

Analisa Kebutuhan Tenaga Kesehatan Menggunakan Algoritma *K-Means*

Brins Leonard Pailan¹, Haviluddin², Masna Wati³, Novianti Puspitasari⁴, Edy Budiman⁵

Program Studi Informatika, Fakultas Teknik, Universitas Mulawarman

Jl. Sambaliung No. 9 Kampus Gn. Kelua, Samarinda, Kalimantan Timur, Indonesia 75123

Email: ¹brinsleonard21@gmail.com; ²haviluddin@unmul.ac.id; ³masnawati@fkti.unmul.ac.id;

⁴miechan.novianti@gmail.com; ⁵edy.budiman@fkti.unmul.ac.id

INFORMASI ARTIKEL

ABSTRAK

Histori Artikel

Diterima : 21 September 2020

Direvisi : 24 Februari 2021

Diterbitkan : 08 Maret 2021

Kata Kunci:

Clustering

Tenaga Kesehatan

K-Means

Penempatan tenaga kesehatan yang bersesuaian keahlian pada Pusat Kesehatan Masyarakat (Puskesmas) sangat diperlukan namun belum merata terdistribusi. Penelitian ini bertujuan untuk menganalisa tenaga kesehatan yang sesuai bidang keahliannya untuk ditempatkan pada setiap Puskesmas di Provinsi Kalimantan Timur. Tenaga kesehatan yang bersesuaian keahlian dianalisa untuk ditempatkan pada 179 Puskesmas di kawasan perkotaan, kawasan pedesaan, dan kawasan terpencil/sangat terpencil dengan menggunakan algoritma cerdas yaitu algoritma K-Means. Berdasarkan percobaan, Puskesmas terbagai ke dalam 4 kelompok terdiri dari kelompok 1 sebanyak 82 Puskesmas kawasan pedesaan dengan fasilitas non rawat inap; kelompok 2 sebanyak 23 Puskesmas kawasan perkotaan dengan fasilitas non rawat inap; kelompok 3 sebanyak 59 Puskesmas kawasan pedesaan dengan fasilitas rawat inap; dan kelompok 4 sebanyak 15 Puskesmas kawasan perkotaan dengan fasilitas rawat inap. Hasil penelitian memperlihatkan bahwa Puskesmas kelompok 1 masih kekurangan tenaga kesehatan dokter gigi dan ahli gizi; Puskesmas kelompok 2 masih kekurangan tenaga kesehatan farmasi; Puskesmas kelompok 3 masih kekurangan tenaga kesehatan ahli gizi; dan Puskesmas kelompok 4 sudah memenuhi standar minimal tenaga kesehatan. Hal ini menunjukkan bahwa pengelompokan Puskesmas tersebut dapat dijadikan acuan dalam mengambil kebijakan yang diharapkan dapat menjawab permasalahan distribusi penempatan tenaga kesehatan yang tidak merata di Provinsi Kalimantan Timur.

2021 SAKTI – Sains, Aplikasi, Komputasi dan Teknologi Informasi.

Hak Cipta.

I. Pendahuluan

Di negara Indonesia, Pusat Kesehatan Masyarakat (Puskesmas) adalah salah satu fasilitas pelayanan kesehatan yang memiliki peran sangat strategis dalam upaya mempercepat peningkatan derajat kesehatan masyarakat. Di Provinsi Kalimantan Timur sendiri terdapat 179 Puskesmas yang tercatat dalam Data Dasar Puskesmas Tahun 2018, namun banyaknya jumlah puskesmas tidak membuat masyarakat lebih memilih pelayanan kesehatan di Puskesmas karena dianggap kurang layak dari segi SDM/tenaga kesehatannya serta sarana prasarana yang dimiliki. Kurangnya jumlah tenaga kesehatan yang dimiliki tentu berpengaruh bagi tingkat kelayakan suatu Puskesmas. Berdasarkan Data Dasar Puskesmas Provinsi Kalimantan Timur Tahun 2018, dapat diketahui bahwa ada Puskesmas di Provinsi Kalimantan Timur yang tidak memiliki dokter umum, dokter gigi, tenaga farmasi, dan ahli gizi.

Masalah utama dalam pengelolaan tenaga kesehatan adalah distribusi SDM yang tidak merata. Untuk menunjang seluruh upaya pembangunan kesehatan diperlukan tenaga yang mempunyai sikap nasional, profesional, semangat pengabdian yang tinggi, berilmu dan terampil. Tenaga kesehatan dan tenaga penunjang perlu ditingkatkan kualitas, kemampuan serta persebarannya agar merata dan dapat mendukung penyelenggaraan pembangunan kesehatan disetiap tingkatan khususnya dalam mendukung pelaksanaan otonomi daerah.

Dari permasalahan tersebut, Pemerintah memiliki kewajiban meningkatkan kondisi kelayakan puskesmas di Provinsi Kalimantan Timur dengan mengambil beberapa kebijakan salah satunya dalam hal pendistribusian tenaga kesehatan. Kebijakan-kebijakan tersebut haruslah tepat sasaran, sehingga perlu dilakukan pengklasteran Puskesmas. Salah satu ilmu yang dapat digunakan adalah data mining khususnya Clustering K-Means.

Algoritma K-Means merupakan algoritma pengelompokan iteratif yang sederhana diimplementasikan, relatif cepat, dan mudah beradaptasi [1]. Puskesmas akan dikelompokkan agar pihak pemerintah sebagai pengambil keputusan dapat mengetahui kondisi setiap puskesmas berdasarkan tenaga kesehatannya.

Berbagai penelitian yang menggunakan algoritma K-Means telah dilakukan oleh peneliti. Peneliti [2], telah membandingkan algoritma K-Means dan Fuzzy C-Means Clustering dalam aktivitas lalu lintas jaringan, data yang digunakan berupa 4 dataset klien selama 152 hari dari ICT. Penelitian tersebut menghasilkan informasi bahwa Fuzzy C-Means lebih baik daripada K-means. Peneliti [3], telah menggunakan algoritma K-Means untuk mengelompokkan pemakaian obat pada puskesmas Lempake, data yang digunakan berupa laporan pemakaian 225 obat dari tahun 2016 sampai 2018. Penelitian tersebut menghasilkan 3 kelompok yaitu pemakaian obat tinggi, sedang, dan kurang. Penelitian tersebut juga menyatakan bahwa pengelompokan 3 cluster lebih baik dibandingkan 2 cluster. Peneliti [4], telah menggunakan algoritma K-Means untuk analisis penyakit menular manusia di Kabupaten Majalengka, data yang digunakan adalah data penyakit menular pada tahun 2014 yang didapatkan dari Dinas Kesehatan. Penelitian tersebut menghasilkan informasi yang dapat membantu pengendalian persediaan obat untuk penyakit menular di Kabupaten Majalengka. Peneliti [5], telah membandingkan algoritma K-Means dan K-Medoids terhadap kelayakan puskesmas di DIY, data yang digunakan berupa data sekunder 121 puskesmas di DIY. Penelitian tersebut menghasilkan informasi bahwa metode clustering K-means lebih baik daripada K-medoids. Peneliti [6], telah menggunakan algoritma K-Means untuk mengelompokkan curah hujan di Provinsi Kalimantan Timur, data yang digunakan berupa dataset curah hujan dari tahun 1986-2008 dan berasal dari 13 stasiun pengamatan BMKG Provinsi Kalimantan Timur. Penelitian tersebut berhasil membagi tingkat curah hujan ke dalam tiga kelompok yaitu rendah, sedang dan tinggi. Peneliti [7], Priyatman, H. dkk telah menggunakan algoritma K-Means untuk memprediksi waktu kelulusan mahasiswa menggunakan data akademik sebanyak 100 data. Implementasi algoritma K-Means pada data telah berhasil, namun untuk melihat tingkat kemampuan real-k-means clustering dalam memprediksi waktu kelulusan tergantung mahasiswa itu sendiri. Peneliti [8], M.W. Talakua, dkk telah menggunakan algoritma K-Means untuk mengelompokkan kota berdasarkan indikator Indeks Pembangunan Manusia (IPM). Data IPM diperoleh dari Badan Pusat Statistik (BPS) dengan jumlah data sebanyak 11 kota. Berdasarkan hasil penelitian, dari 11 kota hanya terdapat 1 kota saja yang memiliki nilai IPM tertinggi, sedangkan kelompok ketiga yaitu nilai IPM terendah memiliki anggota yang paling banyak. Artikel ini menggunakan algoritma K-Means untuk menganalisa kebutuhan tenaga kesehatan di Puskesmas. Tujuan penelitian ini adalah untuk membantu pemerintah dalam menempatkan tenaga kesehatan yang diperlukan oleh setiap Puskesmas sehingga pelayanan kesehatan kepada masyarakat dapat terlayani dengan baik dan tepat. Artikel ini terdiri dari motivasi penulisan artikel pada bagian pertama. Kedua, menerangkan model kerja algoritma K-Means. Ketiga, analisa hasil percobaan. Kesimpulan penelitian pada bagian akhir.

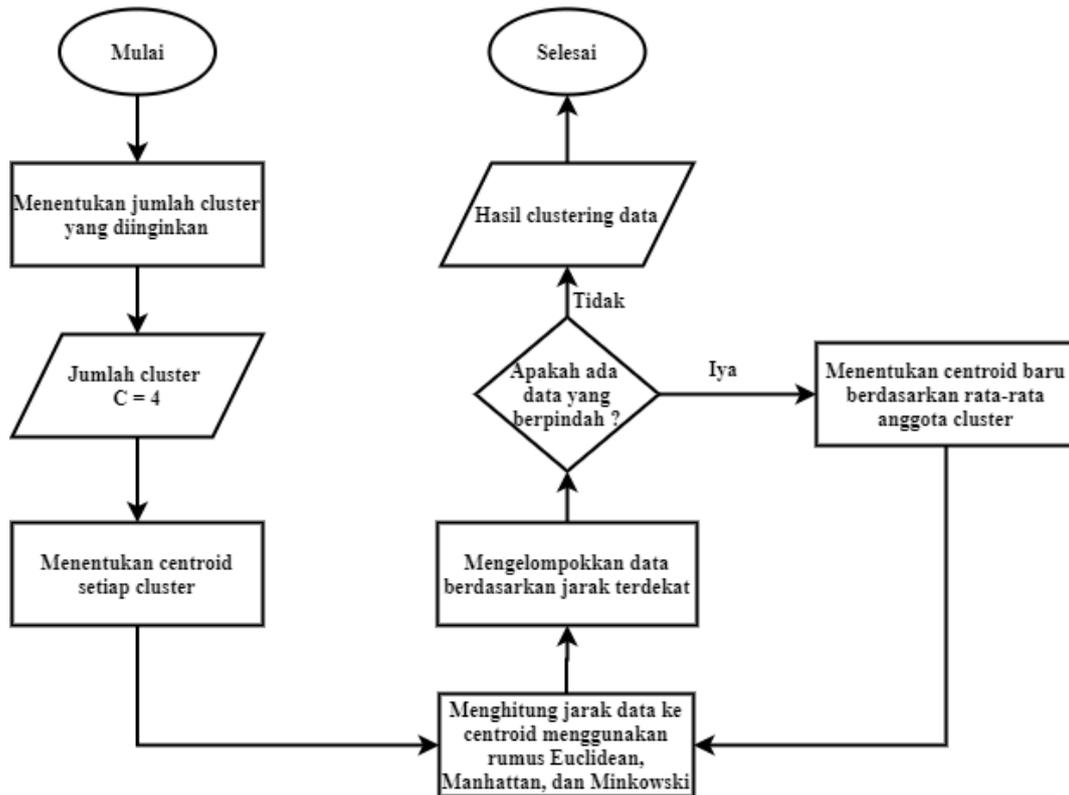
II. Material dan Metode

Dalam penelitian ini, penerapan algoritma K-Means digunakan untuk menganalisa kondisi kebutuhan tenaga kesehatan Puskesmas di Provinsi Kalimantan Timur. Berikut dijelaskan model penerapan algoritma K-Means.

A. Algoritma K-Means

Algoritma K-Means clustering merupakan salah satu metode yang melakukan pengelompokan data dengan sistem partisi. Metode K-Means berusaha mengelompokkan data ke dalam beberapa kelompok yang mempunyai karakteristik yang sama. Secara prinsip, metode K-Means memiliki K sebagai jumlah cluster yang diinginkan. Sedangkan, Means yaitu rata-rata dari suatu kelompok data yang didefinisikan sebagai cluster [9].

Secara umum, metode K-Means memiliki empat tahapan proses terdiri dari (1) menentukan jumlah cluster; (2) mengalokasikan data ke dalam cluster secara random; (3) menghitung centroid/rata-rata dari data yang ada di masing-masing cluster; dan (4) mengalokasikan masing-masing data ke centroid/rata-rata terdekat, jika masih ada data yang berpindah cluster atau apabila perubahan nilai centroid kembali ke tahap tiga [10], [2]. Alur dan penjelasan algoritma K-Means dapat dilihat pada Gambar 1.



Gambar 1. Flowchart K-Means

Algoritma K-Means dimulai dengan menetapkan jumlah cluster dan menentukan titik pusat awal masing-masing cluster. Kemudian, menghitung jarak antara setiap titik data dan ke pusat cluster. Dalam penelitian ini tiga metode perhitungan jarak terdiri dari Euclidean Distance, Manhattan Distance, dan Minkowski Distance telah digunakan untuk melihat keakuratan sebaran data. Rumus metode *Euclidean Distance* ditunjukkan pada persamaan 1, *Manhattan Distance* pada persamaan 2, dan *Minkowski Distance* pada persamaan 3 [11].

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

$$\text{Manhattan Distance} \quad d(x, y) = \sum_{i=1}^n |x_i - y_i| \quad (2)$$

$$\text{Minkowski Distance} \quad d(x, y) = \left(\sum_{i=1}^n |x_i - y_i|^3 \right)^{1/3} \quad (3)$$

Dimana, d adalah jarak antara x dan y ; x adalah data pusat klaster; y adalah data pada atribut; i adalah setiap data; n adalah jumlah data; x_i adalah data pada pusat klaster ke i ; y_i adalah data pada setiap data ke i . Kemudian, menentukan titik data ke pusat cluster yang jaraknya dari pusat cluster paling dekat dari semua pusat cluster. Dilanjutkan dengan menghitung pusat cluster baru menggunakan persamaan 4.

$$C(i) = \frac{x_1 + x_2 + x_{...} + x_{...}}{\sum x} \quad (4)$$

Dimana, X_1 adalah nilai data record ke-1; X_2 adalah nilai data record ke-2; $\sum x$ adalah jumlah data record. Tahap akhir, menghitung jarak antara setiap titik data dan pusat cluster yang baru diperoleh, jika tidak ada titik data yang ditugaskan kembali maka proses berhenti, jika tidak, ulangi langkah 3 sampai 5.

B. Perhitungan Akurasi Jarak

Dalam penelitian ini, pengukuran jarak metode Euclidean, Manhattan, dan Minkowski telah menggunakan nilai simpangan baku, yaitu simpangan baku dalam cluster (S_w) dan simpangan baku antar cluster (S_B). Metode pengukuran jarak terbaik mempunyai nilai rasio simpangan baku dalam cluster (S_w) dan simpangan baku antar cluster (S_B) yang paling kecil. Semakin kecil nilai S_w dan semakin besar nilai S_B maka metode tersebut memiliki kinerja yang baik, artinya mempunyai homogenitas yang tinggi [12]. Simpangan baku dalam cluster (S_w) dan simpangan baku antar cluster (S_B) menggunakan Persamaan 5 dan 6.

$$S_w = K^{-1} \sum_{k=1}^K S_k \quad (5)$$

Dimana, K adalah banyaknya cluster yang terbentuk, S_k adalah simpangan baku cluster ke- k .

$$S_B = [(K - 1)^{-1} \sum_{k=1}^K (\bar{X}_k - \bar{X})^2]^{\frac{1}{2}} \quad (6)$$

Dimana, \bar{X}_k adalah rata-rata cluster ke- k , \bar{X} adalah rata-rata keseluruhan cluster. Sedangkan, rasio simpangan baku dalam dan antar cluster menggunakan Persamaan 7.

$$\text{Rasio} = \frac{S_w}{S_B} \quad (7)$$

Oleh karena itu, jarak terbaik dari ketiga metode adalah jarak yang memiliki nilai rasio terkecil. Hal ini berarti bahwa cluster yang baik adalah mempunyai homogenitas (kesamaan) yang tinggi antar anggota dalam satu cluster (within cluster) dan heterogenitas yang tinggi antar cluster yang satu dengan cluster yang lain (between cluster) [13], [14], [15].

C. Puskesmas

Dalam penelitian ini, analisa kebutuhan tenaga kesehatan dilakukan pada 179 Puskesmas yang berada di Provinsi Kalimantan Timur. Kemudian, terdapat 10 jenis tenaga kesehatan dengan 3 kawasan Puskesmas yaitu perkotaan, pedesaan, dan terpencil/sangat terpencil. Dimana, jenis tenaga kesehatan pada Puskesmas telah ditetapkan dalam Peraturan Menteri Kesehatan Republik Indonesia Nomor 75 Tahun 2014. Adapun, standar ketenagaan Puskesmas dapat dilihat pada Tabel 1.

Tabel 1. Standar Ketenagaan Puskesmas

No	Jenis Tenaga	Kawasan					
		Perkotaan		Pedesaan		Terpencil/Sangat Terpencil	
		Non Rawat Inap	Rawat Inap	Non Rawat Inap	Rawat Inap	Non Rawat Inap	Rawat Inap
1	Dokter Umum	1	2	1	2	1	2
2	Dokter Gigi	1	1	1	1	1	1
3	Perawat	5	8	5	8	5	8
4	Bidan	4	7	4	7	4	7
5	Tenaga Kesehatan Masyarakat	2	2	1	1	1	1
6	Tenaga Kesehatan Lingkungan	1	1	1	1	1	1
7	Ahli Teknologi Laboratorium Medik	1	1	1	1	1	1
8	Tenaga Gizi	1	2	1	2	1	2
9	Tenaga Kefarmasian	1	2	1	1	1	1
10	Tenaga Pendukung	5	5	3	3	3	3

III. Hasil dan Pembahasan

Pada sub bagian ini, akan diuraikan proses kerja clustering K-Means yang diterapkan pada data tenaga kesehatan Puskesmas Provinsi Kalimantan Timur. Proses pertama pada perhitungan adalah menentukan banyaknya cluster yang akan dibentuk. Penetapan jumlah cluster pada penelitian ini sebanyak 4 cluster yaitu Puskesmas kawasan Pedesaan dengan fasilitas Non Rawat Inap (C1), Puskesmas kawasan Perkotaan dengan fasilitas Non Rawat Inap (C2), Puskesmas kawasan Pedesaan dengan fasilitas Rawat Inap (C3), dan

Puskesmas kawasan Perkotaan dengan fasilitas Rawat Inap (C4). Kemudian menentukan titik centroid awal masing-masing cluster dengan mengacu pada Permenkes RI Nomor 75 Tahun 2014. Penentuan titik centroid awal dapat dilihat pada Tabel 2.

Tabel 2. Centroid Awal

C1	1	1	5	4	1	1	1	1	1	3
C2	1	1	5	4	2	1	1	1	1	5
C3	2	1	8	7	1	1	1	2	1	3
C4	2	1	8	7	2	1	1	2	2	5

Proses selanjutnya adalah menghitung jarak setiap data dengan centroid setiap cluster menggunakan perhitungan jarak Euclidean, Manhattan, dan Minkowski. Pada tahap ini telah dilakukan penentuan jarak setiap Puskesmas ke pusat cluster C1, C2, C3, dan C4 menggunakan Persamaan (1).

Jarak C1

$$d(1,1) = \sqrt{(2-1)^2 + (1-1)^2 + (13-5)^2 + (16-4)^2 + (0-1)^2 + (2-1)^2}$$

$$\sqrt{(2-1)^2 + (1-1)^2 + (0-1)^2 + (5-3)^2}$$

$$d(1,1) = 14,731$$

Jarak C2

$$d(1,2) = \sqrt{(2-1)^2 + (1-1)^2 + (13-5)^2 + (16-4)^2 + (0-2)^2 + (2-1)^2}$$

$$\sqrt{(2-1)^2 + (1-1)^2 + (0-1)^2 + (5-5)^2}$$

$$d(1,2) = 14,697$$

Jarak C3

$$d(1,3) = \sqrt{(2-2)^2 + (1-1)^2 + (13-8)^2 + (16-7)^2 + (0-1)^2 + (2-1)^2}$$

$$\sqrt{(2-1)^2 + (1-2)^2 + (0-1)^2 + (5-3)^2}$$

$$d(1,3) = 10,724$$

Jarak C4

$$d(1,4) = \sqrt{(2-2)^2 + (1-1)^2 + (13-8)^2 + (16-7)^2 + (0-2)^2 + (2-1)^2}$$

$$\sqrt{(2-1)^2 + (1-2)^2 + (0-2)^2 + (5-5)^2}$$

$$d(1,4) = 10,817$$

Berdasarkan perhitungan jarak kelompok C1, C2, C3, C4 didapatkan hasil jarak data ke-1 dengan centroid kelompok C1 sebesar 14,731, centroid kelompok C2 sebesar 14,697, centroid kelompok C3 sebesar 10,724, dan centroid kelompok C4 sebesar 10,817. Sehingga Puskesmas dikelompokkan ke dalam cluster berdasarkan jarak nilai terkecil. Selanjutnya, mencari nilai centroid baru menggunakan Persamaan 4, sehingga didapatkan nilai centroid yang ditampilkan pada Tabel 3.

Tabel 3. Nilai centroid baru

C1	2	1	3	2,5	0,5	0,5	1	0,5	1	3,5
C2	2,4	1,1	3,6	3,5	0,5	0,9	1,3	0,8	1	7,7
C3	1,5	1,1	12,5	12,2	1,1	1,2	1,4	1	1	3,4
C4	3,1	1,1	17,9	16,6	2,4	1,3	2,7	1,5	1,6	11,2

Tabel 3 memperlihatkan bahwa nilai centroid baru dan centroid lama berbeda sehingga proses perhitungan jarak harus berlanjut ke iterasi selanjutnya. Proses pengukuran jarak Euclidean berhenti pada iterasi ke-14, jarak Manhattan berhenti pada iterasi ke-15, dan jarak Minkowski berhenti pada iterasi ke-22. Nilai ketiga jarak dapat dilihat pada Tabel 4.

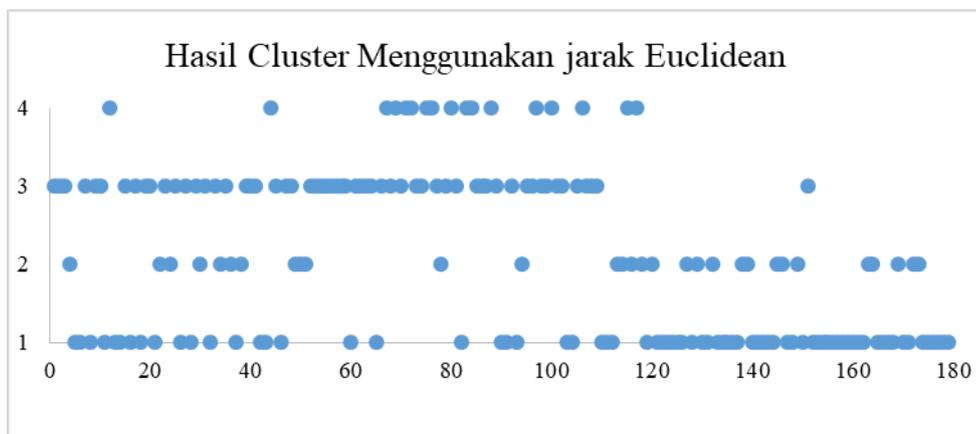
Tabel 4. Nilai Euclidean, Manhattan dan Minkowski pada iterasi terakhir

No	Nama Puskesmas	Euclidean Distance				Manhattan Distance				Minkowski Distance			
		C1	C2	C3	C4	C1	C2	C3	C4	C1	C2	C3	C4
1	Batu Kajang	11,03	14,55	7,704	36,994	17,878	30,522	24,627	76,2	10,041	10,498	7,350	32,831
2	Muser	9,526	12,172	6,84	37,245	13,073	27,304	23,678	76,2	8,820	8,366	7,180	33,237
3	Kerang	12,19	10,375	7,5	35,733	17,683	24,304	16,678	69,2	11,657	6,634	7,857	32,429
4	Tanjung Aru	10,345	8,845	9,976	38,616	18,61	21,304	23,169	76,867	9,233	5,265	9,601	34,695
5	Suatang Baru	7,608	13,492	10,971	40,97	8,902	21,957	26,678	79,2	7,686	9,620	11,885	36,856
6	Pasir Belengkong	7,084	9,898	9,442	39,753	12,341	20	24,525	76,2	6,432	5,842	9,689	35,717
7	Suliliran Baru	13,461	11,403	6,487	34,609	20,537	29,435	16,271	67,2	12,968	7,277	7,664	31,421
8	Padang Pengrapat	6,401	11,14	10,109	40,499	11,634	21,435	26,373	77,2	5,905	6,981	10,339	36,368
9	Tanah Grogot	20,316	11,794	7,065	27,041	36,122	34,913	11,051	55,867	17,893	9,446	5,905	24,386
10	Kuaro	21,525	15,899	6,493	25,185	34,415	41,435	13,898	53,2	19,511	11,751	4,033	22,553
179	Lok Bahu	6,816	20,193	22,085	52,727	16,195	32,261	47,678	100,2	5,213	15,098	20,993	47,079

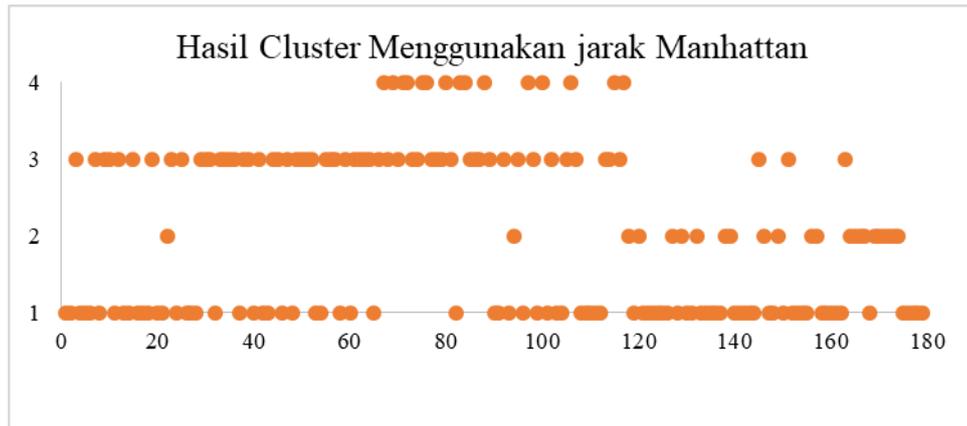
Setelah mendapatkan nilai Euclidean Distance maka proses selanjutnya adalah menentukan cluster berdasarkan nilai jarak terkecil. Hasil cluster setiap data dapat ditampilkan pada Tabel 5 yang menunjukkan bahwa terdapat 4 cluster pada setiap jarak. Euclidean Distance telah memperoleh kelompok C1 sebanyak 73 Puskesmas, kelompok C2 sebanyak 30 Puskesmas, kelompok C3 sebanyak 59 Puskesmas, dan kelompok C4 sebanyak 17 Puskesmas. Manhattan Distance telah memperoleh kelompok C1 sebanyak 82 Puskesmas, kelompok C2 sebanyak 23 Puskesmas, kelompok C3 sebanyak 59 Puskesmas, dan kelompok C4 sebanyak 15 Puskesmas. Sedangkan, Minkowski Distance telah memperoleh kelompok C1 sebanyak 67 Puskesmas, kelompok C2 sebanyak 50 Puskesmas, kelompok C3 sebanyak 46 Puskesmas, dan kelompok C4 sebanyak 16 Puskesmas.

Tabel 5. Perbandingan Hasil Cluster Ketiga Jarak

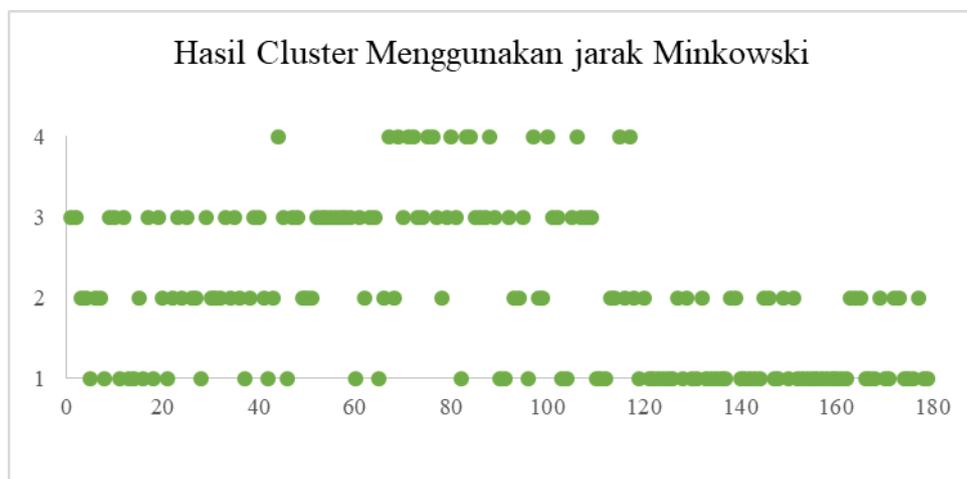
No	Nama Puskesmas	Euclidean Distance				Manhattan Distance				Minkowski Distance			
		C1	C2	C3	C4	C1	C2	C3	C4	C1	C2	C3	C4
1	Batu Kajang			*		*						*	
2	Muser			*		*						*	
3	Kerang			*			*				*		
4	Tanjung Aru		*			*					*		
5	Suatang Baru	*				*				*			
6	Pasir Belengkong	*				*				*			
7	Suliliran Baru			*			*				*		
8	Padang Pengrapat	*				*				*			
9	Tanah Grogot			*			*				*		
10	Kuaro			*			*				*		
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
179	Lok Bahu	*				*				*			
Jumlah		73	30	59	17	82	23	59	15	67	50	46	16



Gambar 2. Grafik hasil cluster menggunakan jarak Euclidean

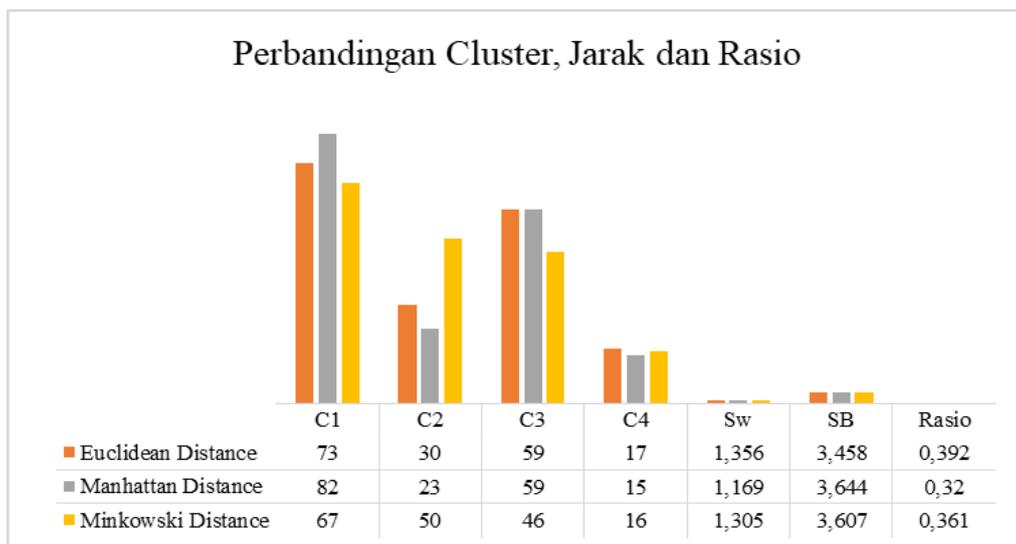


Gambar 3. Grafik hasil cluster menggunakan jarak *Manhattan*



Gambar 4. Grafik hasil cluster menggunakan jarak *Minkowski*

Dalam penelitian ini, Rasio Simpangan baku dalam cluster (S_w) dan Simpangan baku antar cluster (S_B) telah digunakan untuk melihat akurasi terbaik ketiga metode dengan menggunakan Persamaan (7).



Gambar 5. Perbandingan Cluster, Jarak dan Rasio

Gambar 5 didapatkan bahwa Manhattan Distance merupakan metode pengukuran jarak terbaik dengan nilai rasio sebesar 0.32 dibanding Euclidean Distance sebesar 0.392 dan Minkowski Distance sebesar 0.361. Oleh sebab itu, hasil dari clustering dengan metode pengukuran jarak Manhattan Distance dapat digunakan sebagai model perhitungan. Hasil uji ini juga telah dibandingkan dengan data dasar sehingga sebanyak 101 Puskesmas dari 179 Puskesmas mempunyai kelompok yang sama dengan tingkat akurasi sebesar 56.42%. Hal ini berarti penyebaran tenaga kesehatan Puskesmas di Provinsi Kalimantan Timur masih belum merata.

Dari ketiga perhitungan jarak telah diperoleh jarak Manhattan Distance memiliki tingkat akurasi yang baik sehingga dijadikan acuan dalam penyebaran tenaga kesehatan. Sebagai contoh dokter umum pada kelompok C1 bernilai 2.195. Jadi bisa disimpulkan bahwa puskesmas yg berada pada kelompok C1 memiliki rata-rata 2 dokter per Puskesmasnya, Tabel 6.

Tabel 6. Centroid akhir Manhattan Distance

Cluster	Dokter Umum	Dokter Gigi	Perawat	Bidan	Tenaga Kesehatan Masyarakat	Tenaga Kesehatan Lingkungan	Ahli Teknologi Laboratorium Medik	Tenaga Gizi	Tenaga Kefarmasian	Tenaga Penunjang
C1	2,195	0,963	9,146	8,390	1,024	1,610	1,171	0,988	1,402	7,232
C2	4,870	1,609	9,870	8,609	1,391	3,522	1,652	1,043	1,783	15,913
C3	2,780	1,051	20,475	18,881	1,525	2,881	1,576	1,627	2,627	12,254
C4	5,267	2,133	41,133	40,933	2,067	4,133	3,067	3,667	4,133	11,667

Sehingga dapat disimpulkan bahwa puskesmas-puskesmas yang berada pada kelompok C1 masih kekurangan tenaga kesehatan dokter gigi dan ahli gizi, dimana nilai rata-rata pada variabel tersebut masih dibawah nilai standar yaitu minimal 1 orang. Sedangkan untuk variabel lainnya sudah memenuhi standar, yang berarti bahwa 8 (delapan) tenaga kesehatan lainnya sudah terpenuhi. Sedangkan Puskesmas-puskesmas yang berada pada kelompok C2 masih kekurangan tenaga farmasi dan kelompok C3 masih kekurangan ahli gizi yang memiliki nilai rata-rata pada variabel masih dibawah nilai standar yaitu minimal 2 orang. Sedangkan, Puskesmas-puskesmas yang berada pada kelompok C4 yaitu Puskesmas-puskesmas di kawasan perkotaan dengan fasilitas rawat inap dapat dilihat sudah memenuhi standar minimal tenaga kesehatan. Hal ini bersesuaian dengan Permenkes RI Nomor 75 Tahun 2014.

IV. Kesimpulan

Penerapan algoritma K-Means dalam menganalisa tenaga kesehatan Puskesmas di Provinsi Kalimantan Timur telah dikerjakan. Berdasarkan hasil percobaan telah didapatkan kelompok C1 yaitu Puskesmas kawasan pedesaan dengan fasilitas non rawat inap yang beranggotakan 82 Puskesmas, kelompok C2 yaitu Puskesmas kawasan perkotaan dengan fasilitas non rawat inap yang beranggotakan 23 Puskesmas, kelompok C3 yaitu Puskesmas kawasan pedesaan dengan fasilitas rawat inap yang beranggotakan 59 Puskesmas, dan kelompok C4 yaitu Puskesmas kawasan perkotaan dengan fasilitas rawat inap yang beranggotakan 15 Puskesmas. Sedangkan, ketiga metode pengukuran Euclidean Distance, Manhattan Distance, dan Minkowski Distance menggunakan rasio S_w dan S_B telah menghasilkan nilai terkecil untuk Manhattan Distance yaitu 0,32 dibandingkan dengan Euclidean Distance yaitu 0,392 dan Minkowski Distance yaitu 0,361. Dengan kata lain, algoritma K-Means menggunakan Manhattan Distance lebih baik daripada Euclidean Distance dan Minkowski Distance. Berdasarkan hasil analisis pengelompokan data tenaga kesehatan setiap Puskesmas di Provinsi Kalimantan Timur Tahun 2018, diharapkan Dinas Kesehatan Kota maupun Pemerintah Provinsi Kalimantan Timur dapat menjadikan bahan pendukung untuk membuat keputusan atau kebijakan mengenai pendistribusian tenaga kesehatan di setiap Puskesmas. Komparasi dan optimalisasi algoritma K-Means untuk menghasilkan berbagai akurasi menjadi penelitian selanjutnya.

Daftar Pustaka

- [1] S. Ghosh and S. K. Dubey, "A Comparative Analysis of Fuzzy C-Means Clustering and K Means Clustering Algorithms," *Int. J. Comput. Eng. Res.*, vol. 2, no. 3, pp. 737-739, 2012.
- [2] Purnawansyah and Haviluddin, "K-Means clustering implementation in network traffic activities," in *Proceedings - CYBERNETICSCOM 2016: International Conference on Computational Intelligence and Cybernetics*, 2017. doi: 10.1109/CyberneticsCom.2016.7892566.
- [3] O. A. Murtioso, "Clustering Pemakaian Obat Menggunakan Algoritma K-Means Pada Puskesmas Lempake," vol. X, no. X, pp. 1-7, 2019.
- [4] A. Bastian, H. Sujadi, and G. Febrianto, "Penerapan Algoritma K-Means Clustering Analysis Pada Penyakit Menular Manusia (Studi Kasus Kabupaten Majalengka)," *J. Sist. Inf.*, vol. 14, no. 1, 2018.
- [5] Y. H. Susanti and E. Widodo, "Perbandingan K-Means dan K-Medoids Clustering terhadap Kelayakan Puskesmas di DIY Tahun 2015," vol. 1, no. 1, pp. 116-122, 2017.

- [6] N. Puspitasari and Haviluddin, "Penerapan metode k-means dalam pengelompokkan curah hujan di kalimantan timur," *Proc. SNRIK 2016*, vol. Volume 1, 2016.
- [7] H. Priyatman, F. Sajid, and D. Haldivany, "Klasterisasi Menggunakan Algoritma K-Means Clustering untuk Memprediksi Waktu Kelulusan Mahasiswa," *J. Edukasi dan Penelit. Inform.*, vol. 5, no. 1, p. 62, 2019, doi: 10.26418/jp.v5i1.29611.
- [8] M. W. Talakua, Z. A. Leleury, and A. W. Taluta, "Analisis Cluster Dengan Menggunakan Metode K-Means Untuk Pengelompokkan Kabupaten/Kota Di Provinsi Maluku Berdasarkan Indikator Indeks Pembangunan Manusia Tahun 2014," *BAREKENG J. Ilmu Mat. dan Terap.*, vol. 11, no. 2, pp. 119–128, 2017, doi: 10.30598/barekengvol11iss2pp119-128.
- [9] F. Nasari and S. Darma, "Seminar Nasional Teknologi Informasi dan Multimedia 2015 PENERAPAN K-MEANS CLUSTERING PADA DATA PENERIMAAN MAHASISWA BARU (STUDI KASUS : UNIVERSITAS POTENSI UTAMA)."
- [10] Yudi Agusta, "K-Means – Penerapan, Permasalahan dan Metode Terkait," *J. Sist. dan Inform.*, vol. 3, no. Februari, pp. 47–60, 2007.
- [11] M. Nishom, "Perbandingan Akurasi Euclidean Distance, Minkowski Distance, dan Manhattan Distance pada Algoritma K-Means Clustering berbasis Chi-Square," *J. Inform. J. Pengemb. IT*, vol. 4, no. 1, pp. 20–24, 2019, doi: 10.30591/jpit.v4i1.1253.
- [12] M. J. Bunkers, "Definition of Climate Regions in the Nothern Plains Using an Objective Cluster Modification Technique," *J. Clim.*, 1996, doi: 10.1017/CBO9781107415324.004.
- [13] J. Pérez-Ortega, N. N. Almanza-Ortega, A. Vega-Villalobos, R. Pazos-Rangel, C. Zavala-Díaz, and A. Martínez-Rebollar, "The K-Means Algorithm Evolution," in *In Introduction to Data Science and Machine Learning*, IntechOpen, 2019.
- [14] A. Agrawal and H. Gupta, "Global K-Means (GKM) Clustering Algorithm: A Survey," *Int. J. Comput. Appl.*, vol. 79, no. 2, pp. 0975–8887, 2013, doi: 10.5120/13713-1472.
- [15] K. Kaur, D. S. Dhaliwal, and R. K. Vohra, "Statistically Refining the Initial Points for K-means Clustering Algorithm," *Int. J. Adv. Res. Comput. Eng. Technol.*, vol. 2, no. 11, pp. 2972–2977, 2013.