

Penerapan Algoritma K-Means Clustering Analysis Penyakit Stroke Ringan Di Rumah Sakit

¹Muhammad Rickyu Yudhistira Firmansyah, ²Muhammad Yasser Arafat

^{1,2}Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Pamulang,
Tangerang Selatan, Indonesia

¹rickyumuhammad2@gmail.com, ²dosen00680@unpam.ac.id

Info Artikel

Riwayat Artikel:

Received Nov 02, 2025

Revised Nov 20, 2025

Accepted Nov 24, 2025

Abstract – This study implements the K-Means Clustering Algorithm for the analysis of mild stroke disease in hospitals. Stemming from the high incidence of stroke in Indonesia and the limitations of conventional medical recording systems which are manual and subjective, this research aims to design and build an information technology-based decision support system capable of classifying patients objectively, measurably, and efficiently. The method used is applied research with a quantitative and descriptive approach, and system development utilizing the K-Means method. The data analyzed includes specific clinical and risk factors such as age, blood pressure, cholesterol levels, and medical history, obtained through hospital medical records and interviews with medical personnel. The application of the K-Means Clustering algorithm enables the grouping of mild stroke patients into clusters based on the similarity of characteristics, the results of which can identify patient categories with low, moderate, or high-risk levels. The proposed system is designed using the Unified Modeling Language (UML) and Database Design (ERD, LRS), and implemented using supporting applications such as Visual Studio Code, HTML, CSS, MySQL, and Python. The results of Black Box Testing and White Box Testing show that the functional system and internal logic run valid and are in accordance with the expected requirement specifications, thereby capable of increasing effectiveness and efficiency in patient data management and accelerating the process of risk level identification.

Keywords: Data Mining, Decision Support, K-Means Clustering, Mild Stroke, Risk Analysis

Corresponding Author:

Suttichai Premrudeeprechacharn

Email: suttichai@mail.com



This is an open access article under the [CC BY 4.0](#) license.

Abstrak Indonesia – Penelitian ini mengimplementasikan Algoritma K-Means Clustering untuk analisis penyakit stroke ringan di rumah sakit. Berangkat dari tingginya angka kejadian stroke di Indonesia dan keterbatasan sistem pencatatan medis konvensional yang bersifat manual dan subjektif, penelitian ini bertujuan untuk merancang dan membangun sistem pendukung keputusan berbasis teknologi informasi yang mampu mengelompokkan pasien secara objektif, terukur, dan efisien. Metode yang digunakan adalah applied research dengan pendekatan kuantitatif dan deskriptif, serta pengembangan sistem menggunakan metode K-Means. Data yang dianalisis mencakup faktor klinis dan risiko tertentu seperti usia, tekanan darah, kadar kolesterol, dan riwayat penyakit, yang diperoleh melalui rekam medis rumah sakit dan wawancara dengan tenaga medis. Penerapan algoritma K-Means Clustering memungkinkan pengelompokan pasien stroke ringan ke dalam klaster-klaster berdasarkan kemiripan karakteristik, yang hasilnya dapat mengidentifikasi kategori pasien dengan tingkat risiko rendah, sedang, atau tinggi. Sistem yang diusulkan dan dirancang menggunakan Unified Modeling Language (UML) dan Perancangan Basis Data (ERD, LRS), serta diimplementasikan menggunakan aplikasi pendukung seperti Visual Studio Code, HTML, CSS,

MySQL, dan Python. Hasil pengujian Black Box Testing dan White Box Testing menunjukkan bahwa sistem fungsional dan logika internal berjalan valid dan sesuai dengan spesifikasi kebutuhan yang diharapkan, sehingga mampu meningkatkan efektivitas dan efisiensi dalam pengelolaan data pasien serta mempercepat proses identifikasi tingkat risiko.

Kata Kunci: Analisis Resiko, Data Mining, K-Means Clustering, Sistem Pendukung Keputusan, Stroke ringan,

1. PENDAHULUAN

Stroke merupakan salah satu penyakit tidak menular yang memberikan ancaman serius bagi kesehatan masyarakat dunia. Kondisi ini terjadi akibat adanya gangguan aliran darah menuju otak yang mengakibatkan kerusakan fungsi neurologis secara cepat. Menurut laporan dari World Health Organization (WHO), lebih dari 15 juta orang di seluruh dunia mengalami stroke setiap tahunnya, di mana sekitar 5 juta meninggal dan 5 juta lainnya menderita kecacatan permanen. Angka tersebut menjadikan stroke sebagai salah satu penyebab utama kematian dan kecacatan jangka panjang, sehingga berdampak besar terhadap kualitas hidup pasien dan beban ekonomi keluarga serta negara. Tidak hanya itu, stroke juga memberi pengaruh sosial dan psikologis yang cukup besar terhadap pasien maupun lingkungan sekitarnya karena adanya keterbatasan aktivitas yang muncul pasca kejadian stroke.

Sementara itu di Indonesia sendiri, prevalensi stroke menunjukkan tren peningkatan yang signifikan dalam beberapa tahun terakhir. Faktor penyebabnya antara lain perubahan gaya hidup masyarakat yang cenderung kurang aktif, meningkatnya prevalensi hipertensi, obesitas, hipercolesterolemia, dan diabetes mellitus sebagai faktor risiko utama stroke. Data dari Riset Kesehatan Dasar (Riskesdas) juga memperlihatkan bahwa stroke merupakan penyakit penyebab kematian nomor satu di Indonesia dalam kategori penyakit tidak menular. Situasi tersebut menjadi tantangan besar bagi sistem pelayanan kesehatan nasional dalam menyediakan layanan pencegahan, diagnosis, serta penanganan stroke yang cepat dan tepat.

Salah satu kondisi yang perlu mendapatkan perhatian serius adalah stroke ringan (minor stroke). Meskipun gejala neurologis yang muncul bersifat sementara dan dinilai tidak terlalu parah, stroke ringan memiliki potensi tinggi berkembang menjadi stroke berat dalam rentang waktu yang singkat apabila tidak dilakukan tindakan medis dan pemantauan yang menyeluruh. Oleh karena itu, proses identifikasi tingkat risiko pada pasien stroke ringan sangat penting untuk dilakukan sedini mungkin. Identifikasi ini bertujuan untuk menentukan kategori pasien yang membutuhkan penanganan intensif sebagai upaya pencegahan kemungkinan terjadinya stroke ulang dengan dampak yang lebih serius.

Namun, dalam praktiknya, sebagian besar rumah sakit di Indonesia masih mengandalkan rekam medis manual atau sistem pencatatan yang belum sepenuhnya terintegrasi dengan kemampuan analitik berbasis teknologi. Kondisi ini membuat proses penilaian risiko masih sangat bergantung pada pengalaman subjektif tenaga medis sehingga memiliki potensi menimbulkan ketidakstabilan dalam pengambilan keputusan klinis. Hal ini diperparah dengan keterbatasan waktu dan sumber daya tenaga kesehatan dalam melakukan penilaian risiko secara menyeluruh terhadap seluruh pasien stroke.

Perkembangan teknologi informasi dalam bidang kesehatan membuka peluang besar untuk menghadirkan sistem analitik cerdas yang mampu membantu tenaga medis dalam melakukan pemetaan risiko pasien secara lebih objektif, cepat, dan akurat. Salah satu pendekatan analisis data yang banyak digunakan adalah data mining, khususnya metode clustering, yang berfungsi untuk mengelompokkan data berdasarkan tingkat kesamaan karakteristik. Dalam konteks ini, algoritma K-Means Clustering dipilih karena memiliki keunggulan dalam hal efisiensi komputasi dan kemampuannya dalam menghasilkan pengelompokan yang jelas berdasarkan jarak kedekatan antar data.

Dengan memanfaatkan data rekam medis seperti usia, tekanan darah, kadar kolesterol, serta riwayat penyakit penyerta, algoritma K-Means Clustering dapat mengelompokkan pasien stroke ringan ke dalam kategori risiko rendah, sedang, dan tinggi. Hasil pengelompokan tersebut berpotensi besar dijadikan dasar bagi Sistem Pendukung Keputusan (Decision Support System/DSS) di rumah sakit, sehingga dapat digunakan untuk:

2. PENELITIAN YANG TERKAIT

Beberapa jumlah penelitian sebelumnya telah memanfaatkan teknik data mining, khususnya algoritma K-Means Clustering, dalam mendukung analisis data kesehatan dan data rekam medis pasien. Pemanfaatan metode clustering tersebut dilakukan dengan tujuan memberikan insight tambahan terkait karakteristik pasien, tingkat risiko penyakit, serta pengelompokan kondisi klinis yang dapat digunakan sebagai dasar pengambilan keputusan medis. Meskipun demikian, setiap penelitian memiliki fokus dan ruang lingkup yang berbeda sehingga diperlukan analisis komparatif untuk memahami posisi serta kontribusi penelitian ini dalam konteks kajian yang sudah ada.

Penelitian oleh Payon (2024) menunjukkan bahwa algoritma K-Means efektif dalam segmentasi citra CT-Scan untuk mendeteksi area kerusakan otak pada pasien stroke. Dalam penelitiannya, Payon melakukan perbandingan antara metode thresholding dan K-Means dengan fokus pada citra medis (image processing) untuk identifikasi jaringan otak yang terdampak stroke. Walaupun hasilnya mendukung pemanfaatan K-Means dalam bidang kesehatan, penelitian tersebut tidak menargetkan aspek karakteristik klinis pasien stroke ringan, sehingga implementasinya tidak secara langsung berperan dalam proses prediksi tingkat risiko pasien berdasarkan rekam medis.

Berbeda dengan pendekatan berbasis citra, Damayanti (2022) meneliti clustering faktor risiko dan gejala stroke guna menentukan tingkat keparahan kondisi pasien. Penelitian ini telah mengarah pada pemetaan tingkat risiko, namun belum mengintegrasikan hasil clustering ke dalam sistem pendukung keputusan bagi tenaga medis. Analisis pada penelitian tersebut masih bersifat deskriptif sehingga rekomendasi klinis yang dihasilkan masih belum bersifat implementatif terhadap alur pelayanan kesehatan rumah sakit.

Penelitian berfokus lebih luas dilakukan oleh Widianto dan Kuswiadji (2024) yang menerapkan K-Means Clustering untuk mengelompokkan pasien berdasarkan tingkat keparahan penyakit. Walaupun memberikan gambaran umum mengenai potensi metode clustering pada domain medis, kasus yang digunakan tidak spesifik untuk stroke ringan, sehingga belum memberikan pemetaan tingkat risiko yang fokus pada kondisi pasca-stroke ringan yang memiliki kecenderungan berkembang menjadi stroke berat.

Selain itu, Dilawati (2024) mengkaji pemanfaatan clustering dalam pengelompokan data rekam medis guna mendukung manajemen rumah sakit. Penelitian tersebut menunjukkan bahwa hasil clustering dapat meningkatkan efektivitas pengelolaan data dan penentuan strategi layanan. Akan tetapi, fokus kajian lebih menekankan pada pengelolaan administratif, bukan pada pemanfaatan data klinis untuk deteksi dini penyakit tertentu, termasuk stroke ringan.

Pada tataran riset internasional, Shin et al. (2023) mengembangkan model unsupervised learning berbasis clustering untuk memprediksi pola pemulihan pasien stroke. Penelitian tersebut menunjukkan kontribusi signifikan dalam memahami perkembangan klinis jangka panjang. Namun, fokus penelitian lebih ditekankan pada tahap rehabilitasi pasca-stroke, bukan pada proses awal identifikasi tingkat risiko stroke pada saat pertama kali terdiagnosis sebagai stroke ringan.

Selain penelitian-penelitian yang secara langsung berhubungan dengan stroke, terdapat juga kajian metodologis yang memanfaatkan K-Means Clustering untuk pendekripsi dan pengelompokan penyakit tidak menular lainnya seperti diabetes, penyakit jantung koroner, dan gagal ginjal. Pendekatan ini membuktikan bahwa pemrosesan data rekam medis berbasis clustering dapat memberikan dukungan klinis berbasis data, sehingga praktik serupa memiliki peluang besar untuk ditransformasikan ke konteks penyakit stroke ringan. Namun, hingga saat ini, masih teridentifikasi keterbatasan dalam pengembangan model yang spesifik menggunakan parameter klinis stroke ringan dalam konteks rumah sakit, terutama terkait pemanfaatan hasil clustering untuk deteksi dini risiko stroke berulang.

Dengan demikian, terdapat research gap yang menjadi dasar urgensi penelitian ini, yaitu belum adanya model clustering berbasis karakteristik klinis pasien stroke ringan yang bertujuan memberikan dukungan aksi klinis secara cepat kepada tenaga medis dalam deteksi dini kemungkinan peningkatan risiko stroke lebih berat.

Penelitian ini hadir dengan kontribusi baru berupa pengembangan model K-Means Clustering yang diterapkan pada data rekam medis pasien stroke ringan sebagai komponen penilaian risiko secara objektif dan terukur, sehingga hasil pengelompokan dapat dimanfaatkan dalam sistem pendukung keputusan (Decision Support System). Dengan adanya inovasi ini, diharapkan rumah sakit akan mampu meningkatkan

efektivitas pelayanan, mengoptimalkan penanganan pasien berisiko tinggi, serta memperkuat upaya pencegahan kejadian stroke berulang yang lebih fatal.

3. METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif deskriptif, di mana seluruh analisis dilakukan terhadap data numerik yang diperoleh dari rekam medis pasien stroke ringan. Pendekatan kuantitatif digunakan karena penelitian ini bertujuan untuk melakukan pengukuran karakteristik klinis pasien secara objektif dan menghasilkan informasi berupa pola pengelompokan tingkat risiko stroke berdasarkan data yang tersedia. Sementara itu, sifat deskriptif dalam penelitian ini digunakan untuk menjelaskan hasil pengelompokan yang diperoleh dari proses data mining sehingga dapat memberikan gambaran mengenai kondisi klinis pasien pada masing-masing kelompok risiko yang terbentuk.

Metode utama yang diterapkan dalam penelitian ini adalah data mining, khususnya pada teknik clustering, yang bertujuan untuk mengelompokkan data pasien stroke ringan berdasarkan tingkat kesamaan karakteristik klinis. Tahapan dalam penelitian ini berpedoman pada alur Knowledge Discovery in Database (KDD), karena kerangka kerja tersebut mencakup seluruh proses pengolahan data mulai dari bentuk awal yang belum terstruktur hingga menghasilkan pengetahuan baru yang dapat diinterpretasikan dan dimanfaatkan untuk mendukung keputusan klinis. Secara umum, tahapan KDD yang dilakukan dalam penelitian ini meliputi: pengumpulan data, preprocessing dan pembersihan data, transformasi data, penerapan algoritma K-Means, evaluasi hasil clustering, serta interpretasi klinis terhadap hasil yang diperoleh. Setiap tahapan saling berhubungan dan sangat menentukan kualitas akhir dari hasil pengelompokan.

Tahap awal penelitian dimulai dengan pengumpulan data pasien stroke ringan, dimana data yang dikumpulkan merupakan data sekunder yang bersumber dari rekam medis. Data yang diperoleh berupa informasi klinis yang berhubungan langsung dengan faktor risiko stroke, seperti usia, tekanan darah, kadar kolesterol, serta riwayat penyakit penyerta seperti hipertensi atau diabetes. Atribut tersebut dipilih karena telah banyak dikaitkan dalam literatur kesehatan sebagai indikator penting dalam menentukan tingkat risiko stroke berulang. Pada tahapan ini, proses perizinan dan penerapan etika penelitian menjadi prioritas agar kerahasiaan identitas serta hak pasien tetap terlindungi sesuai dengan standar rumah sakit.

Setelah data terkumpul, langkah berikutnya adalah preprocessing, yang bertujuan untuk memastikan kualitas data sebelum dilakukan analisis lebih lanjut. Preprocessing meliputi pengecekan konsistensi format data, penanganan nilai hilang (missing value), identifikasi data ekstrem (outlier), serta penghapusan data duplikat apabila ditemukan. Tahap ini sangat penting karena kualitas data yang buruk dapat menyebabkan algoritma clustering menghasilkan pembentukan kelompok yang keliru dan tidak dapat diinterpretasikan secara klinis. Selain itu, dilakukan juga pemilihan variabel yang benar-benar relevan dengan tujuan penelitian agar analisis lebih fokus dan akurat.

Selanjutnya, dilakukan proses transformasi data, di mana seluruh atribut yang memiliki skala pengukuran berbeda dinormalisasi sehingga memiliki rentang nilai yang seragam. Normalisasi dilakukan karena algoritma K-Means sangat bergantung pada perhitungan jarak Euclidean, sehingga atribut dengan skala lebih besar dapat mendominasi perhitungan jarak jika tidak dilakukan penyeragaman nilai. Transformasi ini membantu algoritma menghasilkan pengelompokan yang lebih seimbang serta mencerminkan kontribusi masing-masing variabel klinis terhadap pembentukan klaster.

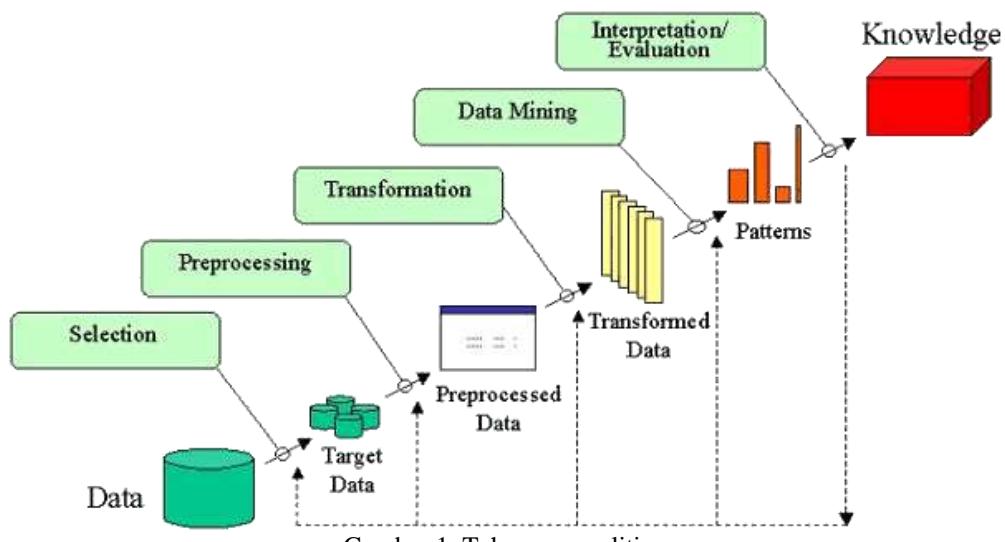
Tahapan inti dari penelitian ini adalah penerapan algoritma K-Means Clustering sebagai metode untuk membentuk kelompok pasien stroke ringan berdasarkan kedekatan karakteristik klinis. Algoritma K-Means bekerja dengan menentukan titik pusat kelompok (centroid) secara acak pada awal proses, kemudian melakukan pembaruan posisi centroid secara berulang hingga tidak ada lagi perubahan signifikan pada posisi centroid. Pada penelitian ini, jumlah klaster ditentukan dengan mempertimbangkan tujuan penelitian, yaitu untuk mengelompokkan pasien ke dalam tingkatan risiko rendah, sedang, dan tinggi, sehingga jumlah klaster (k) yang digunakan adalah tiga. Setiap iterasi algoritma bertujuan untuk meminimalkan jarak antar data dalam satu klaster dan memaksimalkan perbedaan antar klaster yang satu dengan yang lain.

Setelah proses clustering selesai, tahap selanjutnya adalah evaluasi hasil clustering untuk memastikan bahwa pengelompokan yang terbentuk telah sesuai dan layak untuk diinterpretasikan secara

klinis. Evaluasi dilakukan dengan menganalisis penyebaran data pada setiap klaster serta karakteristik rata-rata nilai atribut yang terdapat pada centroid masing-masing klaster. Selain itu, hasil clustering kemudian dibandingkan dengan pengetahuan klinis yang telah ada untuk memastikan bahwa kelompok risiko yang terbentuk memiliki dasar yang logis dalam dunia kedokteran.

Tahap akhir dari penelitian ini adalah interpretasi hasil, yaitu menjelaskan karakteristik klinis pada masing-masing klaster untuk menentukan tingkat risiko stroke ringan yang berpotensi berkembang menjadi stroke berat. Interpretasi ini bertujuan untuk menghasilkan informasi yang dapat digunakan oleh tenaga medis sebagai dasar pendukung dalam proses pengambilan keputusan di rumah sakit. Dengan hasil pengelompokan yang diperoleh, tenaga medis dapat melakukan monitoring lebih optimal serta menentukan prioritas penanganan bagi pasien yang tergolong ke dalam kelompok risiko sedang dan tinggi.

Dengan demikian, metodologi penelitian yang digunakan dalam penelitian ini tidak hanya berfungsi untuk menganalisis data secara komputasional, namun juga diarahkan agar hasilnya dapat diimplementasikan dalam konteks klinis dan memberikan manfaat nyata dalam peningkatan kualitas pelayanan kesehatan, khususnya pada upaya pencegahan terjadinya stroke berulang pada pasien stroke ringan.



Gambar 1. Tahapan penelitian

Gambar di atas menggambarkan proses Knowledge Discovery in Databases (KDD), yaitu tahapan yang dilakukan untuk memperoleh pengetahuan (knowledge) dari data mentah melalui beberapa proses data mining.

Proses dimulai dengan pemilihan (selection) data yang relevan, dilanjutkan dengan pra-pemrosesan (preprocessing) untuk membersihkan data dari noise, duplikasi, atau data yang tidak konsisten. Setelah itu, data ditransformasikan (transformation) menjadi format yang sesuai untuk dianalisis. Data yang sudah siap kemudian dianalisis menggunakan teknik data mining untuk menemukan pola-pola yang tersembunyi (data mining). Pola-pola hasil analisis selanjutnya dievaluasi dan diinterpretasikan (interpretation/evaluation) agar benar-benar menghasilkan pengetahuan yang bermanfaat (knowledge). Seluruh rangkaian proses ini bertujuan agar data mentah dapat diubah menjadi pengetahuan yang dapat digunakan secara nyata untuk mendukung pengambilan keputusan atau analisis lanjutan.

Tabel 1. Informasi dataset

No.	Atribut	Jenis Data	Skala	Keterangan
1	Usia	Numerik	Rasio	Tahun
2	Tekanan Darah Sistolik	Numerik	Rasio	mmHg
3	Tekanan Darah Diastolik	Numerik	Rasio	mmHg

4	Kadar Kolesterol	Numerik	Rasio	mg/dL
5	Riwayat Hipertensi	Kategorikal	Nominal	Ya/Tidak
6	Riwayat Diabetes	Kategorikal	Nominal	Ya/Tidak
7	Kebiasaan Merokok	Kategorikal	Nominal	Ya/Tidak
8	Aktivitas Fisik	Kategorikal	Ordinal	Rendah/Sedang/Tinggi

4. HASIL DAN PEMBAHASAN

Pada bagian penelitian ini, algoritma K-Means Clustering dimanfaatkan sebagai pendekatan analitik untuk mengelompokkan data pasien stroke ringan berdasarkan kesamaan karakteristik klinis tertentu. Adapun variabel yang digunakan dalam proses pengelompokan meliputi usia, tekanan darah, kadar kolesterol, serta riwayat penyakit penyerta seperti hipertensi atau diabetes. Seluruh data yang digunakan terlebih dahulu melalui serangkaian tahapan preprocessing, termasuk pembersihan data, normalisasi, dan pemilihan atribut yang dianggap relevan dalam penilaian risiko stroke tinggi.

Setelah proses preprocessing selesai, algoritma K-Means diterapkan untuk membagi pasien ke dalam beberapa kelompok (*cluster*) yang merepresentasikan tingkat risiko yang bervariasi. Hasil clustering menunjukkan bahwa setiap kelompok memiliki karakteristik klinis unik yang dapat dijadikan indikator dalam menentukan tingkat kewaspadaan dan pola penanganan medis. Sebagai contoh, ditemukan sebuah cluster dengan kondisi pasien yang memiliki tekanan darah serta kadar kolesterol masih berada pada batas normal sehingga dapat dikategorikan sebagai kelompok risiko lebih rendah. Sebaliknya, terdapat pula cluster lain yang didominasi pasien dengan kondisi hipertensi atau hiperkolesterolemia, sehingga kelompok tersebut berpotensi masuk pada kategori risiko sedang hingga tinggi karena memiliki faktor pemicu stroke yang lebih serius.

Temuan tersebut memperkuat teori bahwa stroke ringan tidak dapat dianggap sebagai kondisi yang sepenuhnya aman. Jika tidak dilakukan pemantauan serta penanganan medis yang optimal, kondisi ini dapat berkembang menjadi stroke berat yang berdampak lebih fatal terhadap kesehatan pasien. Dengan adanya hasil pengelompokan ini, tenaga medis diharapkan dapat lebih mudah mengenali pasien yang berada dalam kelompok risiko tinggi dan membutuhkan intervensi lebih cepat, terarah, dan intensif..

Hasil penelitian ini juga sejalan dengan berbagai studi sebelumnya yang menegaskan bahwa pemanfaatan teknik data mining, khususnya metode clustering, mampu menyediakan informasi tambahan yang mendukung proses pengambilan keputusan klinis. Dalam konteks operasional rumah sakit, hasil K-Means Clustering dapat digunakan untuk membantu pemetaan risiko sejak dini, pengelompokan prioritas pemantauan, pengaturan strategi intervensi, hingga peningkatan kualitas layanan berbasis data rekam medis yang tersedia. Dengan demikian, pendekatan ini dapat menjadi alat bantu pendukung keputusan (decision support) bagi pihak medis dalam melakukan perencanaan manajemen pasien.

Secara keseluruhan, meskipun penelitian ini masih memiliki keterbatasan seperti ukuran data yang belum terlalu besar, variabel klinis yang masih bisa diperluas, serta belum dilakukannya validasi hasil clustering secara langsung dengan tenaga kesehatan yang berpengalaman, namun hasil awal menunjukkan bahwa penerapan algoritma K-Means memiliki potensi signifikan dalam membantu proses analisis dan evaluasi kondisi pasien stroke ringan di lingkungan rumah sakit. Pengembangan penelitian lanjutan dengan peningkatan kualitas dan kuantitas data diharapkan dapat semakin memaksimalkan kontribusi metode ini dalam dunia kesehatan, khususnya dalam pengelolaan dan pencegahan risiko stroke yang lebih baik dan berbasis data.

5. KESIMPULAN

Kesimpulan dari penelitian ini menunjukkan bahwa penerapan teknik data mining, khususnya algoritma K-Means Clustering, pada data klinis pasien stroke ringan mampu memberikan gambaran pengelompokan risiko yang lebih terstruktur dan sistematis. Dengan memanfaatkan variabel-variabel klinis seperti tekanan darah, kadar kolesterol, indeks massa tubuh (BMI), usia, riwayat hipertensi, serta kebiasaan gaya hidup yang memengaruhi kesehatan pasien, sistem clustering berhasil memisahkan pasien ke dalam

beberapa kelompok risiko yang berbeda. Pembagian kelompok ini memberikan informasi tambahan bagi tenaga medis untuk memahami karakteristik pasien secara lebih spesifik berdasarkan profil kesehatannya.

Secara umum, hasil penelitian ini memperlihatkan bahwa model clustering yang diterapkan mampu menghasilkan pengelompokan yang sesuai dengan penilaian klinis terhadap faktor risiko stroke. Pasien dengan tekanan darah dan kolesterol tinggi cenderung terkелompok dalam kategori risiko lebih berat, sementara pasien yang memiliki nilai variabel lebih rendah berada pada kelompok risiko yang lebih ringan. Temuan ini menegaskan bahwa K-Means Clustering memiliki potensi besar sebagai alat bantu analisis dalam mendukung upaya deteksi dini serta langkah preventif terhadap progresivitas penyakit stroke. Dengan pengelompokan risiko yang akurat, intervensi medis dapat diberikan lebih cepat dan tepat sasaran untuk mencegah berkembangnya stroke ringan menjadi stroke berat yang berpotensi meningkatkan angka kecacatan bahkan kematian.

Meskipun memberikan hasil yang menjanjikan, penelitian ini masih memiliki keterbatasan yang perlu diperhatikan dalam pengembangan selanjutnya. Jumlah data yang digunakan masih relatif terbatas, sehingga mungkin belum sepenuhnya merepresentasikan karakteristik populasi pasien stroke ringan secara lebih luas. Selain itu, parameter klinis yang dipakai sebagai dasar pengelompokan masih dapat ditingkatkan, misalnya dengan menambahkan variabel seperti kadar gula darah, tingkat aktivitas fisik, riwayat penyakit jantung, dan indikator biomarker lain yang relevan dengan kejadian stroke. Penggunaan metode evaluasi clustering yang lebih komprehensif juga dapat meningkatkan kualitas hasil analisis.

Ke depannya, penelitian dapat diarahkan pada integrasi model clustering ke dalam Sistem Pendukung Keputusan (*Decision Support System*) yang dapat digunakan secara langsung oleh tenaga medis di rumah sakit. Dengan demikian, proses penilaian risiko pasien dapat dilakukan secara otomatis dan real-time, sehingga membantu tenaga medis dalam proses triase, perencanaan perawatan, serta rekomendasi intervensi klinis yang lebih personal. Selain itu, pengembangan antarmuka visualisasi data berbasis dashboard interaktif juga dapat meningkatkan kemudahan interpretasi hasil clustering, baik oleh dokter, perawat, maupun pengambil kebijakan rumah sakit.

Berdasarkan keseluruhan hasil dan potensi pengembangannya, dapat disimpulkan bahwa penerapan algoritma K-Means Clustering pada data klinis pasien stroke ringan merupakan pendekatan yang bernilai inovatif dan aplikatif dalam mendukung upaya pencegahan, pemantauan, dan pengelolaan penyakit stroke secara lebih efektif. Dengan peningkatan kualitas dataset dan pengembangan sistem yang lebih terintegrasi, metode ini berpeluang besar memberikan kontribusi nyata dalam peningkatan kualitas layanan kesehatan khususnya pada bidang neurologi dan manajemen risiko penyakit stroke.

DAFTAR PUSTAKA

- [1] Adiwijaya, F. F., Amaruloh, D. S., & Mulya, A. R. (2021). Sistem Registrasi Surat Perintah Tugas (Spt) Di Dinas Pekerjaan Umum, Penataan Ruang Dan Pertanahan Provinsi Kepulauan Riau. *Komputa : Jurnal Ilmiah Komputer Dan Informatika*, 10(2), 70–77. <https://doi.org/10.34010/komputa.v10i2.6806>
- [2] Afni, N., Salim, A., Maulana, Y. I., Nugraha, A., & Komarudin, R. (2022). Information System Program Design Of Panti Asuhan Perancangan Program Sistem Informasi Panti Asuhan. *Journal of Information System, Informatics and Computing*, 6(2), 486–496. <https://doi.org/10.5236/jjisicom.v6i2.962>
- [3] Alasi, G., & Putri, R. A. (2025). Penerapan K-Means Clustering untuk Pengelompokan Pasien Rumah Sakit berdasarkan Tingkat Keparahan Penyakit. 07(03), 475–486.
- [4] Ananta Maulida, R., Anisa Aulia, S., Dzulfahmi Yulianto, S., Clara Efendi, S., & Fansyuri, M. (2025). *Pengambilan Keputusan Medis Berbasis Algoritma K-Nearest Neighbor (KNN) Dalam Klasifikasi Pasien Stroke*. 1(1), 118–121.
- [5] Ananda, P. L., Wardhani, N. I., & Nurhayati, E. (2024). Pemanfaatan Bahasa Pemograman Web Untuk Meningkatkan Pemahaman Teknologi Informasi : Studi Kasus Penggunaan Visual Studio Code. *Jurnal Multidisiplin Saintek*, 5(9), 1–11. <https://ejournal.warunayama.org/kohesi%0AKohesi>
- [6] Choi, J., Park, H., Chie, E. K., Choi, S. W., Lee, H. Y., Yoo, S., Kim, B. J., & Ryu, B. (2023). Current Status and Key Issues of Data Management in Tertiary Hospitals: A Case Study of Seoul National University Hospital. *Healthcare Informatics Research*, 29(3), 209–217. <https://doi.org/10.4258/hir.2023.29.3.209>
- [7] Feigin, V. L., Brainin, M., Norrving, B., Martins, S. O., Pandian, J., Lindsay, P., F Grupper, M., & Rautalin, I. (2025). World Stroke Organization: Global Stroke Fact Sheet 2025. *International Journal of Stroke*, 20(2), 132–144. <https://doi.org/10.1177/17474930241308142>

- [8] Nengsih, Y. G., Samosir, K., & Satria, B. (2025). Segmentasi Pasien Rumah Sakit Berdasarkan Pola Kunjungan Menggunakan Algoritma K-Means Clustering untuk Optimasi Layanan Medis. *Jurnal Pendidikan Sains Dan Komputer*, 5(01), 59–69. <https://doi.org/10.47709/jpsk.v5i01.5492>
- [9] Sandffreni, S., Ulum, M. B., & Azizah, A. H. (2021). Analisis Perancangan Sistem Informasi Pusat Studi Pada Fakultas Ilmu Komputer Universitas Esa Unggul. *Sebatik*, 25(2), 345–356. <https://doi.org/10.46984/sebatik.v25i2.1587>
- [10] Setiawan, K. E., & Kurniawan, A. (2023). Pengelompokan Rumah Sakit di Jakarta Menggunakan Model DBSCAN, Gaussian Mixture, dan Hierarchical Clustering. *Jurnal Informatika Terpadu*, 9(2), 149–156. <https://doi.org/10.54914/jit.v9i2.995>
- [11] Sie, J. B. L., Izmy Alwiah Musdar, & Syamsul Bahri. (2022). Pengujian White Box Testing Terhadap Website Room Menggunakan Teknik Basis Path. *KHARISMA Tech*, 17(2), 45–57. <https://doi.org/10.55645/kharismatech.v17i2.235>
- [12] Purba, W., Sembiring, G. A., Saputra, A., Turnip, T., Jua, B., & Manihuruk, I. (2023). Penerapan Data Mining Untuk Pengolahan Data Rekam Medis Menggunakan Metode K-Means Clustering Pada Rumah Sakit Royal Prima Medan. *Jurnal TEKINKOM*, 6(1), 158–168. <https://doi.org/10.37600/tekinkom.v6i1.857>
- [13] Raihan, H., & Voutama, A. (2023). Black box testing on college database applications with equivalence partition techniques [Pengujian black box pada aplikasi database perguruan tinggi dengan teknik equivalence partition]. *Antivirus : Jurnal Ilmiah Teknik Informatika*, 17(1), 1–18.
- [14] Raihani, K. (2025). Pembuatan Website Penjualan Toko Aksesoris Dengan Menggunakan PHP Dan MySQL. *Jurnal Minfo Polgan*, 13(2), 2299–2308. <https://doi.org/10.33395/jmp.v13i2.14445>
- [15] Puan Maharani. (2025). Pengembangan Website PT. Rantangin Digital Indonesia Menggunakan Framework Next Js dan Tailwind CSS. *Repeater : Publikasi Teknik Informatika Dan Jaringan*, 3(1), 129–137. <https://doi.org/10.62951/repeater.v3i1.355>