



Implementasi Algoritma K-Means untuk Mengelompokkan Data Tingkat Kemiskinan di Sulawesi Selatan Berdasarkan Kota/Kabupaten

Irfan¹, Lut Faizal²

Program Studi Sistem dan Teknologi Informasi, Universitas Muhammadiyah Sinjai^{1,2}
Jl. Teuku Umar No.8 B, Sinjai Utara, Sinjai, Sulawesi Selatan, 92615, Indonesia
Irfan@umsi.ac.id^{*1}, lutfaizal@umsi.ac.id²

Kata Kunci:

Kemiskinan;
Algoritma K-
Means;
RapidMiner.

ABSTRAK

Kemiskinan adalah suatu kondisi dimana individu atau kelompok tidak memiliki akses yang memadai terhadap sumber daya ekonomi, sosial, dan budaya yang dibutuhkan untuk mencapai standar kehidupan yang wajar. Data Badan Pusat Statistik (BPS) memberikan gambaran umum mengenai tingkat kemiskinan di Provinsi Sulawesi Selatan, meskipun hanya menyajikan tingkat kemiskinan yang dikategorikan berdasarkan kabupaten/kota. Untuk menentukan prioritas, pemerintah harus memiliki pengetahuan tentang daerah dengan tingkat kemiskinan tertinggi dan terendah agar dapat mengatasi kemiskinan secara efektif. Oleh karena itu, sangat penting untuk mengkategorikan kabupaten atau kota di Sulawesi Selatan berdasarkan tingkat kemiskinannya. Hal ini akan memungkinkan pemerintah untuk mengembangkan kebijakan atau strategi yang tepat untuk mengetasi kemiskinan dengan mempertimbangkan tingkat kemiskinan di setiap kabupaten atau kota. Pengelompokan data dalam penelitian ini dilakukan dengan menggunakan algoritma K-Means dan program RapidMiner. Temuan penelitian ini menunjukkan adanya empat klaster kabupaten atau kota yang berbeda, yang dikategorikan berdasarkan tingkat kemiskinannya. Klaster-klaster tersebut dikategorikan sebagai berikut: Klaster 0 terdiri dari Jeneponto, Gowa, Pangkep, Luwu, dan Luwu Utara. Klaster 1 meliputi Bulukumba, Takalar, Maros, Wajo, Pinrang, Enrekang, Tana Toraja, dan Toraja Utara. Klaster 2 terdiri dari Bone dan Makassar. Terakhir, Klaster 3 terdiri dari Kepulauan Selayar, Bantaeng, Sinjai, Barru, Soppeng, Sidrap, Luwu Timur, Pare-Pare, dan Palopo. Klaster 2 memiliki tingkat atau jumlah kemiskinan terbesar di antara semua klaster.

Keywords

Poverty;
K-Means
algorithm;
RapidMiner;

ABSTRACT

Poverty is a state in which individuals or groups lack sufficient access to the economic, social, and cultural resources required to attain a reasonable standard of living. The Central Bureau of Statistics (BPS) data offers an overview of the poverty rate in South Sulawesi province, although it only presents the poverty rate categorized by district or city. In order to prioritize their answers, the government must possess knowledge about the regions with the highest and lowest poverty rates to effectively address poverty. Hence, it is crucial to categorize districts or municipalities in South Sulawesi according to their poverty rate. This would enable the government to develop suitable policies or strategies to alleviate poverty while considering the poverty rate in each district or municipality. The data clustering in this study was performed using the K-Means algorithm and the RapidMiner program. The findings revealed the presence of four distinct clusters of districts or municipalities, which were categorized according to their poverty levels. The clusters are categorized as follows: Cluster 0 consists of Jeneponto, Gowa, Pangkep, Luwu, and



North Luwu. Cluster 1 includes Bulukumba, Takalar, Maros, Wajo, Pinrang, Enrekang, Tana Toraja, and North Toraja. Cluster 2 comprises Bone and Makassar. Lastly, Cluster 3 consists of Selayar Islands, Bantaeng, Sinjai, Barru, Soppeng, Sidrap, East Luwu, Pare-Pare, and Palopo. Cluster 2 has the greatest poverty rate or number among all the clusters.

---Jurnal JISTI @2024---

PENDAHULUAN

Kemiskinan adalah kurangnya sumber daya keuangan untuk memenuhi kebutuhan makanan dan non-makanan yang esensial, yang biasanya dinilai melalui tingkat pengeluaran. Kemiskinan mempengaruhi berbagai aspek kehidupan masyarakat, termasuk sektor ekonomi dan sosial. Dampak terhadap berbagai aspek kehidupan masyarakat akan berkaitan erat dengan tingkat kemiskinan di suatu wilayah. Indikator-indikator seperti tingkat kemiskinan, ambang batas kemiskinan, tingkat pengangguran, dan pengeluaran per kapita dapat memberikan gambaran kemiskinan di suatu wilayah.

Kondisi ekonomi dan sosial setiap daerah memiliki karakteristik yang berbeda-beda. Karakteristik berdasarkan keadaan yang mempengaruhi kemiskinan di setiap kabupaten/kota dapat dikelompokkan berdasarkan kemiripannya satu sama lain. Pengelompokan setiap kabupaten/kota berguna untuk melihat bagaimana kemiskinan di wilayah tersebut sehingga pemerintah dapat menerapkan kebijakan yang baik.

Sistem komputerisasi dapat dimanfaatkan membantu pemerintah dalam menganalisis data yang besar dan mengambil keputusan mengenai klasifikasi statistik kemiskinan di provinsi Sulawesi Selatan. Teknik data mining merupakan salah satu metodenya. Data mining adalah teknik yang digunakan dalam pengolahan data dalam jumlah besar. Data yang telah dianalisis menggunakan teknik data mining dapat menghasilkan wawasan baru, yang dapat digunakan untuk membuat keputusan yang tepat. Data mining melibatkan berbagai metode seperti Klasifikasi, Cluster, Estimasi, Prediksi, dan Asosiasi (Andini dkk., 2022). Tingkat kemiskinan di setiap kabupaten/kota dapat dikelompokkan menggunakan analisis cluster.

Analisis cluster adalah metode data mining yang dapat digunakan untuk menemukan informasi dalam data multidimensi. Tujuan dari analisis cluster adalah untuk mengidentifikasi pola atau kelompok objek yang mirip dalam kumpulan data yang digunakan. Pengelompokan data dilakukan dengan memilih objek-objek yang mirip dalam satu kelompok dari setiap cluster yang mewakili jenis atau kategori objek. Teknik clustering yang paling sering digunakan dalam data mining adalah algoritma K-Means dan K-Medoids (Hastari dkk., 2023; Hidayat dkk., 2023; Prasetyaningrum & Susanti, 2023). K-means clustering adalah pendekatan partisi yang sangat efektif yang banyak digunakan dalam berbagai aplikasi karena kesederhanaannya dan kemampuannya untuk cepat konvergen. Penelitian oleh (Fajriani, 2019) telah menunjukkan bahwa algoritma k-means menghasilkan hasil klaster yang lebih optimal daripada algoritma tradisional dalam analisis clustering. Algoritma k-means mengungguli algoritma k-medoids (Sugianto dkk., 2020). Penelitian sebelumnya tentang klasifikasi data kemiskinan (Suhartini & Yuliani, 2021), yang berjudul Penerapan Data Mining untuk Mengelompokkan Data Warga Miskin Menerapkan Algoritma K-Means di Dusun Bagik Endep, Kelurahan Sukamulia Timur, untuk mengelompokkan keluarga miskin untuk distribusi bantuan yang efisien oleh pemerintah daerah dengan menggunakan teknik K-Means. Dataset terdiri dari 200 entri, masing-masing dengan 9 atribut. Metode K-Means mengkategorikan 18 orang di Cluster 1 memiliki status ekonomi tinggi, 72 orang di Cluster 2 memiliki status ekonomi sedang, dan 110 orang di Cluster 3 memiliki status ekonomi rendah.



Kemudian penelitian (Waworzuntu & Amin, 2018) menggunakan algoritma k-means untuk memetakan calon penerima Jamkesda untuk klasifikasi kemiskinan, sehingga membantu pihak Kelurahan dalam membuat keputusan yang tepat untuk menghindari kesalahan pemilihan penerima Jamkesda. Sementara itu, penelitian (Widayani & Harliana, 2020) membandingkan algoritma K-Means dan Fuzzy Subtractive Clustering (SFCM) untuk klasifikasi keluarga miskin, Standar deviasi dan validitas dari hasil pengelompokan yang diperoleh akan digunakan untuk membandingkan kedua teknik tersebut. Berdasarkan enam kali percobaan, algoritma SFCM mengungguli pendekatan K-Means dalam hal akurasi dan standar deviasi antarkelompok, akan tetapi algoritma K-Means dapat melakukan clustering dengan lebih cepat.

KAJIAN PUSTAKA

1. Kemiskinan

Kemiskinan adalah suatu kondisi kekurangan yang dialami oleh individu atau masyarakat karena keadaan di luar kendali mereka, dan bukan karena pilihan mereka. Kemiskinan dapat muncul karena faktor lingkungan, ekonomi, struktural, sosial, dan budaya. Kemiskinan alam dan ekonomi muncul dari kelangkaan sumber daya, seperti sumber daya alam dan sumber daya manusia, yang menyebabkan terbatasnya peluang produksi dan menghambat pembangunan. Kemiskinan struktural dan sosial muncul dari pembangunan yang tidak merata, karena daerah-daerah membutuhkan kemajuan nasional dan ekonomi, yang mengharuskan adanya pemerataan pembangunan. Kemiskinan kultural muncul dari budaya masyarakat yang menumbuhkan rasa kecukupan, yang pada akhirnya mengakibatkan kemiskinan (Anwar, 2023).

2. Data Mining

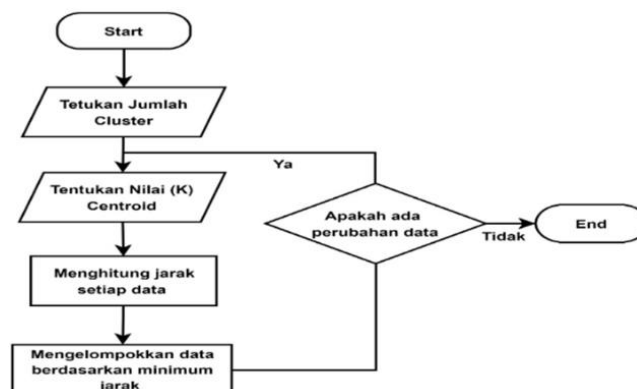
Data mining adalah mengumpulkan informasi dan pengetahuan yang berguna yang tersembunyi di dalam kumpulan data yang luas atau rumit. Data mining adalah proses yang menggunakan metode statistik, matematika, kecerdasan buatan, dan pembelajaran mesin untuk mengekstrak dan menemukan informasi berharga dan pengetahuan terkait dari beragam basis data yang sangat besar (Lesmana, 2021).

3. Clustering

Clustering adalah teknik untuk mengelompokkan data menggunakan sistem partisi dan juga merupakan teknik analisis data atau data mining teknik pemodelan unsupervised learning (Ordila dkk., 2020). Clustering adalah kumpulan atau kelompok objek data yang berbeda dari objek di cluster lain tetapi sebanding satu sama lain di dalam cluster tersebut. Untuk memastikan bahwa objek dalam sebuah cluster memiliki tingkat kemiripan yang tinggi satu sama lain, maka objek-objek tersebut akan diorganisasikan ke dalam satu atau beberapa cluster (Aulia, 2021).

4. Algoritma K-Means

K-Mean adalah teknik analisis kumpulan data yang mengelompokkan data ke dalam klaster, sehingga setiap subjek yang diamati merupakan bagian dari klaster data yang rata-ratanya cukup dekat satu sama lain. K-Means Clustering adalah teknik analisis data yang menggunakan skema partisi untuk pengelompokan data, K digunakan untuk konstanta pengelompokan secara keseluruhan dan Means berarti rata-rata dataset, dalam hal ini sebagai cluster (Regina dkk., 2021). K-Means mengelompokkan data menjadi dua atau lebih (Adha dkk., 2021).



Gambar 1 Flowchar K-Means (Mohammad Ferdiansyah & Umi Chotijah, 2024)

Algoritma K-means dapat diselesaikan dengan mengikuti langkah-langkah (Fauzi & Dana, 2023):

- a. Tentukan cluster dengan memilih beberapa centroid secara acak.
- b. Gunakan rumus euclidean distance untuk menghitung jarak antara setiap data ke centroid terdekatnya dengan rumus berikut:

$$d(i,j) = \sqrt{\sum_j^m (C_{ij} - C_{kj})^2}$$

keterangan :

x_i = data kriteria

μ_j = Centroid Pada Cluster ke-j

- c. Dari setiap data dalam cluster, kumpulkan jarak terpendek ke centroid menggunakan rumus:

$$\text{Min} \sum_{k=1}^k dik = \sqrt{\sum_j^m (C_{ij} - C_{kj})^2}$$

- d. Perbarui nilai centroid setiap cluster dengan menghitung nilai rata-rata dari data dalam cluster tersebut. Nilai baru centroid dihitung dengan rumus:

$$C_{kj} = \frac{\sum_{i=1}^p x_{ij}}{p}$$

- e. Ulangi langkah 1 hingga 4 hingga seluruh data cluster stabil dan tidak terjadi perubahan lagi pada anggota cluster.
- f. Catat nilai centroid akhir dari tiap cluster sebagai hasil dari pengelompokan data yang telah memenuhi kriteria kestabilan seperti dijelaskan di langkah 1, 2, dan 3.

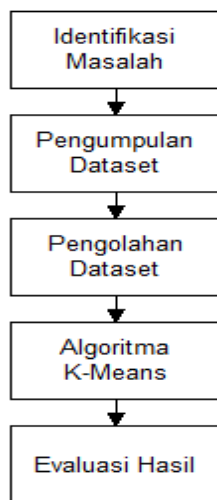
5. Rapidminer

Rapidminer merupakan salah satu platform perangkat lunak yang kuat untuk ilmu data dan pembelajaran mesin. Platform ini menawarkan banyak pilihan alat untuk pemodelan, evaluasi, dan penerapan data. Bahkan mereka yang tidak memiliki pengetahuan pemrograman dapat membuat dan menguji berbagai macam model dengan mudah berkat arsitektur RapidMiner yang ramah pengguna. Dengan antarmuka drag-and-drop yang disediakan oleh RapidMiner, pengguna dapat membuat alur kerja untuk memproses dan menganalisis data. Banyak sumber data yang didukung, termasuk basis data, berkas datar, dan sistem data besar seperti Hadoop dan Spark. Selain itu, program ini dilengkapi dengan berbagai macam operator yang telah dibuat sebelumnya yang berfungsi sebagai blok bangunan alur kerja dan mencakup setiap langkah penggalan data, termasuk pemilihan fitur, pembersihan data, dan pemodelan (Rafi Nahjan dkk., 2023).



METODE PENELITIAN

Metode penelitian ini terdiri dari serangkaian tahapan yang dimaksudkan untuk membantu peneliti dalam melakukan analisis cluster terhadap tingkat kemiskinan di Provinsi Sulawesi Selatan.



Gambar 2 Metode Penelitian

Penelitian ini melibatkan beberapa langkah yang telah diidentifikasi yang perlu diselesaikan, sebagai berikut:

1. Identifikasi Masalah

Para tahap ini, penulis berusaha untuk mengenali dan menguraikan masalah-masalah yang berkaitan dengan analisis pengelompokan / cluster tingkat kemiskinan di Provinsi Sulawesi Selatan.

2. Pengumpulan Dataset

Penulis mengumpulkan data dari Badan Pusat Statistik (BPS) Sulawesi Selatan yang diakses pada tanggal 29 Januari 2024. Data tersebut digunakan untuk memfasilitasi pelaksanaan penelitian terkait clustering tingkat kemiskinan di Provinsi Sulawesi Selatan (*Badan Pusat Statistik Provinsi Sulawesi Selatan, t.t.*).

3. Pengolahan Dataset

Data yang diperoleh akan diproses untuk mendapatkan hasil yang kemudian akan diterapkan pada tahap berikutnya untuk menghasilkan informasi yang akurat. Hasil pengolah dataset dapat dilihat pada tabel berikut

Tabel 1. Data Jumlah Penduduk Miskin(*Badan Pusat Statistik Provinsi Sulawesi Selatan*)

No	Kabupaten	Jumlah Penduduk Miskin (Ribuan Jiwa)					
		2018	2019	2020	2021	2022	2023
1	Kepulauan Selayar	17.59	17.36	17.04	16.93	16.74	16.89
2	Bulukumba	31.25	30.49	30.00	31.31	31.29	30.65
3	Bantaeng	17.20	16.91	16.84	17.78	17.22	17.49
4	Jeneponto	55.95	54.05	53.24	52.35	50.59	48.32
5	Takalar	26.57	25.93	25.38	24.60	24.75	25.01
6	Gowa	59.34	57.99	57.68	58.66	57.96	59.10
7	Sinjai	22.48	22.27	22.06	21.69	21.67	21.14
8	Maros	35.97	34.85	34.62	34.11	33.90	34.96
9	Pangkep	50.12	47.07	47.12	48.40	47.53	46.06
10	Barru	15.68	14.92	14.44	15.18	14.73	14.88
11	Bone	79.57	76.25	81.33	79.64	80.34	80.12



12	Soppeng	17.00	16.45	17.23	17.27	17.21	17.21
13	Wajo	29.73	27.48	27.69	26.22	26.75	27.54
14	Sidrap	15.41	14.44	15.36	15.25	15.56	15.75
15	Pinrang	32.94	31.85	33.56	33.51	33.64	34.23
16	Enrekang	25.53	25.40	25.25	26.13	26.15	26.97
17	Luwu	47.91	46.18	46.04	46.26	46.50	47.67
18	Tana Toraja	29.65	28.87	28.41	29.33	29.31	30.23
19	Luwu Utara	42.43	42.48	42.20	43.15	42.29	40.79
20	Luwu Timur	21.15	20.83	20.82	20.99	20.89	21.57
21	Toraja Utara	30.68	28.64	27.88	28.39	27.79	29.14
22	Makassar	66.22	65.12	69.98	74.69	71.83	80.32
23	Pare Pare	8.01	7.62	7.96	7.93	8.01	7.98
24	Palopo	14.27	14.37	14.71	15.21	14.78	14.85

4. K-Mens

K-Means adalah teknik non-hierarki untuk mengelompokkan data ke dalam satu atau lebih cluster atau kelompok (Asih dkk., 2024). Data yang memiliki kualitas yang sama dikelompokkan dalam cluster yang sama, sedangkan data yang memiliki karakteristik berbeda ditempatkan bersama dalam kelompok yang berbeda (Muningsih dkk., 2021). Pada tahap ini digunakan untuk mengelompokkan data tingkat kemiskinan menggunakan berdasarkan jumlah penduduk miskin.

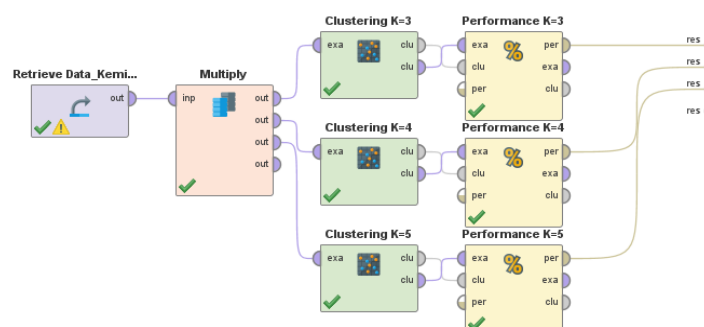
5. Evaluasi Hasil

Salah satu teknik evaluasi yang dapat membantu dalam menilai akurasi cluster dalam penelitian ini adalah menggunakan Davies Bouldin index (DBI). Jumlah K (cluster) yang dihasilkan oleh metode K-Means dapat diukur dengan menggunakan indeks Davies-Bouldin untuk mengetahui keakuratan atau kekuatannya (Kristanto dkk., 2023). Tahap pendekatan pengujian DBI bertujuan untuk memperkecil jarak antar data dokumen yang terdapat pada cluster yang sama dan memaksimalkan jarak antara satu cluster dengan cluster lainnya (Hotijah dkk., 2023).

HASIL DAN PEMBAHASAN

RapidMiner adalah software pengolah data yang menyediakan beberapa algoritma data mining untuk mengekstrak pola dari kumpulan data yang luas dengan mengintegrasikan teknik statistik, basis data, dan kecerdasan buatan. Tersedianya berbagai macam operator dapat mempermudah pengguna RapidMiner dalam melakukan perhitungan data yang kompleks. Operator-operator ini digunakan untuk memanipulasi data. Data dihubungkan ke node-node pada operator, dan hasilnya dapat diamati dengan menghubungkan ke node hasil. Kemampuan RapidMiner untuk menghasilkan grafik visual memungkinkannya untuk dipilih sebagai software dalam melakukan ekstraksi data dengan menggunakan teknik data mining.

Pada kasus analisis data mengenai tingkat kemiskinan, konfigurasi data dan operator yang kami gunakan dapat dilihat pada Gambar 3.



Gambar 3. Konfigurasi RapidMiner

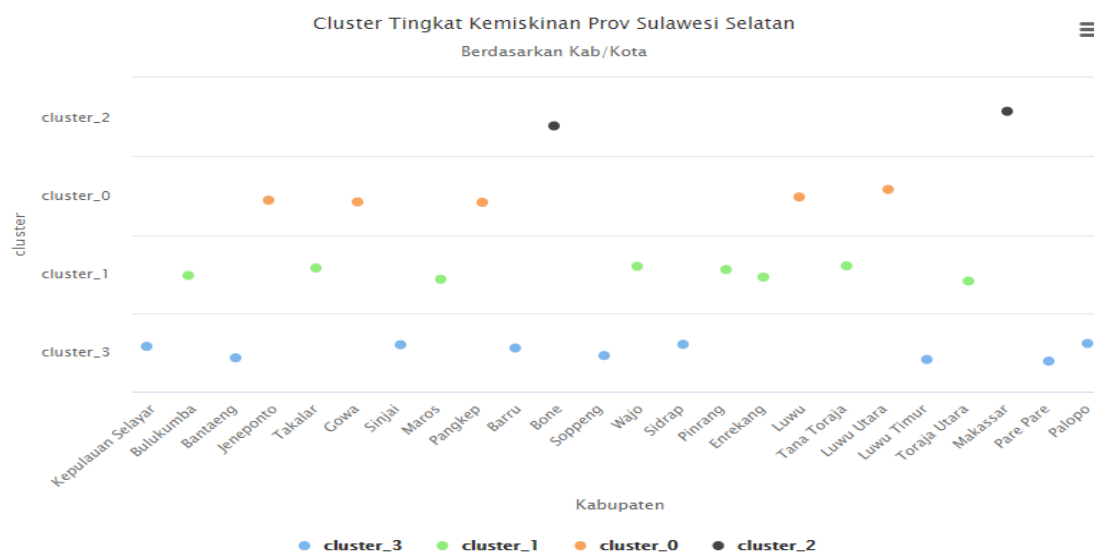


Pada penelitian tentang clustering tingkat kemiskinan di Sulawesi Selatan berdasarkan kabupaten/kota menggunakan RapidMiner, peneliti membandingkan parameter jumlah cluster atau K dengan nilai 3, 4, dan 5. Setelah itu, kinerja cluster dari setiap proses K-means diuji menggunakan DBI. Hasil uji performa dari masing-masing cluster dapat ditemukan dalam tabel yang disajikan berikut ini.

Tabel 2. Nilai DBI

No	Jumlah Cluster	DBI
1	3	0,405
2	4	0,398
3	5	0,474

Hasil perhitungan penelitian menggunakan akurasi dengan menggunakan Davies–Bouldin Index (DBI) bertujuan untuk mengevaluasi kualitas clustering yang telah dilakukan dengan mengukur kuantitas dan fitur turunan dari dataset. Secara esensial, DBI menunjukkan preferensi terhadap nilai yang lebih rendah untuk menilai kualitas clustering. Indeks tersebut dapat digunakan sebagai dasar dalam pengambilan keputusan terkait dengan pemilihan jumlah cluster yang paling sesuai. Hasil uji kinerja menggunakan DBI menunjukkan nilai sebesar 0,398, yang mengindikasikan bahwa K=4 merupakan jumlah cluster optimal untuk mengelompokkan tingkat kemiskinan di Provinsi Sulawesi Selatan. Hasil pengelompokan menunjukkan bahwa terdapat 5 kabupaten/kota dalam cluster 0, 8 kabupaten/kota dalam cluster 1, 2 kabupaten/kota dalam cluster 2, dan 9 kabupaten/kota dalam cluster 3. Hasil cluster dapat ditampilkan pada Gambar 4.



Gambar 4. Hasil Cluster

Gambar diatas mengilustrasikan hirarki tingkat kemiskinan dari yang tertinggi hingga terendah di dalam setiap klaster: klaster_2, klaster 0, klaster 1, dan klaster 3. Klaster_2 terdiri dari 2 kabupaten/kota (Bone, Makassar), Klaster 0 terdiri dari 5 kabupaten/kota (Jeneponto, Gowa, Pangkep, Luwu, dan Luwu Utara), Klaster 1 mencakup 8 kabupaten/kota (Bulukumba, Takalar, Maros, Wajo, Pinrang, Enrekang, Tana Toraja, dan Toraja Utara), sedangkan Klaster 3 melibatkan 9 kabupaten/kota (Kepulauan Selayar, Bantaeng, Sinjai, Barru, Soppeng, Sidrap, Luwu Timur, Pare-Pare, dan Palopo).

SIMPULAN DAN SARAN

Berdasarkan hasil penelitian, dapat disimpulkan bahwa analisis data tingkat kemiskinan di Provinsi Sulawesi Selatan, yang menggunakan aplikasi RapidMiner dengan algoritma k-means, mampu menghasilkan ekstraksi dan pengelompokan data tingkat kemiskinan menjadi empat klaster



dari total 24 data kabupaten/kota di wilayah tersebut. Klaster 0 terdiri dari 5 kabupaten/kota, Klaster 1 terdiri dari 8 kabupaten/kota, Klaster 2 terdiri dari 2 kabupaten/kota, dan Klaster 3 terdiri dari 9 kabupaten/kota. Dengan demikian, analisis ini memberikan gambaran yang jelas mengenai pola tingkat kemiskinan di Provinsi Sulawesi Selatan dan menunjukkan potensi algoritma k-means dalam konteks pemetaan dan pemahaman lebih lanjut terhadap permasalahan kemiskinan.

DAFTAR PUSTAKA

- Adha, R., Nurhaliza, N., & Soleha, U. (2021). *Perbandingan Algoritma DBSCAN dan K-Means Clustering untuk Pengelompokan Kasus Covid-19 di Dunia*. 18(2).
- Andini, Y., Hardinata, J. T., & Purba, Y. P. (2022). Penerapan Data Mining pada Tata Letak Buku Di Perpustakaan Sintong Bingei Pematangsiantar dengan Metode Apriori. *Jurasik (Jurnal Riset Sistem Informasi dan Teknik Informatika)*, 7(1), 13. <https://doi.org/10.30645/jurasik.v7i1.410>
- Anwar, A. A. (2023). *Faktor-Faktor Yang Mempengaruhi Kemiskinan Di Jawa Tengah Periode 2002-2021 Analisis Data Time Series*.
- Asih, D. P. A., Irawan, B., & Bahtiar, A. (2024). *PENGELOMPOKAN DATA TRANSAKSI DALAM MENENTUKAN STRATEGI PENJUALAN MENGGUNAKAN ALGORITMA K-MEANS*. 8(1).
- Aulia, S. (2021). KLASTERISASI POLA PENJUALAN PESTISIDA MENGGUNAKAN METODE K-MEANS CLUSTERING (STUDI KASUS DI TOKO JUANDA TANI KECAMATAN HUTABAYU RAJA). *Djtechno: Jurnal Teknologi Informasi*, 1(1), 1–5. <https://doi.org/10.46576/djtechno.v1i1.964>
- Badan Pusat Statistik Provinsi Sulawesi Selatan. (t.t.). Diambil 29 Februari 2024, dari <https://sulsel.bps.go.id/indicator/23/454/1/jumlah-penduduk-miskin-ribu-jiwa-menurut-kabupaten-kota-se-sulawesi-selatan.html>
- Fajriani, F. (2019). *Persebaran Tingkat Kemiskinan di Sulawesi Selatan Menggunakan K-Means Clustering Analysis*.
- Fauzi, I. A., & Dana, R. D. (2023). *Implementasi Data Mining Clustering Dalam Mengelompokan Kasus Perceraian Yang Terjadi Di Provinsi Jawa Barat Menggunakan Algoritma K-Means*. 1.
- Hastari, D., Nurunnisa, F., Winanda, S., & Aprillia, D. D. (2023). *Application of K-Means and K-Medoids Algorithms for Grouping Country Data Based on Socio-Economic and Health Factors*.
- Hidayat, F. P., Putra, R. P., Alfitriah, M. D., & Widodo, E. (2023). Implementasi Clustering K-Medoids dalam Pengelompokan Kabupaten di Provinsi Aceh Berdasarkan Faktor yang Mempengaruhi Kemiskinan. *Indonesian Journal of Applied Statistics*, 5(2), 121. <https://doi.org/10.13057/ijas.v5i2.55080>
- Kristanto, B., Turmudi Zy, A., & M. Fatchan. (2023). Analisis Penentuan Karyawan Tetap Dengan Algoritma K-Means Dan Davies Bouldin Index. *Bulletin of Information Technology (BIT)*, 4(1), 112–120. <https://doi.org/10.47065/bit.v4i1.521>
- Lesmana, B. (2021). *Pengelompokan Pengiriman Hasil Kelapa Sawit Berdasarkan Tonase dan Kualitas Menggunakan Metode Clustering (Studi Kasus: KUD Bumi Pusaka)*.



- Mohammad Ferdiansyah & Umi Chotijah. (2024). Implementasi Algoritme K-Means++ Untuk Clustering Penjualan Bahan Bangunan. *Jurnal Ilmiah Teknik Informatika dan Komunikasi*, 4(1), 181–193. <https://doi.org/10.55606/juitik.v4i1.767>
- Muningsih, E., Maryani, I., & Handayani, V. R. (2021). Penerapan Metode K-Means dan Optimasi Jumlah Cluster dengan Index Davies Bouldin untuk Clustering Propinsi Berdasarkan Potensi Desa. 9(1).
- Ordila, R., Wahyuni, R., Irawan, Y., & Yulia Sari, M. (2020). PENERAPAN DATA MINING UNTUK PENGELOMPOKAN DATA REKAM MEDIS PASIEN BERDASARKAN JENIS PENYAKIT DENGAN ALGORITMA CLUSTERING (Studi Kasus: Poli Klinik PT.Inecda). *Jurnal Ilmu Komputer*, 9(2), 148–153. <https://doi.org/10.33060/JIK/2020/Vo19.Iss2.181>
- Prasetyaningrum, E., & Susanti, P. (2023). Perbandingan Algoritma K-Means Dan K-Medoids Untuk Pemetaan Hasil Produksi Buah-Buahan. 7.
- Rafi Nahjan, M., Nono Heryana, & Apriade Voutama. (2023). IMPLEMENTASI RAPIDMINER DENGAN METODE CLUSTERING K-MEANS UNTUK ANALISA PENJUALAN PADA TOKO OJ CELL. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 7(1), 101–104. <https://doi.org/10.36040/jati.v7i1.6094>
- Regina, S., Sutinah, E., & Agustina, N. (2021). Clustering Kualitas Kinerja Karyawan Pada Perusahaan Bahan Kimia Menggunakan Algoritma K-Means. *JURNAL MEDIA INFORMATIKA BUDIDARMA*, 5(2), 573. <https://doi.org/10.30865/mib.v5i2.2909>
- Sugianto, C. A., Rahayu, A. H., & Gusman, A. (2020). Algoritma K-Means untuk Pengelompokan Penyakit Pasien pada Puskesmas Cigugur Tengah. *Journal of Information Technology*, 2(2), 39–44. <https://doi.org/10.47292/joint.v2i2.30>
- Suhartini, S., & Yuliani, R. (2021). Penerapan Data Mining untuk Mengcluster Data Penduduk Miskin Menggunakan Algoritma K-Means di Dusun Bagik Endep Sukamulia Timur. *Infotek : Jurnal Informatika dan Teknologi*, 4(1), 39–50. <https://doi.org/10.29408/jit.v4i1.2986>
- Waworuntu, M. N. V., & Amin, M. F. (2018). PENERAPAN METODE K-MEANS PEMETAAN CALON PENERIMA JAMKESDA. 05.
- Widayani, W., & Harliana, H. (2020). Perbandingan Algoritma K-Means dan SFCM Pada Pengelompokan Rumah Tangga Miskin. *Jurnal Sains dan Informatika*, 6(1), 1–9. <https://doi.org/10.34128/jsi.v6i1.200>