

Penggunaan Algoritma *K-Means Clustering* Aplikasi Rapid Miner untuk Menganalisis Tingkat Kematian Pasien Diabetes Mellitus di Rumah Sakit Umum Daerah Ibnu Sina Gresik

Suci Ariani ^{1*}, Resta Dwi Yuliani ², Auliya Rabbani ³

¹⁻³ Manajemen Informasi Kesehatan, Universitas Muhammadiyah Sidoarjo, Indonesia

Email : Suciariani@umsida.ac.id

*Penulis Korespondensi: Suciariani@umsida.ac.id

Abstract. *Diabetes Mellitus is one of the chronic diseases with high morbidity and mortality rates, making data-driven analysis necessary to understand patient mortality patterns. This study aims to analyze the mortality rate of Diabetes Mellitus patients based on age and length of hospitalization using a data mining approach with the K-Means Clustering method. The study employs a quantitative approach using secondary data obtained from the medical records of Diabetes Mellitus patients at Ibnu Sina Regional General Hospital, Gresik Regency, in December 2022. The dataset consists of 266 patient records with variables including age, length of stay, and final patient status. Data analysis was conducted through preprocessing stages, including data cleaning, transformation, and normalization, followed by the clustering process using the K-Means algorithm with the assistance of the RapidMiner application. The results show that patient data are divided into three clusters based on age ranges: 0–40 years, 41–55 years, and 56–90 years. The cluster with the age range of 56–90 years has the highest number of patient deaths compared to the other clusters. Meanwhile, the length of hospitalization does not show a significant effect on patient mortality. This study is expected to serve as a consideration for hospitals and health institutions in efforts to prevent and manage Diabetes Mellitus, particularly among the elderly population.*

Keywords: *Data Mining, Diabetes Mellitus, K-Means Clustering, Mortality Rate, Rapid Miner.*

Abstrak. Diabetes Mellitus merupakan salah satu penyakit kronis yang memiliki angka morbiditas dan mortalitas yang tinggi, sehingga diperlukan analisis berbasis data untuk memahami pola tingkat kematian pasien. Penelitian ini bertujuan untuk menganalisis tingkat kematian pasien Diabetes Mellitus berdasarkan faktor usia dan lama perawatan menggunakan pendekatan data mining dengan metode K-Means Clustering. Penelitian ini menggunakan pendekatan kuantitatif dengan data sekunder yang diperoleh dari rekam medis pasien Diabetes Mellitus di Rumah Sakit Umum Daerah Ibnu Sina Kabupaten Gresik pada bulan Desember 2022. Dataset yang digunakan berjumlah 266 data pasien dengan variabel usia, lama rawat inap, dan status akhir pasien. Analisis data dilakukan melalui tahapan preprocessing yang meliputi pembersihan data, transformasi, dan normalisasi, kemudian dilanjutkan dengan proses clustering menggunakan algoritma K-Means dengan bantuan aplikasi RapidMiner. Hasil penelitian menunjukkan bahwa data pasien terbagi ke dalam tiga kluster berdasarkan rentang usia, yaitu usia 0–40 tahun, 41–55 tahun, dan 56–90 tahun. Kluster dengan rentang usia 56–90 tahun memiliki jumlah pasien meninggal dunia tertinggi dibandingkan kluster lainnya. Sementara itu, lama rawat inap tidak menunjukkan pengaruh yang signifikan terhadap tingkat kematian pasien. Penelitian ini diharapkan dapat menjadi bahan pertimbangan bagi pihak rumah sakit dan instansi kesehatan dalam upaya pencegahan serta penanganan Diabetes Mellitus, khususnya pada kelompok usia lanjut.

Kata kunci: Data Mining, Diabetes Mellitus, K-Means Clustering, Rapid Miner, Tingkat Kematian.

1. LATAR BELAKANG

Penyakit Diabetes Mellitus telah menjadi masalah kesehatan serius dengan angka kematian yang terus meningkat di berbagai fasilitas pelayanan kesehatan. Beban klinis akibat komplikasi diabetes dipengaruhi oleh usia, durasi perawatan, dan kondisi komorbid pasien. Data rekam medis yang tersimpan di rumah sakit sering kali belum dimanfaatkan secara optimal untuk analisis pola kematian. Pendekatan data mining telah digunakan untuk

mengelola data kesehatan berukuran besar secara sistematis. Algoritma K-Means Clustering mampu mengelompokkan data berdasarkan kesamaan karakteristik tanpa label awal. Penggunaan aplikasi RapidMiner mempermudah proses analisis data secara visual dan terstruktur. Efektivitas metode tersebut telah dibuktikan pada analisis risiko serangan jantung menggunakan K-Means Clustering berbasis RapidMiner (Aditiya et al., 2025).

Penyebaran dan tingkat keparahan penyakit kronis dapat dipetakan melalui teknik pengelompokan data pasien. Pola tertentu pada data kesehatan telah berhasil diidentifikasi menggunakan algoritma K-Means. Pemanfaatan RapidMiner telah mempercepat proses komputasi dan interpretasi hasil klaster. Algoritma tersebut diaplikasikan untuk mendeteksi penyebaran penyakit tuberkulosis berdasarkan karakteristik pasien. Hasil klaster memberikan gambaran tingkat risiko penyakit secara tersegmentasi. Pendekatan analitik ini dinilai efektif untuk mendukung pengambilan keputusan medis. Keunggulan tersebut ditunjukkan pada penelitian deteksi penyebaran tuberkulosis berbasis K-Means Clustering (Kusuma & Firmansyah, 2022).

Analisis risiko penyakit kardiovaskular telah dikembangkan melalui pendekatan clustering data. Hubungan antara variabel klinis dan tingkat risiko dapat dijelaskan melalui pembentukan klaster pasien. Teknik K-Means mampu mengelompokkan data sentimen dan parameter medis secara efisien. Proses analisis dilakukan dengan bantuan RapidMiner untuk meningkatkan akurasi pemetaan risiko. Hasil pengelompokan memperlihatkan perbedaan tingkat risiko secara signifikan. Pola tersebut digunakan sebagai dasar pemahaman kondisi kesehatan pasien. Keberhasilan metode ini dilaporkan pada analisis sentimen risiko serangan jantung menggunakan K-Means Clustering (Aditiya et al., 2025).

Pemetaan risiko penyakit stroke telah dilakukan melalui pengelompokan data pasien. Variabel klinis digunakan sebagai dasar pembentukan klaster risiko. Algoritma K-Means diterapkan untuk mengidentifikasi kelompok pasien berisiko tinggi. Hasil klaster disajikan secara sistematis untuk memudahkan interpretasi. Pendekatan ini memberikan gambaran distribusi risiko penyakit secara objektif. Pengelompokan data membantu tenaga medis memahami prioritas penanganan pasien. Penerapan metode tersebut dibuktikan pada analisis risiko stroke menggunakan K-Means Clustering (Nurhakim et al., 2024).

Tingkat kematian pasien penyakit menular juga dianalisis menggunakan teknik clustering. Faktor usia dan lama perawatan dijadikan variabel utama dalam pembentukan klaster. Algoritma K-Means mampu mengelompokkan pasien berdasarkan pola mortalitas. RapidMiner digunakan untuk mengolah data secara terstruktur dan efisien. Hasil klaster menunjukkan perbedaan tingkat kematian antar kelompok pasien. Informasi tersebut

memberikan gambaran kondisi pelayanan kesehatan rumah sakit. Pendekatan ini diterapkan pada analisis mortalitas pasien tuberkulosis menggunakan K-Means Clustering (Upuy, 2022).

Analisis risiko pasien diabetes telah dikembangkan melalui teknik pengelompokan data. Karakteristik pasien dikelompokkan berdasarkan kesamaan variabel klinis. Algoritma K-Means digunakan untuk membentuk klaster risiko diabetes. Hasil klaster membantu mengidentifikasi kelompok pasien dengan tingkat risiko berbeda. Data kesehatan diolah secara sistematis untuk menghasilkan informasi bermakna. Pendekatan clustering memperkuat pemahaman terhadap distribusi risiko pasien. Penerapan metode tersebut dilaporkan pada analisis clustering pasien berisiko diabetes (Firmansyah, 2025).

Pengelompokan penyakit pasien di fasilitas kesehatan tingkat pertama telah dilakukan menggunakan data mining. Algoritma K-Means digunakan untuk mengelompokkan jenis penyakit berdasarkan data kunjungan pasien. Proses clustering membantu petugas kesehatan memahami dominasi penyakit tertentu. Data dianalisis secara terstruktur melalui tahapan preprocessing dan klasterisasi. Hasil pengelompokan memberikan gambaran distribusi penyakit secara jelas. Informasi tersebut mendukung peningkatan kualitas pelayanan kesehatan. Implementasi metode ini ditunjukkan pada pengelompokan penyakit pasien menggunakan K-Means Clustering (Halawa et al., 2023).

Penerapan data mining pada kasus penyakit diare telah menghasilkan pengelompokan data yang informatif. Variabel pasien dianalisis untuk membentuk klaster penyakit. Algoritma K-Means digunakan untuk mengidentifikasi pola sebaran kasus. Hasil klaster membantu memahami karakteristik kelompok pasien. Proses analisis dilakukan secara sistematis melalui tahapan data mining. Informasi yang dihasilkan memperkuat dasar pengambilan keputusan kesehatan. Pendekatan ini dibuktikan pada clustering penyakit diare menggunakan K-Means (Kholil et al., 2024).

Pemanfaatan RapidMiner pada bidang kesehatan telah diperluas ke berbagai objek analisis. Teknik K-Means diterapkan untuk mengelompokkan fasilitas kesehatan berdasarkan capaian layanan. Proses clustering mempermudah pemetaan kinerja pelayanan kesehatan. Data diolah secara visual sehingga hasil lebih mudah dipahami. Pendekatan ini menunjukkan fleksibilitas algoritma K-Means. Analisis berbasis klaster memberikan gambaran perbedaan karakteristik antar kelompok. Penerapan RapidMiner dengan metode K-Means ditunjukkan pada pengelompokan puskesmas berdasarkan cakupan imunisasi (Ramadhan & Saprudin, 2022).

Penerapan algoritma K-Means telah banyak digunakan pada berbagai bidang analisis data. Perbandingan metode clustering menunjukkan keunggulan K-Means dalam efisiensi

komputasi. Pengelompokan data berbasis kesamaan karakteristik menghasilkan informasi terstruktur. Metode ini dinilai efektif untuk analisis data berskala besar. RapidMiner mendukung proses clustering melalui antarmuka visual yang intuitif. Penerapan K-Means pada pemetaan wilayah rawan penyakit menunjukkan hasil signifikan. Keandalan metode ini ditegaskan pada penelitian pemetaan daerah rawan demam berdarah menggunakan K-Means Clustering (Suprihatin et al., 2019).

2. KAJIAN TEORITIS

Diabetes Mellitus dan Tingkat Kematian Pasien

Diabetes Mellitus merupakan penyakit metabolik kronis yang ditandai oleh peningkatan kadar glukosa darah secara persisten. Komplikasi diabetes berkontribusi besar terhadap peningkatan angka kematian pasien di rumah sakit. Faktor usia, durasi perawatan, serta penyakit penyerta berperan dalam memperberat kondisi pasien. Data klinis pasien diabetes menghasilkan informasi kompleks yang sulit dianalisis secara manual. Oleh karena itu, pendekatan analitik berbasis data diperlukan untuk memahami pola kematian pasien. Pengelompokan data kesehatan memungkinkan identifikasi karakteristik pasien secara sistematis. Pendekatan serupa telah diterapkan pada analisis risiko dan mortalitas penyakit menggunakan metode clustering (Upuy, 2022).

Data Mining dalam Bidang Kesehatan

Data mining merupakan proses penggalian pengetahuan dari kumpulan data berukuran besar. Data kesehatan yang tersimpan pada sistem rumah sakit memiliki potensi besar untuk dianalisis. Proses data mining meliputi tahapan seleksi data, pembersihan data, transformasi, dan pemodelan. Informasi tersembunyi dapat diungkap melalui pola yang terbentuk dari data pasien. Pendekatan ini telah banyak digunakan pada analisis penyakit dan risiko kesehatan. Pemanfaatan data mining membantu pengambilan keputusan berbasis bukti. Efektivitas data mining dalam bidang kesehatan telah dibuktikan melalui berbagai studi clustering penyakit (Kholil et al., 2024).

Konsep Clustering dalam Analisis Data

Clustering merupakan teknik pengelompokan data berdasarkan tingkat kemiripan tertentu. Objek data dalam satu kluster memiliki karakteristik yang lebih mirip dibandingkan dengan kluster lain. Teknik ini bersifat unsupervised karena tidak memerlukan label kelas. Clustering digunakan untuk menemukan struktur alami pada data. Analisis kluster membantu penyederhanaan data kompleks menjadi kelompok yang mudah dipahami. Metode clustering

banyak diterapkan pada bidang kesehatan untuk pemetaan risiko penyakit. Penerapan clustering telah terbukti efektif dalam pengelompokan data medis (Handoyo et al., 2014).

Algoritma K-Means Clustering

Algoritma K-Means merupakan metode clustering yang mengelompokkan data berdasarkan jarak terdekat ke pusat klaster. Proses K-Means diawali dengan penentuan jumlah klaster. Setiap data ditempatkan pada klaster terdekat berdasarkan nilai centroid. Centroid diperbarui secara iteratif hingga mencapai kondisi stabil. Algoritma ini dikenal karena kesederhanaan dan efisiensi komputasinya. K-Means banyak digunakan pada analisis data berukuran besar. Keunggulan K-Means telah dibuktikan pada berbagai studi analisis kesehatan dan risiko penyakit (Alfina et al., 2012).

RapidMiner sebagai Alat Analisis Data

RapidMiner merupakan perangkat lunak data mining berbasis visual. Aplikasi ini menyediakan berbagai operator analisis data tanpa pemrograman kompleks. Pengguna dapat melakukan preprocessing, clustering, dan evaluasi model secara terintegrasi. RapidMiner mendukung implementasi algoritma K-Means secara efisien. Hasil analisis disajikan dalam bentuk visual yang mudah dipahami. Aplikasi ini banyak digunakan pada penelitian kesehatan dan sistem informasi. Keandalan RapidMiner telah ditunjukkan pada berbagai penelitian clustering berbasis K-Means (Ramadhan & Saprudin, 2022).

Penerapan K-Means Clustering pada Data Kesehatan

Algoritma K-Means telah diterapkan pada berbagai kasus penyakit. Analisis penyebaran penyakit menular dilakukan melalui pembentukan klaster pasien. Data usia, durasi perawatan, dan kondisi klinis digunakan sebagai variabel clustering. Hasil klaster memberikan gambaran tingkat risiko penyakit. Pendekatan ini membantu tenaga kesehatan dalam pengelompokan pasien. Analisis berbasis klaster mendukung efisiensi pelayanan medis. Penerapan tersebut dibuktikan pada deteksi penyebaran tuberkulosis menggunakan K-Means (Kusuma & Firmansyah, 2022).

Analisis Risiko dan Mortalitas Berbasis Clustering

Clustering digunakan untuk menganalisis risiko dan tingkat kematian pasien. Pengelompokan pasien memungkinkan identifikasi kelompok dengan risiko kematian tinggi. Variabel klinis dianalisis secara simultan untuk membentuk klaster. Hasil analisis membantu perencanaan strategi penanganan pasien. Pendekatan ini memberikan gambaran objektif terhadap kondisi pasien. Pengelompokan berbasis data meningkatkan akurasi analisis risiko. Metode ini diterapkan pada analisis mortalitas pasien tuberkulosis menggunakan K-Means (Upuy, 2022).

Relevansi K-Means Clustering pada Pasien Diabetes Mellitus

Pasien Diabetes Mellitus memiliki karakteristik klinis yang beragam. Variasi kondisi pasien memerlukan pendekatan analisis yang sistematis. Algoritma K-Means mampu mengelompokkan pasien berdasarkan kesamaan karakteristik. Hasil klaster membantu identifikasi kelompok pasien dengan risiko tinggi. Analisis ini mendukung pengambilan keputusan berbasis data kesehatan. Pendekatan clustering memperkuat pemahaman terhadap distribusi risiko pasien diabetes. Penerapan K-Means pada analisis risiko diabetes telah dilakukan secara efektif (Firmansyah, 2025).

3. METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif dengan metode data mining untuk menganalisis data rekam medis pasien Diabetes Mellitus. Pendekatan kuantitatif dipilih karena penelitian berfokus pada pengolahan data numerik dan kategorikal secara sistematis untuk menghasilkan temuan yang objektif dan terukur (Sihotang, 2023; Zulfikar et al., 2024). Metode data mining digunakan untuk menggali pola tersembunyi dalam data kesehatan yang berukuran besar. Analisis dilakukan tanpa label kelas awal sehingga teknik clustering dianggap sesuai. Pendekatan ini memungkinkan pengelompokan pasien berdasarkan karakteristik tertentu secara sistematis dan terstruktur (Waruwu et al., 2025).

Penelitian dilaksanakan di Rumah Sakit Umum Daerah Ibnu Sina Gresik dengan objek penelitian berupa data pasien Diabetes Mellitus yang menjalani perawatan. Pemilihan lokasi didasarkan pada ketersediaan data rekam medis yang lengkap dan terstruktur. Data yang digunakan merupakan data sekunder yang diperoleh dari bagian rekam medis rumah sakit, sesuai dengan karakteristik penelitian kuantitatif yang memanfaatkan data terdokumentasi (Zulfikar et al., 2024). Variabel penelitian meliputi usia, jenis kelamin, lama perawatan, dan status akhir pasien. Seluruh data telah melalui proses seleksi, serta kerahasiaan identitas pasien dijaga sesuai etika penelitian.

Pengumpulan data dilakukan melalui metode dokumentasi terhadap arsip rekam medis pasien. Metode dokumentasi dipilih karena mampu menyediakan data yang objektif, akurat, dan dapat diverifikasi kembali (Sihotang, 2023). Data dikumpulkan sesuai variabel yang telah ditentukan dengan izin resmi dari pihak rumah sakit. Seluruh data kemudian dikompilasi ke dalam satu basis data terpusat. Data yang tidak lengkap atau tidak sesuai dikeluarkan dari proses analisis agar kualitas data tetap terjaga.

Analisis data dilakukan menggunakan metode clustering dengan algoritma K-Means. Tahap awal analisis meliputi preprocessing data yang mencakup pembersihan, transformasi, dan normalisasi data numerik. Data kategorikal dikonversi ke bentuk numerik agar dapat diproses oleh algoritma. Tahapan preprocessing ini bertujuan menghasilkan data yang siap dianalisis dan meningkatkan akurasi hasil pengelompokan, sebagaimana dianjurkan dalam penelitian kuantitatif berbasis analisis data (Waruwu et al., 2025).

Penerapan algoritma K-Means dilakukan dengan menentukan jumlah kluster terlebih dahulu. Setiap data pasien dikelompokkan ke dalam kluster berdasarkan jarak terdekat dengan centroid. Nilai centroid diperbarui secara iteratif hingga tidak terjadi perubahan kluster. Proses ini bertujuan memperoleh kelompok data yang homogen di dalam kluster dan heterogen antar kluster (Zulfikar et al., 2024). Setiap kluster kemudian dianalisis untuk melihat pola tingkat kematian pasien Diabetes Mellitus.

Seluruh proses analisis dibantu menggunakan aplikasi RapidMiner. Aplikasi ini memfasilitasi proses preprocessing, penerapan algoritma K-Means, serta visualisasi hasil clustering melalui workflow yang terstruktur. Hasil analisis ditampilkan dalam bentuk tabel dan grafik untuk memudahkan interpretasi. Penggunaan perangkat lunak analisis data mendukung efisiensi dan keakuratan pengolahan data kuantitatif (Waruwu et al., 2025).

Tahapan penelitian dilakukan secara sistematis mulai dari identifikasi masalah hingga penarikan kesimpulan. Proses dimulai dengan pengumpulan dan pengolahan data, dilanjutkan dengan analisis clustering menggunakan algoritma K-Means. Hasil kluster dianalisis untuk mengidentifikasi pola tingkat kematian pasien. Interpretasi hasil dilakukan berdasarkan karakteristik setiap kluster yang terbentuk, kemudian disimpulkan sesuai tujuan penelitian (Sihotang, 2023).

4. HASIL DAN PEMBAHASAN

Bab ini membahas hasil analisis data pasien Diabetes Mellitus yang dirawat di Rumah Sakit Umum Daerah Ibnu Sina Kabupaten Gresik pada bulan Desember 2022. Data yang digunakan berupa dataset rekam medis pasien rawat inap sebanyak 266 data, yang mencakup informasi riwayat pasien (hidup atau meninggal), usia, dan lama rawat inap. Dataset ini digunakan sebagai dasar untuk menganalisis tingkat kematian pasien Diabetes Mellitus berdasarkan faktor usia dan lama perawatan melalui pendekatan data mining menggunakan metode clustering.

Atribut yang digunakan dalam penelitian ini terdiri dari tiga variabel utama, yaitu riwayat pasien, usia, dan lama rawat inap. Atribut riwayat bertipe data polynominal yang menunjukkan kondisi akhir pasien hidup atau meninggal dunia. Atribut usia dan lama rawat inap bertipe data integer yang masing-masing merepresentasikan umur pasien serta durasi perawatan di rumah sakit. Ketiga atribut ini dipilih karena dianggap relevan dalam menggambarkan karakteristik klinis pasien Diabetes Mellitus dan berpotensi memengaruhi tingkat kematian.

Implementasi metode K-Means Clustering dilakukan menggunakan aplikasi RapidMiner Studio versi 10.1 sebagai alat bantu analisis. Proses dimulai dengan mengimpor data pasien dalam format file Excel ke dalam RapidMiner melalui operator Read Excel. Pada tahap ini dilakukan penentuan tipe data dan atribut agar sesuai dengan kebutuhan analisis clustering. Setelah proses impor berhasil dan data dinyatakan siap digunakan tanpa error, operator K-Means ditambahkan dan dihubungkan dengan dataset untuk melakukan proses pengelompokan data.

Pada tahap clustering, nilai parameter K ditentukan sebanyak tiga klaster. Penentuan jumlah klaster ini bertujuan untuk mengelompokkan pasien ke dalam kategori usia yang berbeda. Proses clustering dilakukan secara iteratif hingga centroid setiap klaster stabil. Berdasarkan hasil pengolahan data terhadap 266 dataset pasien, diperoleh tiga klaster, yaitu Cluster 0 dengan jumlah 85 data, Cluster 1 dengan jumlah 158 data, dan Cluster 2 dengan jumlah 23 data. Rentang usia pada Cluster 0 berada pada usia 41–55 tahun, Cluster 1 pada usia 56–90 tahun, dan Cluster 2 pada usia 0–40 tahun.

Hasil clustering kemudian dianalisis untuk melihat distribusi pasien hidup dan meninggal dunia pada setiap klaster. Cluster 0 terdiri dari 68 pasien hidup dan 20 pasien meninggal dunia. Cluster 1 memiliki jumlah pasien terbanyak dengan 133 pasien hidup dan 22 pasien meninggal dunia. Sementara itu, Cluster 2 terdiri dari 19 pasien hidup dan 4 pasien meninggal dunia. Berdasarkan hasil tersebut, jumlah pasien meninggal dunia terbanyak berada pada Cluster 1 dengan rentang usia 56–90 tahun, disusul oleh Cluster 0 dengan rentang usia 41–55 tahun, dan paling sedikit berada pada Cluster 2 dengan rentang usia 0–40 tahun.

Visualisasi data menunjukkan bahwa tingkat kematian pasien Diabetes Mellitus cenderung meningkat pada kelompok usia lanjut. Grafik tingkat kematian berdasarkan faktor usia memperlihatkan bahwa pasien yang meninggal dunia rata-rata berada pada rentang usia tua, sedangkan pasien yang hidup tersebar pada berbagai kelompok usia. Hal ini menunjukkan bahwa usia merupakan faktor yang berpengaruh terhadap tingkat kematian pasien Diabetes Mellitus.

Sementara itu, analisis berdasarkan faktor lama rawat inap menunjukkan bahwa durasi perawatan tidak memiliki pengaruh signifikan terhadap tingkat kematian pasien. Pasien yang meninggal dunia tercatat memiliki lama rawat inap yang bervariasi, mulai dari 1 hari hingga lebih dari 15 hari. Di sisi lain, sebagian besar pasien yang sembuh dirawat selama kurang dari 10 hari. Temuan ini mengindikasikan bahwa lama perawatan bukan faktor utama yang menentukan tingkat kematian pasien Diabetes Mellitus di RSUD Ibnu Sina Kabupaten Gresik.

Berdasarkan hasil analisis tersebut, dapat disimpulkan bahwa tingkat kematian pasien Diabetes Mellitus tertinggi berada pada kelompok usia 56–90 tahun, sedangkan lama rawat inap tidak menunjukkan pengaruh yang signifikan terhadap tingkat kematian. Oleh karena itu, hasil penelitian ini diharapkan dapat menjadi referensi bagi pemerintah, khususnya Dinas Kesehatan, dalam melakukan upaya pencegahan dan sosialisasi gaya hidup sehat kepada masyarakat, terutama pada kelompok usia lanjut. Upaya tersebut dapat meliputi penerapan pola makan seimbang, konsumsi air putih yang cukup, istirahat teratur, aktivitas fisik yang rutin, serta pengelolaan stres yang baik guna menekan risiko komplikasi dan kematian akibat Diabetes Mellitus.

5. KESIMPULAN DAN SARAN

Dari 266 pasien Diabetes Mellitus yang dirawat inap di RSUD Ibnu Sina Gresik pada bulan Desember 2022, sebanyak 220 orang sembuh dan pasien meninggal sebanyak 46 orang, dimana tingkat kematian pasien tertinggi berada pada usia tua (lansia), yaitu berada pada cluster 1 dengan rentang usia 56-90 tahun. Kemudian, disusul ke posisi ke-2 berada pada cluster 0 dengan rentang usia 41-55 tahun, dan yang terakhir pada cluster 2 dengan jumlah pasien meninggal dunia berada pada rentang usia 0-40 tahun. Lama rawat inap pasien di RSUD Ibnu Sina Gresik bervariasi mulai dari 1 hari sampai 28 hari dan dialami oleh pasien yang sembuh dan juga meninggal, maka faktor lama perawatan tidak mempengaruhi tingkat kematian pasien Diabetes Mellitus.

Diharapkan kepada peneliti yang selanjutnya untuk dapat menambah variabel, mengkombinasikan beberapa algoritma berbeda sehingga dapat membandingkan hasil dan metode terbaik yang digunakan.

DAFTAR REFERENSI

Aditiya, B., Kurniaty, A., Muslim, A., Rachayudiza, A., Manullang, D., & Fansyuri, M. (2025). Analisis sentimen risiko serangan jantung menggunakan K-means clustering dengan

- Rapidminer. *Journal of Information Technology and Informatics Engineering*, 1(1), 108-112.
- Aditiya, B., Kurniaty, A., Muslim, A., Rachayudiza, A., Manullang, D., & Fansyuri, M. (2025). Analisis sentimen risiko serangan jantung menggunakan K-means clustering dengan Rapidminer. *Journal of Information Technology and Informatics Engineering*, 1(1), 108-112.
- Alfina, T., Santosa, B., & Hakim, J. A. R. (2012). *Analisa perbandingan metode hierarchical clustering, K-means dan gabungan keduanya dalam cluster data (Studi kasus: Problem kerja praktek jurusan teknik industri ITS)*, vol. 1.
- Firmansyah, H. (2025). Analysis of clustering of diabetes risk patients using the K-means algorithm: Analisis clustering pasien berisiko diabetes dengan algoritma K-means. *SITEDI (Sistem Informasi dan Teknologi Digital)*, 2(3), 110-113. <https://doi.org/10.70888/sitedi.v2i1.44>
- Halawa, H., Harmaja, O. J., & Sandi Hulu, W. (2023). Implementasi algoritma K-means clustering untuk pengelompokan penyakit pasien pada Puskesmas Pulo Brayan. *Jurnal Sains dan Teknologi*, 5(1), 150-157.
- Handoko, K. (2016). *Penerapan data mining dalam meningkatkan mutu pembelajaran pada instansi perguruan tinggi menggunakan metode K-means clustering (Studi Kasus di Program Studi TKJ Akademi Komunitas Solok Selatan)*, vol. 02, no. 03.
- Handoyo, R., Mangkudjaja, R. M., & Nasution, S. M. (2014). Perbandingan metode clustering menggunakan metode single linkage dan K-means pada pengelompokan dokumen. *J. SIFO Mikroskil*, 15(2), 73-82. <https://doi.org/10.55601/jsm.v15i2.161>
- Kholil, K. A., Rahaningsih, N., & Dana, R. D. (2024). Penerapan data mining untuk clustering penyakit diare menggunakan algoritma K-means. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 8(3), 3124-3131. <https://doi.org/10.36040/jati.v8i3.9616>
- 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. <https://doi.org/10.37012/jtik.v8i2.1173>
- Mara, M. N., & Satyahadewi, N. (n.d.). *Pengklasifikasian karakteristik dengan metode K-means cluster analysis*.
- Nurhakim, B., Septiani, I., Anam, K., & Pratama, D. (2024). Penarapan algoritma K-means clustering dalam menganalisis resiko penyakit stroke. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 8(1), 318-322. <https://doi.org/10.36040/jati.v8i1.8452>
- Pudjiantoro, T. H., & Renaldi, F. (2011). *Penerapan data mining untuk menganalisa kemungkinan pengunduran diri calon mahasiswa baru*.
- Ramadhan, R. R., & Saprudin, U. (2022). Penerapan Rapidminer menggunakan metode K-means untuk pengelompokan Puskesmas pada cakupan imunisasi dasar (Studi Kasus: Kota Bandung). *Jurnal Teknologi Informatika dan Komputer*, 8(2), 176-187. <https://doi.org/10.37012/jtik.v8i2.1238>

Sihotang, H. (2023). *Metode penelitian kuantitatif*.

Siregar, A. H., Sihotang, D. D., Wijaya, B. A., & Siregar, S. D. (2024). Implementasi algoritma K-means menggunakan RapidMiner untuk klasterisasi data obat pada Rumah Sakit Royal Prima. *Jurnal Teknologi Dan Ilmu Komputer Prima (JUTIKOMP)*, 7(2), 200-211.

Suprihatin, S., Utami, Y. R. W., & Nugroho, D. (2019). *K-means clustering untuk pemetaan daerah rawan demam berdarah*. *J. Teknol. Inf. Dan Komun.TIKomSiN*, 7(1). <https://doi.org/10.30646/tikomsin.v7i1.408>

Upuy, D. (2022). Analysis mortality rate of tuberculosis patients seen from age and length of treatment at RSUD Dr. M. Haulussy Ambon using the K-means clustering algorithm for the Rapidminer application. *Telematika: Jurnal Informatika dan Teknologi Informasi*, 19(3), 337-346. <https://doi.org/10.31315/telematika.v19i3.7709>

Waruwu, M., Puat, S. N., Utami, P. R., Yanti, E., & Rusydiana, M. (2025). Metode penelitian kuantitatif: Konsep, jenis, tahapan dan kelebihan. *Jurnal Ilmiah Profesi Pendidikan*, 10(1), 917-932. <https://doi.org/10.29303/jipp.v10i1.3057>

Widodo, E., & Hadikristanto, W. (2023). Pengelompokan untuk penjualan obat dengan menggunakan algoritma K-means. *Bulletin of Information Technology (BIT)*, 4(3), 408-413. <https://doi.org/10.47065/bit.v4i3.848>

Zulfikar, R., Sari, F. P., Fatmayati, A., Wandini, K., Haryati, T., Jumini, S., ... & Fadilah, H. (2024). *Metode penelitian kuantitatif (Teori, metode dan praktik)*.