

# Clustering Fasilitas Kesehatan Berdasarkan Kecamatan di Jakarta dengan Algoritma K-Medoids

Yahya Aliya Rohim<sup>1\*</sup>, Tariska Zidny Fatikhah<sup>2</sup>, Ramadhan Qodri Hardiansyah<sup>3</sup>, Dwi Hartanti<sup>4</sup>

<sup>1,2,3,4</sup>Teknik Informatika  
Universitas Duta Bangsa Surakarta

<sup>1\*</sup>202030334@mhs.udb.ac.id, <sup>2</sup>202030048@mhs.udb.ac.id, <sup>3</sup>202020430@mhs.udb.ac.id, <sup>4</sup>dwhartanti@udb.ac.id

*Abstrak*— Pada penelitian ini, kami menggunakan algoritma K-Medoids untuk melakukan clustering (pengelompokan) fasilitas kesehatan berdasarkan kecamatan di Jakarta pada tahun 2020. Tujuan dari penelitian ini adalah untuk mengidentifikasi pola-pola kemiripan antara kecamatan berdasarkan jenis dan jumlah fasilitas kesehatan yang tersedia di setiap kecamatan. K-Medoids adalah salah satu metode clustering yang berfokus pada pemilihan titik-titik tengah yang mewakili setiap kelompok. Algoritma ini bekerja dengan menginisialisasi k-titik awal sebagai medoids, kemudian mengalokasikan setiap data ke medoid terdekat dan menghitung total jarak antara data dan medoid. Proses ini diulang untuk mencari kombinasi medoid yang memiliki total jarak minimum. Hasilnya adalah partisi data ke dalam kelompok-kelompok yang saling terpisah. Data yang digunakan dalam penelitian ini mencakup informasi tentang jenis fasilitas kesehatan (seperti rumah sakit, klinik, dan puskesmas) yang ada di setiap kecamatan di Jakarta pada tahun 2020. Data ini diperoleh dari sumber terpercaya yang menyediakan data terkait kesehatan di Jakarta. Setelah mengimplementasikan algoritma K-Medoids pada data fasilitas kesehatan di setiap kecamatan, kami memperoleh hasil clustering yang menunjukkan kelompok-kelompok kecamatan yang memiliki pola kemiripan dalam hal jenis dan jumlah fasilitas kesehatan yang tersedia. Hasil ini dapat memberikan wawasan yang berharga bagi pemangku kepentingan dalam perencanaan dan pengembangan sistem kesehatan di Jakarta.

*Kata kunci*— Clustering Fasilitas Kesehatan, Jakarta, Algoritma K-Medoids.

*Abstract*— In this study, we used the K-Medoids algorithm to cluster health facilities by sub-district in Jakarta in 2020. The purpose of this study was to identify patterns of similarity between sub-districts based on the type and number of health facilities available in each sub-district. K-Medoids is one of the clustering methods that focuses on selecting the midpoints that represent each group. The algorithm works by initializing the starting k-points as medoids, then allocating each data to the nearest medoid and calculating the total distance between the data and the medoid. This process is repeated to look for medoid combinations that have a minimum total distance. The result is partitioning data into mutually separate groups. The data used in this study included information on the types of health facilities (such as hospitals, clinics, and public health center) in each sub-district in Jakarta in 2020. This data is obtained from reliable sources that provide health-related data in Jakarta. After implementing the K-Medoids algorithm on health facility data in each sub-district, we obtained clustering results that showed sub-district groups that had similar patterns in terms of the type and number of health facilities available. These results can provide valuable insights for stakeholders in health system planning and development in Jakarta.

*Keywords*— Health Facility Clustering, Jakarta, K-Medoids Algorithm.

## I. PENDAHULUAN

Isu kesehatan pada masyarakat khususnya di negara berkembang seperti Indonesia, dipengaruhi oleh dua aspek, yakni aspek fisik serta aspek non-fisik. Faktor fisik meliputi fasilitas kesehatan dan perawatan penyakit, sedangkan faktor non-fisik melibatkan aspek kesehatan manusia seperti pengetahuan, sikap, dan perilaku [1]. Dalam pembangunan kesehatan, pelayanan kesehatan dianggap sebagai hak dasar masyarakat yang harus dijamin. [2]. Tujuan dari pembangunan kesehatan adalah untuk meningkatkan kesadaran, motivasi, dan kemampuan individu dalam menjalani gaya hidup sehat bertujuan untuk mencapai tingkat kesehatan masyarakat yang optimal. Namun, perlu diperhatikan bahwa fasilitas

kesehatan di Indonesia, terutama di beberapa kecamatan di Jakarta, masih tergolong kurang. Misalnya, rumah sakit bersalin, rumah sakit umum, klinik, dan fasilitas pengobatan lainnya mungkin belum mencukupi. Kurangnya unit kesehatan di beberapa kecamatan Jakarta ini dapat menyulitkan masyarakat dalam menjalani hidup sehat dan mendapatkan pengobatan yang dibutuhkan. Sebagai ilustrasi, seorang perempuan yang sedang mengandung dan akan menjalani proses persalinan mungkin akan mengalami kesulitan jika sarana dan fasilitas yang tersedia tidak memadai. Karena itu, diharapkan pemerintah dapat memberikan solusi atas kondisi fasilitas kesehatan yang belum memadai di Kota Jakarta. Penelitian ini memiliki tujuan untuk melakukan pengelompokan wilayah

unit kesehatan secara cepat dan efektif. Metode yang digunakan dalam penelitian ini adalah metode pengelompokan, yang memiliki tujuan untuk mengelompokkan data menjadi beberapa kluster yang mempunyai kesamaan sifat. Memaksimalkan kesamaan antar data dalam satu kluster dan meminimalkan kesamaan antara data dari kluster yang berbeda adalah prinsip dasar dari clustering. Analisis clustering mempunyai fungsi sebagai alat untuk memisahkan objek ke dalam kelompok yang memiliki sifat berbeda antar kelompoknya.

## II. METODOLOGI PENELITIAN

Penelitian ini menggunakan data fasilitas kesehatan di Kota Jakarta dan mengaplikasikan algoritma K-Medoids untuk melakukan pengelompokan. Tahapan klustering fasilitas kesehatan dengan algoritma K-Medoids, antara lain:

### A. Pengumpulan Data

Tahap pertama yaitu melakukan pengumpulan data. Data tersebut diperoleh dari situs resmi Open Data Jakarta ([data.jakarta.go.id](http://data.jakarta.go.id)), yang merupakan sumber data terbuka yang disediakan oleh Pemerintah Kota Jakarta.

### B. Data Selection

Sebelum memulai proses data mining, langkah penting yang perlu dilakukan adalah membersihkan data. Proses pembersihan data (data cleaning) meliputi tindakan seperti menghapus data duplikat, memeriksa dan memperbaiki ketidaksesuaian data, serta memperbaiki kesalahan penulisan. Dalam konteks ini, terdapat beberapa atribut yang tidak akan digunakan dalam pengolahan data, yaitu tahun, lokasi, dan kecamatan. Oleh karena itu, atribut- atribut tersebut dapat dihapus sehingga hanya menyisakan dua atribut yang akan digunakan dalam analisis, yaitu "fasilitas kesehatan" dan "jumlah fasilitas kesehatan".

### C. Data Transformasi

Tahap ini melibatkan perubahan bentuk data asli menjadi bentuk yang lebih sesuai untuk mempermudah dalam pengolahan data. Salah satu contoh transformasi adalah menginisialisasi atribut fasilitas kesehatan dari data nominal menjadi

bentuk numerikal. Inisialisasi ini dapat dilakukan dengan mengurutkan angka berdasarkan frekuensi kemunculan atribut. Tabel 1 menunjukkan contoh hasil inisialisasi atribut fasilitas kesehatan menggunakan angka sesuai dengan frekuensinya di Microsoft Excel.

Tabel 1. Data Transformasi

Fasilitas Kesehatan	Jumlah	Inisialisasi
RSU	1	1
RSU	0	1
RSU	4	1
RSU	5	1
RSU	4	1
RSU	5	1
RSU	1	1
RSU	4	1
RSU	4	1
RSU	3	1
RSU	6	1
RSU	2	1
RSU	3	1
RSU	1	1
RSU	7	1
RSU	1	1
RSU	3	1
RSU	7	1
RSU	4	1
...	...	...
Klinik Balai Kesehatan	5	24

### D. Pengolahan Data

Algoritma PAM, yang dikenal sebagai algoritma K-Medoids, adalah salah satu algoritma pengelompokan (clustering) yang menggunakan medoid sebagai representasi pusat cluster. Perbedaan utama pada algoritma K-Medoids dan algoritma K-Means terletak pada pemilihan pusat kluster. Dalam algoritma K-Medoids, objek data digunakan sebagai medoid yang mewakili pusat kluster untuk setiap kluster, sedangkan pada algoritma K-Means, nilai rata-rata (mean) digunakan sebagai center kluster [5][6]. Langkah-langkah algoritma K-Medoids :

1. Memilih medoid awal secara acak dari objek data yang ada.

2. Tentukan setiap data (objek) ke kluster terdekat menggunakan persamaan ukuran jarak Euc lidian Distance dengan persamaan:

$$d(\mathbf{p}, \mathbf{q}) = \sqrt{\sum_{i=1}^n (q_i - p_i)^2}$$

3. Pilih secara acak objek pada masing-masing kluster sebagai anggota medoids baru.
4. Hitung jarak antara tiap-tiap objek dalam kluster dengan kandidat medoid baru yang sesuai.
5. Hitung total perubahan (S) dengan menghitung selisih antara total jarak baru dan total jarak sebelumnya. Jika  $S < 0$ , maka tukar objek dengan data kluster untuk membangun sekumpulan k objek baru sebagai medoid.
6. Ulangi langkah 3 hingga 5 sampai tidak terjadi perubahan medoid. Dalam proses ini, teruslah membentuk cluster baru berdasarkan medoid yang terbaru dan anggota-anggota cluster yang sesuai.

#### E. Data Mining

Aktivitas mencari pengetahuan yang tersembunyi dari sekumpulan data yang sangat banyak merupakan arti dari data mining [8]. Hasil atau output dari proses data mining dapat digunakan sebagai dasar untuk mengambil keputusan dan merencanakan langkah di masa depan [9]. Data mining dapat dibagi menjadi enam metode utama berdasarkan tujuan dan tugasnya, yaitu deskripsi, estimasi, prediksi, klasifikasi, klasterisasi, dan asosiasi [9].

#### F. Clustering

Clustering sebagai salah satu teknik dalam data mining, yang digunakan untuk mengidentifikasi dan membentuk kelompok objek yang memiliki sifat yang serupa dalam dataset yang besar [4]. kluster berbeda dari klasifikasi dalam hal bahwa kluster tidak melibatkan variabel target. Tujuan dari clustering adalah untuk mengelompokkan objek data ke dalam kelompok yang mempunyai kesamaan sifat atau pola tertentu, tanpa

memperhatikan variabel target yang mungkin ada dalam data. [1].

#### G. Davies Bouldin Index (DBI)

Merupakan sebuah metrik yang digunakan untuk menilai hasil dari algoritma clustering. DBI dikenalkan oleh David L. Davies dan Donald W. Bouldin pada tahun 1979 [12]. Dalam penggunaan DBI, suatu kluster dianggap mempunyai skema clustering yang baik jika memiliki nilai DBI mendekati 0 (positif).

### III. HASIL DAN PEMBAHASAN

Hasil penelitian hendaknya dituliskan secara jelas dan padat. Diskusi hendaknya menguraikan arti pentingnya hasil penelitian, bukan mengulanginya.

#### A. Perhitungan Manual Algoritma K-Medoids

##### Memilih medoid awal secara acak

Setiap medoid dipilih secara acak (random) dalam proses pemilihan medoid yang telah di transformasi. Diasumsikan Kepulauan Seribu Utara dan Kepulauan Seribu Selatan sebagai medoid awal.

Tabel 2. Medoids Awal

Cluster	Kecamatan	Fasilitas Kesehatan	Jumlah
1	Kepulauan Seribu Utara	Rumah Sakit Umum	1
2	Kepulauan Seribu Selatan	Rumah Sakit Umum	0

#### Hitung jarak menggunakan persamaan Euclidian Distance

Dalam perhitungan jarak antara setiap objek dengan medoid awal menggunakan persamaan Euclidean Distance sebagai berikut :

Tabel 3. Iterasi 1

wilayah	cost1	cost2	kedekatan	cluster
Kepulauan Seribu	0	0,011627907	0	1
Kepulauan Seribu	0,011627907	0	0	2
Cilandak	0,034883721	0,046511628	0,034883721	1
Jagakarsa	0,046511628	0,058139535	0,046511628	1
Kebayoran Baru	0,034883721	0,046511628	0,034883721	1
Kebayoran Lama	0,046511628	0,058139535	0,046511628	1
Mampang Prapatan	0	0,011627907	0	1
Pancoran	0,034883721	0,046511628	0,034883721	1

Pasar Minggu	0,034883721	0,046511628	0,034883721	1
Pesanggrahan	0,023255814	0,034883721	0,023255814	1
Setiabudi	0,058139535	0,069767442	0,058139535	1
Tebet	0,011627907	0,023255814	0,011627907	1
Jatinegara	0,023255814	0,034883721	0,023255814	1
Matraman	0	0,011627907	0	1
Pulogadung	0,069767442	0,081395349	0,069767442	1
Cakung	0	0,011627907	0	1
Duren Sawit	0,023255814	0,034883721	0,023255814	1
Kramat Jati	0,069767442	0,081395349	0,069767442	1
Makasar	0,034883721	0,046511628	0,034883721	1
...	...	...	...	...
Cilincing	1,035144989	1,038209967	1,035144989	1
Jumlah Cost	114,9957418	115,6794697		
Total Cost	230,6752115			

Tabel 4. Iterasi 4

Wilayah	Cost1	Cost2	Kedekatan	Cluster
Kepulauan Seribu	1,406015432	1,035144989	1,035144989	1
Kepulauan Seribu	1,414213562	1,038209967	1,038209967	1
Cilandak	1,381716351	1,026685584	1,026685584	1
Jagakarsa	1,373717997	1,024114333	1,024114333	1
Kebayoran Baru	1,381716351	1,026685584	1,026685584	1
Kebayoran Lama	1,373717997	1,024114333	1,024114333	1
Mampang Prapatan	1,406015432	1,035144989	1,035144989	1
Pancoran	1,381716351	1,026685584	1,026685584	1
Pasar Minggu	1,381716351	1,026685584	1,026685584	1
Pesanggrahan	1,389765963	1,029381768	1,029381768	1
Setiabudi	1,365771801	1,02166896	1,02166896	1
Tebet	1,397865947	1,032201908	1,032201908	1
Jatinegara	1,389765963	1,029381768	1,029381768	1
Matraman	1,406015432	1,035144989	1,035144989	1
Pulogadung	1,357878678	1,019350369	1,019350369	1
Cakung	1,406015432	1,035144989	1,035144989	1
Duren Sawit	1,389765963	1,029381768	1,029381768	1
Kramat Jati	1,357878678	1,019350369	1,019350369	1
Makasar	1,381716351	1,026685584	1,026685584	1
...	...	...	...	...
Cilincing	0,720930233	0	0	1
Jumlah Cost	234,5760268	127,7051394		
Total Cost	362,2811663			

Setelah mendapatkan jumlah jarak (cost) pada setiap objek di iterasi 1, langkah selanjutnya iterasi 2. Pada iterasi 2, Medoid baru akan dipilih. Medoid baru dipilih dengan mempertimbangkan objek-objek yang bukan medoid pada iterasi sebelumnya. Lakukan perhitungan jarak kembali untuk setiap objek di iterasi kedua dengan menggunakan medoid yang baru.

Tabel 5. Medoids Baru

Cluster	Kecamatan	Fasilitas Kesehatan	Jumlah
2	Kelapa Gading	Klinik Balai Kesehatan	86
1	Cilincing	Klinik Balai Kesehatan	24

### Hitung Total Simpangan (S)

Setelah mendapatkan jumlah jarak dari iterasi sebelumnya, langkah berikutnya adalah menghitung (S) dengan mengambil selisih antara total cost yang baru dan total cost sebelumnya. Jika jumlah S kurang dari 0, maka dilakukan pertukaran jumlah objek untuk menentukan medoid baru.

$$\begin{aligned}
 S &= \text{Total jarak 2} - \text{Total jarak 1} \\
 &= 362,2811663 - 230,6752115 \\
 &= 131,605954817405
 \end{aligned}$$

Karena total simpangan lebih dari 0, maka perhitungan sampai disini, dan hasil akhir dari proses clustering adalah sebagai berikut:

Tabel 6. Hasil Cluster Akhir

Fasilitas Kesehatan	Jumlah	Cluster
RSU	1	1
RSU	0	2
RSU	4	1
RSU	5	1
RSU	4	1
RSU	5	1
RSU	1	1
RSU	4	1
RSU	4	1
RSU	3	1
RSU	6	1
RSU	2	1
RSU	3	1
RSU	1	1
RSU	7	1

IV. KESIMPULAN

Berdasarkan hasil perhitungan di atas diperoleh 2 kluster, yaitu kluster 0 dengan fasilitas kesehatan sejumlah 43 wilayah dan kluster 1 dengan fasilitas kesehatan sedikit sejumlah 177 wilayah. Dari hasil kluster masih banyak kecamatan di Jakarta yang belum memiliki fasilitas yang cukup. Dari hasil clustering, terlihat masih banyak kecamatan di Jakarta yang belum memiliki fasilitas kesehatan yang cukup. Hal ini memberikan acuan penting bagi pemerintah untuk memperhatikan pemerataan fasilitas kesehatan, sehingga masyarakat dapat dengan mudah mendapatkan pelayanan tanpa harus melakukan perjalanan jauh dari tempat tinggal mereka. Dengan demikian, upaya untuk meningkatkan aksesibilitas dan ketersediaan fasilitas kesehatan di kecamatan-kecamatan tersebut menjadi penting dalam upaya memenuhi kebutuhan masyarakat akan pelayanan kesehatan yang memadai.

UCAPAN TERIMA KASIH

Bagian ini memberikan apresiasi kepada perorangan maupun organisasi yang memberikan bantuan kepada penulis dan atau kepada pihak sponsor apabila ada.

REFERENSI

- [1] Mayangsarim, Sekar D.S., Solikhun & Irawan. Pengelompokan Jumlah Desa/Kelurahan yang Memiliki Sarana Kesehatan menurut Provinsi dengan Menggunakan Metode K-Means Cluster. Konferensi Nasional Teknologi Informasi dan Komputer (KOMIK), vol. 3, pp. 370-377, 2019.
- [2] Sadewo & Gading M. Pemanfaatan Algoritma Clustering dalam Mengelompokkan Jumlah Desa/Kelurahan yang Memiliki Sarana Kesehatan Menurut Provinsi dengan K-Means. Konferensi Nasional Teknologi Informasi dan Komputer (KOMIK), vol. 1, pp. 124-131, 2017.
- [3] Han, J., Kamber, M. & Pei, J. Data mining: concepts and techniques Third Edition. Morgan Kaufmann Publishers, San Francisco, 2011.
- [4] Hantoro, Kusdarnowo., Andi S. & Siti A. Pencarian Lokasi Perumahan Berdekatan dengan Fasilitas Kesehatan dan Belanja menggunakan Algoritma K-Means. Jurnal IKRA-ITH Informatika, vol 4, pp.25-34, 2020.
- [5] Kaur, Noor K et al. "K-Medoids Clustering Algorithm". Vol. 1, Issue 1. April 2014.
- [6] IBRAHIM, L. Fattouh. Using Of Clustering And Ant-Colony Algorithms CWSP-PAM-ANT In Network Planning. In: Digital Telecommunications,, 2006. ICdT'06. International Conference on. IEEE, 2006. p. 63-63.
- [7] Andini, A. D., & Arifin, T. (2020). IMPLEMENTASI ALGORITMA K-MEDOIDS UNTUK KLASTERISASI DATA PENYAKIT PASIEN DI RSUD KOTA BANDUNG. JURNAL RESPONSIF, 2(2), 128-138. <http://ejournal.ars.ac.id/index.php/jti>
- [8] R. Setiawan, "Penerapan Data Mining Menggunakan Algoritma K-Means Clustering untuk Menentukan Strategi Promosi Mahasiswa Baru (Studi Kasus: Politeknik LP3I Jakarta)," J. Lentera ICT, vol. 3, no. 1, pp. 76-92, 2016.

RSU	1	1
RSU	3	1
RSU	7	1
RSU	4	1
...	...	...
Klinik Balai Kesehatan	24	1

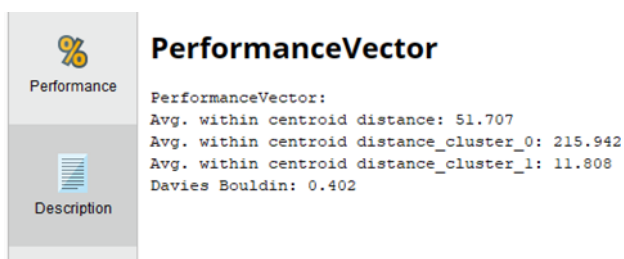
B. Perhitungan Algoritma K-Medoids dengan Rapidminer



Gambar 1. Design Pengujian

Selanjutnya dilakukan pengujian menggunakan aplikasi RapidMiner. Dalam pengujian tersebut, penelitian ini menggunakan Davies-Bouldin Index (DBI) sebagai metode untuk memeriksa hasil klustering. Pada Gambar 1, terlihat desain pengujian yang menggunakan algoritma K-medoid.

Setelah dataset diimpor ke RapidMiner, dataset tersebut akan disambungkan ke operator algoritma K-medoids. Selanjutnya, output dari operator K-medoids akan disambungkan ke operator performance untuk melakukan evaluasi menggunakan metrik DBI. Hasil evaluasi akan terhubung ke operator Result untuk menampilkan hasil akhir dari proses clustering.



Gambar 2. Hasil DBI

Pada gambar 2 didapat DBI sejumlah 0,402. Semakin rendah DBI yang dihasilkan (lebih dari 0) maka akan semakin baik cluster yang didapat.

- [9] Y. Mardi, "Data Mining: Klasifikasi Menggunakan Algoritma C4.5," J. Edik Inform., vol. 2, no. 2, pp. 213–219, 2017.
- [10] A. F. Lestari and M. Hafiz, "Penerapan Algoritma Apriori Pada Data Penjualan Barbar Warehouse," J. INOVTEK Polbeng Seri Inform., vol. 5, no. 1, pp. 96–105, 2020.
- [11] Marlina, D., Fauzer Putri, N., Fernando, A., & Ramadhan, A. (2018). Implementasi Algoritma K-Medoids dan K-Means untuk Pengelompokan Wilayah Sebaran Cacat pada Anak. Jurnal CoreIT, 4(2).
- [12] Irhamni F, Damayanti F, Khusnul K B & A Miftachul. Optimalisasi Pengelompokan Kecamatan Berdasarkan Indikator Pendidikan Menggunakan Metode Clustering Dan Davies Bouldin Index. Seminar Nasional Sains dan Teknologi 2014. Jakarta. 2014 : 1