

## Pengelompokan Masyarakat Kurang Mampu Dengan Menggunakan Algoritma K-Means Data Mining

Evan Edward Siagian<sup>1</sup>, Irfansyah Nuddin Lubis<sup>1</sup>, Monita Setya<sup>1</sup>, Ade Dermawan Sijabat<sup>1</sup>,  
David JM Sembiring<sup>1,\*</sup>, Meiliyani Br Ginting<sup>2</sup>

<sup>1</sup>Program Studi Teknik Informatika, Fakultas Sains dan Teknologi, Institut Teknologi dan Bisnis Indonesia, Indonesia

<sup>2</sup>Program Studi Rekayasa Perangkat Lunak, Fakultas Sains dan Teknologi, Institut Teknologi dan Bisnis Indonesia, Indonesia

Jl. Sei Batang Hari No. 84A, Babura Sunggal, Kec. Medan Sunggal, Kota Medan, Sumatera Utara, Indonesia

Email: <sup>1</sup>evaneds@gmail.com, <sup>2</sup>irfansyahnuddinlubis@gmail.com, <sup>3</sup>monitasetya44@gmail.com, <sup>4</sup>adsijabat@gmail.com,

<sup>5,\*</sup>davidjmsembiring@gmail.com, <sup>6</sup>meiliyaniginting@gmail.com

(\* : coresponding author)

**Abstrak**—Beberapa desa sering mengalami kesulitan dalam mengklasifikasikan masyarakat yang kurang mampu secara ekonomi, sehingga penyaluran bansos terkadang salah sasaran. Berbagai hibah yang diterima, seperti pemberian bantuan kepada masyarakat miskin dalam bentuk subsidi. Permasalahan yang dihadapi antara lain adalah data masyarakat yang belum terkelola dengan baik sehingga menyulitkan proses analisis, ketiadaan metode pengelompokan yang terukur sehingga bantuan kerap salah sasaran. Tanpa adanya sistem pengelompokan yang objektif dan berbasis data, kesalahan distribusi bantuan akan terus berulang, mengakibatkan bantuan tidak tepat sasaran. Untuk mengatasi permasalahan tersebut, salah satu solusi yang dapat diterapkan adalah penggunaan teknik data mining. Di masa lampau, pengelolaan data besar sering dilakukan secara manual atau menggunakan metode konvensional yang memerlukan banyak waktu, tenaga, dan biaya. Data mining adalah proses menggali dan menganalisis kumpulan data yang besar untuk menemukan pola, hubungan, atau informasi penting yang dapat mendukung pengambilan keputusan. Algoritma K-Means merupakan salah satu metode clustering dalam data mining yang digunakan untuk mengelompokkan data ke dalam sejumlah kelompok (cluster) berdasarkan kemiripan karakteristik. Tujuan penelitian ini adalah untuk merancang dan mengimplementasikan sistem pengelompokan masyarakat miskin berbasis algoritma K-Means yang dapat membantu pemerintah desa dalam menyalurkan bantuan secara tepat sasaran, mempercepat proses analisis data, dan mengurangi kesalahan distribusi bantuan. Penelitian ini menggunakan 30 data penduduk dengan 5 atribut yaitu pekerjaan, penghasilan, tanggungan, kepemilikan rumah, dan aset. Metode yang digunakan dalam penelitian ini adalah Algoritma K-Means. Dari perhitungan yang telah dilakukan, direkomendasikan sebanyak 3 cluster dengan hasil yang sama, yaitu cluster 1 sebanyak 10 penduduk, cluster 2 sebanyak 10 penduduk, dan cluster 3 sebanyak 10 penduduk juga.

**Kata Kunci:** Data Mining; Masyarakat Kurang Mampu; Pengelompokan; Algoritma K-Means.

**Abstract**—Some villages often experience difficulties in classifying economically disadvantaged communities, resulting in the distribution of social assistance sometimes being misdirected. Various grants are received, such as subsidies provided to the poor. Problems encountered include poorly managed community data, which complicates the analysis process, and the lack of a measurable grouping method, which often misdirects aid. Without an objective, data-driven grouping system, aid distribution errors will continue to recur, resulting in misdirected aid. To address these issues, one solution is the use of data mining techniques. In the past, big data management was often done manually or using conventional methods that required significant time, effort, and expense. Data mining is the process of exploring and analyzing large data sets to discover patterns, relationships, or important information that can support decision-making. The K-Means algorithm is a clustering method in data mining used to group data into groups (clusters) based on similar characteristics. The purpose of this study is to design and implement a system for grouping poor communities based on the K-Means algorithm that can assist village governments in distributing aid precisely to targets, accelerate the data analysis process, and reduce aid distribution errors. This study uses 30 population data with 5 attributes: occupation, income, dependents, home ownership, and assets. The method used in this study is the K-Means Algorithm. From the calculations that have been carried out, it is recommended that there are 3 clusters with the same results, namely cluster 1 with 10 residents, cluster 2 with 10 residents, and cluster 3 with 10 residents as well.

**Keywords:** Data Mining; Underprivileged Communities; Clustering; K-Means Algorithm.

### 1. PENDAHULUAN

Masyarakat adalah sekelompok makhluk hidup yang berhubungan erat dan menjalani kehidupan kolektif karena aturan tertentu, tradisi tertentu, konvensi dan hukum tertentu yang sama. Sistem-sistem masyarakat tersebut saling berhubungan antara manusia dengan manusia lainnya, membentuk suatu kesatuan[1]. Beberapa desa sering mengalami kesulitan dalam mengklasifikasikan masyarakat yang kurang mampu secara ekonomi, sehingga penyaluran bansos terkadang salah sasaran. Berbagai hibah yang diterima, seperti pemberian bantuan kepada masyarakat miskin dalam bentuk subsidi, beras untuk masyarakat miskin (Raskin), bantuan biaya pendidikan (BBPPPA), dan program lainnya, terkadang salah alokasi. Oleh karena itu, diperlukan suatu teknik untuk mengelompokkan masyarakat yang kurang mampu agar pendistribusian bantuan dapat berhasil di masa yang akan datang[2].

Permasalahan yang dihadapi antara lain adalah data masyarakat yang belum terkelola dengan baik sehingga menyulitkan proses analisis, ketiadaan metode pengelompokan yang terukur sehingga bantuan kerap salah sasaran, perbedaan kriteria masyarakat kurang mampu di tiap daerah yang masih bersifat subjektif, serta minimnya pemanfaatan teknologi analisis data di tingkat desa atau kecamatan. Kondisi ini menimbulkan ketidakefisienan dalam distribusi bantuan, bahkan berpotensi menimbulkan ketidakadilan sosial. Urgensi penelitian ini terletak pada ancaman yang akan terjadi jika permasalahan tersebut tidak segera diatasi. Tanpa adanya sistem pengelompokan yang objektif dan berbasis data, kesalahan distribusi bantuan akan terus berulang, mengakibatkan bantuan tidak tepat sasaran, meningkatnya

ketimpangan sosial, dan berkurangnya kepercayaan masyarakat terhadap pemerintah. Selain itu, alokasi anggaran menjadi tidak efisien karena bantuan yang seharusnya diterima oleh masyarakat yang benar-benar membutuhkan justru jatuh ke pihak yang tidak layak, sehingga tujuan program kesejahteraan tidak tercapai secara optimal.

Untuk mengatasi permasalahan tersebut, salah satu solusi yang dapat diterapkan adalah penggunaan teknik data mining. Melalui teknik ini, data masyarakat yang sebelumnya tersebar dan tidak terkelola dengan baik dapat diolah menjadi informasi yang terstruktur dan mudah dianalisis. Proses ini memungkinkan pengelompokan masyarakat secara objektif berdasarkan indikator yang jelas, sehingga meminimalkan kesalahan dalam penentuan penerima bantuan dan memastikan distribusi bantuan menjadi lebih tepat sasaran[3]. Di masa lampau, pengelolaan data besar sering dilakukan secara manual atau menggunakan metode konvensional yang memerlukan banyak waktu, tenaga, dan biaya. Proses ini tidak hanya lambat, tetapi juga rawan kesalahan dalam pencatatan maupun analisis. Seiring bertambahnya volume data, metode lama menjadi semakin tidak efisien dan sulit diandalkan untuk mendukung pengambilan keputusan yang cepat dan akurat. Oleh karena itu, data mining menjadi solusi modern yang mampu mengolah data dalam jumlah besar secara otomatis, cepat, dan dengan tingkat akurasi yang lebih tinggi[4].

Data mining adalah proses menggali dan menganalisis kumpulan data yang besar untuk menemukan pola, hubungan, atau informasi penting yang dapat mendukung pengambilan keputusan. Dengan memanfaatkan data mining, informasi yang sebelumnya tersembunyi di dalam data dapat diungkap dan digunakan sebagai dasar yang kuat dalam perencanaan, evaluasi, dan penetapan kebijakan. Penerapan teknik ini tidak hanya membantu dalam mengelompokkan data, tetapi juga meningkatkan efisiensi pengelolaan informasi dan transparansi dalam setiap proses pengambilan keputusan berbasis data[5]. Algoritma K-Means merupakan salah satu metode clustering dalam data mining yang digunakan untuk mengelompokkan data ke dalam sejumlah kelompok (cluster) berdasarkan kemiripan karakteristik. Proses kerja K-Means dimulai dengan menentukan jumlah cluster yang diinginkan, kemudian algoritma akan mengelompokkan data sedemikian rupa sehingga data dalam satu kelompok memiliki kesamaan yang tinggi, sementara perbedaan antar kelompok menjadi maksimal[6]. Keunggulan K-Means terletak pada kemampuannya memproses data dalam jumlah besar dengan waktu yang relatif singkat, sehingga cocok digunakan untuk pengelompokan masyarakat berdasarkan indikator ekonomi. Dengan menerapkan algoritma ini, pola dan kelompok masyarakat yang memiliki tingkat kesejahteraan serupa dapat diidentifikasi secara lebih jelas, sehingga mendukung penyaluran bantuan yang tepat sasaran dan terukur[7].

Berdasarkan penelitian yang dilakukan oleh S Ramadhani, Ambarita, A M H Pardede tentang “Metode K-Means Untuk Pengelompokan Masyarakat Miskin Dengan Menggunakan Jarak Kedekatan Manhattan City Dan Euclidean (Studi Kasus Kota Binjai)” dapat disimpulkan bahwa penelitian ini berhasil melakukan Klastering Dari 20 data diperoleh 3 group, Cluster 1 terdapat 6 data penjualan, Cluster 2 terdapat 8 data kemiskinan, Cluster 3 terdapat 6 data kemiskinan. Dan diperoleh group terbanyak adalah cluster 2[8]. Berdasarkan penelitian yang dilakukan oleh Sudi suryadi tentang “Penerapan Metode Clustering K-Means Untuk Pengelompokan Kelulusan Mahasiswa Berbasis Kompetensi” dapat disimpulkan bahwa algoritma K-Means dapat digunakan untuk mengelompokkan kelulusan mahasiswa berbasis kompetensi dengan memanfaatkan nilai indeks prestasi kumulatif dan nilai kompetensi mahasiswa itu sendiri[9].

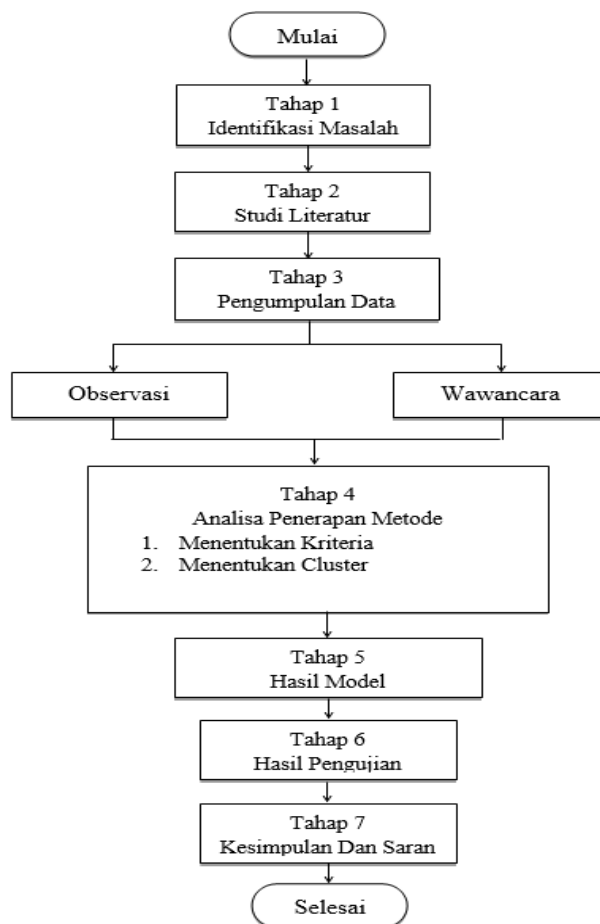
Berdasarkan penelitian yang dilakukan oleh Dina Sunia, Kurniabudi, Pareza Alam Jusia tentang “Penerapan Data Mining Untuk Clustering Data Penduduk Miskin Menggunakan Algoritma K-means” dapat disimpulkan bahwa Penelitian ini menggunakan data penduduk Jambi Selatan yang diambil di Dinas Sosial Kota Jambi dengan jumlah data yang digunakan 512 dan mempunyai 16 atribut yaitu: tanggungan dalam keluarga, anggota keluarga yang cacat, lanjut usia, penguasaan bangunan, luas bangunan, jenis lantai, jenis atap, jenis dinding, pekerjaan utama kepala keluarga, dan pendapatan perbulan. Metode yang digunakan dalam penelitian ini adalah K-Means[10]. Berdasarkan penelitian yang dilakukan oleh Irmanita Nasution, Agus Perdana Windarto, M. Fauzan tentang “Penerapan Algoritma K-Means Dalam Pengelompokan Data Penduduk Miskin Menurut Provinsi” Jumlah record yang digunakan sebanyak 34 provinsi dengan menghasilkan 2 cluster yakni cluster tinggi sebanyak 8 provinsi dan cluster rendah sebanyak 26 provinsi. Berdasarkan hasil pengujian k-means untuk kasus presentase data penduduk miskin menggunakan tools RapidMiner versi 5.3[11].

Berdasarkan studi literatur yang telah dilakukan, sebagian besar penelitian sebelumnya fokus pada penerapan algoritma K-Means dalam pengelompokan data penduduk miskin pada lingkup wilayah yang luas, seperti provinsi atau kota, serta menggunakan jumlah data dan atribut yang relatif besar. Namun, belum banyak penelitian yang mengkaji penerapan algoritma K-Means pada skala desa dengan jumlah data terbatas tetapi relevan dengan kebutuhan lokal, serta mempertimbangkan indikator ekonomi yang spesifik pada kondisi masyarakat setempat. Selain itu, penelitian terdahulu umumnya hanya berfokus pada proses clustering tanpa mengintegrasikan hasilnya ke dalam sistem pendukung keputusan yang dapat digunakan langsung oleh aparat desa untuk mempermudah proses distribusi bantuan. Hal ini menciptakan peluang untuk melakukan penelitian yang tidak hanya menghasilkan model pengelompokan, tetapi juga dapat diimplementasikan secara praktis dalam pengelolaan bantuan sosial di tingkat desa. Dari kesimpulan penelitian tersebut maka dalam penelitian ini akan dilakukan pengelompokan data untuk membantu kelurahan desa dalam pengelompokan penduduk kurang mampu sehingga bantuan dari pemerintahan desa dapat tersalurkan dengan tepat. Dengan demikian perlu dilakukan penelitian untuk pengelompokan masalah masyarakat kurang mampu

## **2. METODOLOGI PENELITIAN**

### **2.1 Kerangka Kerja Penelitian**

Metodologi penelitian adalah suatu langkah yang dimiliki dan dilakukan peneliti dalam rangka mengumpulkan informasi atau data dan melakukan penelitian terhadap data yang diperoleh. Metode penelitian memberikan gambaran tentang desain penelitian, termasuk tetapi tidak terbatas pada: prosedur dan langkah-langkah yang harus diikuti, waktu penelitian, sumber data, dan langkah-langkah dimana data diperoleh dan kemudian dikumpulkan, mengolah dan menganalisis, serta pengumpulan data, yang berkaitan dengan penelitian guna memperoleh data yang dibutuhkan dan memperoleh informasi. Sambil mengumpulkan informasi dan data yang diperlukan.



Gambar 1. Kerangka Kerja Penelitian

Kerangka kerja penelitian pada Gambar 1 menggambarkan alur tahapan yang dilakukan mulai dari identifikasi permasalahan hingga penarikan kesimpulan. Proses dimulai dengan pengumpulan data penduduk yang berisi atribut-atribut relevan seperti pekerjaan, penghasilan, tanggungan, kepemilikan rumah, dan aset. Data tersebut kemudian melalui tahap pra-pemrosesan untuk memastikan kelengkapan dan konsistensi sebelum dianalisis menggunakan algoritma K-Means. Hasil clustering yang diperoleh selanjutnya dianalisis untuk menentukan kategori masyarakat berdasarkan tingkat kesejahteraan. Tahap akhir dari kerangka kerja ini adalah penyusunan rekomendasi dan implementasi hasil pengelompokan ke dalam sistem pendukung keputusan yang dapat membantu pemerintah desa menyalurkan bantuan secara tepat sasaran. Dengan demikian, Gambar 1 menjadi panduan visual yang memudahkan pemahaman alur penelitian secara keseluruhan.

## 2.2 Data Mining

Data mining adalah proses mengekstrak informasi penting dari data. Informasi penting ini diperoleh melalui proses yang sangat rumit seperti penggunaan kecerdasan buatan, teknik statistik, matematika, pembelajaran mesin, dll. Teknik kompleks ini nantinya akan mengidentifikasi dan mengekstrak informasi yang berguna dari database besar. Contoh sederhana adalah membaca data dari buku telepon, misalnya [12]. Dari penjelasan di atas dapat disimpulkan bahwa data mining adalah suatu proses dimana data diekstraksi dalam jumlah yang sangat besar dari data dengan metode matematika statistik menggunakan teknologi kecerdasan buatan terbaru. Istilah data mining memiliki beberapa persamaan, seperti penemuan pengetahuan atau pengenalan pola. Bahkan, mereka memiliki istilah mereka sendiri. [13]

## 2.3 Algoritma K-Means

K-Means adalah metode analisis kelompok yang menghasilkan pembagian n objek pengamatan menjadi k kelompok (cluster), dengan setiap objek pengamatan yang tergabung dalam satu grup dengan mean (mean) terdekat berusaha

memusatkan grup pada data sehingga banyak iterasi perbaikan yang dibuat oleh kedua algoritma. K-Means adalah metode pengelompokan data non-hierarki yang mencoba membagi data yang ada menjadi berkelompok [14].

K-Means adalah algoritma yang mengambil k parameter input dan membagi sekumpulan n objek menjadi k grup sehingga derajat kemiripan antar anggota suatu grup tinggi, sedangkan derajat kemiripan dengan anggota grup lain sangat rendah. Kesamaan anggota dengan kelompok diukur dengan kedekatan benda dengan nilai rata-rata dalam kelompok atau dapat disebut pusat gravitasi atau pusat massa kelompok. [9]. Pengelompokan data menurut metode kmeans biasanya dilakukan sesuai dengan algoritma berikut [15], [16], [17]:

- a. Menentukan jumlah grup
- b. Membagi data menjadi beberapa grup secara acak
- c. Pusat grup (pusat gravitasi / mean) dari data masing-masing

Jika masih ada data yang memindahkan grup, atau jika nilai centroid berubah di atas ambang batas yang ditentukan, atau jika nilai fungsi tujuan yang digunakan masih di atas ambang batas yang ditentukan. Pada langkah 3 dari algoritma Kmeans di atas, posisi titik berat (center) setiap kelompok harus dihitung ulang dari rata-rata (mean) semua nilai data untuk setiap karakteristik. Jika M mewakili jumlah data dalam suatu grup, i mewakili karakteristik ke i dalam suatu grup, p mewakili dimensi data untuk menghitung centroid dari karakteristik. Adapun rumus pada algoritma K-Means adalah sebagai berikut [18], [19], [20]:

$$d_{ij} = \sqrt{(x_{1i} - x_{1j})^2 + (x_{2i} - x_{2j})^2 + (x_{ki} - x_{kj})^2} \tag{1}$$

### 3. HASIL DAN PEMBAHASAN

#### 3.1 Analisa Masalah

Algoritma K-Means merupakan salah satu teknik dalam data mining untuk mengelompokkan data menjadi beberapa kelompok berdasarkan jarak, kriteria, kondisi atau karakteristik (pengelompokan). Data dalam suatu kelompok harus memiliki jarak, kriteria, kondisi, atau karakteristik terpendek yang sama atau hampir sama satu sama lain. Algoritma K-Means dapat mengelompokkan objek dengan nilai. Saat mengelompokkan masyarakat kurang mampu digunakan data masyarakat desa Hutagodang dengan algoritma K-Means. Algoritma K-Means pada proses cluster tidak dapat menggunakan data berupa huruf, sehingga kriteria konversi berupa angka pengelompokan data dengan algoritma K-Means.

- a. Menyiapkan sampling data
- b. Menentukan jumlah cluster
- c. Menentukan nilai centroid/titik pusat
- d. Menghitung jarak masing masing nilai centroid
- e. Mengelompokkan data berdasarkan jarak terpendek

Untuk menentukan jarak setiap centroid berdasarkan jarak terpendek, kriteria, kondisi atau karakteristik yang sama atau hampir sama menggunakan model Euclidean, yang dirumuskan sebagai berikut:

$$d_{ij} = \sqrt{(x_{1i} - x_{1j})^2 + (x_{2i} - x_{2j})^2 + (x_{ki} - x_{kj})^2} \tag{1}$$

Keterangan

$d_{ij}$  = Jarak dari data ke i ke pusat cluster j

$x_{ki}$  = Data dari ke - i pada *attribute* data ke - k

$x_{kj}$  = Data dari ke - j pada *attribute* data ke - k

Tabel 1. Data Masyarakat

No	Nama	Kriteria Dan Penilaian				
		Pekerjaan	Penghasilan	Tanggung	Kepemilikan Rumah	Aset
1.	Basri	Petani	Rp.2.200.000	2	1	2
2.	Iswan	Petani	Rp.2.300.000	2	1	2
3.	Puli	Pedagang	Rp.2.100.000	1	2	1
4.	Asru	Pedagang	Rp.2.400.000	1	1	1
5.	Leman	Petani	Rp.2.250.000	3	2	1
6.	Gini Pa	Pedagang	Rp.2.350.000	4	2	2
7.	Pangih	Pedagang	Rp.2.500.000	2	1	2
8.	Rendi	Petani	Rp.2.200.000	3	2	1
9.	Surya	Petani	Rp.2.100.000	3	2	1
10.	Dadang	Pedagang	Rp.2000.000	4	2	2
11.	Aman L	Petani	Rp.2.150.000	3	1	1
12.	Hendra	Pedagang	Rp.2.400.000	3	1	1
13.	Sulaiman	Petani	Rp.2.450.000	3	2	1
14.	Zulkifli	Pedagang	Rp.2.550.000	3	1	1

No	Nama	Kriteria Dan Penilaian				
		Pekerjaan	Penghasilan	Tanggung	Kepemilikan Rumah	Aset
15.	Sakti	Petani	Rp.2.600.000	3	2	1
16.	Kokman	Petani	RP.2.300.000	3	2	1
17.	Boy	Petani	RP.2000.000	3	1	1
18.	Udin	Pedagang	RP.2.450.000	3	2	1
19.	Amir	Petani	RP.2.150.000	3	2	1
20.	Akhir	Petani	RP.2.500.000	3	1	1
21.	Naja	Petani	RP.2.350.000	4	1	2
22.	Fery	Petani	RP.2.100.000	3	2	1
23.	Daulat	Pedagang	RP.2.200.000	3	2	1
24.	Togu	Petani	RP.2.650.000	2	1	2
25.	Misdi	Petani	RP.2.600.000	4	2	2
26.	Ulong	Petani	RP.2.400.000	3	2	1
27.	Lela	Pedagang	RP.2.345.000	1	1	1
28.	Barrik	Pedagang	RP.2.540.000	1	2	1
29.	Justan	Petani	RP.2.500.000	2	1	2
30.	Ali	Petani	RP.2.400.000	2	1	2

Data yang digunakan dalam penelitian ini ditampilkan pada Tabel 1 yang memuat 30 data penduduk dengan lima atribut penilaian, yaitu pekerjaan, penghasilan, tanggungan, kepemilikan rumah, dan aset. Atribut-atribut tersebut dipilih karena dinilai relevan untuk menentukan tingkat kesejahteraan masyarakat. Nilai penghasilan dinyatakan dalam rupiah, sedangkan atribut tanggungan, kepemilikan rumah, dan aset direpresentasikan dengan skala penilaian numerik yang memudahkan proses analisis. Variasi data pada Tabel 1 menunjukkan adanya perbedaan signifikan pada masing-masing atribut antarpenduduk, sehingga menjadi dasar dalam proses pengelompokan (clustering) menggunakan algoritma K-Means. Data ini kemudian akan melalui tahap pra-pemrosesan untuk memastikan kelengkapan dan konsistensinya sebelum digunakan dalam analisis lebih lanjut.

### 3.1.1 Penerapan Algoritma K-Means

Setelah didapatkan data pada analisis masalah pada penelitian, maka tahapan selanjutnya adalah melakukan proses penyelesaian dengan menggunakan Algoritma K-Means. Adapun proses dapat dilihat berikut:

- Iterasi Ke-1: Jumlah cluster / kelompok yaitu 3 (tiga) sesuai dengan jumlah attribute kriteria.
- Menentukan Nilai Centroid Awal  
Menentukan nilai titik berat awal / titik berat awal dengan 3 (tiga) record sesuai dengan jumlah kriteria atribut. Nilai center of gravity awal / pusat diperoleh secara acak dari data masyarakat.

**Tabel 2.** Nilai Centroid Awal

Basri	C1	2	1	2
Gini Pa	C2	4	2	2
Aman L	C3	3	1	1

Pada Tabel 2 menampilkan nilai centroid awal yang digunakan sebagai titik referensi dalam proses pengelompokan data menggunakan algoritma K-Means. Centroid awal ini dipilih dari data yang ada untuk mewakili masing-masing cluster secara sementara sebelum proses iterasi dimulai.

- Menghitung jarak dengan menggunakan model Euclidean

#### Cluster C1:

$$d(x1,c1) = \sqrt{(a1 - c1a)^2 + (a1 - c1b)^2 + (a1 - c1c)^2}$$

$$= \sqrt{(2 - 2)^2 + (1 - 1)^2 + (2 - 2)^2} = 0$$

$$d(x2,c1) = \sqrt{(a2 - c1a)^2 + (b2 - c1b)^2 + (c2 - c1c)^2}$$

$$= \sqrt{(2 - 2)^2 + (2 - 1)^2 + (2 - 2)^2} = 1$$

$$d(x3,c1) = \sqrt{(a3 - c1a)^2 + (b3 - c1b)^2 + (c3 - c1c)^2}$$

$$= \sqrt{(1 - 2)^2 + (1 - 1)^2 + (1 - 2)^2} = 2$$

$$d(x4,c1) = \sqrt{(a4 - c1a)^2 + (b4 - c1b)^2 + (c4 - c1c)^2}$$

$$= \sqrt{(1 - 2)^2 + (1 - 1)^2 + (1 - 2)^2} = 2$$

$$d(x5,c1) = \sqrt{(a5 - c1a)^2 + (b5 - c1b)^2 + (c5 - c1c)^2}$$

$$= \sqrt{(3 - 2)^2 + (2 - 1)^2 + (1 - 2)^2} = 3$$

#### Cluster C2:

$$d(x1,c2) = \sqrt{(a1 - c2a)^2 + (a1 - c2b)^2 + (a1 - c2c)^2}$$

$$= \sqrt{(2 - 4)^2 + (2 - 2)^2 + (2 - 2)^2} = 2$$

$$d(x2,c2) = \sqrt{(a2 - c2a)^2 + (b2 - c2b)^2 + (c2 - c2c)^2} \\ = \sqrt{(3 - 4)^2 + (3 - 2)^2 + (3 - 2)^2} = 2$$

$$d(x3,c2) = \sqrt{(a3 - c2a)^2 + (b3 - c2b)^2 + (c3 - c2c)^2} \\ = \sqrt{(1 - 4)^2 + (1 - 2)^2 + (1 - 2)^2} = 11$$

$$d(x4,c2) = \sqrt{(a4 - c2a)^2 + (b4 - c2b)^2 + (c4 - c2c)^2} \\ = \sqrt{(1 - 4)^2 + (1 - 2)^2 + (1 - 2)^2} = 11$$

$$d(x5,c2) = \sqrt{(a5 - c2a)^2 + (b5 - c2b)^2 + (c5 - c2c)^2} \\ = \sqrt{(3 - 4)^2 + (2 - 2)^2 + (1 - 2)^2} = 2$$

**Cluster C3:**

$$d(x1,c3) = \sqrt{(a1 - c3a)^2 + (a1 - c3b)^2 + (a1 - c3c)^2} \\ = \sqrt{(2 - 3)^2 + (2 - 1)^2 + (2 - 1)^2} = 3$$

$$d(x2,c3) = \sqrt{(a2 - c3a)^2 + (b2 - c3b)^2 + (c2 - c3c)^2} \\ = \sqrt{(2 - 3)^2 + (2 - 1)^2 + (2 - 1)^2} = 3$$

$$d(x3,c3) = \sqrt{(a3 - c3a)^2 + (b3 - c3b)^2 + (c3 - c3c)^2} \\ = \sqrt{(1 - 3)^2 + (1 - 1)^2 + (1 - 1)^2} = 4$$

$$d(x4,c3) = \sqrt{(a4 - c3a)^2 + (b4 - c3b)^2 + (c4 - c3c)^2} \\ = \sqrt{(1 - 3)^2 + (1 - 1)^2 + (1 - 1)^2} = 4$$

$$d(x5,c3) = \sqrt{(a5 - c3a)^2 + (b5 - c3b)^2 + (c5 - c3c)^2} \\ = \sqrt{(3 - 3)^2 + (2 - 1)^2 + (1 - 1)^2} = 1$$

Rangkuman hasil perhitungan cluster pada iterasi pertama adalah sebagai berikut: Yang dihasilkan dari penjumlahan anggota setiap cluster, setelah itu dibagi dengan jumlah anggota masing-masing cluster.

**Tabel 3.** Hasil Perhitungan Iterasi I

No	Nama	Tanggungungan	Kepemilikan Rumah	Aset	dc1	dc2	dc3	Cluster
1.	Basri	2	1	2	0	2	3	1
2.	Iswan	2	1	2	1	2	3	1
3.	Puli	1	2	1	2	11	4	1
4.	Asru	1	1	1	2	11	4	1
5.	Leman	3	2	1	3	2	1	3
6.	Gini Pa	4	2	2	9	0	3	2
7.	Pangih	2	1	2	1	4	3	1
8.	Rendi	3	2	1	3	2	1	3
9.	Surya	3	2	1	3	2	1	3
10.	Dadang	4	2	2	9	0	3	2
11.	Aman L	3	1	1	6	3	8	2
12.	Hendra	3	1	1	6	3	8	2
13.	Sulaiman	3	2	1	3	2	1	3
14.	Zulkifli	3	1	1	6	3	8	2
15.	Sakti	3	2	1	3	2	1	3
16.	Kokman	3	2	1	3	2	1	3
17.	Boy	3	1	1	6	3	8	2
18.	Udin	3	2	1	3	2	1	3
19.	Amir	3	2	1	6	3	8	2
20.	Akhir	3	1	1	6	3	8	2
21.	Naja	4	1	2	9	0	3	2
22.	Fery	3	2	1	3	2	1	3
23.	Daulat	3	2	1	3	2	1	3
24.	Togu	2	1	2	1	4	3	1
25.	Misdi	4	2	2	9	0	3	2
26.	Ulong	3	2	1	3	2	1	3
27.	Lela	1	1	1	2	11	4	1
28.	Barrik	1	2	1	2	11	4	1
29.	Justan	2	1	2	1	2	3	1
30.	Ali	2	1	2	0	2	3	1

Pada Tabel 3 menyajikan hasil perhitungan pada Iterasi I dalam proses pengelompokan data menggunakan algoritma K-Means. Pada tahap ini, setiap data dihitung jaraknya terhadap masing-masing centroid awal yang telah ditentukan sebelumnya. Aturan pengelompokan data berdasarkan hasil perhitungan jarak adalah Jika

1. Jika  $dc1 \rightarrow < dc2$  dan  $dc1 < dc3$  maka cluster =1
2. Jika  $dc2 \rightarrow < dc1$  dan  $dc2 < dc3$  maka cluster =2
3. Jika  $dc3 \rightarrow < dc1$  dan  $dc3 < dc2$  maka cluster =3

d. Iterasi Ke-2:

1. Pencarian nilai centroid/ titik pusat awal pada iterasi kedua diperoleh dari nilai mean jumlah anggota tiap kelompok.

Perolehan nilai C1:

**Tabel 4.** Perolehan Nilai C1 Untuk Iterasi Kedua

No.	Nama	Tanggung	Kepemilikan Rumah	Aset	Cluster
1.	Basri	2	1	2	1
2.	Iswan	2	1	2	1
3.	Puli	1	2	1	1
4.	Asrul	1	1	1	1
5.	Pangih	2	1	2	1
6.	Togu	2	1	2	1
7.	Lela	1	1	1	1
8.	Barrik	1	2	1	1
9.	Justan	2	1	2	1
10.	Ali	2	1	2	1
	C1	$(2+2+1+1+2+2+1+1+2+2)/10 = 1,6$	$(1+1+2+1+1+1+1+2+1+1)/10 = 1,2$	$(2+2+1+1+2+2+1+1+2+2)/10 = 1,6$	

Padat Tabel 4 menampilkan hasil perhitungan jarak data terhadap centroid C1 pada Iterasi Kedua dalam proses pengelompokan menggunakan algoritma K-Means. Pada tahap ini, nilai centroid telah diperbarui berdasarkan rata-rata data yang tergabung dalam cluster pada Iterasi Pertama. Perolehan nilai C2:

**Tabel 5.** Perolehan Nilai C2 Untuk Iterasi Kedua

No.	Nama	Tanggung	Kepemilikan Rumah	Aset	Cluster
1.	Gini Pa	4	2	2	2
2.	Dadang	4	2	2	2
3.	Aman L	3	1	1	2
4.	Hendra	3	1	1	2
5.	Zulkifli	3	1	1	2
6.	Boy	3	1	1	2
7.	Amir	3	2	1	2
8.	Akhir	3	1	1	2
9.	Naja	4	1	2	2
10.	Misdi	4	2	2	2
	C1	$(4+4+3+3+3+3+3+3+4+4)/10 = 3,4$	$(2+2+1+1+1+1+2+1+1+2)/10 = 1,4$	$(2+2+1+1+1+1+1+1+2+2)/10 = 1,4$	

Pada Tabel 5 menampilkan hasil perhitungan jarak data terhadap centroid C2 pada Iterasi Kedua dalam proses pengelompokan menggunakan algoritma K-Means. Pada tahap ini, nilai centroid telah diperbarui berdasarkan rata-rata data yang tergabung dalam cluster pada Iterasi Pertama. Perolehan nilai C3:

**Tabel 6.** Perolehan Nilai C3 Untuk Iterasi Kedua

No.	Nama	Tanggung	Kepemilikan Rumah	Aset	Cluster
1.	Leman	3	2	1	3
2.	Rendi	3	2	1	3
3.	Surya	3	2	1	3
4.	Sulaiman	3	2	1	3
5.	Sakti	3	2	1	3
6.	Kokman	3	2	1	3
7.	Udin	3	2	1	3
8.	Fery	3	2	1	3
9.	Daulat	3	2	1	3
10.	Ulong	3	2	1	3

No.	Nama	Tanggungjan	Kepemilikan Rumah	Aset	Cluster
	C1	(3+3+3+3+3+3+3+3+3+3)/10 = 3	(2+2+2+2+2+2+2+2+2)/10 = 2	(1+1+1+1+1+1+1+1+1+1)/10 = 1	

Pada Tabel 5 menampilkan hasil perhitungan jarak data terhadap centroid C2 pada Iterasi Kedua dalam proses pengelompokan menggunakan algoritma K-Means. Pada tahap ini, nilai centroid telah diperbarui berdasarkan rata-rata data yang tergabung dalam cluster pada Iterasi Pertama. Sehingga nilai centroid awal / titik pusat awal adalah sebagai berikut:

Tabel 7. Nilai Centroid Awal Untuk Iterasi Kedua

C1	1,6	1,2	1,6
C2	3,4	1,4	1,4
C3	3	2	1

Pada Tabel 7 menampilkan nilai centroid awal yang digunakan pada Iterasi Kedua dalam proses clustering menggunakan algoritma K-Means. Nilai-nilai ini diperoleh dari rata-rata setiap atribut pada data yang tergabung dalam masing-masing cluster pada Iterasi Pertama.

- Menghitung jarak dengan menggunakan model Euclidean pada iterasi ke-2 adalah:

**Cluster C1:**

$$d(x1,c1) = \sqrt{(a1 - c1a)^2 + (a1 - c1b)^2 + (a1 - c1c)^2}$$

$$= \sqrt{(2 - 1,6)^2 + (1 - 1,2)^2 + (2 - 1,6)^2} = 0,36$$

$$d(x2,c1) = \sqrt{(a2 - c1a)^2 + (b2 - c1b)^2 + (c2 - c1c)^2}$$

$$= \sqrt{(2 - 1,6)^2 + (1 - 1,2)^2 + (2 - 1,6)^2} = 0,36$$

$$d(x3,c1) = \sqrt{(a3 - c1a)^2 + (b3 - c1b)^2 + (c3 - c1c)^2}$$

$$= \sqrt{(1 - 1,6)^2 + (2 - 1,2)^2 + (1 - 1,6)^2} = 1,36$$

$$d(x4,c1) = \sqrt{(a4 - c1a)^2 + (b4 - c1b)^2 + (c4 - c1c)^2}$$

$$= \sqrt{(1 - 1,6)^2 + (1 - 1,2)^2 + (1 - 1,6)^2} = 0,76$$

$$d(x5,c1) = \sqrt{(a5 - c1a)^2 + (b5 - c1b)^2 + (c5 - c1c)^2}$$

$$= \sqrt{(3 - 1,6)^2 + (2 - 1,2)^2 + (1 - 1,6)^2} = 2,96$$

**Cluster C2:**

$$d(x1,c2) = \sqrt{(a1 - c2a)^2 + (a1 - c2b)^2 + (a1 - c2c)^2}$$

$$= \sqrt{(2 - 3,4)^2 + (1 - 1,4)^2 + (2 - 1,4)^2} = 2,48$$

$$d(x2,c2) = \sqrt{(a2 - c2a)^2 + (b2 - c2b)^2 + (c2 - c2c)^2}$$

$$= \sqrt{(2 - 3,4)^2 + (1 - 1,4)^2 + (2 - 1,4)^2} = 2,48$$

$$d(x3,c2) = \sqrt{(a3 - c2a)^2 + (b3 - c2b)^2 + (c3 - c2c)^2}$$

$$= \sqrt{(1 - 3,4)^2 + (2 - 1,4)^2 + (1 - 1,4)^2} = 6,28$$

$$d(x4,c2) = \sqrt{(a4 - c2a)^2 + (b4 - c2b)^2 + (c4 - c2c)^2}$$

$$= \sqrt{(1 - 3,4)^2 + (1 - 1,4)^2 + (1 - 1,4)^2} = 6,08$$

$$d(x5,c2) = \sqrt{(a5 - c2a)^2 + (b5 - c2b)^2 + (c5 - c2c)^2}$$

$$= \sqrt{(3 - 3,4)^2 + (2 - 1,4)^2 + (1 - 1,4)^2} = 0,68$$

**Cluster C3:**

$$d(x1,c3) = \sqrt{(a1 - c3a)^2 + (a1 - c3b)^2 + (a1 - c3c)^2}$$

$$= \sqrt{(2 - 3)^2 + (1 - 2)^2 + (2 - 1)^2} = 3$$

$$d(x2,c3) = \sqrt{(a2 - c3a)^2 + (b2 - c3b)^2 + (c2 - c3c)^2}$$

$$= \sqrt{(2 - 3)^2 + (1 - 2)^2 + (2 - 1)^2} = 3$$

$$d(x3,c3) = \sqrt{(a3 - c3a)^2 + (b3 - c3b)^2 + (c3 - c3c)^2}$$

$$= \sqrt{(1 - 3)^2 + (2 - 2)^2 + (1 - 1)^2} = 4$$

$$d(x4,c3) = \sqrt{(a4 - c3a)^2 + (b4 - c3b)^2 + (c4 - c3c)^2}$$

$$= \sqrt{(1 - 3)^2 + (1 - 2)^2 + (1 - 1)^2} = 5$$

$$d(x5,c3) = \sqrt{(a5 - c3a)^2 + (b5 - c3b)^2 + (c5 - c3c)^2}$$

$$= \sqrt{(3 - 3)^2 + (2 - 2)^2 + (1 - 1)^2} = 0$$

Rangkuman hasil perhitungan cluster pada iterasi kedua adalah sebagai berikut: Didapatkan dari jumlah anggota setiap cluster, kemudian dibagi dengan jumlah anggota setiap cluster.

**Tabel 8.** Rekapitulasi Hasil Perhitungan pada Iterasi Kedua

No	Nama	Tanggungungan	Kepemilikan Rumah	Aset	dc1	dc2	dc3	Cluster
1.	Basri	2	1	2	0,36	2,48	3	1
2.	Iswan	2	1	2	0,36	2,48	3	1
3.	Puli	1	2	1	1,36	6,28	4	1
4.	Asru	1	1	1	0,76	6,08	5	1
5.	Leman	3	2	1	2,96	0,68	0	3
6.	Gini Pa	4	2	2	6,56	1,08	2	2
7.	Pangih	2	1	2	0,36	2,48	3	1
8.	Rendi	3	2	1	2,96	0,68	0	3
9.	Surya	3	2	1	2,96	0,68	0	3
10.	Dadang	4	2	2	6,56	1,08	2	2
11.	Aman L	3	1	1	2,36	0,48	2	2
12.	Hendra	3	1	1	2,36	0,48	2	2
13.	Sulaiman	3	2	1	2,96	0,68	0	3
14.	Zulkifli	3	1	1	2,36	0,48	2	2
15.	Sakti	3	2	1	2,96	0,68	0	3
16.	Kokman	3	2	1	2,96	0,68	0	3
17.	Boy	3	1	1	2,36	0,48	2	2
18.	Udin	3	2	1	2,96	0,68	0	3
19.	Amir	3	2	1	2,36	0,48	2	2
20.	Akhir	3	1	1	2,36	0,48	2	2
21.	Naja	4	1	2	6,56	1,08	2	2
22.	Fery	3	2	1	2,96	0,68	0	3
23.	Daulat	3	2	1	2,96	0,68	0	3
24.	Togu	2	1	2	0,36	2,48	3	1
25.	Misdi	4	2	2	6,56	1,08	2	2
26.	Ulong	3	2	1	2,96	0,68	0	3
27.	Lela	1	1	1	0,76	6,08	5	1
28.	Barrik	1	2	1	1,36	6,28	4	1
29.	Justan	2	1	2	0,36	2,48	3	1
30.	Ali	2	1	2	0,36	2,48	3	1

Pada Tabel 8 menyajikan rekapitulasi hasil perhitungan jarak setiap data terhadap centroid pada Iterasi Kedua. Dalam tabel ini, terlihat distribusi anggota setiap cluster berdasarkan jarak terdekat antara data dengan pusat cluster (centroid) yang telah diperbarui dari Iterasi Pertama. Proses perhitungan ini bertujuan untuk memastikan setiap data ditempatkan pada cluster yang paling sesuai dengan karakteristiknya. Hasil rekapitulasi menunjukkan adanya pergeseran beberapa anggota cluster dibandingkan iterasi sebelumnya, yang menandakan proses penyesuaian menuju kondisi konvergen di mana pembagian anggota cluster tidak lagi berubah secara signifikan. Aturan pengelompokan / pengelompokan data berdasarkan hasil perhitungan jarak adalah Jika

- a.  $dc1 < dc2$  dan  $dc1 < dc3$  maka cluster =1
- b. Jika  $dc2 < dc1$  dan  $dc2 < dc3$  maka cluster =2
- c. Jika  $dc3 < dc1$  dan  $dc3 < dc2$  maka cluster =3

**Tabel 9.** Pengelompokan / Cluster Data Berdasarkan Jarak Terdekat

Cluster 1	Cluster 2	Cluster 3
Basri	Gini Pa	Leman
Iswan	Dadang	Rendi
Puli	Aman L	Surya
Asrul	Hendra	Sulaiman
Pangih	Zulkifli	Sakti
Togu	Boy	Kokman
Lela	Amir	Udin
Barrik	Akhir	Fery
Justan	Naja	Daulat
Ali	Misdi	Ulong

Pada Tabel 9 menampilkan hasil akhir pengelompokan data ke dalam cluster berdasarkan jarak terdekat antara setiap data dengan centroid. Pengelompokan ini merupakan tahap penting dalam proses analisis, karena menentukan posisi akhir setiap data pada cluster yang memiliki karakteristik paling mirip.

Hasil cluster pada iterasi ke-1 dengan iterasi kedua sama sehingga tidak dilanjutkan lagi prosesnya keiterasi ke-3. Nama masyarakat tersebut yaitu:

- a. Anggota cluster ke-1 adalah : Basri, Iswan, Puli, Asrul, Pangih, Togu, Lela, Barrik, Justan, Ali
- b. Anggota cluster ke-2 adalah : Gini Pa, Dadang, Aman L, Hendra, Zulkifli, Boy, Amir, Akhir, Naja, Misdi
- c. Anggota cluster ke-3 adalah : Leman, Rendi, Surya, Sulaiman, Sakti, Kokman, Udin, Fery, Daulat, Ulong

#### 4. KESIMPULAN

Berdasarkan hasil observasi atau investigasi dapat disimpulkan sebagai berikut: Penelitian ini menggunakan data dengan jumlah data yang digunakan sebanyak 30 data penduduk yang mempunyai 5 atribut yaitu, Pekerjaan, penghasilan, tanggungan, kepemilikan rumah, aset. Metode yang digunakan dalam penelitian ini adalah Algoritma K-Means dari perhitungan yang telah dilakukan, Maka direkomendasikan penulis sebanyak 3 cluster. Hasil perbandingan perhitungan sama yaitu, cluster 1 sebanyak 10 penduduk, cluster 2 sebanyak 10 penduduk dan cluster 3 sebanyak 10 penduduk juga.

#### REFERENCES

- [1] R. Aditiya Perdana and P. Aisyiah Rakhma Devi, "Impelementasi Data Mining Untuk Clustering Penduduk Miskin Di Desa Setrohadi Menggunakan Algoritma Pillar K-Means," *JATI (Jurnal Mhs. Tek. Inform.,* vol. 8, no. 3, pp. 3780–3787, 2024, doi: 10.36040/jati.v8i3.9796.
- [2] V. Hutagaol, D. Anggraeni, and A. Nata, "Penerapan Data Mining Untuk Clustering Masyarakat Kurang Mampu Pada Desa Rawang Pasar Vi Dengan Algoritma K-Means," *JUPI (Jurnal Ilm. Penelit. dan Pembelajaran Inform.,* vol. 10, no. 1, pp. 304–312, 2025, doi: 10.29100/jupi.v10i1.5741.
- [3] S. Amaliyah, Jasmir, and S. Rianti, "Penerapan Data Mining Untuk Menentukan Kelompok Prioritas Penerima Bantuan PKH Menggunakan Metode Clustering K-Means Pada Desa Kuala Dendang," *J. Inform. Dan Rekayasa Komputer(JAKAKOM),* vol. 3, no. 1, pp. 453–458, 2023, doi: 10.33998/jakakom.2023.3.1.802.
- [4] 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.
- [5] 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>
- [6] A. Waruwu, M. Yetri, and F. Setiawan, "Implementasi Data Mining Dalam Mengelompokkan Data penduduk Kurang Mampu Menggunakan Metode K-Means Clustering," *J. Sist. Inf. Triguna Dharma (JURSI TGD),* vol. 2, no. 6, p. 945, 2023, doi: 10.53513/jursi.v2i6.8965.
- [7] B. H. Prakoso, E. Rachmawati, D. R. P. Mudiono, V. Vestine, and G. E. J. Suyoso, "Klasterisasi Puskesmas dengan K-Means Berdasarkan Data Kualitas Kesehatan Keluarga dan Gizi Masyarakat," *J. Buana Inform.,* vol. 14, no. 01, pp. 60–68, 2023, doi: 10.24002/jbi.v14i01.7105.
- [8] S. Ramadani, I. Ambarita, and A. M. H. Pardede, "Metode K-Means Untuk Pengelompokan Masyarakat Miskin Dengan Menggunakan Jarak Kedekatan Manhattan City Dan Euclidean ( Studi Kasus Kota Binjai )," *Inf. Syst. Dev.,* vol. 04, no. 2, pp. 15–29, 2019.
- [9] S. Suryadi, "Penerapan Metode Clustering K-Means Untuk Pengelompokan Kelulusan Mahasiswa Berbasis Kompetensi," *J. Inform.,* vol. 6, no. 1, pp. 52–72, 2019, doi: 10.36987/informatika.v6i1.738.
- [10] P. A. J. Dina Sunia, Kurniabudi, "Penerapan Data Mining untuk Clustering Data Penduduk Miskin Menggunakan Algoritma K-Means," *J. Ilm. Mhs. Tek. Inform.,* vol. Vol 1 No 2, no. 2016, pp. 121–134, 2019.
- [11] I. Nasution, A. P. Windarto, and M. Fauzan, "Penerapan Algoritma K-Means Dalam Pengelompokan Data Penduduk Miskin Menurut Provinsi," *Build. Informatics, Technol. Sci.,* vol. 2, no. 2, pp. 76–83, 2020, doi: 10.47065/bits.v2i2.492.
- [12] M. I. Zuhendra, R. Hidayat, and H. Hendrawaty, "Penerapan Data Mining Untuk Klasterisasi Tingkat Kemiskinan Berdasarkan Data Terpadu Kesejahteraan Sosial (Dtks)," *SKANIKA Sist. Komput. dan Tek. Inform.,* vol. 7, no. 1, pp. 32–41, 2024, doi: 10.36080/skanika.v7i1.3149.
- [13] D. L. Fay, "Pengantar Data Mining," *Angew. Chemie Int. Ed. 6(11),* 951–952., 1967.
- [14] A. Hidayat, Nurhidayati, and amri muliawan Nur, "Implementasi Algoritma K-Means Untuk Klasterisasi Peserta," *Print. J. Pengemb. Rekayasa Inform. dan Komput.,* vol. 1, no. 2, pp. 154–166, 2023.
- [15] R. R. A. Aria, S. Susilowati, and I. R. Rahadjeng, "Data Mining Menentukan Cluster Penerima Program Bantuan dengan Metode K-Means," *Remik,* vol. 7, no. 1, pp. 291–300, 2023, doi: 10.33395/remik.v7i1.12030.
- [16] 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 (Internasional 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>
- [17] A. Fatlun, S. S. Hilabi, B. Priyatna, and A. L. Hananto, "Penggunaan Algoritma K-Means Clustering CRISP-DM Dalam Mengelompokkan Drama Korea Sebagai Rekomendasi Film," *JATISI (Jurnal Tek. Inform. dan Sist. Informasi),* vol. 12, no. 1, pp. 130–140, 2025, doi: 10.35957/jatisi.v12i1.11347.
- [18] Egy Andryan, A. Abdullah, and P. Y. Utami, "Clustering Penduduk Kurang Mampu Di Desa Mekar Baru Menggunakan Algoritma K-Means," *J. Tek. Komput. AMIK BSI,* vol. 8, no. 2, pp. 174–180, 2022, doi: 10.31294/jtk.v4i2.
- [19] A. Setyo Nugroho, "Penerapan Metode K-Means Clustering Untuk Penentuan Eligibilitas Bantuan Langsung Tunai Berbasis Mobile," *J. Inf. Technol. Ampera,* vol. 5, no. 1, pp. 2774–2121, 2024, doi: 10.51519/journalita.v5i1.510.
- [20] Dayla May Cytry, S. Defit, and G. Nurcahyo, "Penerapan Metode K-Means dalam Klasterisasi Status Desa terhadap Keluarga Beresiko Stunting," *J. KomtekInfo,* vol. 10, no. 3, pp. 122–127, 2023, doi: 10.35134/komtekinfo.v10i3.423.