

IMPLEMENTASI ALGORITMA *K-MEANS++* UNTUK KLASTERISASI DAERAH RAWAN *STUNTING* DI KABUPATEN TASIKMALAYA

Bintang Ary Pradana, Betha Nurina Sari

Program Studi Informatika, Universitas Singaperbangsa Karawang
Jl. HS. Ronggo Waluyo, Telukjambe Timur, Karawang, Indonesia
2210631170060@student.unsika.ac.id

ABSTRAK

Stunting merupakan masalah kesehatan yang hingga kini masih menjadi tantangan besar di Indonesia, termasuk di Kabupaten Tasikmalaya. Untuk memahami tingkat kerawanan *stunting*, dilakukan klasterisasi kecamatan di Kabupaten Tasikmalaya berdasarkan data jumlah balita dengan kategori pendek dan sangat pendek pada periode 2020 hingga 2022 menggunakan algoritma *K-Means++*. Penelitian ini menggunakan tahapan *Knowledge Discovery in Databases* (KDD). Hasil klasterisasi membagi kecamatan menjadi tiga kategori, yakni rendah, sedang, dan tinggi. Kecamatan Cigalontang secara konsisten berada dalam kategori tinggi selama tiga tahun berturut-turut, sedangkan kecamatan Salopa, Jatiwaras, dan Cisayong berpindah ke kategori sedang pada 2022. Sebagian besar wilayah tetap dalam kategori sedang, menunjukkan bahwa kerawanan *stunting* secara umum masih cukup tinggi. Evaluasi menggunakan *Davies-Bouldin Index* (DBI) menunjukkan nilai 0,7839 yang menandakan kualitas klasterisasi yang baik. Dari temuan ini dapat dijadikan dasar dalam penyusunan strategi penanggulangan *stunting* yang lebih terarah, dengan Kecamatan Cigalontang sebagai prioritas utama, serta kecamatan lain dalam kategori sedang yang juga memerlukan perhatian lebih lanjut.

Kata kunci : *Stunting, K-Means++*, *Clustering*, *Davies-Bouldin Index*, Kabupaten Tasikmalaya.

1. PENDAHULUAN

Stunting (kerdil) adalah isu kesehatan serius yang hingga kini menjadi tantangan dalam meningkatkan kualitas sumber daya manusia di Indonesia. *Stunting* pada balita merujuk pada kondisi dimana pertumbuhan tinggi badan anak lebih rendah dibandingkan dengan standar usia yang seharusnya, yang disebabkan oleh kekurangan gizi dalam 1000 Hari Pertama Kehidupan (HPK) [1].

Dampak dari kondisi ini tidak hanya terlihat pada pertumbuhan secara fisik, tetapi juga dapat mempengaruhi pola perkembangan kognitif dan kemampuan produktivitas anak di masa yang akan datang.

Faktor penyebab *stunting* dapat bersifat multidimensi, mulai dari asupan gizi tidak seimbang, keterbatasan akses layanan kesehatan, pola asuh yang kurang optimal, hingga faktor sosial-ekonomi [2].

Dampak kumulatif dari berbagai faktor risiko *stunting* ini dapat diukur melalui parameter antropometri, khususnya pengukuran Tinggi Badan menurut Umur (TB/U) pada balita. Mengacu pada pedoman Organisasi Kesehatan Dunia (WHO) yang diadopsi dalam Peraturan Menteri Kesehatan Republik Indonesia No. 2 Tahun 2020, status gizi anak dapat diklasifikasikan berdasarkan standar deviasi (SD) dari median standar pertumbuhan. Anak dapat dikategorikan sebagai pendek (*stunted*), jika tinggi badan berada diantara -2 hingga -3 SD, sedangkan anak dikategorikan sangat pendek (*severely stunted*), jika tinggi badan berada di bawah -3 SD [3].

Kedua kategori tersebut dapat digunakan sebagai dasar dalam mengidentifikasi kelompok-kelompok balita dengan risiko *stunting* di suatu wilayah.

Kabupaten Tasikmalaya merupakan salah satu wilayah yang perlu menjadi perhatian dalam upaya penanganan *stunting* di Provinsi Jawa Barat. Pada tahun 2022, prevalensi *stunting* di Kabupaten Tasikmalaya tercatat sebesar 27,2%. Upaya pemerintah daerah setempat untuk menurunkan angka tersebut telah menunjukkan hasil positif, dengan penurunan angka *stunting* menjadi 20,7% pada tahun 2023 [4].

Meskipun demikian, angka tersebut masih belum memenuhi target nasional yang ditetapkan dalam Rencana Pembangunan Jangka Menengah Nasional (RPJMN) 2020-2024, sebesar 14% pada tahun 2024 [5]. Oleh karena itu, upaya penurunan *stunting* di Kabupaten Tasikmalaya perlu dipercepat agar dapat mencapai target yang diharapkan.

Untuk mendukung percepatan penurunan *stunting*, diperlukan pendekatan yang terfokus pada wilayah dengan risiko sedang hingga tinggi. *Clustering* dapat dimanfaatkan untuk mengelompokkan kecamatan berdasarkan tingkat kerawanan. Dengan demikian, penentuan wilayah prioritas intervensi dapat dilakukan secara lebih efisien.

K-Means merupakan algoritma klasterisasi yang umum digunakan untuk mengelompokkan data berdasarkan jarak terdekatnya dengan titik pusat klaster (*centroid*), yang akan diperbarui secara terus-menerus hingga tidak berubah [6].

Namun, kelemahan utama algoritma ini terletak pada inisialisasi *centroid* secara acak, yang dapat menyebabkan konvergensi yang tidak diinginkan sehingga dapat menghasilkan klaster yang kurang optimal [7].

Untuk mengatasi hal ini, *K-Means++* dikembangkan untuk memperbaiki pemilihan *centroid*. Beberapa penelitian menerapkan kedua algoritma tersebut, dan hasilnya menunjukkan bahwa *K-Means++* menghasilkan klaster yang lebih baik, dengan nilai *Davies-Bouldin Index* yang lebih rendah dan *Silhouette Score* yang lebih tinggi dibandingkan dengan *K-Means* [8] [9] [10].

Berdasarkan hasil temuan tersebut, penelitian ini bertujuan untuk mengelompokkan kecamatan di Kabupaten Tasikmalaya berdasarkan data jumlah balita pendek dan sangat pendek menggunakan algoritma *K-Means++*. Dengan hal ini, diharapkan dapat menghasilkan klaster yang lebih optimal, sehingga mendukung pemerintah daerah dalam menentukan wilayah prioritas penurunan *stunting*.

2. TINJAUAN PUSTAKA

2.1. Penelitian Terdahulu

Beberapa penelitian sebelumnya telah menyoroiti kondisi kerawanan *stunting* di Jawa Barat, khususnya Kabupaten Tasikmalaya. Salah satu penelitian menggambarkan secara spasial hubungan antara jumlah keluarga berisiko *stunting* dan prevalensi *stunting* di Provinsi Jawa Barat tahun 2022. Hasilnya menunjukkan bahwa Kabupaten Tasikmalaya termasuk dalam *cluster* signifikan bertipe *High-High* (HH), yaitu wilayah dengan tingkat yang tinggi baik dalam jumlah keluarga berisiko *stunting* maupun prevalensi *stunting* [11].

Selain itu, penelitian lain yang melakukan *clustering* daerah rawan *stunting* di Jawa Barat juga mengidentifikasi bahwa Kabupaten Tasikmalaya tergolong ke dalam kelompok daerah dengan tingkat kerawanan *stunting* yang tinggi [12].

Temuan ini mengindikasikan bahwa Kabupaten Tasikmalaya merupakan daerah yang perlu menjadi prioritas dalam upaya percepatan penurunan *stunting*.

2.2. Clustering

Clustering merupakan salah satu teknik dalam *data science* yang digunakan dalam mengelompokkan data menjadi beberapa grup berdasarkan kemiripan karakteristik. Metode ini termasuk dalam kategori *unsupervised learning* karena mampu mengidentifikasi struktur atau pola tersembunyi dalam data tanpa label. Tujuan utamanya adalah untuk memastikan bahwa data dalam satu kelompok memiliki tingkat kesamaan yang tinggi, sementara data antar kelompok lainnya memiliki tingkat perbedaan yang signifikan [13].

2.3. Algoritma *K-Means++*

K-Means++ adalah variasi dari algoritma *K-Means* yang memperbaiki proses pemilihan *centroid*. Titik pusat pertama dipilih secara acak, sedangkan pusat berikutnya dipilih dengan probabilitas yang dipengaruhi oleh jarak kuadrat antara titik data dan pusat terdekat yang sudah dipilih. Semakin jauh jarak titik data dari pusat yang sudah ada, semakin besar

peluang titik tersebut dipilih sebagai pusat berikutnya. Pendekatan ini membantu distribusi pusat klaster yang lebih merata, sehingga mengurangi risiko hasil klasterisasi yang buruk serta meningkatkan hasil akurasi dan efisiensi waktu dibandingkan *K-Means* standar. Berdasarkan hal ini, algoritma *K-Means++* bekerja sebagai berikut:

- Menentukan jumlah nilai k (jumlah optimal klaster).
- Menentukan titik pusat klaster awal menggunakan rumus:

$$K = \frac{D(x)^2}{\sum_{x \in X} D(x)^2} \quad (1)$$

Keterangan:

$D(x)^2$: Jarak Euclidean dari titik data x ke *centroid* terdekat.

$\sum_{x \in X} D(x)^2$: Total jarak Euclidean dari seluruh titik data.

- Menghitung jarak Euclidean setiap objek data dengan menggunakan rumus:

$$d(x, y) = \sqrt{\sum_{i=0}^n (x_i - y_i)^2} \quad (2)$$

Keterangan:

$d(x, y)$: Jarak antara data x dan *centroid* y .

x_i : Nilai data x pada observasi ke- i .

y_i : Titik pusat y pada observasi ke- i .

n : Jumlah observasi.

$$V_{ij} = \frac{1}{N_i \sum_{k=0}^{N_i} x_{kj}^2} \quad (3)$$

Keterangan:

V_{ij} : Nilai rata-rata *centroid* klaster ke- i untuk variabel ke- j .

N_i : Jumlah anggota dalam klaster ke- i .

i, k : Indeks dari klaster.

j : Indeks variabel.

x_{kj} : Nilai data ke- k dari variabel ke- j pada klaster tersebut.

- Mengelompokkan data berdasarkan jarak minimum ke *centroid*.
- Memperbarui titik *centroid* dengan menghitung rata-rata setiap klaster.
- Mengulangi langkah (c), (d), dan (e) hingga tidak ada objek yang berpindah [14].

2.4. Knowledge Discovery in Database

Knowledge Discovery in Databases (KDD) adalah serangkaian tahapan yang bertujuan untuk menemukan pola dan informasi yang berguna dalam data besar. KDD melibatkan langkah-langkah seperti *Selection*, *Preprocessing*, *Transformation*, *Data Mining*, dan *Interpretation/Evaluation* [15].

2.5. Robust Scaler

Robust Scaler merupakan metode penskalaan data yang digunakan untuk meminimalkan dampak penciran (*outlier*) pada data. Teknik ini akan

menggunakan median sebagai titik pusat data dan rentang interkuartil (IQR) untuk mengatur skala data, sehingga pencilan yang dapat memengaruhi hasil skala dapat dikurangi. Secara matematis, dapat dijelaskan sebagai berikut:

$$x'_i = \frac{x_i - Q_2(x)}{Q_3(x) - Q_1(x)} \quad (4)$$

Keterangan:

x'_i : Nilai setelah diskalakan.
 x_i : Nilai pada data ke- i .
 $Q_2(x)$: Median dari data.
 $Q_1(x)$ dan $Q_3(x)$: Kuartil pertama dan ketiga [16].

2.6. Elbow Method

Metode *Elbow* adalah teknik yang paling sering digunakan dalam menentukan jumlah kluster yang optimal. Titik optimal kluster ditentukan ketika grafik membentuk sudut siku yang tajam atau mengalami penurunan paling signifikan [17].

Hal ini menunjukkan bahwa penambahan jumlah kluster setelah titik tersebut tidak memberikan peningkatan yang signifikan terhadap kualitas kluster.

2.7. Sum of Square Error

Sum of Square Error (SSE) adalah metode statistik yang digunakan untuk mengukur total selisih antara nilai aktual dan nilai yang diperoleh dari hasil model [18]. Metode ini juga sering digunakan dalam penentuan jumlah kluster yang optimal dengan cara melihat penurunan SSE yang paling signifikan pada berbagai nilai k yang diuji [19].

$$SSE = \sum_{i=1}^n (x_i - y_i)^2 \quad (3)$$

Keterangan:

x : Nilai aktual atau sebenarnya.
 y : Nilai yang tercapai atau yang diperoleh [18].

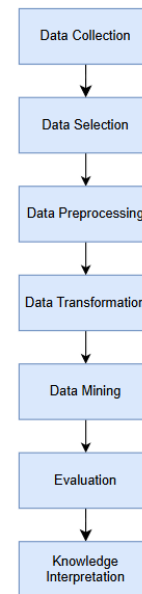
2.8. Davies-Bouldin Index

Davies-Bouldin Index (DBI) adalah metrik evaluasi klusterisasi yang mengukur rasio antara jarak intra-kluster (dalam kluster) dan inter-kluster (antar kluster). Nilai DBI yang lebih rendah menunjukkan bahwa kluster memiliki kohesi internal yang tinggi dan terpisah dengan baik dari kluster lainnya. Dengan demikian, performa klusterisasi akan semakin optimal saat nilai DBI mendekati nol [20] [21].

3. METODE PENELITIAN

3.1. Tahap Perancangan

Proses perancangan data yang digunakan mengacu pada metode *Knowledge Discovery in Database* (KDD), yang diterapkan untuk mengeksplorasi dan menganalisis data secara sistematis guna memperoleh pola atau informasi yang bermakna sebagai dasar pengambilan keputusan.



Gambar 1. Diagram Alir

3.2. Data Collection

Data diperoleh melalui berbagai sumber yaitu dari Open Data Kabupaten Tasikmalaya terkait data kategori balita berisiko *stunting* di setiap kecamatan di Kabupaten Tasikmalaya dari tahun 2020 hingga 2022. Selain itu, data spasial dalam format *GeoJSON* juga digunakan untuk memetakan batas administratif wilayah kecamatan di Kabupaten Tasikmalaya.

3.3. Data Selection

Pada tahap ini, data yang diperoleh dilakukan pemilihan data sesuai dengan tujuan analisis. Proses ini bertujuan untuk mengidentifikasi dan menyaring data yang penting untuk dianalisis secara mendalam.

3.4. Data Preprocessing

Tahap ini meliputi proses pemeriksaan dan penanganan terhadap data yang bermasalah, seperti nilai yang hilang (*missing values*), data duplikat, nilai negatif, distribusi yang tidak seimbang, serta nilai ekstrim (*outlier*) dengan tujuan untuk memastikan kualitas dari data sebelum dilakukan analisis lebih lanjut.

3.5. Data Transformation

Transformasi dilakukan untuk menyesuaikan format data dengan kebutuhan analisis. Pada tahap ini, dilakukan standarisasi data agar setiap atribut memiliki skala yang seragam dan dapat diproses secara optimal dalam algoritma *K-Means++*.

3.6. Data Mining

Tahap utama dalam proses KDD adalah penerapan algoritma *K-Means++ clustering* untuk mengelompokkan kecamatan berdasarkan tingkat kerawanan *stunting*. Hasil dari pengelompokan ini

akan menunjukkan pola distribusi dari wilayah kecamatan berdasarkan variasi tingkat risikonya.

3.7. Evaluation

Evaluasi dilakukan menggunakan *Davies-Bouldin Index* (DBI) untuk mengukur kualitas hasil kluster. Nilai DBI yang lebih rendah dan mendekati nol mengindikasikan kluster yang lebih baik.

3.8. Knowledge Interpretation

Hasil klusterisasi akan divisualisasikan dalam bentuk peta tematik yang menunjukkan pola distribusi kerawanan *stunting* antar kecamatan. Visualisasi ini dapat digunakan dalam identifikasi wilayah prioritas penanganan, sehingga dapat mendukung perumusan kebijakan penanggulangan *stunting* yang lebih tepat sasaran.

4. HASIL DAN PEMBAHASAN

4.1. Data Collection

Data yang dikumpulkan terdiri dari 12 atribut, termasuk kode dan nama provinsi, kabupaten, kecamatan, jumlah balita berisiko *stunting* dalam kategori pendek dan sangat pendek, satuan, dan tahun pencatatan dari 2020 hingga 2022. Selain itu, data spasial juga digunakan untuk memetakan batas administratif kecamatan di Kabupaten Tasikmalaya.

4.2. Data Selection

Selanjutnya dilakukan seleksi atribut sesuai dengan tujuan analisis dan didapatkan empat atribut utama, yaitu tahun, nama kecamatan, jumlah balita dengan kategori pendek, dan jumlah balita dengan kategori sangat pendek. Berikut merupakan tabel kasus *stunting* yang ada di Kabupaten Tasikmalaya setelah dilakukan seleksi atribut.

Tabel 1. Dataset Rawan *Stunting* Setelah Seleksi

Tahun	Nama Kecamatan	Kategori Sangat Pendek	Kategori Pendek
2020	Cipatujah	94	456
2020	Karangnunggal	65	369
...	...	...	...
2022	Pagerageung	84	321
2022	Sukaresik	78	289

4.3. Data Preprocessing

Untuk memastikan kualitas data, tahap pembersihan dilakukan dengan mengecek *missing value*, duplikasi, anomali seperti data negatif, keseimbangan data, dan *outlier*. Hasilnya, tidak ditemukan *missing value* atau duplikasi, serta tidak ada nilai negatif pada data numerik. Distribusi data per tahun juga menunjukkan keseimbangan. Meskipun terdeteksi beberapa *outlier*, yang menunjukkan beberapa nilai ekstrim pada beberapa atribut. Nilai-nilai ini tetap dipertahankan untuk menjaga keutuhan karakteristik data yang merepresentasikan kondisi sesungguhnya di wilayah penelitian.

4.4. Data Transformation

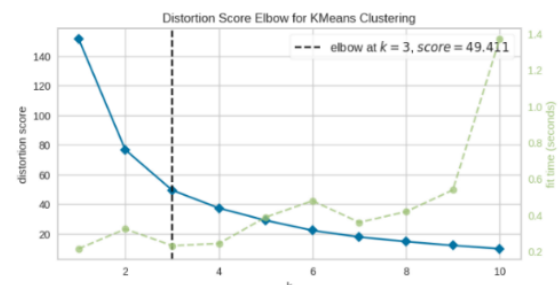
Setelah tahap pembersihan, data diproses standarisasi untuk menyesuaikan skala antar atribut. Standarisasi ini dilakukan agar dapat mengatasi perbedaan skala antar fitur, sehingga tidak mempengaruhi kinerja model yang digunakan. Pada tahap ini, diterapkan metode *Robust Scaler*, yang digunakan untuk mengurangi pengaruh nilai ekstrim. Dengan cara ini, setiap fitur dalam dataset dapat memberikan kontribusi yang seimbang pada proses analisis berikutnya.

Tabel 2. Dataset Akhir Sebelum Masuk Tahap Modeling

Tahun	Nama Kecamatan	Kategori Sangat Pendek	Kategori Pendek
2020	Cipatujah	0.444444	0.930851
2020	Karangnunggal	-0.015873	0.468085
...	...	...	...
2022	Pagerageung	0.285714	0.212766
2022	Sukaresik	0.190476	0.042553

4.5. Data Mining

Pada tahap utama ini, dilakukan klusterisasi menggunakan algoritma *K-Means++*. Penentuan jumlah kluster dilakukan dengan pendekatan *elbow method*, di mana nilai k diuji dari 1 hingga 10. Proses klusterisasi dijalankan sebanyak 100 kali untuk memperoleh hasil yang lebih stabil.



Gambar 2. Hasil k Optimal dengan *Elbow Visualizer*

Berdasarkan Gambar 2, terjadi penurunan nilai yang mulai melambat pada $k = 3$, yang menunjukkan titik *elbow* sebagai kandidat jumlah kluster optimal. Untuk memastikan hasil tersebut, dilakukan perhitungan nilai *Sum of Square Errors* (SSE), sebagaimana ditampilkan pada Tabel 3.

Tabel 3. Nilai *Sum of Square Errors* (SSE)

k	Nilai SSE	Selisih
2	76.64	-
3	49.41	27.23
4	37.31	12.10
5	29.02	8.28
6	22.24	6.79
7	17.80	4.44
8	14.72	3.08
9	12.10	2.62
10	9.90	2.20

Meskipun pengujian dilakukan mulai dari $k = 1$, nilai tersebut tidak dicantumkan karena tidak membentuk kluster (seluruh data berada dalam satu kelompok). Dari tabel tersebut disajikan nilai selisih SSE antar k , yang dihitung berdasarkan pengurangan nilai SSE pada k sebelumnya terhadap k saat ini. Sebagai contoh, nilai SSE pada $k = 2$ adalah 76,64 dan pada $k = 3$ adalah 49,41, sehingga selisihnya sebesar 27,23.

Dari Tabel 3, penurunan SSE paling signifikan terjadi dari $k = 2$ ke $k = 3$, dengan selisih sebesar 27,23. Setelah $k = 3$, nilai SSE terus menurun tetapi dengan selisih yang semakin kecil dan cenderung melandai. Temuan ini memperkuat hasil metode *elbow*, yang menunjukkan bahwa $k = 3$ merupakan jumlah kluster optimal yang digunakan dalam proses klusterisasi ini.

4.6. Evaluation

Evaluasi kualitas hasil klusterisasi menggunakan metrik *Davies-Bouldin Index* (DBI) dengan nilai yang mendekati nol dan tidak negatif menunjukkan bahwa kluster yang terbentuk memiliki kesamaan yang tinggi di dalam kluster dan pemisahan yang jelas antar kluster.

```
# DBI
davies_bouldin = davies_bouldin_score(X, labels)
print(f"Davies-Bouldin Index: {davies_bouldin}")

Davies-Bouldin Index: 0.7839088751920326
```

Gambar 3. Nilai Evaluasi dengan DBI

Nilai *Davies-Bouldin Index* (DBI) yang diperoleh adalah 0,7839 yang menginformasikan bahwa kluster yang terbentuk memiliki kualitas yang baik.

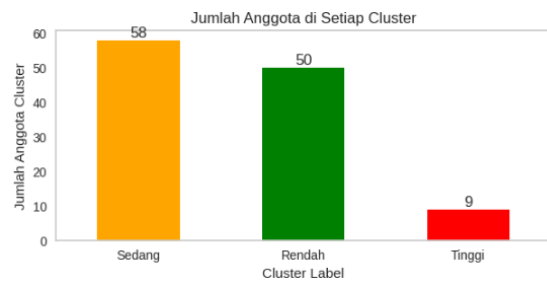
4.7. Knowledge Interpretation

Berdasarkan hasil klusterisasi, daerah rawan *stunting* di Kabupaten Tasikmalaya antara tahun 2020 hingga 2022 terbagi menjadi tiga kluster, yaitu rendah, sedang, dan tinggi. Informasi lebih lanjut mengenai pelabelan kluster dapat dilihat pada Tabel 4.

Tabel 4. Pelabelan Cluster

klaster	Kategori pendek			kategori sangat pendek			Label
	min	max	mean	min	max	mean	
0	228	511	354.41	3	143	87.78	Sedang
1	26	288	160.86	1	80	38.98	Rendah
2	349	895	565.56	144	340	197.11	Tinggi

Penentuan label kluster dilakukan berdasarkan rata-rata nilai dari dua atribut, yaitu kategori pendek dan sangat pendek, dengan nilai minimum dan maksimum yang menggambarkan variasi tingkat kerawanan di setiap kluster.



Gambar 4. Visualisasi hasil clustering

Pada Gambar 4 diatas, menunjukkan distribusi jumlah anggota pada setiap kluster kerawanan *stunting* di Kabupaten Tasikmalaya. Kluster dengan kerawanan *stunting* sedang mencakup 58 kecamatan, sementara kluster dengan kerawanan *stunting* rendah terdiri dari 50 kecamatan. Kluster dengan kerawanan *stunting* tinggi mencakup 9 kecamatan.

Tabel 5. Hasil Cluster Tahun 2020

Label	Nama Kecamatan
Rendah	Cikalong, Cikatomas, Cibalong, Parungponteng, Bantarkalong, Bojongasih, Culamega, Taraju, Tanjungjaya, Cineam, Karangjaya, Gunungtanjung, Leuwisari, Kadipaten
Sedang	Cipatujah, Karangnunggal, Pancatengah, Bojunggambir, Sodonghilir, Salawu, Puspahiang, Sukaraja, Manonjaya, Singaparna, Mangunreja, Sukarame, Padakembang, Sariwangi, Sukaratu, Sukahening, Rajapolah, Jamanis, Ciawi, Pagerageung, Sukaresik
Tinggi	Salopa, Jatiwaras, Cigalontang, Cisayong

Tabel 5 menunjukkan hasil klusterisasi tingkat kerawanan *stunting* di Kabupaten Tasikmalaya pada tahun 2020. Sebanyak 14 kecamatan tergolong tingkat kerawanan rendah, 21 kecamatan tergolong sedang, dan 4 kecamatan tergolong tinggi.

Tabel 6. Hasil Cluster Tahun 2021

Label	Nama Kecamatan
Rendah	Cipatujah, Cikalong, Cikatomas, Cibalong, Parungponteng, Bantarkalong, Bojongasih, Culamega, Taraju, Salawu, Puspahiang, Tanjungjaya, Cineam, Karangjaya, Gunungtanjung, Leuwisari, Sariwangi, Sukahening
Sedang	Karangnunggal, Pancatengah, Bojunggambir, Sodonghilir, Sukaraja, Manonjaya, Singaparna, Mangunreja, Sukarame, Padakembang, Sukaratu, Rajapolah, Jamanis, Ciawi, Kadipaten, Pagerageung, Sukaresik
Tinggi	Salopa, Jatiwaras, Cigalontang, Cisayong

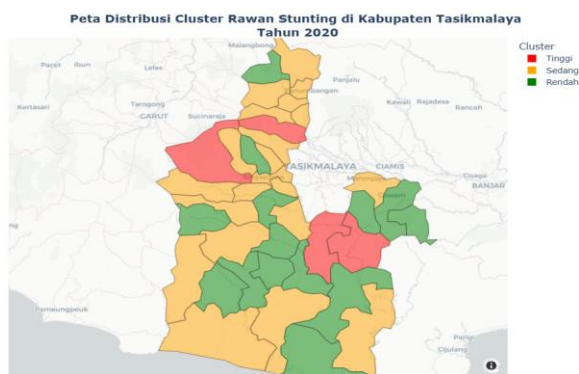
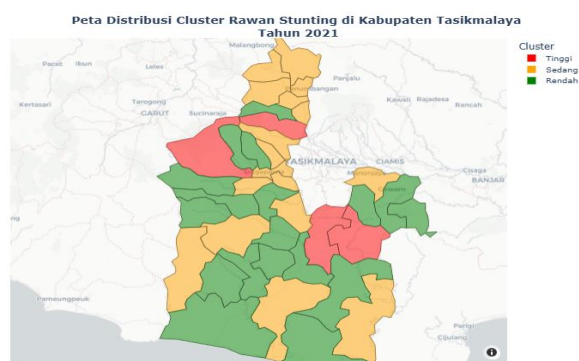
Tabel 6 menunjukkan hasil klusterisasi tingkat kerawanan *stunting* di Kabupaten Tasikmalaya pada tahun 2021. Sebanyak 18 kecamatan tergolong tingkat kerawanan rendah, 17 kecamatan tergolong sedang, dan 4 kecamatan tergolong tinggi.

Tabel 7. Hasil *Cluster* Tahun 2022

Label	Nama Kecamatan
Rendah	Cipatuh, Cikalong, Cikatomas, Cibalong, Parungponteng, Bantarkalong, Bojongasih, Culamega, Taraju, Puspahiang, Tanjungjaya, Cineam, Karangjaya, Gunungtanjung, Leuwisari, Padakembang, Sariwangi, Kadipaten
Sedang	Karangnunggal, Pancatengah, Salawu, Bojonggambir, Sodonghilir, Sukaraja, Salopa, Jatiwaras, Manonjaya, Singaparna, Mangunreja, Sukarame, Sukaratu, Cisayong, Sukahening, Rajapolah, Jamanis, Ciawi, Pagerageung, Sukaresik
Tinggi	Cigalontang

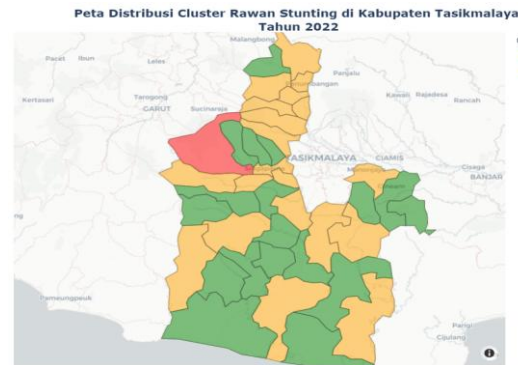
Tabel 7 menunjukkan hasil klasterisasi tingkat kerawanan *stunting* di Kabupaten Tasikmalaya pada tahun 2022. Sebanyak 18 kecamatan tergolong tingkat kerawanan rendah, 20 kecamatan tergolong sedang, dan 1 kecamatan tergolong tinggi.

Untuk mempermudah pemahaman mengenai distribusi geografis dari masing-masing klaster, disajikan visualisasi data yang menggambarkan sebaran klaster rawan *stunting* pada tahun 2020 hingga 2022.

Gambar 5. Distribusi Daerah Rawan *Stunting* di Kabupaten Tasikmalaya Tahun 2020Gambar 6. Distribusi Daerah Rawan *Stunting* di Kabupaten Tasikmalaya Tahun 2021

Gambar 5, 6, 7, menunjukkan pemetaan kecamatan rawan *stunting* di Kabupaten Tasikmalaya yang terbagi dalam tiga kategori warna. Warna merah mewakili kecamatan dengan tingkat kerawanan *stunting* yang tinggi, warna oranye menunjukkan

daerah dengan kerawanan *stunting* yang sedang, dan warna hijau menandakan daerah dengan tingkat kerawanan *stunting* yang rendah. Dari hasil visualisasi ini, diharapkan dapat membantu dalam memetakan wilayah-wilayah yang perlu mendapatkan perhatian prioritas dalam upaya penanggulangan *stunting*.

Gambar 7. Distribusi Daerah Rawan *Stunting* di Kabupaten Tasikmalaya Tahun 2022

Dari hasil tersebut, ditemukan bahwa selama tahun 2020 dan 2021 terdapat empat kecamatan yang secara konsisten berada pada tingkat kerawanan *stunting* tinggi, yaitu Salopa, Jatiwaras, Cigalontang, dan Cisayong. Namun, pada tahun 2022, hanya Cigalontang yang masih menempati kategori tersebut, sedangkan tiga lainnya mengalami perbaikan dan berpindah ke kategori sedang. Hal ini mencerminkan adanya peningkatan kondisi di beberapa wilayah. Meskipun demikian, masih banyak kecamatan yang tetap berada dalam kategori sedang, yang menunjukkan bahwa tantangan *stunting* masih cukup besar di Kabupaten Tasikmalaya. Kondisi kecamatan Cigalontang yang tidak mengalami perubahan selama tiga tahun berturut-turut menekankan pentingnya tindakan prioritas di wilayah tersebut.

5. KESIMPULAN DAN SARAN

Berdasarkan hasil klasterisasi menggunakan algoritma *K-Means++* terhadap data *stunting* di Kabupaten Tasikmalaya tahun 2020 hingga 2022, diperoleh tiga klaster yaitu rendah, sedang, dan tinggi. Kecamatan Cigalontang secara konsisten berada dalam klaster tinggi selama tiga tahun berturut-turut, sementara kecamatan Salopa, Jatiwaras, dan Cisayong menunjukkan perbaikan dengan berpindah ke klaster sedang pada tahun 2022. Meskipun demikian, dominasi wilayah masih berada dalam kategori sedang, menunjukkan bahwa tantangan *stunting* secara umum masih cukup besar. Nilai *Davies-Bouldin Index* (DBI) yang diperoleh sebesar 0,7839 menunjukkan bahwa kualitas klasterisasi yang baik. Temuan ini dapat dijadikan dasar dalam merancang strategi penanggulangan *stunting* yang lebih terarah, khususnya di wilayah dengan tingkat kerawanan yang tinggi.

Penelitian selanjutnya disarankan untuk memperluas variabel yang digunakan dengan

memasukkan aspek-aspek seperti kondisi ekonomi, tingkat pendidikan orang tua, kualitas sanitasi lingkungan, serta ketersediaan dan aksesibilitas pelayanan kesehatan. Hal ini bertujuan untuk memperoleh pemetaan yang lebih menyeluruh mengenai kondisi yang berkontribusi terhadap *stunting*. Selain itu, penelitian selanjutnya juga dapat mengeksplorasi dengan membandingkan algoritma klusterisasi lainnya guna memperoleh hasil pengelompokan yang lebih optimal.

DAFTAR PUSTAKA

- [1] S. R. Bincar Robinson Hutasuhut, "Mencegah Stunting pada 1000 Hari Pertama Kehidupan pada Masyarakat Kelurahan Pasar Merah Barat," *J. Implementa Husada*, vol. 3, no. 2, 2022, doi: 10.30596/jih.v3i2.11591.
- [2] N. W. E. P. Dewi dan N. K. S. Ariani, "Hubungan Pengetahuan Ibu Tentang Gizi Menurunkan Resiko Stunting Pada Balita di Kabupaten Gianyar," *J. Menara Med.*, vol. 3, no. 2, 2021.
- [3] Kemenkes, "Peraturan Menteri Kesehatan Republik Indonesia Nomor 2 Tahun 2020 Tentang Standar Antropometri Anak," 2020. [Daring]. Tersedia pada: <https://peraturan.bpk.go.id/Details/152505/perm-enkes-no-2-tahun-2020>
- [4] Qjabar.com, "Publikasi Data Stunting Kabupaten Tasikmalaya," Qjabar.com. Diakses: 15 April 2025. [Daring]. Tersedia pada: <https://qjabar.com/2024/12/24/publikasi-data-stunting-kabupaten-tasikmalaya/>
- [5] Kemendagri, "Kemendagri Mendukung dalam Upaya Pencapaian Target Penurunan Stunting 14% Tahun 2024," kemendagri.go.id. Diakses: 15 April 2025. [Daring]. Tersedia pada: https://www.bangda.kemendagri.go.id/berita/baca_kontent/541%0A
- [6] N. K. Zuhail, "Study Comparison K-Means Clustering Dengan Algoritma Hierarchical Clustering," in *Prosiding Seminar Nasional Teknologi dan Sains*, 2022.
- [7] M. Ahmed, R. Seraj, dan S. M. S. Islam, "The k-means algorithm: A comprehensive survey and performance evaluation," 2020. doi: 10.3390/electronics9081295.
- [8] Mohammad Ferdiansyah dan Umi Chotijah, "Implementasi Algoritma K-Means++ Untuk Clustering Penjualan Bahan Bangunan," *J. Ilm. Tek. Inform. dan Komun.*, vol. 4, no. 1, 2024, doi: 10.55606/juitik.v4i1.767.
- [9] C. A. Sri Fastaf dan Y. Yamasari, "Analisa Pemetaan Kriminalitas Kabupaten Bangkalan Menggunakan Metode K-Means dan K-Means++," *J. Informatics Comput. Sci.*, vol. 3, no. 04, 2022, doi: 10.26740/jinacs.v3n04.p534-546.
- [10] N. Nugroho dan F. D. Adhinata, "Penggunaan Metode K-Means dan K-Means++ Sebagai Clustering Data Covid-19 di Pulau Jawa," *Teknika*, vol. 11, no. 3, 2022, doi: 10.34148/teknika.v11i3.502.
- [11] M. P. Kenanga, S. D. Rahmawati, T. Surtimanah, Y. Mahwati, dan S. Komalaningsih, "Analisis Spasial Keluarga Berisiko Stunting dengan Kejadian Stunting di Provinsi Jawa Barat Tahun 2022," 2023.
- [12] Taufik Hidayat, Mohamad Jajuli, dan Susilawati, "Clustering daerah rawan stunting di Jawa Barat menggunakan algoritma K-Means," *INFOTECH J. Inform. Teknol.*, vol. 4, no. 2, 2023, doi: 10.37373/infotech.v4i2.642.
- [13] K. P. Sinaga dan M. S. Yang, "Unsupervised K-means clustering algorithm," *IEEE Access*, vol. 8, 2020, doi: 10.1109/ACCESS.2020.2988796.
- [14] I. M. Nur dan Abdurakhman, "K-Means ++ Algorithm for Health Services Clustering Based on Districts in West Java Province," *EKSAKTA J. Sci. Data Anal.*, vol. 5, no. 1, hal. 96–102, 2024, doi: 10.20885/eksakta.vol5.iss1.art11.
- [15] Haris Kurniawan, Sarjon Defit, dan Sumijan, "Data Mining Menggunakan Metode K-Means Clustering Untuk Menentukan Besaran Uang Kuliah Tunggal," *J. Appl. Comput. Sci. Technol.*, vol. 1, no. 2, 2020, doi: 10.52158/jacost.v1i2.102.
- [16] L. B. V. de Amorim, G. D. C. Cavalcanti, dan R. M. O. Cruz, "The choice of scaling technique matters for classification performance," *Appl. Soft Comput.*, vol. 133, 2023, doi: 10.1016/j.asoc.2022.109924.
- [17] F. Fauji dan L. Farokhah, "Penerapan Data Mining Menggunakan K-Means Clustering Dalam Mengelompokkan Tingkat Kesulitan Mata Pelajaran," *J. Inf. Syst. Res.*, vol. 6, no. 3, hal. 1705–1714, 2025, doi: 10.47065/josh.v6i3.6959.
- [18] D. B. A. Pramana, T. Amalina, dan R. I. Adam, "Analisis K-Means Clustering Pada Pengiriman Produk Bearing," *J. Ilm. Wahana Pendidik.*, vol. 8, no. 15, 2022.
- [19] M. N. N. Eric Kamti, Ade Putra Sanjaya, "Clustering Penjualan Terbaik Dengan Sum of Squares Error Dan Menentukan Nilai K Menggunakan Algoritma K-Means," vol. 23, hal. 387–393, 2024.
- [20] Y. A. Wijaya, D. A. Kurniady, E. Setyanto, W. S. Tarihoran, D. Rusmana, dan R. Rahim, "Davies Bouldin Index Algorithm for Optimizing Clustering Case Studies Mapping School Facilities," *TEM J.*, vol. 10, no. 3, 2021, doi: 10.18421/TEM103-13.
- [21] F. Rizki Ani, R. Kurniawan, dan T. Suprpti, "PENERAPAN ALGORITMA K-MEANS CLUSTERING UNTUK PERENCANAAN PERSEDIAAN STOK SEPATU DI TOKO DIOMCLOTHING," *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 7, no. 6, 2024, doi: 10.36040/jati.v7i6.8233.