

PENGELOMPOKAN PENYAKIT ISPA PADA PASIEN BERDASARKAN USIA DENGAN MENGGUNAKAN METODE K-MEANS CLUSTERING

Dedek Andrea¹, Darjat Saripurna², Magdalena Simanjuntak³

Teknik Informatika, STMIK Kaputama, Binjai^{1,3}

Teknik Informatika, STMIK Triguna Dharma, Medan²

E-mail: *dedekandrea8@gmail.com¹, darjatsaripurna@gmail.com²,
magdalena.simanjuntak84@gmail.com³

ABSTRAK

Infeksi Saluran Pernapasan Akut (ISPA) adalah penyakit menular yang dipengaruhi oleh faktor lingkungan dan dapat menyerang saluran pernapasan. Penyakit ini disebabkan oleh berbagai agen infeksi seperti virus, bakteri, dan jamur. Penelitian ini bertujuan untuk mengelompokkan data pasien ISPA di Puskesmas Binjai Estate berdasarkan usia menggunakan metode K-Means Clustering. Pendekatan ini dipilih karena kemampuannya meminimalkan jarak antara data dengan pusat klusternya, sehingga data dengan karakteristik serupa dapat dikelompokkan secara efektif. Data yang digunakan adalah 941 data pasien ISPA dari Puskesmas Binjai Estate, dengan variabel yang dianalisis meliputi usia, gejala, dan diagnosis. Sistem pengelompokan diimplementasikan menggunakan aplikasi MATLAB R2025a. Hasil pengujian menunjukkan bahwa proses klustering berhenti pada iterasi ke-2 karena tidak ada lagi perubahan keanggotaan kluster. Dengan implementasi 3 kluster, diperoleh kluster 1 dengan 387 data (pusat centroid (2.10, 1.1, 1)), kluster 2 dengan 260 data (pusat centroid (3.81, 2.5, 1.62)), dan kluster 3 dengan 294 data (pusat centroid (4.1, 4.44, 3)). Hasil ini memberikan informasi berharga bagi pihak puskesmas untuk memahami karakteristik pasien ISPA dan menyusun protokol penanganan yang lebih efektif.

Kata kunci

Data Mining, K-Means Clustering, Penyakit ISPA, MATLAB

ABSTRACT

Acute Respiratory Tract Infection (ARI) is an infectious disease influenced by environmental factors that can affect the respiratory tract. This disease is caused by various infectious agents such as viruses, bacteria, and fungi. This study aims to cluster ARI patient data at Binjai Estate Health Center based on age using the K-Means Clustering method. This approach was chosen for its ability to minimize the distance between data and its cluster centers, so that data with similar characteristics can be grouped effectively. The data used consisted of 941 ARI patient records from the Binjai Estate Health Center, with analyzed variables including age, symptoms, and diagnosis. The clustering system was implemented using the MATLAB R2025a application. The test results show that the clustering process stopped at the 2nd iteration because there were no more changes in cluster membership. With the implementation of 3 clusters, cluster 1 was obtained with 387 data (centroid center (2.10, 1.1, 1)), cluster 2 with 260 data (centroid center (3.81, 2.5, 1.62)), and cluster 3 with 294 data (centroid center (4.1, 4.44, 3)). These results provide valuable information for the health center to understand the characteristics of ARI patients and develop more effective handling protocols.

Keywords

Data Mining, K-Means Clustering, ARI Disease, MATLAB

1. PENDAHULUAN

Infeksi Saluran Pernapasan Akut (ISPA) merupakan salah satu penyakit menular yang dipengaruhi oleh faktor lingkungan dan dapat menyerang saluran pernapasan, baik bagian atas maupun bawah. Penyakit ini disebabkan oleh beragam agen infeksi seperti virus, bakteri, dan jamur, yang dapat menimbulkan gejala klinis bervariasi. Berdasarkan data dari Kementerian Kesehatan Republik Indonesia, ISPA tetap menjadi salah satu masalah kesehatan masyarakat yang signifikan dengan prevalensi kasus yang tinggi di berbagai wilayah (Wahyudi et al., 2022). Penanganan yang tidak optimal berisiko menyebabkan komplikasi serius, terutama pada kelompok usia rentan.

Meskipun data pasien ISPA di Puskesmas Binjai Estate tersedia, pengelolaannya masih belum terstruktur secara analitis untuk mengidentifikasi pola dan karakteristik dominan. Kurangnya analisis mendalam terhadap data pasien, khususnya terkait pengelompokan berdasarkan usia, menghambat Puskesmas dalam merancang strategi intervensi yang tepat sasaran dan alokasi sumber daya yang efisien. Oleh karena itu, diperlukan suatu pendekatan data mining untuk menggali wawasan dari data historis guna mendukung pengambilan keputusan yang lebih berbasis bukti (Ucha Putri et al., 2021).

Data mining adalah proses menemukan pola, hubungan, dan tren yang signifikan dalam kumpulan data yang besar (Ucha Putri et al., 2021). Salah satu teknik utama dalam data mining adalah clustering, yaitu proses mengelompokkan data ke dalam sub-kelompok berdasarkan karakteristik atau pola tertentu yang serupa (Hendrastuty, 2024). Penelitian ini mengimplementasikan metode K-Means Clustering untuk mengatasi permasalahan tersebut. K-Means dipilih karena efisiensinya dalam mengelompokkan objek data berdasarkan kesamaan atribut, di mana objek dalam satu kluster memiliki kemiripan yang tinggi dan objek antar kluster memiliki perbedaan yang signifikan (Tedju, 2022). Dengan menerapkan metode ini pada data pasien ISPA, penelitian ini bertujuan untuk mengelompokkan pasien berdasarkan usia, yang diharapkan dapat menghasilkan kluster-kluster informatif.

Penelitian ini didukung oleh beberapa studi sebelumnya. Prasetyo et al. (2021) menggunakan algoritma K-Means untuk klasifikasi penyakit ISPA, sementara Prastiwi et al. (2022) mengimplementasikannya untuk menentukan persediaan stok barang. Maulia et al. (2021) juga menggunakan K-Means Clustering untuk mengelompokkan jenis penyakit pasien di Puskesmas, yang menunjukkan bahwa metode ini efektif untuk identifikasi pola penyakit. Selain itu, Agoestina et al. (2022) juga memanfaatkan algoritma K-Means untuk mengelompokkan pasien ISPA di Puskesmas Bah Biak, dan Tedju (2022) menggunakannya untuk pengelompokan penyakit hipertensi.

Hasil pengelompokan ini dapat memberikan kontribusi nyata bagi Puskesmas Binjai Estate dalam memahami distribusi penyakit, merancang program pencegahan yang lebih terfokus, dan meningkatkan kualitas pelayanan kesehatan secara keseluruhan.

2. METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif dengan data sekunder yang diperoleh dari Puskesmas Binjai Estate. Tahapan penelitian mengikuti alur kerja yang sistematis:

- a. Identifikasi Masalah: Menentukan masalah utama yaitu belum adanya sistem pengelompokan data pasien ISPA di puskesmas.
- b. Kajian Teori: Mengumpulkan literatur terkait data mining, clustering, algoritma K-Means, dan ISPA.

- c. Pengumpulan Data: Mengumpulkan 941 data pasien ISPA dari tahun 2022-2024 dari Puskesmas Binjai Estate.
 - d. Transformasi Data: Data non-numerik (Usia, Gejala, Diagnosis) diubah menjadi data numerik (Tabel 1).
 - e. Analisis dan Pengujian: Menerapkan algoritma K-Means Clustering dengan MATLAB R2025a untuk mengelompokkan data.
 - f. Implementasi dan Evaluasi: Menguji coba program dan menganalisis hasil pengelompokan untuk mendapatkan kesimpulan.
- Berikut adalah tabel kriteria transformasi data yang digunakan :

Tabel 1. Kriteria Usia

Kode	Usia	Keterangan
1	< 5 Tahun	Balita
2	6 – 12 Tahun	Anak-anak
3	13 – 19 Tahun	Remaja
4	20 – 59 Tahun	Dewasa
5	> 60 Tahun	Lansia

Tabel 2. Kriteria Gejala

Kode	Gejala
1	Batuk, Demam, Pilek
2	Sakit Tenggorokan, Batuk
3	Batuk Berdahak, Lemas
4	Sesak Nafas, Demam
5	Demam Tinggi, Batuk

Tabel 3. Kriteria Diagnosis

Kode	Diagnosis
1	ISPA Ringan
2	ISPA Sedang
3	ISPA Berat

3. HASIL DAN PEMBAHASAN

3.1 Perhitungan K-Means Clustering

Dalam penelitian ini, k ditentukan sama dengan 3. Titik centroid awal dipilih secara acak dari data yang telah ditransformasi. Proses perhitungan jarak menggunakan Euclidean Distance dilakukan hingga tidak ada lagi pergeseran data antar kluster, yang tercapai pada iterasi ke-2.

Tabel 4. Hasil Iterasi 1

Nilai Jarak							
NAMA	USIA (X)	GEJALA (Y)	DIAGNOSIS (Z)	C1	C2	C3	Group
A1	2	1	1	0.00	1.41	4.9	1
A2	1	2	1	1.41	2.00	4.69	1
A3	4	4	3	4.12	3.00	1.00	3
A4	4	3	2	3.00	1.73	2.24	2
A5	2	1	1	0.00	1.41	4.9	1

Tabel 5. Hasil Iterasi 2

Nilai Jarak							
NAMA	USIA (X)	GEJALA (Y)	DIAGNOSIS (Z)	C1	C2	C3	Group
A1	2	1	1	0.79	2.17	4.50	1
A2	1	2	1	0.79	2.59	4.39	1
A3	4	4	3	4.37	2.17	0.50	3
A4	4	3	2	3.41	0.86	1.80	2
A5	2	1	1	0.79	2.17	4.50	1

Tabel 6. Hasil Iterasi II

No	Nama	Group Iterasi I	Group Iterasi II
1	A1	1	1
2	A2	1	1
3	A3	3	3
4	A4	2	2
5	A5	1	1

Dari Tabel 6, dapat dilihat bahwa hasil pengelompokan pada iterasi pertama dan kedua identik. Oleh karena itu, proses iterasi dihentikan.

3.2 Visualisasi dan Hasil Clustering

Hasil pengelompokan data divisualisasikan menggunakan plot 3D di MATLAB R2025a, yang menunjukkan tiga kluster dengan titik-titik data dan centroidnya. Berdasarkan visualisasi dan perhitungan akhir, diperoleh tiga kluster dengan karakteristik sebagai berikut:

Tabel 7. Hasil Cluster Akhir

Cluster	Pusat Cluster	Variabel			Jumlah Data	Keterangan Karakteristik
		X	Y	Z		
1.	Centroid 1	2.10	1.1	1	387	Anak-anak (usia 6-12 tahun) dengan gejala ringan (batuk, demam, pilek) dan diagnosis ISPA ringan.
2.	Centroid 2	3.81	2.5	1.62	260	Remaja dan Dewasa (usia 13-40 tahun) dengan gejala sedang (sakit tenggorokan, batuk berdahak) dan diagnosis ISPA sedang.
3.	Centroid 3	4.1	4.44	3	294	Dewasa (usia 20-59 tahun) dengan gejala berat (sesak napas, demam tinggi) dan diagnosis ISPA berat.

5. KESIMPULAN

Penelitian ini berhasil menerapkan metode K-Means Clustering untuk mengelompokkan 941 data pasien ISPA di Puskesmas Binjai Estate berdasarkan usia, gejala, dan diagnosis. Hasil pengelompokan menunjukkan adanya tiga kluster yang berbeda: kluster anak-anak dengan ISPA ringan, kluster remaja-dewasa dengan ISPA sedang, dan kluster dewasa dengan ISPA berat. Implementasi ini membuktikan bahwa

algoritma K-Means dapat menjadi alat bantu yang efektif bagi tenaga kesehatan untuk memahami karakteristik pasien dan merancang strategi penanganan yang lebih terfokus dan efisien.

6. DAFTAR PUSTAKA

- Agoestina, F. S., Tambunan, H. S., & Nasution, R. A. (2022). Pemanfaatan Algoritma K-Means untuk Pengelompokan Pasien Penyakit Infeksi Saluran Pernafasan Akut (ISPA). *SATESI: Jurnal Sains Teknologi Dan Sistem Informasi*, 2(1), 60–69. <https://doi.org/10.54259/satesi.v2i1.804>
- Alfarizi, M. R. S., Al-farish, M. Z., Taufiqurrahman, M., Ardiansah, G., & Elgar, M. (2023). Penggunaan Python Sebagai Bahasa Pemrograman untuk Machine Learning dan Deep Learning. *Karya Ilmiah Mahasiswa Bertauhid (KARIMAH TAUHID)*, 2(1), 1–6.
- Hendrastuty, N. (2024). Penerapan Data Mining Menggunakan Algoritma K-Means Clustering Dalam Evaluasi Hasil Pembelajaran Siswa. *Jurnal Ilmiah Informatika Dan Ilmu Komputer (Jima-Ilkom)*, 3(1), 46–56. <https://doi.org/10.58602/jima-ilkom.v3i1.26>
- Maulana, A., Danar Dana, R., & Dienwati Nuris, N. (2024). Implementasi Algoritma K-Means Clustering Dalam Pengelompokan Data Kerusakan Rumah Akibat Bencana Alam Di Kabupaten Cirebon. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 8(2), 1417–1424. <https://doi.org/10.36040/jati.v8i2.9024>
- Maulia, S., Ginting, B. S., & Sihombing, A. (2021). Implementasi Data Mining Pengelompokan Jenis Penyakit Pasien Menggunakan Metode Clustering (Studi Kasus : Puskesmas Sambirejo). *Jurnal Informatika Kaputama (JIK)*, 5(1), 71–80. <https://doi.org/10.59697/jik.v5i1.304>
- Prasetyo, V. R., Lazuardi, H., Mulyono, A. A., & Lauw, C. (2021). Penerapan Aplikasi RapidMiner Untuk Prediksi Nilai Tukar Rupiah Terhadap US Dollar Dengan Metode Linear Regression. *Jurnal Nasional Teknologi Dan Sistem Informasi*, 7(1), 8-17. <https://doi.org/10.25077/teknosi.v7i1.2021.8-17>
- Prastiwi, H., Jeny Pricilia, & Errissya Rasywir. (2022). Implementasi Data Mining Untuk Menentuksn Persediaan Stok Barang Di Mini Market Menggunakan Metode K-Means Clustering. *Jurnal Informatika Dan Rekayasa Komputer(JAKAKOM)*, 2(1), 141–148. <https://doi.org/10.33998/jakakom.2022.2.1.34>
- Pujiono, S., Astuti, R., & Muhamad Basysyar, F. (2024). Implementasi Data Mining Untuk Menentukan Pola Penjualan Produk Menggunakan Algoritma K-Means Clustering. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 8(1), 615–620. <https://doi.org/10.36040/jati.v8i1.8360>
- Putra, E. M., Moh. Adib, & Prayitno, B. (2022). Hubungan Kondisi Fisik Rumah Dan Perilaku Keluarga Terhadap Kejadian Infeksi Ssaluran Pernafasan Akut (ISPA) Pada Balita Di Wilayah Kerja Puskesmas Perumas I Kota Pontianak 2021, *Journal of Environmental Health and Sanitation Technology. Journal of Environmental Health and Sanitation Technology*, 1(1), 32–39.
- SYAMILURRAHIM, M. (2019). Pengembangan Media Pembelajaran Kubus Dan Balok Berbasis Gui Matlab. *AL-Ahya*, 01(01), 219–232.
- Tedju, Y. (2022). Pengelompokan Penyakit Hipertensi Di Kota Kupang Menggunakan Metode K-Means. *HOAQ (High Education of Organization Archive Quality): Jurnal Teknologi Informasi*, 13(2), 98–108. <https://doi.org/10.52972/hoaq.vol13no2.p98-108>
- Ucha Putri, S., Irawan, E., Rizky, F., Tunas Bangsa, S., -Indonesia Jln Sudirman Blok No, P.

- A., & Utara, S. (2021). Implementasi Data Mining Untuk Prediksi Penyakit Diabetes Dengan Algoritma C4.5. Januari, 2(1), 39–46.
- Wahyudi, A., Zaman, C., Studi, P., Kesehatan, M., Ventilasi, L., Risiko, F., & Merokok, A. K. (2022). Analisis kejadian ispa pada anak dalam lingkungan keluarga perokok di wilayah kerja uptd puskesmas x kota palembang. 2(3), 475–482.
- Wulantika Dwi Mulyaningtyas, & Musta'in, M. (2024). Pengelolaan Bersihan Jalan Nafas Tidak Efektif pada Pasien ISPA. Jurnal Keperawatan Berbudaya Sehat, 2(1), 1–6. <https://doi.org/10.35473/jkbs.v2i1.2430>
- Zalukhu, A., Swingly, P., & Darma, D. (2023). Perangkat Lunak Aplikasi Pembelajaran Flowchart. Jurnal Teknologi, Informasi Dan Industri, 4(1), 61–70. <https://ejurnal.istp.ac.id/index.php/jtii/article/view/351>