

Penerapan Normalisasi Data Metode Decimal Scaling Dan Metode K-Means Dalam Mengelompokkan Kasus Demam Berdarah

Ila Yati Beti^{1*}, Hengki Juliansa²

¹ Fakultas Ilmu Komputer, Informatika, Universitas Dehasen Bengkulu, Bengkulu, Indonesia

² Prodi Teknik Informatika, Politeknik Sekayu, Musi Banyuasin, Indonesia

Email: ^{1,*}lilyb@unived.ac.id, ²hengki.juliansa@gmail.com

Email Penulis Korespondensi: lilyb@unived.ac.id

Abstrak— Demam berdarah, disebabkan oleh virus yang ditularkan melalui gigitan nyamuk *Aedes aegypti* dan *Aedes albopictus*, terus menunjukkan tren peningkatan kasus yang mengkhawatirkan. Meskipun upaya-upaya pencegahan dan pengendalian telah diterapkan secara luas, peningkatan ini menimbulkan kekhawatiran serius di kalangan para ahli kesehatan masyarakat dan pemerintah. Penyebab dari peningkatan kasus demam berdarah tahun 2024 ini dapat bervariasi, termasuk perubahan iklim yang mempengaruhi persebaran nyamuk pembawa virus, urbanisasi yang meningkatkan habitat nyamuk, dan perubahan perilaku manusia yang mempengaruhi tingkat kebersihan lingkungan. Namun, data tersebut seringkali tersebar dan memiliki skala yang berbeda-beda, sehingga sulit untuk langsung dianalisis. Analisis data kasus demam berdarah menjadi penting untuk memahami pola penyebaran penyakit dan mengambil langkah-langkah pencegahan dengan metode Decimal Scaling dan pengelompokan data dengan metode K-Means membantu dalam memahami pola-pola kasus demam berdarah. Dimana metode Decimal Scaling untuk menghasilkan data yang lebih baik dan seimbang. Setelah data ternormalisasi maka proses selanjutnya menggali informasi data demam berdarah dengan menerapkan pengelompokan data mining menggunakan Metode K-Means. Berdasarkan hasil pengujian hasil cendroid dimana cluster 0 memiliki nilai yang lebih besar pada masing-masing nilai grade demam berdarah yang terlihat pada gambar 4 sehingga didapatkan hasil pengujian dengan jumlah sampel sebanyak 197 data uji dari tahun 2019 sampai tahun 2023 dengan 2 jumlah cluster dimana cluster 0 memiliki 81 anggota dan cluster 1 memiliki 115 anggota. Sehingga dapat disimpulkan yang mendapatkan prioritas penanganan yang lebih utama pada pada kelompok cluster C0.

Kata Kunci: Demam Berdarah; Data Mining; Normalisasi Data; Decimal Scaling; K-Means

Abstract—Dengue fever, caused by a virus transmitted through the bite of *Aedes aegypti* and *Aedes albopictus* mosquitoes, continues to show an alarming trend of increasing cases. Although prevention and control efforts have been widely implemented, this increase is raising serious concerns among public health experts and governments. The causes of the increase in dengue fever cases in 2024 may vary, including climate change which affects the distribution of mosquitoes that carry the virus, urbanization which increases mosquito habitat, and changes in human behavior that affect the level of environmental cleanliness. However, this data is often scattered and has different scales, making it difficult to directly analyze. Data analysis of dengue fever cases is important to understand the pattern of disease spread and take preventive steps using the Decimal Scaling method and grouping data using the K-Means method helps in understanding patterns of dengue fever cases. Where is the Decimal Scaling method to produce better and balanced data. After the data is normalized, the next process is to explore information on dengue fever data by applying data mining grouping using the K-Means method. Based on the results of the cendroid test results where cluster 0 has a greater value for each dengue fever grade value as seen in Figure 4, the test results were obtained with a total sample of 197 test data from 2019 to 2023 with 2 number of clusters where cluster 0 has 81 members and cluster 1 has 115 members. So it can be concluded that those who get priority for treatment are more important in the C0 cluster group.

Keywords: Dengue Fever; Data Mining; Data Normalization; Decimal Scaling; K-Means

1. PENDAHULUAN

Demam berdarah adalah penyakit yang disebabkan oleh infeksi virus yang ditularkan melalui gigitan nyamuk *Aedes aegypti* dan *Aedes albopictus* [1] [2]. Penyakit ini merupakan masalah kesehatan masyarakat yang serius di banyak negara tropis dan subtropis di seluruh dunia, termasuk di Indonesia. Pada tahun 2024 telah menjadi saksi dari tantangan yang terus berkembang dalam penanggulangan penyakit demam berdarah di Indonesia. Demam berdarah, disebabkan oleh virus yang ditularkan melalui gigitan nyamuk *Aedes aegypti* dan *Aedes albopictus*, terus menunjukkan tren peningkatan kasus yang mengkhawatirkan. Meskipun upaya-upaya pencegahan dan pengendalian telah diterapkan secara luas, peningkatan ini menimbulkan kekhawatiran serius di kalangan para ahli kesehatan masyarakat dan pemerintah.

Penyebab dari peningkatan kasus demam berdarah tahun 2024 ini dapat bervariasi, termasuk perubahan iklim yang mempengaruhi persebaran nyamuk pembawa virus, urbanisasi yang meningkatkan habitat nyamuk, dan perubahan perilaku manusia yang mempengaruhi tingkat kebersihan lingkungan. Sebagai respons terhadap peningkatan ini, perlu dilakukan analisis mendalam untuk memahami pola-pola dan faktor-faktor yang berkontribusi terhadap penyebaran penyakit ini. Pada beberapa kasus demam berdarah, data yang terkait dengan faktor-faktor penyebab dan karakteristik pasien dapat menjadi sangat penting untuk analisis dan pengambilan keputusan. Namun, data tersebut seringkali tersebar dan memiliki skala yang berbeda-beda, sehingga sulit untuk langsung dianalisis. Analisis data kasus demam berdarah menjadi penting untuk memahami pola penyebaran penyakit dan mengambil langkah-langkah pencegahan yang tepat salah satunya dengan memanfaatkan Data Mining.

Data mining adalah proses pengolahan data yang bertujuan untuk menggali informasi yang tersimpan dalam dataset. Informasi yang ditemukan melalui penggalian data ini dapat digunakan dalam pengambilan keputusan. Data mining merupakan sebuah teknik atau bidang ilmu yang relevan dalam berbagai disiplin pengetahuan yang terkait dengan pengolahan data[3][4][5]. Dalam menganalisis data kasus demam berdarah dengan memanfaatkan data mining perlu

dilakukan analisis mendalam untuk memahami pola-pola dan faktor-faktor yang berkontribusi terhadap penyebaran penyakit ini. Sehingga diperlukan data yang baik dan akurat, maka dilakukan normalisasi data agar memperoleh data yang lebih akurat. Dimana data kasus demam berdarah dapat dikelompokkan dengan tepat, maka harus dilakukan normalisasi data. Normalisasi data adalah proses yang memastikan rentang nilai dari beberapa variabel menjadi seragam, tanpa ada yang terlalu besar atau terlalu kecil, sehingga mempermudah analisis statistik[6]. Salah satu teknik normalisasi yang umum digunakan adalah Metode Decimal Scaling, yang merupakan suatu cara untuk mentransformasikan data dengan normalisasi guna memastikan rentang nilai pada setiap atribut menjadi seragam dengan skala tertentu, dengan menggeser nilai desimal dari data sesuai kebutuhan[7]. Hal ini bertujuan untuk menghasilkan data yang lebih baik dan seimbang. Setelah data ternormalisasi maka proses selanjutnya menggali informasi data demam berdarah dengan menerapkan pengelompokan data mining menggunakan Metode K-Means. Metode K-Means adalah merupakan algoritma yang diterapkan dalam proses pengelompokan data atau clustering [8][9]. Tujuan dari penerapan algoritma ini adalah untuk mengelompokkan data ke dalam beberapa kluster atau partisi, dengan maksud untuk mengidentifikasi informasi yang terkandung dalam setiap kelompok. Penerapan metode K-Means membantu mengidentifikasi pola-pola dalam kasus demam berdarah, seperti pola penyebaran geografis dan karakteristik demografis pasien.

Terdapat penelitian terkait yang mendukung penelitian ini, seperti penelitian yang dilakukan oleh Soeb Aripin, dkk pada tahun 2023 tentang penanganan kasus Demam Berdarah dengan Metode K-Medoids Clustering diperoleh 72 anggota dan rendah 28 anggota berdasarkan kaliter penyakit DBD [10]. Penelitian yang dilakukan Muhammad Rafli Kusnaldi, dkk pada tahun 2022 tentang penerapan normalisasi data untuk mengelompokkan data mahasiswa menggunakan metode K-Means dalam menentukan prioritas bantuan UKT berhasil diterapkan dengan sampel data yang terdiri dari 22 orang terdapat 3 cluster dimana jumlah cluster 0 cluster 1 dan cluster 3 dan yang mendapat prioritas yaitu cluster 0 [11]. Penelitian yang dilakukan Daniel Tunggono Saputro dan Wida Pesah Sucihermayanti pada tahun 2021 tentang klaterisasi tingkat kesehatan bayi dan balita menggunakan K-Means mendapatkan tiga klaster meliputi tingkat kesehatan tinggi, sedang dan rendah [12]. Penelitian yang dilakukan. Darnisa Azzahra Nasution dkk pada tahun 2019 tentang perbandingan normalisasi data menyimpulkan terdapat tiga cara yang dilakukan dalam transformasi data dengan normalisasi, yaitu min-max normalization, z-score normalization, dan decimal scaling bahwa akurasi terbaik terletak pada dataset wine yang telah dinormalisasi menggunakan metode min-max normalization[7].

Berdasarkan masalah dan penelitian terdahulu yang menjadi acuan, penelitian ini memiliki tujuan untuk melihat pengaruh normalisasi data dengan metode Decimal Scaling untuk menyamakan skala data sehingga perbedaan skala tidak memengaruhi analisis. Sedangkan, pengelompokkan menggunakan metode K-Means membantu untuk mengidentifikasi penyakit demam berdarah dengan cara membantu mengelompokkan kasus demam berdarah

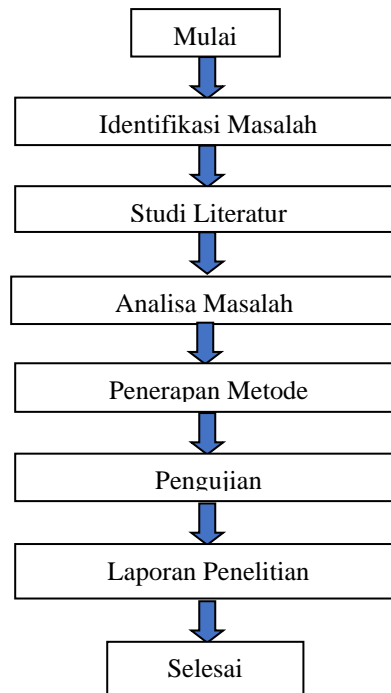
2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Pada bagian tahapan penelitian dalam penerapan normalisasi data metode Decimal Scaling dan metode K-Means dalam mengelompokkan kasus Demam Berdarah dilakukan beberapa tahapan penelitian seperti dibawah ini:

- a. Identifikasi masalah, pada tahapan ini dilakukan untuk mengetahui permasalahan dan metode yang digunakan pada penelitian ini.
- b. Studi literatur, ditahap ini penulis mempelajari beberapa kajian literatur terkait dengan penelitian yang telah dibuat oleh beberapa orang sebelumnya termasuk juga mempelajari jurnal-jurnal atau buku-buku yang berkaitan dengan demam berdarah, data mining, normalisasi data metode Decimal Scaling dan metode K-Means.
- c. Analisa masalah, pada tahap analisa ini penulis melakukan pengumpulan data, mempelajari dan melakukan perumusan demi mendukung penelitian ini dalam melakukan proses pengolahan data.
- d. Penerapan metode, tahap ini merupakan tahap pengelompokan data mahasiswa dengan metode K-Medoids yang dimana hasil akhir dari penelitian ini menghasilkan mengelompokkan kasus Demam Berdarah.
- e. Pengujian merupakan suatu cara atau teknik untuk menentukan data uji yang dapat menguji secara lengkap dan memiliki kemungkinan yang tinggi untuk menemukan masalah yang dimana pada penelitian ini diuji menggunakan Aplikasi Rapid Miner dengan berapa jumlah cluster yang perlu dimasukkan dan hanya memiliki atribut bertipe numeric dengan mengelompokkan data clustering setiap dusun yang mengalami penyakit demam berdarah.
- f. Laporan penelitian bertujuan untuk Menemukan hasil yang diperoleh dari penelitian yang telah dilakukan oleh peneliti tersebut. Pada penelitian ini harus dihasilkan pengelompokan penanganan demam berdarah dengan menggunakan normalisasi data metode Decimal Scaling dan metode K-Means untuk melakukan penanganan demam berdarah pada tahun yang memiliki tingkat tertinggi yang mengalami penyakit demam berdarah

Berdasarkan tahapan penelitian diatas dapat digambarkan dalam bentuk diagram sebagai berikut ini:



Gambar 1. Tahapan Penelitian

2.2 Demam Berdarah

Demam berdarah merupakan sebuah penyakit yang tersebar luas di Indonesia dan disebabkan oleh virus Dengue yang ditularkan oleh nyamuk Aedes dan Aedes Albopictus kepada manusia melalui gigitannya. Kejadian demam berdarah juga dipengaruhi oleh faktor iklim yang mempengaruhi pertumbuhan dan perkembangan nyamuk Aedes dan Aedes Albopictus [13][14]. Penyakit demam berdarah memiliki tingkatan keparahan yang bervariasi, mulai dari derajat 1 hingga 4 yaitu[10]:

- DBD grade I : demam disertai 2 atau lebih tanda : sakit kepala, nyeri di belakang bola mata, pegal pegal dan nyeri sendi dengan uji bendung positif.
- DBD grade II : gejala diatas disertai perdarahan spontan seperti bintik bintik merah di kulit, mimisan, perdarah gusi, muntah darah atau berak hitam.
- DBD grade III : gejala diatas disertai kegagalan sirkulasi (kulit dingin dan lembab serta gelisah)
- DBD grade IV: Renjatan/ syok berat dengan tekanan darah dan nadi tidak terukur

2.3 Normalisasi Data

Normalisasi data adalah proses yang memastikan rentang nilai dari beberapa variabel menjadi seragam, tanpa ada yang terlalu besar atau terlalu kecil, sehingga mempermudah analisis statistik Tujuannya adalah untuk menghilangkan perbedaan skala antar atribut atau variabel dalam dataset, sehingga memungkinkan perbandingan yang adil antara atribut-atribut tersebut. Dengan normalisasi data, nilai-nilai yang awalnya memiliki rentang yang berbeda-beda akan disesuaikan menjadi rentang yang serupa[15][16].

2.4 Metode Decimal Scaling

Metode Decimal Scaling merupakan suatu cara untuk mentransformasikan data dengan normalisasi guna memastikan rentang nilai pada setiap atribut menjadi seragam dengan skala tertentu, dengan menggeser nilai desimal dari data sesuai kebutuhan. Hal ini bertujuan untuk menghasilkan data yang lebih baik dan seimbang. Metode decimal scalling bekerja dengan cara menggeser titik desimal pada nilai data[17]. Banyaknya penggeseran titik desimal bergantung pada nilai maksimum absolut dari masing-masing fitur data atau variabel. Decimal scalling digunakan ketika nilai berada dalam rentang logaritmik, misalnya rentang suatu nilai adalah antara 0 dan 1 sedangkan nilai lain pada rentang 0 dan 1000. Adapun rumus decimal scalling adalah sebagai berikut[11] :

$$\text{Newdata}_i = \text{data}_i / i \quad (1)$$

Keterangan :

Newdata : Data hasil normalisasi

I : Adalah nilai max dari kriteria

2.5 Metode K-Means

Metode K-Means adalah merupakan algoritma yang diterapkan dalam proses pengelompokan data atau clustering. Data yang termasuk dalam satu kelompok harus memiliki jarak yang paling pendek, kriteria, kondisi, atau karakteristik yang serupa atau hampir serupa satu sama lain[18][19][20]. Algoritma K-Means memiliki kemampuan untuk mengelompokkan objek yang memiliki kesamaan. Berikut adalah tahapan-tahapan dalam melakukan pengelompokan data menggunakan Algoritma K-Means[11]:

- Menyiapkan sampling data
- Menentukan jumlah cluster
- Menentukan nilai centroid/titik pusat
- Menghitung jarak masing-masing centroid
- Mengelompokkan data berdasarkan jarak terpendek

Untuk Menentukan jarak masing-masing centroid berdasarkan jarak terpendek menggunakan model Euclidean yang dirumuskan sebagai berikut:

$$d_{ij} = \sqrt{(x_{1i} - x_{1j})^2 + (x_{2i} - x_{2j})^2 + \dots + (x_{ki} - x_{kj})^2} \quad (2)$$

Keterangan:

d_{ij} : Jarak dari data ke i ke pusat cluster j

x_{1i} : Data dari ke - i pada atribut data ke-k

x_{1j} : Data dari ke - j pada atribut data ke-k

3. HASIL DAN PEMBAHASAN

3.1 Pembahasan

Dalam mengelompokkan kasus Demam Berdarah dilakukan dengan normalisasi dengan metode Decimal Scaling dan pengelompokkan data dengan metode K-Means membantu dalam memahami pola-pola kasus demam berdarah. Sampel data yang digunakan sebanyak 54 kasus demam berdarah yang diambil secara acak berdasarkan 4 wilayah. Adapun sampel data yang digunakan dapat dilihat pada tabel dibawah 1 dibawah ini:

Tabel 1. Sampel Data

No	Tahun	Wilayah A	Wilayah B	Wilayah C	Wilayah D	Demam Berdarah Grade 1	Demam Berdarah Grade 2	Demam Berdarah Grade 3	Demam Berdarah Grade 4
1	2019	0	1	0	1	2	1	2	1
2	2019	1	1	1	1	1	2	2	1
3	2019	1	0	1	1	1	2	2	1
4	2019	1	1	1	1	2	1	2	1
5	2019	1	1	1	1	2	1	1	2
6	2019	1	1	0	1	1	2	2	1
7	2019	1	1	1	1	1	2	2	1
8	2019	1	1	1	1	1	2	2	1
..
..
54	2023	0	1	0	0	1	1	2	1

Maka proses selanjutnya melakukan transformasi data berdasarkan wilayah agar sampel data dilakukan normalisasi data. Adapun hasil proses transformasi data berdasarkan wilayah terdapat 196 kasus terdapat pada tabel 2 berikut ini:

Tabel 2. Data Sampel Hasil Transformasi

No	Wilayah	Demam Berdarah Grade 1	Demam Berdarah Grade 2	Demam Berdarah Grade 3	Demam Berdarah Grade 4
1	Wilayah B	2	1	2	1
2	Wilayah D	2	1	2	1
3	Wilayah A	1	2	0	1
4	Wilayah B	1	1	2	1
5	Wilayah C	1	2	0	1
6	Wilayah D	1	5	2	1
7	Wilayah A	1	2	2	1
8	Wilayah C	1	2	2	1
9	Wilayah D	1	2	2	1
10	Wilayah D	1	2	2	1
....
....

196	Wilayah B	1	1	2	1
-----	-----------	---	---	---	---

Setelah sampel data telah didapatkan maka proses yang dilakukan selanjutnya normalisasikan data kasus demam berdarah menggunakan metode Decimal Scaling untuk mendapatkan hasil data demam berdarah yang telah di normalisasikan sebai sampel data uji metode K-Means. Adapun prose normalisasi metode Decimal Scaling berdasarkan rumus 1 sebagai berikut:

1. Normalisaki kriteria Demam Berdarah Grade 1
 $Newdata_i = 2/2 = 1$
2. Normalisaki kriteria Demam Berdarah Grade 2
 $Newdata_i = 1/5 = 0,5$
3. Normalisaki kriteria Demam Berdarah Grade 3
 $Newdata_i = 2/3 = 0,666666667$
4. Normalisaki kriteria Demam Berdarah Grade 4
 $Newdata_i = 1/2 = 0,5$

Hasil normalisasi data menggunakan metode Decimal Scaling untuk mendapatkan hasil normalisasi data demam berdarah dapat dilihat pada tabel 3 berikut ini:

Tabel 3. Data Hasil Decimal Scaling

No	Wilayah	Demam Berdarah Grade 1	Demam Berdarah Grade 2	Demam Berdarah Grade 3	Demam Berdarah Grade 4
1	Wilayah B	1	0,2	0,666667	0,5
2	Wialayh D	1	0,2	0,666667	0,5
3	Wilayah A	0,5	0,4	0	0,5
4	Wilayah B	0,5	0,2	0,666667	0,5
5	Wilayah C	0,5	0,4	0	0,5
6	Wialayh D	0,5	1	0,666667	0,5
7	Wilayah A	0,5	0,4	0,666667	0,5
8	Wilayah C	0,5	0,4	0,666667	0,5
9	Wialayh D	0,5	0,4	0,666667	0,5
10	Wialayh D	0,5	0,4	0,666667	0,5
....
....
196	Wilayah B	0,5	0,2	0,666667	0,5

Setelah normalisasi dilakukan pada data sampel, langkah berikutnya adalah menerapkan metode K-Means untuk mengelompokkan data. Proses penerapan metode K-Means dalam pengelompokan data demam berdarah berdasarkan sampel data yang telah dinormalisasi dengan metode Decimal Scaling adalah sebagai berikut:

- a. Menyiapkan sampling data
 Data sapel yang telah dinomalisasi pada tabel 3 diatas
- b. Menentukan jumlah cluster
 Jumlah cluster / kelompok yaitu 2
- c. Menentukan nilai centroid/titik pusat
 Menentukan nilai centroid awal / titik pusat awal, dengan 2 record data sesuai dengan jumlah attribute kriteria. Nilai centroid awal / titik pusat awal diperoleh dari data demam berdarah secara acak pada tabel 4.

Tabel 4. Nilai Centroid Awal

Pusat Iterasi Ke1	Wilayah	Demam Berdarah Grade 1	Demam Berdarah Grade 2	Demam Berdarah Grade 3	Demam Berdarah Grade 4
C0 [10]	Wialayh D	0,5	0,4	0,666667	0,5
C1 [50]	Wilayah C	0,5	0,4	0,333333	0,5

- d. Menghitung jarak masing-masing centroid
 1. Cluster C0
 $d_{x1,c0} = \sqrt{(1 - 0,5)^2 + (0,2 - 0,4)^2 + (0,666667 - 0,666667)^2 + (0,5 - 0,5)^2} = 0,538516481$
 2. Cluster C1
 $d_{x1,c0} = \sqrt{(1 - 0,5)^2 + (0,2 - 0,4)^2 + (0,666667 - 0,333333)^2 + (0,5 - 0,5)^2} = 0,633333509$

Untuk hasil pada masing cluster dapat dilihat pada tabel 5 dibawah ini:

Tabel 5. Hasil Masing-Masing Centroid

No	Wilayah	Demam Berdarah Grade 1	Demam Berdarah Grade 2	Demam Berdarah Grade 3	Demam Berdarah Grade 4	Cluster C0	Cluster C1
1	Wilayah B	1	0,2	0,666667	0,5	0,538516481	0,633333509
2	Wialayh D	1	0,2	0,666667	0,5	0,538516481	0,633333509
3	Wilayah A	0,5	0,4	0	0,5	0,666657	0,333333
4	Wilayah B	0,5	0,2	0,666667	0,5	0,2	0,388730412
5	Wilayah C	0,5	0,4	0	0,5	0,666657	0,333333
6	Wialayh D	0,5	1	0,666667	0,5	0,6	0,686375505
7	Wilayah A	0,5	0,4	0,666667	0,5	9,66667E-06	0,333333667
8	Wilayah C	0,5	0,4	0,666667	0,5	9,66667E-06	0,333333667
9	Wialayh D	0,5	0,4	0,666667	0,5	9,66667E-06	0,333333667
10	Wialayh D	0,5	0,4	0,666667	0,5	9,66667E-06	0,333333667
....
....
196	Wilayah B	0,5	0,2	0,666667	0,5	0,2	0,38873

e. Mengelompokkan data berdasarkan jarak terpendek

Aturan yang digunakan untuk mengelompokkan / cluster data berdasarkan hasil perhitungan jarak adalah

Jika $dc1 < dc2$ maka cluster =0 dan Jika $dc2 > dc1$ maka cluster =1

Tabel 6. Hasil Cluster 1

No	Wilayah	Demam Berdarah Grade 1	Demam Berdarah Grade 2	Demam Berdarah Grade 3	Demam Berdarah Grade 4	DC0	DC1	Cluster
1	Wilayah B	1	0,2	0,666667	0,5	0,538516481	0,633333509	C1
2	Wialayh D	1	0,2	0,666667	0,5	0,538516481	0,633333509	C1
3	Wilayah A	0,5	0,4	0	0,5	0,666657	0,333333	C0
4	Wilayah B	0,5	0,2	0,666667	0,5	0,2	0,388730412	C1
5	Wilayah C	0,5	0,4	0	0,5	0,666657	0,333333	C0
6	Wialayh D	0,5	1	0,666667	0,5	0,6	0,686375505	C1
7	Wilayah A	0,5	0,4	0,666667	0,5	9,66667E-06	0,333333667	C1
8	Wilayah C	0,5	0,4	0,666667	0,5	9,66667E-06	0,333333667	C1
9	Wialayh D	0,5	0,4	0,666667	0,5	9,66667E-06	0,333333667	C1
10	Wialayh D	0,5	0,4	0,666667	0,5	9,66667E-06	0,333333667	C1
....
....
196	Wilayah B	0,5	0,2	0,666667	0,5	0,2	0,38873	C1

Proses selanjutnya melakukan iterasi ke dua dengan pencarian nilai centroid / titik pusat awal pada iterasi ke-2 diperoleh dari nilai rata-rata penjumlahan anggota masing-masing cluster, setelah itu dibagi dengan jumlah anggota masing-masing cluster. Adapun hasil yang didapat pada tabel dibawah ini

Tabel 7. Nilai Centroid Pada Iterasi Ke-2

Pusat Iterasi Ke1	Demam Berdarah Grade 1	Demam Berdarah Grade 2	Demam Berdarah Grade 3	Demam Berdarah Grade 4
C0	0,619047619	0,292857143	0,265873016	0,55952381
C1	0,777272727	0,28	0,66969697	0,536363636

Maka proses selanjutnya lakukan hal yang sama dari langkah 4 dan 5 sampai ditemukan tidak terjadi perubahan cluster antara cluster sebelumnya dengan cluster selanjutnya, dalam penelitian ini proses pencarian itersai berhenti pada iterasi ke 7 dimana hasil cluster iterasi 7 memiliki hasil yang sama dengan hasil clusteriterasi 6 seperti tabel 8 dibawah ini:

Tabel 8. Hasil Cluste 7

No	Wilayah	Demam Berdarah Grade 1	Demam Berdarah Grade 2	Demam Berdarah Grade 3	Demam Berdarah Grade 4	DC0	DC1	Cluster
1	Wilayah B	1	0,2	0,666667	0,5	0,569418353	0,103003219	C0
2	Wialayh D	1	0,2	0,666667	0,5	0,569418353	0,103003219	C0
3	Wilayah A	0,5	0,4	0	0,5	0,447856479	0,791747945	C1
4	Wilayah B	0,5	0,2	0,666667	0,5	0,272465156	0,510499425	C1

5	Wilayah C	0,5	0,4	0	0,5	0,447856479	0,791747945	C1
6	Wialayh D	0,5	1	0,666667	0,5	0,687131941	0,942739525	C1
7	Wilayah A	0,5	0,4	0,666667	0,5	0,231766093	0,545570069	C1
8	Wilayah C	0,5	0,4	0,666667	0,5	0,231766093	0,545570069	C1
9	Wialayh D	0,5	0,4	0,666667	0,5	0,231766093	0,545570069	C1
10	Wialayh D	0,5	0,4	0,666667	0,5	0,231766093	0,545570069	C1
....	
....	
196	Wilayah B	0,5	0,2	0,666667	0,5	0,5064	0,268754	C1

Berdasarkan data pengujian itersai 7 didapatkan kelompok cluster C0 sebanyak 815 Kasus, C1 sebanyak 115 kasus. Sehingga dapat disimpulkan yang mendapatkan prioritas penangana yang lebih utama pada pada kelompok cluster C0.

3.2 Hasil Pengujian

Adapun hasil pengujian dalam mengelompokkan kasus Demam Berdarah dilakukan dengan metode Decimal Scaling dan pengelompokkan data dengan metode K-Means membantu dalam memahami pola-pola kasus demam berdarah dengan menggunakan aplikasi data Mining yaitu rapidminer dengan hasil Cluster/Pengelompokan pada gambar, bentuk Cluster Model pada gambar 3 dan Hasil Centroid pada gambar 4 berdasarkan hasil normalisasi desimal scaling dan pengelompokan menggunakan metode K-Means dapat dilihat pada dibawah ini:

Row No.	id	Wilayah	cluster	Demam Berdarah Grade 1	Demam Berdarah Grade 2	Demam Berdarah Grade 3	Demam Berdarah Grade 4
1	1	Wilayah B	cluster_0	1	0.200	0.667	0.500
2	2	Wialayh D	cluster_0	1	0.200	0.667	0.500
3	3	Wilayah A	cluster_1	0.500	0.400	0	0.500
4	4	Wilayah B	cluster_1	0.500	0.200	0.667	0.500
5	5	Wilayah C	cluster_1	0.500	0.400	0	0.500
6	6	Wialayh D	cluster_1	0.500	1	0.667	0.500
7	7	Wilayah A	cluster_1	0.500	0.400	0.667	0.500
8	8	Wilayah C	cluster_1	0.500	0.400	0.667	0.500
9	9	Wialayh D	cluster_1	0.500	0.400	0.667	0.500
10	10	Wialayh D	cluster_1	0.500	0.400	0.667	0.500
11	11	Wilayah B	cluster_0	1	0.200	0.333	0.500
12	12	Wialayh D	cluster_0	1	0.200	0.333	0.500
13	13	Wilayah A	cluster_0	1	0.200	0.333	1
14	14	Wilayah C	cluster_0	1	0.200	0.333	0.500
15	15	Wilayah A	cluster_0	1	0.200	0.333	1
16	16	Wilayah C	cluster_0	1	0.200	0.333	0.500
17	17	Wialayh D	cluster_0	1	0.200	0.333	0.500

Gambar 2. Hasil Cluster/Pengelompokan Dengan Aplikasi Rapidminer

Cluster Model

Cluster 0: 81 items
Cluster 1: 115 items
Total number of items: 196

Gambar 3. Hasil Cluster Model Dengan Aplikasi Rapidminer

Pada gambar 2 dan 3 menjelaskan hasil pengujian yang didapat pada aplikasi Rapidminer dalam menampilkan hasil pengelompokan dan jumlah cluster dimana cluster 0 memiliki 81 anggota dan cluster 1 memiliki 115 anggota terlihat hasil yang didapatkan sama dengan hasil pengujian secara manual

Attribute	cluster_0	cluster_1
Demam Berdarah Grade I	1	0.500
Demam Berdarah Grade 2	0.207	0.351
Demam Berdarah Grade 3	0.580	0.443
Demam Berdarah Grade 4	0.556	0.539

Gambar 4. Hasil Centroid Dengan Aplikasi Rapidminer

Berdasarkan hasil pengujian hasil cendroit dimana cluster 0 memilii nilai yang lebih besar pada masing-masing nilai grade demam berdarah yang terlihat pada gambar 4 sehingga didapatkan hasil pengujian dengan jumlah sampel sebanyak 197 data uji dari tahun 2019 sampai tahun 2023 dengan 2 jumlah cluster dimana cluster 0 memiliki 81 anggota dan cluster 1 memiliki 115 anggota. Sehingga dapat disimpulkan yang mendapatkan prioritas penanganan yang lebih utama pada pada kelompok cluster C0.

4. KESIMPULAN

Berdasarkan hasil dan pengujian dalam mengelompokkan kasus Demam Berdarah dilakukan dengan metode Decimal Scaling dan pengelompokkan data dengan metode K-Means membantu dalam memahami pola-pola kasus demam berdarah. Dimana metode Decimal Scaling untuk menghasilkan data yang lebih baik dan seimbang. Setelah data ternormalisasi maka proses selanjutnya menggali informasi data demam berdarah dengan menerapkan pengelompokan data mining menggunakan Metode K-Means. Berdasarkan hasil pengujian pengujian hasil cendroit dimana cluster 0 memilii nilai yang lebih besar pada masing-masing nilai grade demam berdarah yang terlihat pada gambar 4 sehingga didapatkan hasil pengujian dengan jumlah sampel sebanyak 197 data uji dari tahun 2019 sampai tahun 2023 dengan 2 jumlah cluster dimana cluster 0 memiliki 81 anggota dan cluster 1 memiliki 115 anggota. Sehingga dapat disimpulkan yang mendapatkan prioritas penanganan yang lebih utama pada pada kelompok cluster C0.

REFERENCES

- [1] R. T. S. Muhammad Hariyanto, "Clustering pada Data Mining untuk Mengetahui Potensi Penyebaran Penyakit DBD Menggunakan Metode Algoritma K-Means dan Metode Perhitungan Jarak Euclidean Distance," *Sist. Komput. dan Tek. Inform.*, vol. 1, no. 1, hal. 117–122, 2018.
- [2] A. Sukohar, "Demam Berdarah Dengue (DBD)," *Medula*, vol. 2, no. 2, hal. 1–15, 2014.
- [3] E. Buulolo, *Data Mining Untuk Perguruan Tinggi*. Medan: CV Budi Utama, 2020.
- [4] A. Wanto *et al.*, *Data Mining : Algoritma dan Implementasi*. Medan: Yayasan Kita Menulis, 2020.
- [5] Z. Nabila, A. R. Isnain, Permata, dan Z. Abidin, "ANALISIS DATA MINING UNTUK CLUSTERING KASUS COVID-19 DI PROVINSI LAMPUNG DENGAN ALGORITMA K-MEANS," *J. Teknol. dan Sist. Inf.*, vol. 2, no. 2, hal. 100–108, 2021.
- [6] A. Harmain, H. Kurniawan, D. Maulina, dan M. Universitas Amikom Yogyakarta, "Data Normalization for K-Means Efficiency on Groups of Areas With Potential Fores and /Land Fire Based on Heat Spots Distribution," *Teknimedia*, vol. 2, no. 2, hal. 83–89, 2021.
- [7] D. A. Nasution, H. H. Khotimah, dan N. Chamidah, "Perbandingan Normalisasi Data untuk Klasifikasi Wine Menggunakan Algoritma K-NN," *Comput. Eng. Sci. Syst. J.*, vol. 4, no. 1, hal. 78, 2019, doi: 10.24114/cess.v4i1.11458.
- [8] E. M. Fitri, R. R. Suryono, dan A. Wantoro, "Klasterisasi Data Penjualan Berdasarkan Wilayah Menggunakan Metode K-Means Pada Pt Xyz," *J. Komputasi*, vol. 11, no. 2, hal. 157–168, 2023, doi: 10.23960/komputasi.v11i2.12582.
- [9] A. Yudhistira dan R. Andika, "Pengelompokan Data Nilai Siswa Menggunakan Metode K-Means Clustering," *J. Artif. Intell. Technol. Inf.*, vol. 1, no. 1, hal. 20–28, 2023, doi: 10.58602/jaiti.v1i1.22.
- [10] T. Soeb Aripin dan G. P. N. S. P. Angin, "Penerapan Metode K-Medoids Clustering Pada Penanganan Kasus Demam Berdarah," *BEES Bull. Electr. Electron. Eng.*, vol. 3, no. 3, hal. 139–146, 2023, doi: 10.47065/bees.v3i3.3173.
- [11] M. R. Kusnadi, T. Gulo, dan S. Aripin, "Penerapan Normalisasi Data Dalam Mengelompokkan Data Mahasiswa Dengan Menggunakan Metode K-Means Untuk Menentukan Prioritas Bantuan Uang Kuliah Tunggal," *J. Comput. Syst. Informatics*, vol. 3, no. 4, hal. 330–338, 2022, doi: 10.47065/josyc.v3i4.2112.
- [12] A. S. Soetoko dan D. F. Soetoko, "Meningkatkan Kualitas Hidup Penderita Gangguan Metabolik di Klinik Pratama Semarang melalui Pelatihan Pengaturan Diet dan Senam Kaki," *J. ABDIMAS-KU J. Pengabdi. Masy. Kedokt.*, vol. 1, no. 3, hal. 126, 2022, doi: 10.30659/abdimasku.1.3.126-133.
- [13] M. A. Sembiring, "Penerapan Metode Algoritma K-Means Clustering Untuk Pemetaan Penyebaran Penyakit Demam Berdarah Dengue (Dbd)," *J. Sci. Soc. Res.*, vol. 4, no. 3, hal. 336, 2021, doi: 10.54314/jssr.v4i3.712.
- [14] M. Jasri, A. Wijaya, dan R. Sunggara, "Penerapan Data Mining untuk Klasifikasi Penyakit Demam Berdarah Dengue (DBD) Dengan Metode Naïve Bayes (Studi Kasus Puskesmas Taman Krocok)," *SMARTICS J.*, vol. 8, no. 1, hal. 35–42, 2022, [Daring]. Tersedia pada: <https://doi.org/10.21067/smartics.v8i1.7375>.
- [15] M. Sholeh, D. Andayati, dan R. Y. Rachmawati, "Data Mining Model Klasifikasi Menggunakan Algoritma K-Nearest Neighbor

- Dengan Normalisasi Untuk Prediksi Penyakit Diabetes,” *TeIKa*, vol. 12, no. 02, hal. 77–87, 2022, doi: 10.36342/teika.v12i02.2911.
- [16] S. Emalia Saqila, I. Putri Ferina, dan A. Iskandar, “Analisis Perbandingan Kinerja Clustering Data Mining Untuk Normalisasi Dataset,” *J. Sist. Komput. dan Inform. Hal 356–*, vol. 365, no. 2, 2023, doi: 10.30865/json.v5i2.6919.
- [17] Gde Agung Brahmana Suryanegara, Adiwijaya, dan Mahendra Dwifabri Purbolaksono, “Peningkatan Hasil Klasifikasi pada Algoritma Random Forest untuk Deteksi Pasien Penderita Diabetes Menggunakan Metode Normalisasi,” *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 5, no. 1, hal. 114–122, 2021, doi: 10.29207/resti.v5i1.2880.
- [18] N. Afiasari, N. Suarna, dan N. Rahaningsi, “Implementasi Data Mining Transaksi Penjualan Menggunakan Algoritma Clustering dengan Metode K-Means,” *J. SAINTEKOM*, vol. 13, no. 1, hal. 100–110, 2023, doi: 10.33020/saintekom.v13i1.402.
- [19] M. P. A. Ariawan, I. B. A. Peling, dan G. B. Subiksa, “Prediksi Nilai Akhir Matakuliah Mahasiswa Menggunakan Metode K-Means Clustering (Studi Kasus : Matakuliah Pemrograman Dasar),” *J. Nas. Teknol. dan Sist. Inf.*, vol. 9, no. 2, hal. 122–131, 2023, doi: 10.25077/teknosi.v9i2.2023.122-131.
- [20] I. Pii, N. Suarna, dan N. Rahaningsih, “Penerapan Data Mining Pada Penjualan Produk Pakaian Dameyra Fashion Menggunakan Metode K-Means Clustering,” *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 7, no. 1, hal. 423–430, 2023, doi: 10.36040/jati.v7i1.6336.