Bit (Fakultas Teknologi Informasi Universitas Budi Luhur) p-ISSN: 1693-9166
Vol. 21, No. 2, September 2024, hlm. 169-177 e-ISSN: 2685-127X

KLASTERISASI TINGKAT KEMISKINAN MENGGUNAKAN METODE K-MEANS DI DKI JAKARTA

Muhamad Khaerul Rafli[1*], Utomo Budiyanto[2]

[1,2] Teknik Informatika, Fakultas Teknologi Informasi, Universitas Budi Luhur, Jakarta Selatan, Indonesia
Email: [1*]2011501836@student.budiluhur.ac.id, [2]utomo.budiyanto@budiluhur.ac.id

(Naskah masuk: 7 Agustus 2024, diterima untuk diterbitkan: 7 September 2024)

Abstrak

Tingkat kemiskinan dipengaruhi oleh faktor seperti kurangnya lapangan kerja yang memadai ataupun suatu pekerjaan yang tidak mampu mencukupi tanggungan keluarga serta anak-anak. Selain itu, salah satu alasan mengapa masalah kemiskinan belum terselesaikan adalah hasil survei menunjukkan bahwa bantuan pemerintah sering kali tidak sesuai dengan kebutuhan penduduk. Selain itu, dalam pendistribusian bantuan sosial yang dilakukan oleh pemerintah juga tidak tepat sasaran serta tidak merata. Hal tersebut bisa diakibatkan oleh ketidakakuratan data validasi. Tujuan dari penelitian ialah mengidentifikasi hasil *Clustering K-means* dengan data primer yang diperoleh dari Pusat Data dan Teknologi Informasi Dinas Sosial Provinsi DKI Jakarta dan mengelompokkan tingkat kemiskinan di Provinsi DKI Jakarta menggunakan metode *Clustering K-means*. Metode penelitian didasarkan pada penggunaan algoritma *K-means clustering*. Kesimpulan yang diambil ialah penggunaan metode *Clustering K-Means* dapat diterapkan untuk mengelompokan *cluster* kemiskinan di Provinsi DKI Jakarta. Pada proses pengujian terhadap 8000 data penduduk DKI Jakarta. Didapatkan hasil 14 data. Pengujian dilakukan dengan jumlah *cluster* yang diuji $K = 4$. Menghitung jarak setiap data dengan *centroid* awal menggunakan persamaan *Euclidean Distance*. Hasil yang diperoleh *centroid* ke-*1* 14,97 dan *centroid* ke-*2* 22,05. Dari 11 iterasi yang diproses didapatkan hasil *clustering* yaitu iterasi ke-3. Dikarenakan iterasi ke-3 dipilih sebagai yang paling cocok karena beberapa alasan utama. Pertama, *centroid* tidak banyak berubah lagi setelah *iterasi* ini, menunjukkan bahwa algoritma telah mencapai stabilitas. Kedua, terdapat penurunan *inertia* yang signifikan, mengindikasikan bahwa jarak total dari titik-titik ke *centroid* mereka telah diminimalkan secara optimal. Terakhir, validasi visual menunjukkan bahwa data telah dikelompokkan dengan baik sesuai dengan *centroid* yang dihitung.

Kata kunci: *clustering, data mining, Kemiskinan DKI Jakarta, k-means, euclidean distance*

CLUSTERIZATION OF POVERTY LEVELS USING THE K-MEANS METHOD IN DKI JAKARTA

Abstract

The poverty rate is influenced by factors such as lack adequate employment opportunities or a job that is unable to support family and children. In addition, one reasons why the poverty problem has not been resolved is that survey results show that government assistance often does not match the needs population. In addition, the distribution of social assistance carried out by the government is also not on target and uneven. This can be caused by inaccurate validation data. The purpose study was to identify the results of K-means Clustering with primary data obtained from the Data and Information Technology Center DKI Jakarta Provincial Social Service and to group the poverty rate in DKI Jakarta Province using the K-means Clustering method. The research method is based on the use K-means clustering algorithm. The conclusion drawn is that the use K-Means Clustering method can be applied to group poverty clusters in DKI Jakarta Province. In the testing process on 8000 DKI Jakarta population data. The results obtained were 14 data. The test was carried out with the number of clusters tested K = 4. Calculating the distance of each data with the initial centroid using the Euclidean Distance equation. The results obtained centroid 1 14,97 and centroid 2 22,05. From 11 iterations processed, the clustering results obtained were iteration 3. Because iteration 3 was chosen as the most suitable for several main reasons. First, the centroids did not change much after this iteration, indicating that the algorithm had reached stability. Second, there was a significant decrease in inertia, indicating that the total distance from the points to their centroids had been optimally minimized. Finally, visual validation showed that the data had been well clustered according to the calculated centroids.

Keywords*: Clustering, Data Mining, Poverty of DKI Jakarta, k-means, euclidean distance*

1. PENDAHULUAN

Kemiskinan tetap menjadi tantangan kritis di Indonesia. Meskipun pemerintah telah mengimplementasikan berbagai program untuk mengatasinya [1]. Masalah ini persisten dan kompleks, terutama karena berkaitan erat dengan

tingkat pendapatan individu dan rumah tangga [2]. Pendapatan yang tidak memadai menyebabkan individu hanya mampu memenuhi kebutuhan dasar, yang merupakan standar minimum untuk kehidupan yang layak. Tingkat pendapatan yang cukup untuk memenuhi kebutuhan dasar ini menjadi penentu garis kemiskinan, yaitu ambang batas yang membedakan antara kondisi miskin dan tidak miskin [3]. Meskipun upaya-upaya pemerintah terus berlangsung, penyelesaian penuh terhadap masalah kemiskinan membutuhkan pendekatan yang lebih mendalam dan berkelanjutan untuk meningkatkan pendapatan dan kualitas hidup masyarakat [4].

Tingkat kemiskinan dipengaruhi oleh faktor seperti kurangnya lapangan kerja yang memadai ataupun suatu pekerjaan yang tidak mampu mencukupi tanggungan keluarga serta anak-anak [5]. Selain itu, salah satu alasan mengapa masalah kemiskinan belum terselesaikan adalah hasil survei menunjukkan bahwa bantuan pemerintah sering kali tidak sesuai dengan kebutuhan penduduk. Selain itu, dalam pendistribusian bantuan sosial yang dilakukan oleh pemerintah juga tidak tepat sasaran serta tidak merata [6]. Hal tersebut bisa diakibatkan oleh ketidakakuratan data validasi.

Menurut studi literatur yang dilakukan memaparkan tingkat kemiskinan dipengaruhi oleh faktor seperti kurangnya lapangan pekerjaan yang memadai ataupun suatu pekerjaan yang tidak mampu mencukupi tanggungan keluarga serta anak-anak. Masalah kemiskinan yang terjadi dalam ruang lingkup masyarakat belum terselesaikan dengan merata sebab ketidaktepatan bantuan yang diberikan oleh pemerintah. Pemerintah dalam memberikan sebuah bantuan kepada masyarakat miskin belum dilakukan secara merata [7].

Pada implementasi *data mining* terdapat suatu teknik yang disebut *clustering* yang bertujuan dalam mengelompokkan beberapa obyek berdasarkan kesamaan karakteristiknya kedalam beberapa *cluster*. Terdapat dua jenis metode *clustering* yaitu *clustering* hierarkis dan *clustering* partisi. Dalam *clustering* hierarkis data dikelompokkan ke dalam struktur pohon di mana kelompok yang paling mirip digabungkan bertahap hingga semua data tergabung dalam satu cluster [8].

Pada penelitian terdahulu yang berjudul "Analisis *Clustering* Provinsi di Indonesia Berdasarkan Tingkat Kemiskinan Menggunakan Algoritma K-Means" memaparkan provinsi yang termasuk dalam *cluster* 0 atau provinsi yang memiliki tingkat kemiskinan rendah diantaranya adalah Gorontalo, Sulawesi Tengah, BTT, NTB, DI Yogyakarta, Jawa Tengah, Lampung, Bengkulu, dan Aceh. Selanjutnya provinsi yang termasuk dalam *cluster* 1 atau provinsi yang memiliki tingkat kemiskinan sedang memiliki sejumlah 21 provinsi di Indonesia. Terakhir, *cluster* 2 yaitu provinsi yang memiliki tingkat kemiskinan tinggi diantaranya adalah Papua, Papua Barat, dan Maluku [9]. Pada penelitian tersebut telah dipaparkan sejumlah provinsi tingkat kemiskinan di seluruh Indonesia. Sedangkan kebaruan yang diambil dalam penelitian ini, peneliti menypesifikasikan atau mengambil salah satu provinsi di Indonesia untuk dilakukan analisis tingkat kemiskinan dengan menggunakan metode *K-means*. Adapun provinsi yang menjadi tinjauan analisis pada penulisan ini ialah DKI Jakarta.

Pada penelitian terdahulu yang dilakukan pada obyek warga DKI Jakarta dengan menggunakan algoritma K-Means hanya terbatas pada analisis kepadatan penduduk DKI Jakarta. Pengelompokan kepadatan penduduk di Provinsi DKI Jakarta dapat menghasilkan 2 *Cluster* yaitu tertinggi dan terendah. Kepadatan penduduk tertinggi terdapat pada kelurahan Kali Anyar pada tahun 2019 yang berjumlah 95676.10063 (jiwa/km), dan daerah kelurahan terendah terdapat pada kelurahan Pulau Harapan pada tahun 2018 dengan jumlah 1045.276234(jiwa/km) [10]. Pada penelitian ini, peneliti menggunakan *clustering K-means* yang merupakan metode lebih kompleks dan detail dalam mengelompokkan data kemiskinan. Selain itu, penelitian ini berfokus pada data dari Dinas Sosial Provinsi DKI Jakarta yang belum banyak dieksplorasi sebelumnya dengan pendekatan data *mining*. *Gap* penelitian ini mencakup penerapan metode yang lebih inovatif dan spesifik dalam konteks DKI Jakarta, serta penggunaan *dataset* yang lebih akurat.

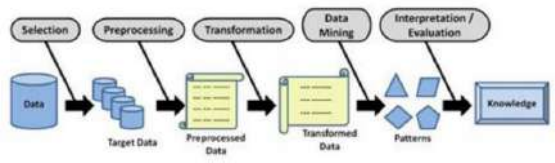
K-means adalah salah satu metode *clustering* yang sering digunakan untuk mengelompokkan data ke dalam beberapa grup berdasarkan kesamaan karakteristik. Untuk aplikasi Klasterisasi tingkat kemiskinan, metode ini dapat sangat berguna dalam mengidentifikasi kelompok-kelompok dengan tingkat kemiskinan yang serupa. Dalam hal ini untuk memanfaatkan teknik *data mining* dan *clustering Kmeans*, dengan tujuan utama mengidentifikasi dan memetakan wilayah-wilayah di DKI Jakarta yang mengalami kondisi kemiskinan. Hasil dari penelitian ini bertujuan untuk memberikan gambaran bagi Dinas Sosial Provinsi DKI Jakarta, membantu dalam merumuskan strategi dan keputusan yang lebih informasi. Melalui pendekatan yang sistematis dan berbasis data, penelitian ini berupaya memfasilitasi peningkatan kebijakan publik di DKI Jakarta terutama dalam upaya mengurangi kemiskinan dan meningkatkan kualitas hidup warga DKI Jakarta.

Dalam penelitian ini, teknik *data mining* yang digunakan untuk mengelompokkan wilayah di DKI Jakarta adalah metode *Clustering K-means*. Metode ini memiliki keunggulan dalam hal kemudahan pemahaman dan implementasi serta proses pembelajarannya yang relatif cepat. Metode *Clustering K-means* diterapkan untuk melakukan klasterisasi wilayah di Provinsi DKI Jakarta guna mengidentifikasi daerah dengan tingkat kemiskinan tertentu [11]. Penentuan jumlah klaster yang optimal dilakukan menggunakan metode *Elbow* [12].

Berdasarkan pemaparan tersebut, peulis mengambil judul "Klasterisasi Tingkat Kemiskinan Menggunakan Metode *K-Means* Di Dki Jakarta". Tujuan dari penelitian ialah mengidentifikasi hasil *Clustering K-means* dengan data primer yang diperoleh dari Pusat Data dan Teknologi Informasi Dinas Sosial Provinsi DKI Jakarta dan mengelompokkan tingkat kemiskinan di Provinsi DKI Jakarta menggunakan metode *Clustering K-means*.

2. METODE PENELITIAN

Metode penelitian didasarkan pada penggunaan algoritma *K-means clustering*. *K-means clustering* adalah metode dalam statistik dan *machine learning* yang digunakan untuk mengelompokkan data ke dalam beberapa grup atau klaster. Tujuannya adalah untuk membagi data yang memiliki kemiripan satu sama lain ke dalam klaster-klaster yang berbeda, sehingga data dalam satu klaster lebih mirip satu sama lain dibandingkan dengan data di klaster lainnya. Data penelitian adalah informasi yang diperoleh langsung dari Dinas Sosial Provinsi DKI Jakarta. Adapun beberapa langkah yang digunakan dalam penerapan metode *clustering k-means* diantaranya (1) menentukan jumlah *cluster*, (2) inisialisasi *centroid* awal, (3) mengelompokkan data, (4) iterasi, (5) evaluasi hasil. Gambar 1 dibawah ini memaparkan terkait dengan metode *Knowledge Discovery in Database* (KDD), yaitu:



Gambar 1. Metode *Knowledge Discovery in Database* (KDD)

Pada tahap *selection* bertujuan untuk memastikan pemilihan variabel yang tepat dalam pengolahan data, sehingga menghindari duplikasi atau perulangan yang tidak diperlukan. Pada tahap *prepocessing* meliputi dua langkah yaitu data *cleaning* dan data *integration*. Pada tahap *transformation* melibatkan pengubahan data ke dalam format yang sesuai untuk dianalisis menggunakan data mining. Pada tahap *data mining* adalah mengekstrak pengetahuan baru dari data yang telah diproses. Pada penelitian ini teknik *Clustering* yang digunakan adalah metode *K-means*. Pada tahap *evaluation* difokuskan pada pengidentifikasian pola-pola menarik dari basis data yang telah ditentukan. Pada tahap *konowledge* bertujuan agar pengetahuan baru yang diperoleh dari proses ini dapat dipahami oleh semua pihak yang terlibat dan digunakan sebagai dasar untuk pengambilan keputusan.

Selanjutnya teknik pengumpulan data yang dilakukan pada penelitian meliputi (1) wawancara dengan KAPUSDATIN KESOS Dinas Sosial Provinsi DKI Jakarta dilakukan untuk memperoleh informasi mendalam tentang prosedur kerja, tantangan yang dihadapi dan strategi yang diterapkan dalam menangani masalah kemiskinan, (2) sudi pustaka dilakukan dengan memilih, menetapkan, dan mempersiapkan daftar pustaka yang terdiri dari jurnal, buku, makalah, dan skripsi yang memiliki korelasi relevan, (3) tinjauan pustaka dilakukan untuk mengumpulkan informasi dari berbagai jurnal yang relevan.

Proses data menggunakan *Knowledge Discovery in Database* (KDD) melalui beberapa tahapan diantaranya *data selection*, *data processing*, *tranformation*, serta *data mining* [13]. Selanjutnya terkait dengan rancangan basisdata meliputi *logical record structure* dan spesifikasi basis data. Rangcangan pengujian sendiri terbagi menjadi beberapa *poin* diantaranya metode pengujian, tahapan pengujian yang terdiri dari pengujian UI sertapengujian fungsi dasar aplikasi [14].

3. HASIL DAN PEMBAHASAN

3.1 Implementasi Metode

Pada pengujian ini peneliti memakai data sebanyak 14 data hanya sebagai sampel untuk proses perhitungan secara manual perhitungan Clustering *K-means*. Tetapi untuk di sistem yang dibangun peneliti memakai semua data yang akan di uji yaitu sebanyak 800 data. Pada proses perhitungan *Clustering K-means* penulis melakukan perhitungan manual. Perhitungan manual dilakukan menggunakan di program yang telah dibuat. Berikut langkah-langkah penyelesaiannya:

a. Menentukan Jumlah *Cluster*
   Jumlah cluster menyesuaikan dengan kebutuhan analisis dalam penelitian ini jumlah cluster yang akan digunakan sebanyak 4 *cluster*s.
b. Hasil Data *Cleaning*
   Proses ini dilakukan untuk membersihkan data dari elemen-elemen yang tidak diperlukan. Termasuk menghapus nilai yang tidak lengkap, mengurangi gangguan dari data yang berisik (*noise*) dan menangani konsistensi serta keterkaitan dalam data. Kemudian, dari hasil *cleaning* 8000 data diperoleh data sebanyak 14.
c. Menentukan *Centroid* Awal
   Centroid awal merupakan titik pusat *cluster* pertama yang diperoleh secara acak dari data sampel dimana dari 14 data, data yang dipilih data ke-7. Pemilihan titik *centroid* awal akan sangat berpengaruh terhadap hasil *cluster*. Sebagai salah satu contoh, disini peneliti atau penulis akan melakukan perhitungan untuk data ke-1 terhadap data ke-7. Selanjutnya, Tabel 1 memaparkan terkait dengan penentuan dari centroid awal yang dilakukan.

Perhitungan penentuan *centroid*

$= (57-53)^2 + (0-0)^2 + (2-1)^2 + (3-2)^2 + (17-4)^2 + (9-3)^2 + (4-3)^2$

$= 4^2 + 0^2 + 1^2 + 13^2 + 6^2 + 1^2$

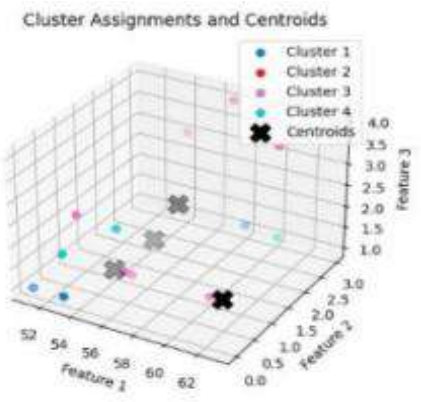
$= 16 + 0 + 1 + 1 + 169 + 36 + 1 = 224$

Tabel 1. Hasil Penentuan Centroid Awal

Centroid	***Feature 1***	***Feature 2***	***Feature 3***	***Feature 4***	***Feature 5***	***Feature 6***	***Feature 7***
C1	53	0	1	2	4	3	3
C2	63	0	2	4	38	7	2
*C*3	54	3	3	1	25	9	1
*C*4	60	3	1	3	31	4	3

d. Menghitung Jarak Data ke Titik Pusat *Cluster*
Sebagai dua contoh, disini peneliti akan melakukan perhitungan untuk data ke-1 terhadap centroid 1 dan centroid 2.

Pada pengujian *cluster* dengan *range* nilai K dilakukan dengan jumlah *cluster* yang akan diuji adalah dari K=4. Pengujian menggunakan data penduduk diperoleh dari Dinas Sosial Provinsi DKI Jakarta periode data 2019. Dari 11 iterasi yang diproses didapatkan hasil *clustering* yaitu iterasi ke-3. Dikarenakan iterasi ke-3 dipilih sebagai yang paling cocok karena beberapa alasan utama. Pertama, centroid tidak banyak berubah lagi setelah iterasi ini, menunjukkan bahwa algoritma telah mencapai stabilitas. *Kedua*, terdapat penurunan inertia yang signifikan, mengindikasikan bahwa jarak total dari titik-titik ke *centroid* mereka telah diminimalkan secara optimal. Terakhir, validasi visual menunjukkan bahwa data telah dikelompokkan dengan baik sesuai dengan centroid yang dihitung. Kombinasi faktor-faktor ini menjadikan iterasi ke-3 sebagai pilihan terbaik dalam analisis *clustering* yang dilakukan. Pada gambar 2 memaparkan terkait dengan visualisasi hasil *clustering*.



Gambar 2. Visualisasi Hasil *Clustering*

Visualisasi hasil *clustering* dalam Gambar 2 yang dilampirkan menunjukkan representasi 3D dari pengelompokan data dengan jelas. Pada gambar tersebut, *centroid* dari setiap *cluster* ditandai dengan simbol 'X', sementara data yang dikelompokkan dalam masing-masing *cluster* ditandai dengan warna yang berbeda. Warna-warna yang berbeda ini memudahkan dalam mengidentifikasi dan membedakan kelompok data yang terbentuk. Visualisasi ini sangat membantu dalam memahami bagaimana data terdistribusi ke dalam *cluster* yang berbeda dan memberikan gambaran yang intuitif tentang struktur internal dari dataset yang dianalisis. Dengan visualisasi 3D ini, hasil *clustering* dapat dievaluasi dengan lebih baik, memastikan bahwa pengelompokan yang dilakukantelah mencerminkan pola yang ada dalam data secara efektif.

3.2 Pengujian Aplikasi

K-means adalah salah satu metode *clustering* yang sering digunakan untuk mengelompokkan data ke dalam beberapa grup berdasarkan kesamaan karakteristik. Untuk aplikasi klasterisasi tingkat kemiskinan, metode ini dapat sangat berguna dalam mengidentifikasi kelompok-kelompok dengan tingkat kemiskinan yang serupa. Pada pengujian aplikasi yang dilakukan dengan penggunaan metode *blackbox testing* memiliki tujuan dalam menghasilkan data yang berkesesuaian antara *output* dan *input* dan pula memiliki fungsi dalam pengujian fungsionalitas [15]. Penulis juga menampilkan grafik dengan kualitas *cluster* yang terbaik menggunakan *Elbow Method.* Tabel 2 dibawah ini memaparkan terkait dengan hasil pengujian *blackbox testing dashboard* yang dihasilkan.

Tabel 2. *Blackbox Testing Dashboard*

No	***Field***	***Scenario***	***Expected Result***	**Status**
1	Halaman *Dashboard*	Klik halaman *Dashboard*	*Success* sistem menampilkan halaman *Dashboard*	OK
2	Halaman *K-Means*	Klik halaman *K-Means*	*Success* sistem menampilkan halaman *K-Means Dataset*	OK

Selanjutnya pada pengujian *blackbox testing* pada halaman *K-Means* dihasilkan data yang tersajikan pada Tabel 3. Kemudian hasil dari pengujian *blackbox testing* pada halaman *K-Means Dataset* yang diperoleh dari penelitian disajikan dalam Tabel 4.

Tabel 3. *Blackbox Testing K-Means* dengan format

No	*Field*	*Scenario*	*Expected Result*	Status
1	Halaman *K-Means*	Klik *Choose File*	*Success* sistem menampilkan file *excel xlxs*	OK
2	Halaman *K-Means*	Pilih *file excel* dengan format selain *xlxs*	*Succes* sistem menampilkan pop up "*Uploadfile excel* dengan format *xlxs"*	OK
3	Halaman *K-Means*	Pilih *file excel* dengan format *xlxs*	*Succes* sistem menampilkan *file excel* dengan format *xlxs*	OK
4	Halaman*K-Means*	*Upload file excel*	*Success* sistem menampilkan data yang telah di *upload*	OK
5	Halaman *K-Means*	Klik tombol button *use this data*	*Succes* sistem menampilkan halaman *K-Means Dataset*	OK

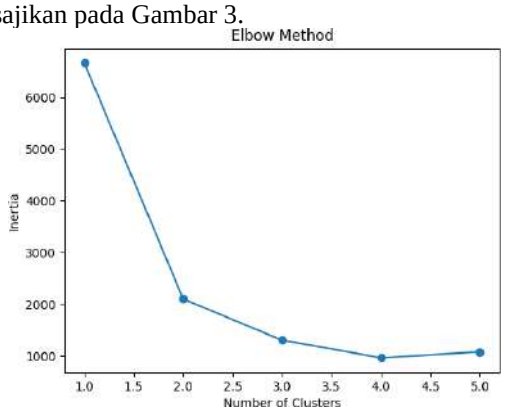
Tabel 4. *Blackbox testing K-Means Dataset*

No	*Field*	*Scenario*	*Expected Result*	Status
1	Halaman *Cluster Config*	*Input* pada *clusters* dan *Maximum Iteranios*	*Success* sistem*input clusters* & *Maximum Iteranios*	OK
2	Halaman *Cluster Config*	Klik tombol button *process* pada data yang telah di *input*	*Succes* sistem menampilkan halaman *Process*	OK
3	Halaman *Process*	Klik button halaman *Process*	*Succes* menampilkan *exploring the K-Means Clustering Methodology*	OK
4	Halaman *Elbow*	*Input* pada *MaxClusters*	*Succes* sistem *Input MaxClusters*	OK
5	Halaman *Elbow*	Klik tombol *button process*	*Succes* sistem menampilkan grafik *ElbowMethod*	OK
6	Halaman *Clstering*	Klik *button* halaman *Clsutering*	*Succes* sistem menampilkan *Clustering Result*	OK

Tabel 5. *Blackbox testing History*

No	*Field*	*Scenario*	*Expected Result*	Status
1	Menu *History*	Klik menu *History*	*Success* sistem menampilkan halaman *History*	OK
2	Halaman *K-Means*	Klik tombol *button detail*	*Success* sistem menampilkan halaman *DetailHistory K-MeansClustering*	OK

Pada tahap ini akan menjelaskan hasil dari pengujian hasil *Elbow Method* pada halaman *Elbow* disajikan pada Gambar 3.



Gambar 3. Hasil *Elbow Method*

Hasil dari pengujian *blackbox testing* pada menu *History*, yang diperoleh dari penelitian disajikan dalam Tabel 5.

Pengujian kualitas *cluster* diterapkan menggunakan data penduduk Provinsi DKI Jakarta memiliki jumlah 800 data, menguji kualitas *cluster* dengan 4 *cluster* menggunakan metode *Clustering K-means* dengan kategori/label yang disajikan pada Tabel 6.

Tabel 6. *Cluster Label*

Cluster	*Label*
1	Tidak Miskin
2	Sedang Miskin
3	Miskin
4	Sangat Miskin

Pengujian untuk menentukan jumlah atau nilai *cluster* terbaik menggunakan metode *Clustering K-means* dengan jumlah *cluster d*imana hasil *cluster* terbaik akan menjadi acuan penggunaan *cluster* untuk penelitian ini. Langkah selanjutnya adalah menentukan *cluster* terbaik dengan label yang telah diperoleh dari *result Clustering*. Pada penduduk dengan label "Tidak Miskin" yang disajikan pada tabel 7 adalah mereka yang tidak mengalami kesulitan ekonomi. Mereka memiliki pendapatan yang cukup dalam pemenuhan kebutuhan primer. Mereka mampu hidup dengan nyaman dan tidak perlu khawatir tentang kebutuhan sehari-hari.

Tabel 7. Detail *Cluster* 1

Nama	Ijazah Tertinggi	Jenis Cacat	Status Pekerjaan	Jumlah Jam Kerja	Lapangan Usaha	Penyakit Kronis
Marwiyah, 60 Tahun	M. Ibtidaiyah	Tuna netra/buta	Berusaha dibantu buruh tidak tetap/ tidak dibayar	4	8	Hipertensi (tekanan darah tinggi)
Abd Manaf, 53 Tahun	SD/SDLB	Tuna daksa/ cacat tubuh	Berusaha dibantu buruh tetap/dibayar	4	3	Asma
Juhri, 51 Tahun	SD/SDLB	Tuna daksa/ cacat tubuh	Buruh/ karyawan/ pegawai swasta	3	4	Masalah jantung
Aggregasi Nilai Tertinggi/Terbanyak	SD/SDLB	Tuna daksa/ cacat tubuh	Berusaha dibantu buruh tidak tetap	4	8	Hipertensi

Tabel 8. Detail *Cluster* 2

Nama, Umur	Ijazah Tertinggi	Jenis Cacat	Status Pekerjaan	Jumlah Jam Kerja	Lapangan Usaha	Penyakit Kronis
IBRAHIM, 63 Tahun	SD/SDLB	Tuna netra/buta	PNS/ TNI/ Polri/ BUMN/ BUMD/ anggota legislatif	38	7	Rematik
Aggregasi Nilai Tertinggi/Terbanyak	SD/SDLB	Tuna netra/buta	PNS/ TNI/ Polri/ BUMN/ BUMD/ anggota legislatif	38	7	Rematik

Tabel 9. Detail *Cluster* 3

Nama, Umur	Ijazah Tertinggi	Jenis Cacat	Status Pekerjaan	Jumlah Jam Kerja	Lapangan Usaha	Penyakit Kronis
Ruslan, 57 Tahun	SD/SDLB	Tuna netra/buta	Buruh/ karyawan/ pegawai swasta	17	9	Masalah jantung
Nursaman BinAbd Kadir, 54 Tahun	SMP/ SMPLB	Tuna rungu	Berusaha dibantu buruh tidaktetap/ tidak dibayar	25	9	Hipertensi (tekanan darah tinggi)
Sarifudin, 55 Tahun	Paket A	Tuna daksa/ cacat tubuh	Berusaha dibantu buruh tetap/dibayar	20	5	Rematik
Junaidi, 60 Tahun	Paket A	Tuna daksa/ cacat tubuh	Buruh/ karyawan/ pegawai swasta	16	1	Rematik
Mat Jamin, 57 Tahun	SMP /SMPLB	Tuna wicara	Buruh/ karyawan/ pegawai swasta	25	7	Asma
Hasanah, 62 Tahun	M. Ibtidaiyah	Tuna wicara	PNS/ TNI/ Polri/ BUMN/ BUMD/ anggota legislatif	20	7	Asma
Uto Masjoyo, 54 Tahun	SD/SDLB	Tuna rungu	Berusaha dibantu buruh tetap/ dibayar	25	7	Asma
Aggregasi Nilai Tertinggi/Terbanyak	SD/SDLB	Tuna daksa/ cacat tubuh	Buruh/ karyawan/ pegawai swasta	25	9	Asma

Tabel 10. Detail *Cluster* 4

Nama, Umur	Ijazah Tertinggi	Jenis Cacat	Status Pekerjaan	Jumlah Jam Kerja	Lapangan Usaha	Penyakit Kronis
Rausin, 54 Tahun	Paket A	Tuna netra/buta	Berusaha dibantu buruh tidaktetap/ tidak dibayar	34	2	Masalah jantung
Sartini, 60 Tahun	SMP/SMPLB	Tuna daksa/ cacattubuh	Buruh/ karyawan/ pegawai swasta	31	4	Asma
Husaini, 53 Tahun	SD/SDLB	Tuna netra/buta	Buruh/ karyawan/ pegawai swasta	29	5	Asma
Aggregasi Nilai Tertinggi/Terbanyak	SD/SDLB	Tuna netra/buta	Buruh/ karyawan/ pegawai swasta	34	5	Asma

Pada penduduk dengan label “Miskin” adalah *cluster* menghadapi kesulitan yang signifikan dalam memenuhi kebutuhan dasar mereka. Mereka sering kali berada dalam kondisi ekonomi yang rentan dan

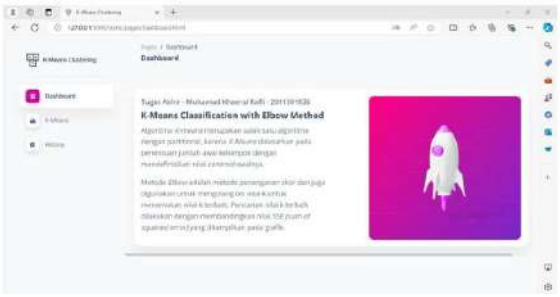
memerlukan bantuan dari luar untuk mencukupi kebutuhan hidup sehari- hari. Kesejahteraan mereka terganggu oleh pendapatan yang tidak mencukupi dan kondisi hidup yang kurang layak. Berikut *cluster* 3 dapat dilihat pada Tabel 9.

Pada penduduk dengan label "Sangat Miskin" adalah *cluster* berada dalam kondisi ekonomi yang sangat sulit. Mereka tidak mampu memenuhi kebutuhan dasar tanpa bantuan signifikan dari pemerintah atau organisasi sosial. Kehidupan mereka sangat dipengaruhi oleh kekurangan ekonomi yang ekstrem, dan mereka sering kali menghadapi masalah kesehatan, pendidikan, dan tempat tinggal yang serius. Hasil *cluster* 4 dapat dilihat pada Tabel 10.

Dari hasil pengujian kualitas *cluster*, *cluster* terbaik adalah *Cluster* 3 (Miskin). Hal ini karena *cluster* ini memberikan gambaran yang jelas mengenai penduduk yang berada dalam kondisi ekonomi yang rentan dan memerlukan perhatian khusus dari Dinas Sosial Provinsi DKI Jakarta. Penduduk dalam *cluster* ini sering kali membutuhkan intervensi yang lebih intensif untuk membantu mereka keluar dari kemiskinan dan mencapai kesejahteraan yang lebih baik. Dengan mengidentifikasi *cluster* ini sebagai *cluster* terbaik, dapat diambil tindakan yang lebih tepat sasaran dalam program bantuan sosial untuk meningkatkan kualitas hidup penduduk yang termasuk dalam kategori miskin

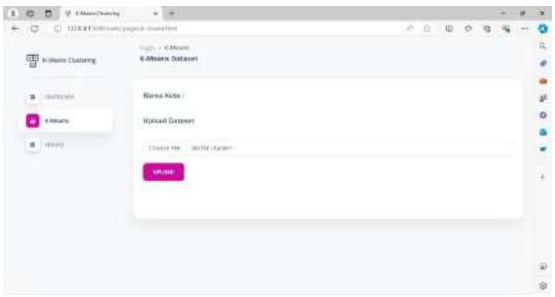
3.3 Tampilan Layar

Tampilan layar *dashboard* pada halaman *dashboard* ini terdapat dua navbar yang tersedia: *Dashboard* dan *K-Means* yang disajikan pada gambar 4.



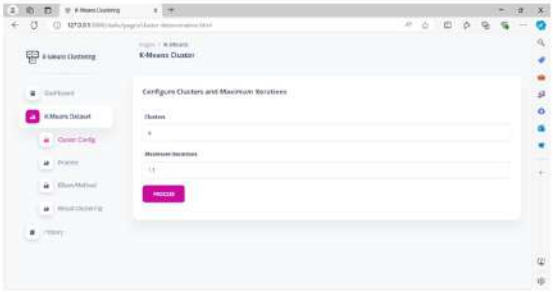
Gambar 4. Tampilan Layar *Dashboard*

Tampilan layar *K-Means* muncul setelah pengguna mengklik menu *K-Means*. Pengguna dapat mengunggah data dalam format *Excel "xlsx"* melalui tombol *"Choose File"* kemudian mengklik tombol *"Upload"* yang disajikan pada gambar 5.



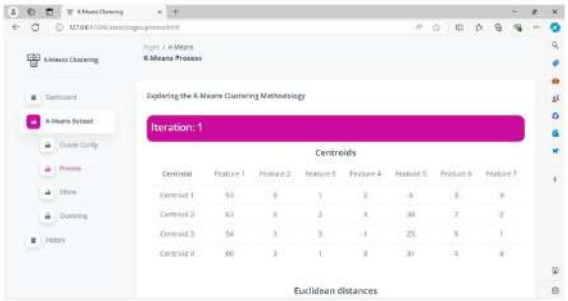
Gambar 5. Tampilan Layar *K-Means*

Tampilan layar *Cluster Config* merupakan sub menu dari menu *K-Means Dataset*. Saat *user* klik button *Use This Data* akan menampilkan 2 kolom yaitu *Clusters* dan *Maximum Iterations*. Langkah selanjutnya user mengisi 2 kolom tersebut untuk menentukan *Clusters* dan *Maximum Iterations* yang disajikan pada gambar 6.



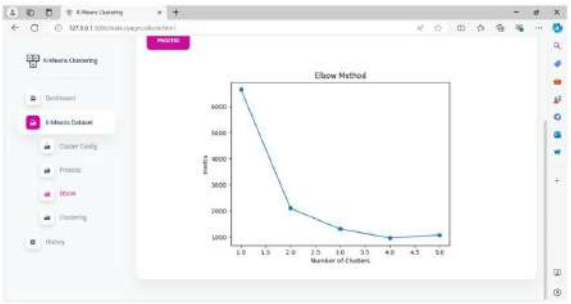
Gambar 6. Tampilan Layar *K-Means Dataset (Cluster Config)*

Tampilan layar *Process* merupakan *sub-menu* dari menu *K-Means*. Ketika *user* telah menentukan *Clusters* dan *Maximum Iterations* langkah selanjutnya user klik button *process* pada menu *Cluster Config* setelah itu akan menampilkan halaman menu *Process* pada gambar 7.



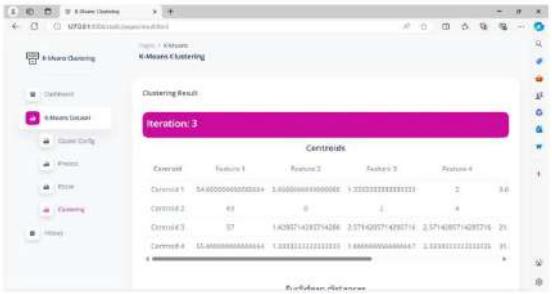
Gambar 7. Tampilan *Layar K-Means Dataset (Process)*

Tampilan layar *Elbow* merupakan *sub-menu* dari menu *K-Means.* Selanjutnya *user* klik menu *Elbow* dan mengisi kolom *Max Clusters* klik button *process.* Menampilkan grafik *Elbow* yang disajikan pada gambar 8.



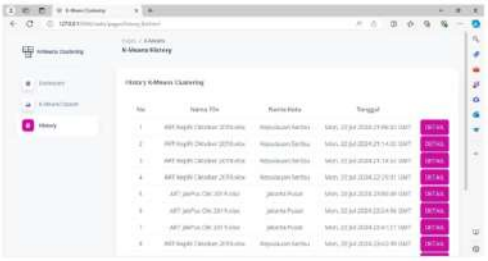
Gambar 8. Tampilan Layar *K-Means Dataset (Elbow)*

Tampilan layar *Clustering* merupakan *sub-menu* dari menu *K-Means.* Selenjutnya *user* klik menu *Clustering.* Menampilkan halaman *Clustering Result* pada Gambar 9.



Gambar 9. Tampilan Layar *K-Means Dataset (Clustering)*

Tampilan layar *History* adalah hasil *clustering*. Selanjutnya user klik history. Menampilkan halaman history yang disajikan pada Gambar 10.



Gambar 10. Tampilan Layar *History*

Hasil dari penelitian senada dengan penelitian terdahulu yang memaparkan metode *K-means* dapat digunakan dalam klasterisasi tingkat kemiskinan. Pada penelitian terdahulu menjelaskan provinsi yang termasuk dalam *cluster* 0 atau provinsi yang memiliki tingkat kemiskinan rendah diantaranya adalah Gorontalo, Sulawesi Tengah, BTT, NTB, DI Yogyakarta, Jawa Tengah, Lampung, Bengkulu, dan Aceh. Selanjutnya provinsi yang termasuk dalam *cluster* 1 atau provinsi yang memiliki tingkat kemiskinan sedang memiliki sejumlah 21 provinsi di Indonesia. Terakhir, *cluster* 2 yaitu provinsi yang memiliki tingkat kemiskinan tinggi diantaranya adalah Papua, Papua Barat, dan Maluku [9].

4. KESIMPULAN

Kesimpulan yang diambil berdasarkan paparan diatas ialah penggunaan metode *Clustering K-Means* dapat diterapkan untuk mengelompokan *cluster* kemiskinan di Provinsi DKI Jakarta. Pada proses pengujian terhadap 8000 data penduduk DKI Jakarta. Didapatkan hasil 14 data. Pengujian dilakukan dengan jumlah *cluster* yang diuji *K* = 4. Menghitung jarak setiap data dengan *centroid* awal menggunakan persamaan *Euclidean Distance*. Hasil yang diperoleh *centroid* ke-*1* 14,97 dan *centroid* ke-*2* 22,05. Dari 11 *iterasi* yang diproses didapatkan hasil *clustering* yaitu *iterasi* ke-3. Dikarenakan *iterasi* ke-3 dipilih sebagai yang paling cocok karena beberapa alasan utama. Pertama, *centroid* tidak banyak berubah lagi setelah *iterasi* ini, menunjukkan bahwa algoritma telah mencapai stabilitas. Kedua, terdapat penurunan *inertia* yang signifikan, mengindikasikan bahwa jarak total dari titik-titik ke *centroid* mereka telah diminimalkan secara optimal. Terakhir, validasi visual menunjukkan bahwa data telah dikelompokkan dengan baik sesuai dengan *centroid* yang dihitung.

DAFTAR PUSTAKA

[1] F. D. Hardini, A. K. Joy, B. Lidya Maharani, A. Rizky Airlangga, and C. Asnanti, "Kebijakan Pajak sebagai Upaya Pengentasan Kemiskinan," *J. Krit. Stud. Huk.*, vol. 9, no. 5, pp. 203–208, 2024, [Online]. Available: https://ojs.co.id/1/index.php/jksh/article/view/1245.

[2] I. N. Pratama, "Analisis Determinan Tingkat Kemiskinan Di Kabupaten Sumbawa," *J. Law Gov.*, vol. 11, no. 2, pp. 143–153, 2023, doi: 10.58406/jeb.v11i2.1314.

[3] P. I. Lestari, B. Robiani, and Sukanto, "Kemiskinan Ekstrem, Ketimpangan Dan Pertumbuhan Ekonomi Di Indonesia," *J. Ilm. Ekon. dan Bisnis*, vol. 11, no. 2, pp. 1739–1752, 2023, [Online]. Available: https://jurnal.unived.ac.id/index.php/er/article/view/4789.

[4] A. Natalia and E. N. Maulidya, "Aktualisasi Empat Pilar Sustainable Development Goals (SDGs) Di Perdesaan Kecamatan Natar Kabupaten Lampung Selatan," *JIIP J. Ilm. Ilmu Pemerintah.*, vol. 8, no. 1, pp. 21–41, 2023, doi: 10.14710/jiip.v8i1.16513.

[5] U. Kencana, Yuswalina, and T. Eza, "Efektivitas Peraturan Daerah yang Berkesejahteraan Sosial di Kota Palembang: Studi Kasus Anak Jalanan, Gelandangan dan Pengemis di Masa Pandemi Covid-19," *Simbur Cahaya*, vol. 27, no. 2, pp. 70–97, 2021, doi: 10.28946/sc.v27i2.1039.

[6] R. Nasution and M. Marliyah, "Analisis Program Pemerintah Dalam Penanggulangan Kemiskinan Dan Pengangguran Di Kecamatan Pulau Rakyat Kabupaten Asahan," *Jesya*, vol. 6, no. 1, pp. 810–823, 2023, doi: 10.36778/jesya.v6i1.1031.

[7] S. Suhartini and R. Yuliani, "Penerapan Data Mining untuk Mengcluster Data Penduduk Miskin Menggunakan Algoritma K-Means di Dusun Bagik Endep Sukamulia Timur," *Infotek J. Inform. dan Teknol.*, vol. 4, no. 1, pp. 39–50, 2021, doi: 10.29408/jit.v4i1.2986.

[8] N. Sepriyanti, R. Sani Nahampun, M. H. Zikri, I.

Ambarani, and A. Rahmadeyan, "Implementation of K-Means Clustering to Group Poverty Levels in Riau Province," *Semin. Nas. Penelit. dan Pengabdi. Masy.*, pp. 59–65, 2022, [Online]. Available: https://journal.irpi.or.id/index.php/sentimas.

[9] A. Bahauddin, A. Fatmawati, and F. Permata Sari, "Analisis Clustering Provinsi Di Indonesia Berdasarkan Tingkat Kemiskinan Menggunakan Algoritma K-Means," *J. Manaj. Inform. dan Sist. Inf.*, vol. 4, no. 1, pp. 1–8, 2021, doi: 10.36595/misi.v4i1.216.

[10] A. Khalif, A. N. Hasanah, M. H. Ridwan, and B. N. Sari, "Klasterisasi Tingkat Kemiskinan di Indonesia menggunakan Algoritma K-Means," *Gener. J.*, vol. 8, no. 1, pp. 54–62, 2024, doi: 10.29407/gj.v8i1.21470.

[11] N. Oktaviany, N. Suarna, and W. Prihartono, "Implementasi Algoritma K-Means Clustering Dalam Mengelompokkan Kepadatan Penduduk Di Provinsi Dki Jakarta," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 8, no. 1, pp. 119–126, 2024, doi: 10.36040/jati.v8i1.8241.

[12] N. A. Maori and E. Evanita, "Metode *Elbow* dalam Optimasi Jumlah Cluster pada K-Means Clustering," *Simetris J. Tek. Mesin, Elektro dan Ilmu Komput.*, vol. 14, no. 2, pp. 277–288, 2023, doi: 10.24176/simet.v14i2.9630.

[13] M. A. S. Fazrin, A. Voutama, and Y. Umaidah, "Penerapan Algoritma C4.5 Untuk Prediksi Penyakit Paru-Paru Menggunakan Rapidminer," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 7, no. 2, pp. 1409–1415, 2023, doi: 10.36040/jati.v7i2.6810.

[14] N. K. Ningrum, I. U. W. Mulyono, and Z. Umami, "Pengujian UI/UX dengan System Usability Scale dan Single Ease Question pada Aplikasi Pantau untuk Monitoring Perkembangan Penanaman Tanaman di Lahan Hijau," *Proceeding Sci. Eng. Natl. Semin. 7 (SENS 7)*, pp. 9–16, 2022, [Online]. Available: https://conference.upgris.ac.id/index.php/sens/article/view/3466%0Ahttps://conference.upgris.ac.id/index.php/sens/article/download/3466/2146.

[15] S. Catrio, M. Rachmanto, H. A. Agatha, T. Ramdani, and A. Yusuf, "Pengujian Aplikasi Sapawarga (Jabar Super Apps) Menggunakan Metode Black Box Testing," *J. Inform. dan Tek. Elektro Terap.*, vol. 12, no. 3, pp. 2319–2334, 2024.