

# Klasterisasi Tingkat Kemiskinan di Indonesia Menggunakan Algoritma K-Means

Assyifa Khalif<sup>1</sup>, Anisa Nur Hasanah<sup>2</sup>, Muhammad Hafizh Ridwan<sup>3</sup>,  
Betha Nurina Sari<sup>4</sup>

<sup>1,2,3,4</sup>Informatika, Fakultas Ilmu Komputer, Universitas Singaperbangsa Karawang

E-mail: <sup>1</sup>[2010631170058@student.unsika.ac.id](mailto:2010631170058@student.unsika.ac.id),

<sup>2</sup>[2010631170052@student.unsika.ac.id](mailto:2010631170052@student.unsika.ac.id), <sup>3</sup>[1910631170107@student.unsika.ac.id](mailto:1910631170107@student.unsika.ac.id),

<sup>4</sup>[betha.nurina@staff.unsika.ac.id](mailto:betha.nurina@staff.unsika.ac.id)

**Corresponden Author:** [2010631170058@student.unsika.ac.id](mailto:2010631170058@student.unsika.ac.id)

Diterima Redaksi: 09 September 2023 Revisi Akhir: 27 Februari 2024 Diterbitkan Online: 12 Maret 2024

**Abstrak** – Kemiskinan adalah salah satu tantangan sosial yang mendalam di seluruh dunia dan menjadi fokus utama dalam agenda pembangunan global. Penelitian ini menerapkan metode klasterisasi dalam menganalisis dan memahami masalah kemiskinan. Penelitian ini menggunakan data dari Badan Pusat Statistik (BPS) mengenai 34 provinsi di Indonesia untuk mengelompokkan daerah berdasarkan tingkat kemiskinan dengan 12 variabel. Analisis klasterisasi membantu untuk mengidentifikasi karakteristik yang mungkin terabaikan oleh pendekatan konvensional, yang pada gilirannya memungkinkan pengembangan solusi yang lebih terarah dan efektif dalam upaya mengatasi kemiskinan. Penelitian ini menggunakan algoritma K-Means dengan kerangka metodologi CRISP-DM. Penentuan  $k$  pada  $k$ -means dioptimasi menggunakan metode Elbow, mendapatkan  $k$  optimum = 3. Tiga klaster menunjukkan kelompok sangat miskin, miskin, dan hampir miskin. Hasilnya menunjukkan bahwa hampir 95% penduduk miskin di Indonesia termasuk dalam kelompok 'Miskin'. Hasil evaluasi menggunakan nilai Silhouette Score sebesar 0.7416 mengindikasikan bahwa klasterisasi ini sangat efektif dalam memisahkan kelompok-kelompok tersebut.

**Kata Kunci** — Kemiskinan, Klasterisasi, K-Means, Metodologi CRISP-DM, Pembangunan Berkelanjutan.

**Abstract** – Poverty is one of the deep social challenges around the world and is a major focus in the global development agenda. This research applies the clustering method in analysing and understanding poverty issues. This research uses data from the Central Bureau of Statistics (BPS) on 34 provinces in Indonesia to cluster regions based on poverty rates with 12 variables. Clustering analysis helps to identify characteristics that may be overlooked by conventional approaches, which in turn enables the development of more targeted and effective solutions to poverty. This research uses the K-Means algorithm with the CRISP-DM methodological framework. The determination of  $k$  in  $k$ -means is optimised using the Elbow method, obtaining the optimum  $k = 3$ . Three clusters indicate the very poor, poor, and near poor groups. The results show that almost 95% of the poor in Indonesia belong to the 'Poor' group. The evaluation result using the Silhouette Score value of 0.7416 indicates that this clustering is very effective in separating these groups.

**Keywords** — Poverty, Clustering, K-Means, CRISP-DM Methodology, Sustainable Development.



## 1. PENDAHULUAN

Kemiskinan, sebagai salah satu tantangan sosial yang paling mendalam di dunia, terus menjadi fokus utama dalam agenda pembangunan global. Sejak tahun 2000, Perserikatan Bangsa-bangsa telah secara aktif berupaya mengatasi masalah ini melalui berbagai inisiatif dan komitmen yang tegas. Namun, upaya untuk mengakhiri kemiskinan tetap menjadi pekerjaan besar, khususnya di negara-negara yang masih dalam tahap awal perkembangan[1]. Pentingnya isu kemiskinan tak hanya sebatas masalah statistik atau ekonomi semata. Kemiskinan meresap ke dalam berbagai aspek kehidupan manusia, mengganggu stabilitas ekonomi, menghambat akses pendidikan yang berkualitas, dan berdampak negatif pada kesehatan masyarakat. Oleh karena itu, untuk mencapai Sustainable Development Goals (SDGs) tahun 2030 yang telah ditetapkan oleh PBB, perlu adanya

pendekatan yang holistik dalam memahami dan mengatasi masalah ini[2]. Dengan melibatkan berbagai disiplin ilmu seperti ekonomi, sosial, pendidikan, dan kesehatan, tulisan ini bertujuan untuk memberikan kontribusi nyata dalam upaya global untuk mengakhiri kemiskinan. Melalui analisis data yang mendalam dan pendekatan multidimensional, penelitian ini diharapkan dapat memberikan saran-saran konkret yang dapat digunakan sebagai landasan untuk kebijakan-kebijakan yang lebih efektif dalam menghapuskan kemiskinan di Indonesia, mewujudkan visi pembangunan berkelanjutan yang adil dan inklusif. Landasan hukum untuk pendekatan ini dapat ditemukan Dalam Undang-Undang Tahun 2004 Nomor 25, yang mengatur Sistem Perencanaan Pembangunan Nasional. Undang-Undang ini menetapkan dasar perencanaan pembangunan nasional, yang mencakup tujuan pembangunan berkelanjutan yang sesuai dengan SDGs dan komitmen Indonesia dalam mengatasi masalah kemiskinan secara komprehensif[3].

Penelitian ini menerapkan metode klusterisasi dalam menganalisis dan memahami permasalahan kemiskinan. Klusterisasi adalah alat penting yang dapat membantu mengidentifikasi pola-pola kompleks dalam data kemiskinan, sehingga memungkinkan pengambilan keputusan yang lebih tepat dan efisien dalam merumuskan kebijakan dan program-program intervensi[4]. Analisis klusterisasi akan membantu untuk mengidentifikasi karakteristik yang mungkin terabaikan oleh pendekatan konvensional, sehingga memungkinkan kita untuk mengembangkan solusi yang lebih terarah dan efektif dalam upaya mengatasi kemiskinan. Hui Liu menerapkan metode klustering untuk mengidentifikasi karakteristik kemiskinan pada rumah tangga miskin di Cina[5]. Penelitian Nanda,dkk [6] menunjukkan bahwa dalam melakukan klusterisasi data kemiskinan di Indonesia lebih baik menggunakan algoritma K-Medoids dibanding K-Means. Penelitian Mariah Abdul Rahman, dkk [7] menerapkan K-means mengidentifikasi indikator kemiskinan multidimensi untuk kelompok 40 persen terbawah di Malaysia. Clark [8] menerapkan algoritma k-means dan pohon keputusan untuk mengidentifikasi faktor-faktor yang menyebabkan tunawisma di Amerika Serikat. Endartyana [9] mengelompokkan kabupaten/kota di Papua berdasarkan indikator kemiskinan, dan Repollo [10] menganalisis kemiskinan pada sebuah komunitas di Filipina menggunakan algoritma k-means.

Penelitian mengenai klustering data kemiskinan sudah dilakukan beberapa peneliti di Indonesia sebelumnya menggunakan algoritma K-Means, diantaranya Nurhafiza [11] di provinsi Riau, Matius Tadi [12] di provinsi Banten, Fitriani [13] di Jawa tengah, Sofyan [14] di Kecamatan Samadua Aceh Selatan, Zunaida Sitorus [15] di kota Tanjungbalai. Selain itu, Achmad Bahauddin [16] telah menerapkan klustering provinsi di Indonesia menggunakan algoritma k-means untuk membentuk tiga klaster berdasarkan tingkat kemiskinan, yaitu rendah, sedang, dan tinggi. Pada penelitian tersebut hanya menggunakan empat variabel, yaitu persentase penduduk miskin, garis kemiskinan, P1, dan P2 sehingga perlu dikembangkan lagi dengan menambah penggunaan variabel pada datasetnya. Penelitian Khoirun Niswati, dkk [17] menghasilkan 8 klaster dalam pengelompokkan kabupaten dan kota di Indonesia berdasarkan indikator kemiskinan. Penelitian Istiqamah,dkk [18] mendapatkan tiga klaster kemiskinan status di Hulu Sungai Tengah, yaitu tingkat rendah, menengah dan tinggi.

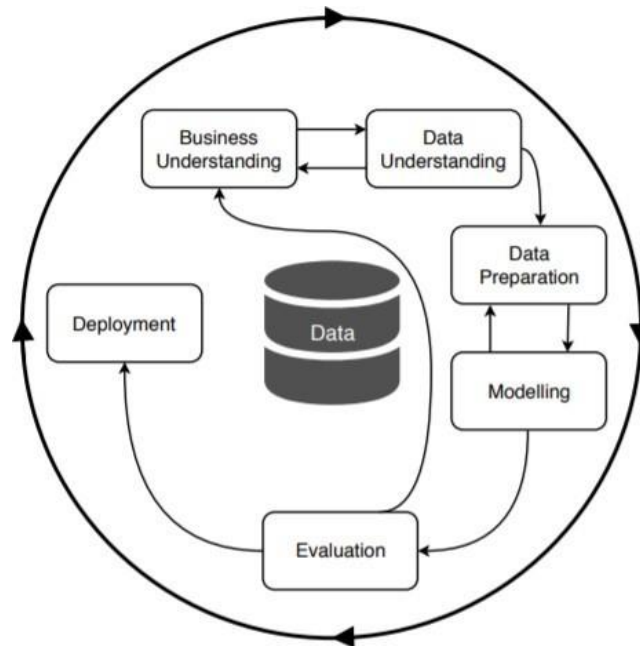
## 2. METODE PENELITIAN

Penelitian ini bersumber dari data Badan Pusat Statistik (BPS). Dataset diambil pada 1 Juni 2023 yaitu Data berdasarkan 34 Provinsi di Indonesia dari tahun 2015 sampai 2022 Dengan total baris data sebanyak 272 baris dan 10 kolom.

Tabel 1. Parameter dan Deskripsi dari Dataset Kemiskinan

<b>Parameter</b>	<b>Deskripsi</b>
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 Ratio.
ratarata_upah_perjam	Rata – Rata Upah per Jam.
PDRB	Produk Domestik Regional Bruto (PDRB) per Kapita.

## 2.1. Metodologi CRISP-DM



Gambar 1. Model CRIPS-DM

CRISP-DM merupakan singkatan dari *Cross-Industry Standart Process for Data Mining*[19]. Metodologi ini adalah metodologi data mining yang lengkap dengan menawarkan model proses untuk pemula dan pakar.,diantaranya:

1. Pemahaman Bisnis (Business Understanding) berkonsentrasi pada memperoleh pemahaman tentang tujuan proyek dari sudut pandang bisnis.
2. Pemahaman Data (Data Understanding) meliputi empat langkah yaitu pengumpulan data primer untuk meningkatkan pemahaman data, deskripsi, eksplorasi, dan verifikasi kualitas.
3. Persiapan Data (Data Preparation) menggabungkan seluruh proses yang diperlukan untuk menghasilkan kumpulan data akhir atau yang akan digunakan dalam tahap pemodelan data awal.
4. Pemodelan (Modeling) melibatkan pemilihan dan penerapan teknik pemodelan yang berbeda dengan atribut atau parameter yang diubah untuk mendapatkan hasil terbaik.
5. Evaluasi (Evaluation) adalah tahap peninjauan dan evaluasi hasil model dan pembuatan tahap awal untuk mengevaluasi pencapaian tujuan dalam bisnis.
6. Penyebaran (Deployment) yaitu merupakan tahap pengorganisasian dan penyajian hasil penemuan pengetahuan dari proses data mining sebelumnya agar mudah dipahami dan disajikan dengan jelas.

## 2.2. Klasterisasi

Klasterisasi juga dikenal sebagai pengelompokan data, adalah proses menggolongkan himpunan data yang atribut kelasnya belum dijelaskan. Tujuan dari clusterisasi adalah untuk memaksimalkan dan meminimumkan kesamaan antar kelas. Misalnya, dalam kasus di mana ada suatu himpunan obyek pada proses awal bisa dikelompokkan menjadi beberapa kategori, lalu menjadi himpunan yang teratur, dan selanjutnya dapat diferensiasi berdasarkan dengan klasifikasi tertentu[20].

## 2.3. Metode *K-Means*

Teknik klasterisasi berbasis prototipe menciptakan pembagian satu tingkat dari objek data. *K-Means* mendefinisikan prototipe dalam bentuk pusat massa (*centroid*), yang biasanya adalah rerata dari sekelompok titik data, dan umumnya digunakan untuk objek dalam ruang n-dimensi yang kontinu. Teknik klasterisasi *K-Means* adalah pendekatan yang sederhana, dan kita akan mulai dengan menjelaskan algoritma dasarnya. Pertama memilih *K centroid* awal, di mana *K* adalah parameter yang ditentukan oleh peneliti, yaitu jumlah klaster yang diharapkan. Setiap titik selanjutnya diberikan kepada *centroid* terdekat, dan setiap kelompok titik yang diberikan kepada

*centroid* membentuk sebuah kluster. Kemudian, *centroid* dari setiap kluster diperbaharui berdasarkan titik-titik yang diberikan ke kluster tersebut. Langkah-langkah pengalokasian dan pembaruan diulang sampai tidak ada titik yang berpindah kluster, atau dengan kata lain, sampai *centroid* tetap sama[21].

### 2.3.1. Metode Siku (*Elbow Method*)

Metode Siku (*Elbow Method*) merupakan sebuah teknik yang dipakai dalam analisis klusterisasi untuk penentuan jumlah *cluster* yang optimal. Untuk mencapai konvergensi pada algoritma *K-Means*, pemilihan nilai *K* yang baik sangat penting. *Elbow Method* mengukur distorsi kuadrat yang terkait dengan berbagai nilai *K*. Ketika nilai *K* meningkat, distorsi rata-rata menurun, tetapi ini juga berarti bahwa setiap kluster akan berisi lebih sedikit sampel yang semakin mendekati pusat massa. Oleh karena itu, saat memilih nilai *K* yang optimal, kita mencari titik di mana penurunan distorsi tidak lagi signifikan, yang sering disebut sebagai "siku" pada grafik[22].

## 3. HASIL DAN PEMBAHASAN

### 3.1. *Business Understanding* (Pemahaman Bisnis)

Tujuan pada tahapan ini adalah untuk memahami dengan mendalam pola dan karakteristik kemiskinan di Indonesia serta mengidentifikasi kluster-kluster kemiskinan yang berbeda. Tujuan utama adalah untuk membantu pemerintah dan lembaga terkait dalam pengembangan kebijakan yang lebih efektif untuk mengurangi kemiskinan dan memberikan informasi yang relevan kepada organisasi non-pemerintah dan lembaga tertentu untuk mengarahkan program bantuan dan bantuan sosial dengan lebih efektif.

### 3.2. *Data Understanding* (Pemahaman Data)

Penelitian ini mendapatkan sumber data dari Badan Pusat Statistik (BPS). Dataset yang diambil yaitu Data berdasarkan 34 Provinsi di Indonesia dari tahun 2015 sampai 2022 Dengan total baris data sebanyak 272 baris dan 10 kolom terdiri dari atribut seperti persentase penduduk miskin (P0), indeks kedalaman kemiskinan (P1), indeks keparahan kemiskinan (P2), tingkat pengangguran, rata-rata upah per jam, pengeluaran per kapita, PDRB per kapita, RLS (tingkat rata-rata lama sekolah), UHH (tingkat rata-rata rumah tangga), IPM (Indeks Pembangunan Manusia), gini ratio, dan jumlah penduduk miskin.

### 3.3. *Data Preparation* (Persiapan Data)

Dalam tahapan ini, dataset yang digunakan untuk analisis klusterisasi ini diimpor dan dipersiapkan.

```
from sklearn.cluster import KMeans
from sklearn.preprocessing import MinMaxScaler
#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 2. *Data Preparation*

Pada tahap ini mengimpor dua modul yang diperlukan dari pustaka scikit-learn, yaitu *K-Means* dari modul *cluster* dan *MinMaxScaler* dari modul *preprocessing*. Modul *K-Means* akan digunakan untuk menjalankan algoritma klustering *K-Means*, sedangkan *MinMaxScaler* akan digunakan untuk penskalaan data. Selanjutnya, kode ini mendefinisikan variabel (*x*) yang akan digunakan untuk menyimpan fitur-fitur yang akan digunakan dalam analisis klustering. Fitur-fitur ini diambil dari sebuah *DataFrame* yang disebut (*df*). Fitur-fitur ini mencakup berbagai aspek seperti persentase penduduk miskin, indikator kedalaman kemiskinan, indikator keparahan kemiskinan, tingkat pengangguran, upah per jam rata-rata, pengeluaran per kapita, PDRB per kapita, RLS, UHH, IPM, rasio gini, dan jumlah penduduk miskin.

### 3.4. Data Modelling (Pemodelan Data)

Dalam proses pemodelan data menggunakan metode klusterisasi, algoritma *K-Means* dari pustaka *scikit-learn* (*sklearn*) digunakan

```

Elbow Method

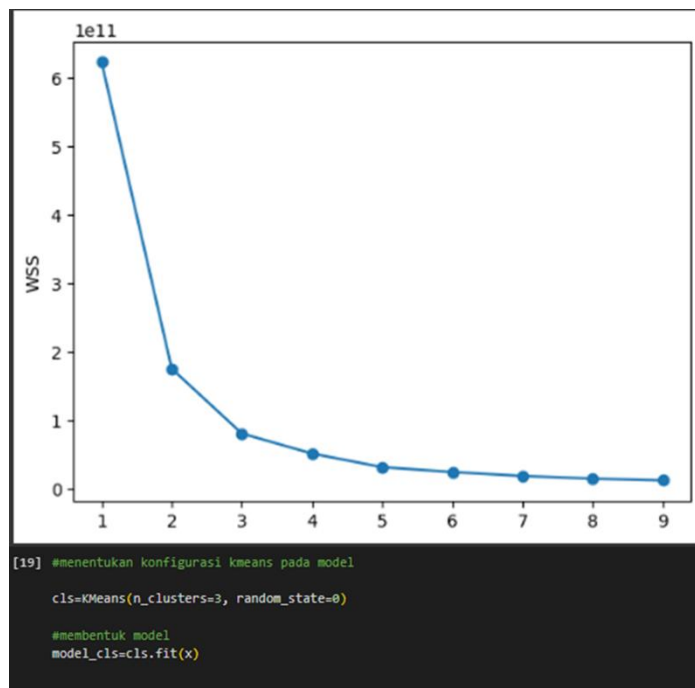
[8] cluster_range=range(1,10)

cluster_wss=[]
#wss=within some square
for number_cluster in cluster_range:
    clusters=KMeans(number_cluster)
    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 3. Data Modelling

Penggunaan *elbow method* pada penelitian ini bertujuan untuk mengoptimasi nilai *k* pada algoritma *k-means* [23], agar tahap pemodelan data dapat menghasilkan cluster terbaik [24]. Penggunaan *Elbow Method* dalam klusterisasi, menggunakan *KMeans* dari *sklearn.cluster* untuk mengelompokkan data ke dalam kluster. Tahapannya meliputi impor modul yang dibutuhkan, menentukan rentang jumlah kluster, membuat list kosong untuk menyimpan nilai *WSS* (*Within Sum of Squares*), dan melakukan perulangan untuk setiap jumlah kluster. Selanjutnya, melatih model *K-Means*, menghitung nilai *WSS*, dan menyimpannya. Terakhir, tampilkan hasilnya dalam plot yang memperlihatkan perubahan nilai *WSS* dengan jumlah kluster. Tujuan utamanya menentukan jumlah kluster yang ideal untuk hasil klusterisasi yang optimal.



Gambar 4. Visualisasi *Elbow Method*

Hasil dari perhitungan *Elbow Method* didapat bahwa nilai yang paling optimal untuk menentukan jumlah *n* kluster adalah 3. Dalam *Elbow Method* nilai atau point yang optimal untuk digunakan terdapat pada patahan atau lekukan kedua. Pada titik ini, penambahan jumlah kluster tidak memberikan penurunan variansi yang signifikan. *Elbow point* pada titik ke-3 menandakan jumlah kluster optimal yang mencapai keseimbangan antara kompleksitas model (jumlah kluster) dan variasi yang dijelaskan oleh model.

### 3.5. Evaluation (Evaluasi)

*Silhouette Score* (disebut juga *Silhouette Coefficient*) menghitung rata-rata *Silhouette Coefficient* untuk semua sampel dalam data[25]. *Silhouette Score* merupakan metrik evaluasi yang dipakai dalam pengukuran untuk mengetahui sejauh mana hasil dari analisis klusterisasi efektif dalam mencari nilai k terbaik[26]. Nilai *Silhouette Score* berkisar antara -1 hingga 1, dengan nilai mendekati 1 menunjukkan pemisahan yang baik antara kluster, nilai mendekati 0 menunjukkan tumpang tindih antara kluster, dan nilai mendekati -1 menunjukkan pemisahan yang buruk[27].

```
[21] from sklearn.metrics import silhouette_score

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

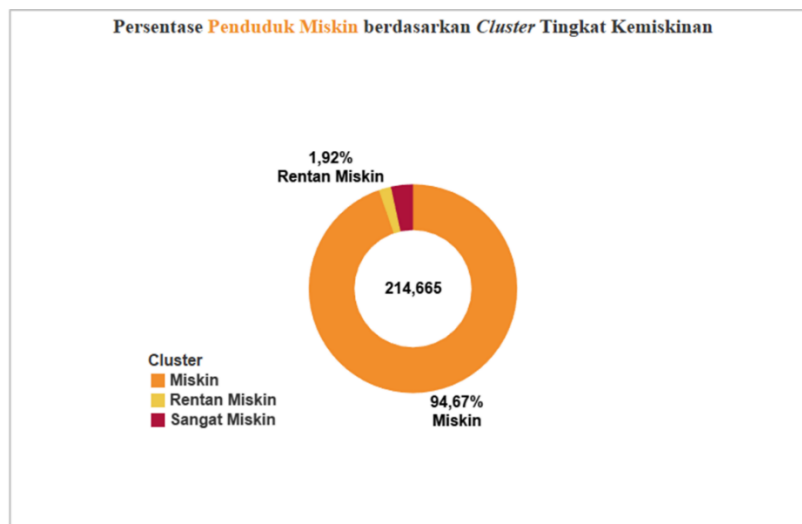
Silhouette Score: 0.7416344989173794
```

Gambar 5. *Silhouette Score*

Dengan nilai *Silhouette Score* sebesar 0.7416, dapat disimpulkan bahwa analisis klusterisasi sangat efektif dalam mengidentifikasi dan memisahkan kelompok yang relevan dalam data. Hasil ini menunjukkan bahwa data poin dalam satu kluster sangat mirip satu sama lain dan sangat berbeda dari data poin dalam kluster lain. Ini adalah indikasi yang sangat baik bahwa analisis klusterisasi berhasil dalam memberikan pemahaman yang mendalam tentang pola-pola dalam data.

### 3.6. Deployment (Penyebaran)

Hasil dari analisis klusterisasi menunjukkan bahwa populasi penduduk Indonesia dari 34 provinsi telah berhasil dikelompokkan menjadi tiga kluster yang berbeda. Keempat kluster ini diberi label sebagai berikut:



Gambar 6. Persentase Jumlah Penduduk Miskin

Klaster 1: Sangat Miskin, Klaster 2: Miskin, Klaster 3: Rentan Miskin. Dari hasil tersebut ditemukan bahwa hampir 95% dari total populasi masyarakat Indonesia yang terdiri dari 34 provinsi termasuk dalam kategori kluster Miskin. Hal ini menggarisbawahi pentingnya perhatian khusus terhadap kelompok ini dalam upaya pengentasan kemiskinan. Penggunaan metode klusterisasi, khususnya *K-Means*, telah membantu dalam mengidentifikasi kelompok miskin dengan karakteristik serupa.

#### 4. SIMPULAN

Dari hasil penelitian yang dilakukan dapat disimpulkan bahwa:

1. Hasil analisis Klasterisasi Tingkat Kemiskinan di Indonesia menggunakan Algoritma *K-Means* dari tahun 2015 hingga 2022 menunjukkan bahwa data penduduk Indonesia dari 34 provinsi berhasil dikelompokkan menjadi tiga klaster yang berbeda, yaitu Sangat Miskin, Miskin, dan Rentan Miskin. Temuan utama adalah bahwa hampir 95% dari total populasi penduduk Indonesia termasuk dalam kelompok klaster Miskin. Hal ini menunjukkan bahwa pentingnya perhatian khusus terhadap kelompok ini dalam upaya pengentasan kemiskinan. Penggunaan metode klasterisasi, terutama *K-Means*, telah membantu mengidentifikasi kelompok miskin dengan karakteristik serupa. Hal ini memungkinkan perancangan kebijakan yang lebih sesuai dan efektif.
2. Hasil analisis ini dapat berfungsi sebagai referensi pemerintah mengurangi kemiskinan seperti:
  - Membantu pemerintah dan lembaga terkait dalam memperoleh pemahaman tentang pola dan ciri-ciri kemiskinan di Indonesia.
  - Membantu dalam pengenalan kelompok miskin yang memiliki karakteristik yang sebanding.
  - Memungkinkan pembuatan kebijakan yang lebih tepat dan efisien untuk mengurangi tingkat kemiskinan.
3. Penerapan metode *Elbow* yang terbaik pada penelitian ini dengan melihat perbandingan klaster pada titik tertentu membentuk siku dengan menggunakan *Sum of Square Error*. Hasil dari *Sum of Square Error* dalam interval dari  $K=1$  hingga  $K=10$ , siku membentuk pada titik  $K=3$ , yang merupakan titik klaster terbaik.
4. Hasil evaluasi menggunakan nilai Silhouette Score sebesar 0.7416 mengindikasikan bahwa klasterisasi ini sangat efektif dalam memisahkan kelompok-kelompok tersebut. Data poin dalam satu klaster memiliki karakteristik yang mirip satu sama lain dan berbeda dari klaster lain.

#### 5. SARAN

Penulis memberikan saran sebagai berikut:

1. Observasi Lanjutan: Seiring dengan hasil klasterisasi yang diperoleh, disarankan untuk melakukan penelitian lanjutan yang lebih mendalam tentang setiap kluster, khususnya kelompok Miskin. Ini akan membantu dalam pemahaman yang lebih baik tentang karakteristik dan kebutuhan kelompok ini.
2. Perumusan Kebijakan Khusus: Pemerintah dan lembaga terkait seharusnya merumuskan kebijakan yang lebih khusus dan terfokus untuk kelompok miskin, mengingat sebagian besar populasi termasuk dalam kelompok ini. Kebijakan ini harus mencakup program pengentasan kemiskinan yang lebih sesuai dengan karakteristik kelompok miskin.
3. Kolaborasi Antar Sektor: Kolaborasi antara pemerintah, Lembaga Swadaya Masyarakat (LSM), dan sektor swasta sangat penting dalam implementasi kebijakan dan program-program pengentasan kemiskinan. Sinergi antar sektor dapat menghasilkan hasil yang lebih efektif.
4. Pemantauan dan Evaluasi Terus-Menerus: Setelah perumusan kebijakan dan implementasi, penting untuk melakukan pemantauan dan evaluasi berkelanjutan untuk menilai efektivitas tindakan yang diambil dalam mengentaskan kemiskinan. Dengan mengikuti saran-saran ini dan memahami hasil analisis klasterisasi, diharapkan upaya pengentasan kemiskinan di Indonesia dapat menjadi lebih terfokus dan berhasil dalam mengurangi tingkat kemiskinan.

#### DAFTAR PUSTAKA

- [1] E. Tania, "Upaya United Nations Development Programmed Melalui Proyek SWARGA dalam Meningkatkan Jumlah Perempuan sebagai Pengambil Keputusan di DPRD Provinsi DKI Jakarta," Universitas Katolik Parahyangan, Bandung, 2017. Accessed: Sep. 30, 2023. [Online]. Available: <https://repository.unpar.ac.id>
- [2] A. S. Alisjahbana and E. Murniningtyas, *Tujuan pembangunan berkelanjutan di Indonesia : konsep, target, dan strategi implementasi*. 2018.

- [3] A. Silomba, Muh. Massyat, and Muh. S. Tajuddin, "Peranan Badan Perencanaan Pembangunan, Penelitian Dan Pengembangan Daerah Dalam Proses Perumusan Kebijakan Pembangunan di Kabupaten Mamasa," *Journal Peqguruang: Conference Series*, vol. 2, no. 2, p. 423, Nov. 2020, doi: 10.35329/jp.v2i2.1519.
- [4] N. I. Febianto and N. Palasara, "Analisa Clustering K-Means Pada Data Informasi Kemiskinan Di Jawa Barat Tahun 2018," *Jurnal Sisfokom (Sistem Informasi dan Komputer)*, vol. 8, no. 2, pp. 130–140, Aug. 2019, doi: 10.32736/sisfokom.v8i2.653.
- [5] H. Liu, Y. Liu, R. Zhang, and X. Wu, "A Clustering Algorithm via Density Perception and Hierarchical Aggregation Based on Urban Multimodal Big Data for Identifying and Analyzing Categories of Poverty-Stricken Households in China," *Sci Program*, vol. 2021, 2021, doi: 10.1155/2021/6692975.
- [6] N. T. Luchia *et al.*, "Perbandingan K-Means dan K-Medoids Pada Pengelompokan Data Miskin di Indonesia," *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 2, no. 2, pp. 35–41, 2022.
- [7] M. A. Rahman, N. S. Sani, R. Hamdan, Z. A. Othman, and A. A. Bakar, "A clustering approach to identify multidimensional poverty indicators for the bottom 40 percent group," *PLoS One*, vol. 16, no. 8 August, Aug. 2021, doi: 10.1371/journal.pone.0255312.
- [8] A. Y. Clark, N. Blumenfeld, E. Lal, S. Darbari, S. Northwood, and A. Wadpey, "Using k-means cluster analysis and decision trees to highlight significant factors leading to homelessness," *Mathematics*, vol. 9, no. 17, Sep. 2021, doi: 10.3390/math9172045.
- [9] R. Y. Endartyana, "Pengelompokan Kabupaten/Kota Di Provinsi Papua Berdasarkan Indikator Angka Kemiskinan," , Institut Teknologi Sepuluh Nopember , Surabaya, 2019.
- [10] M. P. Repollo, R. Aurelius, and C. Robielos, "Applying Clustering Algorithm on Poverty Analysis in a Community in the Philippines," *IEOM Society International*, pp. 1511–1521, 2021.
- [11] N. Sepriyanti, R. Sani Nahampun, M. H. Zikri, I. Ambarani, and A. Rahmadeyan, "Penerapan K-Means Clustering Untuk Mengelompokkan Tingkat Kemiskinan di Provinsi Riau," *SENTIMAS: Seminar Nasional Penelitian dan Pengabdian Masyarakat*, pp. 59–65, 2022, [Online]. Available: <https://journal.irpi.or.id/index.php/sentimas>
- [12] M. Tadi and B. Arnawisuda Ningsi, "Analisis Klaster Kemiskinan Kabupaten Kota Di Provinsi Banten Menggunakan Metode K-Means," *Lebesgue: Jurnal Ilmiah Pendidikan Matematika, Matematika dan Statistika*, vol. 4, no. 1, pp. 374–385, 2023, doi: 10.46306/lb.v4i1.
- [13] F. Dwi Ratna Sari and S. Partawi Ediwijojo, "Clustering Analysis Using K-Medoids on Poverty Level Problems in Central Java by District/City," *KnE Social Sciences*, pp. 78–87, May 2023, doi: 10.18502/kss.v8i9.13321.
- [14] H. Sofyan, M. Iqbal, M. Marzuki, and M. Muhammad, "The comparison of k-modes clustering and ROCK clustering to the poverty indicator in Samadua Subdistrict, South Aceh," *IOP Conf Ser Mater Sci Eng*, vol. 1087, no. 1, p. 012085, Feb. 2021, doi: 10.1088/1757-899x/1087/1/012085.
- [15] Z. Sitorus, "Penerapan Data Mining Untuk Clustering Penduduk Miskin Di Kota Tanjungbalai Menggunakan Metode Algoritma K-Means," *Journal of Science and Social Research*, no. 1, pp. 212–218, 2024, [Online]. Available: <http://jurnal.goretanpena.com/index.php/JSSR>
- [16] A. Bahauddin, "Analisis Clustering Provinsi Di Indonesia Berdasarkan Tingkat Kemiskinan Menggunakan Algoritma K-Means," *MISI (Jurnal Manajemen informatika & Sistem Informasi)*, vol. 4, no. 1, Jan. 2021.
- [17] K. Niswatin, "Clustering Of Districts And Cities In Indonesia Based On Poverty Indicators Using The K-Means Method," *International Conference on Computing, Mathematics and Statistics*, 2021.
- [18] N. Istiqamah, O. Soesanto, and D. Anggraini, "Application of the K-Means algorithm to determine poverty status in Hulu Sungai Tengah," *J Phys Conf Ser*, vol. 2106, no. 1, Nov. 2021, doi: 10.1088/1742-6596/2106/1/012027.
- [19] R. Novita *et al.*, "Penerapan Algoritma K-Means dan Analisisnya untuk Menentukan Kebijakan Strategis Penyelesaian Studi Mahasiswa," *Jurnal SAINTIKOM*, vol. 22, pp. 401–413, 2023, [Online]. Available: <https://ojs.trigunadharma.ac.id/index.php/jis/index>
- [20] M. Atiqul Mutaqin and W. Andriyani, "KLASTERISASI DATA DISABILITAS MENGGUNAKAN ALGORITMA K-MEANS," *IJIR*, vol. 3, no. 1, pp. 25–35, 2022.
- [21] P.-N. Tan, M. Steinbach, and V. Kumar, *Introduction to data mining*.
- [22] M. Cui, "Introduction to the K-Means Clustering Algorithm Based on the Elbow Method," *Clausius Scientific Press, Canada*, vol. 1, pp. 5–8, 2020, doi: 10.23977/accaf.2020.010102.
- [23] V. Annisa Ekasetya and A. Jananto, "Klusterisasi Optimal Dengan Elbow Method Untuk Pengelompokan Data Kecelakaan Lalu Lintas Di Kota Semarang," *Dinamika Informatika*, vol. 12, no. 1, pp. 20–28, 2020.



- [24] N. Syahfitri, E. Budianita, A. Nazir, and I. Afrianty, "Pengelompokan Produk Berdasarkan Data Persediaan Barang Menggunakan Metode Elbow dan K-Medoid," *KLIK: Kajian Ilmiah Informatika dan Komputer*, vol. 4, no. 3, pp. 1668–1675, 2023, doi: 10.30865/klik.v4i3.1525.
- [25] G. Ogbuabor and U. F. N, "Clustering Algorithm for a Healthcare Dataset Using Silhouette Score Value," *International Journal of Computer Science and Information Technology*, vol. 10, no. 2, pp. 27–37, Apr. 2018, doi: 10.5121/ijcsit.2018.10203.
- [26] A. Naghizadeh and D. N. Metaxas, "Condensed silhouette: An optimized filtering process for cluster selection in K-means," in *Procedia Computer Science*, Elsevier B.V., 2020, pp. 205–214. doi: 10.1016/j.procs.2020.08.022.
- [27] K. R. Shahapure and C. Nicholas, "Cluster quality analysis using silhouette score," in *Proceedings - 2020 IEEE 7th International Conference on Data Science and Advanced Analytics, DSAA 2020*, Institute of Electrical and Electronics Engineers Inc., Oct. 2020, pp. 747–748. doi: 10.1109/DSAA49011.2020.00096.