

PENGELOMPOKAN TINGKAT KEMISKINAN DI SETIAP PROVINSI DI INDONESIA MENGGUNAKAN ALGORITMA K-MEDOIDS

Fathimatuz Zahra^{1*}, Assyifa Khalif², Betha Nurina Sari³

^{1,2,3}Universitas Singaperbangsa Karawang; Jl. HS. Ronggo Waluyo, Puseurjaya, Telukjambe Timur, Karawang, Jawa Barat, 41361, Indonesia; Telp : (0267) 641177

Riwayat artikel:

Received: 18 Maret 2024

Accepted: 30 Maret 2024

Published: 2 April 2024

Keywords:

Klasterisasi;
Kemiskinan;
Provinsi;
K-Medoids.

Correspondent Email:

2010631170009@student.unsika.ac.id

Abstrak. Kemiskinan adalah keadaan di mana individu atau sekelompok masyarakat tidak memiliki kemampuan untuk memenuhi kebutuhan dasar mereka, seperti pangan, sandang, tempat tinggal, pendidikan, dan layanan kesehatan. Kemiskinan bervariasi menurut provinsi di Indonesia. Tujuan dari penelitian ini adalah untuk mengidentifikasi kluster kemiskinan di setiap negara bagian dengan menggunakan pendekatan algoritma K-Medoids. K-Medoids adalah teknik pengelompokan partisi yang mengelompokkan kumpulan objek ke dalam *cluster*. Penelitian ini menggunakan data Badan Pusat Statistik (BPS) yang mencakup 34 provinsi di Indonesia pada tahun 2015 hingga 2022. Hasil yang diperoleh adalah terdapat tiga kluster yang berbeda: kluster sangat miskin, kluster miskin, dan kluster rentan miskin. 50% penduduknya tergolong dalam kemiskinan ekstrem, 45% lagi berada dalam kemiskinan, dan sisanya berada dalam kemiskinan rentan. Nilai *Silhouette Score* sebesar 0,4 menunjukkan *clustering* yang dihasilkan sangat baik.

Abstract. Poverty is a situation where individuals or a group of people do not have the ability to fulfill their basic needs, such as food, clothing, shelter, education, and health services. Poverty varies by province in Indonesia. The purpose of this research is to identify poverty clusters in each state by using the K-Medoids algorithm approach. K-Medoids is a partition clustering technique that groups a collection of objects into clusters. This research uses data from the Central Bureau of Statistics (BPS) covering 34 provinces in Indonesia from 2015 to 2022. The results obtained are three different clusters: the very poor cluster, the poor cluster, and the vulnerable poor cluster. 50% of the population belongs to extreme poverty, 45% to poverty, and the rest to vulnerable poverty. The *Silhouette Score* value of 0.4 indicates that the clustering is very good.

1. PENDAHULUAN

Kemiskinan adalah keadaan di mana individu atau sekelompok masyarakat tidak memiliki kemampuan untuk memenuhi kebutuhan dasar mereka, seperti pangan, sandang, tempat tinggal, pendidikan, dan layanan kesehatan. Selain itu, pengangguran dan kurangnya lahan produktif sebagai sumber pendapatan juga menimbulkan tantangan serius bagi masyarakat miskin dalam memenuhi

kebutuhan dasarnya seperti pangan, air, dan tempat tinggal. Tingkat kemiskinan suatu negara atau wilayah juga mencerminkan tingkat kebahagiaan masyarakat yang tinggal di sana. Oleh karena itu, status kemiskinan mencerminkan tingkat kesejahteraan masyarakat yang tinggal di suatu negara atau wilayah [1]. Sebagai negara berkembang, Indonesia masih menghadapi tantangan besar terkait kemiskinan, namun permasalahan ini

perlu mendapat perhatian serius. Pertumbuhan penduduk di perkotaan menimbulkan berbagai permasalahan yang kompleks. Pertumbuhan penduduk di perkotaan berimplikasi pada meningkatnya kejahatan, kemacetan, tingginya angka pengangguran, dan kesulitan dalam menyediakan perumahan yang layak dan infrastruktur sosial lainnya. Namun untuk mengatasi permasalahan tersebut perlu adanya peningkatan kualitas sumber daya manusia. Dengan tenaga kerja berkualitas tinggi, individu mempunyai peluang lebih besar untuk meningkatkan kesejahteraannya [2].

Penelitian Nurohmah, dkk [3] menunjukkan bahwa, analisis data menggunakan K-Means *clustering* dengan algoritma PCA pada data kemiskinan Kabupaten/Kota di Jawa Barat (2019-2022) menunjukkan pembentukan tiga *cluster* dalam dua skenario. Pemahaman tentang pola kemiskinan di Jawa Barat dapat menjadi dasar untuk merancang kebijakan yang lebih efektif dalam penanggulangan kemiskinan.

Menurut penelitian Khalif, A., dkk [4], dalam menerapkan metode Elbow, analisis terbaik dilakukan dengan membandingkan jumlah *cluster* pada titik-titik tertentu di mana siku terbentuk dalam grafik *Sum of Square Error*. Hasil dari perhitungan *Sum of Square Error* dari K=1 hingga K=10 menunjukkan bahwa siku terbentuk pada titik K=3, yang menandakan titik *cluster* optimal.

Berdasarkan penelitian Alfiah, F., dkk. [5] menemukan dua *cluster* menggunakan *clustering* K-Medoids, yakni *cluster* 1 (30 kabupaten/kota) dengan sanitasi baik, AHH tinggi, dan tingkat melek huruf baik, serta *cluster* 2 (8 kabupaten/kota) dengan proporsi rumah tangga miskin penerima manfaat yang tinggi. Provinsi Jawa Timur mempunyai pengeluaran pangan per kapita tertinggi.

Menurut penelitian yang dilakukan oleh Wahyuli, Disty, dkk. [6], menggunakan data mining dengan algoritma K-Medoids untuk mengelompokkan garis kemiskinan menurut negara bagian berdasarkan sekumpulan data uji. Hasil pengujian data di Rapidminer 5.3 menggunakan K-Medoids menunjukkan bahwa dua kelas klasifikasi dapat dibuat dengan akurasi 100%.

Mengingat Indonesia masih memiliki permasalahan kemiskinan yang serius, maka analisis kesenjangan menjadi sangat penting.

Dengan mengidentifikasi perbedaan antara keadaan saat ini dan keadaan yang diharapkan, penelitian ini dapat menyoroti bidang-bidang yang memerlukan perhatian lebih dalam upaya meningkatkan kesejahteraan. Hal ini memungkinkan penggunaan sumber daya yang lebih efisien dan penerapan strategi intervensi yang lebih baik untuk mengatasi masalah kemiskinan.

Tujuan dari penelitian ini adalah untuk mengidentifikasi dan menilai kesenjangan kemiskinan terbesar di berbagai provinsi di Indonesia.

2. TINJAUAN PUSTAKA

2.1. Kemiskinan

Kemiskinan merupakan kondisi individu dan masyarakat yang mengalami kesulitan dalam memenuhi kebutuhan sehari-hari karena terbatasnya sumber daya, rendahnya upah yang tidak cukup untuk menutupi biaya hidup, dan bertambahnya jumlah penduduk sehingga mengurangi persaingan untuk mendapatkan pekerjaan di berbagai bidang yang membutuhkan adalah tantangannya [5]. Masalah tersebut sesuai dengan faktor-faktor penyebab kemiskinan yang dijelaskan oleh Suparlan, dkk. [6], kemiskinan diakibatkan oleh kurangnya pendapatan atau aset untuk memenuhi kebutuhan dasar seperti pangan, sandang, perumahan, serta layanan kesehatan dan pendidikan yang memadai.

Menurut Suparlan, dkk., kemiskinan dibagi menjadi dua jenis, yakni kemiskinan struktural dan kemiskinan budaya. Kemiskinan struktural terjadi sebagai akibat dari struktur sosial, ekonomi, politik, dan budaya yang memungkinkan kelompok tertentu dalam masyarakat mengakses sumber daya. Kemiskinan budaya, sebaliknya, tercermin dalam perilaku boros, rendahnya tabungan, dan sikap pasrah terhadap situasi kemiskinan [7].

2.2. Data Mining

Data Mining adalah konsep yang dimanfaatkan untuk mengungkap pengetahuan yang terdapat dalam *database* yang belum terungkap secara langsung. [8]. Data mining adalah penggunaan teknologi pengenalan pola seperti statistik dan teknik matematika untuk menemukan pola dan korelasi baru yang

berguna dengan menganalisis data dalam jumlah besar dari berbagai sumber. Proses ini melibatkan penerapan satu atau lebih teknik pembelajaran mesin untuk menganalisis data secara otomatis dan mengekstrak pengetahuan darinya. [9]. Berdasarkan berbagai penjelasan tentang data mining, dapat disimpulkan bahwa *data mining* adalah proses pengolahan data untuk menghasilkan informasi yang berguna. Fungsi utama *data mining* adalah membantu dalam memperoleh informasi yang bermakna serta meningkatkan pengetahuan bagi pengguna.

Teknik data mining biasanya terbagi menjadi dua kategori utama, yaitu prediksi dan deskripsi. Teknik prediksi menggunakan data historis untuk membuat estimasi tentang kejadian yang akan terjadi di masa depan. Di sisi lain, teknik deskripsi bertujuan untuk mengenali pola dalam data yang dapat mengungkapkan informasi tentang keterkaitan tersembunyi antara variabel [10].

2.3. Knowledge Discovery in Database (KDD)

Knowledge Discovery in Database (KDD) adalah serangkaian prosedur untuk mengenali dan mengubah data menjadi informasi yang mengandung pola yang bermanfaat. Proses ini dilakukan pada basis data besar di mana potensi manfaatnya sebelumnya tidak teridentifikasi [11]. Tahapan dari KDD diantaranya yaitu *selection data*, *preprocessing data*, *transformation data*, *data mining*, dan *evaluation data* [12].

2.4. Clustering

Clustering merupakan proses pengelompokan objek data yang memiliki ciri atau karakteristik serupa ke dalam suatu *cluster*, sedangkan objek atau data dengan tingkat kesamaan yang rendah terhadap *cluster* lainnya. Tujuan utama dari *clustering* adalah untuk meningkatkan kesamaan antara objek dalam satu *cluster* sambil mengurangi kesamaan antara *cluster* yang berbeda [13]. *Clustering* adalah teknik dalam data mining yang digunakan untuk menganalisis data [14].

2.5. Silhouette Coefficient

Dalam perhitungan *silhouette coefficient*, jarak antara data diukur menggunakan rumus jarak *Euclidean*. Kriteria pengukuran

pengelompokan berdasarkan nilai *silhouette coefficient* dipaparkan dalam Tabel 1.

Table 1. Kriteria Pengukuran *Silhouette Coefficient*

Score	Kriteria
0.71-1.00	Struktur Kuat
0.51-0.70	Struktur Baik
0.26-0.50	Struktur Lemah
≤0.25	Struktur Buruk

2.6. Algoritma K-Medoids

K-Medoids adalah teknik pengelompokan partisi yang mengelompokkan kumpulan objek ke dalam *cluster*. K-Medoid menggunakan medoid (objek yang bertindak sebagai pusat *cluster*) untuk mewakili setiap *cluster*. Objek yang berdekatan dengan medoid dikumpulkan untuk membentuk *cluster* baru. K-Medoids memiliki kelebihan dalam mengatasi kelemahan K-Means yang rentan terhadap *outlier*. Hasil pengelompokan K-Medoid tetap konsisten meskipun data diacak [16].

Algoritma PAM (*Partitioning Around Medoids*) atau dikenal juga dengan nama algoritma K-Medoids merupakan algoritma yang setiap *cluster*nya diwakili oleh sebuah medoid [17]. Metode K-Medoids adalah teknik pengelompokan partisi yang sangat efektif untuk kumpulan data kecil. Algoritma K-Medoids dikembangkan untuk mengatasi sensitivitas algoritma K-Means terhadap *outlier*. Objek yang berbeda secara signifikan dari sebagian besar objek lain dalam kumpulan data [18]. K-Medoids menunjukkan kinerja yang lebih baik saat digunakan untuk dataset yang berukuran kecil. Algoritma ini menggunakan objek dalam himpunan data sebagai representasi *cluster* [19].

3. METODE PENELITIAN

Penelitian ini menggunakan data yang diperoleh dari Badan Pusat Statistik (BPS) yang mencakup 34 provinsi di Indonesia. Dataset ini terdiri dari 272 baris data dan 10 kolom.

Table 2. Parameter dan Deskripsi dari *Dataset Kemiskinan*

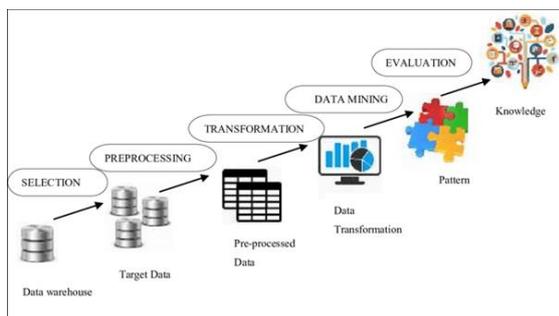
Parameter	Deskripsi
Tahun	Tahun
Provinsi	Nama Provinsi
Jml_pend_miskin	Jumlah Penduduk Miskin

persentase_pend_miskin(P0)	Presentase Penduduk Miskin
ind_kedalaman_kemiskinan(P1)	Indeks Kedalaman Kemiskinan
ind_keparahan_kemiskinan(P2)	Indeks Keparahan Kemiskinan
tk_pengangguran	Tingkat Pengangguran
Pengeluaran_pkt	Pengeluaran per Kapita
RLS	Rata-rata Lama Sekolah
UHH	Umur Harapan Hidup
IPM	Indeks Pembangunan Manusia
GINI	Gini Rasio
ratarata_upah_perjam	Rata – Rata Upah per Jam
PDRB	Produk Domestik Regional Bruto (PDRB) per Kapita

Penelitian ini menggunakan metodologi KDD (*Knowledge Discovery in Database*) yang terdiri dari enam tahapan yaitu seleksi data, *preprocessing* data, transformasi data, data mining, dan evaluasi data.

Dalam definisi data mining, proses ini mengacu pada pencarian pola dan informasi penting dalam kumpulan data yang dipilih menggunakan teknik dan metode tertentu untuk mengekstraksi pengetahuan yang berharga.

Fase-fase ini memungkinkan untuk mengeksplorasi data yang ada, membangun model yang relevan, mengidentifikasi pola-pola yang mungkin belum diketahui sebelumnya, dan mengevaluasi hasil untuk mendapatkan pemahaman yang lebih mendalam.



Gambar 1. Flowchart Metodologi KDD (Sumber : sis.binus.ac.id)

3.1. Selection Data

Pada langkah awal ini, dilakukan pemilihan data atau atribut yang relevan sesuai dengan kebutuhan penelitian, karena tidak semua atribut dalam basis data tersebut diperlukan sesuai dengan tujuan penelitian ini [20].

Gambar 2. Selection Data

3.2. Preprocessing Data

Metode yang digunakan dalam *preprocessing data* adalah *data cleaning*. *Data cleaning* adalah langkah penting untuk menghilangkan *data noise* dan *missing value*. Tujuannya adalah untuk memastikan bahwa atribut yang ada tidak mengandung data yang tidak valid atau nilai yang ilegal [20].

```
#preprocessing data
df.isnull().sum()

Tahun          0
Provinsi        0
persentase_pend_miskin (P0)  0
ind_kedalaman_miskin (P1)  0
ind_keparahan_miskin (P2)  0
tk_pengangguran  0
ratarata_upah_perjam  0
pengeluaran_perkapita  0
PDRB_perkapita  0
RLS             0
UHH             0
IPM             0
gini_ratio      0
jml_pend_miskin  0
dtype: int64

# Menentukan Input Variable
x = df[['persentase_pend_miskin (P0)',
       'ind_kedalaman_miskin (P1)', 'ind_keparahan_miskin (P2)',
       'tk_pengangguran', 'ratarata_upah_perjam', 'pengeluaran_perkapita',
       'PDRB_perkapita', 'RLS', 'UHH', 'IPM', 'gini_ratio', 'jml_pend_miskin']]
```

Gambar 3. Preprocessing Data

3.3. Transformation Data

Pada tahap transformasi data dilakukan normalisasi dan nilai setiap atribut diubah agar sesuai dengan algoritma yang digunakan. Hal ini penting karena tidak semua tipe data dapat ditangani sesuai keinginan, sehingga perlu menyesuaikan tipe data dengan kebutuhan analisis. Proses ini bertujuan untuk mentransformasikan data agar sejalan dengan tujuan analisis yang ingin dicapai [20].

3.4. Data Mining

Tahap data mining menggunakan metode *clustering* K-Medoids atau algoritma *clustering* K-Medoids. K-Medoids adalah metode clustering atau pengelompokan fungsi. Menurut (Larose, 2005), *clustering* mengacu pada pengelompokan data, pengamatan, atau kasus berdasarkan kesamaan objek yang dianalisis. *Cluster* adalah kumpulan data yang mirip satu sama lain atau berbeda dengan kelompok lain [20].

```
from sklearn.preprocessing import MinMaxScaler
from sklearn_extra.cluster import KMedoids # Import KMedoids

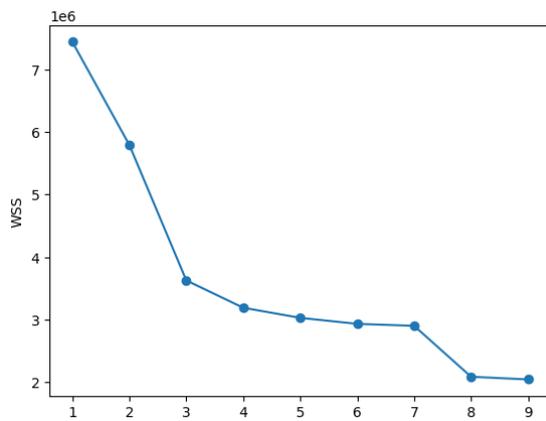
cluster_range = range(1, 10)

cluster_wss = []
# wss=within some square
for number_cluster in cluster_range:
    clusters = KMedoids(n_clusters=number_cluster) # Menggunakan KMedoids
    clusters.fit(x)
    cluster_wss.append(clusters.inertia_)

plt.xlabel("Jumlah Cluster")
plt.ylabel("WSS")
plt.plot(cluster_range, cluster_wss, marker='o')
plt.show()
```

Gambar 4. Data Mining

Proses menggunakan metode Elbow untuk melakukan pengelompokan dengan K-Medoids dalam sklearn.cluster melibatkan serangkaian langkah-langkah. Ini termasuk mengimpor modul yang dibutuhkan, menentukan rentang jumlah *cluster* yang akan diuji, membuat daftar kosong untuk menyimpan nilai WSS (*Within Sum of Squares*), melakukan iterasi untuk setiap jumlah *cluster* yang ditetapkan, melatih model K-Medoids, menghitung dan menyimpan nilai WSS [4].



Gambar 5. Visualisasi *Elbow Method*

```
cls = KMedoids(n_clusters=3, random_state=0)
y_kmed = cls.fit_predict(x)
# Membentuk model
model_cls = cls.fit(x)
```

Gambar 6. Variabel *Cluster*

3.5. Evaluation Data

Pada tahap evaluasi, hasil dari proses pembelajaran data mining dianalisis. Langkah ini melibatkan interpretasi hasil pemodelan data yang telah dilakukan, dengan menggunakan metode *Silhouette Coefficient* untuk mengevaluasi kualitas *cluster* yang terbentuk. Tujuan evaluasi ini adalah untuk menilai apakah pemodelan yang telah dilakukan relevan dengan kasus penelitian dan sesuai dengan

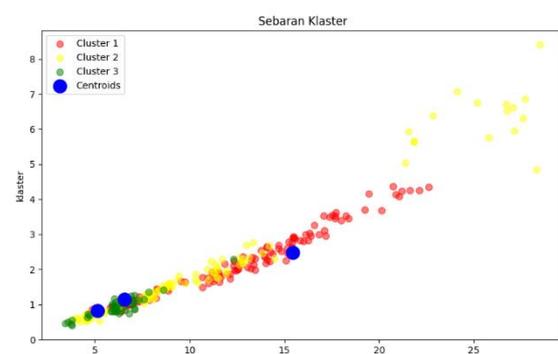
tujuan yang ingin dicapai. Setelah menerima hasil evaluasi, keputusan diambil apakah proses dapat dilanjutkan ke tahap berikutnya atau apakah perlu diulang dari awal jika tidak memenuhi tujuan yang telah ditetapkan [20].

```
from sklearn.metrics import silhouette_score

# Evaluasi menggunakan Silhouette Score
silhouette_avg = silhouette_score(x, model_cls.labels_)
print(f"Silhouette Score: {silhouette_avg}")

Silhouette Score: 0.45736161472654296
```

Gambar 7. Hasil *Silhouette Score*



Gambar 8 Sebaran *Cluster*

3.6. Knowledge

Setelah hasil analisis diperoleh, penelitian ini dapat memberikan kontribusi dalam bentuk pengetahuan yang dapat bermanfaat bagi banyak orang. Setelah diketahui, pemerintah setempat dapat segera mengambil tindakan untuk mengurangi tingkat kemiskinan yang terjadi di beberapa provinsi di Indonesia [20].

4. HASIL DAN PEMBAHASAN

4.1. Analisis Algoritma K-Medoids

Proses klusterisasi atau pengelompokan dilakukan dengan menentukan inerti. Inerti mengukur seberapa baik kumpulan data ketika dikelompokkan dengan mengukur jarak antara setiap titik data dan centroid-nya, lalu jarak tersebut dikuadratkan, dan kuadrat ini dijumlahkan di dalam satu *cluster*. Model *cluster* yang baik adalah model dengan inersia rendah dan jumlah *cluster* (*K*) yang rendah. Setelah memilih lokasi centroid secara acak, setiap sampel data dialokasikan ke centroid terdekat untuk mulai membuat *cluster* yang lebih tepat. Jarak antara setiap sampel data dan setiap centroid dihitung, jarak minimum dipilih,

dan setiap sampel data diberi label yang menunjukkan *cluster* terdekatnya.

Perhitungan kemudian dilakukan dengan menggunakan metode *silhouette coefficient*. Metode ini merupakan kombinasi metode *cohesion and separation* untuk menilai kualitas *cluster*. Rumus metode *silhouette coefficient* adalah:

$$s(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))}$$

Dimana $a(i)$ adalah jarak rata-rata antara i dalam *cluster* yang sama, $b(i)$ adalah rata-rata jarak antara i dan *cluster* terdekat berikutnya. Nilai yang diperoleh dengan metode *silhouette coefficient* adalah dari -1 hingga 1. Jika nilai *silhouette coefficient* mendekati 1 maka diperoleh hasil pengelompokan data yang lebih baik. Sebaliknya, ketika *silhouette coefficient* mendekati -1, hasil *clustering* data menjadi lebih buruk.

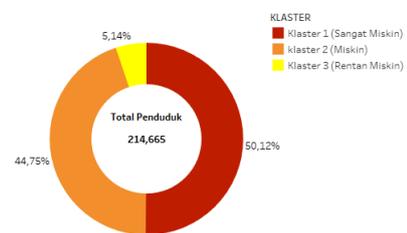
4.2. Hasil

Dengan menggunakan algoritma pengujian, ditemukan bahwa penduduk Indonesia dapat dibagi menjadi tiga kelompok berbeda. Setiap kelompok diberi label untuk memudahkan memahami fungsinya. Dengan *Silhouette Score* 0,4, hasil grup dinilai cukup baik. Berikut adalah interpretasi *Silhouette Score*:

- > 0.5 : *Clustering* yang sangat bagus
- 0.2 - 0.5 : *Clustering* yang bagus
- 0 - 0.2 : *Clustering* yang tidak terlalu bagus
- < 0 : *Clustering* yang buruk

Hasil analisis *cluster* menunjukkan bahwa penduduk Indonesia dari 34 provinsi berhasil dipilah menjadi tiga kelompok berbeda. Ketiga kelompok ini adalah sebagai berikut:

Persentase Penduduk Miskin berdasarkan Klaster Tingkat Kemiskinan



Gambar 9. Persentase Jumlah Penduduk Miskin

Klaster 1: Sangat miskin, Klaster 2: Miskin, Klaster 3: Rentan terhadap kemiskinan. Hasil tersebut menunjukkan bahwa 50% penduduknya tergolong sangat miskin, 45% lainnya tergolong miskin, dan sisanya tergolong miskin rentan. Hal ini menunjukkan perlunya fokus pada kelompok ini secepatnya untuk mengatasi kemiskinan. Penggunaan metode *clustering*, khususnya K-Medoid, membantu mengidentifikasi kelompok-kelompok yang kualitasnya berbeda.

5. KESIMPULAN

- a. Hasil analisis tingkat kemiskinan Indonesia menggunakan algoritma K-Medoids pada tahun 2015 hingga 2022 menunjukkan bahwa data jumlah penduduk 34 provinsi berhasil dikelompokkan menjadi tiga kategori utama, yaitu: sangat miskin, miskin, dan rentan miskin. Hasil tersebut menunjukkan separuh dari total penduduk Indonesia masuk dalam kategori kemiskinan berat, sisanya sebesar 45 persen masuk dalam kategori miskin dan sisanya masuk dalam kategori kemiskinan rentan.
- b. Hasil evaluasi menunjukkan bahwa dengan *Silhouette Score* sebesar 0,4, pengelompokan ini efektif dalam memisahkan kelompok-kelompok tersebut.

DAFTAR PUSTAKA

- [1] N. Zuhdiyaty and D. Kaluge, "Analisis Faktor-faktor yang Mempengaruhi Kemiskinan di Indonesia Selama Lima Tahun Terakhir (Studi Kasus Pada 33 Provinsi)," *Jurnal Jibeka*, Vols. 1, no. 2, p. 27–31, 2017.
- [2] N. Zuhdiyaty, "Analisis Faktor - Faktor Yang Mempengaruhi Kemiskinan Di Indonesia Selama Lima Tahun Terakhir (Studi Kasus Pada 33 Provinsi)," *JIBEK*, Vols. 1, no. 2, pp. 1-5, 2015.
- [3] Y. R. M. a. B. N. S. Nurohmah, "Optimalisasi Performa K-Means Clustering Dengan PCA Dalam Analisis Tingkat Kemiskinan Di Jawa Barat," *JATI (Jurnal Mahasiswa Teknik Informatika)*, Vols. 7, no. 3, pp. 1657-1665, 2023.
- [4] Khalif, A., Hasanah, A. N., Ridwan, M. H., & Sari, B. N., "Klasterisasi Tingkat Kemiskinan di Indonesia menggunakan Algoritma K-Means," *Generation Journal*, Vols. 8, no. 1, pp. 54-62, 2024.
- [5] Alfiah, F., Almadayani, A., Al Farizi, D., & Widodo, E., "Analisis Clustering K-Medoids Berdasarkan Indikator Kemiskinan di Jawa Timur Tahun 2020," *Jurnal Ilmiah Sains*, vol. 22(1), pp. 1-7, 2022.
- [6] Wahyuli, Disty, et al., "Mengelompokkan Garis Kemiskinan Menurut Provinsi Menggunakan Algoritma K-Medoids.," *Prosiding Seminar Nasional Riset Information Science (SENARIS)*, vol. 1, pp. 452-461, 2019.
- [7] Y. A. Sari, "Pengaruh Upah Minimum Tingkat Pengangguran Terbuka Dan Jumlah Penduduk Terhadap Kemiskinan Di Provinsi Jawa Tengah," *Equilibrium*, Vols. 1, no. 2, pp. 121-130, 2021.
- [8] P. Suparlan, *Masyarakat dan Kebudayaan Perkotaan: Perspektif Antropologi Perkotaan*, 2004.
- [9] B. P. S. J. Pusat, *Statistik Indonesia Tahun 2010*, Jakarta Pusat: Badan, 2010.
- [10] M. N. H. S. a. O. S. Y. Fadhillah, "Perancangan Aplikasi Data Mining Untuk Menentukan Pasien Menderita Tifoid Dengan Metode Algoritma C4. 5," *Explorer (Hayward)*, Vols. 1, no. 2, pp. 63-71, 2021.
- [11] M. R. R. a. A. P. T. M. S. Mustafa, "Implementasi Data Mining untuk Evaluasi Kinerja Akademik Mahasiswa Menggunakan Algoritma Naive Bayes Classifier," *Creat. Inf. Technol. J.*, Vols. 4, no. 2, pp. 151-162, 2018.
- [12] F. Nasution, *Penerapan Metode K-Means Clustering untuk Mengelompokkan Ketahanan Tanaman Pangan Kabupaten/Kota di Provinsi Sumatera Utara*, Diss. Universitas Sumatera Utara, 2019.
- [13] Amrullah, A., Purnamasari, I., Sari, B. N., & Voutama, A. , "Analisis Cluster Faktor Penunjang Pendidikan Menggunakan Algoritma K-Means (Studi Kasus: Kabupaten Karawang)," *J. Inform. dan Rekayasa Elektron*, Vols. 5, no. 2, p. 244–252, 2022.
- [14] A. Fahllevi, "Proses Data Mining KDD Pada Pengguna Brand Hp Menggunakan Metode K-Means," *Binus University*, Vols. 1, no. 1, p. 2021, 2021.
- [15] I. N. a. D. Darwis, "Analisis Clustering Pada Pengguna Brand Hp Menggunakan Metode K-Means," *Pros. Semin. Nas. Ilmu Komput.*, Vols. 1, no.1, p. 2021, 2021.
- [16] B. V. A. & H. N. Nurseptia, "Penerapan Algoritma K-Means Untuk Mengelompokkan Kabupaten/Kota Dalam Upaya Pemetaan Lapangan Pekerjaan Baru," *J. Teknol. Informasi*, Vols. 1, no. 2, p. 181–186, 2022.
- [17] S. P. a. H. Abduh, "Penerapan Metode Silhouette Coefficient untuk Evaluasi Clustering Obat," *PENA Tek. J. Ilm. Ilmu-Ilmu Tek.*, Vols. 1, no. 2, p. 48, 2021.
- [18] P. D. & V. S. Arora, "Analysis of K-Means and K-Medoids Algorithm For Big Data," *Procedia Computer Science*, vol. 78, p. 507–512, 2016.
- [19] I. K. U. & M. M. Kamila, "Perbandingan Algoritma KMeans Dan K-Medoids Untuk Pengelompokan Data Transaksi Bongkar Muat," *Jurnal Ilmiah Rekayasa Dan Manajemen Sistem*, Vols. 1, no. 1, p. 119–125, 2019.
- [20] D. A. e. a. Silitonga, "Penerapan Metode KMedoid Pada Pengelompokan Rumah Tangga Dalam Perlakuan Memilah Sampah Menurut Provinsi," *Seminar Nasional Sains dan Teknologi Informasi (SENSASI)*, Vols. 2, No. 1, p. 313–318, 2019.