

KLASTERISASI DATA BALITA STUNTING DI KECAMATAN WILAYAH KABUPATEN CIAMIS BERDASARKAN PREVALENSI DENGAN MENGUNAKAN ALGORITMA K-MEANS

Rizky Rahman Sahputra ¹, Rudi Kurniawan ²

¹ Manajemen Informatika, STMIK IKMI Cirebon

² Rekayasa Perangkat Lunak, STMIK IKMI Cirebon

JL. Perjuangan No. 10B Karyamulya Kec. Kesambi Kota Cirebon

rizkyrahmans012@gmail.com rudi226ikmi@gmail.com

ABSTRAK

Stunting adalah gangguan pertumbuhan dan perkembangan yang terjadi pada anak di bawah usia 5 tahun, ditandai dengan tinggi badan yang tidak sesuai dengan usianya. Stunting masih menjadi masalah gizi utama di Indonesia, dengan angka stunting nasional sebesar 24,4% pada tahun 2021. Berdasarkan data Pemerintahan Provinsi Jawa Barat, pada tahun 2022 Kabupaten Ciamis memiliki persentase balita stunting sebesar 2,74%. Faktor penyebab stunting diantaranya asupan gizi yang tidak mencukupi, infeksi, dan faktor sosial ekonomi. Saat ini di Kabupaten Ciamis, diperlukan klusterisasi kecamatan berdasarkan prevalensi stunting. Ini bertujuan untuk menentukan prioritas kecamatan yang perlu mendapatkan perhatian khusus dalam upaya pencegahan dan penanganan stunting. Metodologi yang digunakan pada penelitian ini yaitu knowledge discovery in database dengan algoritma K-Means. Data yang digunakan adalah data stunting dan prevalensi stunting dari 27 kecamatan di Kabupaten Ciamis, yang bersumber dari Dinas Kesehatan Kabupaten Ciamis. Tujuan penelitian ini adalah untuk mengklusterkan kecamatan di Kabupaten Ciamis berdasarkan prevalensi stunting, menjadi tiga kelompok. Klusterisasi kecamatan berdasarkan prevalensi stunting dapat membantu Pemerintah Daerah untuk menentukan kecamatan mana yang perlu ditangani secara lebih intensif. Hasil penelitian ini adalah klaster data stunting tingkat rendah dengan <10%, tingkat menengah dengan 10%-20% dan tingkat tinggi dengan >20%, yang disajikan dalam bentuk Visualisasi Data sehingga mudah dibaca.

Kata kunci : Algoritma K-Means, Data Mining, Prevalensi Stunting, Klusterisasi Kecamatan, Stunting

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 gagalnya pertumbuhan pada balita, yang ditandai dengan keadaan tubuh balita yang terlalu pendek dibanding balita seusianya. Stunting diakibatkan karena kekurangan gizi kronis terutama dalam periode 1000 hari pertama kehidupan, yaitu sejak bayi berada dalam kandungan sampai dengan usia 23 bulan atau 2 tahun [1]. Algoritma K-Means merupakan salah satu algoritma *clustering* yang banyak digunakan diberbagai bidang seperti bidang pendidikan, kesehatan, sosial, biologi, dan ilmu komputer [2]. Algoritma K-Means melakukan partisi data ke dalam satu atau lebih *cluster* yang memiliki ciri-ciri yang sama.

Algoritma K-Means mampu menghasilkan akurasi atau tingkat ketelitian yang maksimal dengan sifatnya yang linear, waktu eksekusi algoritma ini relatif cepat dan mudah untuk diadaptasikan [1]. Terbukti pada penelitian yang dilakukan oleh Widayani & Harliana yang membandingkan algoritma K-Means dengan *Subtractive Fuzzy C-Means* (SFCM) dalam mengelompokkan data rumah tangga miskin di Kabupaten Cirebon, hasil penelitiannya menunjukkan bahwa algoritma K-Means memberikan waktu eksekusi lebih cepat dibanding SFCM, dimana K-Means memerlukan waktu 9 detik, sedangkan SFCM memerlukan waktu 13.3 detik [3]. Penelitian yang dilakukan oleh Kamila, dkk. yang membandingkan algoritma K-Means dengan K-Medoids dalam mengelompokkan data transaksi bongkar muat di Provinsi Riau, hasil penelitiannya menunjukkan bahwa algoritma K-Means memberikan waktu eksekusi lebih cepat dibanding K-Medoids, dimana K-Means memerlukan waktu rata-rata 1 detik, sedangkan K-Medoids memerlukan waktu rata-rata 1 menit 38 detik [4]. Namun algoritma K-Means mempunyai beberapa kekurangan, salah satunya adalah sensitif dalam menentukan jumlah *cluster*. Winarta & Kurniawan melakukan optimasi *cluster* dengan bantuan *metode elbow* dalam mengelompokkan data pengguna narkoba, hasil penelitiannya menyatakan bahwa *metode elbow* bekerja sangat baik dalam membantu menghasilkan cluster yang optimal, dimana

cluster yang dihasilkan adalah sebanyak 3 cluster dengan nilai selisih SSE sebesar 1257.862 [5].

Berikut pada Tabel 1. diperlihatkan data stunting tahun 2019 sampai tahun 2022 di Kabupaten Ciamis sebagai berikut:

Tabel 1. Data angka Stunting

Kategori	2019	2020	2021	2022
Presentase balita stunting di Ciamis	6,47%	7,33%	11,24%	4,83%
Jumlah balita stunting di Ciamis	4847	4254	7688	3303
Jumlah berat bayi lahir rendah (BBLR) di Ciamis	896	797	863	794
Jumlah Balita kurus dengan indeks bb/tb di Ciamis	1613	2736	1977	2358
Jumlah balita gizi kurang berdasarkan bb/u di Ciamis	306	4810	3697	

Stunting merupakan salah satu permasalahan gizi yang menjadi perhatian dunia dan merupakan masalah gizi utama yang dihadapi Indonesia terutama di Kabupaten Ciamis. Pada tahun 2019, 21.9% balita di dunia mengalami stunting, sebesar 33.1% kasus balita stunting di dunia berasal dari Afrika, 31.9% berasal dari Asia Tenggara, 24.7% berasal dari Mediterania Timur, 6.5% berasal dari Amerika, dan kasus balita stunting terendah yaitu sebesar 6.4% berasal dari Pasifik Barat. Indonesia termasuk dalam urutan ke-34 dari 50 negara dengan kasus balita stunting tertinggi di dunia, dan termasuk dalam urutan ke-6 di Asia Tenggara. Hasil Studi Status Gizi Balita Indonesia (SSGBI) yang terintegrasi dengan Susenas Maret 2019 diperoleh anak balita yang mengalami stunting sebanyak 27,7%. Kemudian pada tahun 2021 diperoleh prevalensi stunting sebesar 24,4% yang menunjukkan bahwa pencapaian target 2024 sebanyak 14% masih sangat jauh untuk pencapaiannya. [6]. Kurangnya skala prioritas untuk pengelompokan daerah rawan stunting juga merupakan salah satu faktor permasalahan untuk penanganan kasus stunting terutama di Kabupaten Ciamis.

Maka berdasarkan uraian tersebut, dalam penelitian akan dilakukan pengelompokan Kecamatan di Kabupaten Ciamis berdasarkan faktor penyebab stunting pada balita menggunakan algoritma k-means. Penelitian ini diharapkan dapat membantu pemerintah dalam mengambil kebijakan yang sesuai terkait penurunan prevalensi stunting pada balita berdasarkan karakteristik dan permasalahan masing-masing cluster.

Dengan melakukan analisis data stunting anak menggunakan algoritma K-Means, maka dapat

mengkategorisasi daerah rawan stunting di Kabupaten Ciamis. Ini akan memberikan pandangan yang lebih mendalam tentang masalah stunting di Kabupaten Ciamis dan dapat membantu pemerintah dan lembaga kesehatan dalam merancang program-program intervensi yang lebih tepat sasaran dan efektif dengan memperhatikan skala prioritasnya.

Melalui penelitian ini, diharapkan dapat memberikan sumbangan penting dalam upaya mengatasi masalah stunting anak di Kabupaten Ciamis dan secara lebih luas di Indonesia. Analisis ini dapat menjadi dasar untuk merumuskan kebijakan kesehatan yang lebih efektif dan program-program pendidikan gizi yang sesuai dengan kebutuhan wilayah ini. Oleh karena itu diusulkan tugas akhir dengan judul “Klasterisasi Data Balita Stunting Di Kecamatan Wilayah Kabupaten Ciamis Berdasarkan Prevalensi Dengan Menggunakan Algoritma K-Means”. Adapun yang menjadi alasan dilakukannya tugas akhir ini dengan judul tersebut adalah untuk membantu pemerintah agar dapat lebih memperhatikan daerah rawan stunting di Kabupaten Ciamis berdasarkan skala prioritas yang telah di klasterisasi.

2. TINJAUAN PUSTAKA

2.1. Stunting

Stunting merupakan kondisi gagalnya pertumbuhan pada balita, yang ditandai dengan keadaan tubuh balita yang terlampau pendek dibanding balita seusianya. Stunting diakibatkan karena kekurangan gizi kronis terutama dalam periode 1000 hari pertama kehidupan, yaitu sejak bayi berada dalam kandungan sampai dengan usia 23 bulan atau 2 tahun. Balita yang menderitastuntingakan lebih mudah terkena penyakit, juga berpotensi menghambat pertumbuhan fisik serta perkembangan kognitif yang akan berpengaruh terhadap tingkat kecerdasan dan produktivitas anak di masa mendatang [1].

2.2. Faktor-Faktor yang Mempengaruhi Stunting

[7] Ada banyak faktor yang dapat mempengaruhi prevalensi stunting, antara lain:

- Akses dan kualitas pelayanan kesehatan ibu dan anak.
- Pengetahuan dan pemahaman masyarakat tentang stunting.
- Akses dan konsumsi makanan bergizi.
- Sanitasi dan kebersihan lingkungan.

2.3. Klasterisasi

Klasterisasi (*Clustering*) merupakan metode yang sangat populer dan paling banyak digunakan dalam pengolahan data karena kesederhanaan dan efektivitasnya dalam mengelompokkan data berukuran besar berdasarkan kecepatan pemrosesan dengan menempatkan objek ke dalam kelas yang mempunyai kesamaan [8].

Clustering merupakan sebuah metode *data mining* yang termasuk dalam kategori *unsupervised learning*. Proses pembelajarannya tidak

membutuhkan kelas. Metode K-Means banyak digunakan untuk menyelesaikan kasus klusterisasi [9].

2.4. Algoritma K-Means

Clustering dengan menggunakan algoritma K-Means merupakan metode data mining yang sangat populer dalam pengolahan data dan sangat efektif dalam mengelompokkan data yang besar ke dalam kelas-kelas yang serupa [10].

K-Means *Clustering* dimana K dimaksudkan sebagai konstanta jumlah cluster yang diinginkan, Means dalam hal ini berarti nilai suatu rata-rata dari suatu grup data yang dalam hal ini didefinisikan sebagai *cluster*, sehingga K-Means *Clustering* adalah suatu metode penganalisaan data atau metode *data mining* yang melakukan proses pemodelan tanpa supervisi (*unsupervised*) dan merupakan salah satu metode yang melakukan pengelompokan data dengan sistem partisi. Metode K-Means berusaha mengelompokkan data yang ada kedalam beberapa kelompok, dimana data dalam satu kelompok mempunyai karakteristik yang sama satu sama lainnya dan mempunyai karakteristik yang berbeda dengan data yang ada didalam kelompok yang lain. Algoritma K-means merupakan algoritma yang membutuhkan parameter input sebanyak k dan membagi sekumpulan objek kedalam k *cluster* sehingga tingkat kemiripan antar anggota dalam suatu *cluster* tinggi sedangkan tingkat kemiripan dengan anggota pada *cluster* lain sangat rendah. Kemiripan anggota terhadap *cluster* diukur dengan kedekatan objek terhadap nilai *mean* pada *cluster* atau dapat disebut sebagai *centroid cluster* atau pusat massa [11].

2.5. Klasterisasi Kecamatan berdasarkan Prevalensi Stunting

Klasterisasi kecamatan berdasarkan prevalensi stunting dapat dilakukan dengan menggunakan algoritma *clustering*. Algoritma *clustering* dapat digunakan untuk mengidentifikasi kecamatan-kecamatan yang memiliki prevalensi stunting tinggi sehingga dapat dilakukan intervensi yang lebih fokus dan terarah.

Berdasarkan landasan teori tersebut, penelitian ini bertujuan untuk melakukan klasterisasi kecamatan berdasarkan prevalensi stunting di Kabupaten Ciamis menggunakan algoritma K-Means. Hasil klasterisasi ini diharapkan dapat memberikan informasi yang bermanfaat bagi penyusunan intervensi penanggulangan stunting di Kabupaten Ciamis.

3. METODE PENELITIAN

3.1. Profile Perusahaan

Perusahaan yang dijadikan sebagai tempat penelitian Tugas Akhir ini adalah Dinas Kesehatan Kabupaten Ciamis.

a. Sejarah

Dinas Kesehatan Kabupaten Ciamis berdiri pada tanggal 17 Agustus 1950. Dinas Kesehatan

Kabupaten Ciamis merupakan salah satu perangkat daerah di lingkungan Pemerintah Kabupaten Ciamis yang mempunyai tugas pokok dan fungsi dalam penyelenggaraan urusan pemerintahan bidang kesehatan.

b. Visi

Visi dari Dinas Kesehatan Ciamis, yaitu “Terwujudnya Masyarakat Kabupaten Ciamis yang Sehat, Mandiri, dan Sejahtera”.

c. Misi

- Meningkatkan derajat kesehatan masyarakat Kabupaten Ciamis
- Meningkatkan kualitas pelayanan kesehatan masyarakat Kabupaten Ciamis
- Meningkatkan kemandirian masyarakat dalam bidang kesehatan
- Meningkatkan kesejahteraan masyarakat dalam bidang Kesehatan

d. Motto Dinkes

Motto yang digunakan oleh Dinas Kesehatan Kabupaten Ciamis, yaitu seperti gambar di bawah ini :



Gambar 1. Motto Dinkes

e. Tujuan

Beberapa tujuan dari Dinkes Kabupaten Ciamis, diantaranya :

- Meningkatkan derajat kesehatan masyarakat Kabupaten Ciamis yang setinggi-tingginya
- Meningkatkan kualitas pelayanan kesehatan masyarakat Kabupaten Ciamis yang bermutu dan terjangkau
- Meningkatkan kemandirian masyarakat dalam bidang Kesehatan

f. Fungsi

Beberapa fungsi dari Dinkes Kabupaten Ciamis, diantaranya :

- Perumusan kebijakan dan strategi pembangunan kesehatan di Kabupaten Ciamis
- Pemberdayaan masyarakat di bidang kesehatan
- Penyelenggaraan pelayanan kesehatan di Kabupaten Ciamis
- Penyelenggaraan pembinaan dan pengawasan kesehatan di Kabupaten Ciamis

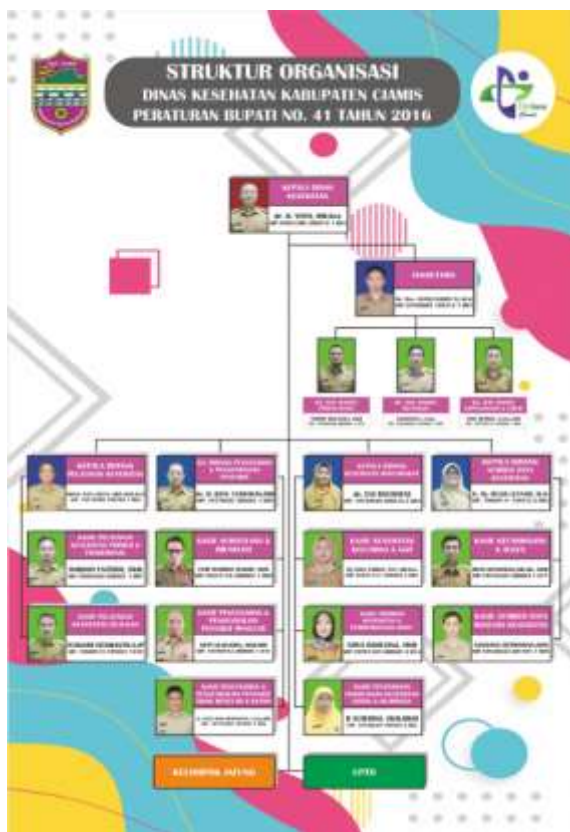
g. Program

Dinas Kesehatan Kabupaten Ciamis melaksanakan berbagai program kesehatan, antara lain:

- Program peningkatan kesehatan ibu dan anak
- Program peningkatan gizi masyarakat
- Program pengendalian penyakit
- Program peningkatan kualitas pelayanan kesehatan
- Program kemandirian masyarakat dalam bidang kesehatan

h. Susunan Organisasi

Dinas Kesehatan Kabupaten Ciamis dipimpin oleh seorang kepala dinas yang bertanggung jawab kepada Bupati. Kepala Dinas Kesehatan Kabupaten Ciamis dibantu oleh 4 orang kepala bidang. Selain itu, Dinas Kesehatan Kabupaten Ciamis juga memiliki 32 unit kerja yang tersebar di seluruh wilayah Kabupaten Ciamis. Berikut gambar susunan organisasinya :



Gambar 2. SOTK Dinkes

h. Kontak

Dinas Kesehatan Kabupaten Ciamis Jl. Pahlawan No. 177, Ciamis, Jawa Barat 46213 Telepon: (0265) 273404 Fax: (0265) 273404. Informasi mengenai Dinas Kesehatan juga dapat di akses pada website <https://dinkes.ciamiskab.go.id/> dan Instagram *dinkesciamis*.

3.2. Sumber Data

Sumber data prevalensi stunting tahun 2022 di Kabupaten Ciamis adalah hasil survei yang dilakukan

oleh Dinas Kesehatan Kabupaten Ciamis. Survei tersebut dilakukan pada bulan April - Mei 2022 dengan menggunakan metode pengukuran langsung. Data prevalensi stunting yang digunakan dalam penelitian ini adalah data prevalensi stunting pada balita usia 0-59 bulan. Data tersebut diperoleh dari hasil survei yang dilakukan di 27 kecamatan di Kabupaten Ciamis. Data prevalensi stunting di Kabupaten Ciamis tahun 2022 menunjukkan bahwa prevalensi stunting di Kabupaten Ciamis adalah sebesar 20,2%.

Sumber data yang digunakan dalam penelitian ini adalah data prevalensi stunting di 27 kecamatan di Kabupaten Ciamis tahun 2022. Data tersebut diperoleh dari Dinas Kesehatan Kabupaten Ciamis.

Data prevalensi stunting tersebut terdiri dari 2 kolom, yaitu:

- Kecamatan: Nama kecamatan di Kabupaten Ciamis
- Prevalensi Stunting: Persentase balita stunting di masing-masing kecamatan

Data tersebut kemudian diolah menggunakan algoritma K-Means untuk melakukan klasifikasi kecamatan berdasarkan prevalensi stunting.

3.3. Teknik Pengumpulan Data

Dalam penelitian "Klasterisasi Data Balita Stunting Di Kecamatan Wilayah Kabupaten Ciamis Berdasarkan Prevalensi Dengan Menggunakan Algoritma K-Means", teknik pengumpulan data tidak diperlukan secara langsung karena data yang digunakan sudah tersedia.

Data prevalensi stunting 2022 di 27 kecamatan Kabupaten Ciamis diperoleh secara tidak langsung melalui dokumentasi dari Dinas Kesehatan Kabupaten Ciamis. Dokumentasi ini berupa laporan, tabel, atau database elektronik yang memuat data tersebut.

Oleh karena itu, teknik pengumpulan data dalam penelitian ini tidak melibatkan wawancara, observasi, atau kuesioner, melainkan hanya mengambil dan mengolah data yang sudah ada dari sumber terpercaya.

3.4. Tahapan Perancangan

Dalam menunjang model, tugas akhir ini mempunyai beberapa tahapan yang mengacu pada *knowledge discovery in database* sebagai berikut:

- Data Selection* akan membuat kumpulan data target, memilih kumpulan data atau fokus pada subkumpulan variabel atau sampel data, tempat penemuan akan dilakukan. Hasil seleksi disimpan dalam file terpisah dari database operasional.
- Pre-processing / Cleaning* data adalah operasi dasar yang dilakukan seperti penghilangan *noise*. Proses pembersihan tersebut antara lain menghapus data duplikat, memeriksa ketidakkonsistenan data, dan memperbaiki kesalahan pada data, seperti kesalahan ketik. Data tersebut dapat diperkaya dengan data atau informasi eksternal yang relevan.

- c. *Transformation* data merupakan proses *pengintegrasian* data terpilih sehingga data tersebut sesuai untuk proses data mining. Ini adalah proses yang sangat bergantung pada jenis atau pola informasi yang perlu dicari dalam database.
- d. *Data mining*, pemilihan tugas *data mining* adalah pemilihan tujuan dari proses KDD, misalnya: karakterisasi, klasifikasi, regresi, *clustering*, asosiasi, dll. Pemilihan tugas penambangan data adalah pemilihan tujuan proses KDD, misalnya karakterisasi, klasifikasi, regresi, pengelompokan, asosiasi, dll. Pemilihan teknik, metode atau algoritma yang tepat sangat bergantung pada tujuan dan proses KDD secara keseluruhan.
- e. *Interpretation/Evaluation* adalah *penerjemahan* pola yang dihasilkan dari data mining. Model informasi yang dihasilkan harus ditampilkan dalam bentuk yang dapat dimengerti. Langkah ini memeriksa apakah model atau informasi yang ditemukan bertentangan dengan fakta atau hipotesis yang sudah ada sebelumnya.

4. HASIL DAN PEMBAHASAN

4.1. Menganalisis Perbedaan Prevalensi Stunting antar Kecamatan di Kabupaten Ciamis

Untuk menganalisis perbedaan prevalensi stunting antar Kecamatan di Kabupaten Ciamis. Peeliti menggunakan metode analisis data pada aplikasi IBM SPSS Statistics. *Statistical Product and Service Solutions* atau disingkat SPSS adalah sebuah program aplikasi untuk analisis statistik dan sistem pengelolaan data dengan menggunakan menu-menu sederhana dan kotak dialog yang mudah dipahami oleh pengguna [12].

Sebelum melakukan analisis, buat tabel distribusi frekuensi untuk penyusunan data ke dalam kelas-kelas tertentu dimana setiap data dimasukkan kedalam interval tertentu dan kemudian input data [12].

Tabel 2. Uji normalitas
Tests of Normality

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Persentase Balita Stunting	.170	265	.000	.840	265	.000

Setelah membuat tabel distribusi frekuensi dan input data, maka selanjutnya melakukan uji normalitas. Uji normalitas bertujuan untuk menguji apakah dalam model regresi, variabel independen dan dependennya berdistribusi normal atau tidak [13]. Uji Normalitas adalah uji yang dilakukan untuk menganalisis sebaran data apakah data berdistribusi normal atau tidaknya. Uji normalitas dengan program IBM SPSS Statistics 23 dengan menggunakan pendekatan Kolmogorov-Smirnov dan Shapiro-Wilk [12].

4.4.1. Lilliefors Significance Correction

Metode Shapiro-Wilk adalah uji yang dilakukan untuk mengetahui sebaran data acak suatu sampel data yang kurang dari 50 sampel. Pengujian ini dikatakan berdistribusi normal apabila nilai signifikansi lebih dari 0.05 (sig. > 0.05). Sementara itu metode Kolmogorov-Smirnov merupakan pengujian yang dilakukan untuk mengetahui sebaran data dengan ukuran data 20 sampai 1000 ($20 \leq N \leq 1000$). Pengujian ini dikatakan berdistribusi normal apabila nilai signifikansi lebih dari 0.05 (sig. > 0.05) [12]. Maka dapat disimpulkan bahwa data hasil uji normalitas di atas, tidak berdistribusi normal.

Setelah mengetahui nilai data tidak terdistribusi normal, maka selanjutnya dilakukan metode analisis non-parametric. Metode nonparametrik dalam statistik adalah teknik bebas distribusi yang tidak bergantung pada asumsi tentang distribusi populasi atau bentuk data. Tidak seperti metode parametrik, yang mengasumsikan distribusi tertentu atau parameter tetap, metode nonparametrik membuat asumsi minimal tentang data. Metode ini berguna dalam situasi di mana data tidak mengikuti distribusi normal, seperti dengan variabel ordinal, peringkat, atau ketika ada outlier [14].

Selanjutnya melakukan uji non-parametrik menggunakan metode Kolmogorov-Smirnov. Uji Kolmogorov-Smirnov dilakukan dengan membandingkan D_{hit} dan D_{tabel} .

$$D_{hitung} = \max |F_0(x) - S_n(x)|$$

Keterangan, $F_0(x)$ = distribusi frekuensi kumulatif teoritis. $S_n(x)$ = distribusi frekuensi kumulatif skor observasi. Uji Kolmogorov-Smirnov satu sampel digunakan untuk menguji apakah suatu sampel berasal dari distribusi tertentu. Prosedur ini digunakan untuk menentukan apakah sampel berasal dari populasi yang terdistribusi normal [15].

Tabel 3. Hasil *one-sample kolmogorov-smirnov test*

Hypothesis Test Summary				
	Null Hypothesis	Test	Sig.	Decision
1	The distribution of Persentase Balita Stunting is normal with a mean of 40.22 and a standard deviation of 34.895.	One-Sample Kolmogorov-Smirnov Test	1.083 E-20 ¹	Reject the null hypothesis.
Asymptotic significances are displayed. The significance level is .05.				
¹ Lilliefors Corrected				

Selain dengan menggunakan *One-Sample Kolmogorov-Smirnov Test*, dapat dilakukan juga dengan *Independent-Samples Kruskal-Wallis Test*. [14] Tes ini digunakan untuk membandingkan tiga atau lebih kelompok independen. dengan memberi peringkat data dalam setiap kelompok dan menentukan apakah ada perbedaan yang signifikan di antara kelompok-kelompok tersebut. Ini adalah alternatif

nonparametrik untuk *one-way ANOVA*. Statistik uji diberikan sebagai berikut:

$$H = \frac{12}{N(N+1)} \left(\frac{R^2_1}{n_1} + \frac{R^2_2}{n_2} + \dots + \frac{R^2_k}{n_k} \right) - 3(N+1)$$

Gambar 3. Rumus *independent-samples kruskal-wallis test*

Dalam persamaan di atas, K mewakili jumlah sampel, n_i menunjukkan ukuran sampel ke- i , N mewakili jumlah ukuran sampel, dan R_i menandakan jumlah peringkat sampel ke- i .

Tabel 4. Hasil *independent-samples kruskal-wallis test*

Hypothesis Test Summary				
	Null Hypothesis	Test	Sig.	Decision
1	The distribution of Persentase Balita Stunting is the same across categories of Nama Kecamatan.	Independent-Samples Kruskal-Wallis Test	.000	Reject the null hypothesis.
Asymptotic significances are displayed. The significance level is .05.				

Berdasarkan hasil uji non-parametrik yang ditunjukkan di atas, maka dapat disimpulkan bahwa terdapat perbedaan prevalensi stunting antar Kecamatan di Kabupaten Ciamis yang menunjukkan hasil uji di atas bahwa hipotesis null ditolak.

4.2. Menentukan cara pengelompokan kecamatan di Kabupaten Ciamis pada balita stunting menggunakan algoritma K-Means berdasarkan prevalensi

Berikut merupakan *output* dari pengeelompokan Kecamatan di Kabupaten Ciamis pada balita stunting menggunakan algoritma K-Means melalui proses yang dilakukan dengan menggunakan *tools RapidMiner*.

4.2.1. Data Selection

Data yang diolah dalam tugas akhir ini merujuk pada data prevalensi stunting antar Kecamatan di Kabupaten Ciamis. Dataset ini berjumlah 265 *record* dengan 8 atribut yang mencakup no, kecamatan, puskesmas, desa, jumlah stunting, persentase balita stunting, jumlah balita diukur, dan persentase jumlah balita yang diukur. Berikut Tabel yang belum di proses yang terdapat pada table 5

Tabel 5. Data original

No	Kecamatan	Puskesmas	Desa	Jumlah stunting	Persentase balita stunting	Jumlah balita diukur	Persentase balita diukur
1	Ciamis	Ciamis	Ciamis	6	100%	588	9700%
2	Ciamis	Ciamis	Kertasari	12	280%	432	9130%
4	Ciamis	Ciamis	Benteng	7	230%	308	9970%
....							
264	Lakbok	Sidharja	Sindangang in	6	270%	220	8660%
265	Lakbok	Sidharja	CintaRatu	9	270%	337	8800%

Dari 8 atribut yang ada di dalam table 4.4., yang akan di seleksi hanya 4 atribut saja diantaranya adalah kecamatan, jumlah stunting, persentase balita stunting, dan jumlah balita diukur.

Tabel 6. Data setelah diseleksi

KECAMATAN	JUMLAH STUNTING	PERSENTASE BALITA STUNTING	JUMLAH BALITA YANG DIUKUR
Ciamis	6	100%	588
Ciamis	11	230%	485
Ciamis	11	240%	459
...			
Lakbok	14	370%	375
Lakbok	9	270%	337

Langkah awal dalam *data selection* melibatkan operator *read excel*. Fungsi dari operator ini adalah untuk membaca informasi terkait pengelompokan data balita stunting di Kecamatan wilayah Kabupaten

Ciamis berdasarkan prevalensi yang disimpan dalam file MS-Excel. Dan langkah berikutnya yaitu import data.

Sebelum data diimpor, sistem akan menampilkan pratinjau data pada tahap *import data*. Hal ini bertujuan untuk meninjau kembali apakah data yang dimasukkan sudah benar. Setelah data sudah benar, pilih opsi "*Finish*" untuk menjalankan proses selanjutnya.

Langkah berikutnya yaitu menerapkan operator *select*. Fungsinya adalah untuk memisahkan atribut-atribut yang diperlukan dan tidak diperlukan dalam proses data mining.

Tahapan pemilihan atribut dapat terlihat pada gambar. Atribut yang dipertimbangkan mencakup kecamatan, jumlah stunting, persentase balita stunting, dan jumlah balita yang diukur. Atribut-atribut ini akan diisolasi dan disusun sebagai parameter dalam proses seleksi atribut.

Row No.	KECAMATAN	JUMLAH ST...	PERSENTAS...	JUMLAH BA...
1	Ciamis	5	1	585
2	Ciamis	12	2.883	432
3	Ciamis	8	2.480	254
4	Ciamis	7	2.380	388
5	Ciamis	5	1.480	440
6	Ciamis	11	2.380	440
7	Ciamis	6	1.720	441
8	Ciamis	8	2.180	367
9	Ciamis	12	2.480	438
10	Ciamis	5	1.480	421
11	Ciamis	6	2.720	340
12	Ciamis	10	6.480	386
13	Wanagiri	15	4.480	214
14	Wanagiri	10	0	335
15	Wanagiri	5	4.280	190

Gambar 4. Hasil *select attribute* pada rapidminer

Hasilnya menunjukkan bahwa dari 8 atribut, hanya 4 atribut yang akan digunakan dalam proses data mining.

4.2.2. Pre-processing / Cleaning Data

Dengan potensi adanya informasi yang tidak berhubungan. Langkah pembersihan data tetap sangat penting. Data yang tidak lengkap, duplikat, dan hilang akan dihapus. Karena keberadaan data yang tidak relevan dapat menurunkan tingkat akurasi dalam proses data mining. Operator yang digunakan dalam konteks ini yaitu *replace missing value*.

Setelah melalui tahap *cleaning data* menggunakan operator *replace missing value*, data yang digunakan dalam hal ini adalah data yang lengkap, tidak ada nilai yang hilang. Oleh karena itu,

jumlah data prevalensi balita stunting di Kabupaten Ciamis tahun 2022 tetap konsisten sebanyak 265 data. Proses selanjutnya akan melibatkan transformasi data dalam tahap data mining.

4.2.3. Data Transformation

Transformasi data melibatkan operator *real to integer*. Melalui penerapan operator *real to integer*, data yang awalnya memiliki format bilangan real akan diubah menjadi bilangan integer.

Setelah operator *real to integer* diterapkan, atribut persentase balita stunting mengalami transformasi menjadi integer. Hasilnya dapat dilihat pada gambar 5.

Row No.	KECAMATAN	JUMLAH ST...	PERSENTAS...	JUMLAH BA...
1	Ciamis	5	1	585
2	Ciamis	12	2.883	432
3	Ciamis	8	2.480	254
4	Ciamis	7	2.380	388
5	Ciamis	5	1.480	440
6	Ciamis	11	2.380	440
7	Ciamis	6	1.720	441
8	Ciamis	8	2.180	367
9	Ciamis	12	2.480	438
10	Ciamis	5	1.480	421
11	Ciamis	6	2.720	340
12	Ciamis	10	6.480	386
13	Wanagiri	15	4.480	214
14	Wanagiri	10	0	335
15	Wanagiri	5	4.280	190

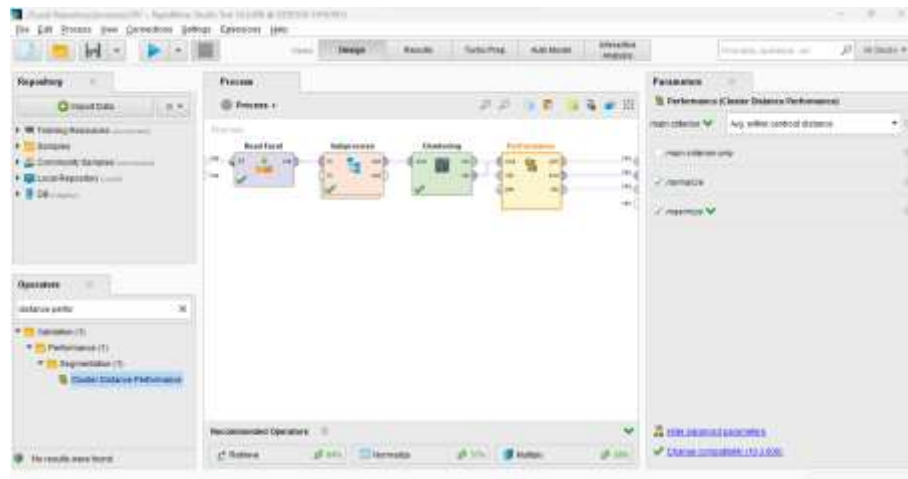
Gambar 5. Hasil *transformation real to integer*

4.2.4. Data Mining

Sebelum melakukan data mining agar terlihat ringkas dan rapi beberapa proses sebelumnya dapat dimasukkan pada operator *subprocess*. Langkah awal dalam data mining adalah mengelompokkan balita stunting antar Kecamatan di wilayah Kabupaten Ciamis berdasarkan prevalensi menggunakan algoritma K-Means. Algoritma K-Means

diimplementasikan dalam proses tersebut, dengan penentuan $k=3$ dan dilakukan hingga maksimal 10 kali.

Selanjutnya untuk melakukan proses penelusuran nilai jarak dan nilai DBI untuk mengukur kinerja suatu sistem atau perangkat komunikasi secara efektif dapat menggunakan operator *cluster distance performance*.



Gambar 6. Rangkaian proses data mining

Setelah melakukan serangkaian tahapan performa yang cermat dan teliti, seperti analisis data, pengukuran kinerja, dan evaluasi berbagai parameter

yang relevan, akhirnya menemukan hasil rata-rata (avg) yang mencerminkan tingkat efisiensi dan efektivitas suatu sistem atau proses.



Gambar 7. Hasil cluster distance performance

4.2.5. Data Interpretation/Evaluation

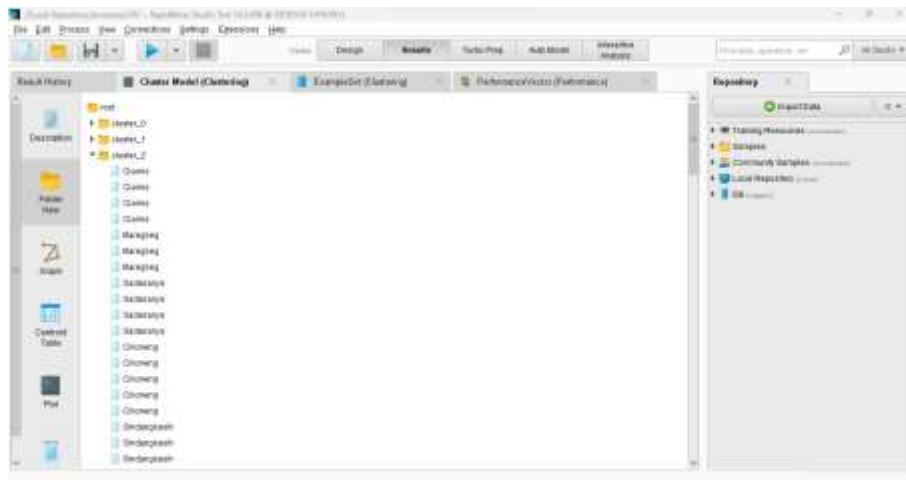
Dalam data interpretation, informasi diubah ke dalam format yang lebih sederhana untuk mempermudah pemahaman. Dalam proses ini

menghasilkan jumlah cluster sebanyak 265 yaitu cluster 0 berjumlah 99 items, cluster 1 berjumlah 54 items, dan cluster 2 berjumlah 112 items.



Gambar 8. Tampilan *cluster model*

Dalam *folder view*, setiap model *cluster* diatur secara terorganisir, memungkinkan pengguna untuk dengan mudah menavigasi dan mengelola entitas-entitas data yang terkait dengan setiap kelompok.



Gambar 9. Tampilan folder

Konsep dasar dalam grafik ini menyatakan bahwa semakin banyak elemen yang terlibat, menciptakan kompleksitas dan keterkaitan yang lebih besar di antara mereka.



Gambar 10. Tampilan grafik

Average Centroid merupakan titik pusat geometris dari sekelompok data atau objek, dihitung dengan merata-ratakan posisi koordinat semua titik dalam himpunan tersebut.

Cluster	Cluster_1	Cluster_2	Cluster_3	Cluster_4
PERSENTASE DATA STUNTING	5.345	1.805	2.630	1.805
Jumlah Stunting	6.181	10.130	6.700	6.700
Jumlah Anak Usia Dini	101.400	201.000	270.000	270.000

Gambar 11. Tampilan *centroid table*

Hasil dari sebuah visualisasi merupakan sebagai seni mewujudkan data menjadi bentuk yang mudah dipahami, memainkan peran kunci dalam

membuka jalan menuju pemahaman mendalam terhadap informasi yang kompleks.



Gambar 12. Tampilan visualisasi

4.3. Mengidentifikasi karakteristik masing-masing kluster kecamatan berdasarkan prevalensi stunting di Kabupaten Ciamis.

Hasil penelitian menunjukkan bahwa karakteristik masing-masing kluster kecamatan berdasarkan prevalensi stunting di Kabupaten Ciamis adalah sebagai berikut:

Tabel 7. Karakteristik *cluster*

Kluster	Prevalensi	Karakteristik Umum
1	<10%	Kecamatan dengan tingkat kesejahteraan yang relatif tinggi, akses terhadap layanan kesehatan dan pangan bergizi yang relatif baik, serta tingkat pendidikan masyarakat yang relatif tinggi.
2	10-20%	Kecamatan dengan tingkat kesejahteraan yang relatif sedang, akses terhadap layanan kesehatan dan pangan bergizi yang relatif cukup baik, serta tingkat pendidikan masyarakat yang relatif sedang.
3	>20%	Kecamatan dengan tingkat kesejahteraan yang relatif rendah, akses terhadap layanan kesehatan dan pangan bergizi yang relatif kurang baik, serta tingkat pendidikan masyarakat yang relatif rendah.

Secara umum, prevalensi stunting di Kabupaten Ciamis cenderung berkorelasi positif dengan tingkat kesejahteraan masyarakat, akses terhadap layanan kesehatan dan pangan bergizi, serta tingkat pendidikan masyarakat.

5. KESIMPULAN DAN SARAN

Berdasarkan penelitian yang sudah dilakukan di atas, ada beberapa kesimpulan yang dapat di ambil, diantaranya : Terdapat perbedaan prevalensi stunting antar kecamatan di Kabupaten Ciamis. Hal ini ditunjukkan oleh hasil uji statistic non-parametrik yang signifikan. Berdasarkan uji non-parametrik *Independent-Samples Kruskal-Wallis Test* yang menyatakan hipotesis “Sebaran Persentase Balita Stunting sama antar kategori Nama Kecamatan” ditolak. Jumlah kluster yang optimal untuk pengelompokan kecamatan di Kabupaten Ciamis berdasarkan prevalensi stunting adalah 3. Dengan menggunakan nilai $K = 3$, maka diperoleh tiga kluster kecamatan, yaitu: Kluster 1: Kecamatan dengan prevalensi stunting rendah (<10%), Kluster 2: Kecamatan dengan prevalensi stunting sedang (10-20%), Kluster 3: Kecamatan dengan prevalensi stunting tinggi (>20%), Karakteristik masing-masing kluster kecamatan berdasarkan prevalensi stunting di Kabupaten Ciamis tertera pada table 7. Karakteristik *cluster*.

Penulis menyadari ada beberapa kekurangan dalam hal penyusunan peneltia ini. Oleh karena itu, ada beberapa saran yang diharapkan akan lebih baik ke depannya dalam penyusunan penelitian, diantaranya: Meningkatkan jumlah data dan variabel yang dianalisis. Jumlah data yang digunakan dalam penelitian ini adalah 265 data dari 27 Kecamatan di Kabupaten Ciamis. Jumlah data ini cukup untuk melakukan analisis klasterisasi dengan algoritma K-Means. Namun, untuk mendapatkan hasil yang lebih akurat, disarankan untuk meningkatkan jumlah data dengan menambahkan kecamatan-kecamatan lain atau wilayah lainnya di Kabupaten Ciamis. Selain itu, disarankan untuk menambahkan variabel-variabel lain yang dapat mempengaruhi prevalensi stunting, seperti, karakteristik demografis masyarakat, kondisi ekonomi

masyarakat, dan kondisi lingkungan. Variabel-variabel ini dapat memberikan gambaran yang lebih lengkap tentang faktor-faktor yang mempengaruhi prevalensi stunting di Kabupaten Ciamis. Melakukan analisis lebih lanjut terhadap karakteristik masing-masing klaster Kecamatan. Hasil penelitian menunjukkan bahwa terdapat perbedaan prevalensi stunting antar kecamatan di Kabupaten Ciamis. Namun, penelitian ini hanya memberikan gambaran umum tentang karakteristik masing-masing klaster Kecamatan. Untuk mendapatkan informasi yang lebih mendalam, disarankan untuk melakukan analisis lebih lanjut terhadap karakteristik masing-masing klaster Kecamatan. Analisis lebih lanjut dapat dilakukan dengan menggunakan metode analisis statistik lainnya, seperti, analisis regresi atau analisis korelasi. Metode analisis statistik ini dapat digunakan untuk menganalisis hubungan antara prevalensi stunting dengan faktor-faktor yang mempengaruhinya. Melakukan analisis longitudinal. Penelitian ini hanya menggunakan data prevalensi stunting tahun 2022. Untuk mendapatkan gambaran yang lebih lengkap tentang prevalensi stunting di Kabupaten Ciamis, disarankan untuk melakukan analisis longitudinal. Analisis longitudinal dapat dilakukan dengan menggunakan data prevalensi stunting dari tahun-tahun sebelumnya. Analisis ini dapat digunakan untuk melihat tren prevalensi stunting di Kabupaten Ciamis dari waktu ke waktu.

DAFTAR PUSTAKA

- [1] A. Fadilah, M. Nurfaizy P, S. Lumbanbatu, and S. Defiyanti, "PENGELOMPOKAN KABUPATEN/KOTA DI INDONESIA BERDASARKAN FAKTOR PENYEBAB STUNTING PADA BALITA MENGGUNAKAN ALGORITMA K-MEANS," *Jurnal Informatika dan Komputer*, vol. 6, no. 2, pp. 223–230, 2022.
- [2] M. Arief Soeleman and F. Ilmu Komputer, "Penentuan Centroid Awal Pada Algoritma K-Means Dengan Dynamic Artificial Chromosomes Genetic Algorithm Untuk Tuberculosis Dataset Pre-Centroid Determination in K-Means Algorithm using Dynamic Artificial Chromosomes Genetic Algorithm for Tuberculosis Dataset," 2021.
- [3] W. Widayani and H. Harliana, "Perbandingan Algoritma K-Means dan SFCM Pada Pengelompokan Rumah Tangga Miskin," *Jurnal Sains dan Informatika*, vol. 6, no. 1, pp. 1–9, Jun. 2020, doi: 10.34128/jsi.v6i1.200.
- [4] D. Transaksi Bongkar Muat di Provinsi Riau, I. Kamila, U. Khairunnisa, P. Studi Sistem Informasi, and F. Sains dan Teknologi UIN Sultan Syarif Kasim Riau, "Perbandingan Algoritma K-Means dan K-Medoids untuk Pengelompokan," *Jurnal Ilmiah Rekayasa dan Manajemen Sistem Informasi*, vol. 5, no. 1, pp. 119–125, 2019.
- [5] A. Winarta and W. J. Kurniawan, "OPTIMASI CLUSTER K-MEANS MENGGUNAKAN METODE ELBOW PADA DATA PENGGUNA NARKOBA DENGAN PEMROGRAMAN PYTHON," *Jurnal Teknik Informatika Kaputama (JTIK)*, vol. 5, no. 1, 2021.
- [6] P. Sitti Patimah, Mk. Suchi Avnalurini Sharief, and Mk. C. Penerbit Cahaya Bintang Cemerlang, *PENCEGAHAN ANAK STUNTING BERBASIS DESA*. 2022.
- [7] Priyono, "Strategi Percepatan Penurunan Stunting Perdesaan (Studi Kasus Pendampingan Aksi Cegah Stunting di Desa Banyumundu, Kabupaten Pandeglang)," *Good Governance*, vol. 16, no. Vol 16, No 2 (2020): September, 2020.
- [8] 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. 20–33, Mar. 2023, doi: 10.56211/helloworld.v2i1.230.
- [9] B. Hadi Prakoso *et al.*, "Klasterisasi Puskesmas dengan K-Means Berdasarkan Data Kualitas Kesehatan Keluarga dan Gizi Masyarakat," 2023.
- [10] E. Irfiani, S. Sulistia Rani, S. Nusa Mandiri Jl Kramat Raya No, and J. Pusat, "Algoritma K-Means Clustering untuk Menentukan Nilai Gizi Balita," vol. 6, no. 4, pp. 17–27, 2018.
- [11] F. Yunita, "PENERAPAN DATA MINING MENGGUNAKAN ALGORITMA K-MEANS CLUSTERING PADA PENERIMAAN MAHASISWA BARU (STUDI KASUS : UNIVERSITAS ISLAM INDRAGIRI)," 2018.
- [12] E. Haryono, M. Slamet, and S. Damar, *STATISTIKA SPSS* 28. 2023. [Online]. Available: www.penerbitwidina.com
- [13] Niken Nanincova, "PENGARUH KUALITAS LAYANAN TERHADAP KEPUASAN PELANGGAN NOACH CAFE AND BISTRO," 2019.
- [14] S. Jhade and M. Muthusamy, *Parametric and Non-Parametric Analysis*, vol. 1. 2023. [Online]. Available: <https://www.researchgate.net/publication/372487275>
- [15] Quraisy Andi, "Normalitas Data Menggunakan Uji Kolmogorov-Smirnov dan Saphiro-Wilk," *J-HEST: Journal of Health, Education, Economics, Science, and Technology*, vol. 3, no. 1, pp. 7–11, 2020.