

Clustering Keluarga Miskin Desa Bina Baru dengan Metode K-Medoids

Felina Amelia, Iwan Iskandar, Siska Kurnia Gusti, Elin Haerani, Yusra
Fakultas Sains dan Teknologi, Teknik Informatika, Universitas Islam Negeri Sultan Syarif
Kasim Riau, Pekanbaru, Indonesia

*E-mail koresponden: : iwan.iskandar@uin-suska.ac.id

Diserahkan 18 April 2023; Direview 1 Mei 2023; Dipublikasikan 30 Mei 2023

Abstrak

Kemiskinan di Indonesia terjadi di berbagai daerah, mulai pedesaan hingga perkotaan memiliki permasalahan kemiskinan masing – masing. Masalah kemiskinan juga dialami oleh Desa Bina Baru. Desa Bina Baru yang memiliki jumlah penduduk sebanyak 5.760 jiwa dengan total 1.742 keluarga, yang tersebar dalam 30 Rukun Tetangga (RT) dan 8 Rukun Warga (RW). Upaya dalam penurunan angka kemiskinan dapat dilakukan dengan berbagai cara, mulai pembangunan yang merata, penyaluran bantuan yang tepat sasaran, pemberian kebijakan yang tepat, dan lain sebagainya. Pengelompokan kemiskinan menjadikan salah satu upaya untuk menurunkan angka kemiskinan agar dapat memberikan informasi kepada pemerintahan daerah dalam memberikan kebijakan yang lebih tepat guna. Clustering merupakan teknik data mining yang bertujuan untuk mengelompokkan objek-objek data menjadi beberapa Cluster. Pada penelitian ini pengelompokan dilakukan dengan teknik pengolahan data mining dengan algoritme K-Medoids dari data Desa Bina Baru tahun 2020 berjumlah 1.005. Hasil perbandingan perhitungan untuk Cluster 1 (kaya) sebanyak 527 penduduk, Cluster 2 (menengah) sebanyak 248 penduduk, dan Cluster 3 (miskin) sebanyak 225 penduduk, Hasil evaluasi dari algoritme k-Medoids adalah 0,991 yang menunjukan cluster yang dibentuk memberikan pengelompokan informasi yang baik. Hasil pengelompokan ini dapat dijadikan acuan untuk informasi kelompok keluarga miskin yang diperlukan pemerintah agar bantuan yang diberikan tepat sasaran.

Kata kunci: *Clustering; Data Mining; K-Medoids; Kemiskinan; Knowledge Discovery Database.*

Abstract

In Indonesia, poverty can be found in both rural and urban areas, and it affects each of these locations differently. Bina Baru Village is likewise plagued by the issue of poverty. A total of 1,742 families make up the 5,760 inhabitants of Bina Baru Village, which is divided into 30 Neighborhood Units (RT) and 8 Residential Units (RW). There are numerous strategies to combat poverty, including equitable development distribution, targeted aid distribution, the implementation of suitable laws, and so on. Poverty grouping is one of the efforts to reduce the poverty rate by providing information to regional governments to deliver more effective policies. A data mining approach called clustering seeks to divide data objects into various groupings. This grouping research was carried out using data mining techniques with the K-Medoids algorithm from the Bina Baru Village data for 2020, totalling 1,005. In comparison, there are 527 people in Cluster 1 (the wealthy), 248 in Cluster 2 (the middle class), and 225 in Cluster 3 (the poor), The evaluation result of the k-Medoids algorithm is 0.991, which indicates that the clusters formed to provide good information grouping. The results of this grouping can

be used as a reference for information on poor family groups needed by the government so that the assistance provided is right on target.

Keywords: *Clustering; Data Mining; K-Medoids; Knowledge Discovery Database; Poverty.*

PENDAHULUAN

Kemiskinan menjadi isu yang terjadi di hampir seluruh negara di dunia. Kemiskinan dapat diartikan sebagai kondisi dimana individu tidak mampu memenuhi kebutuhan dasar yang minimal untuk hidup yang layak, termasuk kebutuhan makanan dan non-makanan [1]. Berdasarkan laporan Badan Pusat Statistik (BPS) mencatat pada tahun 2021 bulan Maret, angka kemiskinan di Indonesia mencapai sebanyak 10,14% terdapat penduduk miskin 27,54 juta orang [2].

Masalah kemiskinan juga dialami oleh Desa Bina Baru. Desa Bina Baru dihadapkan pada permasalahan kemiskinan yang signifikan di daerah tersebut. Berdasarkan observasi dan wawancara kepada pihak desa, kemiskinan yang terjadi di Desa Bina Baru terjadi karena pembangunan infrastruktur dan pemberian kebijakan yang kurang tepat. Di Desa Bina Baru terdapat 4 dusun yaitu Dusun Sukajadi, Dusun Suka Damai, Dusun Suka Maju, dan Dusun Suka Baru. Salah satu langkah untuk mengurangi angka kemiskinan adalah melalui pengelompokan kemiskinan untuk mengetahui jumlah kemiskinan di suatu wilayah, sehingga pemerintah daerah dapat mengambil kebijakan yang lebih akurat dan sesuai untuk mengurangi angka kemiskinan.

Metode *clustering* K-Medoids merupakan salah satu teknik pengolahan data yang termasuk dalam data mining. Tujuannya adalah untuk mengubah data yang besar menjadi informasi penting yang disebut *Knowledge Discovery Database* (KDD). Data dapat dianalisis dengan teknik dalam data mining yaitu untuk mendapatkan koneksi yang jelas dan mencapai kesimpulan yang sebelumnya tidak diketahui secara mudah untuk dipahami dan bermanfaat bagi pemilik data [3].

Clustering merupakan teknik data mining yang bertujuan untuk mengelompokkan objek-objek data menjadi beberapa *cluster*. *Cluster* adalah sekelompok atau sekumpulan objek-objek data yang mempunyai kemiripan satu sama lain di dalam satu *cluster*, tetapi memiliki perbedaan yang signifikan dengan objek-objek data di *cluster* yang lain. Melalui proses pengelompokan, objek-objek data akan dikelompokkan ke dalam satu atau beberapa *cluster*, sehingga objek-objek data yang berada dalam satu *cluster* akan memiliki tingkat kesamaan yang signifikan di antara satu dengan yang lainnya [4].

Metode *clustering* K-Medoids merupakan salah satu algoritme dalam *clustering* yang memiliki kemiripan dengan algoritme K-Means. Namun, terdapat perbedaan antara kedua algoritme tersebut, yaitu K-Medoids atau *Partitioning Around Medoids* (PAM) menggunakan objek sebagai perwakilan (*medoid*) untuk menjadi pusat *cluster* pada setiap *cluster*, sedangkan K-Means menggunakan nilai rata-rata (*mean*) sebagai pusat *cluster*. Metode K-Medoids lebih efektif pada *dataset* yang ukurannya kecil dan memiliki kelebihan untuk mengatasi kelemahan pada algoritme K-Means yang sensitif terhadap *noise* dan *outlier*, dimana objek dengan nilai besar dapat menyimpang dari distribusi data [5]. Penelitian ini menggunakan K-Medoids untuk pengelompokan keluarga miskin di Desa Bina Baru. Hasil dari penelitian ini mampu memberikan pengaruh dalam upaya penurunan angka kemiskinan di Desa Bina Baru, Kampar Kiri Tengah, Kampar, Riau.

PENELITIAN TERKAIT

Penelitian *clustering* dengan algoritme K-Medoids telah banyak diteliti seperti penelitian untuk mengetahui tingkat kemiskinan tinggi dan tingkat kemiskinan rendah di Indonesia [6].

Penelitian lainnya meliputi pengelompokan berdasarkan indikator sosial ekonomi dasar [7]. Literatur mengenai penduduk kurang mampu di Desa Sambirejo Timur [8] menjadi dasar teori keluarga miskin. Selain itu, Pengelompokan data menggunakan Algoritme K-Means *clustering* dan Algoritme K-Medoids [9][16] dan penentuan jumlah *cluster* didasarkan dengan menemukan 2 *cluster* berdasarkan nilai tingkat kemiskinan di Jawa Timur [10].

Penelitian dengan algoritme K-Medoids menggunakan dataset Penduduk Miskin Tahun 2020 memiliki *cluster* optimal sebanyak 2 *cluster* [11]. Penelitian K-Medoids *clustering* didapatkan hasil analisis kecamatan prioritas yang memerlukan bantuan pendidikan dan lapangan pekerjaan [12]. Algoritme K-Medoids dalam penelitian yang telah dilakukan mendapatkan 2 *cluster*, dibandingkan dengan algoritme CLARA mendapatkan 3 *cluster* [13] disesuaikan dengan penentu keputusan dalam mengelompokkan kategori studi kasus. Penelitian metode K-Medoids berdasarkan hasil *cluster* dapat memberikan pertimbangan dari kriteria yang digunakan yaitu Pekerjaan Orang Tua, Status Kepemilikan Rumah dan Penghasilan Orang Tua [14]. Penelitian lainnya ialah mengevaluasi *cluster* menggunakan *Davies-Bouldin Index* (DBI) [15].

METODE PENELITIAN

Proses KDD

Knowledge Discovery in Database Process (KDD) adalah salah satu metode yang dapat digunakan dalam melakukan data mining. KDD merupakan proses dari menggunakan data mining untuk mencari informasi-informasi yang berharga, pola yang ada di dalam data, yang melibatkan algoritme untuk mengidentifikasi pola pada data [20]. Proses KDD dapat melibatkan iterasi yang signifikan dan dapat terjadi iterasi disetiap antar langkah prosesnya. Alur langkah dasar pada proses KDD diilustrasikan pada Gambar 1. Selain itu, hasil dari proses KDD tersebut akan dinilai berdasarkan keakuratan data model yang telah diperoleh melalui perhitungan, dimana dalam penelitian ini menggunakan algoritme K-Medoids.

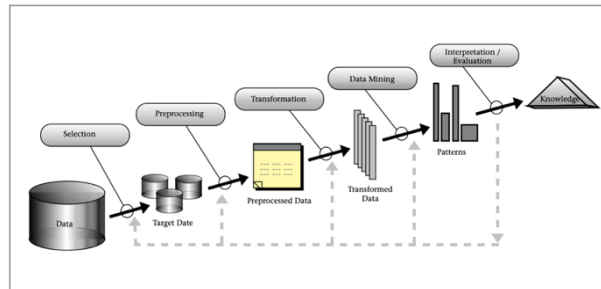
Data Mining adalah suatu istilah yang digunakan untuk menggambarkan proses penggalian nilai dari *database* yang besar. Hal ini dilakukan dengan teknik-teknik otomatisasi untuk mengeksplorasi dan membawa ke permukaan relasi-relasi yang kompleks pada kumpulan data yang sangat besar, seperti data tabulasi pada teknologi manajemen basis data relasional atau domain data spasial, berbasis *text*, dan multimedia. Data Mining terdiri dari beberapa tahapan, yaitu *cleaning and integration*, *selection and transformation*, data mining, *evaluation and presentation*, dan *knowledge*. Dalam tahap-tahap tersebut, dilakukan proses pembersihan data, penggabungan data, seleksi data yang relevan, transformasi data ke format yang sesuai, serta evaluasi dan presentasi hasil data mining untuk mendapatkan pengetahuan yang berharga dan tersembunyi dari data [17].

Data model tersebut akan digunakan untuk mengelola data penduduk terkait kemiskinan di wilayah yang diteliti. Gambar 2 merupakan tahapan kerangka kerja yang digunakan untuk melaksanakan penelitian ini.

Clustering

Clustering adalah sebuah metode dalam data mining yang digunakan untuk mengelompokkan objek-objek berdasarkan karakteristik yang serupa ke dalam beberapa *cluster* [21]. Tujuannya adalah untuk memisahkan data ke dalam kelompok dengan karakteristik yang serupa dan memberikan label pada setiap kelompok. Proses pengelompokan dalam *clustering* label dari setiap data tidak diketahui sebelumnya. *Clustering* memungkinkan objek-objek dalam satu *cluster* memiliki kesamaan yang signifikan dan dapat digunakan untuk analisis lebih lanjut

dalam data mining. Selain itu, *clustering* juga mampu mengatasi kelemahan algoritme K-Means yang rentan terhadap gangguan data (*noise*) dan data yang terpencil (*outlier*) [18].



Gambar 1 Langkah-langkah Proses KDD [20]

Algoritme K-Medoids

Metode K-Medoids merupakan sebuah algoritme pada *clustering* yang sangat efektif digunakan pada *dataset* yang berukuran kecil. Langkah awal dalam metode ini adalah dengan mencari titik representatif pada sebuah *dataset* dengan menghitung jarak antara titik pada setiap kelompok dari semua kemungkinan kombinasi *medoids*. Tujuannya adalah agar jarak pada titik dalam suatu kelompok menjadi kecil, sedangkan jarak pada kelompok menjadi besar. Keuntungan dari algoritme K-Medoids adalah kemampuannya dalam mengatasi kelemahan algoritme K-Means, yaitu sensitif terhadap *noise* dan *outlier*. Dalam K-Medoids, objek dengan nilai yang besar dapat diatasi dan kemungkinan tidak menyimpang dari distribusi data [19].

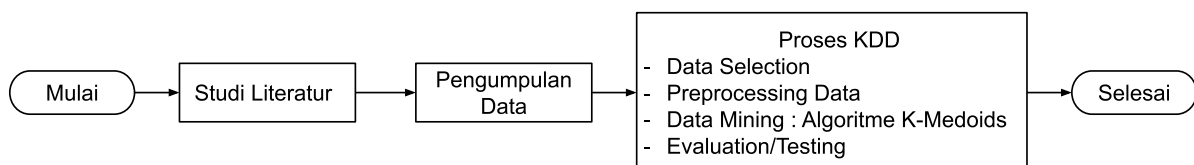
Langkah-langkah dalam algoritme K-Medoids adalah sebagai berikut:

- Langkah pertama adalah melakukan inisialisasi sebanyak k *cluster*.
- Setiap data atau objek akan ditempatkan ke dalam *cluster* terdekat menggunakan rumus *Euclidean Distance* untuk mengukur jarak antara objek-objek tersebut (Persamaan 1).

$$d_{ij} = \sqrt{\sum_{a=1}^p (x_{ia} - x_{ja})^2} \quad (1)$$

dimana $i = 1, \dots, n; j = 1, \dots, n$ dan p adalah banyak variabel, serta V adalah matrik varian kovarian.

- Setelah itu, objek pada setiap *cluster* akan dipilih secara acak sebagai kandidat *medoid bar*.
- Kemudian, jarak antara setiap objek pada masing-masing *cluster* dengan kandidat *medoid* baru dihitung.
- Hitung total simpangan (S) dengan menghitung nilai total *distance* baru dan total *distance* lama. Jika $S < 0$, maka tukar objek dengan data *cluster* untuk membentuk sekumpulan k objek baru sebagai *medoid*.



Gambar 2 Tahapan Penelitian

Studi Literatur

Pada penelitian ini dilakukan penelusuran literatur yang mencakup buku referensi, artikel, dan penelitian sebelumnya yang terkait dengan permasalahan penelitian. Tujuan dari studi literatur ini adalah untuk memperoleh pemahaman teoritis yang mendalam tentang permasalahan yang akan diteliti, seperti teknik data mining, proses *clustering* menggunakan algoritme K-Medoids, *tool* data mining yang akan digunakan, dan data populasi miskin di Desa Bina Baru Kecamatan Kampar Kiri Tengah. Tahap ini penelitian dilakukan untuk memahami masalah yang diteliti secara benar dan menjalankan pembahasan secara tepat.

Pengumpulan Data

Data yang digunakan dalam penelitian ini merupakan data sekunder. Data didapatkan dari *website* SDGS. Data yang digunakan merupakan data Desa Bina Baru, Kecamatan Kampar Kiri Tengah Tahun 2022. Data diambil dari kuesioner keluarga dan kuesioner individu.

HASIL DAN PEMBAHASAN

Hasil dan pembahasan berisikan hasil penelitian yang diperoleh disajikan dalam bentuk uraian teoritik baik secara kualitatif maupun kuantitatif. Hasil penelitian dapat ditampilkan dalam bentuk grafik, gambar, tabel dan lainnya yang dapat membuat pembaca dapat lebih mudah memahami.

Selection Data

Penelitian ini menggunakan data SDGS Desa Bina Baru Tahun 2022, dimana data tersebut terakhir diperbarui pada bulan September 2022. Data yang digunakan diambil dari kuesioner keluarga dan kuesioner individu. Berdasarkan data tersebut, didapat 1.005 data keluarga dengan 64 atribut dan 1.005 data individu dengan 84 atribut.

Preprocessing Data

Data *preprocessing* melibatkan proses pembersihan dan persiapan data dengan tujuan mengatasi masalah konsistensi, data yang tidak lengkap, dan data yang berulang di data awal. Selain itu, Data *preprocessing* juga mencakup pengubahan status barang dari format awal "true/false" menjadi "1/0", yang kemudian dapat digunakan dalam proses asosiasi.

Cleaning Data

Tujuan pada tahap *cleaning*, yaitu untuk menghilangkan *noise* dan data yang tidak relevan, membersihkan data yang kosong, data yang kurang jelas, dan lengkap agar tidak mempengaruhi perhitungan. Data yang tidak relevan itu lebih baik dibuang. Data yang awalnya berjumlah 1.005 pada Tabel 1 merupakan data yang akan melalui tahap *cleaning*.

Pada tahap data *cleaning* data yang kosong atau *missing value* akan dihapus. Setelah melalui tahap data *cleaning*, terdapat 5 data yang dihapus, sehingga data yang digunakan berjumlah 1.000 data. Data yang telah melalui tahap *cleaning* dapat dilihat pada Tabel 2.

Transformation Data

Pada tahap ini, variabel data harus ditentukan terlebih dahulu. Variabel tersebut yang nantinya digunakan untuk mendapatkan informasi survey. Variabel atau atribut yang digunakan ada pada Tabel 3, yang nanti akan digunakan untuk mengisi inisial nilai pada Tabel 4.

Tabel 1 Data kosong kuisioner

NO	PENGHASILAN	PENGELUARAN	ANGGOTA KELUARGA	SUMBER PENGHASILAN	DINDING RUMAH	LANTAI RUMAH	STATUS RUMAH	ATAP RUMAH	NAMA	TEMPAT BUANG AIR BESAR	MCK	SUMBER AIR MINUM	PENERANGAN UTAMA	BAHAN BAKAR MEMASAK	PEMBUANGAN SAMPAH	PEMBUANGAN LIMBAH CAIR	KONDISI RUMAH
1	2500000	2000000	2	BURUH TANI	SEMEN	SEMEN	SENDIRI	GENTENG	SUDARSONO	JAMBAN SENDIRI	SENDIRI	AIR ISI ULANG	LISTRIK PLN	GAS/LPG	DIBAKAR	LUBANG DI TANAH	TIDAK KUMUH
2	8000000	4000000	3	PEMILIK KEBUN SAWIT	SEMEN	KERAMIK	SENDIRI	GENTENG	SUKADI	JAMBAN SENDIRI	SENDIRI	AIR ISI ULANG	LISTRIK PLN	GAS/LPG	DIBAKAR	LUBANG DI TANAH	TIDAK KUMUH
3	1500000	1000000	2	BURUH TANI	SEMEN	SEMEN	SENDIRI	GENTENG	IDIL FITRI	JAMBAN SENDIRI	SENDIRI	AIR ISI ULANG	LISTRIK PLN	MINYAK TANAH	DIBAKAR	LUBANG DI TANAH	TIDAK KUMUH
4	3500000	3000000	6	BURUH TANI	SEMEN	KERAMIK	SENDIRI	GENTENG	A'AH	JAMBAN SENDIRI	SENDIRI	AIR ISI ULANG	LISTRIK PLN	GAS/LPG	DIBAKAR	LUBANG DI TANAH	TIDAK KUMUH
5	3000000	2500000	3	BURUH TANI	SEMEN	SEMEN	BEBAS SEWA	GENTENG	AMAD KURNIAWAN	JAMBAN SENDIRI	SENDIRI	MATA AIR/ SUMUR	LISTRIK PLN	GAS/LPG	DIBAKAR	LUBANG DI TANAH	TIDAK KUMUH
6	3500000	3000000	5	BURUH TANI	SEMEN	KERAMIK	SENDIRI	GENTENG	MUKADIR	JAMBAN SENDIRI	SENDIRI	AIR ISI ULANG	LISTRIK PLN	GAS/LPG	TEMPAT SAMPAH	TANGKI	TIDAK KUMUH
7	4500000	3000000	4	PEDAGANG	SEMEN	SEMEN	SENDIRI	GENTENG	MUSENI	JAMBAN SENDIRI	SENDIRI	AIR ISI ULANG	LISTRIK PLN	GAS/LPG	DIBAKAR	LUBANG DI TANAH	TIDAK KUMUH
8	3000000	2500000	2	BURUH TANI	SEMEN	SEMEN	SENDIRI	GENTENG	SLAMET ARIFIN	JAMBAN SENDIRI	SENDIRI	AIR ISI ULANG	LISTRIK PLN	GAS/LPG	TEMPAT SAMPAH	LUBANG DI TANAH	TIDAK KUMUH
9	3500000	3000000	5	BURUH TANI	SEMEN	SEMEN	SENDIRI	GENTENG	MIFTAHUL HUDA	JAMBAN SENDIRI	SENDIRI	AIR ISI ULANG	LISTRIK PLN	MINYAK TANAH	TEMPAT SAMPAH	TANGKI	TIDAK KUMUH
10	4000000	3000000	5	PEDAGANG	SEMEN	SEMEN	SENDIRI	GENTENG	NARI	JAMBAN SENDIRI	SENDIRI	AIR ISI ULANG	LISTRIK PLN	GAS/LPG	TEMPAT SAMPAH	TANGKI	TIDAK KUMUH
11	4500000	3000000	3	KARYAWAN	SEMEN	SEMEN	SENDIRI	GENTENG	NURYANTO	JAMBAN SENDIRI	SENDIRI	AIR ISI ULANG	LISTRIK PLN	GAS/LPG	DIBAKAR	LUBANG DI TANAH	TIDAK KUMUH
12	2500000	1500000	1	BURUH TANI	SEMEN	KERAMIK	SENDIRI	GENTENG	SUMARJO	JAMBAN SENDIRI	SENDIRI	AIR ISI ULANG	LISTRIK PLN	GAS/LPG	DIBAKAR	LUBANG DI TANAH	TIDAK KUMUH
13	3500000	2500000	3	KARYAWAN	SEMEN	KERAMIK	SENDIRI	GENTENG	A.ZAMRUD SITORUS	JAMBAN SENDIRI	SENDIRI	AIR ISI ULANG	LISTRIK PLN	GAS/LPG	TEMPAT SAMPAH	LUBANG DI TANAH	TIDAK KUMUH
14	5000000	3000000	2	PEMILIK KEBUN SAWIT	SEMEN	SEMEN	SENDIRI	GENTENG	MUHAMMAD ABDHAL	JAMBAN SENDIRI	SENDIRI	AIR ISI ULANG	LISTRIK PLN	GAS/LPG	TEMPAT SAMPAH	LUBANG DI TANAH	TIDAK KUMUH
15	2500000	1500000	5	BURUH TANI	SEMEN	SEMEN	SENDIRI	SENG	YUSRIL MANTO	JAMBAN SENDIRI	SENDIRI	MATA AIR/ SUMUR	LISTRIK PLN	GAS/LPG	TEMPAT SAMPAH	TANGKI	TIDAK KUMUH
16	80000000	60000000	4	PEMILIK KEBUN SAWIT	SEMEN	SEMEN	SENDIRI	SENG	SUPRIYADI	JAMBAN SENDIRI	SENDIRI	AIR ISI ULANG	LISTRIK PLN	GAS/LPG	TEMPAT SAMPAH	TANGKI	TIDAK KUMUH
17	3000000	2000000	2	PEDAGANG	SEMEN	KERAMIK	SENDIRI	GENTENG	KUATO	JAMBAN SENDIRI	SENDIRI	AIR ISI ULANG	LISTRIK PLN	GAS/LPG	DIBAKAR	LUBANG DI TANAH	TIDAK KUMUH
18	2000000	1500000	3	-	SEMEN	SEMEN	SENDIRI	SENG	DENI SYARIFUDIN	JAMBAN SENDIRI	SENDIRI	MATA AIR/ SUMUR	LISTRIK PLN	MINYAK TANAH	DIBAKAR	LUBANG DI TANAH	TIDAK KUMUH
19	2000000	1000000	6	BURUH TANI	SEMEN	SEMEN	SENDIRI	GENTENG	P. YAHYA HARAHAP	JAMBAN SENDIRI	SENDIRI	AIR ISI ULANG	LISTRIK PLN	GAS/LPG	DIBAKAR	LUBANG DI TANAH	TIDAK KUMUH

Tabel 2 Data kuisioner setelah melalui tahap *cleaning*

NO	PENGHASILAN	PENGELUARAN	ANGGOTA KELUARGA	SUMBER PENGHASILAN	DINDING RUMAH	LANTAI RUMAH	STATUS RUMAH	ATAP RUMAH	NAMA	TEMPAT BUANG AIR BESAR	MCK	SUMBER AIR MINUM	PENERANGAN UTAMA	BAHAN BAKAR MEMASAK	PEMBUANGAN SAMPAH	PEMBUANGAN LIMBAH CAIR	KONDISI RUMAH
1	2500000	2000000	2	BURUH TANI	SEMEN	SEMEN	SENDIRI	GENTENG	SUDARSONO	JAMBAN SENDIRI	SENDIRI	AIR ISI ULANG	LISTRIK PLN	GAS/LPG	DIBAKAR	LUBANG DI TANAH	TIDAK KUMUH
2	8000000	4000000	3	PEMILIK KEBUN SAWIT	SEMEN	KERAMIK	SENDIRI	GENTENG	SUKADI	JAMBAN SENDIRI	SENDIRI	AIR ISI ULANG	LISTRIK PLN	GAS/LPG	DIBAKAR	LUBANG DI TANAH	TIDAK KUMUH
3	1500000	1000000	2	BURUH TANI	SEMEN	SEMEN	SENDIRI	GENTENG	IDIL FITRI	JAMBAN SENDIRI	SENDIRI	AIR ISI ULANG	LISTRIK PLN	MINYAK TANAH	DIBAKAR	LUBANG DI TANAH	TIDAK KUMUH
4	3500000	3000000	6	BURUH TANI	SEMEN	KERAMIK	SENDIRI	GENTENG	A'AH	JAMBAN SENDIRI	SENDIRI	AIR ISI ULANG	LISTRIK PLN	GAS/LPG	DIBAKAR	LUBANG DI TANAH	TIDAK KUMUH
5	3000000	2500000	3	BURUH TANI	SEMEN	SEMEN	BEBAS SEWA	GENTENG	AMAD KURNIAWAN	JAMBAN SENDIRI	SENDIRI	MATA AIR/ SUMUR	LISTRIK PLN	GAS/LPG	DIBAKAR	LUBANG DI TANAH	TIDAK KUMUH
6	3500000	3000000	5	BURUH TANI	SEMEN	KERAMIK	SENDIRI	GENTENG	MUKADIR	JAMBAN SENDIRI	SENDIRI	AIR ISI ULANG	LISTRIK PLN	GAS/LPG	DIBAKAR	TANGKI	TIDAK KUMUH
7	4500000	3000000	4	PEDAGANG	SEMEN	SEMEN	SENDIRI	GENTENG	MUSENI	JAMBAN SENDIRI	SENDIRI	AIR ISI ULANG	LISTRIK PLN	GAS/LPG	DIBAKAR	LUBANG DI TANAH	TIDAK KUMUH
8	3000000	2500000	2	BURUH TANI	SEMEN	SEMEN	SENDIRI	GENTENG	SLAMET ARIFIN	JAMBAN SENDIRI	SENDIRI	AIR ISI ULANG	LISTRIK PLN	GAS/LPG	DIBAKAR	LUBANG DI TANAH	TIDAK KUMUH
9	3500000	3000000	5	BURUH TANI	SEMEN	SEMEN	SENDIRI	GENTENG	MIFTAHUL HUDA	JAMBAN SENDIRI	SENDIRI	AIR ISI ULANG	LISTRIK PLN	MINYAK TANAH	TEMPAT SAMPAH	TANGKI	TIDAK KUMUH
10	4000000	3000000	5	PEDAGANG	SEMEN	SEMEN	SENDIRI	GENTENG	NARI	JAMBAN SENDIRI	SENDIRI	AIR ISI ULANG	LISTRIK PLN	GAS/LPG	TEMPAT SAMPAH	TANGKI	TIDAK KUMUH
11	4500000	3000000	3	KARYAWAN	SEMEN	SEMEN	SENDIRI	GENTENG	NURYANTO	JAMBAN SENDIRI	SENDIRI	AIR ISI ULANG	LISTRIK PLN	GAS/LPG	DIBAKAR	LUBANG DI TANAH	TIDAK KUMUH
12	2500000	1500000	1	BURUH TANI	SEMEN	KERAMIK	SENDIRI	GENTENG	SUMARJO	JAMBAN SENDIRI	SENDIRI	AIR ISI ULANG	LISTRIK PLN	GAS/LPG	DIBAKAR	LUBANG DI TANAH	TIDAK KUMUH
13	3500000	2500000	3	KARYAWAN	SEMEN	KERAMIK	SENDIRI	GENTENG	A.ZAMRUD SITORUS	JAMBAN SENDIRI	SENDIRI	AIR ISI ULANG	LISTRIK PLN	GAS/LPG	TEMPAT SAMPAH	LUBANG DI TANAH	TIDAK KUMUH
14	5000000	3000000	2	PEMILIK KEBUN SAWIT	SEMEN	SEMEN	SENDIRI	GENTENG	MUHAMMAD ABDHAL	JAMBAN SENDIRI	SENDIRI	AIR ISI ULANG	LISTRIK PLN	GAS/LPG	TEMPAT SAMPAH	LUBANG DI TANAH	TIDAK KUMUH
15	2500000	1500000	5	BURUH TANI	SEMEN	SEMEN	SENDIRI	SENG	YUSRIL MANTO	JAMBAN SENDIRI	SENDIRI	MATA AIR/ SUMUR	LISTRIK PLN	GAS/LPG	TEMPAT SAMPAH	TANGKI	TIDAK KUMUH
16	80000000	60000000	4	PEMILIK KEBUN SAWIT	SEMEN	SEMEN	SENDIRI	SENG	SUPRIYADI	JAMBAN SENDIRI	SENDIRI	AIR ISI ULANG	LISTRIK PLN	GAS/LPG	TEMPAT SAMPAH	TANGKI	TIDAK KUMUH
17	3000000	2000000	2	PEDAGANG	SEMEN	KERAMIK	SENDIRI	GENTENG	KUATO	JAMBAN SENDIRI	SENDIRI	AIR ISI ULANG	LISTRIK PLN	GAS/LPG	DIBAKAR	LUBANG DI TANAH	TIDAK KUMUH
18	2000000	1000000	6	BURUH TANI	SEMEN	SEMEN	SENDIRI	GENTENG	P. YAHYA HARAHAP	JAMBAN SENDIRI	SENDIRI	MATA AIR/ SUMUR	LISTRIK PLN	MINYAK TANAH	DIBAKAR	LUBANG DI TANAH	TIDAK KUMUH
19	3500000	2500000	4	KARYAWAN	SEMEN	SEMEN	BEBAS SEWA	SENG	IIN ANDY SAPUTRA	JAMBAN SENDIRI	SENDIRI	AIR ISI ULANG	LISTRIK PLN	GAS/LPG	DIBAKAR	LUBANG DI TANAH	TIDAK KUMUH

Tabel 3 Variabel Informasi Survey

NO	NAMA VARIABEL/ATRIBUT	KETERANGAN
1	Penghasilan keluarga perbulan	Gaji perbulan
2	Pengeluaran keluarga perbulan	Berapa banyak pengeluaran perbulan
3	Anggota keluarga	Berapa orang menjadi tanggungan oleh Kepala keluarga
4	Sumber penghasilan	1. Buruh tani. 2. Pemilik Kebun Sawit. 3. Pedagang. 4. Karyawan. 5. PNS. 6. Guru Bantu. 7. Peternakan. 8. Bengkel. 9. Walet. 10. Satpam.
5	Jenis dinding rumah	1. Semen. 2. Kayu.
6	Jenis Lantai Rumah	1. Semen. 2. Keramik. 3. Kayu. 4. Ubin. 5. Marmer.
7	Status Penguasaan Rumah	1. Sendiri. 2. Bebas Sewa. 3. Kontrak. 4. Dinas
8	Jenis Atap Rumah	1. Genteng. 2. Seng. 3. Ijuk.
9	Fasilitas Tempat Buang Air Besar	1. Jamban Sendiri. 2. Jamban umum.
10	Fasilitas MCK	1. Sendiri. 2. MCK umum.
11	Sumber Air Minum	1. Air isi ulang. 2. Mata air/sumur.
12	Sumber Penerangan Utama	1. Listrik PLN. 2. Lilin.
13	Jenis Bahan Bakar untuk Memasak	1. Gas kota/LPG. 2. Minyak Tanah.
14	Tempat Pembuangan Sampah	1. Dibakar. 2. Tempat Sampah.
15	Tempat Pembuangan Limbah Cair	1. Lubang Ditanah. 2. Tangki.
16	Secara Keseluruhan Kondisi Rumah	1. Kumuh. 2. Tidak Kumuh.

Normalisasi Data

Normalisasi yaitu suatu proses transformasi yang mengubah atribut numerik skala yang lebih kecil, seperti 0,1 sampai 1,0. Normalisasikan dilakukan untuk menghasilkan keseimbangan antara nilai yang rendah dan nilai tinggi dalam data. Hasil normalisasi data dapat dilihat pada Tabel 5 yang merupakan data sampel untuk contoh perhitungan K-Medoids pada penelitian ini.

K-Medoids.

Dalam penelitian ini, pelaksanaan K-Medoids mengacu pada studi kasus dengan inisialisasi pusat *cluster* sebanyak 3 (tiga) yaitu *Cluster 1* (kaya), *Cluster 2* (menengah), dan *Cluster 3* (miskin) yang dipilih secara acak untuk setiap *medoid* (Tabel 6).

Dalam penelitian ini, digunakan rumus *Euclidean Distance* untuk menentukan *cluster* yang paling dekat dengan setiap data (objek) yang ada :

$$\begin{aligned}
 Cost\ 1 &= \sqrt{((0,019 - 0,019)^2 + (0,025 - 0,02)^2 + (0,167 - 0,167)^2 + (0,000 - 0,000)^2 + (0,000 - 0,000)^2 + (0,000 - 0,000)^2 + (0,000 - 0,000)^2 + (0,000 - 0,000)^2 + (0,000 - 0,000)^2 + (0,000 - 0,000)^2)} = 0 \\
 Cost\ 2 &= \sqrt{((0,089 - 0,019)^2 + (0,059 - 0,025)^2 + (0,333 - 0,167)^2 + (-1,000 - 0,000)^2 + (0,000 - 0,000)^2 + (0,250 - 0,000)^2 + (0,000 - 0,000)^2 + (0,000 - 0,000)^2 + (0,000 - 0,000)^2 + (0,000 - 0,000)^2)} = 1,047022 \\
 Cost\ 3 &= \sqrt{((0,006 - 0,019)^2 + (0,008 - 0,025)^2 + (0,167 - 0,167)^2 + (0,000 - 0,000)^2 + (0,000 - 0,000)^2 + (0,000 - 0,000)^2 + (0,000 - 0,000)^2 + (0,000 - 0,000)^2 + (0,000 - 0,000)^2 + (0,000 - 0,000)^2)} = 1,000221
 \end{aligned}$$

Penghitungan tersebut terus berlanjut hingga mencapai data ke-20. Tabel 7 menunjukkan hasil perhitungan setelah proses tersebut dilakukan. Hasil perhitungan dengan algoritme K-Medoids secara keseluruhan pada iterasi 1 dapat ditemukan pada Tabel 8.

Tabel 4 Tabel Inisialisasi Data

NO	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	2500000	2000000	2	1	1	1	1	1	1	1	1	1	1	1	1	2
2	8000000	4000000	3	2	1	2	1	1	1	1	1	1	1	1	1	2
3	1500000	1000000	2	1	1	1	1	1	1	1	1	1	2	1	1	2
4	3500000	3000000	6	1	1	2	1	1	1	1	1	1	1	1	1	2
5	3000000	2500000	3	1	1	1	2	1	1	1	2	1	1	1	1	2
6	3500000	3000000	5	1	1	2	1	1	1	1	1	1	1	1	2	2
7	4500000	3000000	4	3	1	1	1	1	1	1	1	1	1	1	1	2
8	3000000	2500000	2	1	1	1	1	1	1	1	1	1	1	1	1	2
9	3500000	3000000	5	1	1	1	1	1	1	1	1	1	2	2	2	2
10	4000000	3000000	5	3	1	1	1	1	1	1	1	1	1	2	2	2
11	4500000	3000000	3	4	1	1	1	1	1	1	1	1	1	1	1	2
12	2500000	1500000	1	1	1	2	1	1	1	1	1	1	1	1	1	2
13	3500000	2500000	3	4	1	2	1	1	1	1	1	1	1	2	1	2
14	5000000	3000000	2	2	1	1	1	1	1	1	1	1	1	2	1	2
15	2500000	1500000	5	1	1	1	1	2	1	1	2	1	1	2	2	2
16	8000000	6000000	4	2	1	1	1	2	1	1	1	1	1	2	2	2
17	3000000	2000000	2	3	1	2	1	1	1	1	1	1	1	1	1	2
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
997	2000000	3000000	7	1	1	3	1	1	1	1	2	1	2	1	1	1
998	3000000	2750000	3	1	1	1	1	1	1	1	1	1	1	1	1	2
999	3500000	3000000	3	1	1	1	1	1	1	1	1	1	1	1	1	2
1000	4000000	3500000	3	1	1	1	1	1	1	1	1	1	1	1	1	2

Tabel 5 Proses normalisasi data

NO	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	0,019	0,025	0,167	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	1,000
2	0,089	0,059	0,333	-1,000	0,000	0,250	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	1,000
3	0,006	0,008	0,167	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	1,000	0,000	0,000	1,000
4	0,032	0,042	0,833	0,000	0,000	0,250	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	1,000
5	0,025	0,034	0,333	0,000	0,000	0,000	0,333	0,000	0,000	0,000	1,000	0,000	0,000	0,000	0,000	1,000
6	0,032	0,042	0,667	0,000	0,000	0,250	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	1,000	1,000
7	0,044	0,042	0,500	-2,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	1,000
8	0,025	0,034	0,167	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	1,000
9	0,032	0,042	0,667	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	1,000	1,000	1,000	1,000
10	0,038	0,042	0,667	-2,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	1,000	1,000	1,000
11	0,044	0,042	0,333	-3,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	1,000
12	0,019	0,017	0,000	0,000	0,000	0,250	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	1,000
13	0,032	0,034	0,333	-3,000	0,000	0,250	0,000	0,000	0,000	0,000	0,000	0,000	0,000	1,000	0,000	1,000
14	0,051	0,042	0,167	-1,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	1,000	0,000	1,000
15	0,019	0,017	0,667	0,000	0,000	0,000	0,000	1,000	0,000	0,000	1,000	0,000	0,000	1,000	1,000	1,000
16	1,000	1,000	0,500	-1,000	0,000	0,000	0,000	1,000	0,000	0,000	0,000	0,000	0,000	1,000	1,000	1,000
17	0,025	0,025	0,167	-2,000	0,000	0,250	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	1,000
18	0,013	0,008	0,833	0,000	0,000	0,000	0,000	0,000	0,000	0,000	1,000	0,000	1,000	0,000	0,000	1,000
19	0,032	0,034	0,500	-3,000	0,000	0,000	0,333	1,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	1,000
20	0,025	0,034	0,500	0,000	0,000	0,000	0,000	1,000	0,000	0,000	0,000	0,000	0,000	1,000	1,000	1,000

Tabel 6 Tampilan Pusat Cluster Medoid 1.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	0,019	0,025	0,167	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	1,000
2	0,089	0,059	0,333	-1,000	0,000	0,250	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	1,000
3	0,006	0,008	0,167	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	1,000	0,000	0,000	1,000

Tabel 7 Hasil Perhitungan

Cost 1	Cost 2	Cost 3
0	1,047022	1,000221
1,047022	0	1,448996
1,000221	1,448996	0
0,712311	1,11961	1,228298
1,067239	1,47589	1,462835
1,145837	1,45418	1,521273
2,027815	1,045238	2,261345
0,01052	1,046384	1,000498
1,802898	2,043805	1,50059
2,500129	1,782262	2,692978
3,00478	2,016121	3,167073
0,30058	1,057224	1,044274
3,176556	2,236935	3,330399
1,414667	1,044989	1,732943
2,06157	2,27601	2,291338
2,454325	2,192425	2,661157
2,015574	1,016323	2,250143
1,563575	1,822309	1,201867
3,197257	2,281506	3,350149
1,763866	2,023591	2,027833

Tabel 8 Hasil Perhitungan Algoritme *K-Medoids* Iterasi ke-1

Cost 1	Cost 2	Cost 3	Kedekatan	Cluster
0	1,047022	1,000221	0	1
1,047022	0	1,448996	0	2
1,000221	1,448996	0	0	3
0,712311	1,11961	1,228298	0,71231113	1
1,067239	1,47589	1,462835	1,06723922	1
1,145837	1,45418	1,521273	1,14583712	2
2,027815	1,045238	2,261345	1,04523828	2
0,01052	1,046384	1,000498	0,01052018	1
1,802898	2,043805	1,50059	1,50059015	3
2,500129	1,782262	2,692978	1,78226184	2
3,00478	2,016121	3,167073	2,01612135	2
0,30058	1,057224	1,044274	0,3005801	1
3,176556	2,236935	3,330399	2,23693545	2
1,414667	1,044989	1,732943	1,04498915	2
2,06157	2,27601	2,291338	2,06156994	1
2,454325	2,192425	2,661157	2,19242492	2
2,015574	1,016323	2,250143	1,01632348	2
1,563575	1,822309	1,201867	1,20186709	3
3,197257	2,281506	3,350149	2,28150589	2
1,763866	2,023591	2,027833	1,76386558	1
Jumlah Kedekatan			23,3801809	

Tahap selanjutnya adalah tahap ketika objek baru dijadikan sebagai kandidat *medoid* yang dipilih secara acak dari setiap *cluster*. Pusat *cluster* yang dibentuk ditampilkan pada Tabel 9.

Tabel 9 Pusat *Cluster Medoids II*

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
18	0,013	0,008	0,833	0,000	0,000	0,000	0,000	0,000	0,000	0,000	1,000	0,000	1,000	0,000	0,000	1,000
19	0,032	0,034	0,500	-3,000	0,000	0,000	0,333	1,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	1,000
20	0,025	0,034	0,500	0,000	0,000	0,000	0,000	1,000	0,000	0,000	0,000	0,000	0,000	1,000	1,000	1,000

Melakukan perhitungan jarak antara setiap objek yang terdapat pada masing-masing *cluster* dengan kandidat *medoid* yang baru dipilih secara acak. Langkah-langkah yang dilakukan pada iterasi kedua masih sama dengan langkah-langkah pada iterasi sebelumnya. Hasil perhitungan pada iterasi kedua dapat dilihat pada Tabel 10. Perhitungan dilakukan hingga hasil seluruh data dapat dilihat pada Tabel 11.

Perhitungan untuk mendapatkan total simpangan (S) ditunjukkan pada Persamaan 2. Jika nilai S kurang dari nol, maka dilakukan pertukaran objek dengan data *cluster* untuk membentuk sekumpulan *k* objek baru sebagai *medoid*. Total simpangan memiliki nilai lebih dari 0, maka iterasi yang digunakan adalah iterasi pertama, begitu juga jika nilai lebih kecil dari 0, maka yang digunakan ialah iterasi selain iterasi kedua.

$$S = \text{total jarak terdekat baru} - \text{total jarak terdekat lama} \quad (2)$$

Hasil iterasi pertama dan kedua dihitung nilai simpangan (S) berdasarkan Persamaan 2 didapatkan hasil :

$$S = 23,3801809 - 25,3117055 = -1,931524677$$

Hasil simpangan yang didapat kurang dari 0, sehingga iterasi ke 2 yang digunakan sebagai acuan yaitu *cluster* 1 sebanyak 8 orang, *Cluster* 2 sebanyak 6 orang, *Cluster* 3 sebanyak 6 orang.

Tabel 10 Hasil Perhitungan 2

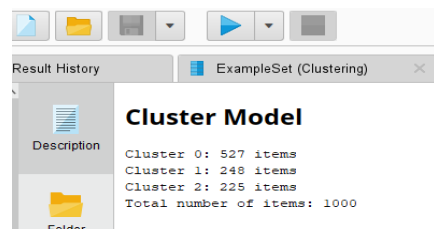
Cost 1	Cost 2	Cost 3
1,563575	3,197257	1,763866
1,822309	2,281506	2,023591
1,201867	3,350149	2,027833
1,436659	3,206991	1,781494
1,167008	3,320816	2,034426
1,758342	3,346858	1,445818
2,472497	1,453046	2,645833
1,563726	3,197227	1,763834
1,740479	3,62477	1,424039
2,833646	2,034453	2,242322
3,35442	1,067296	3,468171
1,660438	3,22863	1,820116
3,509059	1,483708	3,330213
2,108795	2,494525	1,764036
2,00696	3,484154	1,013953
2,840637	2,825372	1,698208
2,550958	1,511566	2,678373
0	3,496172	2,260953
3,496172	0	3,333339
2,260953	3,333339	0

Tabel 11 Hasil Perhitungan Algoritme *K-Medoids* Iterasi ke-2

Cost 1	Cost 2	Cost 3	Kedekatan	Cluster
1,563575	3,197257	1,763866	1,56357506	1
1,822309	2,281506	2,023591	1,82230911	1
1,201867	3,350149	2,027833	1,20186709	1
1,436659	3,206991	1,781494	1,43665945	1
1,167008	3,320816	2,034426	1,16700767	1
1,758342	3,346858	1,445818	1,44581757	3
2,472497	1,453046	2,645833	1,45304575	2
1,563726	3,197227	1,763834	1,56372639	1
1,740479	3,62477	1,424039	1,42403948	3
2,833646	2,034453	2,242322	2,03445314	2
3,35442	1,067296	3,468171	1,06729552	2
1,660438	3,22863	1,820116	1,66043823	1
3,509059	1,483708	3,330213	1,48370782	2
2,108795	2,494525	1,764036	1,7640359	3
2,00696	3,484154	1,013953	1,01395281	3
2,840637	2,825372	1,698208	1,69820817	3
2,550958	1,511566	2,678373	1,51156637	2
0	3,496172	2,260953	0	1
3,496172	0	3,333339	0	2
2,260953	3,333339	0	0	3
Jumlah Kedekatan			25,3117055	

Hasil Pengujian

Setelah persiapan data selesai dan pengujian manual telah didapatkan, langkah selanjutnya yaitu pemodelan menggunakan algoritme *K-Medoids*. Aplikasi *Rapid Miner* digunakan untuk melakukan pemodelan tersebut. Dalam studi kasus ini, terdapat 1.000 data penduduk di Desa Bina Baru yang dikelompokkan menjadi tiga *cluster* berdasarkan tingkat kemiskinan yaitu kaya, menengah, dan miskin.

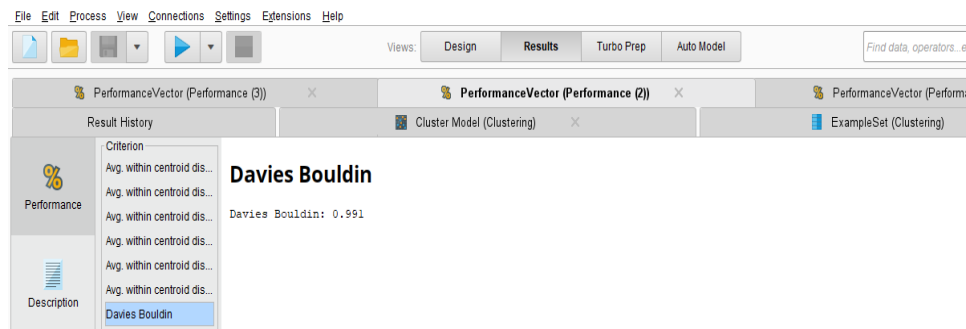


Gambar 3 Rapid Miner menampilkan kluster model

Gambar 4 merupakan hasil pengelompokan *cluster* yang didapat dengan pengujian *Rapid Miner*. Kelompok pada *cluster* pertama (kaya) berjumlah 527 *items*, dan kelompok pada *cluster* kedua (menengah) berjumlah 248 *items*, dan kelompok pada *cluster* ketiga (miskin) berjumlah 225 *items*.

Evaluasi

Setelah dilakukan pengujian dengan menggunakan data penduduk di Desa Bina Baru, didapatkan nilai DBI sebesar 0,991. Semakin kecil nilai DBI, maka evaluasi terhadap *cluster* dapat dikatakan semakin baik, dan dalam hal ini evaluasi terhadap *cluster* dikatakan baik karena mendekati nol.



Gambar 4 Nilai DBI (*Davies Bouldin Index*) pada Rapid Miner

Dilihat dari Gambar 5 hasil, evaluasi *cluster* pada penelitian ini membentuk tiga cluster dengan nilai 0,991.

KESIMPULAN

Penelitian ini adalah mengelompokkan keluarga miskin di Desa Bina Baru menggunakan algoritme K-Medoids, jumlah *cluster* yang terbentuk yaitu 3 cluster. Pengujian data menggunakan *tools Rapid Miner* pada 1.000 data penduduk yang ada di Desa Bina Baru. Hasil perbandingan perhitungan untuk *cluster* 1 (kaya) terdiri dari 527 penduduk, *cluster* 2 (menengah) terdiri dari 248 penduduk, dan *cluster* 3 (miskin) terdiri dari 225 penduduk. Berdasarkan hasil pengujian nilai *cluster* menggunakan DBI diperoleh nilai yaitu 0,991, sehingga dapat dikatakan bahwa pengujian *cluster* cukup baik, karena semakin kecil nilai DBI dan tidak bernilai negatif, semakin baik pengujian sebuah *cluster*. Penelitian ini memiliki dampak mengurangi tingkat kemiskinan di Desa Bina Baru, Kampar Kiri Tengah, Kampar, Riau.

DAFTAR PUSTAKA

1. Kompas.com, "Kemiskinan: Definisi, Jenis, dan Faktor Penyebabnya," kompas.com/skola/read/2020/11/24/172143169/kemiskinan-definisi-jenis-dan-faktor-penyebabnya?page=all, Jan. 21, 2021.
2. kemenkopmk.com, "Program Perlindungan Sosial Dirancang Kurangi Angka Kemiskinan," kemenkopmk.go.id/, Jul. 19, 2021.
3. Sampoerna University, "Pengertian Data Mining, Fungsi, Metode dan Contoh Penerapan," sampoernauniversity.ac.id/id/data-mining/, Jun. 22, 2022.
4. Binus University, "CLUSTERING," <https://socs.binus.ac.id/2017/03/09/Clustering/>, Mar. 09, 2017.
5. N. R. Aeni, A. Nilogiri, and R. Umilasari, "Algoritme Partitioning Around Medoids Dalam Mengelompokkan Provinsi Di Indonesia Berdasarkan Indeks Kinerja Davies Bouldin Pada Kasus Penyakit Hiv."
6. D. Wahyuli, I. Parlina, A. Perdana Windarto, D. Suhendro, A. Wanto, and S. A. Tunas Bangsa Pematangsiantar Jln Jendral Sudirman Blok No, "Prosiding Seminar Nasional Riset Information Science (SENARIS) Mengelompokkan Garis Kemiskinan Menurut Provinsi Menggunakan Algoritme K-Medoids," 2019, [Online]. Available: <http://www.bps.go.id>.
7. E. R. Tirayo, "Pengelompokan Kabupaten/Kota di Sulawesi Tengah Menurut Indikator Sosial Ekonomi Menggunakan K-Medoids Clustering".
8. J. Arianto, "Penerapan Data Mining Untuk Pengelompokan Penduduk Kurang Mampu Desa Sambirejo Timur Dengan Algoritme K-Medoids (Studi Kasus Kantor Kepala Desa Sambirejo Timur)," KOMIK (Konferensi Nasional Teknologi Informasi dan Komputer),

- vol. 3, no. 1, Dec. 2019, doi: 10.30865/komik.v3i1.1660.
9. J. Manurung, P. Sari Ramadhan, and M. Suryanata, "Perbandingan Algoritme K-Means Dan K-Medoids Untuk Pengelompokan Data Masyarakat Miskin Pada Kantor Camat Hatonduhan STMIK Triguna Dharma ** Program Studi Sistem Informasi, STMIK Triguna Dharma *** Program Studi Sistem Informasi, STMIK Triguna Dharma," *Jurnal CyberTech*, vol. 3, no. 9, pp. 1522–1531, 2020, [Online]. Available: <https://ojs.trigunadharma.ac.id/>
 10. F. Alfiah, A. Almadayani, D. al Farizi, and E. Widodo, "Analisis Clustering K-Medoids Berdasarkan Indikator Kemiskinan di Jawa Timur Tahun 2020," *JURNAL ILMIAH SAINS*, vol. 22, no. 1, p. 1, Dec. 2021, doi: 10.35799/jis.v22i1.35911.
 11. N. Widiawati, B. Nurina Sari, and T. Nur Padilah, "Clustering Data Penduduk Miskin Dampak Covid-19 Menggunakan Algoritme K-Medoids," 2022. [Online]. Available: <http://jurnal.polibatam.ac.id/index.php/JAIC>
 12. Sahasrakirana, R. I., Purwandari, T., & Hendrawati, T. (2022). Pengelompokan Kecamatan Di Kabupaten Bandung Berdasarkan Faktor Kemiskinan Menggunakan Metode K-Medoids Clustering. *E-Journal BIAStatistics| Departemen Statistika FMIPA Universitas Padjadjaran*, 16(1), 1-16.
 13. A. Rizki, D. Ardini, H. S. Program, S. S1, and S. J. Matematika, "Perbandingan Algoritme K-Medoids Dan Clara Dalam Analisis Clustering Kemiskinan Di Indonesia."
 14. R. Andrea and N. Nursobah, "Penerapan Algoritme K-Medoids Untuk Pengelompokan Data Penerima Bantuan Uang Kuliah Tunggal Bagi Mahasiswa Terdampak Covid-19," *Building of Informatics, Technology and Science (BITS)*, vol. 3, no. 4, pp. 632–638, Mar. 2022, doi: 10.47065/bits.v3i4.1294.
 15. Luth Fimawahib, "Algoritme K-Medoids untuk Pengelompokan Produksi Padi dan Beras sebagai Upaya Optimalisasi Ketahanan Pangan di Provinsi Riau," *SATIN - Sains dan Teknologi Informasi*, vol. 8, no. 2, Nov. 2022, doi: 10.33372/stn.v8i2.877.
 16. Tb. A. Munandar, "Penerapan Algoritme Clustering Untuk Pengelompokan Tingkat Kemiskinan Provinsi Banten," *JSiI (Jurnal Sistem Informasi)*, vol. 9, no. 2, pp. 109–114, Sep. 2022, doi: 10.30656/jsii.v9i2.5099.
 17. Arora P, Dr Deepali, Varshney S. Analysis of K-Means and K-Medoids Algorithm For Big Data. *International Conference on Information Security & Privacy (ICISP2015)*. Nagpur. 2016; 507-512.
 18. Kamila, I., Khairunnisa, U., & Mustakim, M. (2019). Perbandingan Algoritme K-Means dan K-Medoids untuk Pengelompokan Data Transaksi Bongkar Muat di Provinsi Riau. *Jurnal Ilmiah Rekayasa dan Manajemen Sistem Informasi*, 5(1), 119-125.
 19. Singh K, Malik D & Naveen Sharma. Evolving Limitations in K-Means Algorithm in Data Mining and Their Removal. *IJCEM International Journal of Computational Engineering & Management*. 2011; 12: 105-109.
 20. Fayyad U, Piatetsky-Shapiro G, Smyth P. From Data Mining to Knowledge Discovery in Databases. *AI Magazine*, Volume 17 Number 3, pages 37-54. 1996.
 21. I Damayanti, E Hermawan, N Kamilah. Analisis Spasial Clustering Zona Potensi Ikan Konsumsi Air Tawar di Kabupaten Bogor. *Krea-TIF: Jurnal Teknik Informatika Vol 9 No. 2* pp. 11-20. 2021