



Penerapan Algoritma K-Means untuk Segmentasi Pola Penyakit Berdasarkan Data Kunjungan Pasien di Klinik Nabiya Medika

Muhamad Haikal, Samsoni, Aprinia Handayani

Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Pamulang, Tangerang Selatan, Indonesia, 15316
muhamadhaikal.me@gmail.com, dosen00388@unpam.ac.id, dosen02719@unpam.ac.id

Abstrak

Klinik Nabiya Medika menghasilkan data kunjungan pasien rawat jalan dalam jumlah besar setiap harinya, namun belum dimanfaatkan secara optimal untuk menganalisis pola penyakit. Penelitian ini bertujuan menerapkan algoritma K-Means Clustering untuk mengelompokkan jenis penyakit berdasarkan frekuensi kunjungan serta membangun sistem informasi segmentasi penyakit berbasis web. Sebanyak 412 data kunjungan pasien periode 10–11 Mei 2026 diproses melalui tahap preprocessing yang meliputi deteksi header otomatis, pemetaan kolom, penghapusan data kosong dan duplikat, serta normalisasi menggunakan Min-Max Scaling. Hasil preprocessing menghasilkan 340 data bersih dengan 103 jenis penyakit unik. Proses penelitian menggunakan pendekatan CRISP-DM yang meliputi Business Understanding, Data Understanding, Data Preparation, Modeling, Evaluation, dan Deployment. Hasil clustering menggunakan nilai $k=3$ mencapai konvergensi pada iterasi ke-4 dan menghasilkan tiga cluster, yaitu Rendah, Sedang, dan Tinggi, dengan centroid akhir masing-masing sebesar 1,54, 8,83, dan 36,23 kunjungan. Cluster Tinggi terdiri dari 4 jenis penyakit dengan total 145 kunjungan dan didominasi oleh Dyspepsia serta penyakit saluran pernapasan atas, cluster Sedang terdiri dari 8 jenis penyakit dengan total 63 kunjungan, sedangkan cluster Rendah terdiri dari 91 jenis penyakit dengan total 132 kunjungan. Sistem dibangun menggunakan framework Flask berbasis Python dan database MySQL serta mendukung pengolahan data dalam format CSV, XLSX, dan XLS. Pengujian Black Box dan White Box terhadap 95 skenario menghasilkan tingkat keberhasilan 100%. Hasil penelitian menunjukkan bahwa sistem dapat membantu pihak Klinik Nabiya Medika dalam memahami pola penyakit serta mendukung perencanaan obat, penjadwalan tenaga medis, dan penyusunan program promosi kesehatan

Kata kunci: K-Means Clustering; Segmentasi Penyakit; Data Mining; Min-Max Scaling; Sistem Informasi Kesehatan

Abstract

Klinik Nabiya Medika generates a large amount of outpatient visit data every day, but the data has not been utilized optimally to analyze disease patterns. This study aims to apply the K-Means Clustering algorithm to group diseases based on visit frequency and develop a web-based disease segmentation information system. A total of 412 patient visit records from May 10–11, 2026 were processed through preprocessing stages, including automatic header detection, column mapping, removal of empty and duplicate data, and normalization using Min-Max Scaling. The preprocessing results produced 340 clean records with 103 unique disease types. The research process applied the CRISP-DM approach, consisting of Business Understanding, Data Understanding, Data Preparation, Modeling, Evaluation, and Deployment. The clustering process using $k=3$ converged at the fourth iteration and produced three clusters: Low, Medium, and High, with final centroids of 1.54, 8.83, and 36.23 visits, respectively. The High cluster consisted of 4 disease types with a total of 145 visits and was dominated by Dyspepsia and upper respiratory tract diseases, while the Medium cluster consisted of 8 disease types with a total of 63 visits, and the Low cluster consisted of 91 disease types with a total of 132 visits. The system was developed using the Flask framework based on Python and MySQL database and supports data processing in CSV, XLSX, and XLS formats. Black Box and White Box testing of 95 scenarios resulted in a 100% success rate. The results indicate that the system can assist Klinik Nabiya Medika in understanding disease patterns and support medication planning, medical staff scheduling, and the development of health promotion programs.

Keywords: K-Means Clustering, Disease Segmentation, Data Mining, Min-Max Scaling

1. Pendahuluan

Kesehatan merupakan salah satu aspek penting dalam kehidupan masyarakat dan menjadi faktor utama dalam menunjang kualitas hidup. Seiring dengan perkembangan teknologi informasi, pemanfaatan teknologi dalam bidang kesehatan semakin berkembang, khususnya dalam pengelolaan data (Syahputra & Hanifah, 2024). Klinik pratama sebagai fasilitas pelayanan kesehatan primer menghasilkan data kunjungan pasien dalam jumlah besar

setiap harinya. Namun pada kenyataannya, masih banyak instansi kesehatan seperti klinik yang belum memanfaatkan data secara optimal, termasuk Klinik Nabiya Medika (Dilawati et al., 2024)

Berdasarkan hasil observasi awal di Klinik Nabiya Medika, data kunjungan pasien yang tersedia hanya digunakan sebagai laporan administrasi tanpa dilakukan pengelolaan lebih lanjut. Hal ini menyebabkan pihak klinik kesulitan dalam mengetahui pola penyakit yang sering terjadi, sehingga belum dapat mendukung proses pengambilan keputusan secara efektif, seperti perencanaan pengadaan obat dan penjadwalan tenaga medis (Rizqi & Hayati, 2024)

Dalam data mining, teknik clustering merupakan metode pengelompokan data berdasarkan tingkat kemiripan tanpa menggunakan label (*unsupervised learning*). Salah satu algoritma clustering yang paling banyak digunakan adalah K-Means karena kesederhanaan implementasi, efisiensi komputasi, dan skalabilitasnya (Keparahan et al., 2025). (Lingga, 2024) menyatakan bahwa penerapan K-Means pada data pasien dapat membantu mengelompokkan jenis penyakit sehingga memudahkan analisis dan pengambilan keputusan di fasilitas kesehatan.

(Khalish et al., 2025) menjelaskan bahwa data mining merupakan bagian dari proses *knowledge discovery in database (KDD)* yang bertujuan untuk mengekstraksi pengetahuan dari data. Salah satu teknik yang sering digunakan dalam data mining adalah *clustering*, yaitu proses pengelompokan data berdasarkan kemiripan karakteristik, algoritma *K-Means* menjadi salah satu metode *clustering* yang paling banyak digunakan karena mampu mengidentifikasi pola dan struktur data secara efektif, meskipun memiliki kelemahan seperti sensitive terhadap centroid awal dan data outlier.

Menurut (Handayani, 2022), dalam Pembangunan aplikasi data *mining* berbasis *K-Means* diperlukan perancangan sistem yang terstruktur, termasuk penggunaan basis data untuk menyimpan dan mengelola data secara efisien. Oleh karena itu, dalam penelitian ini basis data dirancang ke dalam beberapa tabel yang saling berelasi, seperti tabel user, sesi_segmentasi, dan hasil_segmentasi. Untuk memastikan struktur data lebih efisien dan tidak terjadi redundansi, maka dilakukan proses normalisasi yang akan dijelaskan pada subbab berikutnya

Beberapa penelitian terdahulu telah mengaplikasikan K-Means pada data kesehatan. (Rizqi & Hayati, 2024) berhasil mengelompokkan kunjungan pasien rumah sakit di wilayah Bekasi menjadi tiga kategori (rendah, sedang, tinggi) dengan hasil evaluasi DBI yang baik. (Herdiawan, 2024) menghasilkan tiga cluster penyakit pada RSUD Ciamis yang dapat meningkatkan efisiensi layanan. Namun, penelitian yang berfokus pada segmentasi jenis penyakit berdasarkan frekuensi kunjungan pada klinik pratama dan dikemas dalam sistem berbasis web yang lengkap masih terbatas.

(Nafis Sjamsuddin et al., 2025) menerapkan algoritma K-Means untuk mengelompokkan pasien rawat inap BPJS berdasarkan karakteristik pasien, sedangkan (Mayola et al., 2025) mengelompokkan rumah sakit berdasarkan intensitas kunjungan pasien. Kedua penelitian tersebut menghasilkan tiga cluster yang dapat mendukung pengambilan keputusan di bidang kesehatan. Persamaannya dengan penelitian ini adalah penggunaan algoritma K-Means, sedangkan perbedaannya terletak pada objek pengelompokan, yaitu pasien rawat inap dan rumah sakit, sementara penelitian ini berfokus pada data kunjungan pasien untuk mengidentifikasi pola penyakit..

(Lakeisyah et al., 2025) menggunakan K-Means untuk mengelompokkan wilayah berdasarkan jumlah penderita penyakit, sedangkan Hamami dan (Hamami & Dahlan, 2024) menerapkannya untuk memetakan persebaran fasilitas dan tenaga kesehatan. Kedua penelitian tersebut bertujuan mendukung pemerataan layanan kesehatan. Persamaannya dengan penelitian ini adalah penggunaan K-Means dalam bidang kesehatan, sedangkan perbedaannya terletak pada objek yang diteliti, yaitu wilayah dan fasilitas kesehatan, bukan data kunjungan pasien.

(Frangky et al., 2025) menggunakan algoritma clustering untuk mengelompokkan pasien berdasarkan persepsi kualitas pelayanan, sedangkan (Zerlinda Ardini, 2025) menerapkan K-Means untuk menganalisis pola penyakit kronis pada lansia. Kedua penelitian tersebut menunjukkan manfaat klasterisasi dalam mendukung peningkatan pelayanan kesehatan. Persamaannya dengan penelitian ini adalah penggunaan K-Means untuk analisis data kesehatan, sedangkan perbedaannya terletak pada fokus penelitian, yaitu kualitas pelayanan dan penyakit kronis pada lansia, sementara penelitian ini berfokus pada pengelompokan data kunjungan pasien untuk mengidentifikasi pola penyakit.

Penelitian ini bertujuan untuk: (1) menerapkan algoritma K-Means Clustering pada data kunjungan pasien Klinik Nabiya Medika untuk mengelompokkan jenis penyakit berdasarkan frekuensi kunjungan, (2) menganalisis hasil klasterisasi sehingga dapat diketahui pola pengelompokan penyakit yang dominan, dan (3) merancang serta membangun sistem klasterisasi berbasis web yang dapat digunakan oleh petugas klinik dalam proses pengolahan dan pengelompokan data secara cepat, akurat, dan efisien. Penelitian ini merupakan hasil original yang mengintegrasikan data mining dengan rekayasa perangkat lunak berbasis web untuk kebutuhan nyata di Klinik Nabiya Medika.

2. Metode Penelitian

2.1 Metodologi Pengembangan

Penelitian ini menggunakan pendekatan **CRISP-DM** (*Cross-Industry Standard Process for Data Mining*) yang merupakan metodologi standar dalam pengembangan sistem berbasis data mining. Metode ini dipilih karena sesuai untuk proses pengelompokan data pasien menggunakan algoritma K-Means, yang melibatkan tahapan analisis data, pemrosesan data, penerapan model, dan evaluasi hasil. Tahapan CRISP-DM yang diterapkan meliputi: (1) *Business Understanding* - analisis kebutuhan sistem dan tujuan penelitian; (2) *Data Understanding* - pemahaman terhadap atribut data kunjungan; (3) *Data Preparation* - pembersihan dan normalisasi data; (4) *Modeling* - penerapan K-Means; (5) *Evaluation* - evaluasi hasil clustering; dan (6) *Deployment* - implementasi dalam sistem berbasis web.

2.2 Data yang Digunakan

Data yang digunakan adalah data kunjungan pasien rawat jalan Klinik Nabiya Medika periode 10–11 Mei 2026, berjumlah **412 record**. Dataset memiliki atribut: No, Tgl, Pasien, Norm, Klinik, Kostumer, Dokter, Kddiagnosa, Diagnosa, dan Satu Sehat. Atribut yang relevan untuk penelitian adalah tanggal kunjungan, kode diagnosa, dan jenis diagnosa penyakit.

2.3 Data yang Digunakan

Sebelum clustering dilakukan, data melalui enam tahapan preprocessing:

1. **Deteksi Header Otomatis:** Sistem mendeteksi baris header secara otomatis menggunakan pencocokan kata kunci, karena file laporan klinik memiliki baris judul di atas header kolom.
2. **Pemetaan Kolom:** Nama kolom dinormalisasi dan alias kolom dideteksi otomatis (contoh: `tgl` → `tanggal\_kunjungan`, `diagnosa` → `jenis\_penyakit`).
3. **Penghapusan Data Kosong:** Baris tanpa nilai diagnosa dihapus. Dari 412 record awal, terdapat 72 record tanpa diagnosa sehingga diperoleh ****340 data bersih****.
4. **Penghapusan Duplikat:** Baris data yang identik dihapus untuk mencegah bias perhitungan.
5. **Normalisasi Min-Max Scaling:** Nilai `jumlah\_kunjungan` dinormalisasi ke rentang [0, 1] menggunakan rumus:
$$X' = \frac{X - X_{min}}{X_{max} - X_{min}}$$
6. **Agregasi per Jenis Penyakit:** Data diagregasi berdasarkan `jenis\_penyakit` dengan menjumlahkan seluruh kunjungan, menghasilkan **103 jenis penyakit unik** dengan rentang 1–46 kunjungan.

2.4 Algoritma K-Means Clustering

Menurut (Putra et al., 2025), data mining melalui Teknik *clustering* mampu mengekstraksi informasi penting dari data dalam jumlah besar dan mengubahnya menjadi pengetahuan yang berguna. Dalam bidang Kesehatan, *clustering* dapat dimanfaatkan untuk mengelompokkan data pasien berdasarkan pola tertentu, sehingga

memudahkan dalam analisis dan pengambilan Keputusan. Oleh karena itu, pada penelitian ini clustering digunakan untuk mengelompokkan pola penyakit berdasarkan data kunjungan pasien di Klinik Nabiya Medika.

Algoritma K-Means bertujuan membagi n data point ke dalam K cluster sehingga meminimalkan *Within-Cluster Sum of Squares* (WCSS). Nilai K ditetapkan **K=3** berdasarkan kebutuhan manajemen klinik untuk mengklasifikasikan penyakit ke dalam tiga tingkat prioritas: **Rendah, Sedang, dan Tinggi**.

Langkah-langkah K-Means yang diterapkan:

1. Menentukan centroid awal menggunakan strategi K-Means++ ($n_init=10$, $random_state=42$).
2. Menghitung jarak setiap data ke centroid menggunakan Euclidean Distance: $d = |X' - C|$
3. Mengelompokkan data ke cluster dengan jarak terdekat.
4. Menghitung ulang centroid sebagai rata-rata anggota cluster: $c_j = \frac{1}{n} \sum_{i=1}^n x_i$
5. Mengulangi langkah 2–4 hingga centroid tidak berubah (konvergen).

Label cluster ditentukan otomatis berdasarkan urutan nilai centroid: terkecil → Rendah, tengah → Sedang, terbesar → Tinggi. Setelah konvergensi, centroid dikembalikan ke skala asli menggunakan **inverse transform**:

$$X = X' \times (X_{max} - X_{min}) + X_{min}$$

2.5 Arsitektur Sistem

Sistem dibangun dengan arsitektur *three-tier* menggunakan stack teknologi berikut:

Table 1 Stack Teknologi Sistem

Komponen	Teknologi	Detail / Fungsi
Backend Framework	Flask (Python)	Server-side logic dan API endpoint
Algoritma Clustering	scikit-learn KMeans	Proses segmentasi / pengelompokan data
Manipulasi Data	pandas, NumPy	Pembersihan dan pengolahan data array/dataframe
Normalisasi	MinMaxScaler (scikit-learn)	Skalasi fitur ke rentang 0 sampai 1
Basis Data	MySQL 8.x	Penyimpanan data relasional
Frontend	HTML5, Bootstrap, Chart.js	Antarmuka pengguna dan visualisasi grafik
Koneksi DB	PyMySQL	Driver penghubung Flask ke MySQL
Format File Input	CSV, XLSX, XLS	Format berkas data yang didukung untuk diunggah

Sistem memiliki fitur: login dengan verifikasi admin, upload data, preprocessing otomatis, clustering K-Means, visualisasi hasil (pie chart, bar chart, line chart), ekspor CSV/Excel, dan laporan yang dapat dicetak sebagai PDF.

3. Hasil dan Pembahasan

3.1 Hasil Preprocessing

Dari 412 record data kunjungan rawat jalan periode 10–11 Mei 2026, proses preprocessing menghasilkan:

1. **340 data bersih** (72 record dihapus karena tidak memiliki diagnosa)
2. **103 jenis penyakit unik** setelah agregasi
3. Nilai kunjungan minimum: 1 kunjungan; maksimum: 46 kunjungan (Dyspepsia)
4. Nilai normalisasi: 0,000 (1 kunjungan) hingga 1,000 (46 kunjungan)

Sistem berhasil mendeteksi header otomatis pada baris ke-1 file XLS klinik, memetakan kolom tgl → tanggal_kunjungan dan diagnosa → jenis_penyakit, serta menghitung jumlah_kunjungan secara otomatis dari frekuensi kemunculan setiap diagnosis. Centroid awal yang diperoleh dari K-Means++ adalah C1 = 0,000, C2 = 0,222, C3 = 1,000 (skala normalisasi). Algoritma melakukan 4 iterasi hingga mencapai konvergensi. Pergerakan centroid selama iterasi ditunjukkan pada Tabel 2.

Table 2. Pergerakan Centroid Selama Iterasi

Iterasi	Centroid 1 (C1)	Centroid 2 (C2)
Awal	0,000	0,222
Iterasi 1	0,012	0,229
Iterasi 2	0,013	0,187

Centroid akhir dikonversikan ke skala kunjungan asli menggunakan inverse Min-Max Scaling:

1. C1 (Rendah): $0,012 \times 45 + 1 = 1,45$ kunjungan
2. C2 (Sedang): $0,174 \times 45 + 1 = 7,88$ kunjungan
3. C3 (Tinggi): $0,783 \times 45 + 1 = 36,25$ kunjungan

3.2 Hasil Preprocessing

Berdasarkan proses clustering, 103 jenis penyakit berhasil dikelompokkan ke dalam tiga cluster. Ringkasan hasil clustering ditampilkan pada Tabel 3.

Table 3 Ringkasan Hasil Clustering K-Means (K=3)

Cluster	Jumlah Penyakit	Jenis	Total Kunjungan	Kunjungan Min	Kunjungan Max	Nilai Centroid Akhir
Tinggi	4		145	26	46	36,25
Sedang	8		63	5	11	7,88
Rendah	91		132	1	4	1,45
Total	103		340	-	-	-

Cluster Tinggi terdiri dari 4 jenis penyakit dengan total 145 kunjungan, yaitu: Dyspepsia (46 kunjungan), Acute nasopharyngitis/common cold (43 kunjungan), Acute upper respiratory infection unspecified (30 kunjungan), dan Acute upper respiratory infections of multiple and unspecified sites (26 kunjungan). Penyakit-penyakit ini merupakan kelompok paling dominan yang memerlukan perhatian utama dari pihak klinik dalam penyediaan obat dan kesiapan tenaga medis.

Secara keseluruhan, kelompok penyakit saluran pernapasan atas (Acute nasopharyngitis, Acute upper respiratory infection unspecified, dan Acute upper respiratory infections of multiple sites) mencapai total **99 kunjungan**, menjadi penyakit yang paling dominan selama periode pengamatan.

Cluster Sedang terdiri dari 8 jenis penyakit dengan total 63 kunjungan, meliputi: Myalgia (11), Other dermatitis (10), Fever of other and unknown origin (10), Diarrhoea and gastroenteritis (8), Typhoid fever (8), Other arthritis (6), Cutaneous abscess furuncle and carbuncle (5), dan Fever unspecified (5). Penyakit dalam cluster ini perlu dipantau secara berkala karena berpotensi meningkat pada periode berikutnya.

Cluster Rendah terdiri dari 91 jenis penyakit dengan total 132 kunjungan dan rentang 1–4 kunjungan. Meskipun frekuensinya rendah, beberapa penyakit dalam cluster ini tetap memerlukan perhatian khusus karena bersifat kronis atau serius, antara lain: Essential (primary) hypertension, Non-insulin-dependent diabetes mellitus, Cerebral infarction, Congestive heart failure, Schizophrenia, dan Tuberculosis of other organs.

3.3 Implementasi Sistem

Sistem berhasil diimplementasikan sebagai aplikasi web berbasis Flask dengan antarmuka yang responsif. Halaman utama sistem meliputi: Login/Register dengan verifikasi admin, Upload Data, Preprocessing, Hasil Clustering dengan 4 jenis visualisasi grafik, dan Laporan.

Table 4 Rekapitulasi Hasil Pengujian Sistem

No	Fitur / Fungsi yang Diuji	Jumlah Skenario	Berhasil	Gagal	tingkat Keberhasilan
1	Login	6	6	0	100%
2	Register	5	5	0	100%
3	Dashboard Admin	5	5	0	100%
4	Upload Data	6	6	0	100%
5	Preprocessing	5	5	0	100%
6	Clustering	4	4	0	100%
7	Export dan Laporan	5	5	0	100%
-	Subtotal Black Box	36	36	0	100%
8	_find_header_row()	2	2	0	100%
9	preprocess()	9	9	0	100%
10	Normalisasi MinMaxScaler	5	5	0	100%
11	run_kmeans()	11	11	0	100%
12	get_descriptive_stats()	7	7	0	100%
13	get_summary()	6	6	0	100%
14	get_report_data()	9	9	0	100%
-	Subtotal White Box	49	49	0	100%
Total	Total Keseluruhan	85	85	0	100%

Dari total 95 skenario pengujian (36 Black Box + 59 White Box), seluruhnya menghasilkan status berhasil dengan **tingkat keberhasilan 100%**, membuktikan bahwa sistem telah berjalan sesuai kebutuhan fungsional dan logika algoritma yang dirancang

4. Kesimpulan

Penelitian ini berhasil menerapkan algoritma K-Means Clustering pada data kunjungan pasien rawat jalan Klinik Nabiya Medika dan membangun sistem informasi segmentasi penyakit berbasis web. Beberapa kesimpulan yang

dapat ditarik: 1). Algoritma K-Means berhasil diterapkan pada 412 data kunjungan periode 10–11 Mei 2026. Setelah preprocessing, diperoleh 340 data bersih dengan 103 jenis penyakit unik. Preprocessing meliputi deteksi header otomatis, pemetaan kolom, penghapusan data kosong dan duplikat, serta normalisasi Min-Max Scaling. 2). Clustering dengan K=3 mencapai konvergensi pada **iterasi ke-4** dengan centroid akhir 1,45 (Rendah), 7,88 (Sedang), dan 36,25 (Tinggi) kunjungan. Cluster Tinggi terdiri dari 4 jenis penyakit (145 kunjungan), Cluster Sedang 8 jenis penyakit (63 kunjungan), dan Cluster Rendah 91 jenis penyakit (132 kunjungan). 3). asil segmentasi menunjukkan penyakit saluran pernapasan atas mendominasi dengan total 99 kunjungan, diikuti Dyspepsia sebagai penyakit dengan kunjungan tertinggi (46 kunjungan). Informasi ini dapat digunakan sebagai dasar perencanaan kebutuhan obat, pengaturan tenaga medis, dan penyusunan program pencegahan penyakit. 4). Sistem berbasis web yang dibangun menggunakan Flask dan MySQL memiliki fitur lengkap: upload data, preprocessing otomatis, clustering, visualisasi interaktif, ekspor data, dan laporan. Pengujian Black Box dan White Box pada 95 skenario menghasilkan tingkat keberhasilan 100%.

Reference

1. Dilawati, H., Widiyanto, H., & Kuswadiji, A. (2024). Klasterisasi Data Rekam Medis Pasien Menggunakan Metode K-Means Clustering Di Rumah Sakit Widodo Ngawi. *BIOS: Jurnal Teknologi Informasi Dan Rekayasa Komputer*, 5(2), 139–147. <https://doi.org/10.37148/bios.v5i2.134>
2. Frangky, F., Sinaga, R., & Raihansyah, M. (2025). Analisis Segmentasi Pasien Berdasarkan Persepsi Kualitas Pelayanan dengan Algoritma Clustering. *Explorer*, 5(1), 52–58. <https://doi.org/10.47065/explorer.v5i1.1818>
3. Hamami, F., & Dahlan, A. (2024). SWADHARMA (JRIS) PENERAPAN ALGORITMA K-MEANS UNTUK MEMETAKAN PERSEBARAN FASILITAS DAN TENAGA KESEHATAN DI KOTA BANDUNG.
4. Handayani, F. (2022). Aplikasi Data Mining Menggunakan Algoritma K-Means Clustering untuk Mengelompokkan Mahasiswa Berdasarkan Gaya Belajar. *Jurnal Teknologi Dan Informasi*. <https://doi.org/10.34010/jati.v12i1>
5. Herdriaman, E. A. (2024). KLASERISASI PASIEN PADA RSUD CIAMIS MENGGUNAKAN METODE K-MEANS. *Jurnal Informatika Dan Teknik Elektro Terapan*, 12(3S1). <https://doi.org/10.23960/jitet.v12i3s1.5124>
6. Keparahan, T., Alasi, P.-G., Al, E., Alasi, G., & Putri, R. A. (2025). Penerapan K-Means Clustering untuk Pengelompokan Pasien Rumah Sakit berdasarkan Penerapan K-Means Clustering untuk Pengelompokan Pasien Rumah Sakit berdasarkan Tingkat Keparahatan Penyakit. 07. <https://doi.org/10.54209/jatilima.v7i03.1598>
7. Khalish, F., Piranti, N. M., & Martadireja, O. (2025). Implementasi Data Mining Menggunakan Teknik Clustering dengan Metode K-Means (Vol. 8, Number 5). <http://jiip.stkipyapisdompnu.ac.id>
8. Lakeisyah, E. T., Marshella, S. H., Putri, N. R., Rahman, M. F., Risyahputri, A., Maulana, R., Tania, K. D., & Sari, W. K. (2025). ANALISIS KLASERISASI JUMLAH PENDERITA PENYAKIT MENGGUNAKAN K-MEANS SEBAGAI DASAR DISTRIBUSI LAYANAN RUMAH SAKIT UMUM DI SUMATERA SELATAN. *JOISIE (Journal Of Information Systems And Informatics Engineering)*, 9(1), 123. <https://doi.org/10.35145/joisie.v9i1.4959>
9. Lingga, A. (2024). Penerapan Algoritma K-Means Clustering Untuk Klasterisasi Penyakit Pasien. *ADA Journal of Information System Research*, 1(3), 93–102. <https://doi.org/10.64366/adajisr.v1i3.43>
10. Mayola, L., Hafizhb, M., & Syahputra, H. S. (2025). Klasterisasi Rumah Sakit berdasarkan Kunjungan Pasien menggunakan Algoritma K-Means: Data 2019-2023. *Jurnal Teknologi Dan Sistem Informasi Bisnis*, 7(1), 15–21. <https://doi.org/10.47233/jteksis.v7i1.1703>
11. Nafis Sjamsuddin, I., Firmansyah, D. A., & Laferani, Y. (2025). Klasterisasi Pasien Rawat Inap BPJS pada RS Islam Assyifa Sukabumi menggunakan Metode K-Means. In *Pasien Rawat Inap BPJS pada RS Islam Assyifa Sukabumi menggunakan Metode K-Means* (Vol. 7).
12. Putra, S. H., Jamaludin, J., Simbolon, H., Ibrahim, I., & Rubianto, R. (2025). Data Mining untuk Pola Perbaikan Pelayanan Pasien dengan Metode Klustering K-Means pada Rumah Sakit Mitra Sejati. *Remik*, 9(3), 1000–1012. <https://doi.org/10.33395/remik.v9i3.15192>
13. Rizqi, M. F., & Hayati, U. (2024). CLUSTERING KUNJUNGAN PASIEN MENGGUNAKAN ALGORITMA K-MEANS PADA RUMAH SAKIT DI WILAYAH BEKASI. In *Jurnal Mahasiswa Teknik Informatika* (Vol. 8, Number 1). <https://opendata.jabarprov.go.id/>
14. Syahputra, R. A., & Hanifah, M. R. (2024). Metode Analisis Kesehatan Dengan Mengguakan Mechine Learning Atau Artificial Inteligenci Atau Data Mining Literature Review. <http://jurnal.utu.ac.id/invasi/>
15. Zerlinda Ardini, D. (2025). Analisis Pola Penyakit Kronis pada Lansia Menggunakan K-Means Clustering di Puskesmas Kelurahan Semper Barat. In *Jurnal Indonesia : Manajemen Informatika dan Komunikasi (JIMIK)* (Vol. 6, Number 3). <https://journal.stmiki.ac.id>