

Analisis Pola Penyakit Kronis pada Lansia Menggunakan *K-Means Clustering* di Puskesmas Kelurahan Semper Barat

Dea Zerlinda Ardini ^{1*}, Rhodiyah ²

^{1,2} Program Studi Sistem Informasi, Sekolah Tinggi Ilmu Komputer Cipta Karya Informatika, Kota Jakarta Timur, Daerah Khusus Ibukota Jakarta, Indonesia.

Corresponding Email: deazerlindaandini@gmail.com ^{1*}

Histori Artikel:

Dikirim 30 Juli 2025; *Diterima dalam bentuk revisi* 10 Agustus 2025; *Diterima* 30 Agustus 2025; *Diterbitkan* 10 September 2025. Semua hak dilindungi oleh Lembaga Penelitian dan Pengabdian Masyarakat (LPPM) STMIK Indonesia Banda Aceh.

Abstrak

Penyakit kronis telah menjadi permasalahan kesehatan sekaligus menjadi penyebab kematian terbesar di dunia. Penyakit kronis penyebab utama kematian secara global adalah penyakit jantung dan pembuluh darah, penyakit pernapasan kronis dan penyakit metabolik seperti diabetes (WHO, 2018). Kelurahan Semper Barat merupakan wilayah dengan jumlah lansia yang cukup signifikan dan akses pelayanan kesehatan yang berpusat di Puskesmas Kelurahan Semper Barat. Namun, sejauh ini belum terdapat studi yang secara khusus menganalisis pola penyakit kronis lansia di wilayah tersebut menggunakan pendekatan data mining. Dalam penelitian ini memiliki keterbaruan berupa sebuah pembelajaran kasus klusterisasi penyakit kronis masyarakat wilayah kelurahan menggunakan metode data mining dengan algoritma K-Means yang hasilnya didapatkan model ini mampu memperlihatkan nilai yang presisi dalam melakukan klusterisasi penyakit kronis. Hasil klusterisasi yang diperoleh dapat dimanfaatkan oleh pihak Puskesmas Kelurahan Semper Barat sebagai dasar keputusan melakukan penyuluhan dan penanganan tepat sasaran sehingga mempermudah penjangkauan kepada lansia yang sudah memiliki riwayat penyakit kronis sesuai dengan kelompok penyakitnya. Hasil pengujian dari Cluster Distance Performance sebelumnya menunjukkan nilai evaluasi 0.579 untuk 2 cluster, yang paling mendekati nol dibandingkan jumlah cluster lain dalam konteks algoritma K-Means, nilai yang semakin mendekati 0 menunjukkan bahwa data dalam satu cluster memiliki banyak kemiripan data dan jarak antar cluster cukup terpisah.

Kata Kunci: Analisa; Pola Penyakit Kronis; Lansia; Puskesmas; Klusterisasi; K-Means Clustering; Data Mining.

Abstract

Chronic diseases have become a major health problem as well as the leading cause of death worldwide. The main chronic diseases causing death globally are cardiovascular diseases, chronic respiratory diseases, and metabolic diseases such as diabetes (WHO, 2018). Semper Barat Subdistrict is an area with a significant number of elderly residents, with healthcare services centered at the Semper Barat Community Health Center (Puskesmas). However, so far, there has been no study that specifically analyzes the patterns of chronic diseases among the elderly in this area using a data mining approach. This study presents a novelty in the form of a case study on clustering chronic diseases within the community of the subdistrict using data mining with the K-Means algorithm. The results show that this model is capable of providing precise values in clustering chronic diseases. The clustering results can be utilized by the Semper Barat Community Health Center as a basis for decision-making in conducting targeted outreach and treatment, thereby facilitating better access to elderly individuals who already have a history of chronic diseases according to their disease group. The testing results from the previous Cluster Distance Performance showed an evaluation value of 0.579 for two clusters, which was the closest to zero compared to other numbers of clusters. In the context of the K-Means algorithm, values closer to zero indicate that the data within a cluster have greater similarity, and the distance between clusters is sufficiently distinct.

Keyword: Analysis; Chronic Disease Patterns; Elderly; Community Health Center; Clustering; K-Means Clustering; Data Mining.

1. Pendahuluan

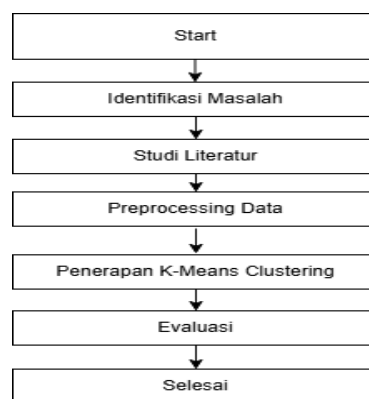
Penyakit kronis telah menjadi masalah utama yang dihadapi banyak negara di dunia, termasuk Indonesia. Menurut data *World Health Organization* (WHO, 2018), penyakit jantung, stroke, diabetes, dan gangguan pernapasan kronis telah menjadi penyebab kematian terbesar secara global, bertanggung jawab atas hampir 74% kematian yang tercatat setiap tahun. Selain itu, lebih dari 15 juta orang meninggal sebelum mencapai usia 70 tahun akibat penyakit tidak menular ini. Kejadian kematian akibat penyakit kronis ini semakin memperjelas dampaknya terhadap kualitas hidup masyarakat dan menunjukkan tantangan besar yang dihadapi oleh sistem pelayanan kesehatan di seluruh dunia. Indonesia, sebagai negara dengan populasi terbesar, turut mengalami dampak signifikan dari peningkatan jumlah lansia yang menderita penyakit kronis. Berdasarkan data dari Badan Pusat Statistik (BPS, 2011) dan Kementerian Kesehatan Indonesia, jumlah lansia di Indonesia pada tahun 2010 tercatat sebanyak 18,1 juta jiwa, atau sekitar 9,6% dari total populasi. Proyeksi pada tahun 2030 jumlah lansia diperkirakan akan meningkat menjadi 36 juta jiwa. Kenaikan angka ini menunjukkan perubahan besar dalam struktur demografis Indonesia dan menuntut perhatian lebih terhadap penanganan penyakit yang dihadapi oleh kelompok lansia, terutama penyakit kronis yang dapat mempengaruhi kualitas hidup mereka. Di sisi lain, beban ini juga menambah tantangan bagi sistem kesehatan, terutama pada layanan primer seperti Puskesmas. Menurut data BPS (2011), sekitar 28,35% lansia berusia 60-69 tahun di Indonesia mengalami gangguan kesehatan terkait penyakit kronis, angka yang terus meningkat pada kelompok usia lebih tua. Masalah kesehatan ini tidak hanya memengaruhi kualitas hidup individu, tetapi juga meningkatkan beban sistem kesehatan, terutama di Puskesmas yang menjadi ujung tombak pelayanan medis masyarakat. Sayangnya, data yang dikumpulkan selama kunjungan pasien sering kali hanya dimanfaatkan untuk keperluan administratif, tanpa ada upaya analisis lebih lanjut yang dapat mengungkapkan pola penyakit dan kecenderungan yang berkembang di kalangan lansia. Hal ini menghambat kemampuan Puskesmas untuk merespons secara efektif terhadap kebutuhan kesehatan lansia. Keberagaman jenis penyakit kronis yang diderita lansia, serta peningkatan jumlah kasus, menunjukkan bahwa sudah saatnya untuk melakukan analisis yang lebih sistematis terhadap data kesehatan lansia. Salah satu metode yang dapat digunakan untuk menganalisis dan mengelompokkan pola penyakit ini adalah data mining, dengan memanfaatkan teknik clustering. K-Means Clustering adalah salah satu algoritma yang dapat digunakan untuk mengelompokkan data penyakit berdasarkan karakteristik tertentu, misalnya pola penyakit yang ada pada lansia. Pendekatan ini memungkinkan Puskesmas untuk mengelompokkan lansia berdasarkan jenis penyakit yang mereka derita, sehingga strategi intervensi dapat lebih disesuaikan dengan kondisi tiap kelompok.

Penelitian sebelumnya menunjukkan bahwa data mining, terutama menggunakan K-Means Clustering, dapat membantu mengidentifikasi pola penyakit pada berbagai kelompok usia, termasuk lansia. Beberapa studi mengungkapkan bahwa algoritma ini efektif untuk mengelompokkan data penyakit secara lebih terstruktur, memudahkan pemahaman terkait penyebaran penyakit, serta memberikan gambaran yang lebih jelas mengenai distribusi penyakit di kalangan populasi tertentu (Ariyanto, 2022; Apriyani *et al.*, 2023). Dengan mengelompokkan lansia berdasarkan kondisi kesehatan mereka, seperti tekanan darah atau kadar gula darah, Puskesmas dapat merancang intervensi kesehatan yang lebih tepat sasaran, serta memberikan perawatan yang lebih sesuai dengan kondisi masing-masing kelompok lansia (Aulia, 2021; Gustrianda & Mulyana, 2022). Puskesmas memainkan peran kunci dalam menyediakan layanan kesehatan dasar kepada masyarakat, terutama lansia yang lebih rentan terhadap penyakit kronis. Oleh karena itu, dengan mengaplikasikan K-Means Clustering untuk menganalisis pola penyakit kronis di Puskesmas Kelurahan Semper Barat, dapat diperoleh gambaran lebih jelas tentang karakteristik kesehatan lansia. Proses ini diharapkan dapat membantu Puskesmas untuk merancang program pencegahan dan perawatan yang lebih tepat sasaran, serta memberikan edukasi kepada lansia mengenai cara mencegah dan mengelola penyakit kronis mereka. Misalnya, kelompok lansia yang memiliki hipertensi tunggal dapat diberi informasi lebih lanjut mengenai pengelolaan tekanan darah, sedangkan kelompok dengan penyakit komorbid,

seperti hipertensi dan diabetes, memerlukan perhatian khusus dalam perawatan (Narulita *et al.*, 2023). Dengan menggunakan metode K-Means Clustering dalam penelitian ini, diharapkan dapat ditemukan pola penyakit kronis yang lebih jelas di kalangan lansia di Puskesmas Kelurahan Semper Barat. Analisis yang lebih terperinci terhadap data kesehatan lansia ini akan memberikan informasi penting untuk perencanaan pelayanan kesehatan yang lebih efektif dan berbasis data. Di samping itu, penelitian ini dapat menjadi acuan bagi Puskesmas dalam merumuskan kebijakan yang lebih responsif terhadap kebutuhan kesehatan lansia, yang pada akhirnya akan berkontribusi pada pengurangan angka kematian dan peningkatan kualitas hidup lansia di Indonesia (Sugianto *et al.*, 2020; Sudrajat *et al.*, 2022).

2. Metode Penelitian

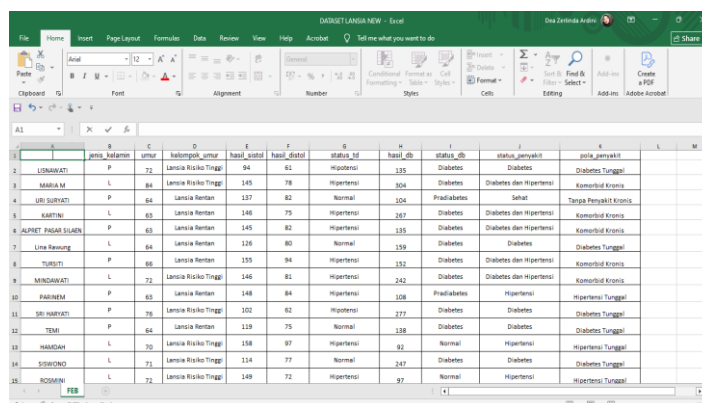
Penelitian ini mengikuti serangkaian langkah yang dimulai dengan pemilihan data dari Puskesmas Kelurahan Semper Barat. Data yang digunakan mencakup informasi mengenai kesehatan lansia, seperti usia, jenis kelamin, tekanan darah, indeks massa tubuh (IMT), dan riwayat penyakit. Langkah pertama adalah mengidentifikasi data yang relevan untuk analisis, kemudian diikuti oleh pembersihan untuk menghilangkan nilai kosong atau data yang tidak konsisten. Pembersihan ini memastikan bahwa data yang digunakan dalam penelitian memiliki kualitas yang baik, sehingga hasil yang diperoleh dapat diandalkan. Setelah data dibersihkan, langkah berikutnya adalah transformasi data. Pada tahap ini, data yang terpilih diubah menjadi format yang sesuai dengan algoritma yang akan digunakan. Misalnya, variabel kategori seperti jenis kelamin diubah menjadi bentuk numerik agar dapat diproses dengan baik. Proses ini juga mencakup normalisasi, sehingga data yang memiliki skala berbeda bisa dianalisis tanpa adanya bias yang disebabkan oleh perbedaan rentang nilai. Langkah berikutnya adalah penerapan algoritma K-Means Clustering, yang digunakan untuk mengelompokkan lansia berdasarkan pola penyakit yang mereka alami. K-Means Clustering adalah metode yang membagi data ke dalam beberapa kelompok atau kluster berdasarkan kemiripan karakteristik di antara data. Algoritma ini memungkinkan pemetaan pola penyakit pada lansia, seperti hipertensi atau diabetes, ke dalam grup-grup yang serupa. Setelah kluster terbentuk, hasilnya dievaluasi untuk memastikan pengelompokan yang dihasilkan memiliki kualitas yang sesuai, dengan membandingkan kedekatan data dalam satu kluster dan jarak antar kluster. Untuk mendapatkan kluster yang paling relevan, penelitian ini juga menggunakan *RapidMiner 2025.2* sebagai alat untuk memproses data dan menentukan jumlah kluster yang optimal. Data yang telah dipersiapkan diimpor ke dalam *RapidMiner*, kemudian diproses menggunakan algoritma K-Means. Penentuan jumlah kluster disesuaikan dengan struktur data dan tujuan analisis. Proses ini memungkinkan peneliti untuk mengevaluasi performa algoritma dan memilih jumlah kluster yang memberikan hasil yang paling representatif.



Gambar 1. Tahapan Penelitian

3.1.2 Data Selection

Pada tahap ini, dilakukan pemilihan atribut data yang relevan dari kumpulan data untuk digunakan dalam analisis atau pemrosesan lebih lanjut. Proses ini bertujuan untuk memastikan bahwa hanya data yang memiliki hubungan langsung dengan tujuan penelitian yang digunakan. Pemilihan atribut yang tepat sangat penting untuk meningkatkan kualitas analisis dan hasil yang diperoleh. Dalam konteks penelitian ini, atribut yang dipilih mencakup informasi seperti jenis kelamin, usia, hasil tekanan darah (sistol dan diastol), serta pola penyakit yang dialami oleh lansia. Dengan menyaring atribut yang relevan, proses analisis dapat dilakukan dengan lebih fokus dan efisien, sehingga menghasilkan kluster yang lebih bermakna dan sesuai dengan tujuan penelitian.



Gambar 3. Tampilan Data Selection

3.1.3 Data Transformasi

Pada tahap ini, data yang telah dipilih akan diubah menjadi format yang lebih sesuai untuk proses pengelompokan. Misalnya, untuk data tekanan darah (sistol dan diastol), dilakukan transformasi berdasarkan nilai yang terukur. Jika nilai sistol dan diastol lebih tinggi dari 140, maka statusnya akan dikategorikan sebagai hipertensi. Jika nilai sistol dan diastol berada dalam kisaran 140 hingga 100, statusnya akan dianggap normal, dan jika nilainya di bawah 100, maka statusnya akan dianggap hipotensi. Begitu juga dengan data terkait diabetes, di mana jika nilai glukosa darah (status_db) menunjukkan angka di bawah 130, maka statusnya akan dikategorikan sebagai diabetes, sementara nilai di bawah 100 akan dianggap normal. Setelah transformasi ini, kedua status tersebut akan digabungkan untuk mendapatkan kategori akhir mengenai pola penyakit pada lansia. Kategori yang terbentuk antara lain meliputi diabetes tunggal, hipertensi tunggal, serta kombinasi keduanya, yaitu diabetes dan hipertensi yang dikenal sebagai komorbid kronis. Sebaliknya, jika seorang lansia tidak memiliki riwayat kedua penyakit ini, statusnya akan ditandai sebagai tanpa penyakit kronis. Hasil dari transformasi ini akan memberikan label akhir pada setiap data yang digunakan dalam analisis klusterisasi.

jenis_kelamin	umur	hasil_sistol	hasil_diastol	hasil_db	status_penyakit	pola_penyakit
P	72	94	61	135	Diabetes	Diabetes Tunggal
L	84	145	78	304	Diabetes dan Hipertensi	Komorbid Kronis
P	64	137	82	104	Sehat	Tanpa Penyakit Kronis
L	63	146	75	267	Diabetes dan Hipertensi	Komorbid Kronis
P	63	145	82	135	Diabetes dan Hipertensi	Komorbid Kronis
L	64	126	80	159	Diabetes	Diabetes Tunggal
P	66	155	94	152	Diabetes dan Hipertensi	Komorbid Kronis
L	72	146	81	242	Diabetes dan Hipertensi	Komorbid Kronis
P	63	148	84	108	Hipertensi	Hipertensi Tunggal
P	76	102	62	277	Diabetes	Diabetes Tunggal
P	64	119	75	138	Diabetes	Diabetes Tunggal
L	70	158	97	92	Hipertensi	Hipertensi Tunggal
L	71	114	77	247	Diabetes	Diabetes Tunggal

Gambar 4. Tampilan Data Hasil Tranformasi

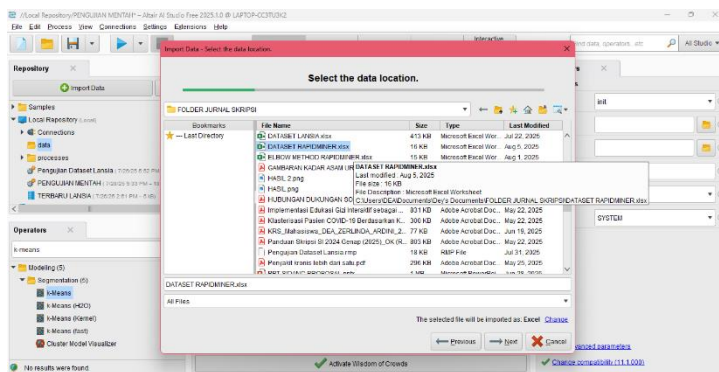
3.1.4 Implementasi RapidMiner

Proses implementasi metode K-Means Clustering menggunakan *RapidMiner* dilakukan melalui serangkaian tahapan dalam workflow analisis data. Langkah pertama adalah mengimpor data yang telah ditransformasi ke dalam *RapidMiner* menggunakan operator *Read Excel*. Setelah data berhasil diimpor, operator *Multiply* digunakan untuk menjalankan beberapa eksperimen dengan jumlah kluster yang berbeda, yaitu 2 kluster, 3 kluster, dan 4 kluster. Setelah itu, operator *K-Means Clustering* diterapkan untuk mengelompokkan data berdasarkan atribut yang telah ditentukan. Selanjutnya, pengaturan dilakukan dengan memilih *mixed measures* dan menggunakan *Euclidean distance* sebagai metode pengukuran jarak antar data. Proses ini kemudian dijalankan untuk menghasilkan sejumlah kluster, posisi centroid dari masing-masing kluster, serta visualisasi distribusi data yang dapat dilihat melalui bentuk *scatter plot*. Hasil dari proses klusterisasi ini dianalisis lebih lanjut menggunakan *Cluster Model* dan *Result Viewer* dalam *RapidMiner*. Dengan menggunakan alat ini, peneliti dapat memeriksa karakteristik kesamaan data dalam setiap kluster, serta membandingkan hasil klusterisasi. Selain itu, *Description* digunakan untuk menilai kluster mana yang memiliki nilai evaluasi paling optimal, yang menunjukkan pengelompokan yang paling sesuai dengan data yang ada.

3.1.5 Tahap Pengujian Menggunakan RapidMiner

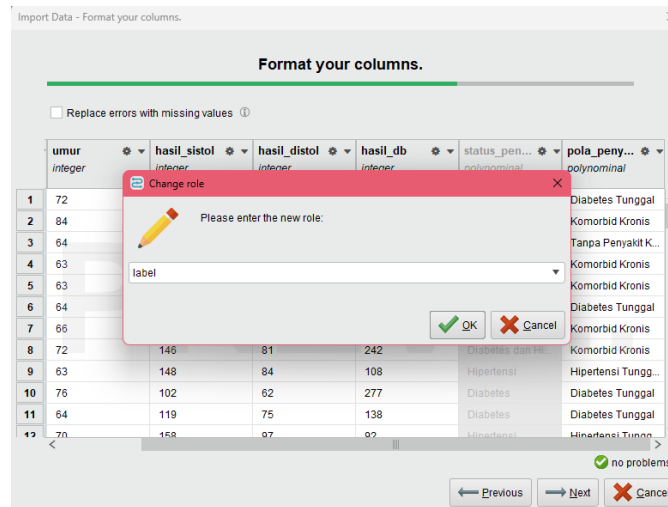
1) Tahap pertama

Langkah pertama dalam pengujian adalah menyiapkan data rekap bulanan kesehatan lansia yang telah ditransformasi dalam format Excel. Data ini akan diimpor ke dalam *RapidMiner* untuk diproses lebih lanjut. Proses transformasi sebelumnya memastikan bahwa data yang digunakan sudah dalam format yang sesuai untuk analisis menggunakan algoritma K-Means. Setelah data siap, file Excel tersebut akan dimuat ke dalam *RapidMiner* menggunakan operator *Read Excel* agar dapat digunakan dalam tahap berikutnya dari proses pengolahan data.



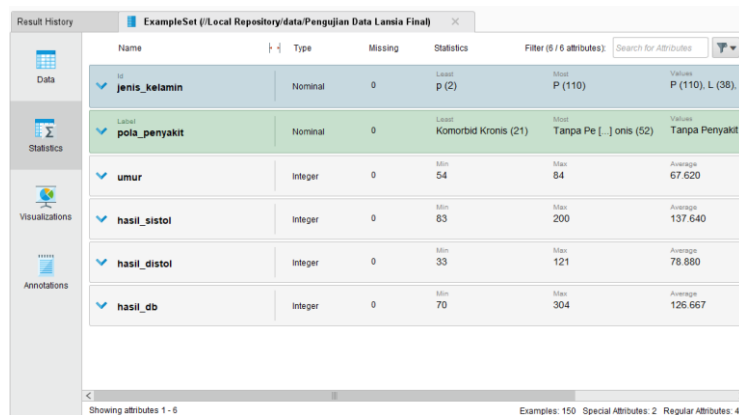
Gambar 5. Tampilan Import Data

Setelah data berhasil diimpor, langkah berikutnya adalah melakukan tahap *preprocessing*. Proses ini melibatkan beberapa tahapan penting, yaitu pembersihan data dari *missing values*, normalisasi variabel numerik agar semua atribut memiliki skala yang sebanding, serta konversi data kategori menjadi format numerik jika diperlukan. Pembersihan data bertujuan untuk memastikan bahwa tidak ada nilai yang hilang atau tidak konsisten yang dapat mempengaruhi hasil analisis. Selanjutnya, normalisasi dilakukan untuk menyamakan skala variabel-variabel numerik, sehingga atribut yang memiliki rentang nilai berbeda tidak mendominasi hasil klusterisasi. Selain itu, untuk mempersiapkan data kategori agar dapat diproses oleh algoritma, variabel-variabel kategorikal seperti jenis kelamin dan pola penyakit diubah menjadi format numerik. Misalnya, *role* pada jenis kelamin diubah menjadi "id", sementara pola penyakit diberi label yang sesuai, seperti "label" untuk mendefinisikan kategori penyakit pada setiap individu. Proses-proses ini memastikan bahwa data yang digunakan dalam analisis sudah siap dan sesuai untuk diterapkan dalam model K-Means.



Gambar 6. Type Data

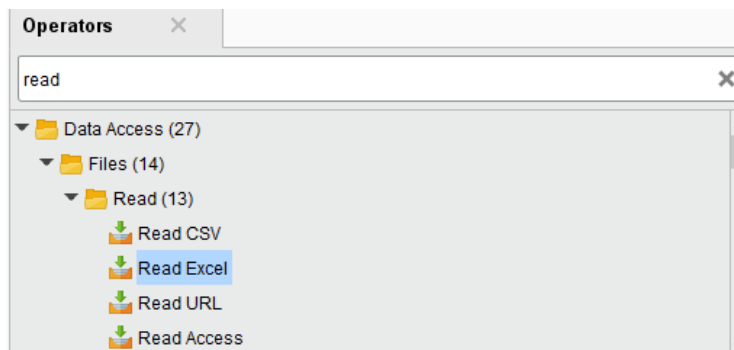
Setelah tahap *preprocessing* selesai dilakukan pada data lansia, hasil yang diperoleh sudah sesuai dengan format dan tipe data yang diperlukan untuk proses analisis lebih lanjut. Hal ini dapat dilihat pada gambar berikut, yang menunjukkan bahwa data telah diproses dan disesuaikan dengan tipe yang disarankan, sehingga siap untuk digunakan dalam penerapan algoritma *K-Means Clustering*.



Gambar 7. Example Set

2) Tahap Kedua

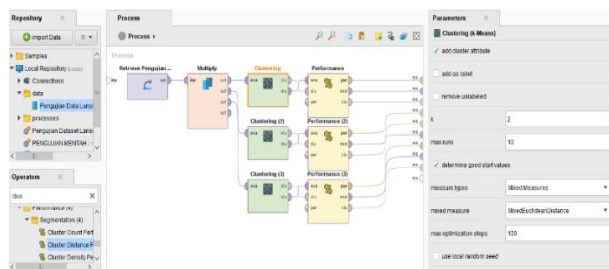
Pada tahap kedua, proses *K-Means Clustering* dilakukan dengan menggunakan beberapa operator di *RapidMiner* yang telah dipersiapkan sebelumnya. Operator pertama yang digunakan adalah *Read Excel*, yang berfungsi untuk membaca file Excel yang telah diimpor ke dalam penyimpanan *RapidMiner*. Setelah data berhasil dibaca, operator *Multiply* digunakan untuk memungkinkan pengujian dengan lebih dari satu operator secara bersamaan, sehingga dapat menjalankan beberapa proses klusterisasi dalam satu waktu. Selanjutnya, operator *K-Means* diterapkan untuk melakukan proses pengelompokan data berdasarkan kemiripan atribut yang relevan. Terakhir, operator *Cluster Distance Performance* digunakan untuk menentukan nilai rata-rata dan titik pusat (centroid) masing-masing kluster, yang berguna untuk mengevaluasi kualitas pengelompokan yang telah dilakukan dan memastikan bahwa hasil klusterisasi sesuai dengan tujuan penelitian.



Gambar 8. Operator

3) Tahap Ketiga

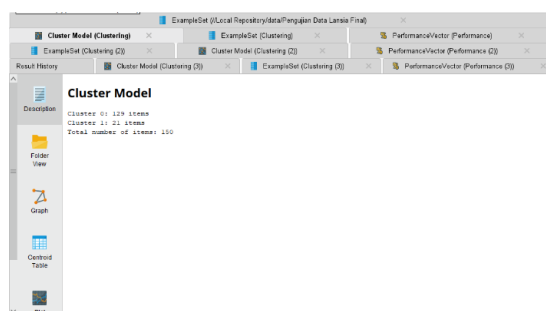
Pada tahap ketiga, langkah yang dilakukan adalah *drag and drop* operator ke dalam halaman desain pada kolom *Process* di *RapidMiner*. Operator *Multiply* digunakan untuk memungkinkan pengujian pada data dengan jumlah kluster yang berbeda, sehingga hasil klusterisasi yang paling optimal dapat ditemukan. Setiap operator yang digunakan disambungkan dengan garis koneksi agar proses dapat dijalankan secara bersamaan. Di kolom parameter, jumlah kluster yang akan digunakan untuk pengujian ditentukan, dan pengaturan pengukuran dilakukan dengan memilih jenis pengukuran yang sesuai, yaitu *Mixed Measure* dengan menggunakan *Mixed Euclidean Distance*. Hal ini dilakukan untuk memastikan bahwa pengukuran jarak antar data dapat dilakukan secara tepat dan sesuai dengan tujuan analisis klusterisasi.



Gambar 9. Tampilan Process

4) Tahap Keempat

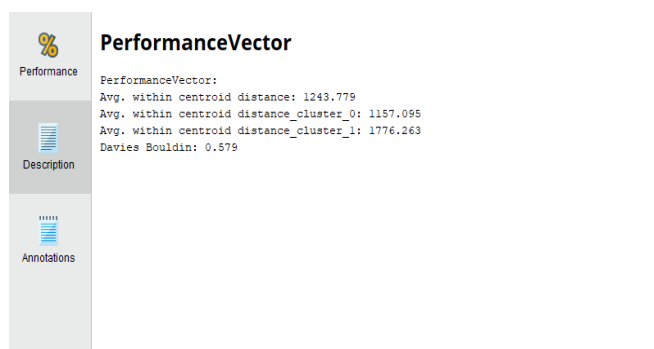
Pada tahap keempat, setelah proses klusterisasi berhasil dijalankan, hasilnya dapat dilihat pada bagian *Results* di *RapidMiner*. Seperti yang terlihat pada gambar berikut, hasil pengujian dengan jumlah 2 kluster menunjukkan bahwa kluster 0 berisi 129 data, sementara kluster 1 berisi 21 data, dengan total data yang terklusterisasi sebanyak 150. Hasil ini memberikan gambaran tentang bagaimana data terbagi ke dalam kluster-kluster berdasarkan karakteristik yang telah dianalisis, yang selanjutnya dapat digunakan untuk evaluasi lebih lanjut mengenai pola penyakit pada lansia.



Gambar 10. Cluster Model

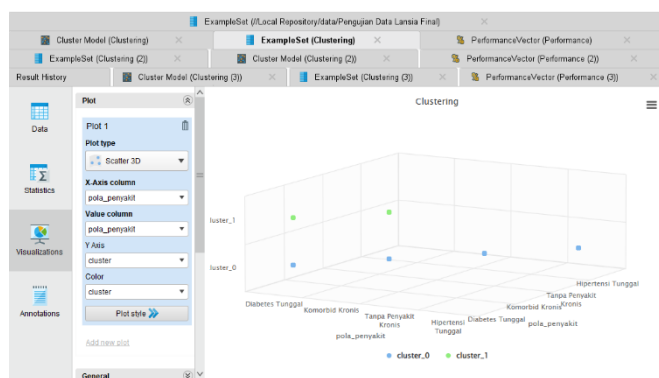
5) Tahapa Kelima

Pada tahap kelima, hasil penggunaan operator *Cluster Distance Performance* dapat dilihat pada gambar berikut, yang menunjukkan perbandingan hasil dari beberapa proses klasterisasi. Pengujian dilakukan dengan jumlah klaster yang berbeda, dan tujuan utamanya adalah untuk mencari jumlah klaster yang menghasilkan nilai yang paling mendekati angka 0, yang menandakan pengelompokan yang lebih terpisah dan data dalam setiap klaster memiliki kemiripan tinggi. Pada pengujian dengan 2 klaster, hasil evaluasi menunjukkan nilai 0.579, sementara pengujian dengan 3 klaster menghasilkan nilai 0.928, dan pengujian dengan 4 klaster menghasilkan nilai 0.931. Berdasarkan hasil ini, klaster dengan jumlah 2 menunjukkan nilai yang paling mendekati 0, yang berarti bahwa jumlah 2 klaster adalah jumlah yang paling optimal untuk analisis ini.



Gambar 11. Performance Vector

Pada tahap selanjutnya, untuk memvisualisasikan hasil klasterisasi, dipilihlah plot *scatter 3D*. Pada plot ini, sumbu X diisi dengan atribut *pola_penyakit*, yang menggambarkan jenis penyakit yang diderita oleh lansia. Sumbu Y diisi dengan nilai klaster, yang menunjukkan kelompok klaster hasil pengelompokan data. Setiap titik pada plot tersebut diberi warna yang berbeda berdasarkan klaster masing-masing, sehingga memudahkan untuk melihat distribusi data dan bagaimana setiap kelompok terpisah satu sama lain dalam ruang tiga dimensi. Visualisasi ini memungkinkan peneliti untuk lebih mudah mengidentifikasi karakteristik dan perbedaan antar klaster yang terbentuk.



Gambar 12. Scatter 3D

3.2 Pembahasan

Pengelompokan data lansia berdasarkan hasil pemantauan kesehatan bulanan, seperti usia, tekanan darah, dan kadar gula darah, dilakukan menggunakan metode *K-Means Clustering*. Proses ini dimulai dengan pengumpulan data, pemilihan atribut yang relevan, serta transformasi data untuk memastikan kesesuaiannya dengan kebutuhan analisis. Data yang digunakan berasal dari rekap bulanan pemantauan kesehatan lansia, berjumlah 150 data, yang kemudian dianalisis menggunakan

RapidMiner. Penggunaan metode ini terbukti efektif dalam menangani data yang besar dan beragam, seperti yang juga ditemukan dalam studi sebelumnya oleh Sudrajat *et al.* (2022) dan Sugianto *et al.* (2020), yang menerapkan algoritma yang sama untuk pengelompokan data kesehatan. Hasil analisis menunjukkan bahwa dua kluster adalah jumlah yang paling optimal. Kluster pertama (kluster 0) terdiri dari 129 data yang mencakup pola penyakit seperti diabetes tunggal, komorbid kronis, tanpa penyakit kronis, serta hipertensi tunggal. Sementara itu, kluster kedua (kluster 1) berisi 21 data dengan pola penyakit yang lebih terfokus pada diabetes tunggal dan komorbid kronis. Ini membuktikan bahwa data dapat dikelompokkan secara efektif berdasarkan pola penyakit yang ditemukan pada lansia, sejalan dengan temuan Ariyanto (2022), yang menggunakan *K-Means* untuk mengelompokkan penyakit saluran pernapasan. Visualisasi dalam bentuk *scatter 3D* memperlihatkan bahwa titik-titik dalam kluster yang sama cenderung berkelompok di area tertentu, yang menegaskan bahwa algoritma *K-Means* dapat mengelompokkan data dengan baik berdasarkan kemiripan karakteristik penyakit. Temuan ini konsisten dengan penelitian Gustrianda & Mulyana (2022), yang juga menggunakan *K-Means* untuk mengelompokkan data kesehatan dan menemukan hasil yang serupa.

Evaluasi menggunakan *Cluster Distance Performance* menunjukkan nilai 0.579 untuk dua kluster, yang lebih mendekati angka 0 dibandingkan pengujian dengan tiga atau empat kluster. Nilai ini mengindikasikan bahwa data dalam satu kluster memiliki kemiripan yang tinggi, sementara jarak antar kluster cukup jauh. Temuan ini memperkuat efektivitas metode *K-Means* dalam menghasilkan kluster yang terpisah dengan jelas. Penemuan ini juga mendukung penelitian Kusuma & Firmansyah (2022), yang menunjukkan keberhasilan *K-Means* dalam mengelompokkan data kesehatan, seperti tekanan darah dan kolesterol. Berdasarkan hasil klusterisasi ini, pihak Puskesmas dapat memanfaatkan informasi untuk merancang intervensi yang lebih sesuai dengan kondisi kesehatan lansia, sehingga pelayanan kesehatan dapat dilakukan lebih tepat sasaran. Ini juga relevan dengan studi Pengetahuan *et al.* (2024), yang menunjukkan pentingnya analisis data untuk meningkatkan efektivitas program kesehatan bagi lansia, seperti pengelolaan penyakit kronis di Posyandu Lansia. Secara keseluruhan, penggunaan *K-Means Clustering* dalam penelitian ini terbukti efektif dalam mengelompokkan data penyakit kronis pada lansia. Hasil pengelompokan ini dapat menjadi dasar bagi pihak Puskesmas untuk merancang kebijakan yang lebih sesuai dengan kebutuhan lansia, serta meningkatkan kualitas pelayanan kesehatan yang diberikan kepada kelompok tersebut.

4. Kesimpulan

Berdasarkan hasil pengujian terhadap pola penyakit kronis pada lansia menggunakan metode *K-Means Clustering*, dapat disimpulkan bahwa tantangan dalam menganalisis pola penyakit lansia, terutama pada data yang besar dan kompleks, dapat diatasi dengan penerapan teknik data mining. Dalam hal ini, algoritma *K-Means Clustering* terbukti efektif dalam melakukan pengelompokan data lansia secara otomatis dan efisien. Penerapan algoritma ini pada data hasil pemantauan kesehatan lansia yang mencakup variabel seperti usia, tekanan darah, dan kadar gula darah memberikan informasi yang berguna, dengan menghasilkan segmentasi lansia berdasarkan kemiripan kondisi kesehatannya. Segmentasi ini dapat digunakan oleh Puskesmas untuk merancang program dan layanan kesehatan yang lebih tepat sasaran. Selain itu, pengujian menunjukkan bahwa jumlah kluster yang paling optimal adalah dua, dengan nilai evaluasi kinerja sebesar 0.579, yang paling mendekati nilai nol dibandingkan pengujian dengan tiga atau empat kluster. Hal ini menunjukkan bahwa dua kluster memiliki perbedaan pola kesehatan yang jelas, yang dapat dijadikan dasar prioritas dalam merencanakan intervensi medis bagi lansia.

5. Ucapan Terima Kasih

Segala puji penulis panjatkan ke hadirat Tuhan Yang Maha Esa atas berkat dan pertolongan-Nya sehingga penelitian ini dapat terselesaikan dengan baik. Penulis menyampaikan penghargaan dan rasa terima kasih kepada seluruh dosen serta staf Program Studi Sistem Informasi STIKOM CKI atas bimbingan, arahan, dan dukungan yang diberikan selama proses penyusunan penelitian ini. Apresiasi yang tulus juga penulis berikan kepada pihak Puskesmas Kelurahan Semper Barat yang telah memberikan izin serta menyediakan data yang diperlukan. Tidak lupa, penulis mengucapkan terima kasih kepada keluarga dan teman-teman yang selalu memberikan doa, dukungan moral, dan semangat. Semoga segala bantuan yang telah diberikan mendapat balasan yang setimpal dari Tuhan Yang Maha Esa.

6. Daftar Pustaka

- Apriyani, P., Dikananda, A. R., & Ali, I. (2023). Penerapan Algoritma K-Means dalam Klusterisasi Kasus Stunting Balita Desa Tegalwangi. *Hello World Jurnal Ilmu Komputer*, 2(1), 20-33. <https://doi.org/10.56211/helloworld.v2i1.230>.
- Ariyanto, D. (2022). Data Mining Menggunakan Algoritma K-Means untuk Klasifikasi Penyakit Infeksi Saluran Pernafasan Akut. *Jurnal Sistim Informasi dan Teknologi*, 13-18. <https://doi.org/10.37034/jsisfotek.v4i1.117>.
- Aulia, S. (2021). Klusterisasi pola penjualan pestisida menggunakan metode K-Means clustering (Studi Kasus di Toko Juanda Tani Kecamatan Hutabayu Raja). *Djtechno: Jurnal Teknologi Informasi*, 1(1), 1-5. <https://doi.org/10.46576/djtechno.v1i1.964>.
- Gustiane, I. T., Martanto, M., & Suprapti, T. (2024). Clustering Hasil Cek Darah Diabetes Lansia Menggunakan Metode K-Means di Posbindu Kp. Lebakjero Desa Ciherang. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 8(2), 2125-2129.
- Gustrianda, R., & Mulyana, D. I. (2022). Penerapan Data Mining Dalam Pemilihan Produk Unggulan dengan Metode Algoritma K-Means Dan K-Medoids. *J. Media Inform. Budidarma*, 6(1), 27.
- Kusuma, P. A., & Firmansyah, A. U. (2022). Deteksi Penyebaran Penyakit Tuberkulosis dengan Algoritma K-Means Clustering Menggunakan Rapid Miner. *Teknol. Inform. dan Komput*, 8(2), 41-54.
- Nahjan, M. R., Heryana, N., & Voutama, A. (2023). Implementasi Rapidminer Dengan Metode Clustering K-Means Untuk Analisa Penjualan Pada Toko Oj Cell. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 7(1), 101-104. <https://doi.org/10.36040/jati.v7i1.6094>.
- Narulita, S., Prihati, P., Oktaga, A. T., & Widyantoro, A. E. (2023). Performansi algoritma clustering K-means untuk penentuan status malnutrisi pada balita. *Jurnal Informasi, Sains Dan Teknologi*, 6(1), 188-202. <https://doi.org/10.55606/isaintek.v6i02.128>.
- Niko, N. S., Rahman, A., Atmaja, D. M. U., & Basri, A. (2023). Klusterisasi Stok Produk Retail Untuk Menentukan Pergerakan Kebutuhan Konsumen Dengan Algoritma K-Means. *Bulletin of Information Technology (BIT)*, 4(3), 306-312. <https://doi.org/10.47065/bit.v4i3.736>.

- Novita, I., Hr, M., Rosita, S., Martunis, M., Sakdah, N., & Yunita, Y. (2024). HUBUNGAN PENGETAHUAN, PERAN KADER DAN DUKUNGAN KELUARGA TERHADAP KEIKUTSERTAAN LANSIA DALAM PROGRAM PENGELOLAAN PENYAKIT KRONIS (PROLANIS) DI POSYANDU LANSIA GAMPONG MULIA. *JOURNAL OF HEALTHCARE TECHNOLOGY AND MEDICINE*, 10(2), 87-96. <https://doi.org/10.33143/jhtm.v10i2.4291>.
- Novitasari, N., Nuris, N. D., & Herdiana, R. (2023). Penerapan algoritma K-Means untuk clustering data jumlah penduduk miskin berdasarkan kota/kabupaten di Jawa Barat menggunakan RapidMiner. *Jurnal Informatika Terpadu*, 9(1), 68–73. <https://doi.org/10.54914/jit.v9i1.660>.
- Sudrajat, W., Cholid, I., & Petrus, J. (2022). Penerapan algoritma K-Means clustering untuk pengelompokan UMKM menggunakan RapidMiner. *Jurnal JUPITER*, 14(1), 27–36.
- Sugianto, C. A., Rahayu, A. H., & Gusman, A. (2020). Algoritma K-Means untuk pengelompokan penyakit pasien pada Puskesmas Cigugur Tengah. *Journal of Information Technology*, 2(2), 39–44. <https://doi.org/10.47292/joint.v2i2.30>.
- Tinendung, I. S., & Zufria, I. (2023). Pengelompokan status stunting pada anak menggunakan metode K-Means clustering. *Jurnal Media Informatika Budidarma*, 7(4), 2014. <https://doi.org/10.30865/mib.v7i4.6908>.