

Clustering Status Pemberian Imunisasi Dasar Di Dinas Kesehatan Kabupaten Bireuen Menggunakan Metode *K-Medoids*

Nina Ulfauza^{1*}, Zara Yunizar², dan Fajriana³

^{1,2,3} Program Studi Teknik Informatika, Universitas Malikussaleh

Jl. Kampus Unimal Bukit Indah, Blang Pulo, Kec. Muara Satu, Kota Lhokseumawe, Aceh 24355, Indonesia

*Corresponding author. Email: nina.190170091@mhs.unimal.ac.id

Abstrak— Imunisasi dasar merupakan salah satu upaya dalam mencegah penyakit menular pada anak-anak. Penelitian ini berfokus pada analisis status pemberian imunisasi di Dinas Kesehatan Kabupaten Bireuen. Metode yang digunakan adalah *K-Medoids* dengan data imunisasi dasar anak usia 0-5 tahun dari tahun 2020 hingga 2022. Hasilnya mengidentifikasi tiga *cluster*: Selesai, Belum Selesai, dan Tidak Selesai. Aplikasi berbasis *web* dirancang menggunakan *DFD*, *JavaScript*, *Python*, dan *MySQL*. Dari hasil penelitian, terlihat perbedaan status pemberian imunisasi dasar di Kabupaten Bireuen pada tahun-tahun tersebut. Informasi ini dapat menjadi dasar untuk merancang strategi peningkatan cakupan imunisasi dasar di wilayah tersebut. Kesimpulan dari penelitian ini memberikan pandangan yang jelas tentang distribusi dan ketersediaan imunisasi dasar, yang dapat membantu Dinas Kesehatan dalam mengoptimalkan upaya pencegahan penyakit melalui imunisasi.

Kata kunci: Imunisasi dasar, *k-medoids*, status pemberian imunisasi, *clustering*.

I. PENDAHULUAN

Data mining merupakan suatu proses yang melibatkan penggunaan teknik statistik, matematika, kecerdasan buatan, dan *machine learning* untuk mengidentifikasi informasi yang bernilai serta pengetahuan terkait dari berbagai database besar. Proses ini melibatkan penggunaan teknik tertentu untuk menganalisis dan menggali pola atau hubungan yang mungkin tidak terlihat secara langsung. Tujuannya adalah untuk memperoleh pemahaman yang berharga dan mendalam dari dataset yang kompleks dan luas [1].

Keadaan imunisasi menjadi penunjuk penting untuk mengevaluasi Tingkat status kesehatan masyarakat, khususnya di negara-negara berkembang seperti Indonesia. Pencapaian status imunisasi yang optimal dapat membantu mengurangi angka kematian dan mencegah penyebaran penyakit menular yang berpotensi berakibat fatal. Dinas Kesehatan Kabupaten atau Kota memiliki tanggung jawab dalam melaksanakan kewajiban untuk mendukung puskesmas dalam pengelolaan penugasan tenaga-tenaga yang tersedia. Dengan demikian, Dinas Kesehatan memiliki peran strategis dalam menyelenggarakan dan meningkatkan layanan kesehatan masyarakat di tingkat kabupaten atau kota [2].

Pemanfaatan teknologi informasi pada dinas kesehatan terutama di kabupaten Bireuen ini masih belum efektif, terutama dalam mengelompokkan data-data tahunan dalam pemberian imunisasi kepada balita. Selama ini pengelompokkan data masih dilakukan secara manual. Pencatatan pemberian imunisasi umumnya dilakukan berdasarkan laporan yang diterima di semua puskesmas di Kabupaten Bireuen. Oleh karena itu, menjadi permasalahan apabila perlu mengelompokkan data

pemberian imunisasi. Proses ini melibatkan pembukaan buku besar untuk mencatat dan menghitung kembali pemberian imunisasi pada balita secara manual, yang menyebabkan waktu yang cukup lama dan mengakibatkan penyampaian laporan menjadi kurang efektif dan efisien. Untuk tujuan tersebut, diperlukan pengembangan suatu sistem yang dapat mengkategorikan status pemberian imunisasi dasar dengan mudah. Dalam penelitian ini, metode yang dipakai adalah *k-medoids*. *K-medoids* merupakan perhitungan yang dapat dilakukan dalam sistem *clustering* yang memanfaatkan benda-benda sebagai gambaran (*medoids*) komunitas *cluster* pada setiap kelompok.

II. METODE PENELITIAN

A. Data Mining

Data mining adalah serangkaian proses yang bertujuan untuk menggali informasi yang sebelumnya tidak diketahui secara manual dari suatu basis data. Salah satu tujuan utama *data mining* adalah melakukan pengelompokan data ke dalam satu atau lebih *cluster*, di mana objek-objek dalam satu cluster memiliki kesamaan yang tinggi [3].

Secara umum, data mining dapat dibedakan menjadi dua kategori utama: deskriptif dan prediktif. Pada *data mining* deskriptif, fokusnya adalah untuk menemukan pola-pola yang dapat dipahami oleh manusia dan menjelaskan karakteristik dari suatu set data. Sementara itu, pada *data mining* prediktif, tujuannya adalah untuk membangun model pengetahuan yang dapat digunakan untuk melakukan prediksi atau estimasi terhadap data yang belum diketahui [4].

B. Clustering

Clustering atau pengelompokan adalah metode yang digunakan untuk membagi rangkaian data menjadi beberapa kelompok berdasarkan kesamaan yang telah ditentukan sebelumnya. Tujuan dari clustering adalah untuk mengoptimalkan nilai dari objektif function yang diset dalam proses clustering, yang pada umumnya berusaha meminimalisasikan variasi antar cluster, serta dapat meminimalisasikan variasi antar cluster. Pengelompokan data adalah untuk mengoptimalkan nilai dari fungsi tujuan yang ditetapkan dalam proses clustering, yang umumnya berusaha meminimalisasikan variasi antar kelompok dan dapat mengurangi variasi antar kelompok [5].

C. Algoritma K-medoids

Algoritma k-medoids adalah salah satu algoritma mengelompokkan objek-objek dalam data ke dalam beberapa cluster. Algoritma ini merupakan metode partisional, yang berarti bahwa objek-objek dalam data dibagi menjadi beberapa cluster secara bertahap. Perbedaan antara algoritma K-Means dan K-Medoids terletak pada cara penentuan pusat cluster. Dalam K-Means, pusat kluster dihitung sebagai nilai rata-rata dari semua objek data yang terdapat dalam kluster tersebut. Sebaliknya, dalam K-Medoids, pusat cluster direpresentasikan oleh satu objek data yang berfungsi sebagai medoids, yang dipilih karena memiliki kesamaan tertinggi dengan objek-objek lain dalam cluster tersebut. Perbedaan ini dapat mempengaruhi hasil pengelompokan, terutama dalam menghadapi outlier atau objek data yang sangat berbeda di dalam suatu cluster [6]. Algoritma K-Medoid memiliki langkah-langkah seperti berikut [7]:

- Menentukan jumlah cluster (k)
- Menginisialisasi medoid awal
- Mengtribusikan objek ke cluster
- Menghitung jarak rata-rata antar objek dalam cluster

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

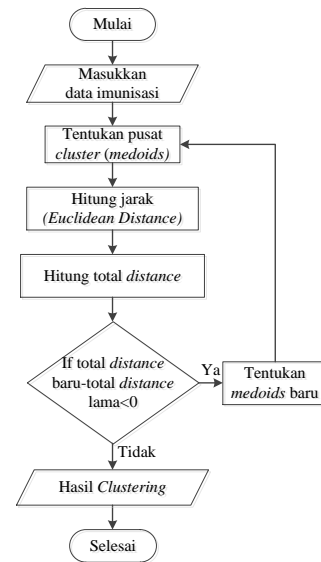
Keterangan;

x_i = Kriteria data

μ_j = Pusat (centroid) pada kluster ke-js.

- Menghitung total simpangan (S) dengan mengurangi nilai total jarak baru dari total jarak sebelumnya. Jika $S < 0$, melakukan pertukaran objek dengan kluster untuk membentuk kelompok k objek baru sebagai medoids
- Melakukan perhitungan langkah 3 hingga 5 secara berulang sampai tidak ada perubahan pada medoids, sehingga diperoleh kluster beserta anggota dari setiap kluster.

D. Skema K-Medoid



Gambar 1. Skema K-medoids

Tahapan dalam membangun sebuah aplikasi berbasis web:

- 1) Mulai, merupakan awal tampilan beranda aplikasi.
- 2) Memasukkan data imunisasi, pada tahap ini admin memasukkan data imunisasi perpuskesmas, ada 5 imunisasi yaitu HB0, BCG, polio1, DPT-HB-HIB1, Campak.
- 3) Inisialisasi pusat cluster (medoids) sebanyak nilai k=3, dipilih secara acak.
- 4) Menghitung jarak menggunakan rumus Euclidean Distance. Pada tahap ini sistem akan melakukan perhitungan data imunisasi kedalam tiga kelompok sesuai dengan nilai K, yaitu selesai, belum selesai dan tidak selesai.
- 5) Jika total distance baru-total distance lama < 0 maka akan menghitung kembali total distance dengan cara menentukan medoids baru hingga selisih lebih dari 0, atau mengulangi langkah 3-5 hingga tidak terjadi lagi perubahan medoids.
- 6) Hasil clustering data imunisasi jika total distance > 0.
- 7) Selesai, merupakan pemberhentian dari sistem.

III. HASIL DAN PEMBAHASAN

A. Dataset Penelitian

Penetapan parameter data yang diolah mencakup variabel-variabel seperti nama puskesmas, HB0, BCG, Polio1, DPT-HB-HIB1, dan campak. Detail dataset dapat dilihat pada tabel 1.

Tabel. 1 Data Rekapitan Laporan Imunisasi Tahun 2020

No.	Nama PKM	HBO	BCG	POLIO1	DPT-HB-HIB1	CAMPAK
1	Samalanga	53	34	37	40	19
2	Simpang Mamplam	80	34	49	39	22
3	Pandrah	54	19	48	11	10
.
.
.
19	Gandapura	37	11	42	12	5
20	Mon Keulayu	12	26	45	20	8

B. Normalisasi Data

Normalisasi data adalah mengatasi ketidakseimbangan rentang nilai antar variabel, sehingga setiap variabel dapat memberikan dampak yang seimbang pada analisis data. Normalisasi data dapat dilihat pada tabel 2.

Tabel 2. Normalisasi Data

No	NAMA PKM	HBO	BCG	POLIO1	DPT-HB-HIB1	CAMPAK
1	Samalanga	0,406	0,535	0,358	0,603	0,400
2	Simpang Mamplam	0,673	0,535	0,585	0,586	0,475
3	Pandrah	0,416	0,186	0,566	0,103	0,175
.
.
.
19	Gandapura	0,248	0,000	0,453	0,121	0,050
20	Mon Keulayu	0,000	0,349	0,509	0,259	0,125

C. Pengujian data

Tahap awal adalah menginisialisasi pusat kluster dengan membentuk 3 kluster dari sampel data, dan setiap *medoid* dipilih secara acak.

Tabel 3. Nilai Pusat Kluster Awal Imunisasi Tahun 2020

No	Nama PKM	HBO	BCG	POLIO1	DPT-HB-HIB1	CAMPAK
18	Makmur	0,673	0,605	0,623	0,621	0,100
19	Gandapura	0,248	0,000	0,453	0,121	0,050
20	Mon Keulayu	0,000	0,349	0,509	0,259	0,125

Langkah kedua mencakup perhitungan jarak terdekat dengan menggunakan rumus *Euclidean Distance* untuk melakukan klusterisasi pada setiap data yang sudah diperoleh.

Menghitung jarak data ke pusat kluster pertama:

$$d_{(18,1)} = \sqrt{(0,673 - 0,46)^2 + (0,605 - 0,535)^2 + (0,623 - 0,358)^2 + (0,621 - 0,603)^2 + (0,100 - 0,400)^2} = 0,486$$

Menghitung jarak data ke pusat kluster kedua:

$$d_{(19,1)} = \sqrt{(0,248 - 0,406)^2 + (0,000 - 0,535)^2 + (0,453 - 0,358)^2 + (0,121 - 0,603)^2 + (0,050 - 0,400)^2} = 0,822$$

Menghitung jarak data ke pusat kluster ketiga:

$$d_{(20,1)} = \sqrt{(0,000 - 0,406)^2 + (0,349 - 0,535)^2 + (0,509 - 0,358)^2 + (0,259 - 0,603)^2 + (0,125 - 0,400)^2} = 0,646$$

Proses perhitungan yang serupa akan diaplikasikan pada seluruh data. Setelah melakukan perhitungan pada seluruh data dan atribut, akan diperoleh jarak terdekat dari setiap data dalam masing-masing setiap kluster, sebagaimana terlihat dalam tabel 4.

Tabel 4. Hasil Perhitungan *Algoritma K-Medoid* ke-1

No	C1	C2	C3	Kedekatan	Cluster
1	0,486	0,822	0,646	0,486	1
2	0,385	0,939	0,851	0,385	1
3	0,720	0,303	0,479	0,303	2
.
.
.
19	0,910	0,000	0,459	0,000	2
20	0,814	0,459	0,000	0,000	3

Langkah berikutnya adalah Menetapkan *medoids* baru dilakukan dengan cara memilihnya secara acak, dengan syarat bahwa setiap *medoids* yang sudah terpilih sebelumnya tidak dapat dijadikan *medoids* baru (*Non Medoids*). Nilai *medoids* baru dapat ditemukan dalam Tabel 5.

Tabel 5. Nilai *Medoids* Baru

No	Nama PKM	HBO	BCG	POLIO1	DPT-HB-HIB1	CAMPAK
2	Simpang Mamplam	0,673	0,535	0,585	0,586	0,475
3	Pandrah	0,416	0,186	0,566	0,103	0,175
4	Jeunieb	0,495	0,163	0,000	0,224	0,150

Hitung kembali jarak dari setiap objek pada iterasi kedua dengan menggunakan *medoids* baru pada tabel 5.

Menghitung jarak data dengan pusat kluster pertama dilakukan dalam proses perhitungan pada analisis kluster.

$$d_{(2,1)} = \sqrt{(0,406 - 0,673)^2 + (0,535 - 0,535)^2 + (0,358 - 0,585)^2 + (0,603 - 0,586)^2 + (0,400 - 0,475)^2} = 0,359$$

Menghitung jarak data dengan pusat kluster kedua merupakan bagian dari proses perhitungan pada analisis kluster.

$$d_{(3,1)} = \sqrt{(0,416 - 0,406)^2 + (0,186 - 0,535)^2 + (0,566 - 0,358)^2 + (0,603 - 0,103)^2 + (0,400 - 0,175)^2} = 0,682$$

Menghitung jarak data dengan pusat kluster ketiga adalah salah satu tahapan dalam perhitungan kluster untuk menentukan kemiripan atau jarak antar data dengan *centroid* dari kluster ketiga.

$$d_{(4,1)} = \sqrt{(0,406 - 0,495)^2 + (0,535 - 0,163)^2 + (0,358 - 0,000)^2 + (0,603 - 0,224)^2 + (0,400 - 0,150)^2} = 0,694$$

Proses perhitungan yang serupa akan terus diterapkan pada seluruh data, sebagaimana tunjukkan pada tabel 6.

Tabel 6. Hasil Perhitungan *Algoritma K-Medoids* Iterasi ke-2

No	C1	C2	C3	Kedekatan	Cluster
1	0,359	0,682	0,694	0,359	1
2	0,000	0,715	0,865	0,000	1
3	0,715	0,000	0,585	0,000	2
.
.
19	0,939	0,303	0,560	0,303	2
20	0,851	0,479	0,736	0,479	2

Setelah memperoleh nilai jarak antara iterasi pertama dan iterasi kedua, total simpangan (S) dihitung dengan mengurangkan nilai total cost baru dari nilai total cost lama. Jika $S < 0$, langkah selanjutnya adalah melakukan pertukaran nilai objek dengan menetapkan *medoids* baru.

$$\begin{aligned}
 S &= \text{total cost baru} - \text{total cost lama} \\
 &= 7,984 - 9,662 \\
 &= -1,679
 \end{aligned}$$

Dikarenakan total simpangan menghasilkan nilai di bawah nol ($S < 0$), dilakukan pertukaran nilai objek dengan menetapkan *medoids* baru. Selanjutnya, langkah berikutnya adalah menetapkan nilai *medoids* baru lagi dengan cara memilihnya secara acak atau acak, dengan catatan bahwa setiap *medoids* yang telah terpilih tidak dapat dijadikan *medoids* baru (Non *Medoids*).

Tabel 7. *Medoids* Baru Lagi

No	Nama PKM	HBO	BCG	POLIO1	DPT-HB-HIB1	CAMPAK
5	Peulimbang	1,000	0,000	0,528	0,000	0,000
7	Jeumpa	0,594	0,279	0,113	0,259	0,050
9	Kuala	0,554	0,512	0,415	0,448	0,100

Menghitung ulang jarak dari setiap objek pada iterasi ketiga dengan menggunakan *medoids* baru. Setelah mendapatkan nilai jarak antara iterasi kedua dan iterasi ketiga, hitung total simpangan (S) dengan mengurangkan nilai total cost baru dari nilai total cost lama. Selanjutnya, dengan ketentuan bahwa jika $S < 0$, dilakukan pertukaran nilai objek dengan menentukan *medoids* baru.

$$\begin{aligned}
 S &= \text{total cost baru} - \text{total cost lama} \\
 &= 9,770 - 7,984 \\
 &= 1,786
 \end{aligned}$$

Dikarenakan total simpangan menghasilkan nilai di atas nol ($S > 0$), pengujian ini diakhiri pada iterasi ketiga. Hasil dari iterasi terakhir akan digunakan sebagai parameter dalam proses pengelompokan. Selanjutnya, anggota kluster yang akan masuk ke dalam kelompok "selesai", "belum selesai", dan "tidak selesai" akan ditentukan berdasarkan *centroid*, yaitu $C1 = \text{selesai}$, $C2 = \text{belum selesai}$, $C3 = \text{tidak selesai}$. Oleh karena itu, anggota dari $C1$ akan termasuk dalam kluster selesai, anggota dari $C2$ akan termasuk dalam kluster belum selesai, dan anggota $C3$ akan termasuk dalam kluster tidak selesai.

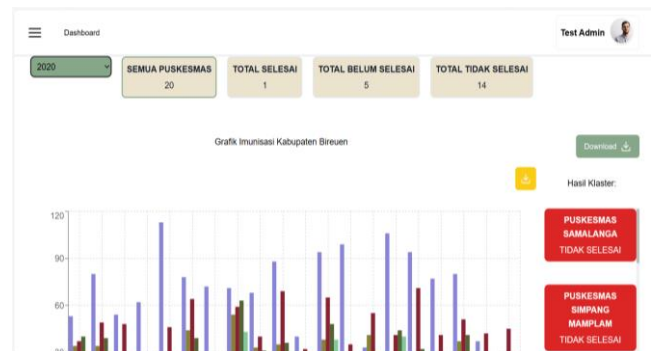
Tabel 8. Hasil Akhir

No	NAMA PKM	Cluster
1	Samalanga	Tidak Selesai

2	Simpang Mamplam	Tidak Selesai
3	Pandrah	Tidak Selesai
4	Jeunieb	Belum Selesai
5	Peulimbang	Selesai
6	Peudada	Tidak Selesai
7	Jeumpa	Belum Selesai
8	Kota Juang	Tidak Selesai
9	Kuala	Tidak Selesai
10	Juli	Tidak Selesai
11	Juli Dua	Tidak Selesai
12	Jangka	Tidak Selesai
13	Peusangan	Belum Selesai
14	Cot Ie Jue	Tidak Selesai
15	Peusangan Selatan	Tidak Selesai
16	Peusangan Siblah Krueng	Tidak Selesai
17	Kuta Blang	Belum Selesai
18	Makmur	Tidak Selesai
19	Gandapura	Belum Selesai
20	Mon Keulayu	Tidak Selesai

D. Penerapan Sistem

Berikut ini adalah tampilan sistem *clustering* Status Pemberian Imunisasi Dasar di Dinas Kesehatan Kabupaten Bireuen Menggunakan Metode *K-Medoids*.



Gambar 2. Tampilan Hasil *K-Medoids* Pada Sistem

IV. KESIMPULAN

Hasil penelitian ini menyimpulkan bahwa *algoritma K-Medoids* dapat diterapkan dengan efektif untuk mengidentifikasi kluster dalam konteks pemberian imunisasi dasar, dengan mengikuti langkah-langkah perhitungan *K-Medoids*. Analisis data mengungkapkan adanya tiga kluster dalam status pemberian imunisasi dasar, yaitu *Cluster Selesai*, *Cluster Belum Selesai*, dan *Cluster Tidak Selesai*, dengan distribusi yang berbeda-beda pada tahun 2020, 2021, dan 2022 di Kabupaten Bireuen. Secara spesifik, pada tahun 2020, terdapat 1 puskesmas dengan kondisi pemberian imunisasi selesai, 5 puskesmas dengan kondisi belum selesai, dan 14 puskesmas dengan kondisi tidak selesai. Pada tahun 2021, terdapat 9 puskesmas dengan kondisi selesai, 6 puskesmas belum selesai, dan 5 puskesmas tidak selesai. Pada tahun 2022, terdapat 11 puskesmas dengan kondisi selesai, 2 puskesmas belum selesai, dan 7 puskesmas tidak selesai. Metode *K-Medoids Clustering* juga menyoroti bahwa pengelompokan pemberian imunisasi dasar dapat dipengaruhi oleh perbedaan *centroid* awal. Hasil penelitian ini dapat menjadi dasar bagi penelitian lanjutan, dengan fokus pada pengembangan untuk meningkatkan cakupan imunisasi dasar.

REFERENSI

- [1] Darmi, Y. D., & Setiawan, A. (2017). Penerapan Metode *Clustering* K-Means Dalam Pengelompokan Penjualan Produk. *Jurnal Media Infotama*, 12(2), 148–157. <https://doi.org/10.37676/jmi.v12i2.418>
- [2] Victor Trismanjaya Hulu, Salman, Supinganto, A., Amalia, L., Khariri, Sianturi, E., Nilasari, Siagian, N., Hastuti, P., & Syamdarniati. (2020). Epidemiologi Penyakit Menular: Riwayat, Penularan dan Pencegahan. In A. Rikki (Ed.), *Paper Knowledge . Toward a Media History of Documents* (1st ed.). Yayasan Kita Menulis.
- [3] Tarigan, P. M. S., Hardinata, J. T., Qurniawan, H., Safii, M., & Winanjaya, R. (2022). Implementasi Data Mining Menggunakan Algoritma Apriori Dalam Menentukan Persediaan Barang. *Jurnal Janitra Informatika Dan Sistem Informasi*, 2(1), 9–19. <https://doi.org/10.25008/janitra.v2i1.142>
- [4] Manurung, E., & Hasugian, P. S. (2019). Data Mining Tingkat Pesanan Inventaris Kantor Menggunakan Algoritma Apriori pada Kepolisian Daerah Sumatera Utara. *Journal Of Informatic Pelita Nusantara*, 4(2), 8–13.
- [5] Muliono, R., & Sembiring, Z. (2019). Data Mining *Clustering* Menggunakan Algoritma K-Means Untuk Klasterisasi Tingkat Tridarma Pengajaran Dosen. *CESS (Journal of Computer Engineering, System and Science)*, 4(2), 2502–2714.
- [6] Kamal Kaur, N., Kaur, U., & Singh, D. (2014). K-Medoid *Clustering* Algorithm-A Review. *International Journal of Computer Application and Technology (IJCAT)*, 1(1). www.technopublications.com
- [7] Fajriana. (2021). Analisis Algoritma K-Medoids pada Sistem Klasterisasi Produksi Perikanan Tangkap Kabupaten Aceh Utara Fajriana. *JEPIN (Jurnal Edukasi Dan Penelitian Informatika) Vol. 7 No. 2 Agustus 2021*, 7(5), 263–269. <https://doi.org/10.52088/ijesty.v2i3.318>