

KLUSTERISASI JUMLAH PENDERITA DEMAM BERDARAH DI KOTA INDONESIA MENGGUNAKAN ALGORITMA K-MEAN

Bibit Sudarsono¹, Umi Faddillah¹, Ayuni Asistiyasari¹, Yosep Nuryaman¹

¹Dosen Jurusan Sistem Informasi Universitas Bina Sarana Informatika
Jl. Kramat Raya No. 98 Senen. Jakarta Pusat

*Koresponden : bibit.bbs@bsi.ac.id

ABSTRAK

Salah satu penyakit paling menakutkan yang tiap tahunnya menyerang penduduk di wilayah Indonesia yaitu Penyakit Demam Berdarah Dengue (DBD). Penyakit Demam Berdarah Dengue (DBD) adalah penyakit yang ditularkan oleh jenis nyamuk *Aedes aegypti*. Penyakit yang penularannya cepat dan mudah ini perlu perencanaan khusus agar pencegahannya lebih optimal. Namun berdasarkan data yang dipublikasikan oleh BPS, terdapat lebih dari 100.000 penderita setiap tahunnya. Oleh sebab itu perlu pembagian atau pengklusteran terhadap kota-kota di Indonesia menjadi kota menjadi beberapa kluster untuk melihat kota yang menjadi prioritas utama dan sebagainya. Algoritma K-Mean merupakan algoritma yang mampu mengklusterisasi data. Berdasarkan penelitian-penelitian sebelumnya maka dilakukan perhitungan data penderita DBD menggunakan Algoritma K-Mean. Dan hasil akhir penelitian dengan rasio 0,000107574 dimana Jawa Tengah, Jawa Timur dan Jawa Barat masuk dalam C3 sehingga memiliki prioritas yang sangat tinggi dalam penanganan DBD.

Kata kunci : *Kmean, Data mining, Demam Berdarah*

1. PENDAHULUAN

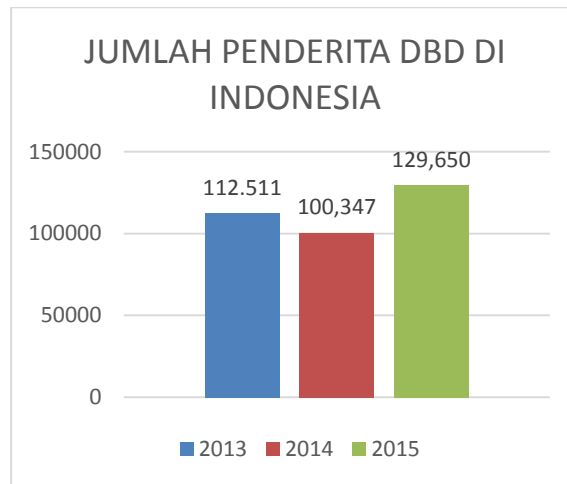
Demam Berdarah Dengue (DBD) adalah salah satu jenis penyakit endemik di wilayah tropis serta sebagian wilayah subtropic. Penyakit Demam Berdarah Dengue (DBD) adalah penyakit yang ditularkan oleh jenis nyamuk *Aedes aegypti* [1]. DBD menjadi salah satu penyakit yang menakutkan dikarenakan mudahnya penularan.

Hal tersebut sesuai dengan data yang dipublikasikan oleh Badan Pusat Statistik Indonesia dimana angka penderita DBD di Indonesia mencapai lebih dari 100.000 kasus per tahun pada tahun 2013-2015 [2].

Dikarenakan mudahnya penyebaran tersebut dibutuhkan perencanaan untuk mengendalikan penyebaran DBD diantaranya melalui fogging hingga kegiatan 4M. terutama Ketika musim hujan datang banyak genangan air yang jadi sarang nyamuk. Selain itu perlu gambaran jelas kota-kota mana saja yang menjadi prioritas dalam penekanan jumlah penderita DBD tersebut.

Oleh sebab itu diperlukan algoritma yang dapat mengklusterisasi kota-kota tersebut untuk menentukan suatu kota termasuk kluster tertentu. Salah satu algoritma yang dapat digunakan untuk pengklusterisasi yaitu Algoritma KMean[3].

Algoritma K-Mean mampu menghasilkan klusterisasi lebih baik dengan rasio 0,000249 dibandingkan algoritma AHP untuk menentukan curah hujan di berbagai kota di Indonesia [4].

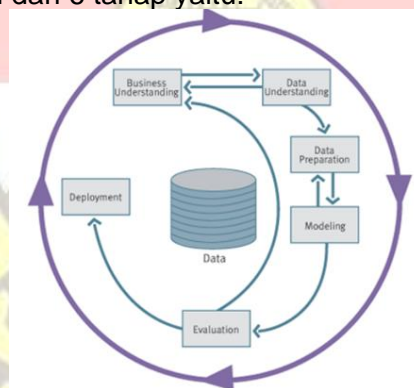


Gambar 1. Data Kasus Penyakit DBD

Walaupun algoritma AHP[5][6] dapat digunakan untuk menghasilkan model kebijakan mengenai penyakit malaria dan disusun dengan menggunakan 3 kriteria prioritas kebijakan.

2. METODE PENELITIAN

Penelitian ini menggunakan model CRISP-DM (Cross-Industry Standard Proses for Data Mining) [4], yang terdiri dari 6 tahap yaitu:



Gambar 2. CRISP-DM

1. Fase Pemahaman Bisnis (*Business/Research Understanding Phase*)

Kasus penyakit DBD di Indonesia cukup banyak di tahun 2013-2015 mencapai lebih 100.000 kasus di setiap tahunnya. Dengan jumlah tersebut perlu penggolongan terhadap kota-kota di Indonesia supaya penanganan pencegahannya lebih maksimal. Dengan pengelompokan tersebut maka nanti dapat ditentukan mana yang menjadi prioritas dari setiap clusternya.

2. Fase Pemahaman Data (*Data Understanding Phase*)

Data yang diambil yaitu data kasus DBD 24 Kota di Indonesia selama dari tahun 2013 hingga tahun 2015. Adapun data tersebut seperti yang tampak pada table dibawah ini

Tabel 1. Data Kasus Penyakit DBD

No	Provinsi	2013	2014	2015
1	Aceh	1.369	2.208	1.509
2	Sumatera Utara	3.223	5.378	5.274
3	Sumatera Barat	2.206	2.328	3.806
4	Riau	1.398	2.342	3.261
5	Jambi	638	1.308	1.354
6	Sumatera Selatan	1.436	1.500	3.396
27	Sulawesi Selatan	4.261	2.904	3.974
28	Sulawesi Tenggara	1.135	838	1.500
29	Gorontalo	238	223	231
30	Sulawesi Barat	500	315	486
31	Maluku	33	12	78
32	Maluku Utara	242	148	119
33	Papua Barat	48	77	66
34	Papua	152	431	647

3. Fase Pengolahan Data (*Data Preparation Phase*)

Pada fase ini dilakukan persiapan terhadap data yang ingin diolah, seperti normalisasi data berupa mengganti value – menjadi 0 agar rumus dapat dijalankan dengan baik.

4. Fase Pemodelan (*Modeling Phase*)

Pada tahap ini dilakukan pemodelan menggunakan algoritma K-Mean

5. Fase Evaluasi (*Evaluation Phase*)

Pada tahap ini pengujian model dilakukan dengan menggunakan MSE , yang nantinya hasil dari pengujian tersebut menjadi dasar seberapa baik hasil algoritma tersebut

6. Fase Penyebaran (*Deployment Phase*)

Setelah diketahui hasil klusterisasinya maka hal tersebut disampaikan kepada pihak berwenang sebagai bahan pertimbangan pengambil keputusan.

3. HASIL DAN PEMBAHASAN

Adapun langkah-langkah perhitungan yang dilakukan adalah sebagai berikut.

A. Penentuan Centroid Awal

Pada fase ini dilakukan pemilihan titik tengah awal yang dilakukan secara acak yaitu seperti table dibawah ini.

Tabel 2. Centroid Awal

K3	W	X	Y
C1	4573	1317	3145
C2	10156	8447	4954
C3	23118	18116	21237

Setelah centroid awal ditentukan, maka dilakukan perhitungan dengan rumus yang ada sehingga menghasilkan clustering seperti table dibawah ini

Tabel 3. Hasil Literasi 1

Data	Jarak ke Centroid			Keanggotaan
	C1	C2	C3	
1	3706	11314	33396	C1
2	4780	7589	28511	C1
3	2657	10098	31471	C1
4	3338	10809	32307	C1
5	4323	12431	34398	C1
11	9235	0	22949	C2
12	30878	22949	0	C3
13	19560	12757	11687	C3
14	1434	9552	31155	C1
15	21415	15928	12125	C3
16	1792	8462	30447	C1
17	10753	6654	21606	C2
30	4966	13391	35483	C1
31	5632	14050	36171	C1
32	5411	13804	35946	C1

Pada table diatas dapat dilihat ada 2 cabang yang masuk katerogri C2 dan 3 cabang masuk kategori C3 dan sisanya masuk C1. Rasio pada literasi 1 yaitu sebesar 0,000050367. Dan kemudian dilakukan perhitungan untuk mencari titik tengah yang baru yang akan digunakan pada literasi selanjutnya. Adapun setelah dilakukan perhitungan berikut adalah titik tengah yang baru

Tabel 4. Centroid Awal Literasi 2

	W	X	Y
C1	1562	1641	2056
C2	8485	8538	7829
C3	17719	12821	19258

Setelah dilakukan perhitungan dengan langkah-langkah sebelumnya, maka pada literasi ke 1 ini didapat pengklusterian sebagai berikut.

Tabel 5. Hasil Perhitungan Literasi 2

Data	Jarak ke Centroid			Keanggotaan
	C1	C2	C3	
1	811	11430	26362	C1
2	5204	6648	21473	C1
3	1987	9704	24280	C1
4	1404	10463	25141	C1
11	11339	3327	16761	C2
12	33226	22038	7817	C3
13	21890	11145	4226	C3
14	2246	9458	24006	C1
15	23727	13898	4620	C3
16	2932	8613	23435	C1
17	12296	3327	14480	C2
32	2779	14063	28850	C1

Berdasarkan hasil perhitungan yang tampak pada table IV, terdapat 2 kota yang masuk C2, 3 kota yang masuk C3 dan sisanya masuk C1. Nilai rasio yang dihasilkan yaitu 0,000107574 atau lebih besar dari rasio sebelumnya yaitu 0,000050367 sehingga perhitungan dilanjutkan ke literasi selanjutnya. Adapun Langkah sebelumnya yaitu mencari titik tengah yang baru dengan rumus yang ada sehingga hasilnya tampak seperti table dibawah ini.

Tabel 6. Centroid Awal Literasi 3

	W	X	Y
C1	1562	1641	2056
C2	8485	8538	7829
C3	17719	12821	19258

Setelah dilakukan perhitungan dengan langkah-langkah sebelumnya, maka pada literasi ke 2 ini didapat pengklusterian sebagai berikut.

Tabel6. Hasil Perhitungan Literasi 3

Data	Jarak ke Centroid			Keanggotaan
	C1	C2	C3	
1	811	11430	26362	C1
2	5204	6648	21473	C1
3	1987	9704	24280	C1
4	1404	10463	25141	C1
11	11339	3327	16761	C2

12	33226	22038	7817	C3
13	21890	11145	4226	C3
14	2246	9458	24006	C1
15	23727	13898	4620	C3
16	2932	8613	23435	C1
17	12296	3327	14480	C2
32	2779	14063	28850	C1

Berdasarkan hasil perhitungan yang tampak pada table IV, terdapat 2 kota yang masuk C2, 3 kota yang masuk C3 dan sisanya masuk C1. Nilai rasio yang dihasilkan yaitu 0,000107574 atau sama dengan rasio sebelumnya yaitu 0,000107574 sehingga perhitungan diakhiri.

4. KESIMPULAN DAN SARAN

Berdasarkan hasil perhitungan menggunakan Algoritma Kmean yang masuk C1 berjumlah 27 kota, C2 berjumlah 2 Kota dan C3 berjumlah 3 Kota. Maka dapat disimpulkan

1. Jawa Tengah, Jawa Timur dan Jawa Barat masuk dalam C3 sehingga memiliki prioritas yang sangat tinggi dalam penanganan DBD
2. Jakarta dan Bali masuk dalam C2 memiliki prioritas yang tinggi dalam penanganan DBD
3. Kota-kota lainnya tetap masuk kategori C1 umumnya jumlah kasusnya berada dibawah rata-rata kasus tiap tahunnya.

DAFTAR PUSTAKA

1. Syamsir & Daramusseng. A, 2018, *Analisi Spasial Efektivitas Fogging di Wilayah Kerja Puskesmas Makroman*, Kota Samarinda.. Jurnal Nasional Ilmu Kesehatan, vol. 1, no. 2,. 2018, pp. 1-7
2. Badan Pusat Statistik, *Jumlah Kasus Penyakit*. [Badan Pusat Statistik \(bps.go.id\)](https://bps.go.id). Diakses tanggal 20 Juli 2022.
3. G. Gustientiedina, M. H. Adiya, and Y. Desnelita, *Penerapan Algoritma K-Means Untuk Clustering Data Obat-Obatan*, J. Nas. Teknol. dan Sist. Inf., vol. 5, no. 1, pp. 17–24, 2019, doi: 10.25077/teknosi.v5i1.2019.17-24
4. Nuryaman, Y, 2017, *Komparasi Algoritma Kmean dan Ahc untuk Klasifikasi Curah Hujan di Indonesia*, Ikraith Informatika, vol. 2, no. 2, 1 Jul. 2018, pp. 70-75.
5. Deny , S., Agus , R., & Yushar , K, 2020, *Analisis Penentuan Tipe Fondasi Pilar Jembatan Dengan Menggunakan Metode Analytical Hierarchy Process (AHP)* (Studi Kasus: Pembangunan Jembatan Walahar Kecamatan Ciampel, Kabupaten Karawang)
6. Armaita, Dedi , H., Eri , B., Indang , D., & Iswandi, U, 2020, *Policy Model of Community Adaptation using AHP in the Malaria Endemic Region of Lahat Regency - Indonesia*. International Journal of Management and Humanities (IJMH), 44-48.