Laman web jurnal: https://ejournal.unama.ac.id/index.php/processor

Jurnal Processor

P-ISSN: 1907-6738 | E-ISSN: 2528-0082

Penerapan Algoritma K-Means Clustering untuk Pengelompokkan Penyebaran Diare di Kabupaten Brebes

Laelatul Barokah[1*], Fadiya Olivia Lorenza[2] , Fitri Ayuning Tyas[3]

[1,2,3]Fakultas Sains Teknologi dan Kesehatan, Program Studi Sistem Informasi, Universitas Muhammadiyah Brebes, Jl. Pangeran Diponegoro Grengseng No.184, Grengseng, Taraban, Kecamatan Paguyangan, Kabupaten Brebes, Jawa Tengah 52276, Indonesia.

[*]Penulisan Korespondensi, Email: laelatulbarokah11@gmail.com

Abstrak−Diare merupakan salah satu penyebab utama kematian balita secara global, termasuk di Kabupaten Brebes, Indonesia. Penelitian ini bertujuan untuk menganalisis pola penyebaran diare di Kabupaten Brebes menggunakan algoritma K-Means Clustering dengan pendekatan Knowledge Discovery in Databases (KDD). Data yang digunakan mencakup jumlah kasus per kecamatan dari tahun 2019 hingga 2022. Proses penelitian terdiri dari tahapan Data Selection, Preprocessing, Transformation, Data Mining dan Evaluation/Interpretation menggunakan Davies Bouldin Index (DBI). Analisis dilakukan menggunakan RapidMiner untuk mengelompokkan wilayah ke dalam tiga klaster berdasarkan tingkat penyebaran diare: tinggi, sedang, dan rendah. Hasil penelitian menunjukkan bahwa algoritma K-Means Clustering menghasilkan performa terbaik pada K=3 dengan nilai DBI 0.100, dibandingkan dengan K=5 sebesar 0.119 dan K=7 sebesar 0.128. Nilai DBI yang lebih kecil ini menunjukkan bahwa K=3 merupakan jumlah klaster optimal, dengan kasus diare kategori tinggi terkonsentrasi di Bantarkawung, sehingga menjadi fokus utama dalam upaya penanganan oleh pemerintah daerah. Pendekatan KDD memastikan analisis berjalan secara sistematis, mulai dari seleksi data hingga interpretasi hasil, sehingga kesimpulan yang dihasilkan lebih terstruktur dan relevan. Untuk meningkatkan manfaat penelitian ini, implementasi hasil dengan melibatkan pihak terkait serta integrasi data mining dengan Sistem Informasi Geografis (SIG) dapat membantu memvisualisasikan pola spasial penyebaran diare dan mendukung strategi pencegahan yang lebih optimal di Kabupaten Brebes.

Kata Kunci: Diare; Kabupaten Brebes; Algoritma K-Means Clustering; Data Mining; Davies Bouldin Index (DBI).

Abstract−Diarrhea is one of the leading causes of child mortality worldwide, including in Brebes Regency, Indonesia. This study aims to analyze the pattern of diarrhea distribution in Brebes Regency using the K-Means Clustering algorithm with a Knowledge Discovery in Databases (KDD) approach. The data includes the number of cases per district from 2019 to 2022. The research process consisted of Data Selection, Preprocessing, Transformation, Data Mining and Evaluation/Interpretation using the Davies Bouldin Index (DBI). The analysis was conducted using RapidMiner to classify regions into three clusters based on the level of diarrhea distribution: high, medium and low. The results showed that the K-Means Clustering algorithm achieved the best performance at K=3 with a DBI value of 0.100, compared to K=5 at 0.119 and K=7 at 0.128. This lower DBI value indicated that K=3 was the optimal number of clusters, with high-category diarrhea cases concentrated in Bantarkawung, making it a primary focus for government intervention. The KDD approach ensured that the analysis was carried out systematically, from data selection to result interpretation, leading to well-structured and relevant conclusions. Implementing the findings by involving relevant stakeholders and integrating data mining with Geographic Information Systems (GIS) helps visualize the spatial distribution of diarrhea and supports more effective prevention strategies in Brebes Regency.

Keywords: Diarrhea; Brebes Regency; K-Means Clustering Algorithm; Data Mining; Davies Bouldin Index (DBI).

1. PENDAHULUAN

Kesehatan merupakan indikator utama yang mencerminkan tingkat kesejahteraan suatu masyarakat atau negara, mencakup kondisi fisik, mental dan sosial yang bebas dari penyakit atau gangguan kesehatan, sekaligus menjadi bagian dari hak asasi manusia secara universal tanpa diskriminasi berdasarkan ras, agama, status sosial atau ekonomi [1]. Pemahaman masyarakat terhadap berbagai jenis penyakit dan pengobatannya sangat penting, terutama dalam menghadapi ancaman kesehatan seperti diare. Diare merupakan kondisi di mana frekuensi buang air besar meningkat hingga tiga kali atau lebih dalam sehari, dengan tinja bertekstur encer dan terkadang disertai darah [2]. Gejala penyakit ini, yang ditandai buang air besar encer atau berair, perut kembung, kram perut, mual, muntah, demam dan sakit kepala merupakan gangguan saluran pencernaan yang terjadi akibat lingkungan tidak higienis dan menjadi penyebab kematian kedua tertinggi pada balita di dunia [3]. Peningkatan kesadaran dan upaya pencegahan terhadap ancaman kesehatan seperti diare menjadi langkah penting untuk mengurangi angka

https://doi.org/10.33998/processor.2025.20.1.2050

Submitted: 07 Januari 2025; Reviewed Round-1: 03 Februari 2025; Reviewed Round-2: 04 Maret 2025;
Accepted; 14 Maret 2025; Published: 30 April 2025

kematian dan meningkatkan kualitas hidup masyarakat, terutama mengingat data global yang menunjukkan tingginya angka kematian anak akibat penyakit ini.

Menurut data *World Health Organization* (WHO), diare menempati posisi sebagai penyebab kematian tertinggi kedua pada anak-anak di bawah usia lima tahun, dengan jumlah kematian mencapai sekitar 525.000 setiap tahunnya [4]. Data *United Nations Children's Fund* (UNICEF) November 2024 mengungkapkan bahwa diare menyumbang sekitar 9% kematian anak di bawah lima tahun pada 2021, terutama akibat minimnya akses udara bersih, sanitasi layak dan lingkungan aman [5]. Berdasarkan data Kementerian Kesehatan Republik Indonesia (KEMENKES) tahun 2020, mencatat diare sebagai penyebab utama kematian anak usia 12-59 bulan (42,83%) dan kedua pada usia 29 hari-11 bulan (14,5%) setelah pneumonia [6]. Data publik dari Badan Pusat Statistik (BPS) Kabupaten Brebes mencatat jumlah kasus penyakit diare sebagai penyakit menular selama periode 2019 hingga 2022 mengalami fluktuasi, dengan 29.108 kasus pada tahun 2019, sedikit menurun menjadi 28.776 kasus pada tahun 2020 dan 2021, lalu turun signifikan menjadi 9.520 kasus pada tahun 2022 [7]. Kondisi ini menuntut Kabupaten Brebes untuk mengambil langkah penanganan efektif guna menurunkan angka penyebaran diare, salah satunya melalui pengelompokan wilayah berdasarkan tingkat penyebaran tertinggi. Langkah penanganan yang efektif pada kasus penyebaran diare dapat dilakukan dengan analisis *data mining* menggunakan metode *Clustering*. *Data mining* dapat menjadi salah satu pencegahan yang efektif untuk menekan angka kasus diare di Kabupaten Brebes. *Data mining* mulai dikenal luas sejak tahun 1990-an, seiring dengan meningkatnya pemanfaatan data di berbagai bidang, seperti pemasaran, bisnis, sains, teknik, seni dan hiburan [8]. *Data Mining* adalah proses untuk menemukan informasi baru dengan mencari pola atau aturan tertentu dari kumpulan data yang besar, sehingga dapat menggali pengetahuan yang sebelumnya tidak diketahui secara manual [9]. Untuk menerapkan teknik *clustering*, penulis dapat mengidentifikasi kelompok-kelompok berisiko tinggi dan merancang intervensi yang lebih tepat sasaran untuk mengurangi penyebaran penyakit diare di masyarakat.

Clustering adalah metode pengelompokkan data berdasarkan kondisi tertentu, di mana kualitas kelompok diukur dari kemiripan tinggi antar objek dalam satu kelompok dan perbedaan yang signifikan antar kelompok lainnya [10]. Algoritma yang dapat digunakan dalam penelitian ini adalah algoritma *K-Means Clustering.* Algoritma *K-Means Clustering* merupakan metode *partitional* yang menggunakan nilai centroid awal untuk menentukan jumlah awal kelompok, sehingga algoritma ini menggunakan jumlah cluster sebagai masukan dan menghasilkan centroid akhir sebagai output, proses dimulai dengan pemilihan centroid awal secara acak, sementara jumlah iterasi bergantung pada pemilihan tersebut [11]. Mengingat efektivitas dalam pengelompokan data, penelitian ini menerapkan algoritma *K-Means.*

Berbagai penelitian mendukung efektivitas penggunaan algoritma *K-Means* dalam *clustering.* Penulis meninjau beberapa penelitian, yaitu pertama menurut (Sari et al., 2021), hasilnya menunjukkan keberhasilan *K-Means* dalam mengidentifikasi wilayah yang memerlukan perhatian lebih dalam penanganan kasus HIV/AIDS [12]. Penelitian kedua yang dilakukan oleh (Syamfithriani et al., 2023), hasil penelitiannya yaitu merekomendasikan pemetaan daerah prioritas penanganan diare dengan algoritma *K-Means* sebagai masukan bagi Dinas Kesehatan Kabupaten Kuningan untuk merumuskan strategi pencegahan dan penanggulangan diare pada balita, dengan penentuan jumlah *cluster* optimum menggunakan metode *Elbow* dan *Silhouette Coefficient* [6]. Penelitian ketiga dilakukan oleh (Qirom et al,. 2024), hasil Penelitian ini mengidentifikasi empat kelompok pasien hipertensi di Puskesmas Rajapolah menggunakan algoritma *K-Means clustering* dengan evaluasi *Davies Bouldin Index* (DBI), dimana kelompok dengan hipertensi tertinggi berada di *cluster* 3, yang terdiri dari 68 pasien berusia 30 hingga 74 tahun dengan stadium 2 hingga krisis hipertensi [13]. Penelitian keempat dilakukan oleh **(**Dikarya & Muharni, 2022), hasil penelitian ini menerapkan algoritma *K-Means* untuk mengelompokkan universitas terbaik menjadi tiga *cluster* berdasarkan atribut *world rank, institution, country* dan *score*, dengan Harvard University berada di *cluster* 2, University of Haifa di *cluster* 1 dan National Chung Cheng University di *cluster* 0, yang memudahkan identifikasi universitas terbaik secara efisien [14]. Penelitian kelima menurut (Indra, 2023), hasil penelitian ini adalah membandingkan metode *K-Means* dan *Hierarchical Clustering* dalam klasterisasi daerah stunting di Indonesia, dengan hasil menunjukkan *K-Means* menghasilkan klaster yang lebih baik berdasarkan nilai *Silhouette Coefficient* 0.48 dan *Calinski-Harabasz index* 10.49, membentuk dua klaster [15]. Penelitian keenam menurut (Setiaji et al,. 2024), hasil penelitian ini menggunakan metode *K-Means* dan *K-Medoids* untuk mengelompokkan harga beras medium di Jawa Tengah, dengan hasil menunjukkan K-Means dengan K=3 dan K=5 memiliki validitas DBI yang lebih unggul dibandingkan *K-Medoids* [16]. Penelitian ketujuh dilakukan oleh (Yolanda & Suhardi, 2023), hasil penelitian ini menggunakan algoritma *K-Means clustering* untuk mengelompokkan data pasien rehabilitasi narkoba di Badan Narkotika Nasional Provinsi Sumatera Utara, menghasilkan tiga kelompok yang memungkinkan penentuan program rehabilitasi yang lebih sesuai dengan kebutuhan masing-masing kelompok, termasuk penerapan program parenting pada kelompok kedua [17]. Penelitian kedelapan dilakukan oleh (Purba et al,. 2021), hasil penelitian ini mengkaji penggunaan Algoritma *K-Means Cluster Analysis* untuk mengelompokkan data Infeksi Saluran Pernafasan Akut di Provinsi

https://doi.org/10.33998/processor.2025.20.1.2050
Submitted: 07 Januari 2025; Reviewed Round-1: 03 Februari 2025; Reviewed Round-2: 04 Maret 2025;
Accepted: 14 Maret 2025; Published: 30 April 2025

Riau berdasarkan variabel per kotamadya, menunjukkan bahwa algoritma ini menghasilkan jumlah keanggotaan *cluster* yang konsisten meskipun titik awal pusat *cluster* berbeda [18]. Penelitian kesembilan dilakukan oleh (Nurfidah et al,. 2024) tentang analisis gempa bumi di Indonesia menggunakan berbagai algoritma *clustering*, studi ini secara komprehensif membandingkan *K-Medoids, K-Means, DBSCAN, Fuzzy C-Means dan K-Affinity Propagation (K-AP)* dengan metrik *Silhouette Score* dan *Cluster Purity* untuk mengidentifikasi algoritma terbaik dalam analisis zona seismik, hasil dari penelitian ini *K-Means* terbukti paling efektif [19]. Penelitian kesepuluh dilakukan oleh (Asyahri et al,. 2022) menerapkan algoritma *K-Means Clustering* untuk menganalisis transaksi penjualan kosmetik guna mengelola stok secara efisien, hasil klasterisasi membantu mengidentifikasi produk kurang laku untuk mencegah penumpukan serta menentukan produk yang harus tersedia, sehingga meningkatkan strategi bisnis dan keuntungan penjualan [20]. Penelitian kesebelas dilakukan oleh (Nurahman et al,. 2022) menerapkan algoritma *K-Means* untuk mengklaster sekolah di Kabupaten Seruyan, menunjukkan 178 sekolah di Cluster 0,3 di Cluster 1 dan 43 di Cluster 2 sebagai yang terendah, dengan evaluasi *Davies Bouldin Index* bernilai -0.695, sehingga direkomendasikan pendampingan dan pengadaan fasilitas bagi sekolah di *Cluster* 2 [21]. Penelitian kedua belas dilakukan oleh (Mawarni et al,. 2022) menerapkan algoritma *K-Means* untuk mengklasifikasikan kedisiplinan siswa berdasarkan absensi, kerapian dan perilaku, dengan analisis terhadap 133 siswa yang menghasilkan tiga *cluster*, yaitu 41 siswa di *Cluster* 1 dan 33 siswa di *Cluster* 2 dan 59 siswa di *Cluster* 3, sehingga algoritma *K-Means* terbukti efektif dalam menilai serta mencegah perilaku indisipliner [22]. Penelitian ketiga belas dilakukan oleh (Rohman and Wibowo, 2024) membandingkan algoritma *K-Means* dan *K-Medoids* untuk segmentasi pelanggan di pusat perbelanjaan dengan metode *elbow* yang menentukan lima *cluster* untuk *K-Means* dan empat *cluster* untuk *K-Medoids*, sementara evaluasi menggunakan *Silhouette Coefficient* menunjukkan bahwa *K-Means* lebih optimal dengan skor 0.553 dibandingkan *K-Medoids* yang hanya 0.485, sehingga *K-Means* menjadi metode terbaik dalam pengelompokan pelanggan [23]. Penelitian keempat belas oleh (Riza et al,. 2021) menerapkan *K-Means* untuk mengelompokkan soal ujian berdasarkan tingkat kesulitan guna memastikan keseimbangan setiap set soal, sementara evaluasi menggunakan statistik deskriptif dan ANOVA menunjukkan bahwa *K-Means* efektif dalam menghasilkan paket soal otomatis yang konsisten dan adil [24]. Penelitian kelima belas oleh (Ariska et al,. 2024) menerapkan *K-Means* dan *K-Medoids* untuk mengelompokkan data UMKM di Kota Pagar Alam, di mana hasil evaluasi menggunakan *Davies Bouldin Index* menunjukkan bahwa *K-Means* memiliki performa lebih baik dengan nilai yang lebih mendekati 0 dibandingkan *K-Medoids*, terbukti pada tahun 2020 dengan DBI *K-Means* sebesar 0.134 dan *K-Medoids* 0.523, serta pada tahun 2022 dengan DBI *K-Means* 0.277 dan *K-Medoids* 0.496 [25].

Penelitian terdahulu telah berhasil menerapkan algoritma *K-Means* dalam klasterisasi berbagai penyakit seperti HIV/AIDS, hipertensi, stunting dan ISPA, namun belum banyak yang mengimplementasikan algoritma ini untuk menganalisis penyebaran diare di Kabupaten Brebes, apalagi dengan menggunakan metode *Knowledge Discovery in Databases* (KDD) sebagai kerangka analisis, yang mencakup tahapan *Data Selection, Preprocessing, Transformation, Data Mining* dan *Evaluation/Interpretation* [10]. Penelitian ini bertujuan untuk mengisi celah tersebut dengan mengaplikasikan *K-Means Clustering* untuk mengelompokkan wilayah di Kabupaten Brebes berdasarkan tingkat penyebaran diare serta mengidentifikasi kecamatan yang membutuhkan perhatian lebih dalam penanganan penyakit tersebut. Harapannya, penelitian ini dapat menghasilkan *cluster* yang akurat, sehingga pemahaman mengenai distribusi diare di Kabupaten Brebes dapat ditingkatkan. Penanganan diare di daerah dengan penyebaran tinggi dapat menjadi lebih tepat sasaran, sekaligus memberikan wawasan mengenai penerapan *K-Means Clustering* untuk intervensi yang lebih efektif.

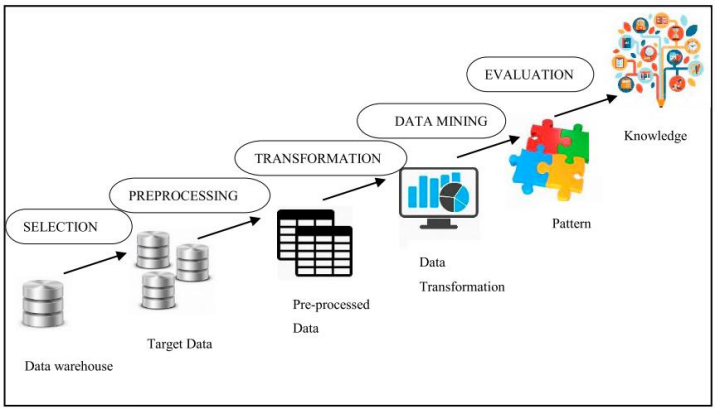
Berdasarkan latar belakang masalah di atas, penelitian ini bertujuan untuk mengelompokkan data penyebaran diare di Kabupaten Brebes menggunakan algoritma *K-Means*, dimana dominasi *cluster* akan mengidentifikasi wilayah dengan jumlah kasus tertinggi berdasarkan per kecamatan. Proses penentuan *cluster* akan mengikuti evaluasi akhir menggunakan *Davies Bouldin Index* (DBI), yang digunakan untuk menentukan jumlah *cluster* terbaik. Aplikasi *RapidMiner* akan digunakan sebagai alat untuk melakukan analisis data yang diperlukan. Diharapkan, penelitian ini dapat memberikan gambaran yang lebih jelas mengenai pola penyebaran diare di Kabupaten Brebes dan berkontribusi pada perencanaan strategi pencegahan yang lebih tepat di Kabupaten Brebes.

2. METODOLOGI PENELITIAN

Penelitian ini menggunakan metode kuantitatif dengan menerapkan tahapan berdasarkan pendekatan *Knowledge Discovery in Databases* (KDD). *Knowledge Discovery in Databases* (KDD) adalah pendekatan praktis dalam proses *data mining* yang bertujuan menemukan dan mengidentifikasi pola dalam data, dengan memastikan pola tersebut memiliki validitas, relevansi dan mudah dipahami [13]. Proses KDD terdiri dari lima tahapan utama, yaitu *Data Selection, Preprocessing, Transformation, Data Mining* dan *Evaluation/Interpretation* [10], yang

https://doi.org/10.33998/processor.2025.20.1.2050
Submitted: 07 Januari 2025; Reviewed Round-1: 03 Februari 2025; Reviewed Round-2: 04 Maret 2025;
Accepted: 14 Maret 2025; Published: 30 April 2025

menjadi kerangka acuan dalam penelitian ini. Setiap tahapan dalam metode KDD dirancang sebagai langkah sistematis untuk mengelola dan menganalisis data secara mendalam, seperti yang diilustrasikan pada gambar 1.



Gambar 1 Tahapan metode penelitian

2.1. Data Selection

Proses ini melibatkan pemilihan data jumlah kasus diare dari publikasi BPS Kabupaten Brebes yang dapat diakses di tautan https://brebeskab.bps.go.id [7]. Data tahun 2019 hingga 2022 tentang kasus diare digunakan untuk menganalisis pola penyebaran penyakit, dengan seluruh informasi diformat dalam file Excel untuk mempermudah analisis.

2.2. Preprocessing

Pada tahap ini, data yang dipilih dibersihkan dari ketidakkonsistenan, data kosong atau nilai yang tidak relevan. Langkah *preprocessing* meliputi pengisian *missing values*, penghapusan duplikasi data dan penyelarasan format data untuk memastikan kualitas data yang siap diolah dalam perhitungan algoritma *K-Means clustering*.

2.3. Transformasi data

Transformasi data dilakukan untuk mengubah data mentah menjadi format yang lebih sesuai untuk proses *data mining*. Tahapan ini melibatkan normalisasi data, pengelompokan atribut atau pembuatan variabel baru yang mendukung analisis.

2.4. Data Mining

Tahap inti dari KDD ini melibatkan penerapan algoritma *K-Means Clustering* untuk mengelompokkan data penyebaran penyakit diare di Kabupaten Brebes. Algoritma *K-Means Clustering* adalah algoritma yang digunakan untuk mengelompokkan data dengan membaginya ke dalam beberapa kelompok berbeda, algoritma ini bekerja dengan mengurangi jarak antara data dan pusat kelompoknya, sehingga setiap data berada dalam kelompok yang paling sesuai [9]. Tahapan dalam penerapan algoritma *K-Means Clustering* mencakup langkah-langkah berikut:

a. Menentukan jumlah *cluster* yang akan dibentuk sebagai dasar pengelompokan data.
b. Menetapkan *centroid* awal untuk setiap *cluster* secara acak.
c. Mengukur jarak antara setiap data observasi dengan *centroid* yang telah ditentukan menggunakan rumus Euclidean Distance, sebagaimana ditunjukkan pada Persamaan.

$$d(x,c) = \sqrt{\sum_{i=1}^{n}(x_i - c_i)^2} \quad (1)$$

d. Mengelompokkan seluruh data observasi berdasarkan kedekatannya dengan *centroid*, yaitu jarak terkecil antara data dan *centroid*.
e. Memperbarui nilai *centroid* dengan menghitung rata-rata (*mean*) dari data dalam masing-masing *cluster yang telah terbentuk.*

https://doi.org/10.33998/processor.2025.20.1.2050
Submitted: 07 Januari 2025; Reviewed Round-1: 03 Februari 2025; Reviewed Round-2: 04 Maret 2025;
Accepted: 14 Maret 2025; Published: 30 April 2025

f. Mengulangi proses mulai dari langkah 3 hingga langkah 5 sampai komposisi anggota dalam setiap *cluster* tidak mengalami perubahan.

Algoritma *K-Means Clustering* dipilih karena efisiensinya dalam mengelompokkan data berdasarkan kesamaan fitur dan kinerjanya yang cepat untuk memproses data besar, dengan proses *clustering* tiga *cluster* untuk mengkategorikan tingkat kasus menjadi tinggi, sedang dan rendah berdasarkan atribut jumlah kasus serta kecamatan di Kabupaten Brebes, sehingga memberikan wawasan mendalam tentang distribusi penyebaran penyakit diare di wilayah tersebut.

2.5. Evaluation/Interpretation

Tahap terakhir ini digunakan untuk mengevaluasi kualitas hasil *clustering* menggunakan *Davies Bouldin Index* (DBI) sebagai metrik evaluasi. *Davies Bouldin index* (DBI) mengevaluasi hasil *clustering* berdasarkan kohesi, yang mengukur kedekatan data dengan *centroid* dan separasi, yang menghitung jarak antar *centroid*, dengan nilai DBI terkecil menunjukkan hasil *clustering* yang paling optimum [15]. Nilai *Davies Bouldin Index* (DBI) dihitung berdasarkan rasio yang diperoleh sebelumnya menggunakan persamaan berikut:

$$DBI = \frac{1}{k}\sum_{i=1}^{k} max_{i \neq j}\,(R_{ij}) \quad (2)$$

Hasil yang dievaluasi kemudian diinterpretasikan untuk memberikan rekomendasi strategi pencegahan dan penanganan penyakit secara lebih efektif.

3. HASIL DAN PEMBAHASAN

Penelitian ini akan menghasilkan pengelompokan kasus diare di Kabupaten Brebes berdasarkan atribut jumlah kasus pada tahun 2019 hingga 2022 dan kecamatan di wilayah tersebut. Proses ini dilakukan melalui KDD menggunakan algoritma *K-Means Clustering* yang diterapkan dengan bantuan tools *RapidMiner*.

3.1 Hasil Perancangan

Berikut adalah hasil yang diperoleh berdasarkan tahapan perancangan dengan pendekatan *Knowledge Discovery in Databases* (KDD).

3.1.1 Data Selection

Proses pemilihan data dalam KDD mencakup identifikasi dan ekstraksi subset data yang relevan dari kumpulan data yang lebih besar. Tujuan utama tahap ini adalah memfokuskan pada bagian data yang berpotensi mengandung pola, informasi atau hubungan yang signifikan. Atribut yang digunakan meliputi jumlah kasus diare dan kecamatan. Sebelum tahap *preprocessing data* dengan perangkat lunak *RapidMiner*, data bertipe nominal diubah terlebih dahulu menjadi format numerik menggunakan MS Excel. Berikut adalah tabel data set penyebaran diare di Kabupaten Brebes:

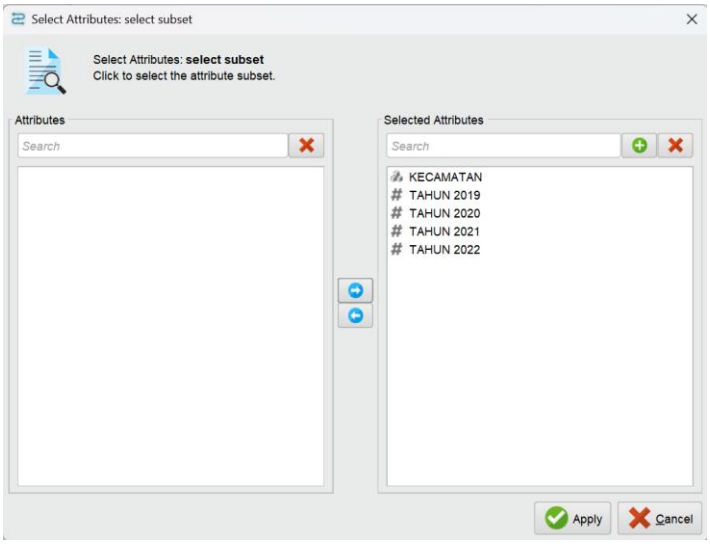
Tabel 1. Data set penyebaran diare di Kabupaten Brebes

Kecamatan	Tahun 2019	Tahun 2020	Tahun 2021	Tahun 2022
Salem	1195	863	863	264
Bantarkawung	4125	4125	4125	289
Bumiayu	2298	2298	2298	532
Paguyangan	1928	1928	1928	703
Sirampog	2427	2427	2427	192
Tonjong	1601	161	1601	895
Larangan	1311	1311	1311	496

https://doi.org/10.33998/processor.2025.20.1.2050
Submitted: 07 Januari 2025; Reviewed Round-1: 03 Februari 2025; Reviewed Round-2: 04 Maret 2025;
Accepted: 14 Maret 2025; Published: 30 April 2025

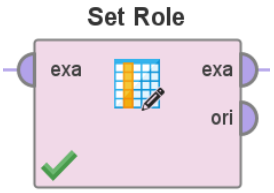
Ketanggungan	1176	1176	1176	1055
Banjarharjo	967	967	967	718
Losari	1844	1844	1844	901
Tanjung	765	765	765	477
Kersana	508	508	508	76
Bulakamba	2266	2266	2266	433
Wanasari	1914	1914	1914	713
Songgom	657	657	657	78
Jatibarang	2342	2342	2342	970
Brebes	1784	1784	1784	728

Gambar 2 menunjukkan tahap pemilihan atribut yang akan digunakan. Berikut adalah gambar dari proses pemilihan atribut *select attribute*:



Gambar 2. Hasil *select attribute*

Gambar 3 menunjukkan tahap pemilihan salah satu atribut untuk dijadikan sebagai ID. Pada tahap ini, atribut Nama "kecamatan" dikategorikan sebagai ID dengan menggunakan bantuan atribut *set role*:



Gambar 3. *Set role*

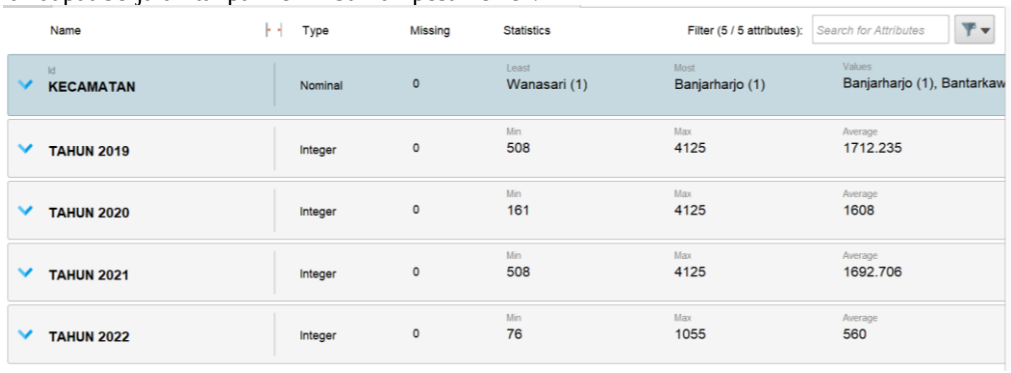
Role	Name	Type	Range	Missings	Comment
	TAHUN 2019	# integer	=[508 – 4125]	= 0	
	TAHUN 2020	# integer	=[161 – 4125]	= 0	
	TAHUN 2021	# integer	=[508 – 4125]	= 0	
	TAHUN 2022	# integer	=[76 – 1055]	= 0	
id	KECAMATAN	nominal	=[Banjarharj...	= 0	

Gambar 4. Hasil set role

https://doi.org/10.33998/processor.2025.20.1.2050
Submitted: 07 Januari 2025; Reviewed Round-1: 03 Februari 2025; Reviewed Round-2: 04 Maret 2025;
Accepted: 14 Maret 2025; Published: 30 April 2025

3.1.2 Preprocessing

Tahap ini dilakukan pemeriksaan terhadap nilai *missing value* untuk memastikan bahwa proses *clustering* di *RapidMiner* dapat berjalan tanpa menimbulkan pesan error:



Name	Type	Missing	Statistics		Filter (5 / 5 attributes): Search for Attributes
Id KECAMATAN	Nominal	0	Least Wanasari (1)	Most Banjarharjo (1)	Values Banjarharjo (1), Bantarkaw
TAHUN 2019	Integer	0	Min 508	Max 4125	Average 1712.235
TAHUN 2020	Integer	0	Min 161	Max 4125	Average 1608
TAHUN 2021	Integer	0	Min 508	Max 4125	Average 1692.706
TAHUN 2022	Integer	0	Min 76	Max 1055	Average 560

Gambar 5. Hasil cek *missing value*

Gambar 5 menunjukkan bahwa kolom *missing* pada setiap atribut bernilai 0, menandakan tidak adanya *missing value*, sehingga data siap diproses ke tahap selanjutnya, yaitu tahap transformasi.

3.1.3 Transformasi data

Pada tahap ini, atribut kecamatan yang bertipe *polinomial* ditambahkan sebagai *ID* menggunakan *change role*, sedangkan atribut jumlah penderita diare per tahun tetap dalam bentuk nilai *integer* untuk memastikan konsistensi dan kesesuaian dalam proses analisis data.

Row No.	KECAMATAN	TAHUN 2019	TAHUN 2020	TAHUN 2021	TAHUN 2022
1	Salem	1195	863	863	264
2	Bantarkawung	4125	4125	4125	289
3	Bumiayu	2298	2298	2298	532
4	Paguyangan	1928	1928	1928	703
5	Sirampog	2427	2427	2427	192
6	Tonjong	1601	161	1601	895
7	Larangan	1311	1311	1311	496
8	Ketanggungan	1176	1176	1176	1055
9	Banjarharjo	967	967	967	718
10	Losari	1844	1844	1844	901
11	Tanjung	765	765	765	477
12	Kersana	508	508	508	76
13	Bulakamba	2266	2266	2266	433
14	Wanasari	1914	1914	1914	713
15	Songgom	657	657	657	78
16	Jatibarang	2342	2342	2342	970
17	Brebes	1784	1784	1784	728

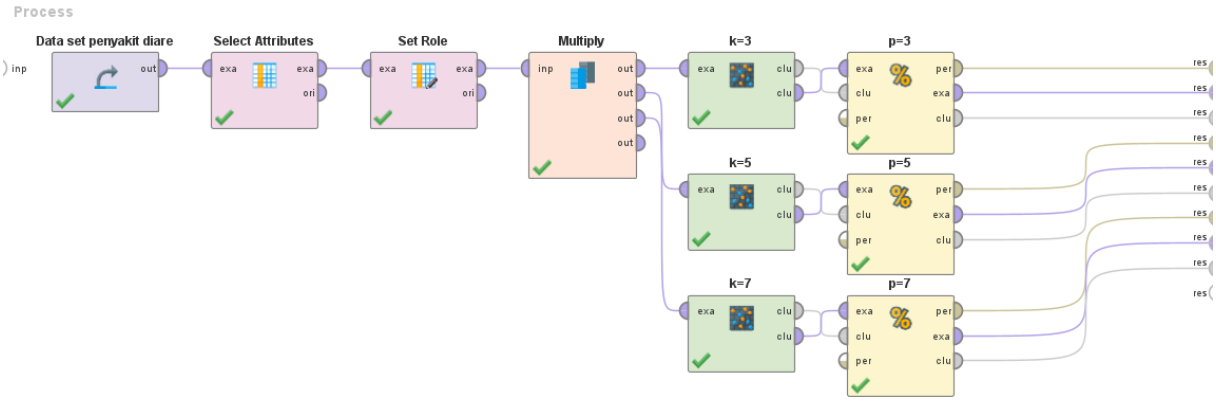
ExampleSet (17 examples,1 special attribute,4 regular attributes)

Gambar 6. Transformasi data atribut kecamatan ditambah ID

3.1.4 Data Mining

https://doi.org/10.33998/processor.2025.20.1.2050
Submitted: 07 Januari 2025; Reviewed Round-1: 03 Februari 2025; Reviewed Round-2: 04 Maret 2025;
Accepted: 14 Maret 2025; Published: 30 April 2025

Gambar 7 menunjukkan langkah pengelompokkan data kasus diare menggunakan algoritma *K-Means Clustering* dengan tiga *cluster*. Pemilihan jumlah *cluster* ini bertujuan untuk mengelompokkan tingkat kasus diare ke dalam kategori tinggi, sedang dan rendah pada aplikasi *RapidMiner*.



Gambar 7. Data mining menggunakan algoritma *k-means clustering*

Gambar 8 menampilkan hasil clustering dengan algoritma *K-Means*, di mana 17 data terbagi ke dalam tiga klaster. Klaster 0 berisi 8 item, klaster 1 terdiri dari 8 item, dan klaster 2 terdiri dari 1 item. Penomoran klaster dimulai dari 0 karena dalam bahasa pemrograman, angka 0 merupakan indeks pertama dalam urutan numerik.

Cluster Model

```
Cluster 0: 8 items
Cluster 1: 8 items
Cluster 2: 1 items
Total number of items: 17
```

Gambar 8. Hasil *clustering* K=3

Gambar 9 menampilkan hasil clustering dengan algoritma *K-Means*, di mana 17 data terbagi ke dalam lima klaster. Klaster 0 berisi 3 item, klaster 1 terdiri dari 1 item, klaster 2 terdiri dari 8 item, klaster 3 terdiri dari 4 item dan klaster 4 terdiri dari 1 item. Penomoran klaster dimulai dari 0 karena dalam bahasa pemrograman, angka 0 merupakan indeks pertama dalam urutan numerik.

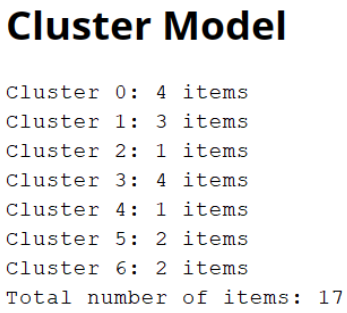
Cluster Model

```
Cluster 0: 3 items
Cluster 1: 1 items
Cluster 2: 8 items
Cluster 3: 4 items
Cluster 4: 1 items
Total number of items: 17
```

Gambar 9. Hasil *clustering* K=5

Gambar 10 menampilkan hasil *clustering* dengan algoritma *K-Means*, di mana 17 data terbagi ke dalam tujuh klaster. Klaster 0 berisi 4 item, klaster 1 terdiri dari 3 item, klaster 2 terdiri dari 1 item, klaster 3 terdiri dari 4 item, klaster 4 terdiri dari 1 item dan klaster 5 terdiri dari 2 item. Penomoran klaster dimulai dari 0 karena dalam bahasa pemrograman, angka 0 merupakan indeks pertama dalam urutan numerik.

https://doi.org/10.33998/processor.2025.20.1.2050
Submitted: 07 Januari 2025; Reviewed Round-1: 03 Februari 2025; Reviewed Round-2: 04 Maret 2025;
Accepted: 14 Maret 2025; Published: 30 April 2025



Gambar 10. Hasil *clustering* K=7

3.1.5 Evaluation/Interpretation

Tahap akhir setelah proses *data mining* di aplikasi *RapidMiner* adalah menekan tombol *run*, yang kemudian akan menampilkan tab hasil *(result)* berisi pengelompokkan meliputi jumlah kasus diare pada tahun 2019 hingga 2020 dan kecamatan dengan nilai K=3.

Row No.	KECAMATAN	cluster	TAHUN 2019	TAHUN 2020	TAHUN 2021	TAHUN 2022
1	Salem	cluster_1	1195	863	863	264
2	Bantarkawung	cluster_2	4125	4125	4125	289
3	Bumiayu	cluster_0	2298	2298	2298	532
4	Paguyangan	cluster_0	1928	1928	1928	703
5	Sirampog	cluster_0	2427	2427	2427	192
6	Tonjong	cluster_1	1601	161	1601	895
7	Larangan	cluster_1	1311	1311	1311	496
8	Ketanggungan	cluster_1	1176	1176	1176	1055
9	Banjarharjo	cluster_1	967	967	967	718
10	Losari	cluster_0	1844	1844	1844	901
11	Tanjung	cluster_1	765	765	765	477
12	Kersana	cluster_1	508	508	508	76
13	Bulakamba	cluster_0	2266	2266	2266	433
14	Wanasari	cluster_0	1914	1914	1914	713
15	Songgom	cluster_1	657	657	657	78
16	Jatibarang	cluster_0	2342	2342	2342	970
17	Brebes	cluster_0	1784	1784	1784	728

ExampleSet (17 examples,2 special attributes,4 regular attributes)

Gambar 11. Hasil evaluasi *cluster* dataset kasus penyakit diare K=3

Gambar 11 menampilkan hasil evaluasi cluster dataset kasus penyakit diare berdasarkan K = 3, di mana *cluster_0* mencakup Kecamatan Bumiayu, Paguyangan, Sirampog, Losari, Bulakamba, Wanasari, Jatibarang serta Brebes, sedangkan *cluster_1* terdiri atas Kecamatan Salem, Tonjong, Larangan, Ketanggungan, Banjarharjo, Tanjung, Kersana maupun Songgom, sementara itu *cluster_2* hanya meliputi Kecamatan Bantarkawung.

3.2 Hasil Pengelompokan

Berikut adalah hasil pengelompokan *K-Means Clustering* yang diperoleh berdasarkan tahapan perancangan dengan pendekatan *Knowledge Discovery in Databases* (KDD).

3.2.1 Clustering Penyebaran Penyakit Diare

https://doi.org/10.33998/processor.2025.20.1.2050
Submitted: 07 Januari 2025; Reviewed Round-1: 03 Februari 2025; Reviewed Round-2: 04 Maret 2025;
Accepted: 14 Maret 2025; Published: 30 April 2025

Berikut adalah hasil dari penerapan algoritma *K-Means Clustering* dalam pengelompokan penyebaran penyakit diare per kecamatan di Kabupaten Brebes. Setelah proses dijalankan, diperoleh nilai Davies-Bouldin Index (DBI) sebagai berikut:

Tabel 2. Hasil cluster penyebaran penyakit diare per kecamatan di Kabupaten Brebes [1]

Cluster/k	**DBI**
K=3	0.100
K=5	0.119
K=7	0.128

Berdasarkan tabel 2, menunjukkan bahwa pengelompokan penyebaran kasus diare di Kabupaten Brebes berdasarkan jumlah per kecamatan memiliki nilai terkecil pada K=3 dengan Davies-Bouldin Index (DBI) sebesar 0.100, yang mengindikasikan kualitas klasterisasi terbaik dibandingkan dengan K=5 dan K=7 sesuai dengan tabel di atas.

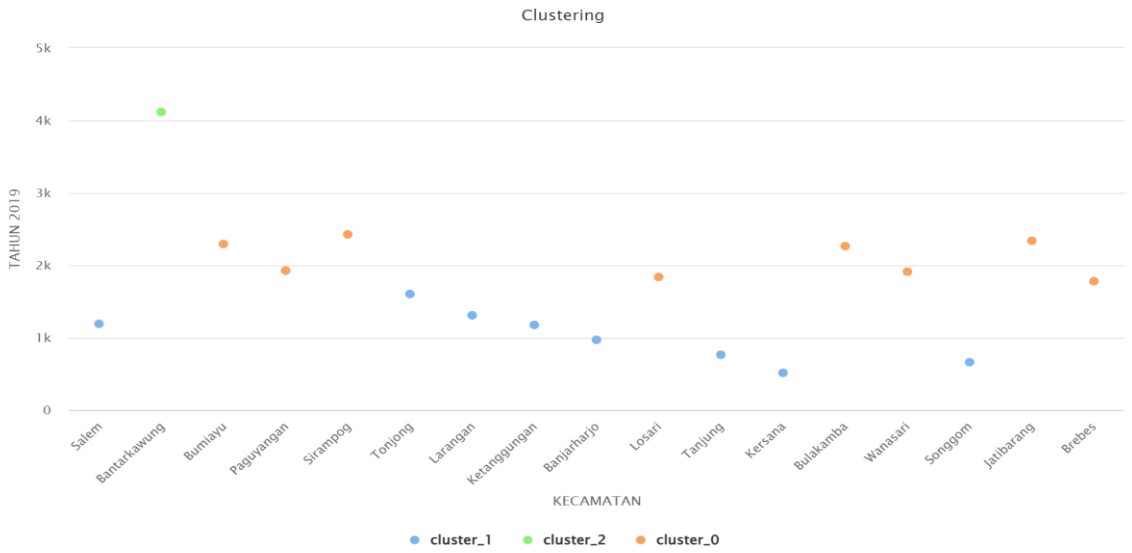
PerformanceVector

```
PerformanceVector:
Avg. within centroid distance: 81896.551
Avg. within centroid distance_cluster_0: 57325.426
Avg. within centroid distance_cluster_1: 116704.746
Avg. within centroid distance_cluster_2: 0.000
Davies Bouldin: 0.100
```

Gambar 12. Hasil *performance vector* K=3

Gambar 12 menampilkan nilai *Davies Bouldin Index* (DBI) sebesar 0.100 dengan metode *K-Means Clustering* pada K=3, yang menunjukkan hasil *PerformanceVector* dalam pengelompokan penyebaran penyakit diare per kecamatan di Kabupaten Brebes.

Hasil *cluster* selanjutnya dapat dilihat pada bagian *Visualizations* di *RapidMiner*.



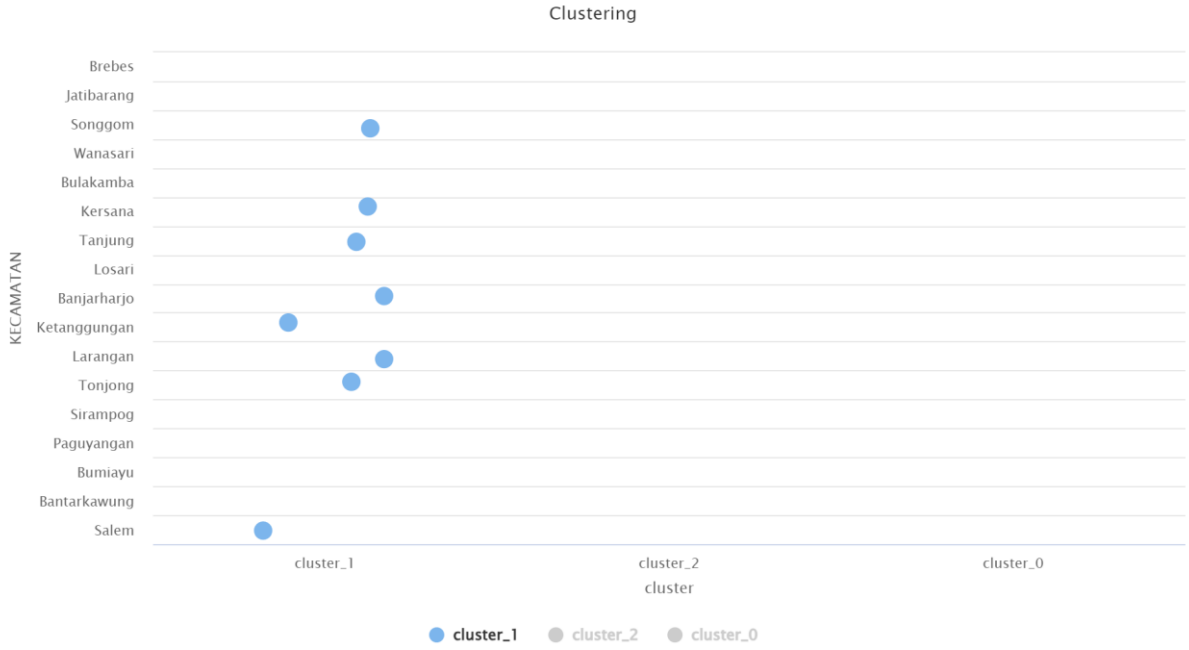
Gambar 13. Hasil *Visualizations* dari Penyebaran diare dalam kategori tinggi, sedang dan rendah

Gambar 13, berdasarkan data penyebaran penyakit diare di Kabupaten Brebes terbagi menjadi tiga *cluster*, yaitu *Cluster_0* yang ditandai dengan warna oranye termasuk kategori sedang, *Cluster_1* yang berwarna biru berada dalam kategori rendah dan *Cluster_2* yang berwarna hijau masuk kategori tinggi.

https://doi.org/10.33998/processor.2025.20.1.2050
Submitted: 07 Januari 2025; Reviewed Round-1: 03 Februari 2025; Reviewed Round-2: 04 Maret 2025;
Accepted: 14 Maret 2025; Published: 30 April 2025

Gambar 14. Hasil *Visualizations* dari Penyebaran diare dalam kategori tinggi

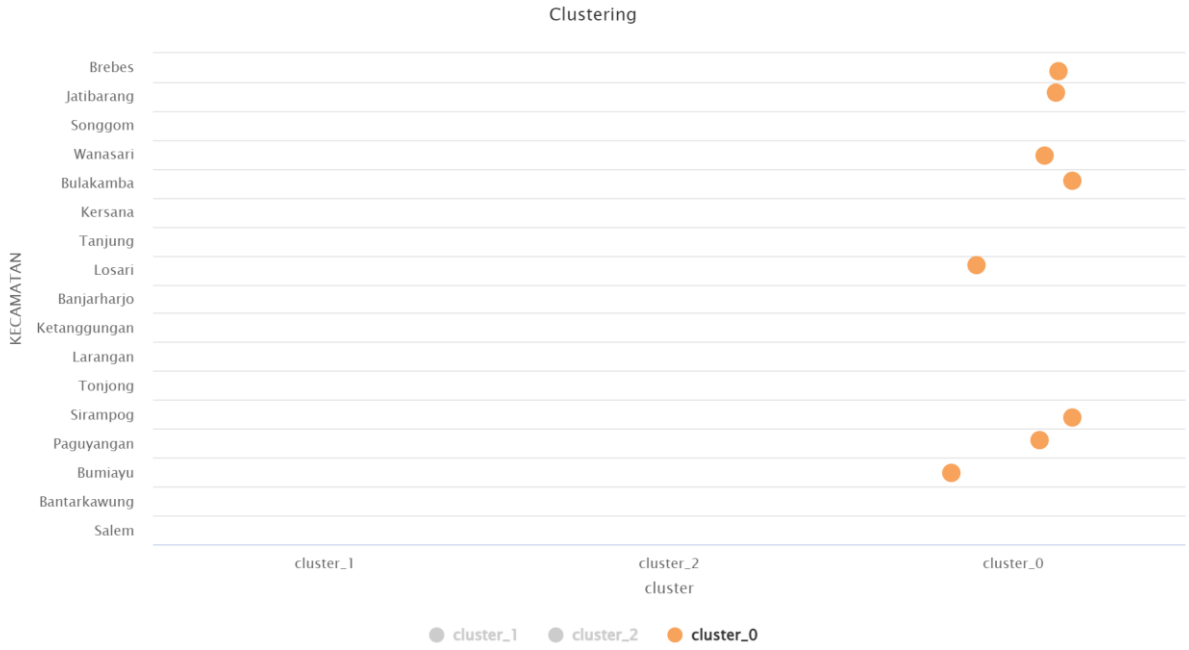
Gambar 14 menunjukkan visualisasi penyebaran kasus diare dalam kategori tinggi di Kabupaten Brebes, di mana kecamatan dengan nilai tertinggi terdapat di *Cluster_2*, yaitu Bantarkawung, dengan jumlah kasus mencapai 4125 dari tahun 2019 hingga 2021 sebelum menurun drastis menjadi 289 pada tahun 2022.



Gambar 15. Hasil *Visualizations* dari Penyebaran diare dalam kategori rendah

Gambar 15 menampilkan hasil visualisasi penyebaran kasus diare dalam kategori rendah di Kabupaten Brebes, di mana kecamatan dengan nilai terendah berada di *Cluster_1*, yaitu Salem, Tonjong, Larangan, Ketanggungan, Banjarharjo, Tanjung, Kersana dan Songgom, dengan jumlah kasus yang selalu di bawah 1601 sejak tahun 2019 dan terus mengalami penurunan hingga tahun 2022 dengan kisaran 264 hingga 718.

https://doi.org/10.33998/processor.2025.20.1.2050
Submitted: 07 Januari 2025; Reviewed Round-1: 03 Februari 2025; Reviewed Round-2: 04 Maret 2025;
Accepted: 14 Maret 2025; Published: 30 April 2025



Gambar 16. Hasil *Visualizations* dari Penyebaran diare dalam kategori sedang

Gambar 16 menampilkan hasil visualisasi penyebaran kasus diare dalam kategori sedang di Kabupaten Brebes, di mana kecamatan yang tergolong dalam

Cluster_0 meliputi Bumiayu, Paguyangan, Sirampog, Losari, Bulakamba, Wanasari, Jatibarang, dan Brebes, dengan jumlah kasus berkisar antara 1784 hingga 2427 pada tahun 2019, yang kemudian mengalami penurunan pada tahun 2022 dengan kisaran 433 hingga 901.

Hasil penelitian menunjukkan bahwa algoritma *K-Means Clustering* berhasil mengelompokkan kasus diare di Kabupaten Brebes ke dalam tiga cluster berdasarkan karakteristik serupa. *Cluster_2* memiliki jumlah kasus tertinggi dengan Kecamatan Bantarkawung mencapai 4125 kasus pada 2019 hingga 2021 sebelum menurun menjadi 289 pada 2022, sementara *Cluster_0* berada pada kategori sedang dengan jumlah kasus berkisar antara 1784 hingga 2427 pada 2019 yang turun menjadi 433 hingga 901 pada 2022 dan *Cluster_1* memiliki jumlah kasus terendah dengan angka di bawah 1601 sejak 2019 yang terus menurun hingga kisaran 264–718 pada 2022. Temuan ini memberikan pemahaman lebih dalam mengenai pola penyebaran diare melalui analisis data mining berbasis algoritma *K-Means Clustering*. Informasi ini dapat dimanfaatkan untuk mengidentifikasi daerah berisiko tinggi, menyusun strategi pencegahan yang lebih efektif dan mengoptimalkan kebijakan kesehatan di Kabupaten Brebes.

4. KESIMPULAN

Berdasarkan hasil analisis penelitian menunjukkan bahwa algoritma *K-Means Clustering* menghasilkan performa *Davies Bouldin Index* (DBI) terbaik pada K=3 dengan nilai 0.100, dibandingkan dengan K=5 sebesar 0.119 dan K=7 sebesar 0.128. Nilai DBI yang mendekati 0 ini menegaskan bahwa K=3 merupakan jumlah klaster optimal, dengan kasus diare kategori tinggi terkonsentrasi di Bantarkawung, sehingga menjadi fokus utama dalam penanganan oleh pemerintah daerah. Metode *Knowledge Discovery in Databases* (KDD) memastikan analisis berjalan sistematis, mulai dari seleksi data hingga interpretasi hasil, sehingga kesimpulan yang diperoleh lebih terstruktur dan relevan. Untuk meningkatkan manfaat penelitian ini, implementasi hasil dengan melibatkan pihak terkait serta integrasi data mining dengan *Sistem Informasi Geografis* (SIG) dapat dilakukan guna memvisualisasikan pola spasial penyebaran diare dan mendukung strategi pencegahan yang lebih optimal di Kabupaten Brebes.

REFERENCES

[1] WHO, “Kesehatan dan Kesejahteraan,” World Health Organization. [Online]. Available:

https://doi.org/10.33998/processor.2025.20.1.2050
Submitted: 07 Januari 2025; Reviewed Round-1: 03 Februari 2025; Reviewed Round-2: 04 Maret 2025;
Accepted: 14 Maret 2025; Published: 30 April 2025

https://www.who.int/data/gho/data/major-themes/health-and-well-being
[2] muthmainnah, Yunita, Sidaria, Latifa, Rahmi, and R. Aprilianty, *Buku Penanganan Diare pada Anak Menggunakan Metode BRAT (Bread, Rice, Applesauce, and Toast)*. penerbit adab, 2023. [Online]. Available: https://www.google.co.id/books/edition/BUKU_PENANGANAN_DIARE_PADA_ANAK_MENGGUNA/8ZLQEAAAQBAJ?hl=id&gbpv=1&dq=diare++terbaru&pg=PA46&printsec=frontcover
[3] D. Setiadi *et al.*, "Penerapan Algoritma K-Means Clustering Pada Pembesaran," vol. 7, no. 6, pp. 3320–3327, 2023.
[4] I. A. R. P. Pratama and D. Ernawati, "Juminten Analisis Persebaran Penyakit Diare di Jawa Barat Menggunakan Data Mining dengan AlgoritmaK-Means Clustering Data Analysis of Diarrhea in West Java Using Data Mining with the K-Means Clustering Algorithm," *Manaj. Ind. dan Teknol.*, vol. 04, no. 01, pp. 1–12, 2023, [Online]. Available: http://juminten.upnjatim.ac.idhttps//doi.org/10.33005/juminten.v4i1.421juminten@upnjatim.ac.id
[5] UNICEF, "Diare," UNICEF. [Online]. Available: https://data.unicef.org/topic/child-health/diarrhoeal-disease/
[6] T. S. Syamfithriani, N. Mirantika, and R. Trisudarmo, "Perbandingan Algoritma K-Means dan K-Medoids Untuk Pemetaan Daerah Penanganan Diare Pada Balita di Kabupaten Kuningan," *J. Sist. Inf. Bisnis*, vol. 12, no. 2, pp. 132–139, 2023, doi: 10.21456/vol12i2pp132-139.
[7] Badan Pusat Statistik Kabupaten Brebes, "Jumlah Kasus HIV/AIDS, IMS, DBD, Diare, TB, dan Malaria Menurut Kecamatan di Kabupaten Brebes," Badan Pusat Statistik Kabupaten Brebes. [Online]. Available: https://brebeskab.bps.go.id/id/statistics-table?subject=522
[8] M. A. Muslim *et al.*, *Data Mining Algoritma C4.5 Disertai contoh kasus dan penerapannya dengan program computer*, vol. 1, no. 13. 2019.
[9] K. Kodratul Munawar and A. Irma Purnamasari, "Implementasi Algoritma K-Means Clustering Pada Klasterisasi Kasus Hiv Di Jawa Barat," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 7, no. 2, pp. 1092–1099, 2023, doi: 10.36040/jati.v7i2.6372.
[10] M. Soni, N. Rahaningsih, and R. Danar Dana, "Komparasi Algoritma K-Means Dan K-Medoids Clustering Pada Data Penyebaran Kasus Hiv Di Provinsi Jawa Barat," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 7, no. 6, pp. 3766–3772, 2024, doi: 10.36040/jati.v7i6.8274.
[11] P. Cahyo and B. Sulpadianti, *pembuatan aplikasi clustering gangguan jaringan menggunakan metode k-means clustering*. Kreatif Industri Nusantara, 2020. [Online]. Available: https://www.google.co.id/books/edition/Pembuatan_aplikasi_clustering_gangguan_j/y8TgDwAAQBAJ?hl=id&gbpv=1&dq=k-means&pg=PA17&printsec=frontcover
[12] S. S. H. Tita Puspita Sari, Ir. April Lia Hananto, Elfina Novalia, Tukino, "Penerapan Algoritma K-Means Dalam Klasterisasi Penyebaran Penyakit Hiv/Aids," *Infotek J. Inform. dan Teknol.*, 2021, doi: https://dx.doi.org/10.29408/jit.v4i1.2999.
[13] D. Saepu Qirom, A. Faqih, and G. Dwilestari, "Implementasi Algoritma K-Means Untuk Klasterisasi Pasien Hipertensi Bersadarkan Karakteristik Pasien," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 8, no. 2, pp. 2056–2063, 2024, doi: 10.36040/jati.v8i2.8314.
[14] F. Dikarya and S. Muharni, "Penerapan Algoritma K-Means Clustering Untuk Pengelompokan Universitas Terbaik Di Dunia," *J. Inform.*, vol. 22, no. 2, pp. 124–131, 2022, doi: 10.30873/ji.v22i2.3324.
[15] I. Indra, N. Nur, M. Iqram, and N. Inayah, "Perbandingan K-Means dan Hierarchical Clustering dalam Pengelompokan Daerah Beresiko Stunting," *INOVTEK Polbeng - Seri Inform.*, vol. 8, no. 2, p. 356, 2023, doi: 10.35314/isi.v8i2.3612.
[16] G. G. Setiaji, A. N. Putri, and D. A. Wicaksana, "Perbandingan Algoritma K-Means dan K-Medoids Untuk Clustering Harga Beras di Provinsi Jawa Tengah," *J. Transform.*, vol. 22, no. 1, pp. 39–45, 2024, doi: 10.26623/transformatika.v22i1.10092.
[17] E. Yolanda, "Penerapan Algoritma K-Means Clustering Untuk Pengelompokan Data Pasien Rehabilitasi Narkoba," *KLIK Kaji. Ilm. Inform. dan Komput.*, vol. 4, no. 1, pp. 182-`191, 2023, doi: 10.30865/klik.v4i1.1107.
[18] N. Purba, P. Poningsih, and H. S. Tambunan, "Penerapan Algoritma K-Means Clustering Pada Penyebaran Penyakit Infeksi Saluran Pernapasan Akut (ISPA) di Provinsi Riau," *J. Inf. Syst. Res.*, vol. 2, no. 3, pp. 220–226, 2021, [Online]. Available: http://ejurnal.seminar-id.com/index.php/josh/article/view/736
[19] F. C-means and K. Algorithms, "Comparative Study of Earthquake Clustering in Relation.pdf," vol. 5, no. 158, pp. 768–778, 2024.
[20] A. H. Nasyuha, Zulham, and I. Rusydi, "Implementation of K-means algorithm in data analysis," *Telkomnika (Telecommunication Comput. Electron. Control.*, vol. 20, no. 2, pp. 307–313, 2022, doi: 10.12928/TELKOMNIKA.v20i2.21986.
[21] N. Nurahman, A. Purwanto, and S. Mulyanto, "Klasterisasi Sekolah Menggunakan Algoritma K-Means berdasarkan Fasilitas, Pendidik, dan Tenaga Pendidik," *MATRIK J. Manajemen, Tek. Inform. dan Rekayasa Komput.*, vol. 21, no. 2, pp. 337–350, 2022, doi: 10.30812/matrik.v21i2.1411.
[22] Q. I. Mawarni and E. S. Budi, "Implementasi Algoritma K-Means Clustering Dalam Penilaian Kedisiplinan Siswa," *J. Sist. Komput. dan Inform.*, vol. 3, no. 4, p. 522, 2022, doi: 10.30865/json.v3i4.4242.
[23] N. Rohman and A. Wibowo, "Perbandingan Metode K-Medoids dan Metode K-Means Dalam Analisis Segmentasi Pelanggan Mall," *SINTECH (Science Inf. Technol. J.*, vol. 7, no. 1, pp. 49–58, 2024, doi: 10.31598/sintechjournal.v7i1.1507.
[24] L. S. Riza, R. A. Rosdiyana, A. Wahyudin, and A. R. Pérez, "The k-means algorithm for generating sets of items in educational assessment," *Indones. J. Sci. Technol.*, vol. 6, no. 1, pp. 93–100, 2021, doi: 10.17509/ijost.v6i1.31523.
[25] S. Ariska, D. Puspita, and I. Anggraini, "Comparison Of K-Means and K-Medoids Algorithm for Clustering Data UMKM in Pagar Alam City," *J. Sisfokom (Sistem Inf. dan Komputer)*, vol. 13, no. 2, pp. 193–199, 2024, doi:

https://doi.org/10.33998/processor.2025.20.1.2050
Submitted: 07 Januari 2025; Reviewed Round-1: 03 Februari 2025; Reviewed Round-2: 04 Maret 2025;
Accepted: 14 Maret 2025; Published: 30 April 2025

10.32736/sisfokom.v13i2.2090.

https://doi.org/10.33998/processor.2025.20.1.2050
Submitted: 07 Januari 2025; Reviewed Round-1: 03 Februari 2025; Reviewed Round-2: 04 Maret 2025;
Accepted: 14 Maret 2025; Published: 30 April 2025