

## Penerapan Algoritma K-Medoids dalam Pengelompokan Balita Stunting di Indonesia

Halimatusakdiah Pohan<sup>1</sup>, Muhammad Zarlis<sup>2</sup>, Eka Irawan<sup>3</sup>, Harly Okprana<sup>4</sup>, Yuegilion Pranayama Purba<sup>5</sup>

<sup>1,2,3,4,5</sup>STIKOM Tunas Bangsa Pematangsiantar, Jl. Jend. Sudirman Blok A No.1-3  
Pematangsiantar, Indonesia  
halimapohan104@gmail.com

**Abstrak.** *Stunting adalah kondisi gagal tumbuh pada balita akibat kekurangan asupan gizi dan infeksi yang berkepanjangan yang mengakibatkan tinggi badan yang lebih pendek dari standar usianya. Indonesia saat ini menjadi urutan ke 4 dalam tingginya kasus prevelensi stunting menurut standar World Health Organization. Adapun tujuan dari penelitian ini untuk mengelompokkan provinsi mana yang mengalami bayi stunting dengan cluster tertinggi maupun cluster terendah yang berguna sebagai masukan bagi pemerintah untuk menangani dengan cepat penurunan stunting di Indonesia. Data yang digunakan dari penelitian ini di dapat dari Badan Pusat Statistika (BPS) dengan nama indikator Prevelensi Stunting tahun 2015-2018. Dalam penelitian ini data di olah dengan menggunakan Algoritma K-Medoids yang merupakan salah satu bagian dari algoritma clustering yang dapat memecahkan dataset ke kelompok-kelompok diantara semua objek data dengan menggunakan objek sebagai perwakilan (medoid) dalam sebuah cluster.*

**Kata Kunci :** *Stunting, Data Mining, K-Medoids.*

**Abstract.** *Stunting is a condition of failure to thrive in toddlers due to lack of nutritional intake and prolonged infection which results in a body height that is shorter than the standard age. Indonesia is currently ranked 4th in the high prevalence of stunting cases according to World Health Organization standards. The purpose of this study is to classify which provinces experience stunting infants with the highest and lowest clusters, which are useful as input for the government to deal with stunting decline in Indonesia quickly. The data used in this study was obtained from the Central Statistics Agency (BPS) under the name of the 2015-2018 Stunting Prevalence indicator. In this study, the data is processed using the K-Medoids Algorithm which is one part of the clustering algorithm that can solve the dataset into groups among all data objects by using the object as a representative (medoid) in a cluster.*

**Keyword :** *Stunting, Data Mining, K-Medoids*

### PENDAHULUAN

Data mining biasanya digunakan untuk proses penambangan data yang besar dari kumpulan fakta yang terekam untuk mengetahui pola yang ada untuk menghasilkan sebuah output yang berupa informasi ataupun pengetahuan. Terdapat beberapa teknik algoritma data mining yang sering dimanfaatkan untuk menemukan pola berupa informasi ataupun pengetahuan yang tersembunyi diantaranya yaitu estimasi, asosiasi, klasifikasi, klusterisasi, dan prediksi[1],[2].

Metode-metode yang terdapat pada data mining untuk Clustering salah satunya yaitu metode K-Medoids. K-Medoids atau Partitioning Around Medoids (PAM) adalah suatu metode pengelompokan pengembangan jenis dari Metode K-Means[3],[4],[5]. Perbedaan dari kedua algoritma ini yaitu algoritma K-Medoids atau PAM menggunakan objek sebagai perwakilan (medoid) sebagai pusat cluster untuk setiap cluster, sedangkan K-Means menggunakan nilai rata-rata (mean) sebagai pusat cluster[6].

Indonesia saat ini menjadi urutan ke 4 dunia dalam tingginya angka prevelensi stunting yang jauh dari target yang sudah ditetapkan oleh World Health Organization (WHO). Stunting merupakan kondisi gagal tumbuh pada balita akibat kekurangan asupan gizi dalam jangka waktu yang lama dan terjadinya infeksi berulang yang mengakibatkan anak mengalami tinggi badan yang lebih pendek dari standar usianya[7]. Stunting berbeda dengan anak yang bertubuh pendek, anak yang bertubuh pendek



mengalami mental dan fisik yang sama dengan anak normal lainnya. Namun faktor genetik dari orang tua yang membuat anak mengalami tubuh pendek dan masih bisa memperbaikinya dengan makanan serta minuman yang bernutrisi dan juga berolahraga, sedangkan penyebab anak mengalami stunting yaitu keterlambatan pertumbuhan fisik serta mental yang disebabkan oleh ketidakcukupan biaya dalam mengontrol kesehatan dan juga asupan gizi selama hamil, penyediaan sarana kesehatan bagi ibu hamil yang terbatas, pemberian ASI yang tidak maksimal, pola pengasuhan yang tidak baik, imunisasi yang tidak lengkap, maupun kurangnya sanitasi dan air bersih.

Dari data yang didapatkan dari Badan Pusat Statistik dengan nama indikator Prevelensi Stunting (Pendek dan Sangat Pendek) pada anak balita pada tahun 2015 – 2018 dimana pada tahun 2018 mengalami kenaikan dari tahun sebelumnya. Pada penelitian ini, metode yang digunakan yaitu K-Medoids. Dimana penulis akan mengelompokkan provinsi mana yang mengalami bayi stunting dengan cluster tertinggi maupun cluster terendah yang berguna sebagai masukan bagi pemerintah untuk menangani dengan cepat penurunan stunting di Indonesia.

Dalam penelitian yang dilakukan oleh [8] mengenai segmentasi pelanggan berdasarkan produk menggunakan metode K-Medoids dan dari penelitian tersebut algoritma K-Medoids dapat diterapkan untuk mengelompokkan pelanggan berdasarkan segmentasi pada produk jumlah pembelian dan area. Dengan data uji sebanyak 6 daerah, 7000 data outlet dan 28 produk. Dapat disimpulkan menggunakan 3 cluster berikut hasilnya C1 menghasilkan 27%, C2 menghasilkan 43% dan C3 menghasilkan 30%.

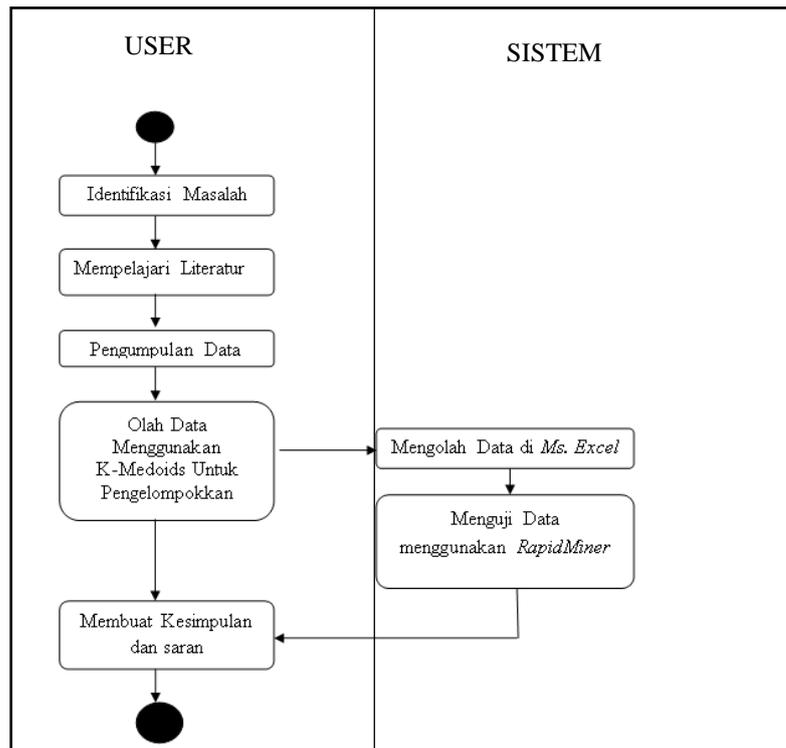
Penelitian selanjutnya yang dilakukan oleh [9] dengan judul “Algoritma K-Medoids untuk menentukan calon mahasiswa yang layak mendapatkan beasiswa bidikmisi di Universitas Budidarma, hasil penelitian ini menunjukkan bahwa algoritma K-Medoids dapat digunakan untuk menentukan calon mahasiswa yang layak mendapatkan beasiswa bidikmisi di Universitas Budi Darma. Proses pengujian yang dilakukan secara manual baik dengan menggunakan aplikasi Rapidminer menghasilkan hasil cluster yang sama.

## METODOLOGI PENELITIAN

Penelitian ini menerapkan data mining dalam mengelompokkan balita yang mengalami stunting untuk mengetahui tingkat cluster balita yang mengalami stunting agar dapat dilakukan penanganan yang ekstra ke provinsi tersebut dengan metode K-Medoids. Pengumpulan data yang digunakan dalam penelitian ini yaitu Penelitian kepustakaan (Library research) yaitu memanfaatkan perpustakaan, buku, prosiding atau jurnal sebagai media untuk bahan referensi dalam menentukan parameter yang digunakan dalam penelitian [10]. Sumber data penelitian diperoleh melalui Badan Pusat Statistik Nasional. Data yang digunakan dalam penelitian ini adalah data prevalensi stunting pada setiap provinsi dari tahun 2015 - 2018 yang terdiri dari 33 provinsi. Variabel yang digunakan adalah prevalensi stunting yang sudah diakumulasikan. Data akan diolah dengan melakukan clustering persentase balita yang mengalami stunting berdasarkan provinsi dalam 2 cluster yakni cluster tertinggi dan terendah

Dari gambar 1 dijelaskan bahwa kegiatan penelitian ini diawali dengan menganalisa masalah yang ada setelah itu menetapkan metode untuk penyelesaian masalah yang sudah di dapat lalu mengumpulkan data, data yang sudah dikumpulkan akan diolah ke Ms.Excel terlebih dahulu apabila data yang diolah tersebut sudah valid kemudian akan diuji dengan menggunakan Rapidminer. Ketika hasil yang didapat pada Rapidminer sudah sesuai dengan hasil yang diolah di excel kemudian dapat menarik sebuah kesimpulan dan saran. . Diagram aktifitas kerja dalam penelitian ini dapat dilihat pada gambar 1 berikut:





**Gambar 1.** Diagram Aktifitas Kerja

Data yang digunakan untuk mendukung terlaksananya penelitian ini ialah data sekunder yang bersumber dari Badan Pusat Statistik (BPS), setelahnya data diolah dengan *Rapidminer* untuk mencari keakuratan data.

**Tabel 1.** Data Balita Stunting Berdasarkan Provinsi

No	PROVINSI	Prevelensi Balita Yang Mengalami Stunting Menurut Provinsi di Indonesia			
		2015	2016	2017	2018
1.	ACEH	31.6	26.4	35.7	37.1
2.	SUMATERA UTARA	33.2	24.44	28.5	32.4
3.	SUMATERA BARAT	27.7	25.54	30.6	29.9
4.	RIAU	23.9	25.07	29.7	27.4
5.	JAMBI	25.9	27	25.2	30.2
6.	SUMATERA SELATAN	23.4	19.24	22.8	31.6
7.	BENGKULU	68.5	22.96	29.4	28
8.	LAMPUNG	22.7	24.78	31.6	27.3
9.	KEP. BANGKA BELITUNG	18.9	21.92	27.3	23.4
10.	KEP. RIAU	22.9	22.85	21	23.6
11.	DKI JAKARTA	23	20.06	22.7	17.6
12.	JAWA BARAT	25.6	25.13	29.2	31.1
13.	JAWA TENGAH	24.8	23.87	28.5	31.3
14.	DI YOGYAKARTA	20.6	21.84	19.8	21.4
15.	JAWA TIMUR	27.1	26.13	26.7	32.8

No	PROVINSI	Prevelensi Balita Yang Mengalami Stunting Menurut Provinsi di Indonesia			
		2015	2016	2017	2018
16.	BANTEN	23.2	26.99	29.6	26.6
17.	BALI	20.6	19.7	19.1	21.9
18.	NUSA TENGGARA BARAT	33.9	29.97	37.2	33.5
19.	NUSA TENGGARA TIMUR	41.2	38.75	40.3	42.7
20.	KALIMANTAN BARAT	34.1	34.93	36.5	33.3
21.	KALIMANTAN TENGAH	33.4	34.11	39	34
22.	KALIMANTAN SELATAN	37.2	31.13	34.2	33.1
23.	KALIMANTAN TIMUR	26.7	27.14	30.6	29.2
24.	KALIMANTAN UTARA	31	31.6	33.4	26.9
25.	SULAWESI UTARA	22.2	21.21	31.4	25.5
26.	SULAWESI TENGAH	35.3	32.04	36.1	32.3
27.	SULAWESI SELATAN	34.1	35.6	34.8	35.7
28.	SULAWESI TENGGARA	31.4	29.57	36.4	28.7
29.	GORONTALO	36.5	33.04	31.7	32.5
30.	SULAWESI BARAT	38.4	39.71	40	41.6
31.	MALUKU	32.3	28.98	30	34
32.	MALUKU UTARA	24.5	24.59	25	31.4
33.	PAPUA BARAT	29.5	30.28	33.3	27.8
34.	PAPUA	28.6	27.99	32.8	33.1

Tabel 1 menunjukkan data yang di dapat dari Badan Pusat Statistik Nasional mengenai data balita stunting yang terdiri dari 34 provinsi dari tahun 2015-2018.

### HASIL DAN PEMBAHASAN

Pengolahan data menggunakan metode K-Medoids yaitu menentukan jumlah cluster (K) dari n objek sebanyak 2 cluster dari data sampel, dengan memilih cluster secara acak untuk menentukan medoid seperti pada tabel 2 berikut :

Tabel 2 Medoids Awal

Nama	Provinsi	2015	2016	2017	2018
C1	Nusa Tenggara Timur	41,2	38,75	40,3	42,7
C2	Gorontalo	36,5	33,04	31,7	32,5

Alokasikan setiap data (objek) ke cluster terdekat menggunakan persamaan ukuran jarak Euclidean Distance. Berikut adalah perhitungan persamaan jarak Euclidean Distance pada data kasus balita yang mengalami stunting :

$$D_{Aceh,C1} = \sqrt{(31,6 - 41,2)^2 + (26,4 - 38,75)^2 + (35,7 - 40,3)^2 + (37,1 - 42,7)^2} = 17,23956$$

$$D_{Aceh,C2} = \sqrt{(31,6 - 36,5)^2 + (26,4 - 33,04)^2 + (35,7 - 31,7)^2 + (37,1 - 32,5)^2} = 10,25961$$

$$D_{Sumut,C1} = \sqrt{(33,2 - 41,2)^2 + (24,44 - 38,75)^2 + (28,5 - 40,3)^2 + (32,4 - 42,7)^2} = 22,67391$$



$$D_{Papua, C2} = \sqrt{(28,6 - 36,5)^2 + (27,99 - 33,04)^2 + (32,8 - 31,7)^2 + (33,1 - 32,5)^2} = 9,45952$$

Untuk hasil perhitungan jarak ke setiap medoid awal bisa dilihat pada tabel 3 berikut ini:

**Tabel 3.** Hasil Perhitungan Iterasi Ke-1

Provinsi	Jarak Ke Medoid		Kedekatan	Cluster Terdekat
	C1	C2		
Aceh	17,23956	10,25961	10,25961	C2
Sumatera Utara	22,67391	9,75192	9,75192	C2
Sumatera Barat	24,79282	11,90210	11,90210	C2
Riau	28,85970	15,88367	15,88367	C2
Jambi	27,50295	14,01362	14,01362	C2
Sumatera Selatan	33,56993	21,02546	21,02546	C2
Bengkulu	36,46250	33,92855	33,92855	C2
Lampung	29,15923	16,90318	16,90318	C2
Kep. Bangka Belitung	36,35972	23,14270	23,14270	C2
Kep. Riau	36,40055	21,96579	21,96579	C2
.....	.....	.....	.....	.....
Papua	20,56569	9,45952	9,45952	C2
<b>JUMLAH</b>	809,99535	467,10045		
<b>TOTAL COST</b>	1277,09580			

Setelah didapatkan hasil jarak dari setiap objek (cost) pada iterasi ke 1 maka lanjut iterasi 2,3 sampai Iterasi n dengan memilih secara acak objek pada masing-masing cluster sebagai kandidat medoid sampai diketahui hasil simpangan sudah  $S > 0$  maka proses perhitungan cluster dihentikan. Hasil cluster yang telah di dapatkan bisa dilihat pada tabel 4 berikut ini :

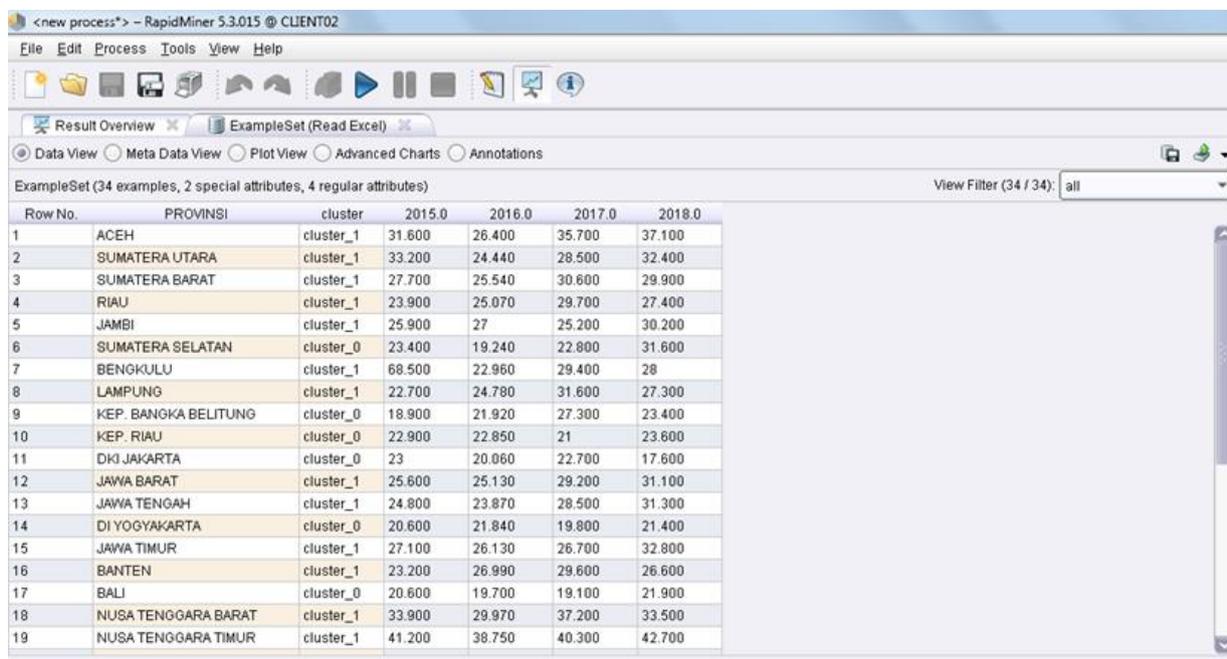
**Tabel 4.** Hasil Cluster

Provinsi	Cluster Terdekat
Aceh	C1
Sumatera Utara	C1
Sumatera Barat	C1
Riau	C1
Jambi	C1
Sumatera Selatan	C2
Bengkulu	C1
Lampung	C1
Kep. Bangka Belitung	C2
Kep. Riau	C2
Dki Jakarta	C2
Jawa Barat	C1
Jawa Tengah	C1
DI Yogyakarta	C2

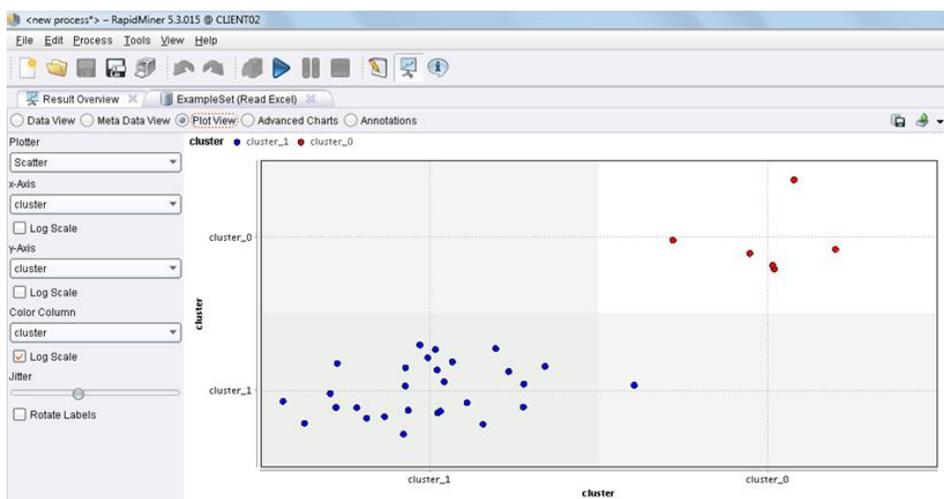


Provinsi	Cluster Terdekat
Jawa Timur	C1
Banten	C1
Bali	C2
Nusa Tenggara Barat	C1
Nusa Tenggara Timur	C1
Kalimantan Barat	C1
Kalimantan Tengah	C1
Kalimantan Selatan	C1
Kalimantan Timur	C1
Kalimantan Utara	C1
Sulawesi Utara	C1
Sulawesi Tengah	C1
Sulawesi Selatan	C1
Sulawesi Tenggara	C1
Gorontalo	C1
Sulawesi Barat	C1
Maluku	C1
Maluku Utara	C1
Papua Barat	C1
Papua	C1

Pengolahan data menggunakan tools rapidminer untuk melakukan pengelompokan dengan algoritma Kmedoids dapat dilihat pada gambar 2 dan gambar 3 berikut:

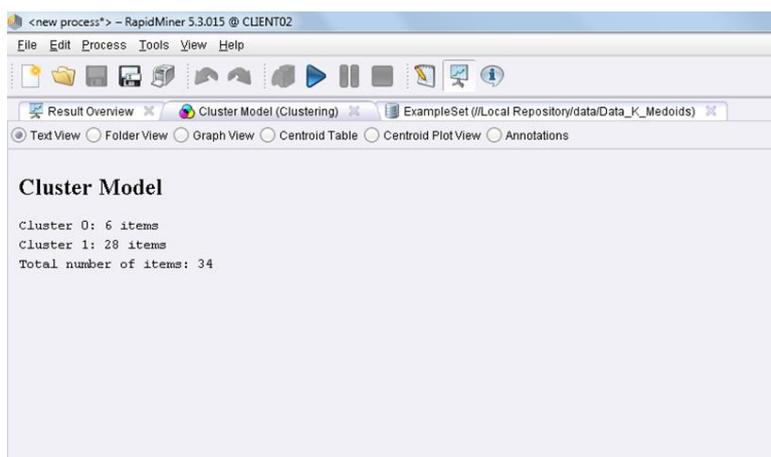


Gambar 2. Pengolahan Data dengan Rapidminer



**Gambar 3.** Grafik Pengelompokan dengan Rapidminer

Pada data yang telah dianalisa dapat disimpulkan bahwa data yang digunakan dalam penelitian ini adalah valid. Hal ini dapat dibuktikan dengan hasil akhir perhitungan manual dengan hasil akhir dari aplikasi Rapidminer 5.3 menampilkan hasil yang sama yaitu cluster tertinggi yaitu 28 dan cluster terendah 6. Berikut ini tampilan dari cluster model berupa teks yang terdapat pada gambar 4 dibawah ini:



**Gambar 4.** Cluster Model

Setelah dilakukan pengujian maka diperoleh 28 Provinsi yang terdapat pada Cluster Tertinggi (C1) yaitu Aceh, Sumatera Utara, Sumatera Barat, Riau, Jambi, Bengkulu, Lampung, Jawa Barat, Jawa Tengah, Jawa Timur, Banten, Nusa Tenggara Barat, Nusa Tenggara Timur, Kalimantan Barat, Kalimantan Tengah, Kalimantan Selatan, Kalimantan Timur, Kalimantan Utara, Sulawesi Utara, Sulawesi Tengah, Sulawesi Selatan, Sulawesi Tenggara, Gorontalo, Sulawesi Barat, Maluku, Maluku Utara, Papua Barat, Papua. Dan 6 Provinsi yang terdapat pada Cluster Terendah (C2) yaitu Sumatera Selatan, Kep. Bangka Belitung, Kep. Riau, Dki Jakarta, DI Yogyakarta, Bali.

## KESIMPULAN

Penelitian yang dilakukan menggunakan metode K-Medoids mampu untuk mengelompokkan balita yang mengalami stunting dengan menggunakan 2 cluster yaitu tinggi dan rendah. Adapun hasil perhitungan manual dan hasil clustering yang dievaluasi menggunakan Software Rapidminer 5.3 mendapatkan hasil yang sama yaitu terdapat 28 provinsi dengan cluster tertinggi dan 6 provinsi

dengan cluster terendah. Data yang didapat pada Badan Pusat Statistik dengan subjek prevalensi stunting pada tahun 2015-2018 pada 34 provinsi di Indonesia dapat dimanfaatkan untuk menghasilkan suatu informasi yang berguna dan dapat diimplementasikan menggunakan metode data miningsalah satunya Algoritma K-Medoids Clustering.

### DAFTAR PUSTAKA

- [1] F. Marisa, "Educational Data Mining (Konsep dan Penerapan)," *Teknol. Inf.*, vol. 4, no. 2, pp. 90–97, 2014, [Online]. Available: <http://ejournal.stimata.ac.id/index.php/TI/article/view/108/148>.
- [2] L. A. W. S, "Penerapan Data Mining Pada Penjualan Tiket Pesawat Menggunakan Algoritma Apriori ( Studi Kasus : PT . Sinar Jagad Indonesia )," *Juisi*, vol. 06, no. 01, 2020.
- [3] D. Marlina, N. Lina, A. Fernando, and A. Ramadhan, "Implementasi Algoritma K-Medoids dan K-Means untuk Pengelompokan Wilayah Sebaran Cacat pada Anak," *J. CoreIT J. Has. Penelit. Ilmu Komput. dan Teknol. Inf.*, vol. 4, no. 2, p. 64, 2018, doi: 10.24014/coreit.v4i2.4498.
- [4] N. R. Aeni *et al.*, "Algoritma Partitioning Around Medoids Dalam Mengelompokan Provinsi Di Indonesia Berdasarkan," pp. 1–9, 2018.
- [5] Y. H. Chrisnanto and G. Abdillah, "Penerapan algoritma partitioning around medoids ( PAM ) clustering untuk melihat gambaran umum kemampuan akademik mahasiswa," *Semin. Nas. Teknol. Inf. dan Komun.*, vol. 2015, no. Sentika, pp. 444–448, 2015.
- [6] D. F. Pramesti, Lahan, M. Tanzil Furqon, and C. Dewi, "Implementasi Metode K-Medoids Clustering Untuk Pengelompokan Data Potensi Kebakaran Hutan/Lahan Berdasarkan Persebaran Titik Panas(Hotspot)," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 1, no. 9, pp. 723–732, 2017.
- [7] D. Rofifah, "濟無No Title No Title No Title," *Pap. Knowl. . Towar. a Media Hist. Doc.*, pp. 12–26, 2020.
- [8] P. E. Prakasawati, Y. H. Chrisnanto, and A. I. Hadiana, "Segmentasi Pelanggan Berdasarkan Produk Menggunakan Metode K- Medoids," *KOMIK (Konferensi Nas. Teknol. Inf. dan Komputer)*, vol. 3, no. 1, pp. 335–339, 2019, doi: 10.30865/komik.v3i1.1610.
- [9] E. Buulolo, R. Syahputra, and A. Fau, "Algoritma K-Medoids Untuk Menentukan Calon Mahasiswa Yang Layak Mendapatkan Beasiswa Bidikmisi di Universitas Budi Darma," *J. MEDIA Inform. BUDIDARMA*, vol. 4, p. 797, Jul. 2020, doi: 10.30865/mib.v4i3.2240.
- [10] M. Sari and Asmendri, "NATURAL SCIENCE : Jurnal Penelitian Bidang IPA dan Pendidikan IPA," *Nat. Sci. J. Penelit. Bid. IPA dan Pendidik. IPA*, vol. 6, no. 1, pp. 41–53, 2020, [Online]. Available: <https://ejournal.uinib.ac.id/jurnal/index.php/naturalscience/article/view/1555/1159>.

