

ANALISIS SEGMENTASI PUSKESMAS DI KABUPATEN CIREBON MENGUNAKAN ALGORITMA *PARTITIONING AROUND MEDOID* BERDASARKAN INDIKATOR PENYEBAB STUNTING

Cahyaningrum Komalasari, Ahmad Faqih, Fatihanursari Dikananda,
Muhamad Sulaeman, Heliyanti Susana

Teknik Informatika, Sekolah Tinggi Manajemen dan Informatika IKMI Cirebon
Jalan Perjuangan No.10B Karyamulya Kec.Kesambi Kota Cirebon Jawa barat 45135, Indonesia
ningrumcahya40@gmail.com

ABSTRAK

Stunting merupakan keadaan dimana seseorang mengalami kekurangan gizi. Namun, hingga saat ini, belum terdapat segmentasi Puskesmas terkait stunting berdasarkan indikatornya menggunakan data mining di Kabupaten Cirebon. Pada penelitian ini bertujuan untuk menganalisis segmentasi puskesmas terkait stunting dengan indikatornya guna memberikan rekomendasi pencegahan dan penyuluhan terkait stunting menggunakan *Algoritma Partitioning Around Medoid (PAM)*. Data bersumber dari website opendata.jabarprov.go.id dan opendata.cirebonkab.go.id dan observasi langsung ke dinas Kesehatan kab. Cirebon. Variabel yang digunakan Jumlah Ibu Hamil yang mengalami KEK (Kekurangan Energi Kronik), Jumlah stunting, Jumlah Ahli Gizi dan Jumlah Puskesmas aktif yang ada di kabupaten Cirebon pada tahun 2022. Proses pengolahan data menggunakan metode *KDD (Knowledge Discovery in Database)*. Implementasi k-medoids dilakukan dengan inisiasi nilai centroid secara acak dalam iterasi. Hasil menunjukkan medoid sementara sebesar 7122,045634, dan medoid awal 6348,966306, dengan selisih 773,0793279 hasil yang signifikan mengakibatkan penghentian iterasi pada tahap pertama. Evaluasi manual dengan rumus *SSW (Sum of Square Within Cluster)* dan *SSB (Sum of Square Between)* menghasilkan rasio cluster 0,8470604, mendekati 0, menunjukkan hasil yang baik dan Evaluasi cluster menggunakan *Davies-Bouldin Index (DBI)* menggunakan RapidMiner menghasilkan nilai 0,569. Cluster 1 (19 puskesmas) memiliki kasus stunting tertinggi, sedangkan Cluster 2 (41 puskesmas) memiliki kasus stunting rendah, memberikan dasar rekomendasi pencegahan serta penyuluhan yang lebih terarah.

Kata kunci: Data Mining, K-Medoid, Segmentasi Stunting, PAM (*Partitioning Around Medoid*)

1. PENDAHULUAN

Dalam era modern ini, kemajuan teknologi telah memberikan kontribusi besar terhadap solusi permasalahan di berbagai sektor, termasuk pendidikan, kesehatan, dan pertanian. Salah satu perhatian pemerintah adalah kasus stunting, yang dapat menghambat pertumbuhan anak secara fisik dan intelektual. Dalam upaya penanggulangan, inovasi teknologi, seperti data mining dengan algoritma PAM (*Partitioning Around Medoid*), telah digunakan untuk melakukan analisis segmentasi terhadap kasus stunting. Penelitian ini, khususnya di Kabupaten Cirebon, mencoba memberikan kontribusi dengan menggali pola, korelasi, dan hubungan antara variabel-variabel terkait stunting untuk memahami kelompok-kelompok risiko tinggi.

Dampak stunting tidak hanya terbatas pada kesehatan fisik anak, tetapi juga dapat memengaruhi kemampuan intelektual mereka di masa depan. Hal ini bisa jadi berpotensi menghambat kontribusi mereka dalam penyaluran sumber daya manusia untuk sektor teknologi dan informatika. Dengan adanya penelitian dan inovasi di bidang teknologi, seperti analisis segmentasi dengan metode PAM, diharapkan dapat memberikan informasi yang lebih mendalam dan relevan untuk merancang intervensi yang tepat sasaran dalam penanggulangan stunting.

Meskipun angka stunting di Kabupaten Cirebon mengalami penurunan, masih diperlukan pendekatan holistik dan upaya lebih lanjut. Dengan fokus pada analisis segmentasi di Puskesmas menggunakan metode PAM, diharapkan pemerintah setempat dapat mengidentifikasi kelompok-kelompok risiko tinggi secara lebih spesifik. Informasi ini akan sangat berharga dalam merancang program intervensi yang lebih efektif dan tepat sasaran untuk mengurangi tingkat stunting di Kabupaten Cirebon.

2. TINJAUAN PUSTAKA

Dalam penelitian ini Pada penelitian ini metode literature review yang digunakan adalah *Systematic Literature Riview* atau biasa disebut SLR dimana penulis secara sistematis mengumpulkan, yang kemudian menganalisis secara kritis dengan menyajikan data-data yang terkait dengan penanggulangan kasus stunting menggunakan data mining. Untuk literature riview sendiri memiliki langkah-langkah yang meliputi beberapa tahapan, yaitu : 1. Melakukan eksplorasi topik, 2. Pencarian dan penyaringan literature, 3. Penelitian dan seleksi artikel, 4. Analisis dan sintesis literature, 5. Menulis literature riview sesuai struktur kepenulisan [1]. Dari literature riview tersebut didapati landasan teori yang digunakan dalam ini adalah sebagai berikut:

2.1. Data mining

Data mining merupakan proses menemukan korelasi, pola, dan kebiasaan dengan memilah-milah data besar dengan menggunakan teknologi pengenalan pola seperti statistik dan matematika [2]. Secara sederhana Data mining merupakan pola untuk menemukan informasi dari database yang besar. Ada 5 fungsi utama dalam tahapan data mining, yaitu Estimasi, Prediksi, Asosiasi, Klasifikasi dan Klasterisasi, dimana dalam penelitian ini penulis menggunakan fungsi Klasterisasi dalam data mining.

Untuk penelitian terdahulu, penulis menggunakan beberapa jurnal dalam penelitian ini, yang digunakan sebagai pembandingan untuk relevansi penelitian yang dilakukan, adapun jurnal tersebut adalah :

- a. Pada *paper* [3] yang mana dalam *paper* tersebut membahas mengenai implementasi algoritma *Partitioning Around Medoids* (PAM) untuk mengelompokkan hasil produksi komoditas perkebunan di Provinsi Kalimantan Timur. Penelitian ini bertujuan untuk mengoptimalkan produksi dengan mengelompokkan area berdasarkan tingkat produksinya. Algoritma PAM dibandingkan menggunakan tiga metode jarak: *Euclidean Distance*, *Manhattan Distance*, dan *Chebyshev Distance* Kualitas kelompok dievaluasi menggunakan *Silhouette Coefficient* (SC).
- b. *Paper* [4] dijadikan *paper* kedua mengenai kasus stunting yang ada di Indonesia, *paper* ini sendiri membahas mengenai masalah stunting di Indonesia dan upaya pemerintah dalam mengatasi masalah tersebut melalui Program 1000 HPK.
- c. *Paper* [5] yang membahas tentang pemanfaatan data mining dalam penentuan penyuluhan penyakit stunting menggunakan metode *Partitioning Around Medoids* (PAM). Penelitian ini bertujuan untuk mengelompokkan Puskesmas berdasarkan karakteristiknya dan menentukan prioritas dalam penyuluhan penyakit stunting. Metode PAM digunakan untuk mengklasifikasikan dan mengurutkan data agar dapat diambil keputusan yang akurat dalam penyuluhan penyakit stunting di Kota Palembang. Pada *paper* yang kelima berjudul “Medoid-based Clustering pada Kecamatan di Kabupaten Lebak dan Pandeglang Provinsi Banten Berdasarkan Trilogi Ketahanan Pangan” [6] yang membahas tentang penggunaan klaster berbasis medoid untuk menganalisis status keamanan pangan di distrik-distrik di Provinsi Banten, Indonesia.

2.2. Algoritma Partitioning Around Medoid (PAM)

Menurut [7] Algoritma *Partitioning Around Medoid* adalah algoritma yang digunakan untuk menemukan k cluster di dalam n objek pertama kali secara acak untuk menemukan medoid (wakil dari objek) untuk setiap cluster yang ada. *Partitioning Around Medoid* juga disebut sebagai K-Medoid,

dimana K-Medoid termasuk ke dalam partitional dengan metode *non-hierarchy*.

2.3. Clustering

Teknik *clustering* pertama kali ditemukan pada 1957 oleh Lloyd, yang mana pertama kalinya di temukan adalah *K-Means Clustering*. Dimana *clustering* sendiri artinya segmentasi dataset yang menjadi inputan tanpa diketahui target classnya. Dataset yang dikelompokkan di kumpulkan berdasarkan kemiripan antara instance yang ada [8]. Beberapa Teknik *clustering* yang sudah di kembangkan semenjak *clustering* pertama di temukan, diantaranya adalah :

- a. *Partitional Clustering*, dimana jenis *clustering* yang setiap elemennya hanya dimiliki oleh cluster itu saja, dan cluster yang lain tidak boleh memilikinya.
- b. *Overlapping Clustering*, disebut juga *soft clustering* dimana setiap elemennya boleh dimiliki oleh cluster lainya.
- c. *Hierarchical Clustering*, yang biasa disebut *multilevel hierarchy* dimana cara kerjanya adalah dengan mengelompokkan cluster yang lebih besar menjadi dua atau lebih cluster menjadi cluster yang lebih kecil sehingga membentuk sebuah *tree diagram* (diagram pohon).
- d. *Density Based Clustering*, cluster ini berkaitan dengan kepadatan sebuah objek, dimana objek yang paling padat akan dipisahkan oleh objek yang lebih memiliki celah.
- e. *Model Based Clustering*, cluster yang model elemennya dibentuk oleh asumsi/model matematika atau statistika standar.

2.4. KDD (Knowledge Discovery in Database)

Knowledge Discovery in Database (DBI) adalah sebuah metode untuk memperoleh pengetahuan dari database, yang mana di dalemnya terdapat tabel-tabel yang saling berkorelasi. Dari informasi atau pengetahuan tersebut bisa digunakan untuk *knowledge base* (basis pengetahuan) yang berfungsi untuk mengambil sebuah keputusan. Istilah *Knowledge Discovery in Database* (KDD) dan data mining sering kali disamakan untuk menggambarkan proses penemuan informasi tersembunyi dalam basis data yang luas. Sebenarnya, keduanya memiliki konsep yang berbeda tetapi saling terkait, dengan data mining sebagai salah satu tahapan dalam proses KDD secara keseluruhan [9]. Dimana tahapan dalam KDD adalah :

- a. *Data Selection*.

Sebelum memulai tahap penggalan informasi dalam *Knowledge Discovery in Database* (KDD), penting untuk melakukan seleksi data dari sekumpulan data operasional. Data yang telah dipilih untuk proses data mining akan disimpan dalam berkas terpisah dari basis data operasional pada fokus *Knowledge Discovery in Database* (KDD).

b. *Pre-processing/Cleaning*

Tahap pembersihan melibatkan langkah-langkah seperti menghapus duplikasi data, memeriksa konsistensi data, dan memperbaiki kesalahan seperti *missing value*. Selain itu, dilakukan proses *enrichment* yang melibatkan penambahan data atau informasi yang relevan dan diperlukan untuk *Knowledge Discovery in Database (KDD)*, seperti data atau informasi eksternal.

c. *Transformation*

Proses *coding/transformation* dalam *Knowledge Discovery in Database (KDD)* adalah suatu proses kreatif yang sangat bergantung pada jenis atau pola informasi yang akan ditemukan dalam basis data yang sesuai dengan proses data mining.

d. *Data Mining*

Dalam proses *KDD* data mining digunakan untuk pencarian pola atau informasi menarik dalam data yang telah dipilih, menggunakan teknik atau metode khusus. Berbagai teknik, metode, atau algoritma dapat digunakan dalam data mining, dan pemilihan metode atau algoritma yang sesuai sangat tergantung pada tujuan serta proses *Knowledge Discovery in Database (KDD)* secara keseluruhan.

e. *Interpretation/ Evaluation.*

Tahap ini melibatkan penelitian untuk mengevaluasi apakah pola atau informasi yang ditemukan konsisten dengan fakta atau hipotesis yang telah ada sebelumnya.

2.5. DBI (Davies Bouldin Index)

David L. Davies dan Donald W. Bouldin adalah orang yang memperkenalkan *DBI* pertama kali pada tahun 1979. *DBI* sendiri merupakan rasio rata-rata jarak antar *cluster* dengan *cluster* terdekatnya. *DBI* bertujuan untuk mengevaluasi kinerja dari setiap segmentasi [10]. *DBI* juga digunakan sebagai metrik segmentasi karena secara umum dari segmentasi berisi dua kategori utama, yaitu validasi internal dan eksternal yang digunakan untuk mengukur hasil kinerja dari setiap *clustering*. Selain itu juga, *DBI* juga memiliki “*within-class*” untuk korelasi positif dan “*between-class*” untuk korelasi negative.

2.6. Stunting

Stunting merupakan keadaan dimana balita dengan masalah gizi, yang penyebabnya dimulai dari dalam kandungan dengan faktor Ibu hamil yang mengalami kekurangan Energi kronik, sehingga menghasilkan manusia yang memiliki kualitas yang kurang baik, karena dapat menyerang motorik otak dan membuat pertumbuhan terhambat baik dari segi mental maupun fisik [5].

3. METODE PENELITIAN

Metode yang digunakan dalam penelitian ini adalah *KDD* (*Knowledge Discovery in Database*), dengan tahapan Data Selection, Data Preprocessing, Data Transformation, Data Mining atau implementasi

dalam algoritma PAM (*Partitioning Around Medoid*) dan *Evaluation and Interpretation*.

3.1. Data selection

Pada proses *data selection* penulis menetapkan lokasi dan dataset yang akan diteliti, dengan memilih dataset terkait kasus stunting di Kabupaten Cirebon berdasarkan indikator tertentu. Indikator tersebut mencakup Jumlah anak yang mengalami stunting usia 0-5 tahun, Jumlah Tenaga Gizi, Jumlah Ibu Hamil yang Mengalami KEK (*Kekurangan Energi Kronik*), Jumlah Posyandu Aktif dengan rentang waktu tahun 2022, yang mana data-data tersebut diambil dari website Open Data Jawa Barat, Open Data Kabupaten Cirebon dan Observasi langsung atau bisa juga disebut sebagai Teknik *netnografi* yaitu penelitian yang menitikberatkan pada pemahaman terhadap ruang siber di mana individu-individu berinteraksi dan mampu membentuk budaya serta sistem masyarakat yang unik di dalamnya [11] dan juga Observasi langsung ke Dinas Kesehatan kab. Cirebon.

3.2. Data Pre-processing

Langkah selanjutnya adalah *pre-processing* data. Dalam dataset sebelumnya, semua data yang didapat secara terpisah, sehingga dalam tahapan ini integrasi data diperlukan, dimana langkah ini dilakukan untuk menghasilkan dataset yang bersih dengan tujuan memperoleh informasi yang bermanfaat. Integrasi data ini juga menghilangkan beberapa komponen, agar sesuai dengan dataset yang diinginkan dengan menghilangkan beberapa komponen dalam pengimplementasian algoritma *K-Medoid*, dan data yang dihilangkan adalah berupa kolom *id* dan satuan.

3.3. Transformation

Pada proses transformasi data penulis hanya mengubah tipe data pada pengimplementasian dalam perhitungan Algoritma *Partitioning Around Medoid* di software RapidMiner 10.1 guna mendapatkan hasil dan informasi yang dibutuhkan.

3.4. Data Mining

Pada tahapan data mining ini, dilakukan perhitungan menggunakan Algoritma *K-Medoid* atau biasa juga disebut *Partitioning Around Medoid*, menggunakan bantuan software Rapid Minner, yang mana perhitungan tersebut juga dapat dilakukan secara manual dengan langkah-langkah sebagai berikut [12] :

- Tentukan Medoid awal.
- Inisialisasi pusat *cluster* sebanyak k (jumlah *cluster*), dimana dalam penelitian ini penulis menggunakan 2 *cluster* yaitu C1 dan C2.
- Alokasikan objek ke dalam *cluster* terdekat dengan menggunakan persamaan ukuran jarak Eulidian Distance, yang digambarkan dalam persamaan berikut :

$$d(x,y) = \sqrt{\sum_{n=1}^i (x_i - y_i)^2}$$

Dengan :

n : jumlah suatu data

k : indeks data

X_k : Nilai atribut ke- k dari X

Y_k : nilai atribut ke- k dari Y

- Memilih secara acak objek untuk dijadikan medoid baru, yang kemudian di hitung jarak masing-masing objek pada *cluster* dengan kandidat medoid baru.
- Menghitung total simpangan, dengan menghitung total *distance*, dengan persamaan :

$$S = \text{total cost baru} - \text{Total cost lama}$$

Dengan :

S = Selisih

Total Cost baru = jumlah cost non-medoids

Total Cost lama = jumlah cost medoids Jika $S < 0$, maka tukar objek dengan cluster untuk membentuk objek menjadi medoid baru.

- Lakukan pengulangan medoid untuk mendapatkan perbedaan kedekatan.

3.5. Evaluation and Interpretation

Pada tahap evaluasi ini, penulis menggunakan DBI (Davies Bouldin Index) untuk mengukur kemiripan dari setiap kelompok-kelompok guna melakukan evaluasi dalam menentukan banyaknya *cluster* yang terbaik agar mengoptimalkan *distance* antara kelompok yang pertama dan pada waktu yang bersamaan mencoba untuk meminimalkan interval antara objek dalam suatu *cluster* [13]

4. HASIL DAN PEMBAHASAN

4.1. HASIL

4.1.1. Data Selection

Seleksi data menjadi tahapan pertama dalam penelitian ini setelah melakukan analisis terhadap kasus-kasus stunting dan membaca literature review dari jurnal-jurnal terkait stunting, dapat disimpulkan bahwa indikator penyebab stunting antara lain adalah Jumlah anak yang mengalami stunting dengan rentang usia 0-5 tahun, Jumlah Ibu hamil yang mengalami KEK (Kekurangan Energi Kronik), Jumlah ahli gizi, dan Jumlah posyandu aktif. Data-data yang didapatkan berupa file excel, yang mana data tersebut dapat dilihat pada gambar 1,2,3 dan 4.

DATA BALITA STUNTING PER DESA HASIL BPB AGUSTUS TH 2022									
NO	Desa/Kelurahan	TBU				TOTAL	Stunting	Prevalensi stunting	Nama Puskesmas
		Sangat Pendek	Pendek	Normal	Tinggi				
6	1 WALED DESA	21	39	181	0	241	60	24,90	WALED
7	2 WALED KOTA	14	45	271	1	331	59	17,82	WALED
8	3 MEKARSARI	11	30	190	0	231	41	17,75	WALED
9	4 WALED ASEH	11	25	129	0	165	36	21,82	WALED
10	5 AMBIT	43	70	960	1	1074	113	23,84	WALED
11	6 GUNUNGISARI	18	58	217	0	293	76	25,94	CIBOGO
12	7 CUYAH	6	34	524	0	564	40	7,09	WALED
13	8 KARANGSARI	5	2	277	0	284	7	2,46	CIBOGO
14	9 CIRULAK KIDUL	4	52	536	0	592	56	9,46	CIBOGO
15	10 CIRULAK	5	46	489	0	540	51	9,44	CIBOGO
16	...	...	...	...	...	...	...	...	...
17	424 GUWA LOR	18	58	226	4	306	76	24,84	KALIWEDI
18	TOTAL	3380	10634	148192	1038	163244	14014	8,58	

Gambar 1. Data Stunting Balita perdesa Tahun 2022

IBU HAMIL KURANG ENERGI KRONIK (KEK) MENDAPAT MAKANAN TAMBAHAN PERPUSKESMAS KABUPATEN CIREBON TAHUN 2022									
No	Puskesmas	Cak. Ibu Hamil Mendapat Tablet Tambahan Darah (TDD) Minimal 90 Tablet selama Masa Kehamilan			Cak. Ibu Hamil Kurang Energi Kronik (KEK) Mendapat Makanan Tambahan			%	%
		Jumlah Ibu Hamil	Jumlah Ibu Hamil Mendapat TDD Minimal 90 Tablet selama Masa Kehamilan		Jumlah Ibu Hamil KEK	Jumlah Ibu Hamil KEK Mendapat Makanan Tambahan			
7	1 WALED	373	373	100,00	45	29	64,44		
8	2 CIBOGO	712	687	92,28	61	18	29,51		
9	3 PASALEMAN	465	409	87,95	51	25	49,02		
10	4 CILEDUG	883	836	94,68	60	40	66,67		
11	5 PABUARAN	864	854	98,84	67	37	55,22		
12	6 LOSARI	720	713	99,03	87	57	65,56		
13	7 ASTANALANGGAR	514	507	98,64	50	46	92,00		
14	8 PABEDILAN	523	523	100,00	64	64	100,00		
15	9 KALIMUKTI	700	674	96,29	59	26	44,07		
16	10 BABAKAN	886	789	89,05	105	105	100,00		
17	...	...	...	...	...	...	...		
18	60 KALIWEDI	115	119	103,48	89	89	100,00		

Gambar 2. Dataset Ibu Hamil Kurang Energi Kronik di Kabupaten Cirebon tahun 2022

DATA TENAGA GIZI TAHUN 2021-2022				
No.	Nama Faskes/ Puskesmas	tahun 2021	tahun 2022	
1	Kalimaro	1	1	
2	WALED	1	1	
3	CIBOGO	1	1	
4	PASALEMAN	1	1	
5	CILEDUG	2	2	
6	PABUARAN	1	1	
7	LOSARI	2	2	
8	ASTANALANGGAR	1	1	
9	PABEDILAN	1	2	
10	KALIMUKTI	1	1	
...	...	...	...	
60	KALIWEDI	1	1	

Gambar 3. Data Tenaga Gizi di Kabupaten Cirebon tahun 2021-2022

A	B	C	D	E	F	G	H	I	J
1	2	3	4	5	6	7	8	9	10
1	tps nama kabupaten kota	tps kode kecamatan	tps nama kecamatan	kependudukan kode kecamatan	kependudukan nama kecamatan	puskesmas	jumlah_posyandu aktif	jumlah tahun	
2	KABUPATEN CIREBON	DENGGI WALED	12.05.01	WALED	WALED	13.00.01	2022		
3	KABUPATEN CIREBON	DENGGI WALED	12.05.01	WALED	CIBOGO	20.00.01	2022		
4	KABUPATEN CIREBON	DENGGI PASALEMAN	12.05.01	PASALEMAN	PASALEMAN	13.00.01	2022		
5	KABUPATEN CIREBON	DENGGI CILEDUG	12.05.01	CILEDUG	CILEDUG	13.00.01	2022		
6	KABUPATEN CIREBON	DENGGI PABUARAN	12.05.01	PABUARAN	PABUARAN	13.00.01	2022		
7	KABUPATEN CIREBON	DENGGI LOSARI	12.05.01	LOSARI	LOSARI	13.00.01	2022		
8	KABUPATEN CIREBON	DENGGI ASTANALANGGAR	12.05.01	ASTANALANGGAR	13.00.01	2022			
9	KABUPATEN CIREBON	DENGGI PABEDILAN	12.05.01	PABEDILAN	PABEDILAN	14.00.01	2022		
10	KABUPATEN CIREBON	DENGGI KALIMUKTI	12.05.01	KALIMUKTI	KALIMUKTI	13.00.01	2022		
...	...	...	...	...	...	...	...		
12	KABUPATEN CIREBON	DENGGI KALIWEDI	12.05.01	KALIWEDI	KALIWEDI	14.00.01	2022		

Gambar 4. Data Posyandu Aktif di Kabupaten Cirebon tahun 2022

4.1.2. Data Preprocessing

Pada dataset Jumlah Anak terindikasi Stunting (usia 0-5 tahun), fokus variabel hanya pada Jumlah Anak Stunting dan Puskesmas. Sementara pada dataset Jumlah Tenaga Gizi Kabupaten Cirebon, variabel yang digunakan adalah Jumlah Tenaga Gizi. Dataset Ibu Hamil Kurang Energi Kronik (KEK) Mendapat Makanan Tambahan difokuskan pada variabel Jumlah Ibu Hamil dengan KEK, dan pada dataset Jumlah Posyandu Aktif, variabel yang diambil adalah Jumlah Posyandu Aktif.

Dari setiap dataset terdapat nama Puskesmas, sehingga memasukkan nama puskesmas sebagai variabel dan labelnya, yang distandarisasikan dengan menambahkan kemendagri nama kecamatan, satuan dan tahun. Hasil dataset yang bersih dapat dilihat dari gambar 5.

A	B	C	D	E	F	G	H	I
kependudukan nama kecamatan	nama puskesmas	jumlah anak terindikasi stunting	jumlah tenaga gizi	jumlah ibu hamil mengalami kek	jumlah posyandu aktif	satuan	tahun	
1 WALED	WALED	345	1	45	31	ORANG	2022	
2 WALED	CIBOGO	265	1	61	29	ORANG	2022	
3 PASALEMAN	PASALEMAN	235	1	51	32	ORANG	2022	
4 CILEDUG	CILEDUG	539	2	60	37	ORANG	2022	
5 PABUARAN	PABUARAN	294	1	67	26	ORANG	2022	
6 LOSARI	LOSARI	313	2	57	15	ORANG	2022	
7 LOSARI	ASTANALANGGAR	113	1	50	10	ORANG	2022	
8 PABEDILAN	PABEDILAN	239	2	64	34	ORANG	2022	
9 PABEDILAN	KALIMUKTI	273	1	55	35	ORANG	2022	
...	...	...	...	...	...	...	...	
12 KALIWEDI	KALIWEDI	456	1	89	46	ORANG	2022	

Gambar 5. Data Stunting Berdasarkan Indikatornya di Kab. Cirebon tahun 2022

4.1.3. Transformasi Data

Dalam melakukan standarisasi data, yang mana format yang akan digunakan dalam implementasi Algoritma Partitioning Around Medoid adalah tidak memasukkan kolom tahun dan satuan serta nama kecamatan kemendagri, dikarenakan hanya menggunakan rentang waktu satu tahun, dengan satuan Orang untuk setiap jumlah variabelnya dan nama kecamatan atau kemendagri tidak disertai karena kurang efektivitasnya pengumpulan informasi jika disertakan dalam proses pengujiannya. Hal ini dilakukan menggunakan RapidMinner.

4.1.4. Data Mining

Pada tahapan data mining dilakukan penghitungan menggunakan algoritma *Partitioning Around Medoid* atau K-Medoid menggunakan Rapid Minner, namun sebelum melakukan tahapan itu, dilakukan pencarian kluster dengan simulasi menggunakan Rapid Minner, yang mana hasil dari *cluster* yang paling optimal adalah 2 *Cluster*, seperti pada gambar 6.



Gambar 6. Hasil Optimasi Cluster

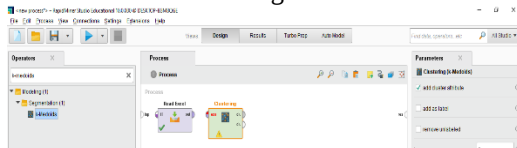
Kemudian dilakukan perhitungan menggunakan rumus eulidian distance, dilanjutkan dengan menghitung simpangan, untuk mengetahui jarak kedekatan atau iterasi.

Tabel 1.1 Hasil dari Simpangan

Rumus = Jumlah kedekatan iterasi 2 - jumlah kedekatan iterasi 1	
Selisih kedekatan	7122,045634 - 6348,966306 = 773,0793279
Jika iterasi > 0, maka iterasi berhenti pada <i>cluster</i> iterasi sebelumnya.	

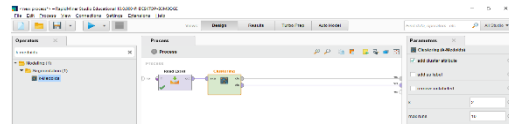
Setelah diketahui nilai simpangannya, diketahui juga hasil cluster setiap puskesmas dalam pengimplementasian ke dalam Rapid Minner dengan Langkah-langkah sebagai berikut :

1. Pada bagian Operators ketika K-Medoid kemudian klik atau drag



Gambar 7. Implementasi Data Mining

2. Kemudian hubungkan antara dataset stunting dan *clustering* lalu hubungkan dengan connectionnya, klik run.



Gambar 8. Connecting Data

3. Setelah itu tentukan k dari data n, yang mana K=2 sesuai dengan nilai *Average Within Centroid* atau pengujian yang dilakukan sebelumnya.
4. Setelah itu diketahui bahwa *cluster* 1 sebanyak 19 item dan *cluster* 2 sebanyak 41 item dari dengan total 60 item.
- 5.



Gambar 9. Hasil Cluster

4.1.5. Evaluation and Interpretation

Untuk melakukan evaluasi dalam menentukan *cluster* terbaik, dilakukan pengujian data stunting menggunakan tools rapid Minner dan atau dengan perhitungan sebagai berikut :

- a. Menghitung Sum of square Within *Cluster* untuk mengetahui kohesi dalam satu kluster [14], rumusnya :

$$SSW_i = \frac{1}{m_i} \cdot \sum_{j=1}^{m_i} d(x_j \cdot c_i)$$

- M_i = Jumlah data dalam kluster ke 1
 X = data dalam kluster
 $D(x,c)$ = Jarak data centroid
 X_j = Data pada kluster tersebut
 C_i = Centroid *cluster* ke-i

Setelah dihitung diketahui SSW dari dataset ini *Cluster* 1 = 73,0401 dan *Cluster* 2 = 122,2041.

- b. Sum of Square Between *Cluster*, untuk mengetahui separasi antar *cluster*, dengan rumus :

$$SSB_{ij} = d(C_i C_j)$$

- C_i = Clutser 1
 C_j = Cluster 2
 $d(C_i.C_j)$ = Jarak antar centroid
 $SSB_{ij} = ((384,85-154,525)^2 + (1,35-1,125)^2 + (67,7-64,425)^2 + (32,02-31,85)^2)$
 $= 230,3485$

Jadi jumlah SSB data ini adalah 230,3485

- c. Menentukan rasio, guna mengetahui seberapa bagus nilai perbandingan antar *cluster* dengan *cluster* lainnya, dengan rumus:

$$R_{ij} = \frac{SSW_i + SSW_j}{SSB_{ij}}$$

- R_{ij} = Rasio antar *cluster*
 SSW_i = Kluster 1
 SSW_j = Kluster 2

$$\begin{aligned} \text{SSBij} &= \text{Separasi dari kluster 1 dan 2} \\ R_{ij} &= 73,0401 + 122,2041 / 230,3485 \\ &= 0,8470604 \end{aligned}$$

Jadi, nilai rasionya adalah 0,8470604.

d. Menghitung *DBI* (Davies Bouldin Index)

$$DBI = \frac{1}{K} \sum_{i=1}^K \max_{j \neq i} (R_{ij})$$

K = Kluster yang ada

R_{i,j} = Rasio antar kluster i dan j

Max = dicari rasio antar kluster yang terbesar

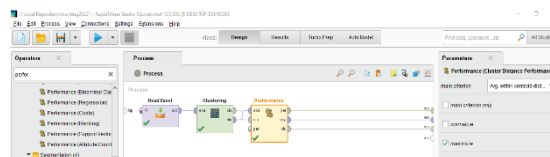
Setelah melakukan perhitungan secara manual dilakukan evaluasi cluster *DBI* (Davies Bouldin Index) menggunakan rapidMiner. Namun untuk mengetahui nilai *DBI* terbaik bagi dataset Stunting di Kabupaten Cirebon berdasarkan indikatornya, dilakukan pengujian secara internal dengan menggunakan rapid miner, dari cluster 2 – 9, dengan hasil seperti tabel 2.

Tabel 1.2 Hasil Pengujian DBI Kluster 2-9

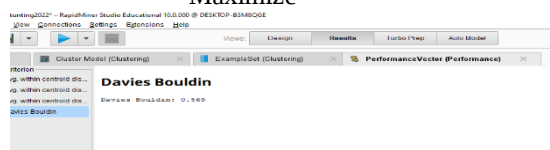
Davies Bouldin Index	
Cluster	Nilai DBI
2	0.569
3	0.707
4	0.890
5	0.900
6	1.019
7	0.990
8	1.170
9	1.159

Dan untuk mengetahui berapa nilai dari *DBI*-nya, yang mana hal itu dapat dilakukan dengan langkah-langkah sebagai berikut :

1. Setelah dilakukan pengclusteran data stunting, pada kolom Operators ketik “Performance” dan klik *Cluster Distance Performance* kemudian drag ke process.
2. Kemudian hubungkan *clustering* ke performance secara menyilang, dan hubungkan performance ke connector, kemudian ceklis maximize agar nilai *DBI* yang keluar tidak minus.
3. Kemudian klik run dan diketahui nilai *Davies Bouldin Index* ini adalah 0,569. Adalah non-negative atau lebih dari sama dengan 0, yang artinya bahwa cluster ini sudah baik karena nilainya mendekati 0.



Gambar 10. Pengujian DBI Connecting dan Maximize



Gambar 11. Hasil Pengujian DBI

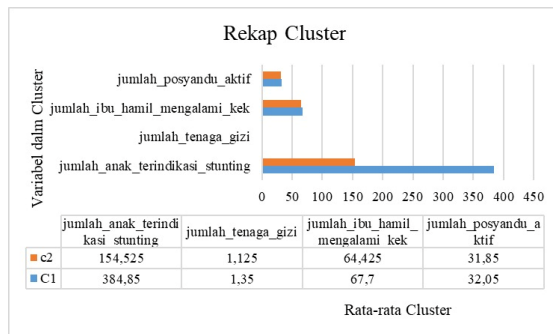
4.2. PEMBAHASAN

Adapun anggota dari setiap puskesmas yang telah tercluster tersebut dapat dilihat pada tabel.

Tabel 1.3 Anggota Cluster

Cluster 1	Cluster 2
WALED	CIBOGO
CILEDUG	PASALEMAN
LOSARI	PABUARAN
GEBANG	ASTANALANGGAR
SUSUKANLEBAK	PABEDILAN
SEDONG	KALIMUKTI
SIDAMULYA	BABAKAN
PANGENAN	GEMBONGAN
MUNDU	KALIMARO
TALUN	KARANGSEMBUNG
SUMBER	KUBANGDELEG
WATUBELAH	SINDANGLAUT
LURAH	ASTANAJAPURA
PLERED	PAMENGGANG
GUNUNGJATI	BEBER
SURANENGGALA	KAMARANG
TEGAL GUBUG	NANGGELA
GESEK	CIPERNA
KALIWEDI	SENDANG
	DUKUPUNTANG
	SINDANGJAWA
	PALIMANAN
	KEPUH
	PLUMBON
	WARUROYOM
	KARANGSARI
	PANGKALAN
	TENGAHTANI
	KEDAWUNG
	MAYUNG
	KEDATON
	KLANGENAN
	BANGODUA
	JAMBLANG
	PANGURAGAN
	CIWARINGIN
	WINONG
	GEMPOL
	SUSUKAN
	BUNDER
	JAGAPURA

Selain itu diketahui juga dalam segmentasi puskesmas di kabupaten Cirebon berdasarkan indikator stunting bahwa rata-rata *clusternya* dapat dilihat dalam gambar 4.12.



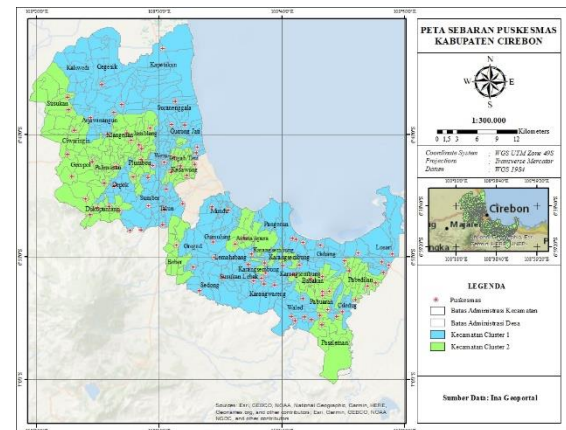
Gambar 12. Rekap Rata-rata Cluster

Dari hasil rekap analisis tersebut dapat dijelaskan makna dari setiap variabel dan *clusternya* yang kemudian mengarah pada pembentukan pernyataan-pernyataan sebagai berikut :

- Jumlah anak yang mengalami stunting**
 Dalam Clutser 1 memiliki jumlah anak yang mengalami stunting lebih tinggi (384,85) dibandingkan dengan Clutser 2 (154,525). Ini menunjukkan bahwa C1 memiliki masalah gizi atau kesehatan anak yang lebih signifikan.
- Jumlah tenaga gizi**
 Kelompok *Cluster* 1 memiliki jumlah tenaga gizi yang lebih banyak (1,35) dibandingkan dengan Clutser 2 (1,125). Hal ini menunjukkan bahwa Clutser 1 memiliki lebih banyak sumber daya untuk mendukung upaya gizi.
- Jumlah ibu hamil yang Mengalam KEK (Kekurangan Energi Kronik)**
 Clutser 1 memiliki jumlah ibu hamil yang mengalami KEK (67,7) lebih tinggi dibandingkan dengan Clutser 2 (64,425). Hal ini mengindikasikan bahwa *Cluster* 1 memiliki masalah gizi ibu hamil yang lebih berat.
- Jumlah Posyandu Aktif**
 Kelompok Clutser 1 memiliki jumlah posyandu aktif yang lebih tinggi (32,05) dibandingkan dengan *Cluster* 2 (31,85). Hal ini menunjukkan bahwa *Cluster* 1 memiliki lebih banyak fasilitas kesehatan atau dukungan kesehatan.

Dari pernyataan-pernyataan berikut dapat disimpulkan sekaligus dijadikan pertimbangan untuk masyarakat dan pemerintah setempat bahwa Clutser 1 cenderung memiliki jumlah anak yang mengalami stunting, jumlah tenaga gizi, jumlah ibu hamil yang mengalami KEK, dan jumlah posyandu aktif yang lebih tinggi dibandingkan dengan C2. Ini bisa mengindikasikan bahwa Clutser 1 memiliki tantangan kesehatan yang lebih besar atau lebih difokuskan untuk melakukan upaya kesehatan masyarakat, salah satunya adalah penyuluhan terhadap stunting.

Dari semua hasil analisis tersebut dapat dilihat pada gambar 13 untuk mengetahui secara jelas pemetaan terhadap penyebaran jumlah tempat-tempat kesehatan yang ada di kabupaten Cirebon pada setiap Kecamatan tahun 2022.



Gambar 13. Sebaran Puskesmas Di Kab. Cirebon

Pemetaan ini dibuat menggunakan aplikasi ArcGIS, di mana data jumlah cluster yang diperoleh diimplementasikan dalam distribusi jumlah puskesmas di Kabupaten Cirebon. Pada pemetaan ini, karakteristik cluster ditandai dengan warna biru sebagai representasi cluster 1, sementara warna hijau digunakan untuk memvisualisasikan cluster 2. Pendekatan ini membantu memberikan gambaran spasial yang jelas terhadap distribusi puskesmas berdasarkan cluster-cluster yang diidentifikasi melalui analisis data menggunakan ArcGIS.

5. KESIMPULAN DAN SARAN

Dalam penelitian ini, ditemukan bahwa segmentasi stunting memerlukan data stunting berdasarkan indikatornya, yang mana data tersebut di dapat dari observasi ke Dinas Kesehatan dan melakukan teknik netnografi dalam website Open Data Jabar dan Open Data Kab. Cirebon dengan data yang diambil adalah data tahun 2022. Hasil segmentasi menggunakan algoritma Partitioning Around Medoid menghasilkan dua cluster. Evaluasi puskesmas menunjukkan nilai yang baik, dengan Cluster 1 (19 puskesmas) memiliki kasus stunting tertinggi, memerlukan perhatian intensif seperti penambahan tenaga gizi. Rekomendasi termasuk perlunya penelitian lanjutan dengan beberapa algoritma dan implementasi hasil penelitian untuk kontribusi dalam penanggulangan kasus stunting, terutama pada puskesmas yang memprioritaskan atau mengalami kasus tertinggi..

DAFTAR PUSTAKA

- [1] Himawat Aryadinata, "Bagaimana Cara Membuat Literature Review yang Baik?," <https://www.dictio.id/>.
- [2] A. Nugraha, O. Nurdian, and G. Dwilestari, "Penerapan Data Mining Metode K-Means Clustering Untuk Analisa Penjualan Pada Toko Yana Sport," 2022.
- [3] D. D. Abdurrahman, F. Agus, and G. M. Putra, "Implementasi Algoritma Partitioning Around Medoids (PAM) untuk Mengelompokkan Hasil Produksi Komoditi Perkebunan (Studi Kasus:

- Dinas Perkebunan Provinsi Kalimantan Timur),” *Informatika Mulawarman : Jurnal Ilmiah Ilmu Komputer*, vol. 16, no. 2, p. 130, Oct. 2021, doi: 10.30872/jim.v16i2.6520.
- [4] I. Fajriyah, Y. B. Kependudukan, D. Keluarga, and B. Nasional, “Segmentasi Wilayah Untuk Menekan Stunting Melalui Program 1000 Hari Pertama Kehidupan (HPK),” *Jurnal Keluarga Berencana*, vol. 5, no. 01, pp. 35–46, 2020, doi: 10.37306/kkb.v4i2.
- [5] A. Meiriza, E. Lestari, P. Putra, N. R. Oktadini, and M. Audya, “Pemanfaatan Data Mining Dalam Penentuan Penyuluhan Penyakit Stunting Menggunakan Partitioning Around Medoids (PAM),” *TeIka*, vol. 12, no. 02, pp. 89–96, 2022, doi: 10.36342/teika.v12i02.2935.
- [6] W. Budiaji *et al.*, “Medoid-based Clustering pada Kecamatan di Kabupaten Lebak dan Pandeglang Provinsi Banten Berdasarkan Trilogi Ketahanan Pangan,” 2022. [Online]. Available: www.unipasby.ac.id
- [7] T. Bimantoro and A. K. Wardhani, “Implementasi Algoritma Partitioning Around Medoids Dalam Pengelompokan Restoran,” *Indonesian Journal of Technology, Informatics and Science (IJTIS)*, vol. 2, no. 1, pp. 33–36, Dec. 2020, doi: 10.24176/ijtis.v2i1.5651.
- [8] M. Orisa and A. Faisol, “Analisis Algoritma Partitioning Around Medoid untuk Penentuan Klasterisasi,” 2021. [Online]. Available: <https://doi.org/10.25047/jtit.v8i2.258>
- [9] Y. Mardi, D. Akademi Perekam dan Informasi Kesehatan Iris Padang Jl Gajah Mada No, and S. Barat, “Jurnal Edik Informatika Data Mining : Klasifikasi Menggunakan Algoritma C4.5,” *Jurnal Edik Informatika*, 2019.
- [10] S. Harini and R. Kusumawati, “Evaluasi Clustering K-Means Dan K-Medoid Pada Persebaran Covid-19 Di Indonesia Dengan Metode Davies-Bouldin Index (DBI),” 2023. [Online]. Available: <https://www.kaggle.com/datasets/hendratno/covid1>
- [11] S. A. J Muntu, J. M. Pingkan Tangkudung, and L. J. H Lotulung, “Studi Netnografi Pada Media Sosial Instagram,” 2021.
- [12] A. ' Ni *et al.*, “Algoritme Partitioning Around Medoid (PAM) Dengan Calinski-Harabasz Index Untuk Clustering Data Outlier (Partitioning Around Medoid (PAM) Algorithm with Calinski-Harabasz Index for Clustering Data Outlier),” 2022. [Online]. Available: <https://magestic.unej.ac.id/>
- [13] S. Ika Murpratiwi, I. Gusti Agung Indrawan, and A. Aranta, “Analisis Pemilihan Cluster Optimal Dalam Segmentasi Pelanggan Toko Retail,” *Jurnal Pendidikan Teknologi dan Kejuruan*, vol. 18, no. 2, 2021.
- [14] M. Wahyudi and L. Pujiastuti, “Komparasi K-Means Clustering dan K-Medoids dalam Mengelompokkan Produksi Susu Segar di Indonesia Comparison of K-Means Clustering and K-Medoids in Clustering Fresh Milk Production in Indonesia,” *Jurnal Bumigora Information Technology (BITe)*, vol. 4, no. 2, pp. 243–254, 2022, doi: 10.30812/bite.v4i2.2104.