

CLUSTERING BANTUAN SOSIAL SESUAI TINGKAT KEMISKINAN DI KELURAHAN GIRIKERTO MENGUNAKAN ALGORITMA *K-MEANS*

Diva Ajeng Brillian Risaychi¹, Bullion Dragon Andah²

^{1,2} Sistem Informasi, Fakultas Teknologi Informasi, Universitas Budi Luhur, Jakarta, Indonesia

Email: ¹ 2012501017@student.budiluhur.ac.id, ² bullion.dragon@budiluhur.ac.id

(* : corresponding author)

Abstrak- Penelitian ini bertujuan untuk mengelompokkan data penduduk miskin di Kelurahan Girikerto pada tahun 2024 berdasarkan tingkat kemiskinan, dengan tujuan mengevaluasi distribusi bantuan sosial yang diterima oleh Penerima Manfaat (PM) sesuai dengan status kemiskinan yang telah ditetapkan oleh Pemerintah Kelurahan Girikerto. Metode yang digunakan adalah klusterisasi dengan algoritma *K-Means*, diterapkan menggunakan aplikasi *RapidMiner*. Data yang digunakan dalam penelitian ini berasal dari Kelurahan Girikerto dan mencakup berbagai jenis bantuan sosial seperti BPNT, PKH, PBI-JK, dan BLT BBM. Proses penelitian mencakup tahap studi literatur untuk memahami dasar teori dan metodologi yang relevan, serta identifikasi dan analisis masalah menggunakan pendekatan *CRISP-DM*. Data yang diperoleh akan dipersiapkan, dibersihkan, dan dinormalisasi sebelum dianalisis menggunakan algoritma *K-Means* untuk menghasilkan kelompok-kelompok yang menunjukkan pola distribusi bantuan yang berbeda-beda sesuai dengan tingkat kemiskinan. Hasil klusterisasi menghasilkan tiga kelompok: *Cluster_0* (penduduk miskin yang tidak menerima bantuan), *Cluster_1* (penduduk menerima 1-2 jenis bantuan), dan *Cluster_2* (penduduk miskin yang menerima 3 jenis bantuan). Visualisasi menunjukkan ketidakseimbangan dalam distribusi bantuan sosial, dengan rekomendasi agar *Cluster_0* menjadi prioritas untuk program bantuan, sementara *Cluster_1* memerlukan evaluasi untuk memastikan kecukupan bantuan, dan *Cluster_2* harus mendapatkan dukungan yang efektif.

Kata Kunci: Algoritma *K-Means*, *Clustering*, *Data mining*

CLUSTERING SOCIAL ASSISTANCE ACCORDING TO POVERTY LEVELS IN GIRIKERTO VILLAGE USING THE *K-MEANS* ALGORITHM

Abstract- This study aims to classify poverty data in the Girikerto Village for the year 2024 based on poverty levels, with the goal of evaluating the distribution of social assistance received by Beneficiaries (PM) according to the poverty status established by the Girikerto Village Government. The method used is Clustering with the *K-Means* algorithm, applied using *RapidMiner* software. The data used in this research comes from Girikerto Village and includes various types of social assistance such as BPNT, PKH, PBI-JK, and BLT BBM. The research process involves a literature review to understand relevant theories and methodologies, as well as problem identification and analysis using the *CRISP-DM* approach. The data obtained will be prepared, cleaned, and normalized before being analyzed using the *K-Means* algorithm to produce Clusters that reveal different patterns of assistance distribution according to poverty levels. The Clustering results in three groups: *Cluster_0* (poor residents not receiving any assistance), *Cluster_1* (residents receiving 1-2 types of assistance), and *Cluster_2* (poor residents receiving 3 types of assistance). Visualization shows an imbalance in the distribution of social assistance, with recommendations for prioritizing *Cluster_0* for aid programs, ensuring that *Cluster_1* receives adequate support, and providing effective assistance to *Cluster_2*.

Keywords: *K-Means Algorithm*, *Clustering*, *Data Mining*

1. PENDAHULUAN

Bantuan sosial adalah bantuan berupa uang, barang, atau jasa untuk individu, keluarga, atau masyarakat yang miskin, tidak mampu, atau rentan terhadap risiko kesehatan, sesuai dengan PERMENSOS NOMOR 1 TAHUN 2019[1]. Bantuan ini diberikan berdasarkan kriteria yang ditetapkan oleh Menteri Sosial dan harus terdaftar dalam Data Terpadu Kesejahteraan Sosial (DTKS), yang mencakup data penerima layanan sosial dan bantuan PERMENSOS NOMOR 3 TAHUN 2021[2]. Jenis bantuan meliputi BPNT (Bantuan Pangan Nontunai), PKH (Program Keluarga Harapan), PBI-JK (Bantuan Iuran Jaminan Kesehatan), dan BLT BBM (Bantuan Langsung Tunai BBM), masing-masing dengan tujuan dan ketentuan berbeda.

Secara umum data *mining* terbagi atas 2 kata, yaitu: Data = Kumpulan fakta yang terekam atau sebuah entitas yang tidak memiliki arti dan selama ini terabaikan. *Mining* = Proses Penambangan. Sehingga data *Mining* itu dapat diartikan sebagai proses Pembangunan data yang menghasilkan sebuah output berupa pengetahuan atau informasi. [3]

Kelurahan Girikerto di Kecamatan Turi, Kabupaten Sleman, melakukan pendataan penduduk miskin dan memberikan status Miskin atau Rentan Miskin. Data ini didaftarkan melalui Sistem Informasi Kesejahteraan Sosial

Next Generation (SIKS-NG) untuk memastikan bantuan yang tepat. Namun, terdapat ketidaksesuaian dalam distribusi bantuan pada tingkat kemiskinan yang sama.

Dalam penelitian sebelumnya yang berjudul Penerapan Algoritma *K-Means* Untuk *Clustering* Data Kemiskinan Provinsi Banten Menggunakan *Rapidminer*, penelitian ini menggunakan aplikasi algoritma *K-Means* dengan menggunakan RapidMiner untuk klasterisasi data kemiskinan di Provinsi Banten. Studi ini menghasilkan tiga klaster yang membedakan tingkat kemiskinan berdasarkan karakteristik penduduk seperti pendidikan dan pengeluaran per kapita [4].

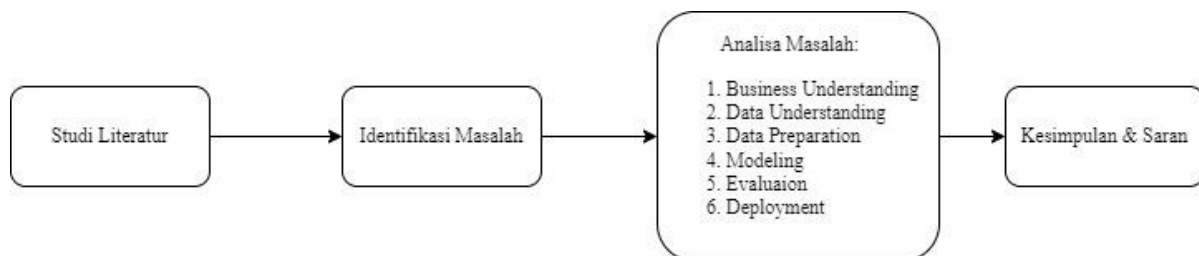
Dalam penelitian yang berjudul Klasterisasi Tingkat Kemiskinan di Indonesia Menggunakan Algoritma *K-Means*, berfokus pada klasterisasi tingkat kemiskinan di Indonesia menggunakan algoritma *K-Means*. Penelitian ini menemukan tiga kategori tingkat kemiskinan di 34 provinsi Indonesia, memungkinkan analisis yang mendalam untuk pembuatan kebijakan yang lebih baik dalam penanggulangan kemiskinan [5].

Penelitian ini bertujuan untuk mengelompokkan data penduduk miskin di Kelurahan Girikerto tahun 2024 untuk mengevaluasi kesesuaian bantuan dengan status kemiskinan menggunakan metode klasterisasi dengan algoritma *K-Means*. *K-Means Clustering* adalah metode pengolahan data yang melakukan proses pemodelan tanpa pengawasan, dan Means adalah nilai rata-rata dari suatu grup data, yang berarti nilai klaster [6]. Klasterisasi membantu mengidentifikasi pola dalam data untuk pengambilan keputusan yang lebih baik dalam kebijakan dan program bantuan sosial.

Dalam penelitian ini ada proses Transformasi data yang bertanggung jawab untuk mengubah data model analitik serta memodelkan data agar sesuai dengan analisis yang diharapkan dan format data yang dibutuhkan [7]. Normalisasi adalah proses penskalaan nilai atribut dari data sehingga bisa terletak pada rentang tertentu. Berikut beberapa normalisasi [8]. *Euclidean Distance*, rumus ini menghitung jarak antara dua titik dalam ruang n-dimensi dengan cara mengambil akar kuadrat dari jumlah kuadrat perbedaan setiap komponen dari vektor p dan q. Jarak ini memberikan indikasi seberapa dekat atau jauh kedua vektor tersebut satu sama lain dalam ruang tersebut [9].

2. METODE PENELITIAN

Proses diperlukan tahapan dengan standarisasi yang dikenal dengan model Cross Standard Industry for Data Mining (CRISP-DM) [10]. Berikut gambaran untuk alur tahapan penelitian:



Gambar 1. Tahapan Penelitian

2.1 Studi Literatur

Studi literatur dilakukan untuk menambah pengetahuan dasar dan teori-teori yang digunakan dalam penelitian. Tahapannya adalah pengumpulan data dan informasi yang berkaitan dengan penelitian yang akan dilakukan.

2.2 Identifikasi Masalah

Tahap pertama dari proses penelitian adalah identifikasi masalah yang relevan dengan subjek. Tahap ini dilakukan dengan menggunakan informasi yang dikumpulkan selama tahap studi literatur, observasi langsung, dan sumber-sumber pendukung penelitian. Tahapan ini adalah langkah pertama dalam proses pemecahan masalah, di mana kita mencoba untuk memahami apa yang menyebabkan masalah itu.

2.3 Analisa Masalah

Analisa masalah bertujuan mengetahui gambaran mengenai masalah yang diteliti. Berikut tahapan pada CRISP-DM:

1. *Business Understanding*

Tahap ini melibatkan pemahaman mendalam tentang masalah yang ingin dipecahkan dengan data mining. Ini melibatkan interaksi dengan pemangku kepentingan untuk mendefinisikan tujuan bisnis, target, dan kendala yang relevan.

2. Data Understanding

Pada tahap ini, mengumpulkan data dan memahami data yang relevan untuk proyek data mining. Dalam mencakup eksploitasi sumber data, eksporasi data awal, dan pemahaman karakteristik data.

3. Data Preparation

Tahap ini melibatkan pembersihan, transformasi, dan pemodelan data yang akan digunakan dalam proses data mining. Proses ini termasuk penghapusan data yang tidak relevan, penggabungan sumber data, dan pelabelan data.

4. Modeling

Di tahap ini, anda memilih Teknik data mining yang sesuai dan mengembangkan model analisis data. Ini dapat melibatkan Teknik seperti regresi, klasifikasi, klasterisasi, atau asosiasi, tergantung pada tujuan proyek.

5. Evaluation

Setelah mengembangkan model, data harus diuji dan di evaluasi untuk memastikan bahwa hal tersebut memenuhi tujuan bisnis dan memiliki kinerja yang baik. Evaluasi melibatkan penggunaan Teknik statistika dan Teknik validasi untuk mengukur kualitas model.

6. Deployment

Pada tahap terakhir dari metodologi CRISP-DM, informasi atau pengetahuan yang diperoleh dari proses pemodelan termasuk analisis setiap *Cluster* yang dihasilkan dari *K-Means Clustering* digunakan untuk menghasilkan laporan yang disusun secara sistematis. Pihak yang bertanggung jawab, seperti manajemen dan tim terkait, dapat memahami dan membuat keputusan lebih mudah dengan laporan ini. Analisis yang dilakukan pada setiap kelompok data dalam laporan dilakukan dengan tujuan untuk memberikan pemahaman yang lebih mendalam tentang karakteristik masing-masing kelompok data, potensi konsekuensi bisnisnya, dan saran untuk tindakan yang tepat berdasarkan hasil analisis. Oleh karena itu, langkah ini tidak hanya menyelesaikan proses analisis data, tetapi juga menghasilkan solusi dan pengetahuan yang bermanfaat untuk organisasi atau konteks yang relevan.

3. HASIL DAN PEMBAHASAN

Hasil dari implementasi pada metode penelitian ini terdapat beberapa tahapan dengan penjabaran sebagai berikut:

3.1 Pengumpulan Data

Penelitian ini menggunakan data penduduk miskin tahun 2024 yang diambil dari Kelurahan Girikerto. Data tersebut berisi 3511 Nama penduduk miskin dengan atribut Kalurahan, Padukuhan, RT, RW, Umur, Status, BPNT, BST, PKH, PBI-JK, BLT-BBM, Bantuan yatim Piatu, RST, Permakanan, Subsidi Adiktif. Data tersebut dikumpulkan melalui dua kali tahapan musyawarah. Musyawarah yang pertama melalui Kepala Padukuhan Bersama RT dan RW dilanjut dengan musyawarah kedua yaitu Kepala Padukuhan Bersama Pamong Kelurahan. Data tersebut digunakan untuk bantuan dari Pemerintah Desa atau pun Pemerintah Pusat. Berikut merupakan data mentah yang di terima dari Kalurhan Girikerto pada tahun 2024.

Tabel 1. Dataset setelah dikumpulkan

KALURAHAN	PADUKUHAN	RT	RW	NAMA	UMUR	STATUS	BPNT	BST	PKH	PBI-JK	BLT-BBM	BYP	RST	P	SA
Girikerto	Babadan	005	024	PM1	60	rentan miskin			1	1					
Girikerto	Babadan	002	023	PM2	39	rentan miskin			1	1					
Girikerto	Babadan	003	010	PM3	26	rentan miskin				1					
Girikerto	Babadan	003	022	PM4	34	rentan miskin	1		1	1					
Girikerto	Babadan	003	008	PM5	19	rentan miskin	1		1	1					
Girikerto	Babadan	002	001	PM6	19	rentan miskin			1	1					
Girikerto	Babadan	004	016	PM7	26	miskin					1				
Girikerto	Babadan	004	014	PM8	47	rentan miskin	1		1	1					
Girikerto	Babadan	003	017	PM9	17	rentan miskin			1	1					
Girikerto	Babadan	001	023	PM10	40	miskin					1				
Girikerto	Babadan	001	007	PM11	18	miskin					1				
Girikerto	Babadan	003	016	PM12	32	rentan miskin					1				

3.2 Data Pre-processing

Tahapan ini memilih kolom yang akan digunakan untuk diolah. Dan data tersebut menjadi seperti Tabel 2.

Tabel 2. Data kolom terpilih

PADUKUHAN	NAMA	STATUS	BPNT	PKH	PBI-JK
Babadan	PM1	rentan miskin		1	1
Babadan	PM2	rentan miskin		1	1
Babadan	PM3	rentan miskin			1
Babadan	PM4	rentan miskin	1	1	1
Babadan	PM5	rentan miskin	1	1	1
Babadan	PM6	rentan miskin		1	1
Babadan	PM7	miskin			1
Babadan	PM8	rentan miskin	1	1	1
Babadan	PM9	rentan miskin		1	1
Babadan	PM10	miskin			1

Dalam penelitian ini dataset di transformasi menjadi seperti tabel 3.

Tabel 3. Dataset setelah di transformasi

PADUKUHAN	PENDUDUK MISKIN	BPNT	PKH	PBI-JK	JUMLAH BANTUAN
Babadan	284	50	109	218	377
Bangunmulyo	240	34	85	161	280
Daleman	248	39	65	160	264
Glagahombo	176	31	45	136	212
Karangawang	226	46	85	180	311
Kemirikebo	270	46	50	175	271
Kloposawit	165	23	36	99	158
Nangsri	198	37	62	142	241
Ngandong	495	91	156	386	633
Nganggring	426	42	91	194	327
Pancoh	243	38	80	186	304
Sukorejo	294	45	67	209	321
Surodadi	245	54	77	176	307

Tabel tersebut merupakan hasil perhitungan pada atribut Penduduk Miskin dan penjumlahan untuk atribut BPNT, PKH, PBI-JK. Atribut Jumlah Bantuan merupakan jumlah dari atribut BPNT, PKH dan PBI-JK. Penjumlahan tersebut disesuaikan dengan setiap Padukuhan yang berada di Girikerto.

Ada beberapa jenis normalisasi dalam data mining, pada penelitian ini normalisasi data menggunakan *min-max normalization* dengan rumus:

$$\frac{v - \min A}{\max A - \min A} \quad (3.1)$$

salah satu contoh perhitungan Normalisasi pada salah satu atribut data

Diketahui:

$$\begin{aligned}
 v & \text{ (Nilai Atribut Penduduk Miskin): } 270 \\
 \text{minA (Nilai Min Atribut Penduduk Miskin): } & 165 \\
 \text{maxA (Nilai Max Atribut Penduduk Miskin): } & 495 \\
 \text{Maka: } v' & = \frac{270 - 165}{495 - 165} \\
 & = \mathbf{0,318}
 \end{aligned}$$

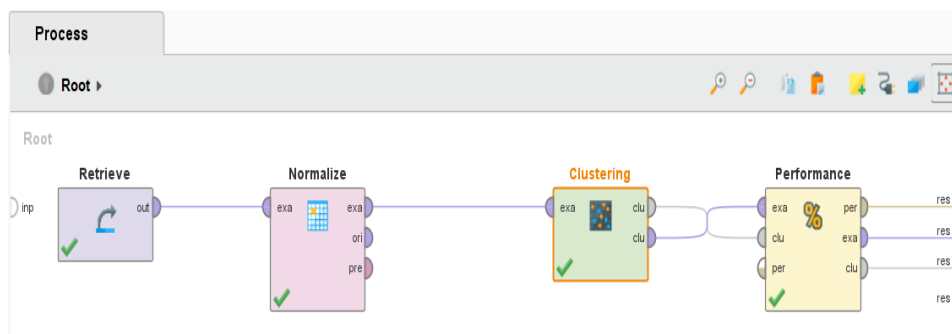
Tabel 4. Dataset Setelah di normalisasi

PADUKUHAN	PENDUDUK MISKIN	BPNT	PKH	PBI-JK	JUMLAH BANTUAN
Babadan	0,361	0,397	0,608	0,415	0,461
Bangunmulyo	0,227	0,162	0,408	0,216	0,257
Daleman	0,252	0,235	0,242	0,213	0,223
Glagahombo	0,033	0,118	0,075	0,129	0,114
Karangawang	0,185	0,338	0,408	0,282	0,322
Kemirikebo	0,318	0,338	0,117	0,265	0,238
Kloposawit	0	0	0	0	0
Nangsri	0,1	0,206	0,217	0,150	0,175
Ngandong	1	1	1	1	1
Nganggriing	0,791	0,279	0,458	0,331	0,356
Pancoh	0,236	0,221	0,367	0,303	0,307
Sukorejo	0,391	0,324	0,258	0,383	0,343
Surodadi	0,242	0,456	0,342	0,268	0,314

Dataset ini sudah dinormalisasi, memungkinkan analisis yang lebih adil dan konsisten antar padukuhan. Normalisasi memastikan bahwa setiap fitur memiliki skala yang sama, yang sangat penting untuk algoritma berbasis jarak seperti *Clustering*. Dengan *dataset* yang dinormalisasi, teknik analisis seperti *K-Means Clustering* dapat diterapkan lebih efektif untuk mengidentifikasi pola dan kelompok dalam data, serta membuat keputusan yang lebih baik berdasarkan distribusi bantuan sosial dan karakteristik penduduk miskin.

3.2.1 Penyajian Model Terbaik

Bagian ini menyajikan algoritma *K-Means* sebagai hasil proses terbaik dengan 3 *Cluster*. Proses ini menggunakan *operator normalize*, *Clustering K-Means*, dan *Cluster distance performance* dengan tools *RapidMiner*. Tujuannya untuk lebih meningkatkan kepastian akan tingkat akurasi yang diperoleh.



Gambar 2. Penyajian Model dengan *RapidMiner*

Merupakan desain arsitektur pada tools *RapidMiner* menggunakan operator *Normalize* Normalisasi digunakan untuk menskalakan nilai agar sesuai dengan rentang tertentu. Menyesuaikan rentang nilai sangat penting saat menangani Atribut dengan unit dan skala yang berbeda. Operator *Cluster Distance Performance* berfungsi sebagai masukan dan mengevaluasi kinerja model berdasarkan *Centeroid Cluster*. Dua ukuran kinerja didukung dari rata – rata dalam jarak *Cluster* dan *indeks Davies Bouldin*.

3.2.2 Perhitungan Dengan Algoritma Terpilih

Pada Tahapan ini dilakukan perhitungan manual dengan algoritma *K-Means* sebagai algoritma terpilih. Pengolahan data penduduk miskin daerah Girikerto dilakukan setelah tahapan transformasi sehingga data tersebut bisa diolah. Perhitungan menggunakan sample dataset yang diambil sebanyak 3 data. Data atribut BPNT, atribut PKH, atribut PBI-JK, dan atribut Jumlah Bantuan yang diproses sebagai berikut:

Tabel 5. Dataset pada algoritma terpilih

PADUKUHAN	PENDUDUK MISKIN	BPNT	PKH	PBI-JK	JUMLAH BANTUAN
Surodadi	0,242	0,456	0,342	0,268	0,314
Glagahomboh	0,033	0,118	0,075	0,129	0,114
Ngandong	1	1	1	1	1
Kloposawit	0	0	0	0	0

Setelah data di inialisasi menjadi angka maka selanjutnya melakukan tahap berikut:

1. Menentukan jumlah *Cluster*, pada penelitian ini memilih 3 *Cluster* ($k=3$) yang akan dibuat.
2. Lalu menentukan *Centeroid* tiap *Cluster*, metode yang digunakan mengambil data secara acak, berikut data *Centeroid* tersebut:

Tabel 6. Titik pusat setiap awal *Cluster*

CENTEROID AWAL	PENDUDUK MISKIN	BPNT	PKH	PBI-JK	JUMLAH BANTUAN
C1	0,185	0,126	0,075	1	0,461
C2	0,1	0,118	0	0,216	0,114
C3	0	0,397	0,117	0,129	0,307

$$d(x, y) = \|x - y\| = \sqrt{\sum (x_i - y_i)^2} \quad (3.1)$$

Berikut jarak data 1 dengan *Centeroid*1:

$$d(1, 1) = \sqrt{(0,242-0,185)^2 + (0,456-0,126)^2 + (0,342-0,075)^2 + (0,268-1)^2 + (0,314-0,461)^2} = 0,859$$

Selanjutnya jarak data 1 dengan *Centeroid* 2:

$$d(1, 2) = \sqrt{(0,242-0,1)^2 + (0,456-0,118)^2 + (0,324-0)^2 + (0,268-0,216)^2 + (0,314-0,114)^2} = 0,542$$

Selanjutnya jarak data 1 dengan *Centeroid* 3:

$$d(1, 3) = \sqrt{(0,242-0)^2 + (0,456-0,397)^2 + (0,324-0,117)^2 + (0,268-0,219)^2 + (0,314-0,307)^2} = 0,365$$

Berikut jarak data 2 dengan *Centeroid* 1:

$$d(2, 1) = \sqrt{(0,033-0,185)^2 + (0,118-0,126)^2 + (0,075-0,075)^2 + (0,129-1)^2 + (0,114-0,462)^2} = 0,951$$

Selanjutnya jarak data 2 dengan *Centeroid* 2:

$$d(2, 2) = \sqrt{(0,033-0,1)^2 + (0,118-0,118)^2 + (0,075-0)^2 + (0,129-0,216)^2 + (0,114-0,114)^2} = 0,133$$

Selanjutnya jarak data 2 dengan *Centeroid* 3:

$$d(2, 3) = \sqrt{(0,033-0)^2 + (0,118-0,397)^2 + (0,075-0,117)^2 + (0,129-0,219)^2 + (0,114-0,307)^2} = 0,344$$

Berikut jarak data 3 dengan *Centeroid* 1:

$$d(3, 1) = \sqrt{(1-0,185)^2 + (1-0,126)^2 + (1-0,075)^2 + (1-1)^2 + (1-0,461)^2} = 1,251$$

Selanjutnya jarak data 3 dengan *Centeroid* 2:

$$d(3, 2) = \sqrt{(1-0,1)^2 + (1-0,118)^2 + (1-0)^2 + (1-0,216)^2 + (1-0,114)^2} = 1,720$$

Selanjutnya jarak data 3 dengan *Centeroid* 3:

$$d(3, 3) = \sqrt{(1-0)^2 + (1-0,397)^2 + (1-0,117)^2 + (1-0,129)^2 + (1-0,307)^2} = 1,834$$

Selanjutnya jarak data 4 dengan *Centeroid* 1:

$$d(4, 1) = \sqrt{(0-0,185)^2 + (0-0,126)^2 + (0-0,75)^2 + (0-1)^2 + (0-0,461)^2} = 1,126$$

Selanjutnya jarak data 4 dengan *Centeroid* 2:

$$d(4, 2) = \sqrt{(0-0,1)^2 + (0-0,118)^2 + (0-0)^2 + (0-0,261)^2 + (0-0,114)^2} = 0,289$$

Selanjutnya jarak data 4 dengan *Centeroid* 3:

$$d(4, 3) = \sqrt{(0-0,1)^2 + (0-0,118)^2 + (0-0)^2 + (0-0,261)^2 + (0-0,114)^2} = 0,531$$

Hasil perhitungan jarak *euclidean distance* dirangkum dalam bentuk tabel.

Tabel 7. Data perhitungan setiap data ke setiap *Centeroid*

ATRIBUT	C0	C1	C2	TITIK TERDEKAT
Surodadi	0,859	0,542	0,365	0,540
Glagahombo	0,951	0,133	0,344	0,133
Ngandong	1,251	1,720	1,834	1,251
Kloposawit	1,126	0,289	0,531	0,289

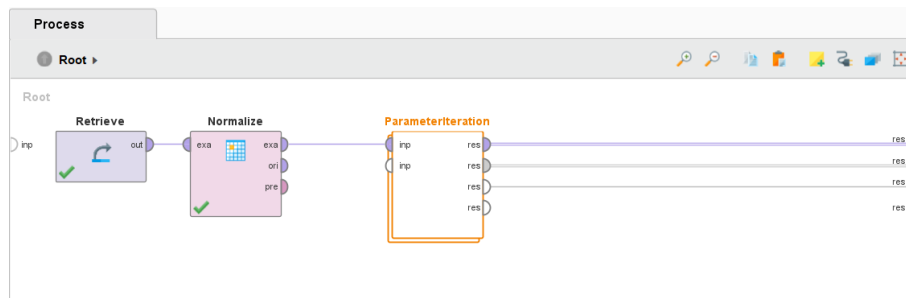
Hasil *Cluster* menurut titik terdekat sesuai dengan hasil *Cluster* menggunakan *tools RapidMiner*.

Tabel 8. Hasil *Cluster* menggunakan *tools RapidMiner*

PADUKUHAN	PENDUDUK MISKIN	BPNT	PKH	PBI-JK	JUMLAH BANTUAN	CLUSTER
Babadan	0,361	0,397	0,608	0,415	0,461	Cluster_1
Bangunmulyo	0,227	0,162	0,408	0,216	0,257	Cluster_1
Daleman	0,252	0,235	0,242	0,213	0,223	Cluster_1
Glagahombo	0,033	0,118	0,075	0,129	0,114	Cluster_0
Karanggawang	0,185	0,338	0,408	0,282	0,322	Cluster_1
Kemirikebo	0,318	0,338	0,117	0,265	0,238	Cluster_1
Kloposawit	0	0	0	0	0	Cluster_0
Nangsri	0,1	0,206	0,217	0,150	0,175	Cluster_0
Ngandong	1	1	1	1	1	Cluster_2
Nganggring	0,791	0,279	0,458	0,331	0,356	Cluster_1
Pancoh	0,236	0,221	0,367	0,303	0,307	Cluster_1
Sukorejo	0,391	0,324	0,258	0,383	0,343	Cluster_1
Surodadi	0,242	0,456	0,342	0,268	0,314	Cluster_1

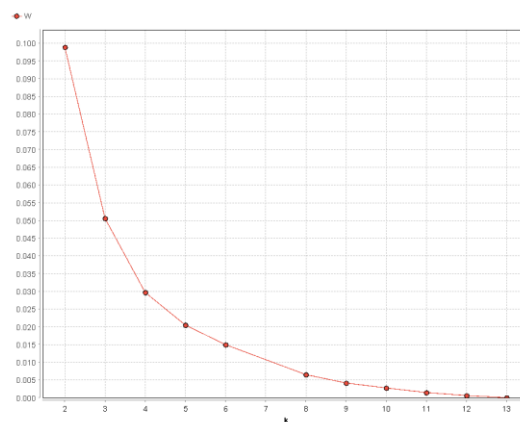
3.3 Pengujian

Pada bagian ini disampaikan mekanisme pengukuran performa model terbaik. Metode uji dari penelitian ini menggunakan *Elbow method*. Berikut proses dari pengujian yang menggunakan *tools RapidMiner*:



Gambar 3. Operator Pengujian

Dari operator pada Gambar 3. menghasilkan visualisasi *elbow method* yang digunakan untuk menentukan *Cluster* optimal.



Gambar 4. Grafik Elbow

Pada Gambar 4 menggambarkan grafik elbow yang dihasilkan dari operator pengujian. Meskipun grafik siku menunjukkan penurunan WSS yang signifikan pada $k=4$, $k=5$, dan $k=6$, namun ini tidak selalu berarti bahwa $k=4$, $k=5$, atau $k=6$ adalah jumlah kluster yang optimal. Penurunan yang signifikan ini hanya menunjukkan bahwa pembagian data menjadi kluster-kluster ini mengurangi variansi data dalam kluster, tetapi tidak menjamin bahwa jumlah kluster tersebut adalah yang terbaik untuk tujuan analisis yang spesifik.

Pemilihan jumlah kluster optimal lebih didasarkan pada interpretasi penurunan WSS yang signifikan dan mempertimbangkan tujuan analisis dan konteks data. Meskipun $k=4$, $k=5$, dan $k=6$ memberikan penurunan WSS yang signifikan, jika analisis atau tujuan praktisnya memerlukan hanya tiga kelompok yang jelas dan terpisah, maka $k=3$ dapat dianggap sebagai jumlah kluster yang optimal.

4. KESIMPULAN

Dari pembahasan diatas penulिन menarik Kesimpulan sebagai berikut:

- Hasil jumlah *Cluster* untuk penelitian ini 3 *Cluster* dengan pembagian *Cluster_0* adalah penduduk miskin yang tidak menerima bantuan, *Cluster_1* penduduk menerima 1 sampai 2 bantuan dan *Cluster_2* penduduk miskin yang menerima 3 bantuan.
- Berdasarkan visualisasi yang ada terlihat tidak setara bantuan sosial yang diterima penduduk miskin di daerah Girikerto. *Cluster 0*: Penduduk yang tidak menerima bantuan sebaiknya menjadi prioritas untuk mendapatkan intervensi atau program bantuan yang dapat membantu meningkatkan kesejahteraan mereka. *Cluster 1*: Penduduk dalam *Cluster* ini menerima satu atau dua jenis bantuan. Evaluasi untuk memastikan bahwa bantuan yang diberikan cukup untuk kebutuhan mereka, dan jika perlu, tingkatkan jumlah atau jenis bantuan yang diterima. *Cluster 2*: Pastikan bahwa penduduk yang menerima tiga jenis bantuan tidak hanya menerima bantuan tetapi juga mendapatkan dukungan yang efektif dan efisien dalam penggunaannya.
- Penelitian ini diharapkan dapat memberikan pertimbangan penting bagi pemerintah Kalurahan Girikerto dalam usaha untuk mengoptimalkan distribusi bantuan sosial BPNT, PKH, dan PBI-JK kepada *Cluster 0* dan *Cluster 1* secara lebih adil. Tujuannya adalah untuk memastikan bahwa bantuan sosial yang diberikan benar-benar

mencapai sasaran dan memberikan manfaat nyata bagi penduduk miskin, sehingga dapat secara signifikan dan berkelanjutan meningkatkan kesejahteraan mereka.

DAFTAR PUSTAKA

- [1] Menteri Sosial, “Peraturan Menteri Sosial Nomor 1 Tahun 2019,” 2019, Accessed: Jun. 10, 2024. [Online]. Available: <https://peraturan.bpk.go.id/Details/129443/permensos-no-1-tahun-2019>
- [2] Menteri Sosial, “Peraturan Kementerian Sosial Nomor 3 Tahun 2021.” Accessed: Jun. 15, 2024. [Online]. Available: <https://peraturan.bpk.go.id/Details/171535/permensos-no-3-tahun-2021>
- [3] F. Marisa, A. L. Maukar, and T. M. Akhriza, “*Data Mining Konsep dan Penerapannya*,” Deepublish, 2021. [Online]. Available: <https://books.google.co.id/books?id=BtIVEAAAQBAJ>
- [4] Y. R. Sari, A. Sudewa, D. A. Lestari, and T. I. Jaya, “Penerapan Algoritma K-Means Untuk Clustering Data Kemiskinan Provinsi Banten Menggunakan Rapidminer,” *Journal of Computer Engineering System and Science*, vol. 5, no. 2, pp. 192-198, 2020.
- [5] A. Khalif *et al.*, “Klasterisasi Tingkat Kemiskinan di Indonesia Menggunakan Algoritma K-Means,” *Generation Journal*, vol. 8, no. 1, pp. 54-62, 2024.
- [6] Nurhayati, “*Pemodelan K- Means Algoritma Dan Big Data Analysis (Pemetaan Data Mustahiq)*,” Pascal Books, 2022. [Online]. Available: https://books.google.co.id/books?id=_bJmEAAAQBAJ
- [7] E. Buulolo, “Data Mining Untuk Perguruan Tinggi,” 2020.
- [8] F. Sulianta, “*Basic Data Mining from A to Z*. Feri Sulianta,” 2023. [Online]. Available: <https://books.google.co.id/books?id=JcLhEAAAQBAJ>
- [9] R. Alhapizi, M. Nasir, and I. Effendy, “Penerapan Data Mining Menggunakan Algoritma K-Means Clustering Untuk Menentukan Strategi Promosi Mahasiswa Baru Universitas Bina Darma Palembang,” *SEA: Journal of Software Engineering Ampera*, vol. 1, no. 1, pp. 1-14, 2020. [Online]. Available: <https://journal-computing.org/index.php/journal-sea/index>
- [10] Y. Suhandi, I. Kurniati, and S. Norma, “Penerapan Metode Crisp-DM Dengan Algoritma K-Means Clustering Untuk Segmentasi Mahasiswa Berdasarkan Kualitas Akademik,” *Jurnal Teknologi Informatika dan Komputer*, vol. 6, no. 2, pp. 12–20, 2020, doi: 10.37012/jtik.v6i2.299.