

Analisis Klasterisasi Penyakit Malaria Menggunakan Metode K-Means di Indonesia

Luthfi Ramdhan Faqih¹, Suastika Yulia Riska^{2*}

^{1,2*}Program Studi Teknik Informatika – Institut Teknologi dan Bisnis Asia Malang

¹luthfiramdhanfaqih@gmail.com, ²riska.suastika@asia.ac.id

ABSTRAK. Malaria merupakan penyakit berbahaya dan berpotensi fatal di Indonesia. Penyebaran dan penularan penyakit malaria terjadi dengan sangat cepat. Tujuan dari penelitian ini adalah untuk mengidentifikasi cluster di negara bagian berdasarkan intensitas kasus malaria. Pada penelitian ini diterapkan metode K-Means pada proses clustering dengan menggunakan nilai K=2, K=3, dan K=5. Artinya, mengikuti K=2 dengan nilai indeks Davies-Bouldin sebesar 0,033, K=3 memiliki nilai indeks Davies-Bouldin sebesar 0,034, dan K=5 memiliki nilai indeks Davies-Bouldin sebesar 0,262. Hasil penelitian menunjukkan bahwa penggunaan K-Means dengan K=2 menghasilkan cluster terbaik dengan nilai indeks Davies-Bouldin terendah (0,033). Hal ini dapat membantu pemerintah merencanakan tindakan pencegahan yang lebih efektif di berbagai provinsi di Indonesia pada tahun-tahun mendatang. Oleh karena itu, penelitian ini memberikan kontribusi penting terhadap upaya pengendalian malaria untuk mengurangi kejadian malaria dan dampak kesehatan masyarakat di Indonesia.

Kata Kunci: *Clustering, Penyakit Malaria, K-Means, Davies-Bouldin Indexs*

ABSTRACT. *Malaria is a dangerous and potentially deadly disease in Indonesia. The spread and transmission of malaria occurs very rapidly. The aim of this study was to identify clusters within the state based on the intensity of malaria cases. In this study, K-means was applied to the clustering process using the values of K=2, K=3, and K=5. This means that the Davis-Boldan index value for K=2 is 0.033, the Davis-Boldan index value for K=3 is 0.034, and the Davis-Boldan index value for K=5 is 0.262. The research results show that using K-Means with K=2 yields the best cluster with the lowest Davies-Bouldin index value (0.033). This will help the government plan more effective preventive measures in different provinces of Indonesia in the coming years. Therefore, this study makes an important contribution to malaria control efforts to reduce malaria incidence and public health impact in Indonesia.*

Keyword: *Clustering, Malaria, K-Means, Davies-Bouldin Indexs*

PENDAHULUAN

Indonesia merupakan salah satu negara yang masih menghadapi permasalahan serius terkait penyakit malaria. Berdasarkan data WHO tahun 2022, malaria masih menjadi masalah kesehatan yang perlu mendapat perhatian. Di Indonesia, dua spesies parasit malaria, *Plasmodium falciparum* dan *Plasmodium vivax*, menyebabkan penyakit malaria yang menimbulkan berbagai gejala seperti demam, sakit kepala, muntah, dan diare (Kalua et al., 2022;Yohannes et al., 2020) . Penyebab tingginya angka kejadian malaria di Indonesia adalah iklim tropis. Hal ini menjadikannya habitat pilihan nyamuk penyebab malaria (Karmila et al., 2017). Pada tahun 2019, jumlah infeksi malaria mencapai 250.644 kasus(Sari & Syakurah, 2023). Malaria masih menjadi penyebab utama kesakitan dan kematian di Indonesia, terutama pada anak-anak dan ibu hamil. Meskipun ada upaya pemberantasan, kasus dan kematian akibat malaria masih tetap tinggi di beberapa wilayah di Indonesia. Oleh karena itu, diperlukan pendekatan yang lebih efektif untuk mengatasi masalah ini.

Clustering atau pengelompokan data merupakan suatu metode analisis data yang dapat digunakan untuk mengidentifikasi pola dan hubungan pada data yang besar dan kompleks (Sindi et al., 2020). Dalam konteks malaria, pengelompokan data dapat membantu mengidentifikasi pola sebaran penyakit, faktor risiko, dan karakteristik populasi yang rentan terhadap penyakit ini. Penelitian mengenai wabah malaria di Indonesia mungkin sudah pernah dilakukan sebelumnya, namun masih terfokus pada beberapa provinsi saja (Bete et al., 2023). Oleh karena itu, penelitian ini akan berfokus pada penerapan algoritma K-Means untuk mengelompokkan data penyakit malaria di Indonesia dan memperoleh wawasan baru yang dapat digunakan untuk perencanaan dan pengambilan keputusan dalam pengendalian penyakit ini.

Masalah yang dihadapi dalam penelitian ini adalah menggunakan algoritma K-Means untuk clustering penyakit malaria di Indonesia. Penelitian ini bertujuan untuk mengetahui pola dan distribusi penyakit malaria

berdasarkan jenis, usia, dan bulan kejadian serta mengidentifikasi cluster yang relevan untuk penyakit malaria di berbagai provinsi Indonesia (Karmila et al., 2017)(Sroyer et al., 2022a). Beberapa studi sebelumnya menggunakan metode K-Means untuk analisis cluster penyakit malaria, seperti penelitian yang dilakukan di Aceh (Zohra et al., 2019) dan Papua (Sroyer et al., 2022b).

Sementara itu, penelitian ini membahas penerapan algoritma K-means dalam mengklasifikasikan kasus malaria di Indonesia dan berbagai provinsi. Data diolah dan dianalisis menggunakan algoritma K-Means untuk mengidentifikasi cluster yang terkait dengan malaria. Hasil penelitian ini diharapkan dapat memberikan wawasan yang lebih baik mengenai prevalensi dan pola malaria di Indonesia serta dapat digunakan untuk mengembangkan strategi pencegahan dan pengobatan malaria di negara ini. Penelitian ini bertujuan untuk mengidentifikasi klaster penyakit malaria terkait dan memberikan wawasan yang lebih baik mengenai distribusi dan pola malaria di Indonesia dengan menggunakan algoritma K-Means. Hasil penelitian ini diharapkan dapat digunakan untuk mengembangkan strategi pencegahan dan pengobatan malaria di tanah air, sehingga membantu mengurangi kejadian malaria di Indonesia.

Dalam konteks analisis data sosial dan kesehatan, penggunaan K-Means telah terbukti efektif dalam berbagai studi, seperti dalam mengoptimalkan performa K-Means clustering dengan PCA dalam menganalisis tingkat kemiskinan di Jawa Barat (Nurohmah et al., 2023). Dalam konteks kesehatan, terdapat penelitian yang mengidentifikasi faktor risiko stunting pada anak dengan menggunakan metode K-Means (Ranjawali et al., 2023). Selain itu, K-Means juga digunakan dalam optimalisasi strategi pemasaran dengan segmentasi pelanggan pada transaksi online retail (Febrianty et al., 2023). Dari penelitian-penelitian tersebut dapat disimpulkan bahwa penggunaan K-means dalam konteks clustering terbukti efektif dalam berbagai penelitian yang menganalisis data sosial, kesehatan, dan budaya. Oleh karena itu, penggunaan K-means dalam konteks pengelompokan malaria di Indonesia dapat menjadi pendekatan potensial untuk mengelompokkan data kasus malaria dan berkontribusi pada pemahaman dan pengendalian malaria di Indonesia..

METODE

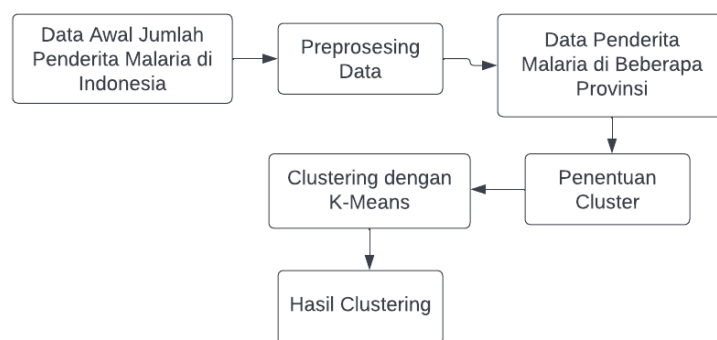
Malaria merupakan masalah kesehatan masyarakat di Indonesia, khususnya di daerah pedesaan yang rentan. Menurut Kementerian Kesehatan RI, jumlah infeksi malaria pada tahun 2015 hingga 2019 masih sangat tinggi dan diperlukan upaya pengendalian yang efektif. Salah satu solusinya adalah dengan mengelompokkan data malaria berdasarkan tiga klaster: tinggi, sedang, dan rendah 4.444. Berdasarkan data kejadian malaria tahun 2015, jumlah kejadian sebanyak 237.025 kasus, dengan jumlah kasus malaria yang terjadi pada tahun 2019 sebanyak 250.644 kasus. Adapun data pada penelitian ditunjukkan pada Tabel 1.

Tabel 1. *Data Penyakit Malaria di Indonesia*

| NO. | PROVINSI | 2015 | 2016 | 2017 | 2018 | 2019 |
|-----|----------------------|-------|-------|-------|-------|-------|
| 1 | Aceh | 80 | 50 | 60 | 20 | 20 |
| 2 | Sumatera Utara | 490 | 270 | 170 | 90 | 70 |
| 3 | Sumatera Barat | 140 | 120 | 100 | 90 | 60 |
| 4 | Riau | 100 | 30 | 30 | 10 | 20 |
| 5 | Jambi | 470 | 140 | 50 | 50 | 20 |
| 6 | Sumatera Selatan | 310 | 280 | 110 | 80 | 70 |
| 7 | Bengkulu | 2.030 | 1.450 | 530 | 160 | 40 |
| 8 | Lampung | 490 | 400 | 520 | 380 | 180 |
| 9 | Kep. Bangka Belitung | 1.080 | 110 | 70 | 160 | 90 |
| 10 | Kep. Riau | 350 | 360 | 170 | 110 | 60 |
| 11 | Dki Jakarta | 0 | 10 | 10 | 10 | 10 |
| 12 | Jawa Barat | 0 | 10 | 10 | 0 | 10 |
| 13 | Jawa Tengah | 60 | 30 | 30 | 20 | 10 |
| 14 | Di Yogyakarta | 30 | 30 | 20 | 10 | 0 |
| 15 | Jawa Timur | 0 | 10 | 0 | 10 | 20 |
| 16 | Banten | 0 | 0 | 0 | 0 | 0 |
| 17 | Bali | 0 | 0 | 10 | 20 | 10 |
| 18 | Nusa Tenggara Barat | 420 | 240 | 150 | 340 | 190 |
| 19 | Nusa Tenggara Timur | 7.040 | 5.410 | 5.760 | 3.420 | 2.370 |
| 20 | Kalimantan Barat | 130 | 60 | 30 | 20 | 0 |

| NO. | PROVINSI | 2015 | 2016 | 2017 | 2018 | 2019 |
|-----|--------------------|--------|--------|--------|--------|--------|
| 21 | Kalimantan Tengah | 420 | 190 | 290 | 250 | 70 |
| 22 | Kalimantan Selatan | 680 | 520 | 280 | 200 | 200 |
| 23 | Kalimantan Timur | 460 | 350 | 440 | 630 | 550 |
| 24 | Kalimantan Utara | 30 | 30 | 90 | 40 | 80 |
| 25 | Sulawesi Utara | 880 | 720 | 370 | 250 | 200 |
| 26 | Sulawesi Tengah | 680 | 500 | 180 | 60 | 60 |
| 27 | Sulawesi Selatan | 100 | 120 | 140 | 150 | 100 |
| 28 | Sulawesi Tenggara | 410 | 440 | 210 | 310 | 300 |
| 29 | Gorontalo | 570 | 150 | 40 | 50 | 30 |
| 30 | Sulawesi Barat | 170 | 90 | 110 | 190 | 140 |
| 31 | Maluku | 5.810 | 3.950 | 2.300 | 1.160 | 720 |
| 32 | Maluku Utara | 2.770 | 2.440 | 790 | 390 | 460 |
| 33 | Papua Barat | 31.290 | 6.790 | 14.970 | 8.490 | 7.380 |
| 34 | Papua | 31.930 | 45.850 | 59.000 | 52.990 | 64.030 |

Tidak hanya itu, pada proses pembentukan cluster dialkuakn analisis dengan membentuk 3 uji coba yaitu K=2, K=3 dan K=5. Uji coba tersebut diterapkan pada Algoritma K-Means. Setelah itu dilakukan evaluasi Davies Bouldin Index (DBI) dengan menentukan nilai K yang terbaik yang mendekati nilai nol, untuk proses clustering dengan menggunakan metode Davies Bouldin Index (DBI). Setelah diketahui nilai DBI nilai cluster, maka nilai terkecil dari hasil cluster yang terbaik dan menunjukkan nilai cluster yang optimal (Febrianty et al., 2023). Adapun Alur kerja penelitian ini ditunjukkan pada Gambar 1.



Gambar 1. Alur Penelitian

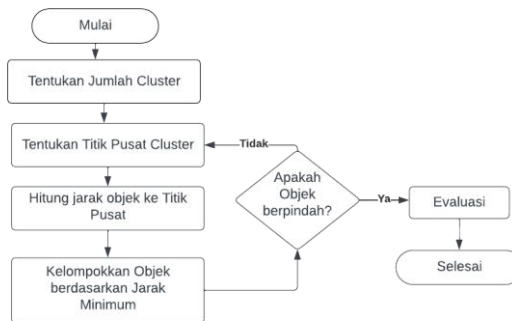
Berdasarkan Gambar 1, terdapat preprosesing data yang digunakan untuk menghilangkan beberapa data yang bisa menjadi permasalahan saat pemrosesan data. Salah satu prosesnya yaitu pembersihan data, membersihkan data dari data yang tidak valid, duplikat, dan data yang hilang atau bernilai nol. Data hasil preprosesing ditunjukkan pada Tabel 2.

Tabel 2. Data Penyakit Malaria di Indonesia Setelah Tahap Preprosesing

| NO. | PROVINSI | 2015 | 2016 | 2017 | 2018 | 2019 |
|-----|----------------------|------|------|------|------|------|
| 1 | ACEH | 80 | 50 | 60 | 20 | 20 |
| 2 | SUMATERA UTARA | 490 | 270 | 170 | 90 | 70 |
| 3 | SUMATERA BARAT | 140 | 120 | 100 | 90 | 60 |
| 4 | RIAU | 100 | 30 | 30 | 10 | 20 |
| 5 | JAMBI | 470 | 140 | 50 | 50 | 20 |
| 6 | SUMATERA SELATAN | 310 | 280 | 110 | 80 | 70 |
| 7 | BENGKULU | 2030 | 1450 | 530 | 160 | 40 |
| 8 | LAMPUNG | 490 | 400 | 520 | 380 | 180 |
| 9 | KEP. BANGKA BELITUNG | 1080 | 110 | 70 | 160 | 90 |
| 10 | KEP. RIAU | 350 | 360 | 170 | 110 | 60 |
| 11 | JAWA TENGAH | 60 | 30 | 30 | 20 | 10 |
| 12 | NUSA TENGGARA BARAT | 420 | 240 | 150 | 340 | 190 |
| 13 | NUSA TENGGARA TIMUR | 7040 | 5410 | 5760 | 3420 | 2370 |
| 14 | KALIMANTAN TENGAH | 420 | 190 | 290 | 250 | 70 |
| 15 | KALIMANTAN SELATAN | 680 | 520 | 280 | 200 | 200 |
| 16 | KALIMANTAN TIMUR | 460 | 350 | 440 | 630 | 550 |

| NO. | PROVINSI | 2015 | 2016 | 2017 | 2018 | 2019 |
|-----|-------------------|-------|-------|-------|-------|-------|
| 17 | KALIMANTAN UTARA | 30 | 30 | 90 | 40 | 80 |
| 18 | SULAWESI UTARA | 880 | 720 | 370 | 250 | 200 |
| 19 | SULAWESI TENGAH | 680 | 500 | 180 | 60 | 60 |
| 20 | SULAWESI SELATAN | 100 | 120 | 140 | 150 | 100 |
| 21 | SULAWESI TENGGARA | 410 | 440 | 210 | 310 | 300 |
| 22 | GORONTALO | 570 | 150 | 40 | 50 | 30 |
| 23 | SULAWESI BARAT | 170 | 90 | 110 | 190 | 140 |
| 24 | MALUKU | 5810 | 3950 | 2300 | 1160 | 720 |
| 25 | MALUKU UTARA | 2770 | 2440 | 790 | 390 | 460 |
| 26 | PAPUA BARAT | 31290 | 6790 | 14970 | 8490 | 7380 |
| 27 | PAPUA | 31930 | 45850 | 59000 | 52990 | 64030 |

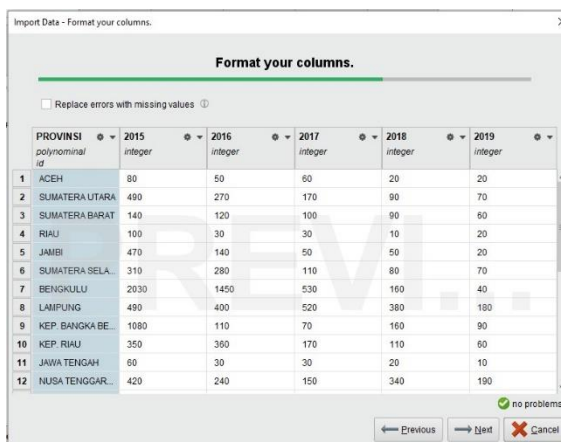
Kemudian dilakukan proses penentuan cluster. Berdasarkan data pada Tabel 1, dilakukan pengelompokan data dengan membagi menjadi 3K, yaitu : 1). K=2 : cluster_0 dan cluster_1; 2) K=3 : cluster_0, cluster_1, cluster_2; dan 3) K=5 : cluster_0, cluster_1, cluster_2, cluster_3 dan cluster_4. Selanjutnya dilakukan proses clustering menggunakan K-Means. Adapun alur dari K-Means ditunjukkan pada Gambar 2.



Gambar 2. Flowchart Proses K-Means Clustering

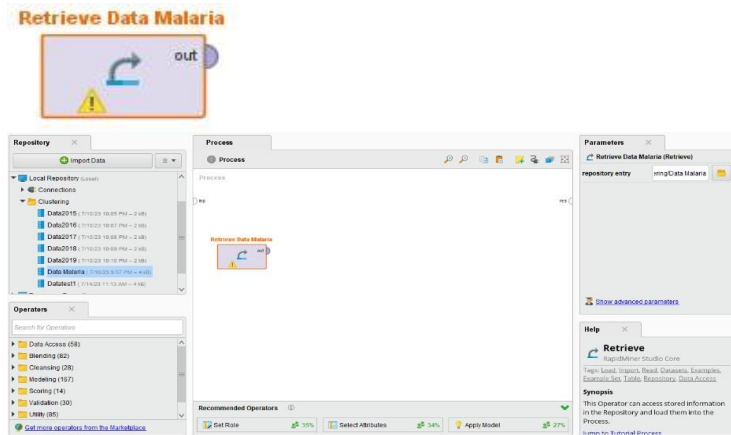
HASIL DAN PEMBAHASAN

Penerapan clustering pada penelitian ini menggunakan algoritma K-means. Data yang digunakan merupakan data persebaran penyakit malaria di 27 provinsi yang ada di Indonesia dari tahun 2015 sampai dengan 2019 yang diperoleh dari Badan Pusat Statistik (BPS). Berdasarkan data pada tabel 2 diatas data malaria tersebut dilakukan pengujian evaluasi dengan DBI (Davies Bouldin Index) dengan tujuan untuk mengetahui jumlah cluster terbaik. Cluster terbaik ditunjukkan dari nilai DBI yang mendekati nilai nol atau lebih kecil. Sehingga setelah itu akan dipilih cluster yang terbaik, pengujian tersebut dibagi menjadi tiga dengan cluster yaitu K=2, K=3 dan K=5. Implementasi clustering pada penelitian ini menggunakan Rapid Miner. Data penelian yang sudah dilakukan tahap preprosesing, diimport pada Rapid Miner yang ditunjukkan pada Gambar 3.



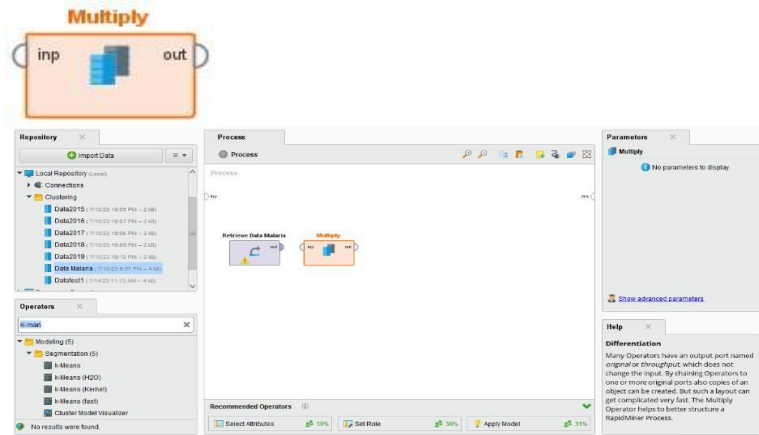
Gambar 3. Hasil Import Data Kasus Penderita Malaria

Selanjutnya dilakukan proses analisis data yaitu dengan drag and drop data (retrieve) yang akan digunakan dari Repository pada halaman Design Process, yang ditunjukkan pada gambar 4.



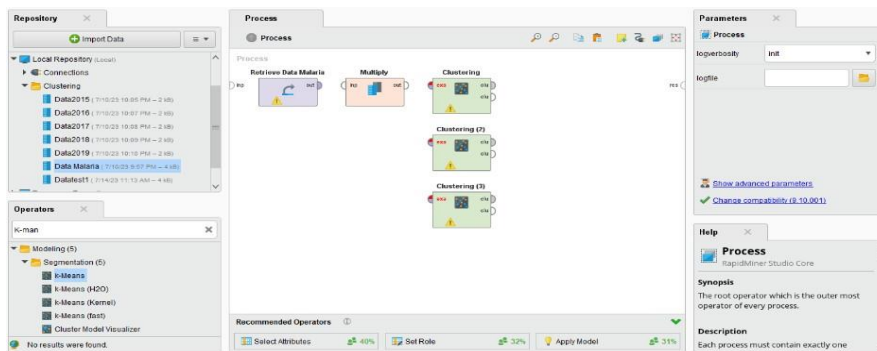
Gambar 4. Design Process Retrieve Data

Memasukkan operator Multiplay yang berfungsi sebagai operator penghubung pada model yang digunakan, ditunjukkan pada Gambar 5.

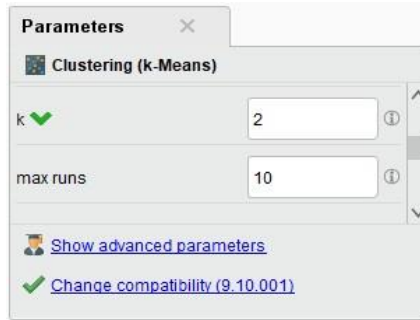


Gambar 5. Multiply Design Process

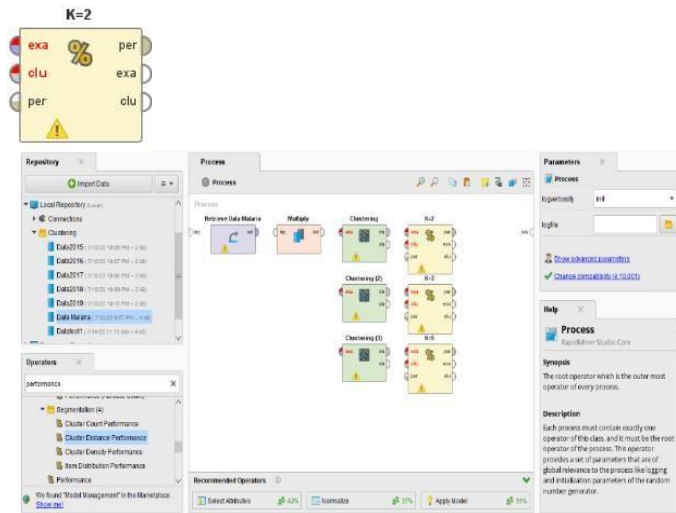
Memasukkan operator Algoritma K-Means, menguji pada K (K=2, K=3 dan K=5). Parameter disesuaikan dengan jumlah K yang diuji, ditunjukkan pada Gambar 6. Nilai parameter diberikan pada setiap dengan berbeda-beda untuk menguji nilai yang ditunjukkan pada Gambar 7. Kemudian memasukkan operator performa (K=2, K=3 dan K=5) parameter kriteria (metode evaluasi) untuk pengujian K yang ditunjukkan pada Gambar 8.



Gambar 6. Modeling Clustering

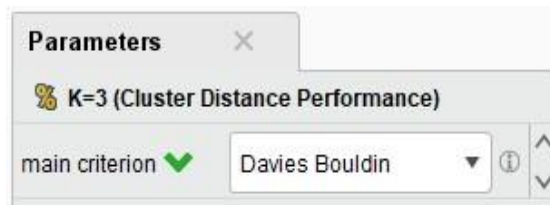


Gambar 7. Parameter

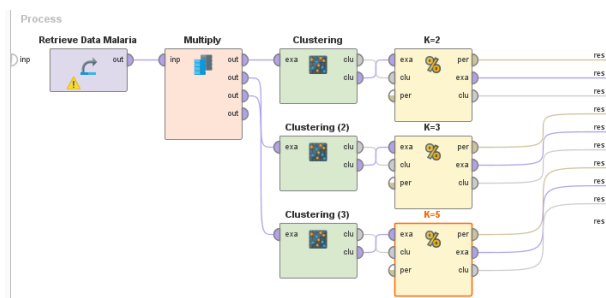


Gambar 8. Performance Operator

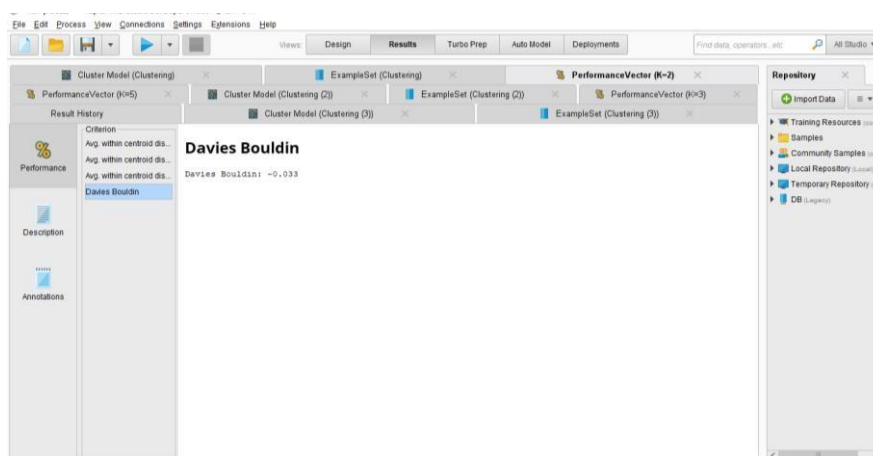
Langkah selanjutnya yaitu menambahkan parameter kriteria untuk evaluasi (davies bouldin index) yang ditunjukkan pada gambar 9. Setelah parameter diberikan nilai, lalu hubungkan semua operator pada bagian process yang digunakan. Setelah itu klik start untuk memulai pengujian, ditunjukkan pada Gambar 10. Hasil pengujian data diperoleh Setelah klik Result, ditunjukkan pada Gambar 11.



Gambar 9. Parameter Kriteria



Gambar 10. Hubungan Operator



Gambar 11. Halaman Result

Hasil pengujian menggunakan evaluasi Davies Bouldin Index ditunjukkan dalam Tabel 3.

Tabel 3. Nilai Evaluasi Davies Bouldin Index

| Jumlah Cluster | Evaluasi Davies Bouldin Index |
|----------------|-------------------------------|
| K=2 | 0.033 |
| K=3 | 0.034 |
| K=5 | 0.262 |

Dari tabel 3.4 maka. Jumlah cluster terbaik data terbagi menjadi 2 kelompok cluster. Evaluasi menggunakan Davies Bouldin Index ketentuan semakin kecil nilainya semakin baik kelompok cluster tersebut. Berdasarkan pengelompokan menjadi 2 cluster menggunakan algoritma K-Means diketahui bahwa jumlah provinsi cluster 1 (cluster_0) sebanyak 26 provinsi dan cluster 2 (cluster_1) ada 1 provinsi. Indeks cluster pada RapidMiner dimulai dari angka nol. Pembagian K=2 merupakan cluster terbaik. Hasil analisis titik centroid kedua cluster ditunjukkan pada tabel 4, dan provinsi hasil clustering ditunjukkan pada tabel 5

Tabel 4. Titik Centroid K-Means dengan K=2

| Tahun | Cluster_0 | Cluster_1 |
|-------|-----------|-----------|
| 2015 | 2205 | 31930 |
| 2016 | 968 | 45850 |
| 2017 | 1075 | 59000 |
| 2018 | 657 | 52990 |
| 2019 | 518 | 64030 |

Tabel 5. Data Provinsi K=2 Setiap Cluster

| K=2 | Cluster_0 | Cluster_1 |
|-----|------------------|-----------|
| | Aceh | Papua |
| | Sumatera Utara | |
| | Sumatera Barat | |
| | Riau | |
| | Jambi | |
| | Sumatera Selatan | |
| | Bengkulu | |
| | Lampung | |

| |
|----------------------|
| Kep. Bangka Belitung |
| Kep. Riau |
| Jawa Tengah |
| Nusa Tenggara Barat |
| Nusa Tenggara Timur |
| Kalimantan Tengah |
| Kalimantan Selatan |
| Kalimantan Timur |
| Kalimantan Utara |
| Sulawesi Utara |
| Sulawesi Tengah |
| Sulawesi Selatan |
| Sulawesi Tenggara |
| Gorontalo |
| Sulawesi Barat |
| Maluku |
| Maluku Utara |
| Papua Barat |

Berdasarkan tabel 5, diketahui bahwa cluster_1 merupakan cluster dengan kasus malaria terbanyak. Disisi lain, pada cluster_0 memiliki jumlah kasus malaria lebih sedikit. Berdasarkan hasil clsturing K=3, diketahui bahwa yang dicluster menjadi 3 terdiri dari cluster_0 (cluster 1) sebanyak 25 provinsi, cluster_1 (cluster 1) sebanyak 1 provinsi dan cluster_2 (cluster 3) sebanyak 1 provinsi. setelah itu, hasil analisis titik centroid kedua cluster ditunjukkan pada tabel 6, dan provinsi hasil clustering ditunjukkan pada tabel 7.

Tabel 6 Titik Centroid K-Means dengan K=3

| Tahun | Cluster_0 | Cluster_1 | Cluster_2 |
|-------|-----------|-----------|-----------|
| 2015 | 1041 | 31930 | 31290 |
| 2016 | 735 | 45850 | 6790 |
| 2017 | 519 | 59000 | 14970 |
| 2018 | 344 | 52990 | 8490 |
| 2019 | 244 | 64030 | 7380 |

Tabel 7 Data Provinsi K=3 Setiap Cluster

| K=3 | Cluster_0 | Cluster_1 | Cluster_2 |
|-----|----------------------|-----------|-------------|
| | Aceh | Papua | Papua Barat |
| | Sumatera Utara | | |
| | Sumatera Barat | | |
| | Riau | | |
| | Jambi | | |
| | Sumatera Selatan | | |
| | Bengkulu | | |
| | Lampung | | |
| | Kep. Bangka Belitung | | |
| | Kep. Riau | | |
| | Jawa Tengah | | |
| | Nusa Tenggara Barat | | |
| | Nusa Tenggara Timur | | |
| | Kalimantan Tengah | | |

Kalimantan Selatan
 Kalimantan Timur
 Kalimantan Utara
 Sulawesi Utara
 Sulawesi Tengah
 Sulawesi Selatan
 Sulawesi Tenggara
 Gorontalo
 Sulawesi Barat
 Maluku
 Maluku Utara

Berdasarkan Tabel 7 diketahui bahwa Cluster 2 (cluster_1) merupakan kelompok provinsi kasus malaria terbanyak di Indonesia, diikuti dengan Cluster 3 (cluster_2) dan Cluster (1cluster_0) sebagai jumlah kasus malaria paling sedikit.

Pada clustering data malaria berdasarkan provinsi menjadi 5 kelompok. Pada hasil clustering K=5, diketahui bahwa jumlah data malaria yang dikelompokkan menjadi 5 cluster terdiri dari Cluster 1 (cluster_0) sebanyak 22 provinsi. Cluster 2 (Cluster_2) sebanyak 1 provinsi. Cluster 3 (cluster_2) sebanyak 1 provinsi. Cluster 4 (clus- ter_3) sebanyak 1 provinsi. Cluster 5 (cluster_4) sebanyak 2 provinsi. Indeks cluster pada rapidminer dimulai dari angka nol. Kemudian, analisis titik centroid kedua cluster ditunjukkan pada Tabel 8 dan provinsi hasil clustering ditunjukkan pada Tabel 9

Tabel 8. Titik Centroid K-Means dengan K=5

| Tahun | Cluster_0 | Cluster_1 | Cluster_2 | Cluster_3 | Cluster_4 |
|-------|-----------|-----------|-----------|-----------|-----------|
| 2015 | 473 | 31930 | 31290 | 7040 | 4290 |
| 2016 | 299 | 45850 | 6790 | 5410 | 3195 |
| 2017 | 188 | 59000 | 14970 | 5760 | 1545 |
| 2018 | 165 | 52990 | 8490 | 3420 | 775 |
| 2019 | 116 | 64030 | 7380 | 2370 | 590 |

Tabel 9. Data Provinsi K=5 Setiap Cluster

| K=5 | Cluster_0 | Cluster_1 | Cluster_2 | Cluster_3 | Cluster_4 |
|-----|----------------------|-----------|-------------|-----------|--------------|
| | Aceh | Papua | Papua Barat | Ntt | Maluku |
| | Sumatera Utara | | | | Maluku Utara |
| | Sumatera Barat | | | | |
| | Riau | | | | |
| | Jambi | | | | |
| | Sumatera Selatan | | | | |
| | Bengkulu | | | | |
| | Lampung | | | | |
| | Kep. Bangka Belitung | | | | |
| | Kep. Riau | | | | |
| | Jawa Tengah | | | | |
| | Nusa Tenggara Barat | | | | |
| | Kalimantan Tengah | | | | |
| | Kalimantan Selatan | | | | |

Kalimantan Timur
Kalimantan Utara
Sulawesi Utara
Sulawesi Tengah
Sulawesi Selatan
Sulawesi Tenggara
Gorontalo
Sulawesi Barat

KESIMPULAN

Penelitian analisis klusterisasi penyakit malaria dengan metode K-Means di Indonesia menyimpulkan bahwa metode ini efektif dalam mengkategorikan dan mengelompokkan data penyakit malaria untuk memahami pola penyebarannya di berbagai wilayah Indonesia. Hasil penelitian ini memungkinkan identifikasi pola penyebaran yang berbeda-beda, memungkinkan tindakan pencegahan yang lebih tepat sasaran dan intervensi yang efektif. Dengan demikian, penelitian ini memberikan kontribusi penting dalam upaya pengendalian dan pencegahan penyakit malaria di Indonesia. Pengujian penelitian ini menggunakan Davies Bouldin Index (DBI), untuk menentukan banyaknya cluster terbaik dari K=2, K=3, dan K=5. Untuk K=2 menunjukkan nilai DBI = 0,033, untuk K=3 menunjukkan nilai DBI = 0,034, dan untuk K=5 menunjukkan nilai DBI = 0,262. Sehingga nilai cluster terbaik ditunjukkan pada nilai K=2, karena memiliki nilai yang minimal.

DAFTAR RUJUKAN

- Bete, Y., Santos, D., Lani, R., Ewal, A., & Lenggu, B. J. (2023). *Menentukan Titik Rawan Malaria Di Provinsi Nusa Tenggara Timur Menggunakan Metode K-Means Clustering*. 1(4).
- Febrianty, E., Awalina, L., & Rahayu, W. I. (2023). Optimalisasi Strategi Pemasaran dengan Segmentasi Pelanggan Menggunakan Penerapan K-Means Clustering pada Transaksi Online Retail Optimizing Marketing Strategies with Customer Segmentation Using K-Means Clustering on Online Retail Transactions. *Jurnal Teknologi Dan Informasi (JATI)*, 13(September), 122–137. <https://doi.org/10.34010/jati.v13i2>
- Kalua, A. L., Veronika H, & Salaki, D. T. (2022). Sistem Pakar Diagnosa Penyakit Malaria dengan Certainty Factor dan Forward Chaining. *Journal of Information Technology, Software Engineering and Computer Science (ITSECS)*, 1(1), 22–34. <https://doi.org/10.58602/itsecs.v1i1.10>
- Karmila, Tambunan, H. S., Sumarno, & Windarto, A. P. (2017). Penerapan Data Mining K-Means dalam Mengelompokkan Kasus Penyakit Malaria Berdasarkan Provinsi dengan Aplikasi RapidMiner. *Regional Development Industry & Health Science, Technology and Art of Life*, 31–40.
- Nurohmah, Y., Mayasari, R., & Nurina Sari, B. (2023). Optimalisasi Performa K-Means Clustering Dengan Pca Dalam Analisis Tingkat Kemiskinan Di Jawa Barat. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 7(3), 1657–1665. <https://doi.org/10.36040/jati.v7i3.6884>
- Ranjawali, R., Talakua, A. C., & Abineno, R. T. (2023). Clustering Stunting Pada Balita Dengan Metode K-Means Di Puskesmas Kanatang. *SATI: Sustainable Agricultural Technology Innovation*, 80–92.
- Sari, E., & Syakurah, R. A. (2023). Analisis Manajemen Pelatihan Kader Malaria Pada Populasi Suku Anak Dalam Di Kabupaten Musi Rawas Utara. *J.Abdimas: Community Health*, 4(1), 01–08. <https://doi.org/10.30590/jach.v4n1.582>
- Sindi, S., Ningse, W. R. O., Sihombing, I. A., R.H.Zer, F. I., & Hartama, D. (2020). Analisis Algoritma K-Medoids Clustering Dalam Pengelompokan Penyebaran Covid-19 Di Indonesia. *Jurnal Teknologi Informasi*, 4(1), 166–173. <https://doi.org/10.36294/jurti.v4i1.1296>
- Sroyer, A. M., Mandowen, S. A., & Reba, F. (2022a). Analisis Cluster Penyakit Malaria Provinsi Papua Menggunakan Metode Single Linkage Dan K-Means. *Jurnal Nasional Teknologi Dan Sistem Informasi*, 7(3), 147–154. <https://doi.org/10.25077/teknosi.v7i3.2021.147-154>

- Sroyer, A. M., Mandowen, S. A., & Reba, F. (2022b). Analisis Cluster Penyakit Malaria Provinsi Papua Menggunakan Metode Single Linkage Dan K-Means. *Jurnal Nasional Teknologi Dan Sistem Informasi*, 7(3), 147–154. <https://doi.org/10.25077/teknosi.v7i3.2021.147-154>
- Yohannes, Y., Devella, S., & Arianto, K. (2020). Deteksi Penyakit Malaria Menggunakan Convolutional Neural Network Berbasis Saliency. *JUITA: Jurnal Informatika*, 8(1), 37. <https://doi.org/10.30595/juita.v8i1.6671>
- Zohra, A. F., Anwar, S., Fitri, A., & Nasution, M. H. (2019). Klasifikasi Wilayah Provinsi Aceh Berdasarkan Tingkat Kerentanan Kasus Malaria Tahun 2015 – 2018. *Jurnal Kesehatan Lingkungan Indonesia*, 18(1), 25. <https://doi.org/10.14710/jkli.18.1.25-33>