



Volume 4, Nomor 2, Desember 2023, hlm 137-146
INFOTECH: Jurnal Informatika & Teknologi
p ISSN 2722-9378 | e ISSN 2722-9386
<http://jurnal.sttcileungsi.ac.id/index.php/infotech>

Clustering daerah rawan *stunting* di Jawa Barat menggunakan algoritma *k-means*

Clustering stunting-prone areas in West Java using k-means algorithm

Taufik Hidayat*, Mohamad Jajuli, Susilawati

*Program Studi Informatika, Fakultas Ilmu Komputer, Universitas Singaperbangsa Karawang, Karawang, Jawa Barat. Jl. HS. Ronggo Waluyo, Puseurjaya, Kec. Telukjambe Timur, Kabupaten Karawang, Jawa Barat 41361

Informasi Artikel

Article History:

Submission: 12-06-2023

Revised: 24-06-2023

Accepted: 25-06-2023

Kata Kunci:

Stunting; Clustering; K-Means; Silhouette Coefficient; CRISP-DM.

Keywords:

Stunting; Clustering; K-Means; Silhouette Coefficient; CRISP-DM.

* Korespondensi:

Taufik Hidayat

taufiik14@gmail.com

Abstrak

Stunting atau kekerdilan pada balita adalah suatu kondisi yang terjadi karena gagal tumbuh yang diakibatkan oleh kekurangan gizi kronis dan infeksi berulang selama periode 1000 hari pertama kehidupan. Angka prevalensi *stunting* di Jawa Barat pada tahun 2021 sebesar 24,5% dan termasuk ke dalam lima provinsi di Indonesia dengan jumlah kasus *stunting* tertinggi. Permasalahan *stunting* tidak bisa dibiarkan dan harus ditangani karena dapat menimbulkan dampak negatif bagi kesehatan, baik jangka panjang maupun jangka pendek. Penelitian ini bertujuan untuk mengelompokkan daerah rawan *stunting* di Jawa Barat pada tahun 2019-2021 menggunakan algoritma *K-Means*. Metode yang digunakan yaitu *Cross Industry Standard Process for Data Mining (CRISP-DM)*. Data yang digunakan adalah data kejadian *stunting* dan faktor risiko *stunting* dari 27 kabupaten/kota di Jawa Barat dan bersumber dari situs resmi Open Data Jawa Barat. Penelitian diawali dengan menentukan *K* terbaik dengan *Elbow Method*, dimana $K=3$ adalah jumlah *cluster* terbaik. Hasil *clustering* dievaluasi dengan *silhouette coefficient* diperoleh nilai evaluasi sebesar 0,61. Nilai tersebut berada dalam kategori *cluster* layak dan sudah sesuai dengan masing-masing *cluster* yang terbentuk.

Abstract

Stunting in toddlers is a condition that occurs due to growth failure caused by chronic malnutrition and repeated infections during the first 1000 days of life. The stunting prevalence rate in West Java in 2021 was 24.5% and was included in the five provinces in Indonesia with the highest number of stunting cases. The problem of stunting cannot be ignored and must be addressed because it can have a negative impact on health, both long and short term. This study aims to classify stunting-prone areas in West Java in 2019-2021 using the K-Means algorithm. The method used is the Cross Industry Standard Process for Data Mining (CRISP-DM). The data used are data on the incidence of stunting and risk factors for stunting from 27 districts / cities in West Java and are sourced from the official Open Data Jawa Barat website. The research begins by determining the best K with the Elbow Method, where $K = 3$ is the best number of clusters. The clustering results were evaluated with the silhouette coefficient and an evaluation value of 0.61 was obtained. This value is in the category of feasible clusters and is in accordance with each cluster formed.



Infotech: Jurnal Informatika & Teknologi is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

1. PENDAHULUAN.

Kesehatan merupakan salah satu hak yang menaungi secara hukum oleh undang-undang bagi semua warga negara, karena kesehatan merupakan kebutuhan dasar manusia. Negara mengakui bahwa kesehatan adalah aset terbesar untuk mencapai kesejahteraan. Tetapi tak dapat dipungkiri bahwa setiap negara memiliki permasalahan mengenai gizi, begitu pun dengan negara berkembang seperti Indonesia. Berdasarkan riset yang dilakukan oleh BPS dan Kementerian Kesehatan melalui hasil Studi Status Gizi Indonesia (SSGI) tahun 2021, terungkap bahwa angka prevalensi *stunting* di Indonesia mencapai 24,4%. Angka ini menunjukkan penurunan sebesar 1,6% per tahun dari 2019 [1].

Stunting merupakan kondisi kerdil/pendek yang dialami oleh anak berusia di bawah lima tahun (balita) karena gagal tumbuh yang disebabkan oleh kekurangan gizi kronis serta infeksi berulang selama periode 1000 Hari Pertama Kehidupan (HPK) terhitung mulai dari janin hingga usia anak berulang tahun yang kedua [2]. Pernyataan yang dikutip dari *World Health Organizations* (WHO) *stunting* dapat diukur berdasarkan indeks PB/U (panjang badan dibanding umur) atau TB/U (tinggi badan dibanding umur) menggunakan batas (z-score) dengan nilai standar deviasi kurang dari -2 [3].

Selaras dengan komitmen negara Indonesia yang terus berupaya menekan laju kejadian *stunting*, sebagaimana yang tertuang dalam Peraturan Presiden No. 72 Tahun 2021 tentang penurunan percepatan *stunting*. Tahun 2021 SSGI mengungkapkan bahwa prevalensi balita dengan kondisi *stunting* di Provinsi Jawa Barat sebesar 24,5% [4]. Perbandingan jumlah penduduk dalam kasus balita yang mengalami *stunting* yakni 2:10, artinya dua dari sepuluh balita di Jawa Barat mengalami *stunting*. Jawa Barat termasuk ke dalam lima provinsi dengan jumlah kasus *stunting* tertinggi yang diikuti oleh provinsi Jawa Timur, Jawa Tengah, Sumatera Utara dan Banten. Hal ini disampaikan oleh Wakil Menteri Kesehatan dalam Forum *Stunting* Nasional 2022 [5]. Dampak buruk yang dapat terjadi akibat *stunting* menyebabkan menurunnya kemampuan produktivitas dan meningkatkan kemiskinan suatu bangsa di masa depan [6].

Berdasarkan penelitian [7] faktor risiko yang menyebabkan *stunting* pada balita di Asia Tenggara dan negara berkembang antara lain bayi berat lahir rendah (BBLR) dan status gizi bayi, pendidikan ibu, pendapatan rumah tangga, dan faktor dominan terhadap risiko *stunting* adalah tingkat sanitasi yang tidak baik.

Data mining adalah teknik menemukan pola, model, dan jenis pengetahuan lain yang menarik dalam sebuah *dataset* [8]. Teknik ini mencakup berbagai metode dan algoritma untuk mengidentifikasi pola, hubungan, serta tren dalam data yang bisa dimanfaatkan untuk mendukung proses pengambilan keputusan. Dalam literatur yang ditulis oleh [8] tujuan utama dari *data mining* adalah untuk mendapatkan dan menggali pengetahuan yang berharga dari data yang terkumpul, sehingga dapat membantu meningkatkan efisiensi dan efektivitas proses bisnis, meningkatkan kualitas produk atau layanan, serta membantu memprediksi tren dan perilaku konsumen. Metode yang diterapkan pada penelitian ini adalah pengelompokan data (*clustering*). *Clustering* didefinisikan sebagai salah satu metode mengelompokkan sampel (*instance*) menjadi beberapa grup atau *subset* atau *cluster* berdasarkan kemiripan dengan *instance* lain [9].

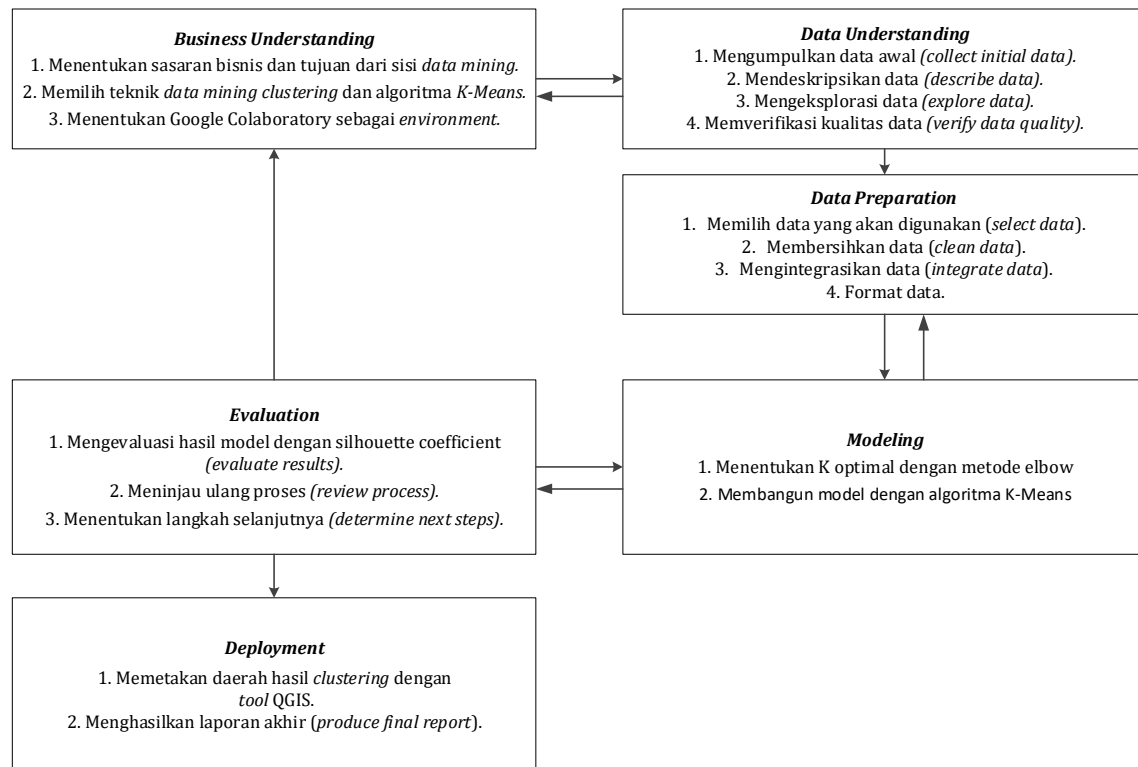
Metode *clustering K-Means* adalah suatu teknik analisis data non hierarki yang digunakan untuk mempartisi sebuah *dataset* ke dalam beberapa kelompok. Teknik ini membagi data menjadi beberapa kelompok yang memiliki karakteristik serupa, dan mengelompokkan data yang memiliki perbedaan karakteristik ke dalam kelompok yang berbeda [10]. Algoritma *K-Means* tergolong ke dalam algoritma sederhana karena dapat melakukan pengelompokan data berjumlah besar dalam waktu komputasi yang cepat dan efisien [11]. Hal ini diperkuat dengan penelitian [12] mengenai analisis pengelompokan ketahanan pangan di setiap provinsi di Indonesia, ditemukan bahwa algoritma *K-Means* memiliki *silhouette coefficient* tertinggi dibandingkan dengan algoritma lain seperti *Self Organizing Maps*, *DBSCAN*, *Louvain Clustering*. Hal ini menunjukkan bahwa *K-Means* lebih efektif dalam mengelompokkan data dan memiliki akurasi yang lebih tinggi dibandingkan dengan algoritma lain yang telah diuji.

Penelitian mengenai *clustering* daerah *stunting* ini sudah pernah dilakukan sebelumnya oleh M. Y. Matdoan, U. A. Matdoan, dan M. Saleh Far-far untuk klasifikasi Provinsi di Indonesia berdasarkan paket pelayanan *stunting* menggunakan *tools* RapidMiner, hasil dari penelitian tersebut mengelompokkan provinsi di Indonesia menjadi tiga *cluster* (0-2). 25 provinsi termasuk ke dalam *cluster* 0. *Cluster* 1 dengan jumlah 4 provinsi dan yang termasuk ke dalam *cluster* 2 berjumlah 5 provinsi.. Penelitian kembali dilakukan oleh [14] dengan judul penerapan algoritma *K-Means* dalam klasterisasi kasus *stunting* balita Desa Tegalwangi yang menghasilkan 2 kelas yaitu *cluster* 0 dengan 392 balita dan *cluster* 1 dengan jumlah 3 balita. Hasil pengujian dievaluasi dengan salah satu evaluasi internal yaitu *Davies-Bouldin Index* (DBI) dengan bantuan *tools* RapidMiner. Pengelompokan daerah rawan *stunting* dengan algoritma *K-Means* terbukti dapat membantu dan menghasilkan informasi untuk pemetaan daerah *stunting*.

Berdasarkan penelitian serupa dan uraian yang telah dijabarkan diatas, agar dapat melakukan analisis dan pemetaan daerah rawan *stunting* di Provinsi Jawa Barat maka diperlukan teknik pengelompokan daerah berdasarkan Kabupaten/Kota. Peneliti bermaksud melakukan penelitian pengelompokan daerah *stunting* dengan algoritma *K-Means clustering* dan hasil pemetaan daerah divisualisasikan dengan aplikasi QGIS.

2. METODE

Metode yang digunakan pada penelitian ini yaitu mengikuti tahapan yang terdapat dalam proses *data mining* dengan menggunakan tahapan CRISP-DM yang dibangun dari 6 tahapan [15] yang direpresentasikan melalui rancangan penelitian pada [Gambar 1](#).



Gambar 1. Rancangan penelitian

2.1 Data Preparation

Setelah melewati tahap *business understanding* dan *data understanding*. Langkah selanjutnya adalah *data preparation*. Tahap ini merupakan tahap persiapan data dari *dataset* awal yang telah dikumpulkan sebelumnya. *Data preparation* melibatkan serangkaian aktivitas pengolahan data mentah atau *preprocessing data* untuk menghasilkan set data final yang akan digunakan dalam proses pemodelan. Melalui tahap *data preparation* ini, dilakukan proses pemilihan atribut dari

keempat *dataset* tersebut. Atribut yang digunakan yaitu “Nama Kabupaten Kota”, “Jumlah Balita *Stunting*”, “Jumlah Bayi Tidak dengan ASI Eksklusif”, “Jumlah Berat Bayi Lahir Rendah”, “Jumlah Kepala Keluarga dengan Akses Sanitasi Tidak Layak”, dan “Tahun”. Pada atribut “Nama Kabupaten Kota” dan “Tahun” digunakan sebagai *identifier* saja sehingga tidak perlu dilakukan proses *encoder*.

Kemudian setelah memilih atribut yang akan digunakan, dilakukan proses pengecekan kualitas data. Hasilnya menunjukkan tidak terdapat *missing value* dan data duplikat dalam data. Tetapi terdapat *outlier* pada atribut “Jumlah Balita *Stunting*”, “Jumlah Bayi Tidak dengan ASI Eksklusif”, dan “Jumlah Kepala Keluarga dengan Akses Sanitasi Tidak Layak”. Untuk menangani *outlier* digunakan teknik *winsorizing*. *Winsorizing* adalah metode populer untuk mengurangi bobot *outlier* dengan menggantinya dengan persentil tertentu dari nilai yang bergantung pada data [15]. Untuk mengatasi rentang data yang memiliki rentang yang berbeda pada setiap atributnya diperlukan teknik normalisasi. Normalisasi yang digunakan menggunakan metode *min-max normalization* dengan mengolah nilai minimum dan maksimum dari setiap atribut sehingga memiliki rentang 0-1 [16]. *Dataset* kasus *stunting* di Jawa Barat pada tahun 2019-2021 yang siap untuk masuk ke tahap berikutnya yaitu modeling disajikan dalam bentuk **Tabel 1**.

Tabel 1. *Dataset* akhir sebelum masuk tahap *modeling*

Nama Kabupaten Kota	Jumlah Balita <i>Stunting</i>	Jumlah Bayi Tidak dengan ASI Eksklusif	Jumlah Bayi Berat Lahir Rendah (BBLR)	Jumlah Kepala Keluarga dengan Akses Sanitasi Tidak Layak	Tahun
KABUPATEN BOGOR	0.91250256	1	0.80324643	1	2019
KABUPATEN SUKABUMI	0.75025599	0.80728634	1	0.70903624	2019
...
KOTA TASIKMALAYA	0.26955765	0.03589985	0.19429415	0.17965994	2021
KOTA BANJAR	0	0.00979831	0	0.00758956	2021

2.2 Modeling

Setelah melakukan tahap pengolahan data yakni berupa *data preprocessing* dan dihasilkan data yang telah sesuai untuk penerapan metode *data mining* yakni metode *clustering*, proses selanjutnya adalah penerapan metode *clustering* menggunakan algoritma *K-Means* untuk menemukan karakteristik atau informasi dalam mengelompokkan daerah rawan *stunting* di provinsi Jawa Barat pada tahun 2019-2021. Untuk menentukan *K* optimal menggunakan metode *elbow*. Metode *elbow* adalah sebuah metode yang digunakan untuk memperoleh informasi dalam menentukan jumlah *cluster* yang optimal dengan melihat perbandingan persentase penurunan jumlah *cluster* yang membentuk sebuah "siku" pada titik tertentu [17].

2.3 Evaluation

Tahap evaluasi digunakan untuk menilai hasil model yang telah dibuat sebelumnya agar sesuai dengan tujuan yang akan dicapai dalam tahap *business understanding*. Teknik evaluasi yang digunakan adalah *Silhouette coefficient* merupakan suatu parameter yang menunjukkan sejauh mana sebuah analisis cocok dengan kelompoknya dalam mengelompokkan data. Kelompok yang terbentuk akan dianggap optimal jika *silhouette coefficient* yang dihasilkan mendekati 1. Sebaliknya, jika koefisien tersebut semakin mendekati angka -1, maka kelompok tersebut dianggap tidak optimal [10]. Persamaan nilai score *silhouette coefficient* dapat dituliskan sebagai rumus berikut [18].

$$S(i) = \frac{b(i) - a(i)}{\max(b(i), a(i))} \quad (1)$$

Keterangan:

$S(i)$: *silhouette score*

$a(i)$: jarak rata-rata antara data i dan semua objek dalam *cluster*

$b(i)$: jarak rata-rata antara data i ke semua objek di *cluster* lain

Untuk kriteria penilaian dari hasil akurasi *silhouette coefficient* dapat dilihat pada **Tabel 2**.

Tabel 2. Interpretasi *Silhouette Coefficient* [19]

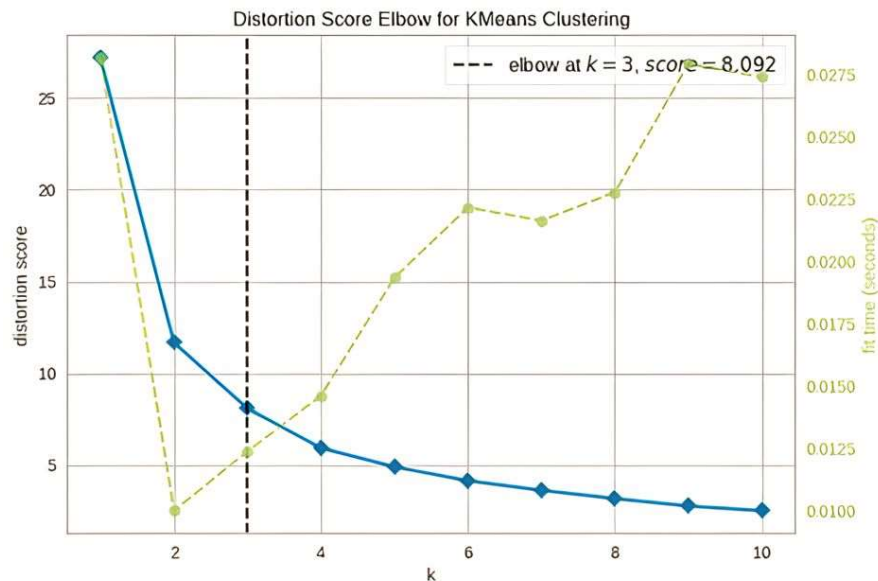
<i>Silhouette Coefficient</i>	Interpretasi
0.71-1.00	<i>Cluster</i> kuat
0.51-0.70	<i>Cluster</i> layak/sesuai
0.26-0.50	<i>Cluster</i> lemah
<0.25	Tidak dapat disebut <i>cluster</i>

2.4 Deployment

Setelah tahap evaluasi selesai dilakukan, tahap selanjutnya adalah *deployment*. Pada tahap ini, hasil analisis dan pemodelan yang telah dilakukan akan dibuat dalam bentuk visualisasi yang mudah dipahami. Tujuan dari tahap *deployment* ini adalah untuk memastikan bahwa hasil analisis dapat digunakan dan dimanfaatkan secara optimal untuk mencapai tujuan yang telah ditentukan pada tahap awal yaitu untuk menekan laju kejadian *stunting* yang ada di Jawa Barat. Visualisasi menggunakan Quantum GIS. QGIS adalah aplikasi *desktop open-source* yang digunakan sebagai sistem informasi geografis (SIG) yang dapat melakukan berbagai fungsi seperti melihat, mengedit, dan menganalisis data. QGIS dapat dijalankan pada berbagai sistem operasi, termasuk Mac OS, Linux, UNIX, dan Microsoft Windows. QGIS diberi lisensi *General Public License* (GPL), yang memungkinkan pengguna untuk memodifikasi aplikasi sesuai kebutuhan atau tujuan spesifik lainnya [20].

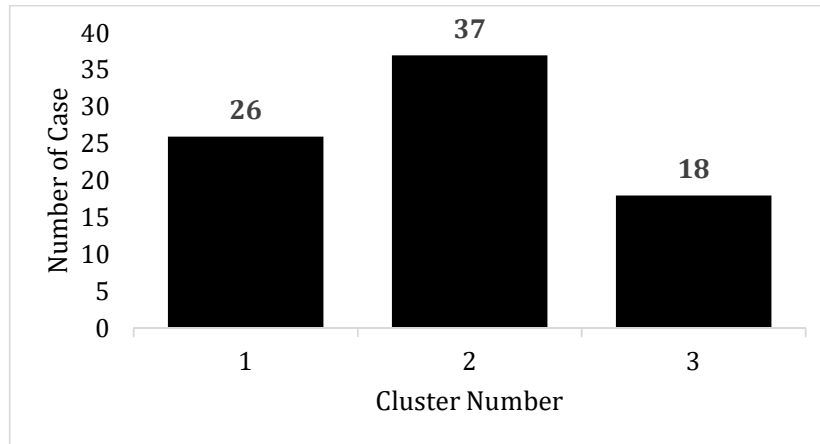
3. HASIL DAN PEMBAHASAN

Penelitian ini menghasilkan tiga *cluster* berdasarkan Kabupaten Kota di Provinsi Jawa Barat pada tahun 2019-2021. Nilai K optimal dihasilkan dari grafik metode *elbow* yang direpresentasikan pada **Gambar 2**.



Gambar 2. Hasil K optimal dengan elbow visualizer

Kemudian dilakukan *modeling* dengan algoritma *K-Means* dan menunjukkan bahwa dari data kasus *stunting* di Jawa Barat pada tahun 2019-2021 yang dibuat menjadi tiga *cluster* yaitu *cluster* 1 dengan jumlah 26 anggota, kemudian *cluster* 2 dengan 37 anggota dan *cluster* 3 dengan jumlah 18 anggota. Jumlah anggota hasil *cluster* dapat dilihat pada **Gambar 3**.



Gambar 3. Visualisasi hasil *clustering K-Means*

Pelabelan *cluster* pada **Tabel 3** digunakan untuk memberi nama dari setiap *cluster* yang dihasilkan oleh algoritma *K-Means clustering* dalam kasus *stunting* di Jawa Barat berdasarkan jumlah rata-rata semua atribut. Rata-rata jumlah atribut keseluruhan yang paling tinggi diberi label daerah rawan *stunting* tinggi.

Tabel 3. Pelabelan *cluster*

<i>Cluster</i>	Jumlah Rata-Rata Setiap Atribut	Label
3	855000.7778	Rawan <i>Stunting</i> Tinggi
2	627221.7027	Rawan <i>Stunting</i> Sedang
1	209432.2308	Rawan <i>Stunting</i> Rendah

Tabel 4 menunjukkan hasil *clustering* pada tahun 2019 terbagi menjadi 3 *cluster*. *Cluster 1* dengan label daerah rawan *stunting* yang rendah ditempati oleh 9 Kabupaten/Kota. Kemudian *cluster 2* dengan label daerah rawan *stunting* sedang ditempati oleh 12 Kabupaten/Kota. Sedangkan untuk daerah rawan *stunting* yang tinggi berada di *cluster 3* ditempati oleh 6 Kabupaten/Kota.

Tabel 4. Hasil *cluster* tahun 2019

Nama Kabupaten Kota		
<i>Cluster 1</i> (Rawan Rendah)	<i>Cluster 2</i> (Rawan Sedang)	<i>Cluster 3</i> (Rawan Tinggi)
Kabupaten Sumedang	Kabupaten Cianjur	Kabupaten Bogor
Kabupaten Purwakarta	Kabupaten Ciamis	Kabupaten Sukabumi
Kabupaten Pangandaran	Kabupaten Kuningan	Kabupaten Bandung
Kota Bogor	Kabupaten Majalengka	Kabupaten Garut
Kota Sukabumi	Kabupaten Indramayu	Kabupaten Tasikmalaya
Kota Cirebon	Kabupaten Subang	Kabupaten Cirebon
Kota Cimahi	Kabupaten Karawang	
Kota Tasikmalaya	Kabupaten Bekasi	
Kota Banjar	Kabupaten Bandung Barat	
	Kota Bandung	
	Kota Bekasi	
	Kota Depok	

Tabel 5 menunjukkan hasil *clustering* pada tahun 2020 yang telah dikelompokkan menjadi 3 *cluster* berdasarkan tingkat keparahan daerah rawan *stunting*. *Cluster 1* yang memiliki label

daerah rawan *stunting* yang rendah terdiri dari 8 Kabupaten/Kota. Sementara itu, 13 Kabupaten/Kota tergabung dalam *cluster 2* yang memiliki label daerah rawan *stunting* sedang. Terakhir, *cluster 3* yang memiliki label daerah rawan *stunting* yang tinggi terdiri dari 6 Kabupaten/Kota.

Tabel 5. Hasil *Cluster* Tahun 2020

Nama Kabupaten Kota		
<i>Cluster 1</i> (Rawan Rendah)	<i>Cluster 2</i> (Rawan Sedang)	<i>Cluster 3</i> (Rawan Tinggi)
Kabupaten Purwakarta	Kabupaten Cianjur	Kabupaten Bogor
Kabupaten Pangandaran	Kabupaten Ciamis	Kabupaten Sukabumi
Kota Bogor	Kabupaten Kuningan	Kabupaten Bandung
Kota Sukabumi	Kabupaten Majalengka	Kabupaten Garut
Kota Cirebon	Kabupaten Sumedang	Kabupaten Tasikmalaya
Kota Cimahi	Kabupaten Indramayu	Kabupaten Cirebon
Kota Tasikmalaya	Kabupaten Subang	
Kota Banjar	Kabupaten Karawang	
	Kabupaten Bekasi	
	Kabupaten Bandung Barat	
	Kota Bandung	
	Kota Bekasi	
	Kota Depok	

Tabel 6 menunjukkan hasil *clustering* pada tahun 2021 terbagi menjadi 3 *cluster*. *Cluster 1* dengan label daerah rawan *stunting* yang rendah ditempati oleh 9 Kabupaten/Kota. Kemudian *cluster 2* dengan label daerah rawan *stunting* sedang ditempati oleh 12 Kabupaten/Kota. Sedangkan untuk daerah rawan *stunting* tinggi berada pada *cluster 3* ditempati oleh 6 Kabupaten/Kota.

Tabel 6. Hasil *Cluster* Tahun 2021

Nama Kabupaten Kota		
<i>Cluster 1</i> (Rawan Rendah)	<i>Cluster 2</i> (Rawan Sedang)	<i>Cluster 3</i> (Rawan Tinggi)
Kabupaten Ciamis	Kabupaten Tasikmalaya	Kabupaten Bogor
Kabupaten Purwakarta	Kabupaten Kuningan	Kabupaten Sukabumi
Kabupaten Pangandaran	Kabupaten Majalengka	Kabupaten Cianjur
Kota Bogor	Kabupaten Sumedang	Kabupaten Bandung
Kota Sukabumi	Kabupaten Indramayu	Kabupaten Garut
Kota Cirebon	Kabupaten Subang	Kabupaten Cirebon
Kota Cimahi	Kabupaten Karawang	
Kota Tasikmalaya	Kabupaten Bekasi	
Kota Banjar	Kabupaten Bandung Barat	
	Kota Bandung	
	Kota Bekasi	
	Kota Depok	

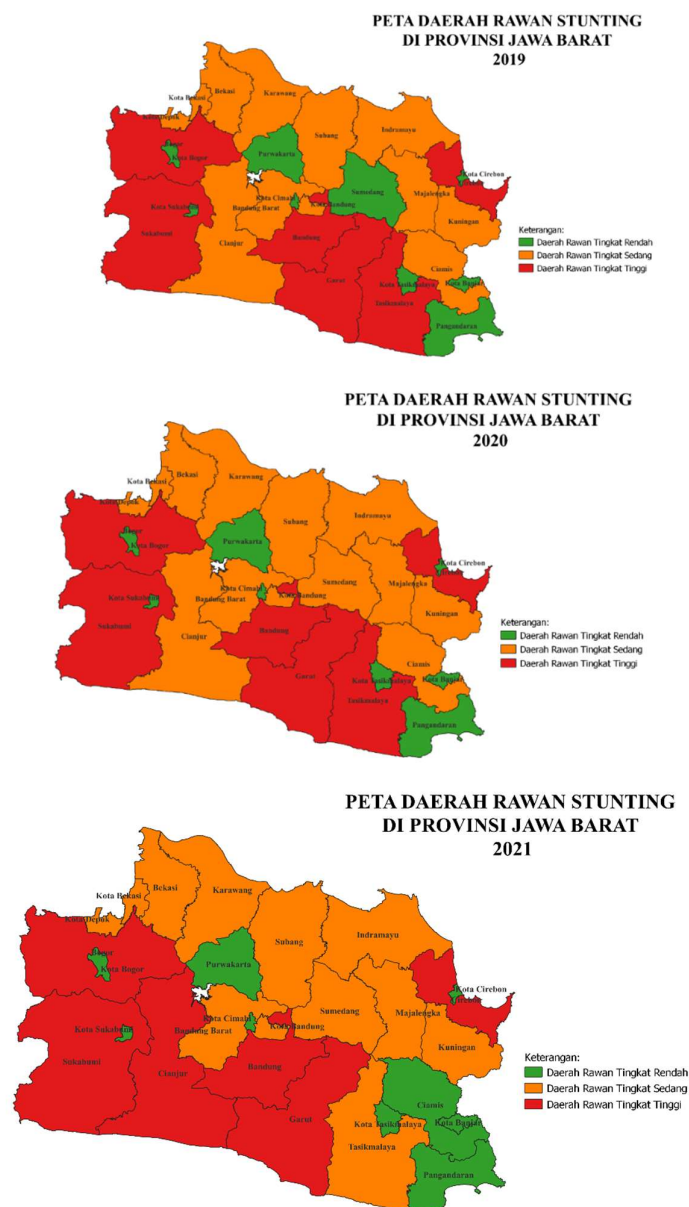
Hasil dari *clustering* akan dilakukan evaluasi guna melihat keakuratan hasil yang dicapai dalam proses modeling dengan menggunakan metode tertentu. Hasil *clustering* yang didapat dari jumlah *cluster K=3* dievaluasi menggunakan *silhouette coefficient*.

```
[24] silhouette_scores = silhouette_score(data, labels)
print("Silhouette Score:", silhouette_scores)
```

Silhouette Score: 0.6162003626003247

Gambar 4. Hasil Pengujian dengan *Silhouette Coefficient*

Hasil uji menggunakan silhouette coefficient atau skor silhouette yang terlihat pada Gambar 4 menunjukkan nilai indeks sebesar 0,61 pada penggunaan algoritma K-Means untuk melakukan clustering data kasus daerah rawan stunting di Jawa Barat tahun 2019-2021. Kualitas cluster yang dihasilkan oleh algoritma K-Means pada penelitian ini dikategorikan sebagai penempatan kluster yang layak atau dapat dikatakan hasil cluster sudah sesuai dengan karakteristiknya. Uji ini dilakukan untuk mengevaluasi hubungan dekat antara objek-objek dan seberapa jauh kluster-kluster terpisah, sehingga dapat diketahui kualitas suatu kluster.



Gambar 5. Pemetaan daerah rawan stunting di Jawa Barat 2019-2021

Pada **Gambar 5** pemetaan menggunakan *tool* QGIS dibagi menjadi 3 warna keterangan untuk menggambarkan visualisasi hasil *cluster*. Masing-masing warna merepresentasikan kategori dari tiap *cluster*. Pada peta bagian kabupaten kota yang berwarna hijau merupakan *cluster* dengan daerah rawan *stunting* tingkat rendah, daerah berwarna oranye merupakan *cluster* dengan daerah rawan *stunting* tingkat sedang, serta *cluster* dengan daerah rawan *stunting* tingkat tinggi ditandai dengan warna merah.

4. SIMPULAN

Penelitian ini menerapkan teknik *data mining clustering* menggunakan algoritma *K-Means* untuk mengelompokkan daerah berdasarkan kasus balita *stunting* di Provinsi Jawa Barat pada tahun 2019, 2020, dan 2021. Jumlah *cluster* yang terbentuk sebanyak 3 *Cluster* diperoleh dari metode *elbow* dan menghasilkan anggota *cluster* yang berbeda di setiap tahunnya. Kemudian dievaluasi menggunakan *silhouette coefficient*, diperoleh nilai 0,61 yang masuk dalam kriteria *cluster* layak atau *cluster* yang terbentuk sudah sesuai dengan karakteristiknya. Visualisasi pemetaan daerah rawan *stunting* menunjukkan adanya Kabupaten/Kota yang konsisten berada di tingkat tinggi, sedang, dan rendah selama tiga tahun berturut-turut. Selain itu, terdapat Kabupaten/Kota yang mengalami perubahan level dari tingkat rendah menjadi sedang, serta Kabupaten/Kota yang mengalami penurunan level dari tingkat tinggi menjadi sedang

DAFTAR PUSTAKA

- [1] Kemenkes and Biro Komunikasi dan Pelayanan Masyarakat, "PENURUNAN PREVALENSI STUNTING TAHUN 2021 SEBAGAI MODAL MENUJU GENERASI EMAS INDONESIA 2045," *Kementrian Kesehatan Republik Indonesia*, Dec. 27, 2021.
- [2] M. Ariani, "Determinan Penyebab Kejadian Stunting Pada Balita: Tinjauan Literatur," *Dinamika Kesehatan Jurnal Kebidanan dan Keperawatan*, vol. 11, no. 1, pp. 172–186, 2020, doi: 10.33859/dksm.v11i1.
- [3] Komalasari, E. Supriati, R. Sanjaya, and H. Ifayanti, "Faktor-Faktor Penyebab Kejadian Stunting Pada Balita," *Majalah Kesehatan Indonesia*, vol. 1, no. 2, pp. 51–56, 2020.
- [4] R. Lestari and Dinkes Jabar, "Komitmen Dinkes Jabar Tangani Kasus Stunting," *Dinas Kesehatan Provinsi Jawa Barat*, Jun. 11, 2022.
- [5] Malini, "Angka Stunting di Indonesia Masih Tinggi, Ini 5 Provinsi dengan Kasus Stunting Terbanyak," <https://nasional.tempo.co/read/1683885/angka-stunting-di-indonesia-masih-tinggi-ini-5-provinsi-dengan-kasus-stunting-terbanyak>, Jakarta, pp. 1–1, Jan. 25, 2023.
- [6] G. R. Suraya and A. W. Wijayanto, "Comparison of Hierarchical Clustering, K-Means, K-Medoids, and Fuzzy C-Means Methods in Grouping Provinces in Indonesia according to the Special Index for Handling Stunting," *Indonesian Journal of Statistics and Its Applications*, vol. 6, no. 2, pp. 180–201, Aug. 2022, doi: 10.29244/ijsa.v6i2p180-201.
- [7] G. Apriluana and S. Fikawati, "Analisis Faktor-Faktor Risiko terhadap Kejadian Stunting pada Balita (0-59 Bulan) di Negara Berkembang dan Asia Tenggara," *Media Penelitian dan Pengembangan Kesehatan*, vol. 28, no. 4, pp. 247–256, Dec. 2018, doi: 10.22435/mpk.v28i4.472.
- [8] J. Han, J. Pei, and H. Tong, "DATA MINING: CONCEPTS AND TECHNIQUES," in *Elsevier*, S. Merken and B. LoGuidice, Eds., 4th ed. Cambridge: Katey Britcher, 2022, pp. 1–752.
- [9] B. Purnama, *PENGANTAR MACHINE LEARNING*, 1st ed., vol. 1. Bandung: Informatika Bandung, 2019.
- [10] N. Thamrin and A. W. Wijayanto, "Comparison of Soft and Hard Clustering: A Case Study on Welfare Level in Cities on Java Island," *Indonesian Journal of Statistics and Its Applications*, vol. 5, no. 1, pp. 141–160, Mar. 2021, doi: 10.29244/ijsa.v5i1p141-160.
- [11] N. K. S. Julyantari, I. K. Budiarta, and N. M. D. K. Putri, "Implementasi K-Means Untuk Pengelompokan Status Gizi Balita (Studi Kasus Banjar Titih)," *Jurnal Janitra Informatika dan Sistem Informasi*, vol. 1, no. 2, pp. 92–101, 2021, doi: 10.25008/janitra.
- [12] A. Ramadhan, K. Prawita, M. A. Izzudin, and G. Amandha, "Analisis Strategi dan Klasterisasi Ketahanan Pangan Nasional dalam Menghadapi Pandemi Covid-19," *Teknologi Pangan* :

- Media Informasi dan Komunikasi Ilmiah Teknologi Pertanian*, vol. 12, no. 1, pp. 110–122, Mar. 2021, doi: 10.35891/tp.v12i1.2179.
- [13] M. Y. Matdoan, U. A. Matdoan, and M. Saleh Far-Far, “Algoritma K-Means Untuk Klasifikasi Provinsi di Indonesia Berdasarkan Paket Pelayanan Stunting,” *PANRITA Journal of Science, Technology, and Arts*, vol. 1, no. 2, pp. 41–46, 2022, [Online]. Available: <https://journal.dedikasi.org/pjsta>
- [14] P. Apriyani, A. R. Dikananda, and I. Ali, “Penerapan Algoritma K-Means dalam Klasterisasi Kasus Stunting Balita Desa Tegalwangi,” *Hello World Jurnal Ilmu Komputer*, vol. 2, no. 1, pp. 21–33, 2023, doi: 10.56211/helloworld.v2i1.230.
- [15] H. Liao, Y. Li, and G. P. Brooks, “Outlier impact and accommodation on power,” *Journal of Modern Applied Statistical Methods*, vol. 16, no. 1, pp. 261–278, 2017, doi: 10.22237/jmasm/1493597640.
- [16] D. Azzahra Nasution, H. H. Khotimah, and N. Chamidah, “PERBANDINGAN NORMALISASI DATA UNTUK KLASIFIKASI WINE MENGGUNAKAN ALGORITMA K-NN,” vol. 4, no. 1, pp. 78–82, 2019.
- [17] M. Cui, “Introduction to the K-Means Clustering Algorithm Based on the Elbow Method,” *Clausius Scientific Press*, vol. 1, pp. 5–8, 2020, doi: 10.23977/accaf.2020.010102.
- [18] F. R. Senduk, Indwiarti, and F. Nhita, “Clustering of Earthquake Prone Areas in Indonesia Using K-Medoids Algorithm,” *Indonesia Journal of Computing*, vol. 4, no. 3, pp. 65–76, 2019, doi: 10.21108/indojc.2019.4.3.359.
- [19] T. D. Harjanto, A. Vatesia, and R. Faurina, “ANALISIS PENETAPAN SKALA PRIORITAS PENANGANAN BALITA STUNTING MENGGUNAKAN METODE DBSCAN CLUSTERING,” *Jurnal Rekursif*, vol. 9, no. 1, pp. 30–42, 2021, [Online]. Available: <http://ejournal.unib.ac.id/index.php/rekursif/30>
- [20] A. A. Sasoeng, S. R. Sentinuwo, and Y. D. Y. Rindengan, “Rancang Bangun Sistem Informasi Geografis Potensi Sumber Daya Alam Di Kabupaten Talaud Berbasis Web,” *Jurnal Teknik Informatika*, vol. 13, no. 1, pp. 1–8, 2018.