Jurnal Minfo Polgan (JMP)
Volume 15, Nomor 3, September-Desember 2026
DOI : https://doi.org/10.33395/jmp.v15i3.16988
e-ISSN : 2797-3298
p-ISSN : 2089-9424

Analisis Klasterisasi Karakteristik Penyakit Kronis pada Lansia Menggunakan *K-Means* untuk Mendukung Prioritas Pelayanan Kesehatan

[1]Fitriana Harahap, [2]Wirhan Fahrozi, [3]Andrian Syahputra
[4]Robiatul Adawiyah, [5]Nur Hayati, [6]Husin Sariangsah
[1*,2]Sistem Informasi, Universitas Pembinaan Masyarakat Indonesia, Medan, Indonesia
[3]Universitas Muhammadiyah Asahan, Kisaran, Indonesia
[4]Sistem Informasi, Universitas Potensi Utama , Medan, Indonesia
[5]Informatika, Universitas Potensi Utama, Medan, Indonesia
[6]Informatika, Universitas Audi Indonesia, Medan, Indonesia
*Korespondensi : fitrianaharahap1@gmail.com

Submit : 22 Sept 2026 | **Diterima** : 25 Okt 2026 | **Terbit** : 29 Okt 2026

ABSTRACT

Chronic diseases are one of the most common health problems experienced by elderly people. Differences in disease characteristics among the elderly require a grouping process to facilitate analysis and decision-making in the health sector. This study aims to apply the K-Means Clustering algorithm to classify chronic diseases among elderly residents in Deli Tua Village. The data used consisted of 20 elderly records with attributes including gender, age, disease symptom code, and duration of illness. The research stages included data collection, data preprocessing, and the implementation of the K-Means algorithm using RapidMiner. The results showed that the K-Means algorithm was able to group elderly data into several clusters based on similar characteristics. The clustering results can help identify patterns of chronic diseases among the elderly and support the development of more targeted health programs in Deli Tua Village.

Keywords: *K-Means Clustering, Chronic disease, Elderly, Data mining, RapidMiner.*

ABSTRAK

Penyakit kronis merupakan salah satu permasalahan kesehatan yang sering dialami oleh kelompok lanjut usia (lansia). Perbedaan karakteristik penyakit yang dialami setiap lansia menyebabkan perlunya dilakukan pengelompokan untuk memudahkan analisis dan pengambilan keputusan dalam bidang kesehatan. Penelitian ini bertujuan untuk menerapkan algoritma *K-Means Clustering* dalam mengelompokkan penyakit kronis pada warga lansia di Desa Deli Tua. Data yang digunakan terdiri dari 20 data lansia dengan atribut jenis kelamin, umur, kode gejala penyakit, dan lama mengidap penyakit. Tahapan penelitian meliputi pengumpulan data, *preprocessing* data, dan penerapan algoritma *K-Means* menggunakan *RapidMiner*. Hasil penelitian menunjukkan bahwa algoritma *K-Means* mampu mengelompokkan data lansia ke dalam beberapa cluster berdasarkan kemiripan karakteristik yang dimiliki. Hasil pengelompokan tersebut dapat membantu dalam mengidentifikasi pola penyakit kronis pada lansia serta mendukung perencanaan program kesehatan yang lebih tepat sasaran di Desa Deli Tua

Kata Kunci*:* K-Means Clustering, Penyakit Kronis, Lansia, Data mining, RapidMiner

PENDAHULUAN

Lanjut usia (lansia) merupakan kelompok penduduk yang rentan mengalami berbagai penyakit kronis akibat penurunan fungsi fisiologis yang terjadi secara bertahap. Penyakit kronis seperti hipertensi, diabetes melitus, penyakit jantung, dan gangguan persendian menjadi masalah kesehatan yang banyak ditemukan pada kelompok lansia. Kondisi tersebut dapat menurunkan kualitas hidup, meningkatkan ketergantungan terhadap orang lain, serta menambah beban pelayanan kesehatan di masyarakat. Berdasarkan hasil penelitian Ardini dan Rhodiyah (2025), penyakit kronis pada lansia menunjukkan pola yang beragam sehingga

@2026 The Author Published By Politeknik Ganesha Medan
This is an open access article under the CC BY licence (http://creativecommons.org/licenses/by/4.0/). which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properlycited.

diperlukan identifikasi karakteristik penyakit untuk mendukung perencanaan layanan kesehatan yang lebih tepat sasaran.

Data kesehatan lansia yang dikumpulkan melalui kegiatan posyandu maupun fasilitas pelayanan kesehatan umumnya memuat informasi mengenai usia, jenis kelamin, riwayat penyakit, gejala yang dialami, serta lama menderita penyakit tertentu. Data tersebut memiliki potensi untuk dimanfaatkan dalam proses pengambilan keputusan, khususnya dalam mengidentifikasi kelompok lansia berdasarkan kesamaan karakteristik penyakit yang diderita. Namun, sebagian besar data kesehatan masih digunakan sebatas dokumentasi sehingga informasi yang terkandung di dalamnya belum dimanfaatkan secara optimal untuk mendukung program kesehatan masyarakat (Aulia et al., 2025).

Salah satu pendekatan yang dapat digunakan untuk mengidentifikasi pola pada data kesehatan adalah teknik clustering. *Clustering* merupakan metode pengelompokan data yang bertujuan menempatkan objek dengan karakteristik serupa ke dalam kelompok yang sama. Dalam bidang kesehatan, teknik clustering telah digunakan untuk mengelompokkan pasien berdasarkan tingkat risiko penyakit, pola diagnosis, maupun karakteristik rekam medis. Hasil pengelompokan tersebut dapat membantu tenaga kesehatan dalam memahami kondisi pasien secara lebih terstruktur serta mendukung penyusunan strategi pelayanan yang sesuai (Wala et al., 2024).

Algoritma *K-Means* merupakan salah satu metode clustering yang banyak digunakan karena memiliki proses perhitungan yang sederhana dan mampu mengelompokkan data dalam jumlah besar secara efektif. Metode ini bekerja dengan membagi data ke dalam beberapa *cluster* berdasarkan kedekatan jarak terhadap titik pusat *cluster* (*centroid*). Penelitian yang dilakukan oleh Juliarti et al. (2023) menunjukkan bahwa algoritma *K-Means* mampu menghasilkan kelompok penyakit lansia berdasarkan karakteristik gejala dan lama menderita penyakit sehingga mempermudah proses identifikasi kondisi kesehatan lansia.

Desa Deli Tua merupakan salah satu wilayah yang memiliki jumlah penduduk lansia cukup besar dengan karakteristik kesehatan yang beragam. Berbagai penyakit kronis ditemukan pada kelompok lansia di wilayah tersebut, namun informasi mengenai pola penyebaran dan pengelompokan penyakit yang dialami masih belum tersedia secara jelas. Kondisi ini menyebabkan proses penyusunan program pencegahan dan pengendalian penyakit kronis belum didukung oleh informasi berbasis data yang memadai. Oleh karena itu, diperlukan suatu metode yang mampu mengelompokkan data kesehatan lansia berdasarkan kesamaan karakteristik penyakit yang diderita.

Beberapa penelitian sebelumnya telah menerapkan algoritma *K-Means* pada data kesehatan, namun sebagian besar berfokus pada pengelompokan pasien rumah sakit atau identifikasi risiko penyakit tertentu. Penelitian yang secara khusus mengkaji pengelompokan penyakit kronis pada warga lansia di tingkat desa masih relatif terbatas. Dengan demikian, penelitian ini memiliki kontribusi dalam memberikan gambaran mengenai karakteristik kelompok penyakit kronis pada lansia berdasarkan kondisi nyata di Desa Deli Tua.

Berdasarkan uraian tersebut, penelitian ini bertujuan untuk menerapkan algoritma *K-Means* dalam mengelompokkan penyakit kronis pada warga lansia di Desa Deli Tua. Hasil pengelompokan yang diperoleh diharapkan dapat memberikan informasi mengenai karakteristik setiap kelompok penyakit kronis sehingga dapat menjadi dasar bagi tenaga kesehatan dan pemerintah setempat dalam merencanakan program pelayanan kesehatan lansia yang lebih efektif dan tepat sasaran.

METODE PENELITIAN

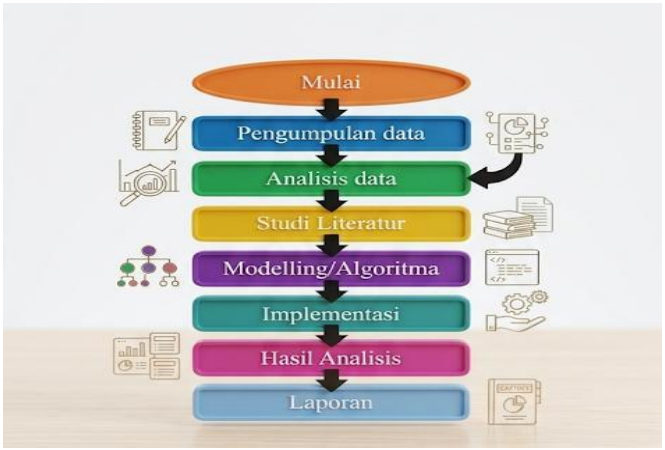
Pendekatan Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan metode data mining untuk mengelompokkan penyakit kronis pada warga lansia di Desa Deli Tua. Teknik yang digunakan adalah *clustering* dengan algoritma *K-Means*. Metode ini dipilih karena mampu mengelompokkan data berdasarkan tingkat kemiripan karakteristik sehingga dapat menghasilkan informasi mengenai pola penyebaran penyakit kronis pada warga lansia. Menurut Maulida et al. (2025), penerapan algoritma *K-Means* dapat membantu proses pengelompokan data secara efektif dan menghasilkan informasi yang mendukung pengambilan keputusan.

Tahapan Penelitian

Tahapan penelitian ini mengikuti model *Knowledge Discovery in Databases* (KDD) yang merupakan kerangka kerja standar dalam penelitian data mining. KDD menyusun serangkaian

proses dari pengumpulan data hingga ekstraksi pengetahuan bermakna dari data mentah (Maulida et al., 2025).



Gambar 1. Tahapan Metode Penelitian

1. Pengumpulan Data
   Pengumpulan data dilakukan melalui observasi, dokumentasi, dan studi pustaka. Data yang digunakan merupakan data kesehatan warga lansia di Desa Deli Tua yang meliputi usia, jenis kelamin, jenis penyakit kronis, dan lama menderita penyakit. Menurut Maulida et al. (2025), kualitas data yang baik sangat mempengaruhi hasil proses *clustering*.
2. Pra-Pemrosesan Data
   Tahap ini dilakukan untuk membersihkan dan menyiapkan data sebelum dianalisis. Proses yang dilakukan meliputi data cleaning, transformasi data, dan normalisasi data. Tujuannya adalah untuk mengurangi kesalahan dan meningkatkan kualitas data yang akan diproses menggunakan algoritma *K-Means*
3. Pemilihan Teknik Clustering
   Penelitian ini menggunakan algoritma *K-Means Clustering*, yaitu salah satu metode *partition-based* yang paling banyak digunakan di berbagai studi ilmiah data mining karena efisiensi dan efektivitas dalam identifikasi kelompok yang jelas (Maulida et al., 2025). Algoritma ini bekerja dengan pembagian data ke dalam sejumlah *cluster* yang ditentukan sebelumnya (K), berdasarkan kemiripan atribut menggunakan ukuran jarak seperti *Euclidean Distance*. Langkah ini dipilih karena praktiknya telah dibuktikan valid dan sering digunakan pada berbagai domain penelitian kuantitatif (Suraya, Sholeh, & Lestari, 2023).
4. Evaluasi Hasil *Clustering*
   Evaluasi dilakukan untuk mengetahui kualitas *cluster* yang terbentuk. Pengujian dilakukan dengan melihat tingkat kedekatan antar anggota dalam cluster dan perbedaan antar *cluster* sehingga dapat diketahui apakah hasil pengelompokan sudah optimal.
5. Interpretasi dan Analisis
   Tahap akhir dilakukan dengan menganalisis hasil pengelompokan yang diperoleh. Setiap cluster diinterpretasikan berdasarkan karakteristik penyakit kronis yang dimiliki warga lansia sehingga dapat memberikan informasi yang bermanfaat bagi pelayanan kesehatan di Desa Deli Tua.

HASIL DAN ANALISIS

Pengumpulan Data

Menjelaskan Pengumpulan data dilakukan untuk memperoleh informasi mengenai penyakit kronis pada warga lansia di Desa Deli Tua. Data diperoleh melalui pendataan langsung terhadap lansia yang menjadi objek penelitian. Data yang dikumpulkan meliputi jenis kelamin, umur, kode gejala penyakit, dan lama mengidap penyakit. Jumlah data yang digunakan dalam penelitian ini sebanyak 20 data lansia. Data tersebut kemudian digunakan sebagai dasar dalam proses pengelompokan menggunakan algoritma *K-Means* untuk mengetahui kelompok lansia dengan karakteristik penyakit kronis yang serupa.

Tabel 1. Data Responden

No	Jenis Kelamin	Umur (th)	Kode Gejala	Lama Mengidap (th)	Gejala
1	laki-laki	68	13	4	Batuk dan Demam
2	laki-laki	61	15	4	Sesak Nafas dan Nyeri Dada
3	perempuan	64	16	7	Pusing dan Lemas
4	laki-laki	63	14	5	Pusing dan Mual
5	perempuan	70	15	4	Sesak Nafas dan Nyeri Dada
6	perempuan	71	11	5	Nyeri dan Sendi Bengkak
7	laki-laki	65	16	3	Pusing dan Lemas
8	perempuan	65	12	6	Penglihatan Buram dan Mati Rasa
9	laki-laki	66	14	4	Pusing dan Mual
10	perempuan	68	19	6	Nyeri Perut dan Sakit saat BAK
11	laki-laki	72	13	5	Batuk dan Demam
12	perempuan	63	11	4	Nyeri dan Sendi Bengkak
13	perempuan	65	15	6	Sesak Nafas dan Nyeri Dada
14	laki-laki	62	12	5	Penglihatan Buram dan Mati Rasa
15	laki-laki	61	14	6	Pusing dan Mual
16	perempuan	67	11	6	Nyeri dan Sendi Bengkak
17	laki-laki	68	19	8	Nyeri Perut dan Sakit saat BAK
18	perempuan	60	13	5	Batuk dan Demam
19	perempuan	73	12	4	Penglihatan Buram dan Mati Rasa
20	laki-laki	71	15	5	Sesak Nafas dan Nyeri Dada

Preprocessing Data

Tahap *preprocessing* data dilakukan untuk menyiapkan data sebelum proses *clustering*. Pada tahap ini dilakukan pengecekan data, pembersihan data yang tidak valid, serta mengubah atribut kategori menjadi numerik agar dapat diproses oleh algoritma *K-Means*. Hasil *preprocessing* menunjukkan bahwa seluruh data telah siap digunakan untuk proses pengelompokan berdasarkan umur, jenis kelamin, kode gejala penyakit, dan lama mengidap penyakit.

Tabel 2. Data Pre-Processing

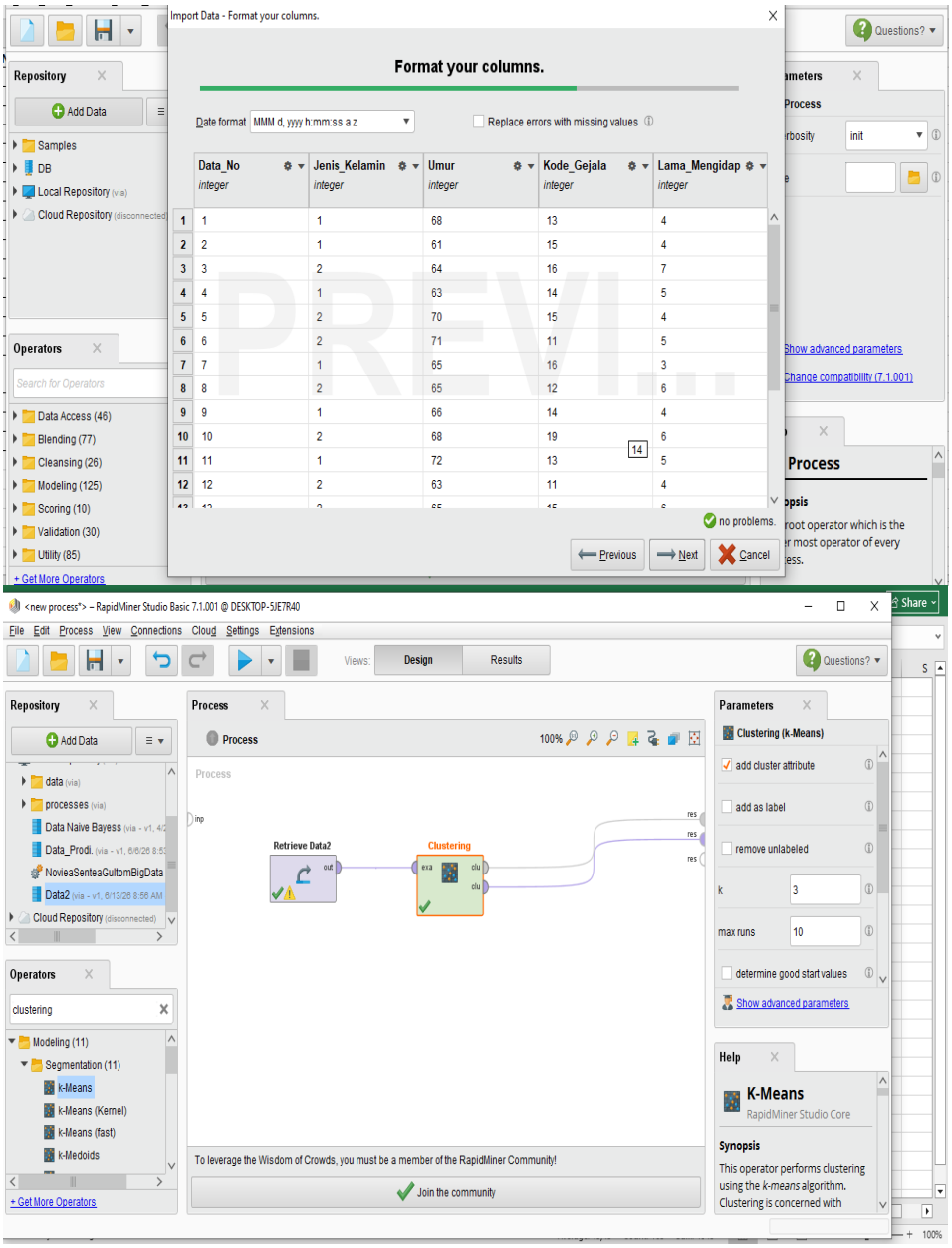
No	Jenis_Kelamin	Umur	Kode_Gejala	Lama_Mengidap
1	1	68	13	4
2	1	61	15	4
3	2	64	16	7
4	1	63	14	5
5	2	70	15	4
6	2	71	11	5
7	1	65	16	3
8	2	65	12	6
9	1	66	14	4
10	2	68	19	6
11	1	72	13	5
12	2	63	11	4
13	2	65	15	6
14	1	62	12	5
15	1	61	14	6
16	2	67	11	6
17	1	68	19	8
18	2	60	13	5

No	Jenis_Kelamin	Umur	Kode_Gejala	Lama_Mengidap
19	2	73	12	4
20	1	71	15	5

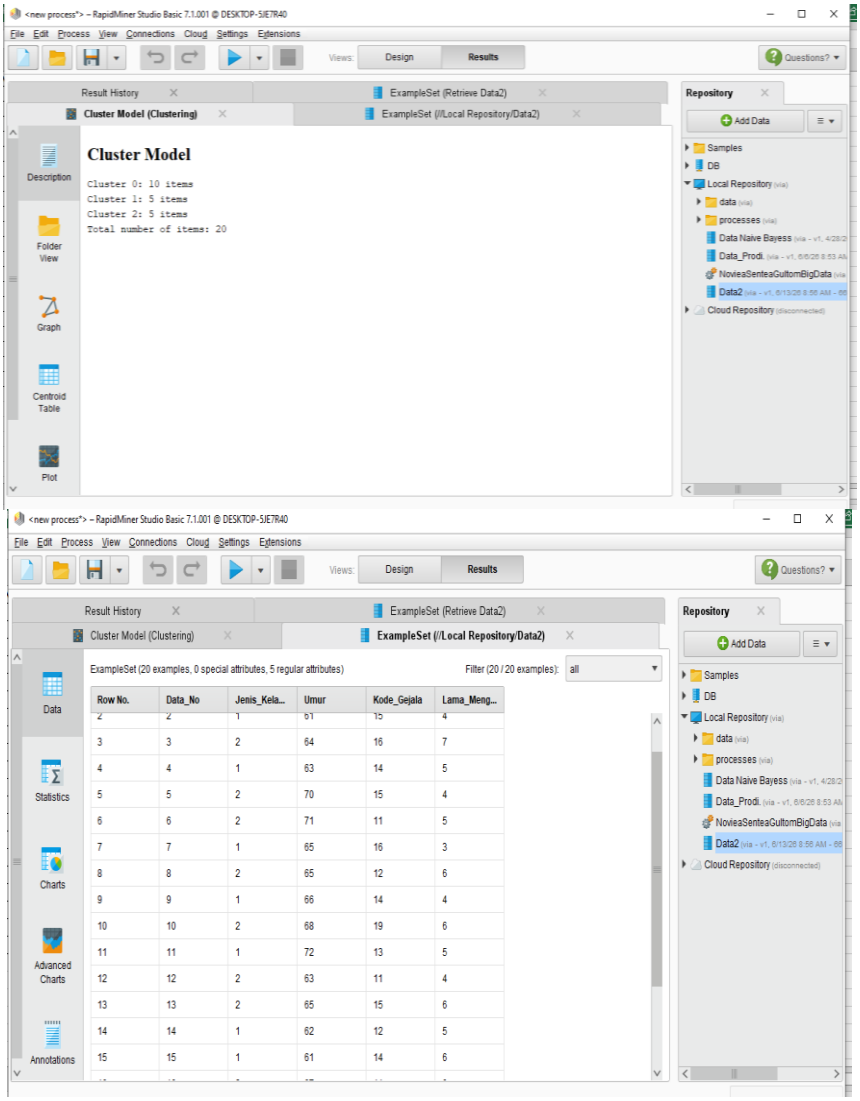
Penerapan Algoritma *K-Means Clustering*

Pada tahap ini, algoritma *K-Means* digunakan untuk mengelompokkan data penyakit kronis pada warga lansia di Desa Deli Tua berdasarkan atribut jenis kelamin, umur, kode gejala penyakit, dan lama mengidap penyakit.

Proses clustering dilakukan dengan menentukan jumlah cluster, kemudian setiap data dikelompokkan berdasarkan kedekatannya dengan *centroid* hingga diperoleh kelompok yang stabil. Hasil pengelompokan menunjukkan bahwa data lansia terbagi ke dalam beberapa *cluster* dengan karakteristik yang berbeda. Pengelompokan ini dapat membantu dalam mengetahui pola penyakit kronis pada lansia dan mendukung perencanaan program kesehatan yang lebih tepat sasaran



Gambar 2. Tahap Pra-Pemrosesan



Gambar 3. *Define Attribute Status*

Pembahasan

Pembahasan dari naskah ini menunjukkan bahwa penerapan algoritma *K-Means Clustering* pada data kesehatan lansia di Desa Deli Tua mampu memberikan gambaran pola penyakit kronis yang dialami oleh responden. Data yang digunakan terdiri dari 20 lansia dengan variasi umur antara 60 hingga 73 tahun, lama mengidap penyakit antara 3 hingga 8 tahun, serta gejala yang beragam seperti batuk dan demam, sesak napas dan nyeri dada, pusing dan mual, nyeri sendi, penglihatan buram, hingga nyeri perut dan gangguan kemih. Proses preprocessing dilakukan dengan mengubah data kategori menjadi numerik sehingga dapat diproses oleh algoritma K-Means menggunakan *Euclidean Distance*. Hasil pengelompokan menunjukkan adanya kecenderungan pola penyakit yang dapat dibagi ke dalam beberapa cluster, misalnya kelompok dengan gejala pernapasan, kelompok dengan gangguan *neurologis*, kelompok dengan gangguan *muskuloskeletal*, kelompok dengan gangguan sensorik, serta kelompok dengan gangguan pencernaan atau kemih. Analisis lebih lanjut memperlihatkan bahwa lansia dengan usia lebih tua cenderung memiliki riwayat penyakit lebih lama, sementara perempuan lebih banyak mengalami gangguan *muskuloskeletal*. Temuan ini memberikan implikasi penting bagi tenaga kesehatan desa untuk menyusun program intervensi yang lebih tepat sasaran, seperti pengendalian penyakit pernapasan, terapi sendi, pemeriksaan mata dan saraf, serta edukasi pola makan dan kesehatan saluran kemih. Dengan demikian, penelitian ini menegaskan bahwa algoritma K-Means dapat menjadi metode efektif dalam mengidentifikasi pola penyakit kronis pada lansia dan mendukung perencanaan pelayanan kesehatan berbasis data di tingkat desa.

KESIMPULAN

Kesimpulan Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa algoritma *K-Means Clustering* dapat diterapkan untuk mengelompokkan penyakit kronis pada warga lansia di Desa Deli Tua berdasarkan karakteristik yang dimiliki, yaitu jenis kelamin, umur, kode gejala penyakit, dan lama mengidap penyakit. Hasil pengelompokan menunjukkan bahwa data lansia dapat dibagi ke dalam beberapa kelompok yang memiliki karakteristik kesehatan yang serupa. Pengelompokan tersebut membantu dalam mengidentifikasi pola penyakit kronis pada lansia sehingga dapat memberikan informasi yang bermanfaat bagi pihak terkait dalam menyusun program pelayanan dan pencegahan penyakit yang lebih tepat sasaran. Dengan demikian, penerapan algoritma *K-Means* dapat menjadi salah satu metode yang efektif untuk mengelompokkan data penyakit kronis pada lansia dan mendukung pengambilan keputusan di bidang kesehatan masyarakat.

REFERENSI

Rahmawati et al. (2023) – Penerapan Data Mining untuk Menentukan Penyebab Kematian di Indonesia Menggunakan Metode Clustering K-Means.

Ardini, D. Z., & Rhodiyah. (2025). Analisis Pola Penyakit Kronis pada Lansia Menggunakan K-Means Clustering di Puskesmas Kelurahan Semper Barat. Jurnal Indonesia Manajemen Informatika dan Komunikasi.

Aulia, W., dkk. (2025). Analisis Algoritma K-Means Clustering dalam Identifikasi Tingkat Risiko Penyakit Berdasarkan Data Rekam Medis Pasien. Journal of Science and Social Research.

Wala, J., Herman, & Umar, R. (2024). Implementasi K-Means Clustering pada Pengelompokan Pasien Penyakit Jantung. JISKA.

Juliarti, K., Negara, A. B. P., & Tursina. (2023). Sistem Pengelompokan Diagnosis Penyakit Pada Lansia Menggunakan Metode K-Means Clustering Berbasis Web. JURISTI.

Lestari, L. S. (2024). Analisis Clustering Penduduk Berdasarkan Kelompok Umur dengan K-Means dan Hierarchical Clustering untuk Perencanaan Demografi. Jurnal Komputer dan Informatika.

M. I. Firdaus and P. A. R. Devi, "Implementation of K-Means Algorithm for Diseases Clustering in Elderly Posyandu Participants," Applied Technology and Computing Science Journal, vol. 5, no. 2, pp. 90–99, 2022, doi: 10.33086/atcsj.v5i2.3691.

Y. Sheng, R. Bond, R. Jaiswal, J. Dinsmore, and J. Doyle, "Augmenting K-Means Clustering With Qualitative Data to Discover the Engagement Patterns of Older Adults With Multimorbidity When Using Digital Health Technologies: Proof-of-Concept Trial," Journal of Medical Internet Research, vol. 26, 2024, doi: 10.2196/46287.

M. N. Kadafi and A. Finandhita, "Penerapan Data Mining dalam Pengelompokan Penyakit untuk Rekomendasi Materi Penyuluhan Kesehatan di Klinik Keluarga Cianjur," Komputa: Jurnal Ilmiah Komputer dan Informatika, vol. 13, no. 1, 2024, doi: 10.34010/komputa.v13i1.11273.

J. Wala, Herman, and R. Umar, "Implementasi K-Means Clustering pada Pengelompokan Pasien Penyakit Jantung," JISKA, 2024.

M. I. Hutagalung and Sriani, "Pengelompokan Data Penyakit THT Menggunakan Algoritma K-Means Clustering," MALCOM: Indonesian Journal of Machine Learning and Computer Science, vol. 4, no. 4, 2024, doi: 10.57152/malcom.v4i4.1692.

"How predictive of future healthcare utilisation and mortality is data-driven population segmentation based on healthcare utilisation and chronic condition comorbidity?" BMC Public Health, 2024.

L. S. Lestari, "Analisis Clustering Penduduk Berdasarkan Kelompok Umur dengan K-Means dan Hierarchical Clustering untuk Perencanaan Demografi," Jurnal Komputer dan Informatika, 2024.

C. F. Tuzahra and Whydiantoro, "Data Mining Approach to the Evaluation of the Most Common Diseases: K-Means Clustering Study in Pustu Lapang Sedar in 2023," Journal of Mechanical and Manufacture, vol. 5, no. 1, 2025, doi: 10.31949/jmm.v5i1.12233.

D. Z. Ardini and R. Rhodiyah, "Analisis Pola Penyakit Kronis pada Lansia Menggunakan K-Means Clustering di Puskesmas Kelurahan Semper Barat," Jurnal Indonesia: Manajemen Informatika dan Komunikasi, vol. 6, no. 3, pp. 1842–1853, 2025, doi: 10.63447/jimik.v6i3.1595.