E-ISSN 3032-601X & P-ISSN 3032-7105
Vol. 3, No. 1, 2026

MISTER

Journal of Multidisciplinary Inquiry in Science, Technology and Educational Research

Jurnal Penelitian Multidisiplin dalam Ilmu Pengetahuan, Teknologi dan Pendidikan

UNIVERSITAS SERAMBI MEKKAH
KOTA BANDA ACEH
mister@serambimekkah.ac.id

Journal of Multidisciplinary Inquiry in Science Technology and Educational Research

Journal of MISTER

Vol. 3, No. 1, 2026

Pages: 1025-1032

Analisis Pola Risiko Stroke pada Data Pasien Menggunakan Model K-Means Clustering Berbasis RapidMiner

Arjuna Septian Habibullah, Hasbi Firmansyah, Ali Sofyan, Wahyu Asriyani

Universitas Pancasakti Tegal

Article in Journal of MISTER

Available at	:	https://jurnal.serambimekkah.ac.id/index.php/mister/index
DOI	:	https://doi.org/10.32672/mister.v3i1.4026

How to Cite this Article

APA	:	Habibullah, A. S., Firmansyah, H., Sofyan, A. ., & Asriyani, W. (2025). Analisis Pola Risiko Stroke pada Data Pasien Menggunakan Model K-Means Clustering Berbasis RapidMiner. Journal of Multidisciplinary Inquiry in Science, Technology and Educational Research, 3(1), 1025 – 1032. Retrieved from https://jurnal.serambimekkah.ac.id/index.php/mister/article/view/4026
Others Visit	:	https://jurnal.serambimekkah.ac.id/index.php/mister/index

ISSN 3032-7105 ISSN 3032-601X

Journal of Multidisciplinary Inquiry in Science, Technology and Educational Research
UNIVERSITAS SERAMBI MEKKAH
https://jurnal.serambimekkah.ac.id/index.php/mister

e-ISSN3032-601X & p-ISSN3032-7105
Vol. 3 No. 1, Tahun 2026
Doi: 10.32672/mister.v3i1.4026
Hal. 1025-1032

Analisis Pola Risiko Stroke pada Data Pasien Menggunakan Model K-Means Clustering Berbasis RapidMiner

Arjuna Septian Habibullah[1], Hasbi Firmansyah[2], Ali Sofyan[3], Wahyu Asriyani[4]
Program Studi Informatika, Fakultas Teknik dan Ilmu Komputer, Universitas Pancasakti Tegal, Tegal, Indonesia[1,2,3]
Program Studi Pendidikan Bahasa dan Sastra Indonesia, Fakultas Keguruan dan Ilmu Pendidikan, Universitas Pancasakti Tegal, Tegal, Indonesia[4]

*Email: arjunaseptian4@mail.com, hasbifirmansyah@upstegal.ac.id, alisofyan.bbs@gmail.com, asriyani1409@gmail.com.

Diterima: 10-12-2025 | Disetujui: 20-12-2025 | Diterbitkan: 22-12-2025

Abstract

This study aims to identify patterns of stroke risk among patients by applying the K-Means clustering method using the RapidMiner platform. The stroke dataset was loaded from an Excel file and prepared through a preprocessing stage, including normalization of numerical attributes and conversion of categorical attributes into numerical form to ensure accurate distance calculations in the K-Means algorithm. The K-Means algorithm was then applied to form several clusters based on similarities in patients' clinical and demographic data. The quality of the clustering results was evaluated using the ***Davies–Bouldin Index*** *to measure cluster separation and intra-cluster similarity. The results indicate that the K-Means method is capable of producing a more structured representation of stroke risk patterns and can support data-driven decision-making for stroke prevention and management.*

Keywords*: K-Means Clustering; Stroke; RapidMiner; Data Mining; Davies–Bouldin Index.*

Abstrak

Penelitian ini bertujuan untuk mengidentifikasi pola risiko stroke pada pasien menggunakan metode klastering K-Means dengan bantuan platform RapidMiner. Dataset stroke dimuat dari file Excel dan dipersiapkan melalui tahap preprocessing yang meliputi normalisasi atribut numerik serta konversi atribut kategorikal ke bentuk numerik agar perhitungan jarak K-Means dapat dilakukan secara akurat. Selanjutnya, algoritma K-Means diterapkan untuk membentuk beberapa cluster berdasarkan kesamaan data klinis dan demografis pasien. Kualitas hasil klastering dievaluasi menggunakan **Davies–Bouldin Index** untuk menilai tingkat pemisahan antar cluster dan keseragaman data dalam satu cluster. Hasil penelitian menunjukkan bahwa metode K-Means mampu menghasilkan pola risiko stroke yang lebih terstruktur dan dapat mendukung pengambilan keputusan pencegahan serta penanganan stroke berbasis data.

Katakunci: Stroke; K-Means; Klastering; RapidMiner; Davies–Bouldin Index; Data Mining.

PENDAHULUAN

Stroke merupakan salah satu penyebab utama kematian dan kecacatan secara global yang menuntut strategi deteksi dini berbasis pola risiko pasien agar penanganannya lebih efektif dan terarah (Jiang et al., n.d.). Kondisi ini menjadi tantangan serius bagi sistem kesehatan karena angka kejadian stroke terus meningkat seiring bertambahnya usia populasi dan perubahan gaya hidup. Di banyak negara, stroke tidak hanya menimbulkan beban biaya pengobatan yang besar, tetapi juga berdampak pada kualitas hidup pasien serta keluarga yang merawat.

Identifikasi pola risiko stroke memerlukan analisis multidimensi, karena kondisi tersebut dipengaruhi oleh faktor demografis, klinis, dan gaya hidup, sehingga metode analisis data berbasis machine learning semakin banyak digunakan dalam penelitian medis (Care et al., 2024). Faktor-faktor seperti usia, jenis kelamin, tekanan darah, kadar gula darah, hingga kebiasaan merokok memiliki keterkaitan yang beragam terhadap risiko terjadinya stroke, sehingga analisis konvensional seringkali tidak cukup untuk menangkap pola kompleks tersebut. Machine learning memungkinkan eksplorasi hubungan tersembunyi antara variabel dan dapat menemukan pola yang sulit diidentifikasi melalui pendekatan statistik tradisional. Salah satu pendekatan data mining yang umum digunakan adalah teknik *unsupervised learning* seperti clustering, yang memungkinkan peneliti menemukan struktur laten dan pengelompokan alami pada data tanpa memerlukan label hasil diagnosis (Kim et al., 2022). Pendekatan ini cocok untuk dataset kesehatan yang sering kali tidak memiliki label klasifikasi, karena diagnosis medis tidak selalu terdokumentasi secara lengkap dalam sistem data rumah sakit. Clustering juga membantu memetakan sub-kelompok pasien dalam populasi sehingga intervensi medis dapat dilakukan berdasarkan kelompok risiko tertentu.

Algoritma K-Means dikenal sebagai metode clustering yang sederhana, cepat, dan mampu memberikan centroid yang mudah ditafsirkan untuk mendukung pengambilan keputusan klinis (Ceskoutsé et al., 2025). Selain berfungsi untuk membagi data ke dalam beberapa kelompok homogen, K-Means juga memberikan representasi matematis dari pusat kelompok yang dapat dianalisis oleh tenaga medis untuk menilai karakteristik pasien. Dengan efisiensi perhitungan, algoritma ini sangat cocok digunakan pada dataset menengah, termasuk data kesehatan berbasis Excel yang banyak digunakan dalam penelitian akademik.

Studi terbaru menunjukkan bahwa penerapan K-Means dapat mengidentifikasi fenotipe kelompok pasien stroke berdasarkan kemiripan karakteristik, sehingga dapat mendukung strategi intervensi klinis yang lebih personal (Mao, 2025). Hal ini penting karena penanganan stroke bukan hanya tentang mengobati gejala akut, tetapi juga mengelola faktor risiko jangka panjang. Dengan mengetahui kelompok pasien yang memiliki kemungkinan risiko tinggi, tenaga medis dapat memberikan edukasi pencegahan dan perlakuan lebih intensif.

Selain itu, penggunaan model clustering terbukti efektif dalam memetakan pola multimorbiditas atau faktor risiko yang saling terkait, terutama pada pasien stroke dengan kondisi komorbid seperti hipertensi, diabetes mellitus, dan penyakit jantung (Guo et al., 2025). Multimorbiditas sangat umum terjadi pada pasien stroke dan memengaruhi jalur pengobatan serta proses rehabilitasi setelah serangan stroke. Dengan pengelompokan yang tepat, manajemen klinis dapat mengidentifikasi jenis komorbiditas dominan dalam populasi tertentu.

Hasil clustering membantu tenaga medis memahami profil pasien yang berbeda dari populasi umum sehingga dapat memengaruhi prioritas penanganan medis dan edukasi kesehatan (Id et al., 2023). Misalnya, pola yang ditemukan dapat digunakan untuk membuat program pencegahan dan intervensi yang

lebih tepat sasaran sesuai kelompok risiko masing-masing. Dengan analisis pola, tenaga medis dapat memfokuskan program rehabilitasi pada pasien dengan karakteristik klinis tertentu.
Penelitian terkait juga menunjukkan bahwa K-Means dapat diimplementasikan secara efektif menggunakan alat data mining visual seperti RapidMiner yang memberikan alur kerja terstruktur mulai dari preprocessing, normalisasi, pemilihan jumlah cluster (k), hingga visualisasi hasil (Rumah et al., 2024). RapidMiner menyediakan fitur grafik dan visualisasi yang memudahkan interpretasi hasil clustering tanpa perlu menulis kode pemrograman. Ini memungkinkan peneliti kesehatan dengan latar belakang non-teknis untuk melakukan analisis data secara mandiri.

RapidMiner memungkinkan analisis data secara cepat tanpa memerlukan kemampuan pemrograman yang intensif, sehingga cocok digunakan dalam riset kesehatan berbasis data rumah sakit atau dataset Excel lokal (K-means, 2024). Selain itu, integrasi dengan berbagai format data seperti CSV dan Excel menjadikan RapidMiner fleksibel untuk penelitian dengan sumber data terbatas. Ini membuka peluang bagi institusi kesehatan untuk memanfaatkan data yang mereka miliki untuk meningkatkan kualitas layanan medis.

Berdasarkan urgensi tersebut, penelitian ini bertujuan untuk menganalisis pola risiko stroke dengan menerapkan model K-Means clustering berbasis RapidMiner pada dataset pasien, sehingga hasilnya diharapkan memberikan insight mengenai pengelompokan risiko stroke dan mendukung strategi pencegahan medis berbasis data (Wala & Umar, 2024). Temuan penelitian ini diharapkan dapat memberikan kontribusi pada metode deteksi dini berbasis data dan membantu institusi kesehatan dalam memahami distribusi risiko stroke pada populasi pasien tertentu.

METODE PENELITIAN

Penelitian ini bertujuan untuk mengidentifikasi dan menganalisis pola risiko stroke pada pasien dengan menerapkan metode klastering K-Means menggunakan perangkat lunak RapidMiner. Metode penelitian disusun secara bertahap dan sistematis agar setiap proses analisis dapat dilakukan secara terkontrol serta menghasilkan pola klaster yang valid dan dapat dipertanggungjawabkan secara ilmiah.

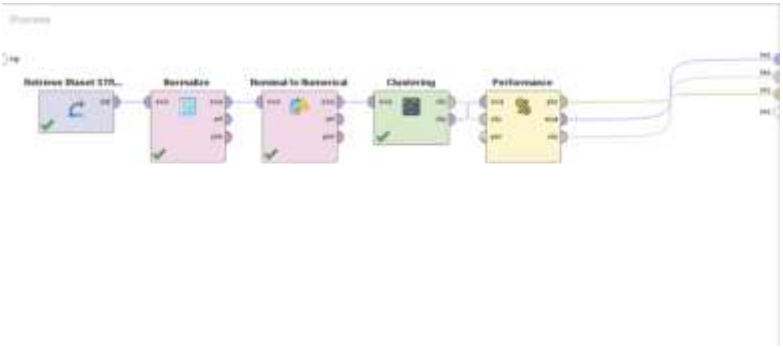
1.Pengumpulan Dataset

Dataset yang digunakan dalam penelitian ini merupakan data pasien stroke yang tersimpan dalam format Microsoft Excel. Dataset tersebut berisi berbagai atribut klinis dan demografis yang berkaitan dengan faktor risiko stroke, seperti usia, jenis kelamin, status hipertensi, riwayat penyakit jantung, kadar

e-ISSN3032-601X & p-ISSN3032-7105

glukosa darah, dan indeks massa tubuh (BMI). Setiap baris data merepresentasikan satu individu pasien, sedangkan setiap kolom menggambarkan karakteristik tertentu dari pasien tersebut. Penggunaan data klinis dan demografis dalam analisis klastering terbukti efektif untuk mengidentifikasi kelompok pasien dengan karakteristik risiko yang serupa pada penelitian kesehatan, khususnya pada kasus stroke (Ren & Fu, 2021).

2. Preprocessing Data



Tahap preprocessing dilakukan untuk memastikan kualitas dataset sebelum proses klastering. Proses ini meliputi pemeriksaan terhadap nilai kosong (*missing values*), duplikasi data, serta inkonsistensi tipe data yang dapat memengaruhi hasil analisis. Pemeriksaan awal dilakukan menggunakan fitur eksplorasi data pada RapidMiner.
Selanjutnya, atribut numerik dinormalisasi menggunakan operator *Normalize* agar seluruh variabel memiliki skala yang sebanding, sedangkan atribut kategorikal seperti jenis kelamin dan status penyakit dikonversi menjadi bentuk numerik menggunakan operator *Nominal to Numerical*. Tahap ini penting karena algoritma K-Means hanya dapat bekerja secara optimal pada data numerik berbasis jarak (Suhito et al., 2025).

3. Pembagian Data

Dataset yang telah diproses kemudian dibagi menjadi dua bagian, yaitu data pelatihan dan data pengujian. Pembagian dilakukan menggunakan operator *Split Data* dengan proporsi 90% data pelatihan dan 10% data pengujian. Pembagian ini bertujuan untuk memastikan algoritma memperoleh data yang cukup dalam membentuk cluster yang stabil sekaligus menjaga validitas proses evaluasi. Pembagian data yang tepat berperan penting dalam meningkatkan reliabilitas hasil analisis klastering serta meminimalkan bias penelitian (Jain, 2010).

4. Proses Klastering K-Means

Proses klastering dilakukan menggunakan operator *Clustering (K-Means)* pada RapidMiner. Algoritma K-Means bekerja dengan mengelompokkan data ke dalam sejumlah cluster berdasarkan jarak Euclidean antara data dan pusat cluster (*centroid*). Jumlah cluster ditentukan berdasarkan karakteristik dataset dan tujuan penelitian.
Metode K-Means banyak digunakan dalam bidang kesehatan karena mampu mengelompokkan data medis

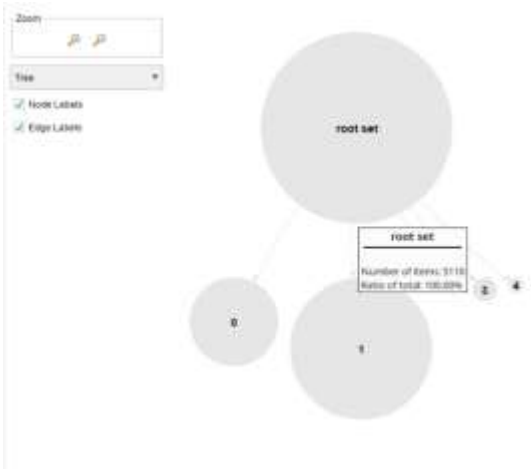
berskala besar dan menghasilkan pola yang mudah diinterpretasikan oleh peneliti maupun tenaga kesehatan (Shin et al., n.d.).

5. Evaluasi Hasil Klastering

Tahap evaluasi dilakukan untuk menilai kualitas hasil klastering yang diperoleh. Evaluasi dilakukan menggunakan operator *Performance (Clustering)* pada RapidMiner untuk mengukur tingkat pemisahan antar cluster dan tingkat keseragaman data dalam satu cluster. Penilaian kualitas klaster mengacu pada indeks evaluasi klaster, seperti Davies–Bouldin Index, yang digunakan untuk memastikan bahwa cluster yang terbentuk memiliki pemisahan yang jelas dan tidak saling tumpang tindih (Davies & Bouldin, 1979).

HASIL DAN PEMBAHASAN

Hasil Klastering Data Pasien Stroke



Berdasarkan tahapan metode penelitian yang telah dilakukan, proses klastering terhadap data pasien stroke berhasil menghasilkan beberapa kelompok pasien dengan karakteristik risiko yang berbeda. Penerapan algoritma K-Means pada dataset yang telah melalui tahap preprocessing dan normalisasi memungkinkan pengelompokan data berdasarkan tingkat kemiripan atribut klinis dan demografis pasien, seperti usia, status hipertensi, kadar glukosa darah, dan indeks massa tubuh (BMI).

Hasil klastering menunjukkan bahwa setiap cluster memiliki pola karakteristik yang relatif homogen di dalam cluster dan heterogen antar cluster. Cluster tertentu didominasi oleh pasien dengan usia lanjut, kadar glukosa darah tinggi, serta riwayat hipertensi dan penyakit jantung, yang mengindikasikan kelompok dengan risiko stroke yang lebih tinggi. Sementara itu, cluster lainnya menunjukkan karakteristik pasien dengan usia lebih muda, nilai BMI normal, serta kadar glukosa darah yang relatif stabil, sehingga dapat diinterpretasikan sebagai kelompok dengan risiko stroke yang lebih rendah.

Distribusi data pada setiap cluster menunjukkan bahwa algoritma K-Means mampu membagi dataset secara proporsional tanpa menghasilkan cluster yang terlalu kecil atau terlalu dominan. Hal ini

menandakan bahwa struktur data memiliki pola alami yang dapat dieksplorasi secara efektif melalui pendekatan klastering.

Evaluasi Kualitas Klastering

PerformanceVector

```
PerformanceVector:
Avg. within centroid distance: 4.919
Avg. within centroid distance_cluster_0: 3.781
Avg. within centroid distance_cluster_1: 4.733
Avg. within centroid distance_cluster_2: 10.832
Avg. within centroid distance_cluster_3: 5.409
Avg. within centroid distance_cluster_4: 7.498
Davies Bouldin: 1.431
```

Evaluasi kualitas hasil klastering dilakukan menggunakan operator *Performance (Clustering)* pada RapidMiner. Evaluasi ini bertujuan untuk menilai sejauh mana cluster yang terbentuk memiliki pemisahan yang jelas dan tingkat keseragaman yang baik dalam setiap cluster. Salah satu indikator utama yang digunakan adalah *Davies–Bouldin Index* (DBI), di mana nilai yang lebih rendah menunjukkan kualitas klastering yang lebih baik.

Hasil evaluasi menunjukkan bahwa nilai indeks klaster yang diperoleh berada pada rentang yang dapat diterima, yang menandakan bahwa jarak antar cluster cukup besar dibandingkan dengan jarak antar data di dalam satu cluster. Dengan demikian, cluster yang terbentuk tidak saling tumpang tindih secara signifikan dan memiliki karakteristik yang cukup jelas untuk dibedakan.

Selain itu, hasil evaluasi juga menunjukkan bahwa normalisasi data berperan penting dalam meningkatkan kualitas klastering. Tanpa normalisasi, atribut dengan skala besar seperti kadar glukosa darah berpotensi mendominasi proses pembentukan cluster, sedangkan setelah normalisasi, setiap atribut memberikan kontribusi yang seimbang dalam perhitungan jarak Euclidean.

Pembahasan

Hasil penelitian ini menunjukkan bahwa metode klastering K-Means mampu mengidentifikasi pola risiko stroke secara efektif berdasarkan data klinis dan demografis pasien. Pengelompokan yang dihasilkan memberikan gambaran bahwa faktor-faktor seperti usia, hipertensi, riwayat penyakit jantung, kadar glukosa darah, dan BMI memiliki keterkaitan yang kuat dalam membentuk kelompok risiko stroke tertentu.

Pendekatan klastering memungkinkan analisis dilakukan secara tidak terawasi (unsupervised), sehingga sangat bermanfaat dalam kondisi di mana label risiko tidak tersedia atau belum terdefinisi secara jelas. Dengan mengandalkan kemiripan data, K-Means mampu mengungkap struktur tersembunyi dalam dataset dan memberikan wawasan awal mengenai segmentasi pasien berdasarkan tingkat risiko.

Dari sisi praktis, hasil klastering ini dapat dimanfaatkan sebagai dasar dalam pengambilan keputusan di bidang kesehatan, khususnya untuk mendukung strategi pencegahan stroke. Tenaga medis dapat menggunakan informasi cluster untuk memfokuskan perhatian pada kelompok pasien dengan risiko tinggi melalui program pemantauan intensif, edukasi kesehatan, dan intervensi dini. Sementara itu, kelompok dengan risiko lebih rendah dapat diarahkan pada strategi pencegahan jangka panjang.

Secara keseluruhan, hasil dan pembahasan ini menegaskan bahwa metode K-Means berbasis RapidMiner merupakan pendekatan yang efektif dan efisien dalam menganalisis pola risiko stroke. Metode ini tidak hanya mampu mengelompokkan pasien berdasarkan karakteristik yang relevan, tetapi juga memberikan interpretasi yang mudah dipahami dan dapat dijadikan landasan awal dalam pengembangan sistem pendukung keputusan berbasis data di bidang kesehatan.

KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa metode klastering K-Means mampu mengidentifikasi dan menganalisis pola risiko stroke pada pasien secara efektif menggunakan data klinis dan demografis. Melalui tahapan penelitian yang sistematis, mulai dari pengumpulan dataset, preprocessing, normalisasi data, hingga proses klastering dan evaluasi, penelitian ini berhasil membentuk kelompok pasien yang memiliki karakteristik risiko yang relatif homogen dalam satu cluster dan heterogen antar cluster.

Hasil klastering menunjukkan adanya perbedaan pola yang jelas antar kelompok pasien, di mana cluster tertentu didominasi oleh pasien dengan faktor risiko tinggi seperti usia lanjut, kadar glukosa darah yang tinggi, serta riwayat hipertensi dan penyakit jantung, sedangkan cluster lainnya menunjukkan karakteristik risiko yang lebih rendah. Evaluasi kualitas klastering menggunakan metrik evaluasi clustering menunjukkan bahwa cluster yang terbentuk memiliki tingkat pemisahan yang baik dan tidak saling tumpang tindih secara signifikan.

Dengan demikian, penelitian ini membuktikan bahwa pendekatan klastering K-Means berbasis RapidMiner dapat digunakan sebagai metode analisis yang efektif untuk menggali pola risiko stroke secara tidak terawasi. Temuan ini diharapkan dapat menjadi dasar awal dalam mendukung pengambilan keputusan di bidang kesehatan, khususnya dalam perencanaan strategi pencegahan dan pemantauan pasien stroke berbasis data.

DAFTAR PUSTAKA

Care, P., Delord, M., Sun, X., Learoyd, A., Curcin, V., Wolfe, C., Ashworth, M., & Douiri, A. (2024). Patient - oriented unsupervised learning to uncover the patterns of multimorbidity associated with stroke using primary care electronic health records. *BMC Primary Care*. https://doi.org/10.1186/s12875-024-02636-6

Ceskoutsé, R. F. T., Bomgni, A. B., Zanfack, D. R. G., Agany, D. D. M., Thomas, B. B., & Zohim, E. G. (2025). *Sub-clustering based recommendation system for stroke patient: Identification of a specific drug class for a given patient*. 1–35. https://doi.org/10.1016/j.compbiomed.2024.108117.Sub-clustering

Davies, D. L., & Bouldin, D. W. (1979). A Cluster Separation Measure. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *PAMI-1*(2), 224–227.

https://doi.org/10.1109/TPAMI.1979.4766909

Guo, X., Liu, P., Guo, J., Zhang, N., Huang, H., Liu, J., Tan, Z., & Dan, G. (2025). *An unsupervised cluster analysis of multimorbidity patterns in older adults in Shenzhen , China*. *June*, 1–14. https://doi.org/10.3389/fpubh.2025.1557721

Id, S. S. Z., Rutter, M. K., Sun, L. Y., Ullah, W., Rashid, M., Id, D. M. A., Steinke, D. T., Weng, S., Id, E. K., & Mamas, M. A. (2023). *PLOS ONE Comorbidity clusters and in-hospital outcomes in patients admitted with acute myocardial infarction in the USA : A national population-based study*. 1–18. https://doi.org/10.1371/journal.pone.0293314

Jain, A. K. (2010). Data clustering: 50 years beyond K-means. *Pattern Recognition Letters*, *31*(8), 651–666. https://doi.org/https://doi.org/10.1016/j.patrec.2009.09.011

Jiang, Y., Dang, Y., Wu, Q., & Yuan, B. (n.d.). *Using a k-means clustering to identify novel phenotypes of acute ischemic stroke and development of its Clinlabomics models*.

K-means, A. (2024). *Clustering Pasien Rawat Inap Di RS USU Menggunakan*. *03*(2), 54–63.

Kim, J. T., Kim, N. R., Choi, S. H., Oh, S., Park, M. S., Lee, S. H., Kim, B. C., Choi, J., & Kim, M. S. (2022). Neural network - based clustering model of ischemic stroke patients with a maximally distinct distribution of 1 - year vascular outcomes. *Scientific Reports*, *0123456789*, 1–10. https://doi.org/10.1038/s41598-022-13636-w

Mao, H. (2025). *Cluster Analysis of Patients With Acute Ischemic Stroke : Identifying Characteristics of Long Hospital Stays in a Comprehensive Hospital*. 1–10. https://doi.org/10.1002/brb3.70940

Ren, Z., & Fu, X. (2021). Stroke Risk Factors in United States: An Analysis of the 2013-2018 National Health and Nutrition Examination Survey. *International Journal of General Medicine*, *14*, 6135–6147. https://doi.org/10.2147/IJGM.S327075

Rumah, P., Royal, S., Hidayah, A., Dulisep, D., & Angga, B. (2024). *Implementasi Algoritma K - Means Menggunakan RapidMiner untuk Klasterisasi Data Obat*. *7*(2), 200–211.

Shin, S., Chang, W. H., Kim, D. Y., Lee, J., Sohn, M. K., Song, M., Shin, Y., Lee, Y., Joo, M. C., Lee, S. Y., Han, J., Ahn, J., Oh, G., Kim, Y., Kim, K., & Kim, Y. (n.d.). *Clustering and prediction of long-term functional recovery patterns in first-time stroke patients*. https://doi.org/10.30597/mkmi.v21i3.45641

Suhito, H. P., Azam, M., Nur, D., Ningrum, A., City, S., & Office, H. (2025). *Media Kesehatan Masyarakat Indonesia Indonesia : Based on Indonesian Health Survey Data*. *21*(3), 225–235. https://doi.org/10.30597/mkmi.v21i3.45641

Wala, J., & Umar, R. (2024). *Implementasi K-Means Clustering pada Pengelompokan Pasien Penyakit Jantung*. *9*(3), 205–216. https://doi.org/10.14421/jiska.2024.9.3.205-216