Indonesia)," vol. 5, no. 1, pp. 55–64, 2025.
- [6] H. G. Simanullang, A. P. Silalahi, and G. P. Pangaribuan, "PENGELOMPOKAN KOMENTAR PENGGUNA JKN MOBILE MENGGUNAKAN METODE K-MEDOIDS," 2025.
- [7] Y. Dodi, R. A. Ekaputri, and I. Izharudin, "Analisa Program Bantuan Beras Sejahtera (Rastra) dalam Meningkatkan Kesejahteraan Masyarakat di Desa Serdang Indah Kecamatan Luas Kabupaten Kaur Provinsi Bengkulu," *Ecoplan*, vol. 7, no. 1, pp. 15–29, 2024, doi: 10.20527/ecoplan.v7i1.688.
- [8] S. N. Mayasari and J. Nugraha, "Implementasi K-Means Cluster Analysis untuk Mengelompokkan Kabupaten/Kota Berdasarkan Data Kemiskinan di Provinsi Jawa Tengah Tahun 2022," *KONSTELASI Konvergensi Teknol. dan Sist. Inf.*, vol. 3, no. 2, pp. 317–329, 2023, doi: 10.24002/konstelasi.v3i2.7200.
- [9] T. Taryadi, "Klasterisasi Data Keluarga Pra Sejahtera Di Kota Pekalongan Dengan Metode K-Means Clustering," *J. Litbang Kota Pekalongan*, vol. 20, no. 1, pp. 70–76, 2022, doi: 10.54911/litbang.v20i1.180.
- [10] 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.
- [11] E. Moruk *et al.*, "ALGORITMA K-MEANS UNTUK CLUSTERING PENERIMA PROGRAM KELUARGA HARAPAN DI NANAENOE NUSA TENGGARA TIMUR," *JATI (Jurnal Mhs. Tek. Inform. Algoritma K-MEANS UNTUK Clust. PENERIMA Progr. Kel. HARAPAN DI NANAENOE NUSA TENGGARA TIMUR)*, vol. 8, no. 2, pp. 1681–1687, 2024.
- [12] B. A. Dwiarni and B. Setiyono, "Akuisisi dan Clustering Data Sosial Media Menggunakan Algoritma K-Means sebagai Dasar untuk Mengetahui Profil Pengguna," *J. Sains dan Seni ITS*, vol. 8, no. 2, 2020, doi: 10.12962/j23373520.v8i2.49815.
- [13] A. Nugrahaini and Rahmadi Yotenka, "Penerapan Algoritma K-Means Clustering untuk

-
- Mengelompokkan Kecamatan di Kabupaten Grobogan Menurut Tingkat Kesejahteraan Keluarga Tahun 2020,” *Emerg. Stat. Data Sci. J.*, vol. 1, no. 3, pp. 290–299, 2023, doi: 10.20885/esds.vol1.iss.3.art36.
- [14] D. Maulana Fathul Aziz, B. Irawan, and A. Bahtiar, “Analisis Dataset Program Keluarga Harapan (Pkh) Desa Sidaharja Menggunakan Algoritma K-Means,” *JATI (Jurnal Mhs. Tek. Inform.*, vol. 7, no. 6, pp. 3370–3376, 2024, doi: 10.36040/jati.v7i6.8200.