

KLASTERISASI PENYAKIT MENGGUNAKAN ALGORITMA K-MEDOIDS PADA DINAS KESEHATAN KABUPATEN AGAM PROVINSI SUMATERA BARAT

Minarni¹, Elsa Indah Sari², Anna Syahrani³, Putri Mandarani²

^{1,2,3}Teknik Informatika, Institut Teknologi Padang, Indonesia

e-mail: minarni1706@gmail.com¹, elsaindahsari5@gmail.com²,
annasyahrani@gmail.com³, pmandarani2@gmail.com⁴

Abstrak

Jumlah kasus penyebaran penyakit *tuberculosis* (TB), *pneumonia*, dan demam berdarah *dengue* (DBD) pada Dinas Kesehatan Kabupaten Agam setiap tahun mengalami peningkatan. Kondisi alam kabupaten Agam dengan wilayah yang luas dan perbukitan menyebabkan sulitnya memantau penyebaran penyakit berdasarkan kasus terbanyak. Hal ini menyebabkan lambatnya pencegahan maupun penanggulangan penyakit tersebut. Untuk itu diperlukan cara pengelompokan atau klusterisasi penyakit berdasarkan jumlah kasus yang banyak terjadi. Penelitian ini bertujuan untuk menerapkan algoritma K-Medoids untuk pengelompokan data penyakit berdasarkan jumlah kasus. Algoritma ini merupakan metode klusterisasi pembatas untuk mengelompokkan sekumpulan n objek menjadi beberapa k kluster. Data yang digunakan berupa data penderita penyakit DBD, TB, dan *pneumonia* dari tahun 2016-2019 sebanyak 370 *record*. Hasil pengujian diperoleh anggota masing-masing kluster untuk tiap-tiap penyakit yaitu *tuberculosis* kluster tinggi 32 dan rendah 60, penyakit *pneumonia* kluster tinggi 33 dan rendah 59, sedangkan untuk penyakit DBD kluster tinggi 10 anggota dan rendah 82 anggota. Pengujian menggunakan perhitungan manual dan aplikasi Rapidminer mendapatkan hasil yang sama dengan sistem. Ini menunjukkan bahwa sistem telah bekerja dengan baik.

Kata kunci: Algoritma K-Medoids, Klusterisasi, Penyakit

Abstract

The number of cases of tuberculosis (TB), pneumonia, and dengue fever (DBD) at the Agam District Health Office has increased every year. The Agam Regency with its wide area and hills makes it difficult to monitor the spread of the disease. This causes prevention and response to be delayed. For that, it is necessary to group or cluster diseases based on cases. The purpose of this study is to apply the K-Medoids algorithm for grouping disease data. This algorithm is a constraint clustering method to group a set of n objects into several k clusters. The data used are data on patients with dengue fever, TB, and pneumonia from 2016-2019 as many as 370 records. The test results obtained members of each cluster for each disease, namely tuberculosis cluster high 32 and low 60, pneumonia disease cluster high 33 and low 59, while for high cluster DBD 10 members and low 82 members. Tests using manual calculations and the Rapidminer application get the same results as the system. This indicates that the system is working properly.

Keywords : K-Medoids Algorithm, Clustering, disease

Diterima Redaksi: 17-6-2021 | Selesai Revisi: 20-10-2021 | Diterbitkan Online: 31-12-2021

DOI: <https://doi.org/10.23887/janapati.v10i3.34904>

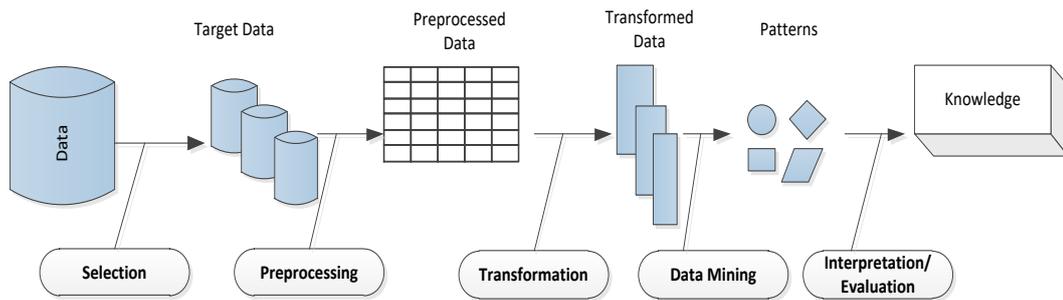
PENDAHULUAN

Kabupaten Agam mempunyai luas 2.232,30 Km² atau 5,29% dari seluruh wilayah Provinsi Sumatera Barat. Kabupaten Agam mempunyai 16 Kecamatan dan 82 Nagari serta 467 Jorong. Jumlah penduduk Kabupaten Agam berdasarkan data BPS tahun 2018 sebanyak 487.914 jiwa[1]. Dinas Kesehatan Kabupaten

Agam merupakan unsur pemerintah Kabupaten yang bertanggung jawab dalam bidang kesehatan meliputi kesehatan masyarakat, pencegahan dan pengendalian penyakit, pelayanan, dan sumber daya kesehatan. Dinas Kesehatan Kabupaten Agam memiliki 23 puskesmas dan 2 rumah sakit yang tersebar di 16 Kecamatan di wilayah Kabupaten Agam.

Berdasarkan data Dinas Kesehatan Kabupaten Agam tahun 2016-2019, jumlah kasus penyebaran penyakit *tuberculosis* (TB), *pneumonia*, dan demam berdarah *dengue* (DBD) mengalami peningkatan. Dengan wilayah yang luas dan kondisi alam perbukitan menyebabkan Dinas Kesehatan Kabupaten Agam sulit untuk memantau penyebaran penyakit dengan jumlah kasus terbanyak. Hal ini menyebabkan lambatnya pencegahan maupun penanggulangan penyakit tersebut. Untuk itu diperlukan cara pengelompokan atau klusterisasi penyakit berdasarkan jumlah kasus yang banyak terjadi.

Data mining menjadi salah satu pilihan untuk menyelesaikan masalah pengelompokan data, karena data mining merupakan proses untuk menemukan pola baru dari sekumpulan data yang sangat besar [2]. Hasil dari proses ini dapat digunakan dalam menentukan keputusan. Di dalam data mining terdapat metode yang digunakan untuk menemukan pola atau pengetahuan baru dikenal dengan *Knowledge Discovery In Database* (KDD) [3]. Proses KDD seperti pada gambar 1.



Gambar 1. Proses Penemuan Pengetahuan dalam Database (KDD)

Proses di atas dapat dijelaskan sebagai berikut:

- a. **Selection**
Selection merupakan langkah awal dalam menentukan variabel yang akan digunakan tidak memiliki kemiripan dan duplikasi.
- b. **Preprocessing**
Pra pengolahan (*preprocessing*) data yang terdiri dari data *cleaning* dan data *integration*.
- c. **Transformation**
Transformasi merupakan proses merubah data menyesuaikan dengan format ekstensi yang diperlukan pada pengolahan data mining.
- d. **Data mining** merupakan proses yang utama untuk mendapatkan pengetahuan baru dari data yang diproses, salah satunya *Clustering*[4]. *Clustering* atau klusterisasi digunakan untuk mengelompokkan data-data ke dalam beberapa kelompok berdasarkan tingkat kemiripan[5]. Dengan tujuan untuk membagi objek (orang, benda, peristiwa, dan lain-lain) ke dalam suatu kelompok, sehingga objek-objek dalam satu kelompok (klaster) memiliki kesamaan yang tinggi namun memiliki perbedaan dengan objek pada kelompok lainnya [6].

- Salah satu algoritma *clustering* yang digunakan yaitu algoritma k-medoids. Algoritma ini merupakan pengembangan dari algoritma k-Means yang bertujuan untuk mengurangi tingkat kepekaan pada bagian yang dihasilkan berkaitan dengan nilai-nilai yang terdapat dalam kumpulan data, dimana memungkinkan adanya penyimpangan distribusi data disebabkan oleh objek dengan nilai yang besar [7]. Algoritma ini merupakan metode klusterisasi partisi untuk mengelompokkan sekumpulan n objek menjadi beberapa k klaster. Objek yang berada pada kumpulan objek mewakili sebuah klasternya, dikenal dengan medoid. Kemudian klaster dihitung dengan cara menghitung kedekatan antara medoid dengan objek non medoid. Medoids digunakan tidak berdasarkan dari pengamatan nilai rata-rata (*mean*) yang dimiliki oleh setiap kelompok [8].
- e. **Interpretasi** merupakan proses mengidentifikasi pola-pola yang unik ke dalam basis pengetahuan yang diidentifikasi.

Beberapa penelitian tentang penerapan algoritma k-medoids dalam bidang kesehatan telah dilakukan diantaranya pengelompokan penyebaran diare [5]. Penelitian ini menghasilkan pengelompokan wilayah di Kota Medan menjadi dua kluster berdasarkan jumlah penderita diare. Selanjutnya penelitian klasterisasi data penyakit pasien di RSUD Kota Bandung [9]. Penelitian ini membahas tentang pengelompokan penyakit yang diderita pasien dikelompokkan menjadi 3 kluster berdasarkan kamar dan jumlah pasien, yaitu kluster 0 untuk penyakit pada klinik gawat darurat sebanyak 18 data, kluster 1 untuk penyakit pada klinik kemuning dan klinik gawat darurat, dan kluster 3 untuk klinik rehab medik sebanyak 20 data.

Penelitian berikutnya tentang pengelompokan kasus penyakit AIDS berdasarkan provinsi [10]. Penelitian ini mengelompokkan kasus penyakit AIDS menjadi dua kluster berdasarkan parameter rata-rata jumlah kasus yang terjadi dan rata-rata jumlah kasus penyakit AIDS yang meninggal. Penelitian untuk pengelompokan penyakit di Pekanbaru Riau [11]. Penelitian ini membahas tentang penyebaran penyakit yang paling banyak diderita dengan cara mengelompokkan menjadi 4 kluster. Penelitian lainnya pengelompokan data imunisasi campak balita di Indonesia [12]. Penelitian ini membahas tentang pengelompokan data imunisasi campak balita, dengan mengelompokkan data imunisasi menjadi 3 kluster, yaitu rendah, sedang, dan tinggi. Penelitian clustering pasien kanker berdasarkan struktur protein dalam tubuh [7], yang membahas tentang pengelompokan jenis kanker berdasarkan kelasnya. Hasil penelitian ini adalah sebuah sistem pengelompokan jenis kanker, dengan nilai persentase kualitas kluster sebesar 77% pada percobaan nilai k 14 menggunakan 116 data.

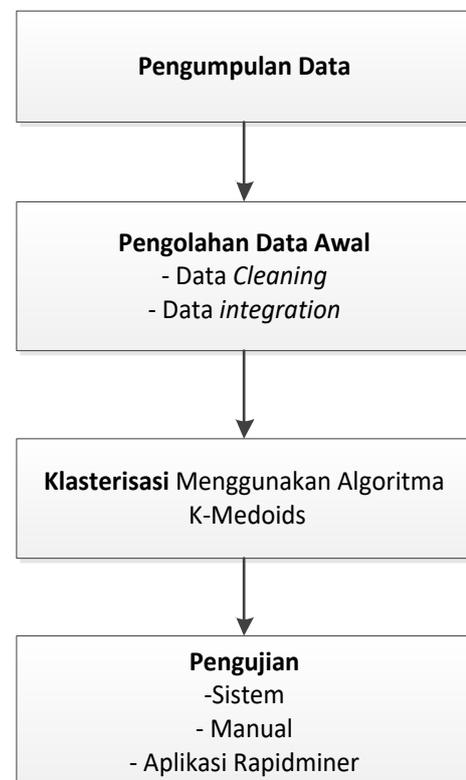
Penelitian tentang pengelompokan data penyakit alergi pada anak dengan mengimplementasikan algoritma k-Medoids membagi kluster menjadi 3, yaitu rendah 21 provinsi, sedang 12 provinsi, dan 1 provinsi untuk kluster tinggi [13]. Penelitian tentang pengelompokan penyebaran covid-19 di Indonesia berdasarkan gejala suhu badan di atas 36,9⁰ C, demam, dan batuk, menghasilkan 3 kluster [14]. Penelitian tentang pengelompokan data pasien kardiovaskuler menggunakan k-Medoids dengan kombinasi teknik reduksi data dengan PCA berdasarkan tingkat komplikasi pasien pada setiap cluster data yang dihasilkan [15]. Penelitian tentang pengelompokan wilayah sebaran cacat pada anak, mendapatkan hasil bahwa algoritma k-

medoids lebih baik dalam mengelompokkan dibandingkan dengan algoritma k-means [16]. Selanjutnya pengelompokan imunisasi lanjutan pada anak usia 2 tahun dengan hasil 3 kluster [17].

Penelitian ini bertujuan untuk menerapkan algoritma k-Medoids untuk pengelompokan data penyakit berdasarkan jumlah kasus pada Dinas Kesehatan Kabupaten Agam. Output yang dihasilkan berupa kelompok atau kluster jumlah kasus tinggi dan rendah untuk penyakit demam berdarah, tuberculosi, dan pneumonia. Hasil pengelompokan kluster penyakit dapat digunakan sebagai bahan pertimbangan dalam mengambil kebijakan untuk pencegahan dan penanggulangan penyakit pada Dinas Kesehatan Kabupaten Agam.

METODE

Penelitian ini dilaksanakan berdasarkan tahapan yang ditampilkan pada gambar 2.



Gambar 2. Tahapan Penelitian

Tahapan-tahapan tersebut dapat dijelaskan sebagai berikut.

1. Pengumpulan Data

Penelitian ini menggunakan data dari Dinas Kesehatan Kabupaten Agam tahun 2016-

2019, berupa data jumlah kasus *tuberculosis* (TB), demam berdarah *dengue* (DBD), dan *pneumonia*.

2. Pengolahan Data Awal

Pengolahan data awal dimulai dengan tahapan pembersihan data (*cleaning data*) dan proses integrasi yang bertujuan untuk menghilangkan nilai-nilai yang keliru atau tidak terisi, duplikasi data, serta memperbaiki data yang tidak konsisten. Data set yang didapatkan adalah sebanyak 370 data penyebaran penyakit. Dari data penyakit yang mempunyai *field* seperti no, kecamatan, puskesmas, sesuai standar, tidak sesuai standar, jumlah pelayanan, laki-laki, perempuan, total kasus, *tuberculosis* anak, jumlah kasus, lintang, bujur dan tahun. Setelah melakukan proses pembersihan data maka data set yang di dapat adalah 368 *record*. Kemudian proses *selection*, dimana proses ini memilih *sampling* data yang akan digunakan sebanyak 368 *record*, yang awalnya 370 *record*, ini disebabkan karena ada 2 data yang nilai dan *valuenya* tidak konsisten. Data yang digunakan berupa data puskesmas, kecamatan, dan jumlah kasus pada penyakit. Hasil *cleaning* dapat dilihat pada tabel 1.

Tabel 1. Hasil *Cleaning* Data Penyakit DBD

| No | Kecamatan | Puskesmas | Kasus | Tahun |
|-----|-----------------|--------------|-------|-------|
| 1 | Baso | Baso | 28 | 2019 |
| 2 | Baso | Padang Tarok | 9 | 2019 |
| 3 | Ampek Angkek | Biaro | 6 | 2019 |
| ... | | | ... | ... |
| ... | | | ... | ... |
| 91 | Tanjung Mutiara | Tiku | 0 | 2016 |

| | | Mutiara | | |
|----|-----------------|--------------|---|------|
| 92 | Tanjung Mutiara | Muaro Putuih | 2 | 2016 |

3. Klasterisasi Menggunakan Algoritma K-Medoids

Langkah-langkah pengelompokan atau klasterisasi penyakit pada gambar 3 [12],

a. Menentukan Jumlah Klaster (k)

Metode yang digunakan untuk menentukan jumlah klaster yaitu Metode Elbow. Metode ini mempertimbangkan nilai perbandingan *Sum of Square Error* (SSE) untuk setiap nilai klasternya. Tujuan dari metode elbow adalah untuk memilih nilai k yang kecil dan masih memiliki SSE yang rendah [18].

SSE dihitung menggunakan rumus berikut:

$$SSE = \sum_{k=1}^k \sum_{xi} |x_i - c_k| \quad (1)$$

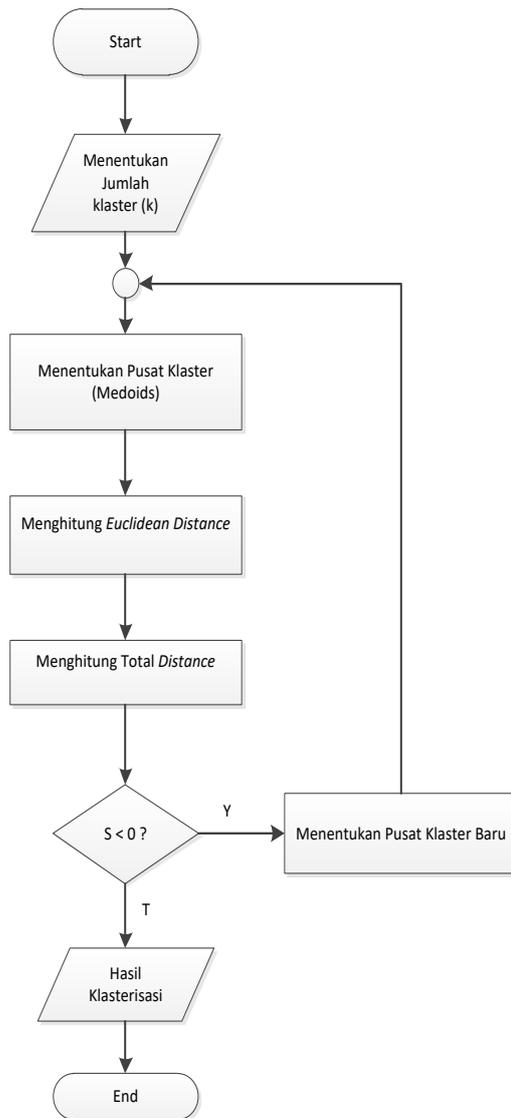
b. Menentukan pusat klaster yang dipilih secara acak.

c. Menghitung *Euclidean Distance* menggunakan rumus berikut.

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

d. Menentukan pusat klaster baru secara acak dan dilanjutkan menghitung *Euclidean Distance*.

e. Menghitung Total Simpangan (S) yaitu selisih total *cost* baru dengan total *cost* lama. Jika $S < 0$, lakukan langkah b sampai e. Dan jika $S > 0$ maka ditemukan hasil klasterisasi.



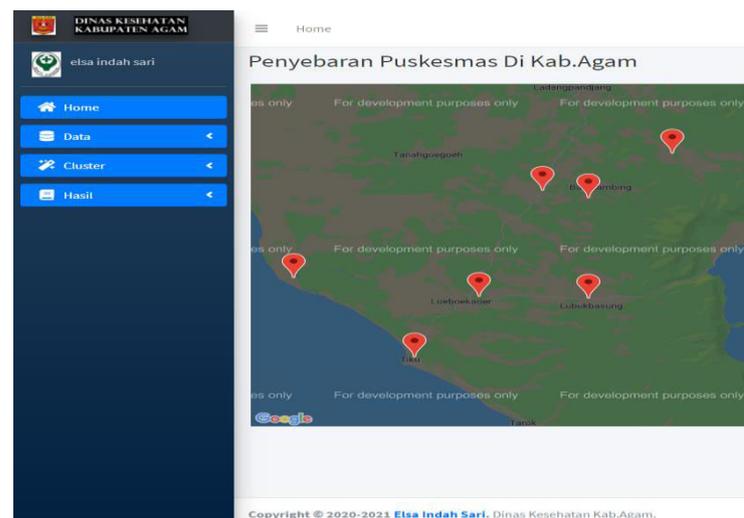
Gambar 3. Flowchart Algoritma K-Medoids

4. Pengujian

Pada tahapan ini, pengujian klasterisasi penyakit dilakukan dengan menggunakan sistem yang dibangun, perhitungan secara manual, dan menggunakan aplikasi *Rapidminer*. Aplikasi ini merupakan salah satu tools untuk melakukan analisis terhadap data mining, text mining, dan analisis prediksi [17]. Pengujian bertujuan untuk mengetahui hasil kinerja klasterisasi dari sistem yang dibangun.

HASIL DAN PEMBAHASAN

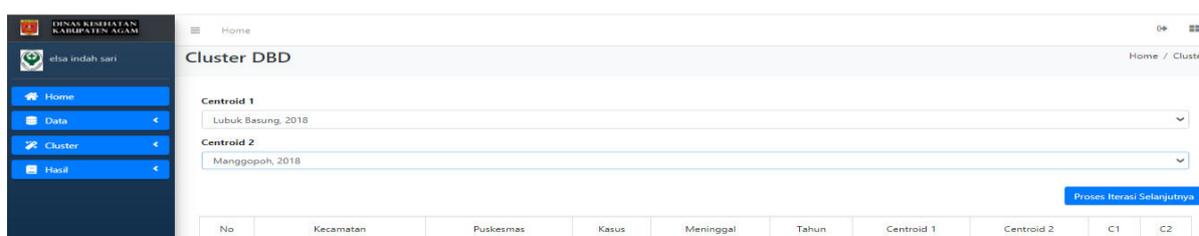
Klasterisasi penyakit dengan menerapkan Algoritma K-Medoids dibangun dengan bahasa pemrograman PHP dan basis data MYSQL. Data penyakit yang digunakan *tuberculosis*, *demam berdarah*, dan *pneumonia*. Parameter klasterisasi ini berupa nama puskesmas, kecamatan, jumlah kasus. Proses klasterisasi digunakan untuk mengetahui pola sebaran penderita penyakit pada suatu daerah. Tampilan halaman utama sistem dapat dilihat pada gambar 4.



Gambar 4. Halaman Utama Sistem Klasterisasi Penyakit

Halaman utama sistem terdiri dari 4 menu, yaitu menu *home* yang akan menampilkan informasi hasil *cluster* dan gambaran penyebarannya. Kemudian menu *data*, di mana menu ini memiliki sub menu, yaitu

user, lokasi, *tuberculosis*, *DBD*, dan *pneumonia*. Kemudian menu *cluster*, dimana pada menu ini proses klasterisasi menggunakan algoritma K-Medoids. Menu hasil terdapat sub menu laporan jenis penyakit berdasarkan tahun yang ada.



Proses klusterisasi dengan algoritma K-Medoids pada menu *cluster* ditampilkan pada gambar 5. Untuk memulai *user* harus memilih salah satu jenis penyakit yang akan dikelompokkan. Kemudian memilih 2 buah pusat klaster secara acak atau centroid 1 dan centroid 2.

Kemudian menekan tombol proses iterasi selanjutnya. Iterasi akan berhenti jika nilai simpangannya besar dari 0. Hasil klusterisasi ditunjukkan pada gambar 6. Di sini terlihat klaster rendah ditandai dengan kotak warna biru dan klaster tinggi dengan warna merah.

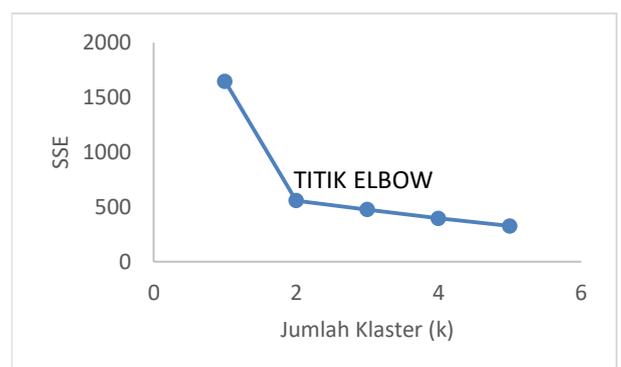
| No | Kecamatan | Puskesmas | Terduga | Kasus | Tahun | Centroid 1 | | Centroid 2 | | Hasil |
|----|-----------------|--------------|---------|-------|-------|------------|----|------------|----|--------|
| | | | | | | 2. | 0. | 35. | 0. | |
| 1 | Tanjung Mutiara | Muaro Putih | 2 | 0 | 2016 | 0 | | 33 | | Rendah |
| 2 | Tanjung Mutiara | Tiku | 0 | 0 | 2016 | 2 | | 35 | | Rendah |
| 3 | Ampek Nagari | Bawan | 35 | 0 | 2016 | 33 | | 0 | | Tinggi |
| 4 | Ampek Nagari | Batu Kambing | 0 | 0 | 2016 | 2 | | 35 | | Rendah |
| 5 | Lubuk Basung | Manggapoh | 2 | 0 | 2016 | 0 | | 33 | | Rendah |
| 6 | Lubuk Basung | Lubuk Basung | 152 | 0 | 2016 | 150 | | 117 | | Tinggi |
| 7 | Tanjung Raya | Pasar Ahad | 8 | 0 | 2016 | 6 | | 27 | | Rendah |
| 8 | Tanjung Raya | Maninjau | 11 | 0 | 2016 | 9 | | 24 | | Rendah |
| 9 | Pelembayan | Koto Alam | 29 | 0 | 2016 | 27 | | 6 | | Tinggi |
| 10 | Pelembayan | Pelembayan | 6 | 0 | 2016 | 4 | | 29 | | Rendah |

Gambar 6. Halaman Hasil Klusterisasi

Berikut proses klusterisasi untuk penyakit DBD menggunakan algoritma K-Medoids secara manual.

a. Menentukan Jumlah klaster.

Jumlah klaster ditentukan menggunakan Metode Elbow, dengan menghitung *Sum of Square Error* (SSE) menggunakan persamaan 1. Hasil perhitungan SSE diperlihatkan pada gambar 7. Pada saat jumlah klaster $k=1$ menunjukkan $SSE= 1645,52$. Saat jumlah klaster $k=2$ nilai $SSE=556,13$. Di sini tampak nilai SSE mengalami penurunan, sehingga garis pada grafik membentuk siku. Nilai SSE terlihat stabil pada saat $k=3$ sampai dengan $k=5$. Sehingga untuk penelitian ini menggunakan klaster $k=2$.



Gambar 7. Penentuan jumlah klaster terbaik dengan 368 data menggunakan Elbow

Di sini digunakan data kasus puskesmas Bawan dan Muaro Putuih tahun 2016. Data tersebut ditampilkan pada Tabel 2.

Tabel 2. Medoids Awal

| Klaster | Puskesmas | Tahun | Kasus |
|---------|--------------|-------|-------|
| C1 | Bawan | 2016 | 35 |
| C2 | Muaro Putuih | 2016 | 2 |

c. Menghitung jarak setiap data ke klaster terdekat menggunakan persamaan 2.

Berikut merupakan contoh perhitungan jarak untuk masing-masing klaster:
Jarak dari semua data ke titik pusat klaster 1 (C1)

$$C1(1) = \sqrt{(28 - 35)^2} = 7$$

$$C1(2) = \sqrt{(9 - 35)^2} = 26$$

$$C1(3) = \sqrt{(6 - 35)^2} = 29$$

.....

.....

$$C1(91) = \sqrt{(0 - 35)^2} = 35$$

$$C1(92) = \sqrt{(2 - 35)^2} = 33$$

Jarak dari semua data ke titik pusat klaster 2 (C2)

$$C2(1) = \sqrt{(28 - 2)^2} = 26$$

$$C2(2) = \sqrt{(9 - 2)^2} = 7$$

$$C2(3) = \sqrt{(6 - 2)^2} = 4$$

.....

.....

$$C2(91) = \sqrt{(0 - 2)^2} = 2$$

$$C2(92) = \sqrt{(2 - 2)^2} = 0$$

Setelah itu menentukan klaster untuk setiap data dengan melihat jarak terdekatnya, sehingga didapatkan hasil iterasi seperti pada Tabel 3.

Tabel 3. Hasil Iterasi Pertama

| Puskesmas | Tahun | Kasus | Jarak Centroid | | Jarak Terdekat | Klaster |
|--------------|-------|-------|----------------|-------|----------------|---------|
| | | | C1 | C2 | | |
| Baso | 2019 | 28 | 7 | 26 | 7 | 1 |
| Padang Tarok | 2019 | 9 | 26 | 7 | 7 | 2 |
| Biaro | 2019 | 6 | 29 | 4 | 4 | 2 |
| | | | | | | |
| | | | | | | |
| Tiku | 2016 | 0 | 35 | 2 | 2 | 2 |
| Muaro Putui | 2016 | 2 | 33 | 0 | 0 | 2 |
| Jumlah | | | 2851 | 757 | | |
| Total | | | 3608 | | | |

d. Menentukan pusat klaster baru secara acak (non medoid)

Selanjutnya melakukan proses perhitungan jarak setiap data ke pusat klaster seperti pada iterasi pertama. Hasil dapat dilihat pada Tabel 5.

Tabel 4. Medoid baru

| Klaster | Puskesmas | Tahun | Kasus |
|---------|--------------|-------|-------|
| C1 | Lubuk Basung | 2019 | 5 |
| C2 | Manggopoh | 2019 | 4 |

Tabel 5. Hasil Iterasi Kedua

| Puskesmas | Tahun | Kasus | Jarak Centroid | | Jarak Terdekat | Klaster |
|--------------|-------|-------|----------------|-------|----------------|---------|
| | | | C1 | C2 | | |
| Baso | 2019 | 28 | 23 | 24 | 23 | 1 |
| Padang Tarok | 2019 | 9 | 2 | 5 | 4 | 1 |
| Biaro | 2019 | 6 | 1 | 2 | 1 | 1 |
| | | | | | | |
| | | | | | | |
| Tiku | 2016 | 0 | 5 | 4 | 4 | 2 |

| | | | | | | |
|-------------|------|---|------|------|---|---|
| Muaro Putui | 2016 | 2 | 3 | 2 | 2 | 2 |
| Jumlah | | | 781 | 4117 | | |
| Total | | | 5552 | | | |

e. Menghitung Total Simpangan

Langkah selanjutnya menghitung Total simpangan (S) yaitu selisih total cost baru dengan total cost lama. Total simpangan (S) dihitung menggunakan persamaan

$$S = total\ cost\ baru - total\ cost\ lama$$

$$S = 5552 - 3608 = 1944$$

karena $S > 0$ maka iterasi berhenti. Sehingga didapatkan hasil klasterisasi penyakit DBD yaitu klaster tinggi terdiri dari 10 anggota dan klaster rendah terdiri dari 82 anggota. Anggota masing-masing klaster tinggi ditunjukkan pada tabel 6 dan klaster rendah pada tabel 7.

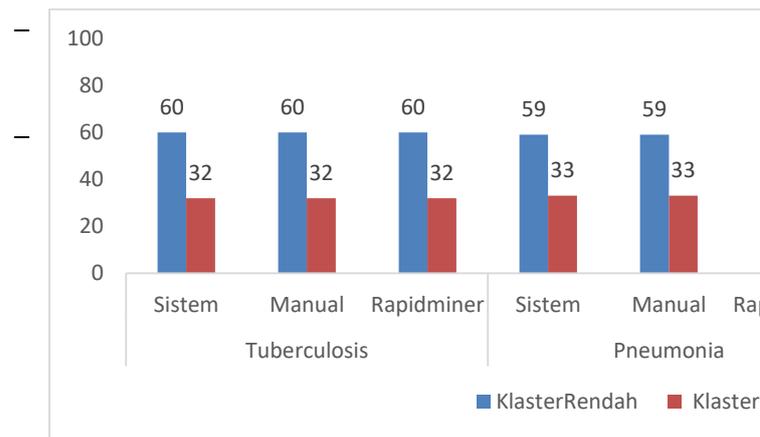
Tabel 6. Anggota Klaster Tinggi (C1)

| No | Puskesmas | Tahun | Kasus |
|----|--------------|-------|-------|
| 1 | Baso | 2019 | 28 |
| 2 | Maninjau | 2019 | 21 |
| 3 | Lubuk Basung | 2018 | 58 |
| 4 | Pakan Kamis | 2017 | 31 |
| 5 | Lubuk Basung | 2017 | 65 |
| 6 | Pakan Kamis | 2016 | 78 |
| 7 | Matur | 2016 | 41 |
| 8 | Koto Alam | 2016 | 29 |
| 9 | Lubuk Basung | 2016 | 152 |
| 10 | Bawan | 2016 | 35 |

Dari hasil klasterisasi data penyakit DBD di atas dapat dilihat yang termasuk dalam

kelompok rendah dengan rata-rata jumlah kasus di bawah 20, sedangkan kelompok tinggi di atas 20 kasus.

Selain penyakit DBD, langkah di atas dilakukan juga untuk klasterisasi penyakit *tuberculosis* dan *pneumonia*.



Gambar 7. Hasil Pengujian Sistem, Manual, dan Rapidminer

| | |
|-------|---------------------|
| | ... |
| | ... |
| 81 | Tiku 2016 0 |
| 82 | Muaro Putuih 2016 2 |

Pengujian

Untuk menguji kinerja sistem maka dilakukan dengan cara membandingkan hasil klasterisasi menggunakan perhitungan manual dan aplikasi *Rapidminer*. Dari gambar 7 menunjukkan hasil klaster yang sama. Ini menunjukkan sistem sudah bekerja dengan baik.

KESIMPULAN

Berdasarkan penelitian ini dapat disimpulkan bahwa klusterisasi penyakit menggunakan algoritma K-Medoids telah diterapkan pada sistem yang dirancang dengan 2 (dua) klaster yaitu rendah dan tinggi. Hasil pengujian diperoleh anggota masing-masing klaster untuk tiap-tiap penyakit yaitu *tuberculosis* klaster tinggi 32 dan rendah 60, penyakit *pneumonia* klaster tinggi 33 dan rendah 59, sedangkan untuk penyakit DBD klaster tinggi 10 anggota dan rendah 82 anggota. Hasil yang sama diperoleh dengan pengujian menggunakan perhitungan manual dan aplikasi Rapidminer. Ini menunjukkan bahwa sistem telah bekerja dengan baik.

Saran untuk penelitian selanjutnya dapat dikembangkan dengan menambahkan jenis penyakit, jumlah klaster, dan atribut lain padanya di Dinas Kesehatan Kabupaten Agam

REFERENSI

- [1] Dinas Kesehatan, *Profil Kesehatan Kabupaten Agam Tahun 2018*. Kab.Agam, 2018.
- [2] Suyanto, *Data Mining untuk Klasifikasi dan Klusterisasi Data*. Bandung: Informatika Bandung, 2017.
- [3] G. Gustientiedina, M. H. Adiya, and Y. Desnelita, "Penerapan Algoritma K-Means Untuk Clustering Data Obat-Obatan," *J. Nas. Teknol. dan Sist. Inf.*, vol. 5, no. 1, pp. 17–24, 2019, doi: 10.25077/teknosi.v5i1.2019.17-24.
- [4] J. Han, M. Kamber, and J. Pei, "Data mining concepts and techniques third edition," *Morgan Kaufmann Ser. Data Manag. Syst.*, vol. 5, no. 4, pp. 83–124, 2011.
- [5] B. Riyanto, "PENERAPAN ALGORITMA K-MEDOIDS CLUSTERING UNTUK PENGELOMPOKKAN PENYEBARAN DIARE DI KOTA MEDAN (STUDI KASUS: KANTOR DINAS KESEHATAN KOTA MEDAN)," vol. 3, pp. 562–568, 2019, doi: 10.30865/komik.v3i1.1659.
- [6] K. Khomsatun, D. Ikhsan, M. Ali, and K. Kursini, "SISTEM PENGAMBILAN KEPUTUSAN PEMILIHAN LAHAN TANAM DI KABUPATEN WONOSOBO DENGAN K-MEANS CLUSTERING DAN TOPSIS," *J. Nas. Pendidik. Tek. Inform. JANAPATI*, vol. 9, no. 1, pp. 55–62, 2020.
- [7] L. P. Rizby, Marji, and L. Muflikhah, "Clustering Pasien Kanker Berdasarkan Struktur Protein Dalam Tubuh," *J. Pengemb. Teknol. Inf. dan Ilmu Komput. Univ. Brawijaya*, vol. 2, no. 10, pp. 3810–3816, 2018.
- [8] I. Kamila, U. Khairunnisa, and M. Mustakim, "Perbandingan Algoritma K-Means dan K-Medoids untuk Pengelompokan Data Transaksi Bongkar Muat di Provinsi Riau," *J. Ilm. Rekayasa dan Manaj. Sist. Inf.*, vol. 5, no. 1, p. 119, 2019, doi: 10.24014/rmsi.v5i1.7381.
- [9] A. D. Andini, T. Arifin, A. R. Sanjaya, A. R. Sanjaya, S. Coefficient, and P. Pasien, "IMPLEMENTASI ALGORITMA K-MEDOIDS UNTUK KLASTERISASI DATA PENYAKIT PASIEN," vol. 2, no. 2, pp. 128–138, 2020.
- [10] L. Purba, S. Saifullah, and R. Dewi, "Pengelompokan Kasus Penyakit Aids Berdasarkan Provinsi Dengan Data Mining K-Medoids Clustering," *KOMIK (Konferensi Nas. Teknol. Inf. dan Komputer)*, vol. 3, no. 1, pp. 687–694, 2019, doi: 10.30865/komik.v3i1.1679.
- [11] T. Juninda and E. Andri, "Penerapan Algoritma K-Medoids untuk Pengelompokan Penyakit di Pekanbaru Riau," no. November, pp. 42–49, 2019.
- [12] S. Sundari, I. S. Damanik, A. P. Windarto, H. S. Tambunan, J. Jalaluddin, and A. Wanto, "Analisis K-Medoids Clustering Dalam Pengelompokan Data Imunisasi Campak Balita di Indonesia," *Pros. Semin. Nas. Ris. Inf. Sci.*, vol. 1, no. September, p. 687, 2019, doi: 10.30645/senaris.v1i0.75.

- [13] H. Ningrum, E. Irawan, and M. R. Lubis, "Implementasi Metode K-Medoids Clustering Dalam Pengelompokan Data Penyakit Alergi Pada Anak," *Jurasik (Jurnal Ris. Sist. Inf. dan Tek. Inform.*, vol. 6, no. 1, pp. 130–139, 2021.
- [14] S. Sindi, W. R. O. Ningse, I. A. Sihombing, F. I. R. H. Zer, and D. Hartama, "Analisis algoritma k-medoids clustering dalam pengelompokan penyebaran covid-19 di indonesia," *JurTI (Jurnal Teknol. Informasi)*, vol. 4, no. 1, pp. 166–173, 2020.
- [15] E. Irwansyah, E. S. Pratama, and M. Ohyver, "Clustering of Cardiovascular Disease Patients Using Data Mining Techniques with Principal Component Analysis and K-Medoids," 2020.
- [16] D. Marlina, A. Fernando, and A. Ramadhan, "Implementasi Algoritma K-Medoids dan K-Means untuk Pengelompokan Wilayah Sebaran Cacat pada Anak," *J. CoreIT J. Has. Penelit. Ilmu Komput. dan Teknol. Inf.*, vol. 4, no. 2, pp. 64–71, 2018.
- [17] H. Kusumah, M. R. Lubis, and H. S. Tambunan, "Penerapan Algoritma K-Medoids Dalam Pengelompokan Imunisasi Lanjutan Pada Anak Usia 2 Tahun," *Resolusi Rekayasa Tek. Inform. dan Inf.*, vol. 1, no. 4, pp. 265–273, 2021.
- [18] D. A. I. C. Dewi and D. A. K. Pramita, "Analisis Perbandingan Metode Elbow dan Silhouette pada Algoritma Clustering K-Medoids dalam Pengelompokan Produksi Kerajinan Bali," *Matrix J. Manaj. Teknol. dan Inform.*, vol. 9, no. 3, pp. 102–109, 2019.