

PENERAPAN ALGORITMA K-MEANS UNTUK KLASTERISASI PENDUDUK MISKIN DI PROVINSI BANTEN

Friska Handayanna

Informatika

Universitas Nusa mandiri

www.nusamandiri.ac.id

frisma.fha@nusamandiri.ac.id



Abstract— People with low incomes are unable to obtain education and other government services. The problem of poverty faced by the government is closely related to people with low incomes who cannot meet their basic needs. The Central Bureau of Statistics describes poverty as the inability to meet basic food and non-food needs as measured by expenditure. This study aims to classify Banten province based on poverty levels, by dividing the number of poor people into high, medium, and low categories. The K-Means clustering method is very fast and easy to use in the K-Means algorithm clustering process. Where the grouping results are formed, namely group one has a moderate number of poor people in three districts/cities, Pandeglang Regency, Lebak Regency, and Tangerang Regency. The second group has the lowest population in one district/city, namely Tangerang City. The third group has the highest number of poor people in the four districts/cities, namely Serang Regency, Cilegon Regency, Serang City, and South Tangerang City. The clustering results show that the Provincial Government of Banten will give priority and special attention to poverty alleviation efforts in the district/city. This will allow for increased revenues and earnings, as well as improved livelihoods and the economy in the area. the K-Means algorithm can classify the poor based on the number of people per district or city in Banten Province.

Keywords: Clustering, Poor Residents, K-Means

Abstrak—Orang-orang dengan pendapatan rendah tidak dapat mendapatkan pendidikan dan layanan pemerintah lainnya. Masyarakat dengan pendapatan rendah yang tidak dapat memenuhi kebutuhan dasar sangat erat terkait dengan masalah kemiskinan yang dihadapi pemerintah. Menurut Badan Pusat Statistik, kemiskinan didefinisikan sebagai ketidakmampuan untuk memenuhi kebutuhan dasar seseorang baik pangan maupun non-pangan, yang diukur dari pengeluaran mereka. Tujuan dari penelitian ini adalah untuk mengklasifikasikan provinsi Banten berdasarkan tingkat kemiskinan, membagi populasi ke dalam kategori tinggi, sedang, dan rendah. Metode clustering K-Means sangat cepat dan mudah digunakan dalam proses clustering algoritma K-Means. Dimana hasil pengelompokan terbentuk yaitu kelompok satu memiliki jumlah penduduk miskin sedang di tiga kabupaten/kota yaitu Kab Pandeglang, Kab Lebak, dan Kab Tangerang. Untuk kelompok kedua memiliki jumlah penduduk miskin terendah di satu kabupaten/kota, yaitu Kota Tangerang. Kelompok ketiga memiliki jumlah penduduk miskin tertinggi di empat kabupaten/kota, yaitu Kabupaten Serang, Kabupaten Cilegon, Kota Serang, dan Kota Tangerang Selatan. Hasil kelompok menunjukkan bahwa pemerintah Provinsi Banten akan memberikan prioritas dan perhatian khusus pada upaya pengentasan kemiskinan di kabupaten/kota tersebut. Ini akan memungkinkan peningkatan pendapatan dan pendapatan, serta peningkatan penghidupan dan perekonomian di daerah tersebut. algoritma K-Means dapat mengelompokkan penduduk miskin berdasarkan jumlah orang per kabupaten atau kota di Provinsi Banten.

Kata kunci: Clustering, Penduduk Miskin, K-Means

PENDAHULUAN

Rendahnya pendapatan berdampak pada ketidakmampuan untuk mendapatkan akses ke pendidikan dan fasilitas pemerintah lainnya. Rendahnya pendapatan masyarakat sangat erat kaitannya dengan masalah kemiskinan yang dihadapi pemerintah (Sari, 2021). Di banyak negara, terutama negara berkembang, kemiskinan

adalah masalah jangka panjang (Pratiwi et al., 2020). Kemiskinan adalah masalah besar, terutama di negara-negara berkembang seperti Indonesia. Kemiskinan di suatu wilayah akan berdampak negatif pada kemajuan negara tersebut dalam jangka panjang (Nugraha, 2023). Kemiskinan adalah masalah di banyak negara, termasuk Indonesia. Angka kemiskinan di Indonesia terus meningkat karena masyarakat tidak dapat

memenuhi kebutuhan pokok mereka seperti pakaian, makanan, dan tempat tinggal (Aprilia & Sembiring, 2021). Menurut BPS, ketidakmampuan seseorang untuk memenuhi kebutuhan pokok mereka, baik makanan maupun bukan makanan, disebut kemiskinan (Setiawan & Zahra, 2023). Tingkat pendapatan nasional yang rendah, laju pertumbuhan ekonomi yang lambat, rendahnya pendapatan perkapita, distribusi pendapatan yang tidak merata, pendidikan yang rendah dan layanan kesehatan yang kurang memadai adalah beberapa penyebab kemiskinan (Mirawati & Feriyanto, 2023). Pada September 2022, jumlah penduduk miskin Provinsi Banten adalah 829,66 ribu, meningkat 15,64 ribu dari Maret 2022 dan turun 22,62 ribu dari September 2021 (BPS, 2023).

Klasterisasi adalah cara untuk mengelompokkan data ke dalam kesamaan tertentu. Ini berbeda dengan klasifikasi, yang merupakan proses pengelompokan data baru berdasarkan kelompok atau klasifikasi yang sudah ada. Dalam clustering, data baru dikelompokkan berdasarkan fitur yang sama (Febriansyah & Muntari, 2023)

Proses kelompokkan data yang sangat mirip satu sama lain ke dalam klaster atau kelompok dikenal sebagai clustering. Tujuan dari clustering sendiri adalah untuk memastikan bahwa Data yang dikumpulkan dalam satu klaster memiliki tingkat kemiripan tertinggi dan bahwa tingkat kemiripan data antara klaster paling rendah (Arkham & Swanjaya, 2020). Algoritma K-Means digunakan untuk mengklasifikasikan rumah tangga miskin berdasarkan provinsi. Studi ini dapat menentukan alasan mengapa setiap provinsi di Indonesia memiliki jumlah penduduk miskin dan memberikan informasi atau informasi kepada pemerintah untuk melakukan hal-hal yang tepat untuk mengurangi jumlah penduduk miskin tersebut (Sembiring & Saifullah, 2021).

Selain itu, mengelompokkan data menggunakan K-Means clustering, yang memaksimalkan kemiripan data antara cluster dan mengurangi kemiripan data antara cluster; ukuran kemiripan cluster didasarkan pada jarak terkecil antara titik pusat data (Aziz et al., 2018). Salah satu metode pengelompokan yang paling sederhana, algoritma K-Means, dapat digunakan dengan mudah dan melakukan proses pengelompokan dengan sangat cepat. Algoritma ini mengelompokkan data-data yang dimasukkan ke dalam komputer tanpa mengetahui target kelasnya (Wanto et al., 2020).

Tingkat persen kemiskinan di setiap provinsi Indonesia dapat dikelompokkan dengan menggunakan metode clustering, yang merupakan bagian dari data mining (Bahauddin et al., 2021).

Karakteristik wilayah dapat dikelompokkan berdasarkan tingkat pendidikan penduduknya, jumlah penduduk miskin, dan komponen tambahan

yang terkait dengan indikator kemiskinan yang ditentukan oleh publikasi data kemiskinan Kota atau Kabupaten yang disajikan oleh Badan Pusat Statistik. Data untuk penelitian ini berasal dari tahun 2015 hingga 2019 (periode Maret). Hasil dari tiga kelompok ini dihasilkan (Sari et al., 2020)

Selama sebelas tahun terakhir (2010–2020), data kemiskinan di delapan Untuk memprofilisasi centroid, kabupaten/kota di Provinsi Banten menggunakan rata-rata. Kabupaten Pandeglang, Lebak, dan Serang termasuk dalam "Kelompok dengan tingkat kemiskinan sedang", sedangkan Kota Tangerang dan Kabupaten Tangerang dimasukkan dalam "Kelompok dengan tingkat kemiskinan tinggi". Oleh karena itu, dalam pengentasan kemiskinan, yang paling penting adalah kelompok yang tinggal di wilayah dengan tingkat kemiskinan tertinggi (Munandar, 2022).

Penelitian yang saat ini dilakukan berdasarkan data Jumlah orang miskin di setiap kota atau kabupaten di Provinsi Banten dari tahun 2020 hingga 2022 dilaporkan oleh BPS. Metode K-Means Clustering digunakan untuk mengelompokkan populasi miskin di wilayah Banten. Tujuan dari studi ini adalah untuk mengelompokkan semua provinsi Banten berdasarkan tingkat kemiskinan mereka, dengan mengelompokkan populasi mereka e dalam kategori tingkat tinggi, tingkat sedang, dan tingkat rendah. Hasil dari kelompok ini akan menunjukkan kabupaten atau kota mana di provinsi Banten yang harus memberikan prioritas dan perhatian khusus untuk mengatasi kemiskinan agar pendapatan dan penghasilan naik, serta penghidupan dan perekonomian meningkat.

BAHAN DAN METODE

Badan Pusat Statistik menyediakan data untuk penelitian ini dan diproses untuk menentukan kemiskinan di Provinsi Banten berdasarkan jumlah penduduk miskin per kabupaten atau kota. Metode penelitian yang dilakukan yaitu terdapat pada Gambar 1.



Sumber: (Handayanna, 2023)

Gambar 1. Metode Penelitian

A. Tahap Pengumpulan Data

Untuk penelitian ini, Data jumlah penduduk miskin menurut kabupaten/kota di Provinsi Banten dari tahun 2020 sampai dengan tahun 2022 diambil dari Publikasi Statistik Indonesia dan diolah oleh Badan Pusat Statistik Nasional .

B. Tahap Pengolahan Data

Untuk memulai proses clustering dengan metode K-Means, pengolahan data diperlukan terlebih dahulu. Untuk membentuk satu data yang baru dengan mengelompokkan sejumlah data yang sama, pengelompokan adalah teknik data mining (Amanda & Sitorus, 2021). K-Means mengelompokkan data dalam satu atau lebih kelompok atau cluster non-hirarki (Hablum et al., 2019). Data tentang jumlah orang miskin dikategorikan berdasarkan kabupaten atau kota di provinsi Banten.

C. Tahap Analisis

Pada tahapan analisis menggunakan tools RapidMiner untuk menganalisis data jumlah orang miskin menurut kabupaten atau kota di Provinsi Banten yang telah dikumpulkan. Selanjutnya, data yang dikumpulkan diolah dengan perhitungan bobot masing-masing indeks. Pada langkah sebelumnya, telah diputuskan bahwa data akan dikelompokkan ke dalam tiga kelompok, yaitu grup dengan tingkat pendistribusian rendah, grup dengan tingkat pendistribusian tinggi, dan grup dengan tingkat pendistribusian sedang. Pada langkah ini, hasilnya akan dianalisis.

HASIL DAN PEMBAHASAN

A. Perhitungan Dengan Algoritma K-Means

Data jumlah penduduk miskin per kota/kabupaten Provinsi Banten dari tahun 2020 hingga 2022 dikelompokkan dengan menggunakan K-Means Clustering, yang melibatkan sampel data jumlah penduduk miskin untuk masing-masing kota atau kabupaten. Perhitungan yang diperlukan untuk melakukan proses kelompokan dilakukan menggunakan algoritma K-Means, dan proses ini dimulai dengan menentukan data yang ingin dikelompokkan untuk mendapatkan hasil penelitian. Data sampel terdapat pada Tabel 1.

Tabel 1. Jumlah Penduduk Miskin Menurut Kabupaten/Kota di Provinsi Banten (Ribuan Jiwa)

No	Kabupaten/Kota	2020	2021	2022
1.	Kab Pandeglang	120,44	131,43	114,65
2.	Kab Lebak	120,83	134,75	117,22
3.	Kab Tangerang	242,16	272,35	270,52
4.	Kab Serang	74,8	83,09	75,45
5.	Kota Tangerang	118,22	134,24	132,88
6.	Kota Cilegon	16,31	18,89	16,46
7.	Kota Serang	42,24	47,91	42,56
8.	Kota Tangerang Selatan	40,99	44,57	44,29

Sumber: <https://banten.bps.go.id>

Perhitungan dilakukan menggunakan algoritma K-Means dalam dua tahap:

1. Pada tahap ini, hasil jumlah cluster sebanyak 3 cluster dan jumlah data sebanyak 8 data dikumpulkan.
2. Nilai centroid ditetapkan secara acak.

Tabel 2. Centroid Data Awal

Kabupaten/Kota	Tahun 2020	Tahun 2021	Tahun 2022
C1 (Kab Pandeglang)	120,44	131,43	114,65
C2 (Kab Lebak)	120,83	134,75	117,22
C3 (Kab Tangerang)	242,16	272,35	270,52

Sumber: <https://banten.bps.go.id>

3. Menghitung centroid terdekat:

Setiap record data akan ditentukan pusat kelompok terdekatnya. Pusat kelompok terdekat adalah kelompok yang terdekat dengan rekaman tersebut. Dengan menggunakan persamaan berikut, hitung jarak antara setiap data terhadap setiap pusat cluster:

$$D_{xy} = \sqrt{(X_1 - Y_1)^2 + (X_n - Y_n)^2} \dots \dots \dots (1)$$

Menghitung centroid terdekat :

Kab Pandeglang

D11

$$D11 = \sqrt{(120,44 - 120,44)^2 + (131,43 - 131,43)^2 + (114,65 - 114,65)^2}$$

D11 = 0,00

D12

$$D12 = \sqrt{(120,83 - 120,44)^2 + (134,75 - 131,43)^2 + (117,22 - 114,65)^2}$$

D12 = 4,22

D13

$$D13 = \sqrt{(242,16 - 120,44)^2 + (272,35 - 131,43)^2 + (270,52 - 114,65)^2}$$

D13=242,84

Kab Lebak

D21

$$D21 = \sqrt{(120,83 - 120,44)^2 + (134,75 - 131,43)^2 + (117,22 - 114,65)^2}$$

D21=4,22

D22

$$D22 = \sqrt{(120,83 - 120,83)^2 + (134,75 - 134,75)^2 + (117,22 - 117,22)^2}$$

D22=0

D23

$$D23 = \sqrt{(120,83 - 242,16)^2 + (134,75 - 272,35)^2 + (117,22 - 270,52)^2}$$

D23=239,07

Kab Tangerang

D31

$$D_{31} = \sqrt{(242,16 - 120,44)^2 + (272,35 - 131,43)^2 + (270,52 - 114,65)^2}$$

D31=242,84

D32

$$D_{32} = \sqrt{(242,16 - 120,83)^2 + (272,35 - 134,75)^2 + (270,52 - 117,22)^2}$$

D32=239,07

D33

$$D_{33} = \sqrt{(242,16 - 242,16)^2 + (272,35 - 272,35)^2 + (270,52 - 270,52)^2}$$

D33=0,00

Pengelompokan dapat dibuat dengan menggunakan iterasi pertama pada untuk menghitung hasil dari perhitungan keseluruhan yang terdapat pada Tabel 3.

Tabel 3. Hasil Perhitungan Jarak Pusat Cluster Iterasi 1

No	Kabupaten/kota	Cluster			Jarak Terpendek
		C1	C2	C3	
1.	Kab Pandeglang	0,00	4,22	242,84	0,00
2.	Kab Lebak	4,22	0,00	239,07	0,00
3.	Kab Tangerang	242,84	239,07	0,00	0,00
4.	Kab Serang	77,18	80,82	319,19	77,18
5.	Kota Tangerang	18,58	15,88	231,04	15,88
6.	Kota Cilegon	182,07	185,74	424,02	182,07
7.	Kota Serang	138,89	138,89	377,24	138,89
8.	Kota Tangerang Selatan	378,86	140,80	378,86	140,80

Sumber: (Handayanna, 2023)

Selanjutnya, posisi data jumlah penduduk miskin Provinsi Banten ditetapkan berdasarkan jarak minimum data terhadap pusat cluster. Kelompok tersebut akan menggabungkan data yang berada di jarak terkecil dari centroid. Tabel 4 menunjukkan posisi data dengan tiap cluster pada iterasi pertama, dengan tanda (1) yang menunjukkan bahwa data termasuk dalam cluster.

Tabel 4. Posisi Data Dengan Tiap Cluster Pada Iterasi Ke-1

No	Kabupaten/kota	Cluster		
		C1	C2	C3
1.	Kab Pandeglang	1		
2.	Kab Lebak		1	

No	Kabupaten/kota	Cluster		
		C1	C2	C3
3.	Kab Tangerang			1
4.	Kab Serang	1		
5.	Kota Tangerang		1	
6.	Kota Cilegon	1		
7.	Kota Serang	1		
8.	Kota Tangerang Selatan			1

Sumber: (Handayanna, 2023)

4. Menghitung Centroid Baru

Untuk menemukan centroid baru, Anda harus menemukan berapa banyak cluster yang dipilih dan kemudian membaginya berdasarkan jumlah cluster yang dipilih. Oleh karena itu, hasil jarak dari setiap objek pada iterasi pertama digunakan untuk iterasi kedua berdasarkan perhitungan di bawah ini.

Tabel 5. Centroid Data Iterasi 2

Kabupaten/Kota	Tahun 2020	Tahun 2021	Tahun 2022
C1 (Kab Pandeglang)	120,44	131,43	114,65
C2 (Kab Tangerang)	242,16	272,35	270,52
C3 (Kab Serang)	74,8	83,09	75,45

Sumber: <https://banten.bps.go.id>

Iterasi 2 menghasilkan hasil pengelompokan dari hasil perhitungan keseluruhan. Dijelaskan pada Tabel 6.

Tabel 6. Hasil Perhitungan Jarak Pusat Cluster Iterasi 2

No	Kabupaten/kota	Cluster			Jarak Terpendek
		C1	C2	C3	
1.	Kab Pandeglang	0,00	242,84	77,18	0,00
2.	Kab Lebak	4,22	239,07	80,82	4,22
3.	Kab Tangerang	242,84	0,00	319,19	0,00
4.	Kab Serang	77,18	319,19	0,00	0,00
5.	Kota Tangerang	18,58	231,04	88,32	18,58
6.	Kota Cilegon	182,07	424,02	104,99	104,99
7.	Kota Serang	135,23	377,24	58,13	58,13
8.	Kota Tangerang Selatan	137,14	378,86	59,98	59,98

Sumber: (Handayanna, 2023)

Menggunakan jarak minimum data terhadap pusat cluster untuk menentukan posisi cluster data jumlah penduduk miskin Provinsi Banten. Dalam hal ini, data akan dimasukkan ke dalam kelompok dengan jarak terkecil dari centroid. Posisi data untuk setiap cluster pada iterasi kedua ditunjukkan dalam Tabel 7.

Tabel 7. Posisi Data Dengan Tiap Cluster Pada Iterasi Ke-2

No	Kabupaten/kota	Cluster		
		C1	C2	C3
1.	Kab Pandeglang	1		
2.	Kab Lebak	1		
3.	Kab Tangerang		1	
4.	Kab Serang			1
5.	Kota Tangerang	1		
6.	Kota Cilegon			1
7.	Kota Serang			1
8.	Kota Tangerang Selatan			1

Sumber: (Handayanna, 2023)

5. Menghitung Centroid Baru

Dengan menentukan centroid baru dengan menghitung jumlah yang terpilih pada cluster kemudian membagi data sebanyak jumlah cluster yang terpilih. Demikian di peroleh hasil jarak dari setiap objek pada iterasi ke-2 maka dilanjutkan ke iterasi ke-3 menurut perhitungan di bawah ini.

Tabel 8. Centroid Data Iterasi 3

Kabupaten/kota	Tahun 2020	Tahun 2021	Tahun 2022
C1 (Kab Pandeglang)	120,44	131,43	114,65
C2 (Kota Tangerang)	118,22	134,24	132,88
C3 (Kota Cilegon)	16,31	18,89	16,46

Sumber: (Handayanna, 2023)

Tabel 9 menunjukkan hasil pengelompokan dari iterasi 3, yang menunjukkan hasil dari perhitungan keseluruhan:

Tabel 9. Hasil Perhitungan Jarak Pusat Cluster Iterasi 3

No	Kabupaten/kota	Cluster			Jarak Terpendek
		C1	C2	C3	
1.	Kab Pandeglang	0,00	4100,02	2892,77	0,00
2.	Kab Lebak	4,22	4096,46	2889,17	4,22
3.	Kab Tangerang	242,84	3859,26	2651,43	242,84
4.	Kab Serang	77,18	4176,90	2969,64	77,18
5.	Kota Tangerang	18,58	4089,46	2881,98	18,58
6.	Kota Cilegon	182,07	4281,66	3074,50	182,07
7.	Kota Serang	76,75	76,75	76,75	76,75
8.	Kota Tangerang Selatan	75,02	75,02	75,02	75,02

Sumber: (Handayanna, 2023)

Posisikan data jumlah penduduk miskin Provinsi Banten berdasarkan jarak minimum dari pusat cluster. Data akan menjadi anggota kelompok jika

mereka berada di jarak terkecil dari centroid. Tabel 10 menunjukkan posisi data untuk tiap kelompok pada iterasi ketiga.

Tabel 10. Posisi Data Dengan Tiap Cluster Pada Iterasi Ke-3

No	Kabupaten/kota	Cluster		
		C1	C2	C3
1.	Kab Pandeglang	1		
2.	Kab Lebak	1		
3.	Kab Tangerang	1		
4.	Kab Serang	1		
5.	Kota Tangerang	1		
6.	Kota Cilegon	1		
7.	Kota Serang	1		
8.	Kota Tangerang Selatan	1		

Sumber: (Handayanna, 2023)

Jika nilai centroid hasil iterasi sama dengan nilai centroid sebelumnya, posisi cluster tidak berubah, maka iterasi berhenti. Dengan mempertimbangkan posisi masing-masing kelompok data penduduk miskin dan nilai dari iterasi ketiga, bahwa kelompok kedua memiliki jumlah penduduk miskin terendah di satu kabupaten/kota, yaitu Kota Tangerang. Kelompok satu memiliki jumlah penduduk miskin sedang di tiga kabupaten/kota yaitu Kab Pandeglang, Kab Lebak, dan Kab Tangerang. Kelompok ketiga memiliki jumlah penduduk miskin tertinggi di empat kabupaten/kota, yaitu Kabupaten Serang, Kabupaten Cilegon, Kota Serang, dan Kota Tangerang Selatan.

B. Analisa Dengan Rapid Miner

Dengan menggunakan tools rapidminer, data hasil clustering menunjukkan bahwa terdapat tiga cluster: cluster 0, cluster 1, dan cluster 2. Cluster 0 adalah cluster tertinggi, cluster 1 adalah cluster sedang, dan cluster 2 adalah cluster terendah. Dengan demikian, folder view dihasilkan dari hasil rapidminer studio yang terdapat pada Gambar 3.

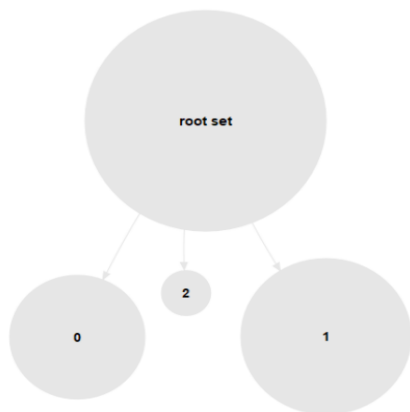


Sumber: (Handayanna, 2023)

Gambar 2. Clustering Kabupaten/Kota Jumlah Penduduk Miskin Provinsi Banten

Cluster penduduk miskin di Provinsi Banten digambarkan pada Gambar 2. Cluster 1 terdiri dari empat Kabupaten/Kota: Serang, Cilegon, Kota

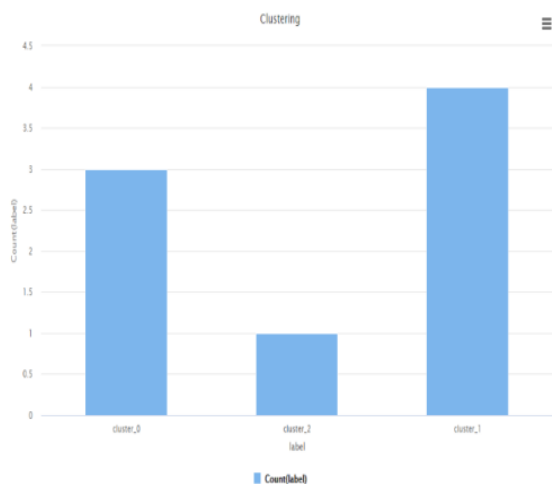
Serang, dan Tangerang Selatan. Cluster 2 terdiri dari satu Kabupaten/Kota, Tangerang, dan Cluster 0 terdiri dari tiga Kabupaten/Kota: Pandeglang, Lebak, dan Kota Tangerang.



Sumber: (Handayanna, 2023)

Gambar 3. Root Set Kabupaten/Kota Jumlah Penduduk Miskin Provinsi Banten

Gambar 3 menunjukkan kelompokan penduduk miskin di setiap kabupaten atau kota di Provinsi Banten berdasarkan jumlah terbesar di Cluster 1 dan Cluster 2 dengan jumlah terkecil.



Sumber: (Handayanna, 2023)

Gambar 4. Visualisasi Grafik Clustering Kabupaten/Kota Jumlah Penduduk Miskin Provinsi Banten

Gambar 4 menunjukkan bahwa cluster 1 memiliki jumlah penduduk miskin tertinggi dan cluster 2 memiliki jumlah penduduk miskin terkecil.

KESIMPULAN

Hasil penelitian menunjukkan bahwa mengelompokkan jumlah penduduk miskin berdasarkan jumlah penduduk yang ada di Kabupaten/Kota di Provinsi Banten dapat dilakukan dengan menggunakan Algoritma K-

Means Clustering dan aplikasi rapidminer. Aplikasi ini menghasilkan hasil yang sama, yaitu mengelompokkan penduduk miskin di Kabupaten/Kota di Provinsi Banten berdasarkan jumlah penduduk miskin tertinggi di daerah tersebut. Dimana hasil pengelompokkan terbentuk yaitu kelompok satu memiliki jumlah penduduk miskin sedang di tiga kabupaten/kota yaitu Kab Pandeglang, Kab Lebak, dan Kab Tangerang. Untuk kelompok kedua memiliki jumlah penduduk miskin terendah di satu kabupaten/kota, yaitu Kota Tangerang. Kelompok ketiga memiliki jumlah penduduk miskin tertinggi di empat kabupaten/kota, yaitu Kabupaten Serang, Kabupaten Cilegon, Kota Serang, dan Kota Tangerang Selatan. Hasil clustering menunjukkan bahwa pemerintah Provinsi Banten harus memprioritaskan pengurangan kemiskinan di kabupaten dan kota tersebut untuk meningkatkan pendapatan dan penghasilan serta meningkatkan penghidupan dan perekonomian. Sehingga algoritma K-Means dapat mengelompokkan penduduk miskin per kabupaten atau kota di Provinsi Banten berdasarkan jumlah penduduknya. Untuk penelitian selanjutnya data yang digunakan sebaiknya bukan berdasarkan jumlah penduduk miskin melainkan dengan kategori lain dan menggunakan tahun yang berbeda.

REFERENSI

- Amanda, & Sitorus, M. V. (2021). Penerapan Algoritma K-Means Clustering Untuk Pengelompokan Konsumsi Produk Kosmetik milik PT Cedefindo. *Jurnal Ilmiah MIKA AMIK Al Muslim*, *V(2)*, 63–68.
- Aprilia, K., & Sembiring, F. (2021). Analisis Garis Kemiskinan Makanan Menggunakan Metode Algoritma K-Means Clustering. *Informasi Dan Manajemen Informatika*, *2(4)*, 1–10.
- Arkham, D., & Swanjaya, D. (2020). K-Means Method For Clustering Public Service Assessment of Government Organization In Kediri City. *Prosiding SEMNAS INOTEK ...*, 155–160. <https://proceeding.unpkediri.ac.id/index.php/inotek/article/view/79>
- Aziz, F. N. R. F. J., Setiawan, B. D., & Arwani, I. (2018). Implementasi Algoritma K-Means untuk Klasterisasi Kinerja Akademik Mahasiswa. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, *2(6)*, 2243–2251.
- Bahauddin, A., Fatmawati, A., & Permata Sari, F. (2021). Analisis Clustering Provinsi Di Indonesia Berdasarkan Tingkat Kemiskinan Menggunakan Algoritma K-Means. *Jurnal Manajemen Informatika Dan Sistem Informasi*, *4(1)*, 1–8.

- <https://doi.org/10.36595/misi.v4i1.216>
Febriansyah, F., & Muntari, S. (2023). Penerapan Algoritma K-Means untuk Klasterisasi Penduduk Miskin pada Kota Pagar Alam. *JISKA (Jurnal Informatika Sunan Kalijaga)*, 8(1), 66–77.
<https://doi.org/10.14421/jiska.2023.8.1.66-77>
- Hablum, R., Khairan, A., & Rosihan, R. (2019). Clustering Hasil Tangkap Ikan Di Pelabuhan Perikanan Nusantara (Ppn) Ternate Menggunakan Algoritma K-Means. *JIKO (Jurnal Informatika Dan Komputer)*, 2(1), 26–33. <https://doi.org/10.33387/jiko.v2i1.1053>
- Mirawati, F., & Feriyanto, N. (2023). *Jurnal Kebijakan Ekonomi dan Keuangan Faktor-Faktor yang memengaruhi jumlah penduduk miskin di Provinsi Banten Tahun 2011-2020*. 2(1), 78–85.
<https://doi.org/10.20885/JKEK.vol2.iss1.art9>
- Munandar, T. A. (2022). Penerapan Algoritma Clustering Untuk Pengelompokan Tingkat Kemiskinan Provinsi Banten. *JSil (Jurnal Sistem Informasi)*, 9(2), 109–114.
<https://doi.org/10.30656/jsii.v9i2.5099>
- Nugraha, I. W. S. A. (2023). Clustering Pemetaan Tingkat Kemiskinan di Provinsi Jawa Barat Menggunakan Algoritma K-Means. *Jurnal Ilmiah Wahana Pendidikan, Januari*, 9(2), 234–244.
<https://doi.org/10.5281/zenodo.7567622>
- Pratiwi, E. D., Ashar, K., & Syafitri, W. (2020). Dampak Kemiskinan Terhadap Pola Mobilitas Tenaga Kerja Antarsektor Di Indonesia. *Jurnal Kependudukan Indonesia*, 15(1), 1.
<https://doi.org/10.14203/jki.v15i1.473>
- Sari, Y. A. (2021). Pengaruh Upah Minimum Tingkat Pengangguran Terbuka Dan Jumlah Penduduk Terhadap Kemiskinan Di Provinsi Jawa Tengah. *Equilibrium : Jurnal Ilmiah Ekonomi, Manajemen Dan Akuntansi*, 10(2), 121–130.
<https://doi.org/10.35906/je001.v10i2.785>
- Sari, Y. R., Sudewa, A., Lestari, D. A., & Jaya, T. I. (2020). Penerapan Algoritma K-Means Untuk Clustering Data Kemiskinan Provinsi Banten Menggunakan Rapidminer. *CESS (Journal of Computer Engineering, System and Science)*, 5(2), 192.
<https://doi.org/10.24114/cess.v5i2.18519>
- Sembiring, Y. R., & Saifullah, R. W. (2021). Implementasi Data Mining Dalam Mengelompokkan Jumlah Penduduk Miskin Berdasarkan Provinsi Menggunakan Algoritma. *Jurnal Penerapan Sistem Informasi (Komputer & Manajemen)*, 2(2), 125–132.
<https://tunasbangsa.ac.id/pkm/index.php/kesatria/article/view/67>
- Wanto, A., Siregar, M. N. H., Windarto, A. P., Hartama, D., Ginantra, N. L. W. S. R., Napitupul, D., Negara, E. S., Lubis, M. R., Dewi, S. V., & Prianto, C. (2020). *Data Mining: Algoritma dan Implementasi*. Yayasan Kita Menulis.