
Pengelompokan Data Keluhan Pasien Pada Layanan Rumah Sakit Berdasarkan Kategori Masalah Menggunakan Metode Clustering

Refa Anggita¹⁾, Realita Buaton²⁾ & Kristina Annatasia Br Sitepu³⁾*

^{1,2,3)}Sistem Informasi, STMIK Kaputama, Indonesia

*Corresponding Email : refaanggita085@gmail.com

Abstrak

Penelitian ini membahas pengelompokan data keluhan pasien pada layanan RSU Artha Medica Binjai berdasarkan kategori masalah menggunakan algoritma K-Means Clustering. Data penelitian mencakup periode 2023–2024 dengan variabel umur pasien, kategori keluhan, dan kategori masalah. Proses pengolahan dilakukan menggunakan perangkat lunak Matlab R2014a, menghasilkan enam cluster dengan karakteristik berbeda. Hasil pengujian menunjukkan konfigurasi enam cluster memiliki nilai cluster variance terendah sebesar 4,5682, menandakan distribusi data paling kompak dibanding konfigurasi lainnya. Secara khusus, cluster keenam memiliki variance 5,0008 dengan Vmin 0,2472 dan Vmaks 10,3912, menunjukkan variasi yang terkendali dan sebaran data yang merapat ke pusat cluster. Temuan ini membuktikan bahwa penerapan K-Means Clustering dapat membantu rumah sakit dalam memahami pola keluhan pasien secara lebih akurat dan menjadi acuan strategis untuk peningkatan kualitas pelayanan.

Kata kunci: K-Means Clustering, keluhan pasien, rumah sakit, kategori masalah, Matlab.

Abstract

This study discusses the clustering of patient complaint data at RSU Artha Medica Binjai based on problem categories using the K-Means Clustering algorithm. The research data covers the period 2023–2024 with variables including patient age, complaint category, and problem category. Data processing was carried out using Matlab R2014a, producing six clusters with distinct characteristics. The testing results indicate that the six-cluster configuration achieved the lowest cluster variance value of 4.5682, representing the most compact data distribution compared to other configurations. Specifically, the sixth cluster recorded a variance of 5.0008 with a minimum value (Vmin) of 0.2472 and a maximum value (Vmax) of 10.3912, showing controlled variation and data points closely concentrated around the cluster centroid. These findings demonstrate that the application of K-Means Clustering can assist hospitals in understanding patient complaint patterns more accurately and serve as a strategic reference for improving service quality.

Keywords: K-Means Clustering, patient complaints, hospital services, problem category, Matlab.

PENDAHULUAN

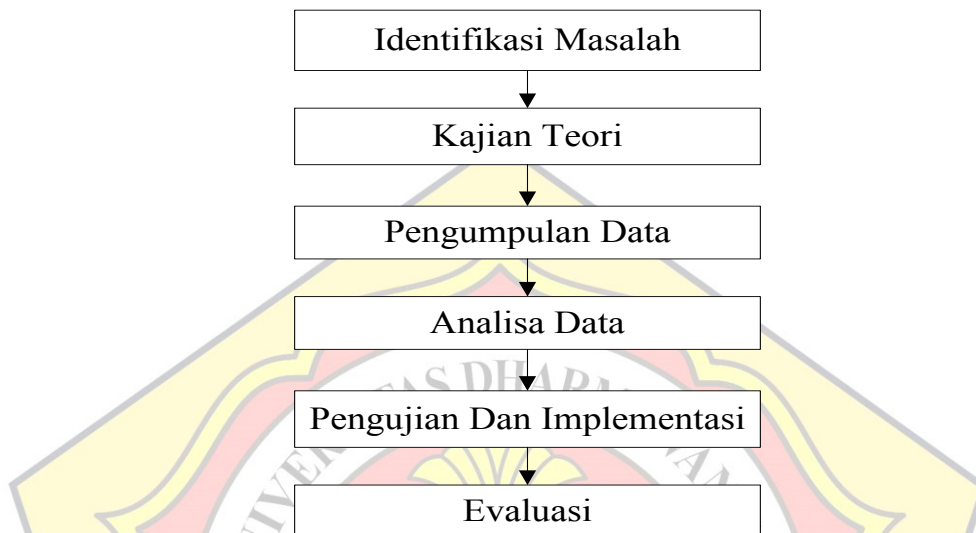
Pelayanan kesehatan yang optimal menjadi faktor penting dalam meningkatkan kepuasan pasien. Rumah sakit dituntut memberikan layanan yang cepat, akurat, dan responsif terhadap keluhan yang mencakup kualitas pelayanan medis, fasilitas, administrasi, maupun sikap tenaga medis. Apabila keluhan tidak dikelola dengan baik, kepercayaan masyarakat dapat menurun dan citra rumah sakit terdampak negatif. RSUD Artha Medica Binjai yang melayani banyak pasien setiap hari juga menghadapi beragam keluhan, sehingga diperlukan sistem yang mampu mengelompokkan keluhan untuk mempermudah analisis dan penanganan secara efektif.

Clustering menjadi salah satu metode yang tepat untuk mengelompokkan keluhan pasien berdasarkan kesamaan karakteristik. Algoritma K-Means, yang umum digunakan dalam clustering, dapat memproses data secara otomatis dan mengelompokkannya berdasarkan kedekatan karakteristik tertentu. Dengan pengelompokan yang jelas, manajemen rumah sakit dapat memahami tren keluhan dan mengambil langkah perbaikan yang strategis. Beberapa penelitian sebelumnya telah membuktikan efektivitas metode ini, seperti pengelompokan pasien BPJS di RSUD Kabupaten Kaur, analisis kecenderungan penyakit di Rumah Sakit Haji Adam Malik Medan, serta pengelolaan data rekam medis di Rumah Sakit Royal Prima Medan.

Berdasarkan hal tersebut, penelitian ini bertujuan menerapkan algoritma K-Means Clustering untuk mengelompokkan keluhan pasien di RSUD Artha Medica Binjai berdasarkan kategori masalah. Penelitian dilakukan pada data tahun 2023–2024 dengan variabel umur pasien, kategori keluhan, dan kategori masalah menggunakan perangkat lunak Matlab. Hasil pengelompokan ini diharapkan dapat menjadi acuan bagi pihak rumah sakit dalam meningkatkan kualitas pelayanan secara terarah dan berbasis data.

METODE PENELITIAN

Ada beberapa tahapan metodologi penelitian yang dilakukan dalam penyelesaian masalah. Tahapan tersebut adalah sebagai berikut :



Gambar 1. Alur Kerja Penelitian

Berdasarkan gambar di atas dapat diketahui bahwa ada beberapa tahapan dalam menyelesaikan penelitian yaitu :

1. Identifikasi Masalah
Tahap ini merupakan tahapan awal dalam penelitian yaitu dengan menentukan latar belakang masalah, tujuan dan manfaat agar tidak keluar dari fokus pembahasan atau penyusunan penelitian.
2. Kajian Teori
Tahap ini adalah mencari informasi, sumber-sumber yang berkaitan dengan masalah keluhan yang ada di Rumah Sakit RSUD Artha Medika,, jurnal dan internet sebagai pendukung dan landasan dasar penulisan penelitian.
3. Pengumpulan Data
Tahap ini merupakan pengumpulan data seperti data keluhan pasien dirumah sakit dan kebutuhan data lainnya yang diperoleh di RSUD Artha Medika.
4. Analisa data

Tahap ini merupakan tahapan mengolah dan menganalisis data yang telah diperoleh sehingga data tersebut dapat di kelompokkan sesuai dengan data yang ada.

5. Pengujian dan Implementasi

Tahap ini merupakan tahapan yang melakukan pengujian validasi dan implementasi data yang telah dianalisa sebelumnya serta penyusunan program.

6. Evaluasi

Tahap ini merupakan tahapan mengambil kesimpulan dan saran yang dapat dilakukan dalam penyusunan penelitian. Dengan adanya kesimpulan maka akan diketahui hasil dari keseluruhan penelitian dan diharapkan dengan saran akan ada perbaikan-perbaikan dan manfaat bagi yang lain.

HASIL DAN PEMBAHASAN

Proses analisis pengelompokan data keluhan pasien pada layanan rumah sakit berdasarkan kategori masalah yang kemudian dilakukan dengan menerapkan algoritma K-Means, sehingga memudahkan proses klasifikasi data. Dengan metode ini, dapat diidentifikasi jumlah kelompok berdasarkan kategori keluhan pasien, kategori masalah yang diperoleh, serta berbagai faktor pendukung lainnya.

4.1 Pembahasan Data Input

Data input pada sistem berupa data yang di dapatkan dari RSUD Artha Medica Binjai. Data input akan di transformasikan berdasarkan nilai transformasi dari masing-masing variabel yang di gunakan, berikut adalah penjelasan data yang dapat di input dan variabel serta nilai transformasi yang dilakukan, yaitu:

1. Data Input

Jumlah data : 800

Variabel : ➡ X Umur
➡ Y Kategori Keluhan
➡ Z Kategori Masalah

2. *Cluster* Pengelompokan : 3 *Cluster*

3. Nilai Transformasi

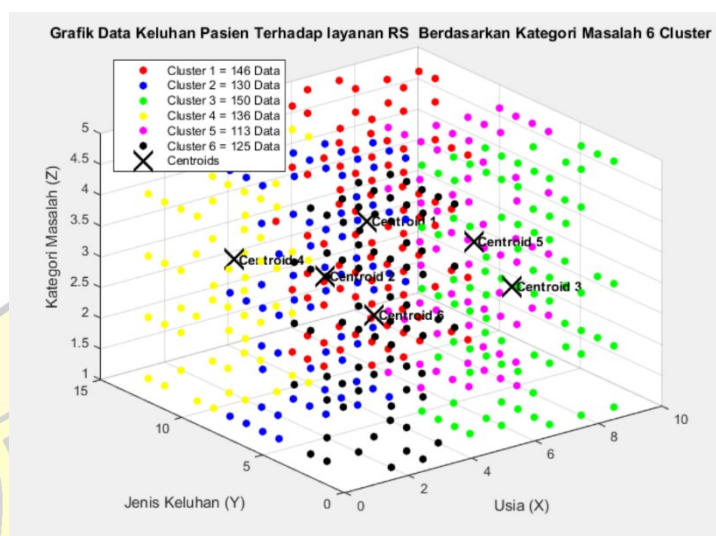
Tabel 1. Nilai Transformasi Data

NO	VARIABEL	TRANSFORMASI	TRANSFORMASI NILAI
1	Umur	17-21 Tahun	1
		22-26 Tahun	2
		27-31 Tahun	3
		32-36 Tahun	4
		37-41 Tahun	5
		42-46 Tahun	6
		47-52 Tahun	7
		53-56 Tahun	8
		57-60 Tahun	9
2	Kategori Keluhan	Cedera akibat prosedur	1
		Fasilitas tidak berfungsi	2
		Infeksi nosokomial	3
		Kebersihan ruangan	4
		Kenyamanan kamar inap	5
		Kesalahan diagnosis	6
		Kesalahan penagihan	7
		Keterlambatan pelayanan	8
		Kurangnya informasi prosedur	9
		Kurangnya komunikasi	10
		Kurangnya respons staf	11
		Kurangnya tindakan pencegahan	12
		Proses administrasi lambat	13
		Sikap staf kurang ramah	14
3	Kategori Masalah	Administrasi	1
		Kebersihan dan Fasilitas	2
		Interaksi Staf	3
		Pelayanan Medis	4
		Keselamatan Pasien	5

4.2 Implementasi

Implementasi pada tahap ini merupakan proses akhir dari penerapan metode Clustering pada sistem pengelompokan data yang telah dirancang. Dalam tahap ini, metode Clustering diintegrasikan ke dalam aplikasi pemrograman menggunakan perangkat lunak Matlab R2014a. Kode program yang dibuat menerapkan algoritma Clustering sehingga sistem mampu melakukan proses pengelompokan data secara efektif dan efisien.

Pengujian dilakukan dengan memproses data input menggunakan jumlah cluster sebanyak 6. Gambar berikut menampilkan hasil proses pengelompokan data dengan 6 cluster menggunakan metode Clustering yang telah diimplementasikan. Grafik hasil pengelompokan memperlihatkan titik-titik koordinat data pada masing-masing cluster serta posisi pusat cluster (centroid) yang dihasilkan.



Gambar 2. Grafik Hasil Pengelompokan 6 Cluster

Pada gambar di atas, dapat diketahui bahwa jumlah *Cluster* 1 sebanyak 146 data, *Cluster* 2 sebanyak 130 data, *Cluster* 3 sebanyak 150 data, *Cluster* 4 sebanyak 136, *Cluster* 5 sebanyak 113 data, dan *Cluster* 6 sebanyak 125 data dengan total 800 data. Keterangan data hasil pengelompokan menggunakan algoritma K-Means Clustering adalah sebagai berikut:

1. Pusat *Cluster* data

Tabel 2. Pusat *Cluster* Data Hasil Pada Matlab

NO	Pusat <i>Cluster</i>	Variabel			Jumlah Data
		X	Y	Z	
1	Centroid 1	6.8562	11.9863	2.9863	146
2	Centroid 2	3.2692	7.5769	3.1308	130
3	Centroid 3	6.7933	2.9333	3.0400	150
4	Centroid 4	2.5956	11.8603	2.9779	136
5	Centroid 5	7.7080	7.0796	3.1327	113
6	Centroid 6	2.5440	3.1360	3.144	125

2. Keterangan Pengelompokan Data

Hasil pengelompokan data menghasilkan enam cluster dengan karakteristik berbeda. Cluster pertama berisi 146 pasien usia 47–52 tahun yang mengeluhkan kebersihan ruangan terkait masalah interaksi staf. Cluster kedua terdiri dari 130 pasien usia 27–31 tahun dengan keluhan keterlambatan pelayanan akibat kurang optimalnya interaksi staf. Cluster ketiga mencakup 150 pasien usia 47–52 tahun yang mengeluhkan infeksi nosokomial, menunjukkan kekhawatiran terhadap risiko medis. Cluster keempat beranggotakan 136 pasien usia 27–31 tahun yang menyoroti kurangnya tindakan pencegahan. Cluster kelima terdiri dari 113 pasien usia 53–56 tahun yang menekankan masalah kesalahan penagihan, sedangkan cluster keenam berisi 125 pasien usia 22–26 tahun yang mengeluhkan infeksi nosokomial, menandakan pentingnya respons cepat tenaga medis terhadap pasien muda.

Tabel 3. Pengelompokan 6 Cluster

Cluster	Centroid	Variance	V_{min}	V_{maks}	Cluster Variance
6	9;12;3	4.3921	0.9937	11.5553	4.5682
	3;7;3	2.6025	0.4224	12.1763	
	7;3;3	5.2646	0.0488	12.7688	
	3;12;3	5.3826	0.1836	11.0365	
	8;7;3	4.7669	0.1092	11.1535	
	2;3;3	5.0008	0.2472	10.3912	

Berdasarkan hasil pengujian pada enam cluster, diperoleh nilai cluster variance paling rendah sebesar 4.5682 yang menunjukkan konfigurasi dengan tingkat kekompakan terbaik di antara seluruh pengujian. Cluster pertama memiliki variance 4.3921 dengan V_{min} 0.9937 dan V_{maks} 11.5553 yang masih dalam batas wajar, sedangkan cluster kedua mencatat variance 2.6025 dengan V_{min} 0.4224 dan

Vmaks 12.1763, mencerminkan distribusi yang lebih merapat. Cluster ketiga menunjukkan variance 5.2646 dengan Vmin 0.0488 dan Vmaks 12.7688, menandakan adanya data yang sangat dekat dengan centroid namun tetap memiliki penyebaran, sementara cluster keempat memiliki variance 5.3826 dengan Vmin 0.1836 dan Vmaks 11.0365 yang menunjukkan distribusi sedang. Adapun cluster kelima mencatat variance 4.7669 dengan Vmin 0.1092 dan Vmaks 11.1535 yang mengindikasikan kekompakan cukup baik, dan cluster keenam memiliki variance 5.0008 dengan Vmin 0.2472 dan Vmaks 10.3912 yang menunjukkan variasi dalam batas terkendali, sehingga secara keseluruhan hasil ini menggambarkan bahwa sebagian besar cluster berada pada rentang distribusi yang wajar dan mendukung analisis pengelompokan secara optimal.

4.3 Pengujian Hasil Cluster

Setelah data dikelompokkan ke dalam sebuah Cluster, peneliti akan melakukan pengujian terhadap hasil pengelompokan tersebut. Pengujian ini bertujuan untuk memastikan bahwa setiap Cluster memiliki kualitas yang baik serta representatif terhadap data yang ada. Dalam pengujian hasil Cluster digunakan analisis Cluster (Cluster Analysis), yaitu dengan mengukur variance atau penyebaran titik data dalam setiap Cluster. Perhitungan ini melibatkan nilai minimum (Vminimum), maksimum (Vmaksimum), dan variance Cluster. Variance untuk setiap Cluster dihitung menggunakan rumus sebagai berikut :

$$\text{Rumus : Variance}_k = \frac{1}{n_k} \sum_{i=1}^n (x_i - c_k) \dots\dots\dots$$

Dimana:

n_k adalah jumlah titik data dalam *Cluster*

x_i adalah titik data ke- i dalam *Cluster*

c_k adalah centroid dari *Cluster*

1. $V_{minimum}$ (Minumum Variance)

$V_{minimum}$ atau minimum variance adalah varians minimum dari semua cluster dalam suatu set data. Varians untuk setiap *Cluster* dihitung terlebih dahulu, kemudian nilai varians minimum dari semua *Cluster* diambil

Rumus : $V_{maksimum} = \min (Variance_1, Variance_2, \dots, Variance_k)$

2. $V_{maksimum}$ (Maximum Variance)

$V_{maksimum}$ atau Maximum Variance adalah variance maksimum dari semua *Cluster* dalam suatu set data. Variance untuk setiap *Cluster* dihitung terlebih dahulu, kemudian nilai variance maksimum dari semua *Cluster* diambil

Rumus : $V_{maksimum} = \max (Variance_1, Variance_2, \dots, Variance_k) \dots\dots\dots$

3. Cluster Variance

Cluster Variance adalah rata-rata dari semua variance *Cluster* dalam suatu set data. Ini memberikan gambaran umum tentang seberapa besar penyebaran titik data dalam *Cluster*.

Rumus : $Cluster\ Variance = \frac{1}{K} \sum_{K=1}^K Variance_k \dots\dots\dots$

4.3.1 Pengujian Hasil Cluster

Cluster 6 (2;3;3) Dengan 125 data

$$Variance\ C6 = \frac{1}{125} (6.5752 + 3.8632 + 6.9992 + \dots + 6.8552) = 5.0008$$

$$V_{min} = \min (6.5752; 3.8632; 6.9992; \dots; 6.8552) = 0.2472$$

$$V_{maks} = \max (6.5752; 3.8632; 6.9992; \dots; 6.8552) = 10.3912$$

$$Cluster\ Variance = \frac{1}{6} (4.3921 + 2.6025 + 5.2646 + 5.3826 + 4.7669 + 5.0008) = 4.5682$$

Berdasarkan hasil pengujian, konfigurasi enam cluster menjadi pilihan optimal dengan nilai cluster variance total terendah sebesar 4,5682, menandakan distribusi data yang paling kompak di antara seluruh pengujian. Fokus pada cluster keenam, diperoleh variance sebesar 5,0008 dengan V_{min} 0,2472 dan V_{maks} 10,3912, yang menunjukkan tingkat variasi data masih terkendali serta sebaran antar titik relatif merapat ke pusat cluster. Hal ini mengindikasikan bahwa cluster

keenam mampu mempertahankan keseimbangan antara kekompakan dan penyebaran data, sehingga mendukung analisis pengelompokan secara efektif.

SIMPULAN

Penelitian ini berhasil menerapkan algoritma K-Means Clustering untuk mengelompokkan data keluhan pasien di RSUD Artha Medica Binjai berdasarkan kategori masalah. Hasil pengujian menunjukkan bahwa konfigurasi enam cluster memberikan hasil terbaik dengan nilai cluster variance terendah sebesar 4,5682, menandakan distribusi data yang paling kompak dibandingkan konfigurasi lainnya. Secara khusus, cluster keenam yang memiliki variance 5,0008, Vmin 0,2472, dan Vmaks 10,3912 menunjukkan variasi data yang terkendali dan sebaran titik yang rapat ke pusat cluster. Temuan ini membuktikan bahwa metode K-Means Clustering efektif dalam mengidentifikasi pola keluhan pasien sehingga dapat menjadi dasar pengambilan keputusan strategis bagi pihak rumah sakit dalam meningkatkan kualitas pelayanan.

DAFTAR PUSTAKA

- Arhami, M., & Nasir, M. (2020). *Data Mining-Algoritma dan Implementasi* (Vol. 1). Andi Offset.
- Carudin, Marisa, Murnawan, Reba, F., Koibur, M. E., Thantawi, A. M., Halim, apriyanto, & Wattimena, F. Y. (2024). *Buku Ajar Data Mining* (Y. Agusdi, Ed.; 1st ed., Vol. 1). PT. Sonpedia Publishing Indonesia. https://www.google.co.id/books/edition/Buku_Ajar_Data_Mining/m-QGEQAAQBAJ?hl=id&gbpv=0
- Dea Syafitri, Yani Maulita, & Lina Arlianan Nur Kadim. (2024). Pengelompokan Data Keluhan Masyarakat Terhadap Fasilitas Umum di Kota Binjai Menggunakan Metode Clustering. *Jurnal Penelitian Teknologi Informasi Dan Sains*, 2(3), 158–170. <https://doi.org/10.54066/jptis.v2i3.2381>
- Fachrurrozi, A., Dimas, A. P., & Mulyanti, D. (2023). Strategi Peningkatan Mutu Pelayanan Kesehatan Di Rumah Sakit: Systematic Literature Review. *Jurnal Riset Rumpun Ilmu Kedokteran (JURRIKE)*, 2(1), 123–134.

- Fatwa, M., Ristu, R., Pandiangan, S., & Supriyadi, E. (2022). Pengaplikasian Matlab Pada Perhitungan Matriks. *Papanda Journal of Mathematics and Sciences Research*, 1(2), 81–93.
- Maharani, P. H., Dewi, L. K. C., & Martini, L. K. B. (2023). Pengaruh Penanganan Keluhan Dan Kualitas Pelayanan Terhadap Loyalitas Pasien Dengan Kepuasan Pasien Sebagai Variabel Mediasi Di Masa Pandemi Covid Di RSK Bedah BIMC, Kuta. *Jurnal Ekonomi Relasi*, 19(1), 134–145. <https://jurnal.itsm.ac.id/index.php/relasi/article/view/667/594>
- Monika, S. (2021). Pengelompokan Pelayanan Pasien Rawat Inap dan Pasien Rawat Jalan Berdasarkan Jenis Penyakit Pasien Menggunakan Metode Clustering (Studi Kasus: Rumah Sakit Delia). *Seminar Nasional Informatika (SENATIKA)*, 1–10.
- Mustika, Ardilla, Y., Manuhutu, A., Ahmad, N., Hasbi, I., Guntoro, Manuhutu, M. A., Ridwan, M., Hozairi, Wardhani, A. K., Alim, S., Romli, I., Religia, Y., Octafian, D. T., Sufandi, U. U., & Ernawati, L. (2021). *Data mining dan aplikasinya* (N. Rismawari, Ed.; 1st ed., Vol. 1). PT. Widina Media Utama.
- Purba, W., Sembiring, G. A., Saputra, A., Turnip, T., Jua, B., Manihuruk, I., Sains, F., & Teknologi, D. (2023). Penerapan Data Mining Untuk Pengelolaan Data Rekam Medis Menggunakan Metode K-Means Clustering Pada Rumah Sakit Royal Prima Medan. *Jurnal Tekinkom*, 6(1). <https://doi.org/10.37600/tekinkom.v6i1.857>
- Ramadhan, F., Chrisnanto, Y. H., & Ningsihm Ade Kania. (2021). Sistem Segmentasi Keluhan Pelanggan di Perumda Air Minum Tirta Raharja Cimahi Menggunakan Metode K-Medoids. *Informatics And Digital Expert (INDEX)*, 3(1), 6–9. <http://index.unper.ac.id>
- Sasmita, & Muntari, S. (2023). Penerapan Algoritma K-Means Clustering Pada Data Keluhan Pelanggan PT. PLN Persero Kota Pagar Alam. *Jurnal Ilmiah Teknosains*, 9(1), 8–12.
- Shepyantoni1, F., Kanedi, I., & Suryana, E. (2024). Penerapan Metode K-Means Clustering Dalam Pengelompokan Data Pasien Rawat Inap Peserta BPJS Di Rumah Sakit Umum. *Jurnal Media Infotama*, 20(2), 493–500.