

## Penerapan Algoritma *K-Medoids* Dalam Pengelompokan Imunisasi Lanjutan Pada Anak Usia 2 Tahun

Hendra Kusumah<sup>1\*</sup>, Muhammad Ridwan Lubis<sup>2</sup>, Heru Satria Tambunan<sup>3</sup>

<sup>1,3</sup> STIKOM Tunas Bangsa, Pematangsiantar, Indonesia

<sup>2</sup> AMIK Tunas Bangsa, Pematangsiantar, Indonesia

Email: <sup>1</sup> [hendrakusuma353@gmail.com](mailto:hendrakusuma353@gmail.com), <sup>2</sup> [ridwanlubis@amiktunasbangsa.ac.id](mailto:ridwanlubis@amiktunasbangsa.ac.id), <sup>3</sup> [heru@amiktunasbangsa.ac.id](mailto:heru@amiktunasbangsa.ac.id)

**Abstrak**—Imunisasi adalah proses untuk membuat seseorang imun atau kebal terhadap suatu penyakit. Proses ini dilakukan dengan pemberian vaksin yang merangsang sistem kekebalan tubuh agar kebal terhadap penyakit tersebut. Imunisasi pada anak perlu dilakukan karena sistem kekebalan tubuh yang belum baik menyebabkan berbagai virus dan bakteri dapat mengganggu kesehatan anak. Imunisasi lanjutan merupakan terusan dari imunisasi dasar, pada imunisasi tahap ini berfungsi sebagai penguat daya tahan tubuh pada anak usia 2 tahun. Penelitian ini bertujuan untuk membuat sebuah model pengelompokan dengan menggunakan algoritma *K-Medoids*. Algoritma *K-Medoids* atau dikenal pula dengan PAM (*Partitioning Around Medoids*) menggunakan metode partisi *clustering* untuk mengelompokkan sekumpulan  $n$  objek menjadi sejumlah  $k$  cluster. Data pada penelitian ini bersumber dari Departemen Kesehatan Tahun 2017 sampai 2019. Pengelompokan dilakukan berdasarkan jumlah penerima vaksin DPT-HB-HiB dan Campak/MR dari 34 Provinsi yang ada di Indonesia. Dari hasil penelitian didapatkan 3 cluster dengan anggota yaitu: cluster 1 di isi 1 Provinsi, cluster 2 di isi 18 Provinsi, cluster 3 di isi 15 Provinsi. Diharapkan penelitian ini dapat menjadi rujukan kepada Pemerintah dalam meningkatkan pelayanan pada program imunisasi khususnya pada imunisasi lanjutan sehingga menjamin tingkat kesehatan anak di Indonesia.

**Kata Kunci:** *Imunisasi Lanjutan Pada Anak Usia 2 Tahun, Algoritma K-Medoids*

**Abstract**—*Immunization is a process to make someone immune or immune to a disease. This process is carried out by administering a vaccine that stimulates the immune system to be immune to the disease. Immunization in children needs to be done because the immune system that has not been good causes various viruses and bacteria that can interfere with children's health. Advanced immunization is a continuation of basic immunization, at this stage of immunization serves as a reinforcement of endurance in children aged 2 years. This study aims to create a grouping model using the K-Medoids algorithm. K-Medoids algorithm or also known as PAM (Partitioning Around Medoids) uses the clustering partition method to group a group of  $n$  objects into a number of  $k$  clusters. The data in this study were sourced from the Ministry of Health from 2017 to 2019. The grouping was based on the number of recipients of the DPT-HB-HiB and Measles / MR vaccine from 34 Provinces in Indonesia. From the research results obtained 3 clusters with members, namely: cluster 1 contains 1 province, cluster 2 contains 18 provinces, cluster 3 contains 15 provinces. It is hoped that this research can be a reference to the Government in improving services in the immunization program, especially in advanced immunizations so as to ensure the health level of children in Indonesia.*

**Keywords:** *Advanced Immunization in Children aged 2 years, K-Medoids Algorithm*

### 1. PENDAHULUAN

Imunisasi adalah proses untuk membuat seseorang imun atau kebal terhadap suatu penyakit. Proses ini dilakukan dengan pemberian vaksin yang merangsang sistem kekebalan tubuh agar kebal terhadap penyakit tersebut. Imunisasi diberikan kepada anak dari mulai dilahirkan, pada usia 2 tahun anak membutuhkan imunisasi lanjutan dengan memberikan vaksin DPT-HB-HiB dan Campak/MR untuk meningkatkan kekebalan tubuh pada anak. Menurut peraturan Menteri Kesehatan Republik Indonesia nomor 12 tahun 2017 tentang penyelenggaraan imunisasi, imunisasi lanjutan adalah ulangan dari imunisasi dasar untuk meningkatkan kekebalan dan masa perlindungan anak yang sudah mendapat imunisasi dasar. Imunisasi yang diberikan kepada anak diusia dua tahun terdiri dari imunisasi terhadap penyakit difteri, pertusis, tetanus, hepatitis b, pneumonia dan meningitis yang disebabkan oleh *Hemophilus Influenza* tipe b (Hib) dan campak.

Menurut Undang-Undang Nomor 36 Tahun 2009 tentang kesehatan, imunisasi adalah bentuk kegiatan prioritas Kementerian Kesehatan RI untuk mencegah terjadinya penyakit menular. Imunisasi merupakan salah satu bentuk nyata komitmen pemerintah untuk mencapai *Sustainable Development Goals (SDGs)* untuk menurunkan angka kematian pada anak. Dengan meningkatnya cakupan imunisasi lanjutan pada anak usia 2 tahun akan menurunkan tingkat kematian pada anak, sehingga perlu dilakukan pendataan untuk mengetahui nilai cakupan imunisasi lanjutan disetiap Provinsi di Indonesia.

Beberapa penelitian pernah membahas masalah pengolahan data pada bidang kesehatan dan diselesaikan menggunakan konsep komputerisasi. Konsep yang dimaksud ialah pengelompokan, metode yang digunakan adalah algoritma *K-Medoids Clustering* seperti pada penelitian yang dilakukan oleh beberapa peneliti, yaitu pada penelitian [1] yang berjudul “Analisis *K-Medoids Clustering* Dalam Pengelompokan Data Imunisasi Campak Balita di Indonesia”. Dari data pengelompokan tersebut diperoleh 3 cluster yakni cluster rendah (2 provinsi), cluster sedang (30 provinsi) dan cluster tinggi (2 provinsi) dari persentase imunisasi campak balita pada setiap provinsi. Pada penelitian [2] yang berjudul “Penerapan Algoritma *K-Medoids* untuk Pengelompokan Penyakit di Pekanbaru Riau”. Berdasarkan penelitian didapatkan hasil 4 cluster, Pada cluster 1 didapatkan 420 record dengan penyakit dominan

adalah Maag sebesar 44,39%, cluster 2 didapatkan 349 record dengan penyakit dominan adalah Diare dan Sakit Perut sebesar 16,98%, pada cluster 3 didapatkan 794 record dengan penyakit dominan adalah Batuk dan Pilek sebesar 65,21% dan pada cluster 4 didapatkan 1248 record dengan penyakit dominan adalah Batuk dan Pilek sebesar 54,10%. Pada penelitian [3] "Pengelompokan Kasus Penyakit Aids Berdasarkan Provinsi Dengan Data Mining K-Medoids Clustering". Penelitian menggunakan 2 cluster dan diperoleh penilaian cluster C1 sebanyak 5 provinsi yaitu Papua, DKI Jakarta, Jawa Barat, Jawa Timur dan Bali, 28 provinsi untuk cluster C2.

Berdasarkan permasalahan yang telah diuraikan di atas, untuk mencari Provinsi mana yang cakupan imunisasi lanjutannya masih rendah melalui data yang di dapatkan dari laman resmi Kementerian Kesehatan berupa jumlah imunisasi lanjutan pada anak usia 2 tahun dapat dilakukan pengelompokan menggunakan teknik *data mining* yaitu Algoritma *K-Medoids Clustering*. Berdasarkan penelitian sebelumnya Algoritma *K-Medoids* dinilai cocok digunakan untuk menyelesaikan masalah pengelompokan. Diharapkan penelitian ini dapat menjadi bahan pertimbangan untuk pemerintah Indonesia dalam hal peningkatan pelayanan imunisasi kepada masyarakat, serta dukungan sosialisasi imunisasi lanjutan kepada masyarakat. Dengan meningkatnya pelayanan dan jumlah anak yang diimunisasi maka tujuan Pemerintah untuk menurunkan angka kematian pada anak di Indonesia dapat terwujud.

## 2. METODOLOGI PENELITIAN

### 2.1 Data Mining

Menurut Gartner Group dalam [4] *data mining* adalah proses menemukan hubungan baru yang mempunyai arti, pola dan kebiasaan dengan memilah-milah sebagian besar data yang disimpan dalam media penyimpanan dengan menggunakan teknologi pengenalan pola seperti teknik statistik dan matematika. *Data mining* merupakan gabungan dari beberapa disiplin ilmu yang menyatukan teknik dari pembelajaran mesin, pengenalan pola, statistik, *database* dan visualisasi untuk penanganan permasalahan pengambilan informasi dari *database* yang besar'.

Menurut Larose dalam [5] *Data mining* semakin menyebar dan berkembang dengan pesat belakangan ini karena kemampuannya dalam menambang pola bermanfaat dan trend dari basis data yang sudah ada. Perusahaan-perusahaan telah menghabiskan dana milyaran untuk mengumpulkan data dalam jumlah *megabytes* atau *terabytes* tapi tidak mendapatkan keuntungan yang bernilai di dalamnya, padahal di dalamnya terbatasi informasi yang berharga namun tersembunyi pada *repositories* data

### 2.2 Clustering

Pengklasteran merupakan satu dari sekian banyak fungsi proses data mining untuk menemukan kelompok atau identifikasi kelompok obyek yang hampir sama. Analisis kluster (*clustering*) merupakan usaha untuk mengidentifikasi kelompok obyek yang mirip-mirip dan membantu menemukan pola penyebaran dan pola hubungan dalam sekumpulan data yang besar. Hal penting dalam proses pengklasteran adalah menyatakan sekumpulan pola ke kelompok yang sesuai yang berguna untuk menemukan kesamaan dan perbedaan sehingga dapat menghasilkan kesimpulan yang berharga [6].

### 2.3 Algoritma K-Medoids

Algoritma *K-Medoids* atau dikenal pula dengan PAM (*Partitioning Around Medoids*) menggunakan metode partisi *clustering* untuk mengelompokkan sekumpulan  $n$  objek menjadi sejumlah  $k$  *cluster*. Algoritma ini menggunakan objek pada kumpulan objek untuk mewakili sebuah *cluster*. Objek yang terpilih untuk mewakili sebuah *cluster* disebut *medoid*. *Cluster* dibangun dengan menghitung kedekatan yang dimiliki antara *medoid* dengan objek *non-medoids* [6].

Langkah-langkah *K-Medoids* adalah [7] :

1. Pilih poin  $k$  sebagai inisial *centroid* / nilai tengah (*medoids*) sebanyak  $k$  *cluster*.
2. Cari semua poin yang paling dekat dengan *medoids*, dengan cara menghitung jarak *vektor* antar dokumen dengan menggunakan *Euclidean Distance*. Rumusnya adalah sebagai berikut :

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}; i = 1, 2, 3, \dots, n \quad (1)$$

Dimana:

$d(x, y)$  = jarak antara data ke- $i$  dan data ke- $j$

$x_i$  = nilai atribut ke satu dari data ke- $i$

$y_j$  = nilai atribut ke satu dari data ke- $j$

$n$  = jumlah atribut yang digunakan

3. Secara acak, pilih poin yang bukan *medoids*.
4. Hitung total jarak antar *medoids*.
5. Jika TD baru < TD awal, tukar posisi *medoids* dengan *medoids* baru, jadilah *medoids* yang baru.

Ulangi langkah 2-5 sampai *medoids* tidak berubah.

### 2.4 RapidMiner

*RapidMiner* merupakan sebuah perangkat lunak untuk melakukan analisis terhadap *data mining*, text mining dan analisis prediksi. *RapidMiner* menggunakan berbagai teknik deskriptif dan prediksi dalam memberikan wawasan

kepada pengguna sehingga dapat membuat keputusan yang paling banyak. *RapidMiner* memiliki kurang lebih 500 operator *data mining* termasuk operator untuk *input*, *output*, *data preprocessing* dan visualisasi [8].

### 3. HASIL DAN PEMBAHASAN

Pengolahan data menggunakan Algoritma *K-Medoids* untuk memperoleh *cluster* imunisasi lanjutan pada anak usia 2 tahun sesuai dengan data skunder yang telah diperoleh pada tabel 1. akan dilakukan proses perhitungan secara manual.

**Tabel 1.** Data Persentase Imunisasi Lanjutan Di Indonesia

No	Provinsi	DPT-HB-HiB 2017	Campak / MR 2017	DPT-HB-HiB 2018	Campak / MR 2018	DPT-HB-HiB 2019	Campak / MR 2019
1	Aceh	26,7	20,4	20,3	13,3	26,3	25
2	Sumatera Utara	56	54,5	54,7	55,4	70,3	64,9
3	Sumatera Barat	49,3	44,3	41,3	36,1	44,6	40,3
4	Riau	38,7	34,8	40,2	34,2	46,1	42,4
5	Jambi	77,7	72,1	78,8	84,4	82,3	78,5
6	Sumatera Selatan	56,6	54,2	85,0	94,1	93,2	94,5
7	Bengkulu	79,1	72,8	66,5	85,2	80,7	72,3
8	Lampung	66,7	62,4	69,2	74	87	80,3
9	Kepulauan Bangka Belitung	59,1	52	55,5	63,6	62,5	51,4
10	Kepulauan Riau	65,6	60,1	69,5	73,2	68,3	66,6
11	DKI Jakarta	86,8	75,5	133,9	112,9	97	96,4
12	Jawa Barat	53,4	74,1	66	59,5	77,4	74,7
13	Jawa Tengah	74,7	63,8	84	81,8	87,9	86,2
14	DI Yogyakarta	82	72,3	59	61,8	100,9	97,8
15	Jawa Timur	68,9	80,7	92,1	58,7	84,6	86
16	Banten	53,8	69,4	66,8	60,1	81,7	77,5
17	Bali	84,1	81,8	81,3	85,1	82,2	84,2
18	Nusa Tenggara Barat	76,4	75,3	64,9	67	78,1	73,7
19	Nusa Tenggara Timur	33,6	31	25	34	43,2	40,6
21	Kalimantan Tengah	50,7	42,5	42,4	57,9	55,6	51,2
22	Kalimantan Selatan	49	44,6	33,5	40,3	60,7	50
23	Kalimantan Timur	60,4	47,2	54,1	49	68,6	58,4
24	Kalimantan Utara	53,5	43,2	54,9	52,2	54,7	46,8
25	Sulawesi Utara	52,4	41,9	45,9	54,5	67,1	57,7
26	Sulawesi Tengah	50,5	50,9	57,4	53,1	71	77,1
27	Sulawesi Selatan	79,3	74,7	65,1	79,9	84,4	81,4
28	Sulawesi Tenggara	51,7	45,6	47,8	54,8	62	53,4
29	Gorontalo	50,6	43,7	37,4	34,9	55,2	49,5
30	Sulawesi Barat	57,7	49,4	41,5	46,8	56,7	47,7
31	Maluku	42,7	38,8	45,9	46,5	55,3	51,7
32	Maluku Utara	46,4	39,1	47,1	38,8	56,3	45,9
33	Papua Barat	30,9	28,6	42,9	266	50,7	27,3
34	Papua	25,6	24,7	23,7	29,9	41,7	37,7

#### 3.1 Proses Perhitungan Algoritma *K-Medoids*

Perhitungan Algoritma *K-Medoids* untuk memperoleh *cluster* dapat diuraikan sebagai berikut :

- Menentukan jumlah *cluster* yang akan digunakan pada penelitian. Pada penelitian ini penulis memilih menggunakan 3 *cluster* untuk mewakili *cluster* tinggi, *cluster* sedang dan *cluster* rendah.
- Menentukan pusat *cluster* sebanyak 3 *cluster* dari dataset yang akan digunakan sebagai *medoids* awal. Data yang akan digunakan sebagai pusat *cluster* (*centroid*) dapat dilihat pada tabel 2.

**Tabel 2.** *Medoids* Awal

<i>Cluster</i>	Provinsi	DPT-HB-HiB 2017	Campak / MR 2017	DPT-HB-HiB 2018	Campak / MR 2018	DPT-HB-HiB 2019	Campak / MR 2019
1	Papua Barat	30,9	28,6	42,9	266	50,7	27,3
2	Jawa Tengah	74,7	63,8	84	81,8	87,9	86,2
3	Papua	25,6	24,7	23,7	29,9	41,7	37,7

- Menghitung nilai jarak (*cost*) dengan persamaan *Euclidean Distance*. Untuk menghitung jarak antara titik *centroid* dengan titik tiap objek data terdekat dapat dilakukan dengan menggunakan persamaan (2).

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}; i = 1, 2, 3, \dots, n \quad (2)$$

Maka perhitungan jarak *centroid* dengan objek data dapat di lihat sebagai berikut ini :

### Iterasi 1

#### Perhitungan jarak *centroid* pertama dengan beberapa data

$$D(1,1) = \sqrt{\frac{(26,7 - 30,9)^2 + (20,4 - 28,6)^2 + (20,3 - 42,9)^2 + (13,3 - 266)^2}{(26,3 - 50,7)^2 + (25 - 27,3)^2}}$$

$$= 255,06$$

$$D(2,1) = \sqrt{\frac{(56 - 30,9)^2 + (54,5 - 28,6)^2 + (54,7 - 42,9)^2 + (55,4 - 266)^2}{(70,3 - 50,7)^2 + (64,9 - 27,3)^2}}$$

$$= 218,15$$

$$D(3,1) = \sqrt{\frac{(49,3 - 30,9)^2 + (44,3 - 28,6)^2 + (41,3 - 42,9)^2 + (36,1 - 266)^2}{(44,6 - 50,7)^2 + (40,3 - 27,3)^2}}$$

$$= 231,62$$

#### Perhitungan jarak *centroid* kedua dengan seluruh data

$$D(1,2) = \sqrt{\frac{(26,7 - 74,7)^2 + (20,4 - 63,8)^2 + (20,3 - 84)^2 + (13,3 - 81,8)^2}{(26,3 - 87,9)^2 + (25 - 86,2)^2}}$$

$$= 143,10$$

$$D(2,2) = \sqrt{\frac{(56 - 74,7)^2 + (54,5 - 63,8)^2 + (54,7 - 84)^2 + (55,4 - 81,8)^2}{(70,3 - 87,9)^2 + (64,9 - 86,2)^2}}$$

$$= 52,49$$

$$D(3,2) = \sqrt{\frac{(49,3 - 74,7)^2 + (44,3 - 63,8)^2 + (41,3 - 84)^2 + (36,1 - 81,8)^2}{(44,6 - 87,9)^2 + (40,3 - 86,2)^2}}$$

#### Perhitungan jarak *centroid* ketiga dengan seluruh data

$$D(1,3) = \sqrt{\frac{(26,7 - 25,6)^2 + (20,4 - 24,7)^2 + (20,3 - 23,7)^2 + (13,3 - 29,9)^2}{(26,3 - 41,7)^2 + (25 - 37,7)^2}}$$

$$= 26,56$$

$$D(2,3) = \sqrt{\frac{(56 - 25,6)^2 + (54,5 - 24,7)^2 + (54,7 - 23,7)^2 + (55,4 - 29,9)^2}{(70,3 - 41,7)^2 + (64,9 - 37,7)^2}}$$

$$= 70,58$$

$$D(3,3) = \sqrt{\frac{(49,3 - 25,6)^2 + (44,3 - 24,7)^2 + (41,3 - 23,7)^2 + (36,1 - 29,9)^2}{(44,6 - 41,7)^2 + (40,3 - 37,7)^2}}$$

$$= 36,18$$

Tabel 3. Hasil Perhitungan Iterasi 1

No	Provinsi	Hasil Perhitungan C1	Hasil Perhitungan C2	Hasil Perhitungan C3
1	Aceh	255,06	143,10	26,56
2	Sumatera Utara	218,15	52,49	70,58
3	Sumatera Barat	231,62	94,44	36,18
4	Riau	232,57	99,93	24,61
5	Jambi	204,87	14,22	119,50
6	Sumatera Selatan	197,38	25,87	124,86
7	Bengkulu	201,12	25,75	113,07
8	Lampung	210,02	19,53	104,84
9	Kepulauan Bangka Belitung	207,82	58,16	68,08
10	Kepulauan Riau	204,77	33,90	91,48
11	DKI Jakarta	209,67	62,67	178,50
12	Jawa Barat	220,71	40,29	92,32
13	Jawa Tengah	208,88	0,00	121,47
14	DI Yogyakarta	232,31	38,13	121,77
15	Jawa Timur	232,70	30,49	121,21
16	Banten	220,54	36,73	93,53
17	Bali	209,98	21,61	129,81
18	Nusa Tenggara Barat	217,35	31,17	104,10
19	Nusa Tenggara Timur	233,22	112,29	11,53
20	Kalimantan Barat	216,64	67,31	55,38
21	Kalimantan Tengah	210,92	74,83	49,56
22	Kalimantan Selatan	228,53	85,74	40,74
23	Kalimantan Timur	222,98	59,96	64,49
24	Kalimantan Utara	216,74	72,47	53,32
25	Sulawesi Utara	215,81	66,47	56,19
26	Sulawesi Tengah	222,06	51,49	73,42

No	Provinsi	Hasil Perhitungan C1	Hasil Perhitungan C2	Hasil Perhitungan C3
27	Sulawesi Selatan	208,94	23,15	115,46
28	Sulawesi Tenggara	214,85	68,16	54,57
29	Gorontalo	233,60	88,16	38,99
30	Sulawesi Barat	222,83	77,35	50,67
31	Maluku	221,47	81,24	40,50
32	Maluku Utara	228,83	85,11	39,33
33	Papua Barat	0,00	208,88	237,37
34	Papua	237,37	121,47	0,00

- d) Menentukan *cluster* awal sebuah objek data berdasarkan nilai terendah dari hasil perhitungan jarak *centroid* dengan objek data pada iterasi 1.

Tabel 4. Cluster Awal Iterasi 1

No	Provinsi	Jarak Ke Medoids			Terdekat	Cluster Yang Diikuti
		C1	C2	C3		
1	Aceh	255,06	143,10	26,56	26,56	3
2	Sumatera Utara	218,15	52,49	70,58	52,49	2
3	Sumatera Barat	231,62	94,44	36,18	36,18	3
4	Riau	232,57	99,93	24,61	24,61	3
5	Jambi	204,87	14,22	119,50	14,22	2
6	Sumatera Selatan	197,38	25,87	124,86	25,87	2
7	Bengkulu	201,12	25,75	113,07	25,75	2
8	Lampung	210,02	19,53	104,84	19,53	2
9	Kepulauan Bangka Belitung	207,82	58,16	68,08	58,16	2
10	Kepulauan Riau	204,77	33,90	91,48	33,90	2
11	DKI Jakarta	209,67	62,67	178,50	62,67	2
12	Jawa Barat	220,71	40,29	92,32	40,29	2
13	Jawa Tengah	208,88	0,00	121,47	0,00	2
14	DI Yogyakarta	232,31	38,13	121,77	38,13	2
15	Jawa Timur	232,70	30,49	121,21	30,49	2
16	Banten	220,54	36,73	93,53	36,73	2
17	Bali	209,98	21,61	129,81	21,61	2
18	Nusa Tenggara Barat	217,35	31,17	104,10	31,17	2
19	Nusa Tenggara Timur	233,22	112,29	11,53	11,53	3
20	Kalimantan Barat	216,64	67,31	55,38	55,38	3
21	Kalimantan Tengah	210,92	74,83	49,56	49,56	3
22	Kalimantan Selatan	228,53	85,74	40,74	40,74	3
23	Kalimantan Timur	222,98	59,96	64,49	59,96	2
24	Kalimantan Utara	216,74	72,47	53,32	53,32	3
25	Sulawesi Utara	215,81	66,47	56,19	56,19	3
26	Sulawesi Tengah	222,06	51,49	73,42	51,49	2
27	Sulawesi Selatan	208,94	23,15	115,46	23,15	2
28	Sulawesi Tenggara	214,85	68,16	54,57	54,57	3
29	Gorontalo	233,60	88,16	38,99	38,99	3
30	Sulawesi Barat	222,83	77,35	50,67	50,67	3
31	Maluku	221,47	81,24	40,50	40,50	3
32	Maluku Utara	228,83	85,11	39,33	39,33	3
33	Papua Barat	0,00	208,88	237,37	0,00	1
34	Papua	237,37	121,47	0,00	0,00	3
Total Jarak Ke Medoids (cost)					1203,75	

- e) Memilih *Medoids* baru dari objek data yang bukan objek data pada *medoids* awal (*Non Medoids*). *Medoids* yang terpilih dapat di lihat pada tabel 5.

Tabel 5. Medoids Baru

Cluster	Provinsi	DPT- HB-HiB	Campak / MR	DPT- HB-HiB	Campak / MR	DPT- HB-HiB	Campak / MR
		2017	2017	2018	2018	2019	2019
1	DKI Jakarta	86,8	75,5	133,9	112,9	97	96,4
2	Nusa Tenggara Barat	76,4	75,3	64,9	67	78,1	73,7
3	Aceh	26,7	20,4	20,3	13,3	26,3	25,0

- f) Menghitung jarak objek data dengan *medoids* baru dengan rumus pada persamaan (2) dan menentukan jarak terdekat setiap objek. Hasil dari perhitungan jarak data dengan *medoids* baru dapat di lihat pada tabel 4.6.

Tabel 6. Hasil Perhitungan Iterasi 2

No	Provinsi	Jarak Ke Medoids			Terdekat	Cluster Yang Diikuti
		C1	C2	C3		
1	Aceh	198,92	124,14	0,00	0,00	3
2	Sumatera Utara	112,58	35,01	92,22	35,01	2

No	Provinsi	Jarak Ke Medoids			Terdekat	Cluster Yang Diikuti
		C1	C2	C3		
3	Sumatera Barat	150,82	73,79	51,11	51,11	3
4	Riau	156,37	82,17	43,35	43,35	3
5	Jambi	66,93	23,42	140,53	23,42	2
6	Sumatera Selatan	64,25	51,35	148,52	51,35	2
7	Bengkulu	78,89	18,87	134,14	18,87	2
8	Lampung	81,45	21,23	127,21	21,23	2
9	Kepulauan Bangka Belitung	114,51	41,02	88,46	41,02	2
10	Kepulauan Riau	90,12	23,53	112,22	23,53	2
11	DKI Jakarta	0,00	88,59	198,92	0,00	1
12	Jawa Barat	97,13	24,28	113,58	24,28	2
13	Jawa Tengah	62,67	31,17	143,10	31,17	2
14	DI Yogyakarta	90,95	34,68	143,06	34,68	2
15	Jawa Timur	72,76	32,98	140,55	32,98	2
16	Banten	94,91	24,98	115,38	24,98	2
17	Bali	62,88	28,73	150,23	28,73	2
18	Nusa Tenggara Barat	88,59	0,00	124,14	0,00	2
19	Nusa Tenggara Timur	170,01	93,75	33,76	33,76	3
20	Kalimantan Barat	125,47	49,75	76,58	49,75	2
21	Kalimantan Tengah	132,46	57,77	71,32	57,77	2
22	Kalimantan Selatan	145,62	65,24	61,60	61,60	3
23	Kalimantan Timur	119,16	42,55	84,71	42,55	2
24	Kalimantan Utara	127,78	56,08	72,32	56,08	2
25	Sulawesi Utara	125,93	50,85	78,80	50,85	2
26	Sulawesi Tengah	111,35	39,72	95,76	39,72	2
27	Sulawesi Selatan	79,14	16,56	136,80	16,56	2
28	Sulawesi Tenggara	126,45	51,04	76,29	51,04	2
29	Gorontalo	147,19	67,53	57,52	57,52	3
30	Sulawesi Barat	135,76	55,77	69,38	55,77	2
31	Maluku	138,50	65,22	62,50	62,50	3
32	Maluku Utara	142,09	67,61	58,68	58,68	3
33	Papua Barat	209,67	217,35	255,06	209,7	1
34	Papua	178,50	104,10	26,56	26,56	3
Total Jarak Ke Medoids (cost)					1416,07	

- g) Menghitung nilai total simpangan ( $S$ ) setelah mendapatkan jarak ke *medoids* pada iterasi ke-1 dan iterasi ke-2 dengan cara menghitung total *cost* baru – total *cost* lama. Jika  $S < 0$  maka tukar posisi *medoids* dengan *medoids* baru. Jika  $S > 0$  maka proses dihentikan.

$$S = \text{Total Cost baru} - \text{Total Cost awal} \quad (3)$$

$$= 1416,07 - 1203,75$$

$$= 212,31$$

- h) Ulangi tahap kedua sampai keenam sampai *Medoids* tidak berubah. Karena di tahap ketujuh nilai  $S = 212,31$  maka proses ini tidak di lanjutkan.

Dari proses perhitungan manual didapatkan hasil yaitu, *cluster* 1 diisi dengan (1 Provinsi), *cluster* 2 diisi dengan (18 Provinsi), *cluster* 3 diisi dengan (15 Provinsi). Pada tabel 7. dapat dilihat pembagian *cluster* tiap Provinsi.

**Tabel 7.** Hasil *Cluster* Perhitungan Manual

No	Provinsi	Cluster
1	Aceh	3
2	Sumatera Utara	2
3	Sumatera Barat	3
4	Riau	3
5	Jambi	2
6	Sumatera Selatan	2
7	Bengkulu	2
8	Lampung	2
9	Kepulauan Bangka Belitung	2
10	Kepulauan Riau	2
11	DKI Jakarta	2
12	Jawa Barat	2
13	Jawa Tengah	2
14	DI Yogyakarta	2
15	Jawa Timur	2
16	Banten	2
17	Bali	2
18	Nusa Tenggara Barat	2
19	Nusa Tenggara Timur	3
20	Kalimantan Barat	3
21	Kalimantan Tengah	3
22	Kalimantan Selatan	3
23	Kalimantan Timur	2

No	Provinsi	Cluster
24	Kalimantan Utara	3
25	Sulawesi Utara	3
26	Sulawesi Tengah	2
27	Sulawesi Selatan	2
28	Sulawesi Tenggara	3
29	Gorontalo	3
30	Sulawesi Barat	3
31	Maluku	3
32	Maluku Utara	3
33	Papua Barat	1
34	Papua	3

### 3.2 Pengujian Data

Untuk melakukan pengujian terhadap pengelompokan dapat dilakukan dengan Aplikasi *RapidMiner* 5.3. Hasil akhir yang akan di tampilkan berupa pengelompokan dimana akan tampak *cluster* dengan masing-masing anggota. Dari proses yang sudah dilakukan sebelumnya maka hasil akhir akan ditampilkan pada aplikasi *RapidMiner*. Ada pun beberapa yang di tampilkan oleh *RapidMiner* sebagai berikut :

Row No.	Provinsi	cluster	DPT-HB-HI	CampakM	DPT-HB-HI	CampakM	DPT-HB-HI	CampakM
1	Aceh	cluster_2	26.700	20.400	20.300	13.300	26.300	25
2	Sumatera Utara	cluster_1	56	54.500	54.700	55.400	70.300	64.900
3	Sumatera Barat	cluster_2	49.300	44.300	41.300	36.100	44.600	40.300
4	Riau	cluster_2	38.700	34.800	40.200	34.200	45.100	42.400
5	Jambi	cluster_1	77.700	72.100	78.800	84.400	82.300	78.500
6	Sumatera Selatan	cluster_1	56.600	54.200	85	94.100	93.200	94.500
7	Bengkulu	cluster_1	79.100	72.800	66.500	85.200	80.700	72.300
8	Lampung	cluster_1	66.700	62.400	69.200	74	87	80.300
9	Kepulauan Bangka Belitung	cluster_1	59.100	52	55.500	63.600	62.500	51.400
10	Kepulauan Riau	cluster_1	65.600	60.100	69.500	73.200	68.300	66.600
11	DKI Jakarta	cluster_1	66.800	75.500	133.900	112.800	97	66.400
12	Jawa Barat	cluster_1	63.400	74.100	66	59.500	77.400	74.700
13	Jawa Tengah	cluster_1	74.700	63.800	84	81.800	87.500	86.200
14	DI Yogyakarta	cluster_1	82	72.300	59	61.800	100.900	97.800
15	Jawa Timur	cluster_1	68.900	80.700	82.100	58.700	84.600	86
16	Banten	cluster_1	53.800	69.400	66.800	60.100	81.700	77.500
17	Bali	cluster_1	84.100	81.800	81.300	85.100	82.200	84.200
18	Nusa Tenggara Barat	cluster_1	76.400	75.300	84.900	67	78.100	73.100
19	Nusa Tenggara Timur	cluster_2	33.600	31	25	34	43.200	40.600
20	Kalimantan Barat	cluster_2	42.200	47.200	48.600	63.100	60.200	64.800

Gambar 1. Hasil Pengolahan RapidMiner

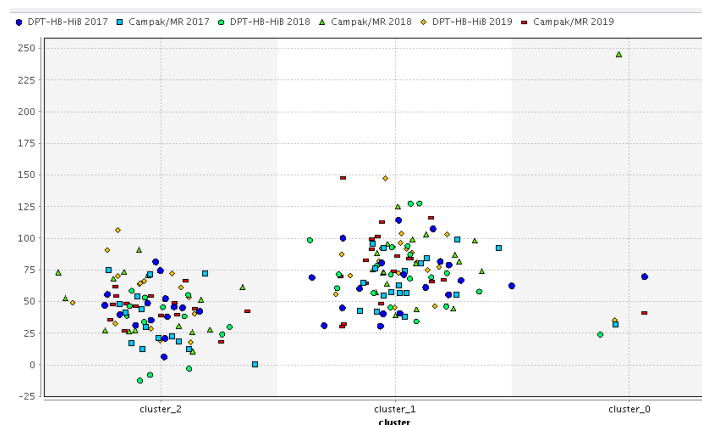
Pada Gambar diatas ditampilkan data setelah di proses oleh *RapidMiner* yang mana ada penambahan 1 kolom setelah kolom Provinsi. Kolom *cluster* ini merupakan hasil dari proses pengelompokan. Index *cluster* pada aplikasi tidak dimulai dari angka 1 melainkan angka 0. Sehingga jika kita lihat provinsi Aceh dan Sumatera Barat bukan berada di *cluster* 2 melainkan di *cluster* 3 begitu juga yang lainnya.

#### Cluster Model

Cluster 0: 1 items  
 Cluster 1: 18 items  
 Cluster 2: 15 items  
 Total number of items: 34

Gambar 2. Jumlah Anggota Cluster

Pada Gambar diatas menampilkan jumlah anggota tiap *cluster*. *Cluster* 1 diisi oleh 1 anggota, *cluster* 2 diisi oleh 18 anggota dan *cluster* 3 diisi oleh 15 anggota.



Gambar 3. Titik Sebaran Cluster

Pada Gambar diatas dapat dilihat titik sebaran dari cluster yang telah didapatkan, *cluster* 2 memiliki titik sebar yang lebih banyak karena memiliki anggota yang lebih banyak yaitu 18 anggota. Dari proses pengolahan data pada perhitungan manual dan aplikasi *RapidMiner* telah terpilih 3 *cluster* dengan masing-masing anggota. *Cluster – cluster* tersebut mewakili 3 criteria yang di tentukan sebelumnya yaitu tinggi sedang dan rendah. Pada Tabel 8. dapat dilihat hasil *cluster* dari perhitungan manual dan aplikasi *RapidMiner*.

Tabel 8. Perbandingan Hasil Cluster

No	Provinsi	Perhitungan Manual	Aplikasi Rapid Miner
1	Aceh	3	3
2	Sumatera Utara	2	2
3	Sumatera Barat	3	3
4	Riau	3	3
5	Jambi	2	2
6	Sumatera Selatan	2	2
7	Bengkulu	2	2
8	Lampung	2	2
9	Kepulauan Bangka Belitung	2	2
10	Kepulauan Riau	2	2
11	DKI Jakarta	2	2
12	Jawa Barat	2	2
13	Jawa Tengah	2	2
14	DI Yogyakarta	2	2
15	Jawa Timur	2	2
16	Banten	2	2
17	Bali	2	2
18	Nusa Tenggara Barat	2	2
19	Nusa Tenggara Timur	3	3
20	Kalimantan Barat	3	3
21	Kalimantan Tengah	3	3
22	Kalimantan Selatan	3	3
23	Kalimantan Timur	2	2
24	Kalimantan Utara	3	3
25	Sulawesi Utara	3	3
26	Sulawesi Tengah	2	2
27	Sulawesi Selatan	2	2
28	Sulawesi Tenggara	3	3
29	Gorontalo	3	3
30	Sulawesi Barat	3	3
31	Maluku	3	3
32	Maluku Utara	3	3
33	Papua Barat	1	1
34	Papua	3	3

Hasil perhitungan manual dan pengolahan dengan Aplikasi *RapidMiner* memiliki hasil *cluster* yang sama, walaupun *centroid* yang digunakan sedikit berbeda. *cluster* 1 diisi oleh Provinsi Papua Barat terpilih sebagai *cluster* tinggi. *Cluster* 2 sebagai *cluster* sedang di isi oleh Provinsi Sumatera Utara, Jambi, Sumatera Selatan, Bengkulu, Lampung, Kepulauan Bangka Belitung, Kepulauan Riau, DKI Jakarta, Jawa Barat, Jawa Tengah, DI Yogyakarta, Jawa Timur, Banten, Bali, Nusa Tenggara Barat, Kalimantan Timur, Sulawesi Tengah, dan Sulawesi Selatan. *Cluster* 3 sebagai *cluster* rendah diisi oleh Provinsi Aceh, Sumatera Barat, Riau, Nusa Tenggara Timur, Kalimantan Barat, Kalimantan Tengah, Kalimantan Selatan, Kalimantan Utara, Sulawesi Utara, Sulawesi Tenggara, Gorontalo, Sulawesi Barat, Maluku, Maluku Utara dan Papua.

#### 4. KESIMPULAN

Dari pembahasan sebelumnya dapat ditarik kesimpulan sebagai berikut :

- Penerapan Data mining dengan menggunakan Algoritma *K-medoids* dalam pengelompokan imunisasi lanjutan menghasilkan 3 *cluster*, *cluster* 1 sebanyak 1, *cluster* 2 sebanyak 18, *cluster* 3 sebanyak 15.
- Cluster* 1 merupakan *cluster* dengan nilai tertinggi diikuti *cluster* 2 sebagai *cluster* sedang dan terakhir *cluster* 3 sebagai rendah. Adapun Provinsi yang ada di *cluster* rendah yaitu Provinsi Aceh, Sumatera Barat, Riau, Nusa Tenggara Timur, Kalimantan Barat, Kalimantan Tengah, Kalimantan Selatan, Kalimantan Utara, Sulawesi Utara, Sulawesi Tenggara, Gorontalo, Sulawesi Barat, Maluku, Maluku Utara dan Papua.
- Pengujian data pada *RapidMiner* 5.3 menggunakan *K-Medoids* Menghasilkan *cluster* yang sama dengan pengujian manual. Dari hasil yang didapatkan tingkat akurasi *RapidMiner* sebesar 100%.

#### REFERENCES

- [1] S. Sundari, I. S. Damanik, A. P. Windarto, H. S. Tambunan, J. Jalaluddin, and A. Wanto, "Analisis K-Medoids Clustering Dalam Pengelompokkan Data Imunisasi Campak Balita di Indonesia," *Pros. Semin. Nas. Ris. Inf. Sci.*, vol. 1, no.

- September, p. 687, 2019, doi: 10.30645/senaris.v1i0.75.
- [2] T. Juninda, Mustasim, and E. Andri, "Penerapan Algoritma K-Medoids untuk Pengelompokan Penyakit di Pekanbaru Riau," *Semin. Nas. Teknol. Informasi, Komun. dan Ind.*, vol. 11, no. 1, pp. 42–49, 2019.
- [3] L. Purba, S. Saifullah, and R. Dewi, "Pengelompokan Kasus Penyakit Aids Berdasarkan Provinsi Dengan Data Mining K-Medoids Clustering," *KOMIK (Konferensi Nas. Teknol. Inf. dan Komputer)*, vol. 3, no. 1, pp. 687–694, 2019, doi: 10.30865/komik.v3i1.1679.
- [4] Y. Mardi, "Data Mining : Klasifikasi Menggunakan Algoritma C4.5," *J. Edik Inform.*, vol. 2, no. 2, pp. 213–219, 2017.
- [5] F. L. Sibuea and A. Sapta, "Pemetaan Siswa Berprestasi Menggunakan Metode K-Means Clustering," *JURTEKSI (Jurnal Teknol. dan Sist. Informasi)*, vol. 4, no. 1, pp. 85–92, 2017, doi: 10.33330/jurteksiv4i1.28.
- [6] D. M. C. Hermanto, "Analisis Algoritma Clustering Dalam Kasus Penentuan Jenis Bunga Iris," *J. MEDIA Apl.*, vol. 9, no. 2, pp. 72–84, 2017.
- [7] S. A. Ritonga, M. Safii, I. Parlina, H. S. Tambunan, and S. Susiani, "Teknik Data Mining dalam Mengelompokkan Produktivitas Padi Menurut Provinsi Menggunakan K-Medoids," *Pros. Semin. Nas. Ris. Inf. Sci.*, vol. 1, no. 1, pp. 651–660, 2019, doi: 10.30645/senaris.v1i0.71.
- [8] I. Oktaria, "RapidMiner," (*online*), 2018. .