

Perbandingan Metode K-Means dan K-Medoids Untuk Pengelompokan Sebaran Covid-19 di Sulawesi Selatan

Ismail^{*1}, M. Afdal Tahir²

^{1,2}Program Studi Sistem Informasi, Universitas Lamappapoleonro
E-mail: ¹ismai@unipol.ac.id, ²afdal.tahir@unipol.ac.id

ABSTRAK

Pandemi COVID-19 merupakan pandemi berskala global yang memiliki dampak hampir keseluruhan negara termasuk Indonesia. Indonesia pertama kali mengonfirmasi kasus COVID-19 pada tanggal 2 maret 2020. Sebaran Pandemi COVID-19 yang menyeluruh di wilayah Indonesia, merupakan penularan yang sangat cepat dan memiliki dampak negatif pada berbagai bidang. Salah satu wilayah yang tak luput dari serangan Covid-19 adalah Provinsi Sulawesi Selatan, seluruh kabupaten dan kota yang ada terpapar covid-19 dengan jumlah kasus yang berbeda-beda di setiap wilayah, permasalahan yang dihadapi petugas masih sulit untuk mengelompokkan wilayah sebaran covid-19 di Sulawesi Selatan. Penelitian ini dilakukan dengan tujuan membuat model perbandingan metode K-Means dan K-Medoids dalam pengelompokan sebaran covid 19 di Sulawesi Selatan. Maka diperlukan pengolahan data untuk pengelompokan wilayah sebaran COVID-19. Pengembangan sistem penelitian ini menggunakan metode SDLC, dimulai dari tahap perencanaan, analisis, desain, implementasi dan uji coba. Hasil perbandingan metode K-Means dan K-Medoids dievaluasi menggunakan metode Davies Bouldin Index. Hasil Evaluasi metode K-Means dan metode K-Medoids dapat disimpulkan bahwa metode K-Means lebih optimal dibandingkan metode K-Medoids. Adapun hasil pengelompokan dari metode K-means diperoleh 4 cluster. Pada cluster 1 2 daerah, pada cluster 2, 3 daerah, pada cluster 3 tiga daerah, pada cluster 4 19 daerah.

Kata Kunci: Covid-19, Cluster, K-Means, K-Medoids.

ABSTRACT

The COVID-19 pandemic is a global-scale pandemic that has an impact on almost all countries, including Indonesia. Indonesia first confirmed a case of COVID-19 on March 2, 2020. The spread of the COVID-19 pandemic, which is comprehensive in Indonesia, is a very fast transmission and has a negative impact on various fields. One of the areas that has not escaped the Covid-19 attack is South Sulawesi Province, all regencies and cities are exposed to Covid-19 with different numbers of cases in each region, the problems faced by officers are still difficult to group the Covid-19 distribution area in South Sulawesi. This study was conducted with the aim of making a comparative model of the K-Means and K-Medoids methods in grouping the spread of COVID-19 in South Sulawesi. Therefore, data processing is needed for grouping COVID-19 distribution areas. The development of this research system uses the SDLC method, starting from the planning, analysis, design, implementation and trial stages. The results of the comparison of the K-Means and K-Medoids methods were evaluated using the Davies Bouldin Index method. The results of the evaluation of the K-Means method and the K-Medoids method can be concluded that the K-Means method is more optimal than the K-Medoids method. The grouping results of the K-means method obtained 4 clusters. In cluster 1 2 regions, in cluster 2, 3 regions, in cluster 3 three regions, in cluster 4 19 regions.

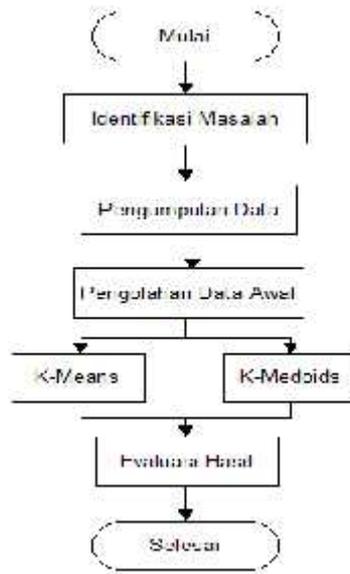
Keywords: Covid-19, Cluster, K-Means, K-Medoids

1. PENDAHULUAN

COVID-19 merupakan penyakit yang disebabkan oleh novel coronavirus. Penyebab COVID-19, yang kemudian dinamai SARS-CoV-2, pertama kali dilaporkan oleh pejabat di Kota Wuhan, Tiongkok, pada Desember 2019. Investigasi retrospektif oleh otoritas Tiongkok telah mengidentifikasi kasus manusia dengan timbulnya gejala pada awal Desember 2019[1]. Sementara beberapa kasus paling awal yang diketahui memiliki kaitan dengan pasar grosir makanan di Wuhan. Virus ini sangat berbahaya karena selain menginfeksi saluran pernapasan juga dapat menyebabkan kematian, orang dengan sistem imun yang lemah dan lanjut usia sangat rentan terhadap infeksi virus ini. Pandemi COVID-19 merupakan pandemi berskala global yang memiliki dampak hampir keseluruhan negara termasuk Indonesia. Indonesia pertama kali mengumumkan kejadian pandemi COVID-19 pada tanggal 2 maret 2020. Penyebaran kasus COVID-19 yang menyebar luas di wilayah Indonesia, merupakan penyebaran yang cukup cepat dan memiliki dampak negatif pada berbagai bidang. Salah satu wilayah yang tak luput dari serangan Covid-19 adalah Provinsi Sulawesi Selatan, Seluruh kabupaten dan kota yang ada terpapar Covid-19 dengan jumlah kasus yang berbeda-beda di setiap wilayah, berdasarkan data tanggal 5 Agustus 2021 laju penyebaran yang dinyatakan sembuh dari Covid ada 73.135 orang, dirawat 11.645 orang, dan meninggal ada 1400 orang. Sebelumnya telah dilakukan pengelompokan jumlah kematian penderita Covid-19 berdasarkan negara di benua Asia dengan metode K-Means. Penelitian ini menghasilkan pengelompokan c1 sebanyak 4 negara, c2 sebanyak 4 negara, dan 41 negara [2]. Selain itu, ada juga penelitian clustering persebaran covid-19 di DKI Jakarta menggunakan k-means yang mendapat hasil cluster 1 dengan jumlah 19 kecamatan, Cluster 2 dengan jumlah 23 kecamatan, dan Cluster 3 sebanyak 2 kecamatan. Di penelitian ini memberikan saran untuk menerapkan metode klusterisasi lainnya seperti Fuzzy k-means, Fuzzy C-Means, K-Medoids, dan sebagainya[3]. Adapun penelitian analisis algoritma K-Medoids Clustering dalam pengelompokan penyebaran COVID-19 di Indonesia yang menggunakan data dengan jumlah 34 didapatkan 1 baris pada cluster pertama, 2 record pada cluster kedua, 31 record pada cluster ketiga[4]. Adapun berbagai penerapan algoritma K-Medoids seperti implementasi metode K-Medoids clustering dalam pengelompokan data penyakit alergi pada anak[5], dan implementasi algoritma K-Medoids untuk klusterisasi data penyakit pasien di RSUD Kota Bandung[6]. Berdasarkan penelitian sebelumnya, maka pada penelitian ini akan mencoba membandingkan hasil clustering dari metode K-Means dan K-Medoids. Adapun perbandingan algoritma K-Means dan K-Medoids untuk Pengelompokan Daerah Rawan Tanah Longsor di Provinsi Jawa Barat didapatkan hasil pengelompokan dengan metode K-Means lebih optimal dengan jumlah $k=6$ [7].

2. METODE PENELITIAN

Metode penelitian menjelaskan tentang tahapan penelitian dalam melakukan perbandingan metode *K-Means* dan *K-Medoids* untuk pengelompokan wilayah sebaran COVID-19 di Sulawesi Selatan. Berikut adalah metodologi pada penelitian ini.



Gambar 1. Flowchart Metode Penelitian

Adapun penjelasan flowchart metode penelitian pada gambar 1 sebagai berikut :

Identifikasi Masalah

Identifikasi masalah pada penelitian untuk membandingkan metode *K-Means* dan *K-Medoids* yang bertujuan untuk mengetahui metode yang lebih efektif dalam melakukan *clustering* COVID-19 di Sulawesi Selatan.

Pengumpulan Data

Penelitian ini menggunakan data yang diambil website dengan Url <https://covid19.sulselprov.go.id/data> diakses pada tanggal 5 Agustus 2021.

Pengolahan Data Awal

Data yang telah diperoleh akan diproses untuk memberikan hasil yang kemudian akan digunakan pada fase selanjutnya sehingga menghasilkan informasi yang tepat. Berikut merupakan data sebaran COVID-19 di Sulawesi Selatan.

Tabel 1 Data Covid-19

No	Kab / Kota	Dirawat	Sembuh	Meninggal	Total
1	Kota Makassar	4453	36728	769	41950
2	Kabupaten Gowa	1054	6027	81	7162
3	Kabupaten Luwu Timur	362	3232	24	3618
4	Kabupaten Maros	552	2650	32	3234
5	Kabupaten Sinjai	173	2915	28	3116

6	Kabupaten Jenepono	165	1730	27	1922
7	Kabupaten Luwu Utara	539	1545	28	2112
8	Kota Pare-Pare	494	1458	40	1992
9	Kabupaten Pangkajene dan Kepulauan	464	1475	46	1985
10	Kota Palopo	276	1717	38	2031
11	Kabupaten Bone	308	1492	12	1812
12	Kabupaten Takalar	343	1463	35	1841
13	Kabupaten Bulukumba	165	1275	23	1463
14	Kabupaten Tana Toraja	543	1008	10	1561
15	Kabupaten Pinrang	89	1052	56	1197
16	Kabupaten Soppeng	279	1262	24	1565
17	Kabupaten Sidenreng Rappang	174	1046	15	1235
18	Kabupaten Barru	246	1025	7	1278
19	Kabupaten Bantaeng	66	943	11	1020
20	Kabupaten Wajo	128	909	22	1059
21	Kabupaten Luwu	126	709	9	844
22	Kabupaten Kepulauan	235	715	34	984

	Selayar				
23	Kabupaten Toraja Utara	336	364	6	706
24	Kabupaten Enrekang	75	395	23	493

K-Means

K-means merupakan algoritma data mining untuk pengelompokkan, pertama kali dipublikasikan oleh Stuart Loyd pada tahun 1984 dan merupakan algoritma clustering yang banyak digunakan. Algoritmanya cukup sederhana untuk diterapkan dan dioperasikan, relatif singkat serta mudah disesuaikan dan banyak digunakan. hal pertama dari cara ini yaitu membentuk sebuah partisi/pusat (*centroid*)/ rata-rata (*mean*) dari kelompok data. Algoritma *K-means* diawali dengan membentuk partisi klaster pertama kemudian secara iteratif partisi klaster ini diperbaiki hingga tidak terjadi perubahan yang banyak pada bagian kelompok data[8]. *Clustering* data dengan algoritma *K-Means* pada dasarnya dilakukan dengan algoritma dasar sebagai berikut[9]:

- 1) Tentukan nilai k sebagai jumlah cluster yang diinginkan.
- 2) Tentukan nilai secara random untuk pusat cluster pertama *centroid* sebanyak k, dengan menggunakan rumus jarak untuk menghitung jarak *EuclideanDistance* yaitu :

$$d(x_i, \mu_j) = \sqrt{\sum_{l=1}^n (x_{il} - \mu_{jl})^2} \tag{1}$$

Dimana: x_i = data kriteria

μ_j = centroid pada cluster ke-j

- 3) Kelompokkan data berdasarkan nilai jarak terkecil setiap data
- 4) Perbaharui *centroid* baru dari rata-rata cluster dengan persamaan:

$$\mu_j(t+1) = \frac{1}{N_{Sj}} \sum_{j \in S_j} x_j \tag{2}$$

Keterangan : $\mu_j(t+1)$ = centroid baru pada iterasi (t+1)

N_{Sj} = Data pada cluster S_j

- 5) Lakukan perulangan dari langkah 2 hingga 5, sampai anggota tiap *cluster* tidak ada yang berubah.

K-Medoids

K-Medoids merupakan suatu algoritma yang digunakan untuk menemukan medoid pada suatu kelompok (*cluster*) yang merupakan titik pusat dari suatu kelompok. Algoritma *K-Medoids* lebih baik dibanding *K-Means* karena pada *K-Medoids* dapat menemukan k sebagai objek yang untuk membatasi jumlah ketidakseimbangan objek data, sementara di *K-Means* memanfaatkan jumlah jarak euclidean untuk objek data. Algoritma *K-Medoids* merupakan suatu algoritma yang mengatasi kelemahan Algoritma *K-Means* yang sensitif terhadap outlier karena objek dengan suatu nilai yang besar mungkin menyimpang dari distribusi data[8]. Adapun langkah-langkah algoritma *K-Medoids* sebagai berikut[10]:

- 1) Inisialisasi pusat cluster sebanyak k (jumlah cluster)

- 2) Alokasikan setiap data (objek) ke cluster terdekat menggunakan persamaan *Euclidian Distance* dengan persamaan dibawah ini:

$$d(x, y) = \sqrt{\sum_{i=1}^n (X_i - Y_i)^2} \quad (3)$$

Keterangan:

d = jarak antara x dan y

x = pusat data *cluster*

y = data pada atribut

i = indeks data

n = jumlah data

X_i = *Medoid Cluster i*

Y_i = nilai atribut data ke i

- 3) Pilih secara acak objek pada masing-masing cluster sebagai kandidat medoid baru.
 4) Hitung jarak setiap objek yang berada pada masing-masing cluster dengan kandidat medoids baru.
 5) Hitung total simpangan (S) dengan menghitung nilai total distance baru – total distance lama. Jika $S < 0$, maka tukar objek dengan data cluster untuk membentuk sekumpulan k objek baru sebagai medoid.
 6) Ulangi langkah 3 sampai 5 hingga tidak terjadi perubahan medoids, sehingga didapatkan cluster beserta anggota clustermasing-masing

Davies Bouldin Index

Metode yang diperkenalkan oleh David L.Davies dan Donald W.Bouldin dan penamaan metode ini menggunakan nama dengan keduanya yakni *Davies-Bouldin* (DBI) diterapkan untuk memberikan evaluasi *cluster*. Evaluasi dengan *Davies-Bouldin Index* mempunyai gambaran penilaian kelompok internal, dimana hasil cluster baik atau tidak dilihat dari kuantitas dan kedekatan antar hasil cluster. *Davies-Bouldin Index* merupakan salah satu metode yang digunakan untuk menilai keakuratan antar hasil cluster dari suatu metode, *cohesion* diartikan menjadi penjumlahan hubungan data dengan titik pusat cluster yang diikuti cluster. Sedangkan pemisahannya berdasarkan jarak antar titik pusat cluster dengan cluster. Pengukuran menggunakan *Davies-Bouldin Indeks* ini memaksimalkan jarak antar *cluster* antara cluster C_i dan C_j dan pada saat yang sama mencoba meminimalkan jarak antar titik dalam suatu cluster. Jika jarak antar cluster maksimal, artinya kesamaan karakteristik antar cluster kecil sehingga perbedaan antar cluster tampak lebih jelas. Jika jarak intra cluster minimum berarti setiap objek dalam cluster memiliki tingkat kesamaan karakteristik yang tinggi[7]. Dengan menggunakan DBI suatu cluster akan dianggap memiliki skema clustering yang optimal adalah yang memiliki DBI minimal[11]. Langkah-langkah perhitungan *Davies Bouldin Index* adalah Sebagai berikut[9]:

- 1) *Sum Of Square Within- Cluster(SSW)*

Rumus yang digunakan untuk mencari matriks kohesi dalam sebuah cluster ke-i dengan rumus sebagai berikut [11][9]:

$$SSW_i = \frac{1}{m_i} \sum_{j=1}^{m_i} d(x_j, c_i) \quad (4)$$

$d(X_j, C_j)$ = jarak setiap data ke centroid i yang dihitung menggunakan jarak euclidiance.

2) *Sum Of Square Between-Cluster (SSB)*

Perhitungan *Sum Of Square Between-Cluster (SSB)* bertujuan untuk mengetahui separasi atau jarak antar cluster dengan rumus perhitungan sebagai berikut [11][9]:

$$SSB_{ij} = d(C_i, C_j) \quad (5)$$

$d(C_i, C_j)$ = jarak antara data ke i dengan data ke j di cluster lain

3) *Ratio (Rasio)*

Perhitungan rasio ($R_{i,j}$) ini bertujuan untuk mengetahui nilai perbandingan antara cluster ke-i dan cluster ke-j untuk menghitung nilai rasio yang dimiliki oleh masing-masing cluster. indeks I dan j

merupakan merepresentasikan jumlah cluster, dimana jika terdapat 4 cluster maka terdapat indeks sebanyak 4 yaitu i,j,k dan l. untuk menentukan nilai rasio dengan rumus sebagai berikut :

$$R_{i,j,\dots,n} = \frac{SSW_i + SSW_j + \dots + SSW_n}{SSB_{i,j} + \dots + SSB_{n,j}} \quad (6)$$

Dimana :

SSW_i = *Sum Of Square Within- Cluster* pada centroid i

$SSB_{i,j}$ = *Sum of Square Between Cluster* data ke i dengan j pada cluster yang berbeda

4) *Davies Bouldin Index (DBI)*

Setelah didapatkan nilai rasio, akan digunakan untuk mencari nilai *Davies Bouldin Index (DBI)* dengan rumus berikut:

$$DBI = \frac{1}{k} \sum_{i=1}^k \max_{t \neq j} (R_{i,j}) \quad (7)$$

Evaluasi Hasil

Dimana, $R_{i,j}$ merupakan ratio dari nilai SSW dan SSB dan k adalah jumlah cluster. Dari perhitungan *Davies Bouldin Index (DBI)* dapat disimpulkan bahwa jika semakin kecil nilai *Davies Bouldin Index (DBI)* yang diperoleh (non negatif ≥ 0) maka cluster tersebut semakin baik.

3. HASIL DAN PEMBAHASAN***1.1 Hasil Penelitian***

Perhitungan Menggunakan Algoritma K-Means.

Berikut merupakan tahapan perhitungan algoritma *K-Means*:

- 1) Tentukan nilai k sebagai jumlah cluster yang diinginkan
- 2) Tentukan nilai acak atau random untuk pusat cluster awal centroid sebanyak k

Tabel 2. Centroid Awal

	Dirawat	Sembuh	Meninggal
C1	4453	36728	769
C2	1054	6027	81
C3	362	3232	24

Setelah menentukan centroid awal maka tahap selanjutnya menghitung jarak *EuclideanDistance* dengan menggunakan persamaan (2)

a) C1 (data 1, centroid 1)

$$= \sqrt{(4453 - 4453)^2 + (36728 - 36728)^2 + (769 - 769)^2}$$

$$= 0$$

b) C2 (data 1, centroid 2)

$$= \sqrt{(4453 - 1054)^2 + (36728 - 6027)^2 + (769 - 81)^2}$$

$$= 30896.24$$

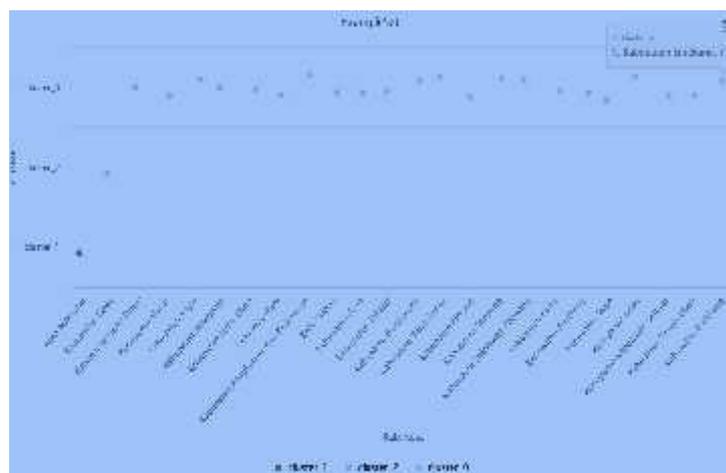
c) C3 (data 1, centroid 3)

$$= \sqrt{(4453 - 362)^2 + (36728 - 3232)^2 + (769 - 24)^2}$$

$$= 33753.12$$

Pengujian Data dengan Rapid Miner

Untuk menghasilkan pembentukan kelompok data, maka pada proses berikutnya dilakukan pembagian kelompok data dengan *Tools Rapidminer*. Hasil akhir yang akan ditampilkan adalah seperti *cluster* kemudian hasil dari uji data kelihatan cluster dengan masing-masing anggota. Berikut merupakan hasil pengujian data dengan *rapid miner*.

Gambar 1 Hasil Pengelompokkan *K-Means*

Pada gambar 1 dapat dilihat bahwa terdapat 3 warna yang berbeda dimana cluster 1 berwarna merah, cluster 2 berwarna kuning, dan cluster 0 berwarna hijau. Dari gambar diatas dapat dilihat bahwa diperoleh 1 kota pada cluster1 yaitu kota Makassar, 1 kabupaten (Gowa) pada cluster2, dan 22 Kabupaten/Kota (Luwu Timur, Sinjai, Maros, Jeneponto, Palopo, Luwu Utara, Bone, Pinrang, Bulukumba, Pangkep, Pare-Pare, Takalar, Soppeng, Sidrap, Bantaeng, Tana Toraja, Barru, Wajo, Luwu, Enrekang, Selayar dan Toraja Utara) pada cluster0. Adapun hasil pengelompokan dengan metode *K-Medoids* sebagai berikut:



Gambar 2 Hasil Pengelompokan *K-Medoids*

Dari gambar 2 dapat dilihat bahwa diperoleh 1 kota pada cluster0 yaitu kota Makassar, 21 kabupaten / kota (Gowa , Luwu Timur, Sinjai, Maros, Jeneponto, Palopo, Luwu Utara, Bone, Pinrang, Bulukumba, Pangkep, Pare-Pare, Takalar, Soppeng, Sidrap, Bantaeng, Tana Toraja, Barru, Wajo, Luwu, dan Selayar) pada cluster1, dan 2 Kabupaten (Enrekang, dan Toraja Utara) pada cluster2.

Evaluasi Hasil

Setelah melakukan proses cluster data maka tahap selanjutnya menentukan cluster yang paling optimal dengan menggunakan Davies Bouldin Index (DBI). Untuk mengetahui cluster terbaik pada DBI adalah dengan mengetahui nilai DBI yang terkecil. Adapun proses perhitungan DBI sebagai berikut:

1) Nilai DBI K-Means

Perhitungan manual davies bouldin index diperlukan data yang sudah terbentuk pada clustering akhir (Tabel 3.11) dan titik pusat/ centroid yang terakhir seperti pada Tabel 3.10.

Berikut ini merupakan proses perhitungan nilai DBI dari algoritma K-Means

a. Sum Of Square Within

$$\begin{aligned}
 SSW_1 &= \frac{0}{1} \\
 &= 0 \\
 SSW_2 &= \frac{0}{1} \\
 &= 0 \\
 SSW_3 &= \frac{1852,951+1298,141+\dots+1006,795}{22}
 \end{aligned}$$

$$= 570,190$$

b. Sum Of Square Between

$$\begin{aligned} \text{SSB}_{1,2} &= \sqrt{(4453 - 1054)^2 + (36728 - 6027)^2 + (769 - 81)^2} \\ &= 30896,245 \end{aligned}$$

$$\begin{aligned} \text{SSB}_{1,3} &= \sqrt{(4453 - 279)^2 + (36728 - 1380,909)^2 + (769 - 25)^2} \\ &= 35600,459 \end{aligned}$$

$$\begin{aligned} \text{SSB}_{2,3} &= \sqrt{(1054 - 279)^2 + (6027 - 1380,909)^2 + (81 - 25)^2} \\ &= 4710,618 \end{aligned}$$

c. Ratio

$$\begin{aligned} R &= \frac{\text{SSW}_1 + \text{SSW}_2 + \text{SSW}_3}{\text{SSB}_{1,2} + \text{SSB}_{1,3} + \text{SSB}_{2,3}} \\ &= \frac{0 + 0 + 570,190}{30896,245 + 35600,459 + 4710,618} \\ &= 0,008 \end{aligned}$$

d. DBI

$$K = 3$$

$$R = 0,008$$

$$\begin{aligned} \text{DBI} &= \frac{R}{K} \\ &= \frac{0,008}{3} \\ &= 0,0027 \end{aligned}$$

2) Nilai DBI K-Medoids

Hasil clustering algoritma K-Medoids dapat dilihat pada tabel 3.18 dan titik pusat/centroid yang terakhir seperti pada Tabel 3.16. Berikut ini merupakan proses perhitungan nilai DBI dari algoritma K-Medoids

a. Sum Of Square Within

$$\begin{aligned} \text{SSW}_1 &= \frac{0}{1} \\ &= 0 \end{aligned}$$

$$\begin{aligned} \text{SSW}_2 &= \frac{5374,971 + 2520,222 + \dots + 0}{21} \\ &= 1012,019 \end{aligned}$$

$$\begin{aligned} \text{SSW}_3 &= \frac{263,384 + 0}{2} \\ &= 131,692 \end{aligned}$$

b. Sum Of Square Between

$$\begin{aligned} \text{SSB}_{1,2} &= \sqrt{(4453 - 235)^2 + (36728 - 715)^2 + (769 - 34)^2} \\ &= 36266,623 \end{aligned}$$

$$\begin{aligned} \text{SSB}_{1,3} &= \sqrt{(4453 - 75)^2 + (36728 - 395)^2 + (769 - 23)^2} \\ &= 36603,419 \end{aligned}$$

$$\begin{aligned}SSB_{2,3} &= \sqrt{(235 - 75)^2 + (715 - 395)^2 + (34 - 23)^2} \\ &= 357,940\end{aligned}$$

c. Ratio

$$\begin{aligned}R &= \frac{SSW1+SSW2+SSW3}{SSB1,2+SSB1,3+SSB2,3} \\ &= \frac{0+1012,019+131,692}{36266,623+36603,419+357,940} \\ &= 0,016\end{aligned}$$

d. DBI

$$K = 3$$

$$R = 0,016$$

$$DBI = \frac{R}{K}$$

$$= \frac{0,017}{3}$$

$$= 0,0052$$

Pada penelitian ini telah dilakukan pengujian dataset dengan jumlah cluster k=3 sampai k=5 dengan menggunakan microsoft excel. Berikut ini merupakan hasil perhitungan nilai *Davies Bouldin Index*

Tabel 6. Hasil Perhitungan DBI

NO	Cluster	K-Means	K-Medoids
1	3	0,0027	0,0052
2	4	0,0015	0,0038
3	5	0,0018	0,0024

4. KESIMPULAN

Adapun kesimpulan yang dapat diambil berdasarkan pembahasan pada penelitian ini adalah sebagai berikut :

- Hasil evaluasi clustering menggunakan Davies Bouldin Index menunjukkan bahwa hasil clustering metode K-Means lebih optimal dibandingkan dengan menggunakan metode K-Medoids.
- Pengelompokan data dari 24 record menggunakan metode K-Means menghasilkan 4cluster yaitu cluster1 sebanyak 1 record, cluster2 sebanyak 1 record , cluster3 sebanyak 3 record dancluster4 sebanyak19 record.

5. SARAN

Untuk melakukan penelitian lebih lanjut dengan tema atau kasus pengelompokan sebaran covid-19 disarankan menggunakan metode K-Means atau dapat melakukan penelitian dengan perbandingan kinerja algoritma K-Means dengan metode yang lain.

DAFTAR PUSTAKA

- [1] World Health Organization, “Considerations in adjusting public health and social measures in the context of COVID-19,” *World Heal. Organ. Interim Guid.*, no. November, pp. 1–13, 2020, [Online]. Available: <https://www.who.int/publications-detail/risk->.
- [2] Noviyanto, “Penerapan Data Mining Dalam Mengelompokkan Jumlah Kematian,” *Paradig. Inform. dan Komput.*, vol. 22, no. 2, 2020.
- [3] A. Solichin and K. Khairunnisa, “Klasterisasi Persebaran Virus Corona (Covid-19) Di DKI Jakarta Menggunakan Metode K-Means,” *Fountain Informatics J.*, vol. 5, no. 2, p. 52, 2020, doi: 10.21111/fij.v5i2.4905.
- [4] S. Sindi, W. Ratnasari, 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,” vol. 4, no. 1, pp. 166–173, 2020.
- [5] 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, p. 130, 2021, doi: 10.30645/jurasik.v6i1.277.
- [6] A. D. Andini and T. Arifin, “Implementasi Algoritma K-Medoids Untuk Klasterisasi Data Penyakit Pasien Di Rsud Kota Bandung,” *J. RESPONSIF Ris. Sains ...*, vol. 2, no. 2, pp. 128–138, 2020, [Online]. Available: <http://ejurnal.ars.ac.id/index.php/jti/article/view/247>.
- [7] M. Herviany, S. P. Delima, and T. Nurhidayah, “Comparison of K-Means and K-Medoids Algorithms for Grouping Landslide Prone Areas in West Java Province Perbandingan Algoritma K-Means dan K-Medoids untuk Pengelompokan Daerah Rawan Tanah Longsor di Provinsi Jawa Barat,” vol. 1, no. April, pp. 34–40, 2021.
- [8] D. Marlina, N. Lina, 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, p. 64, 2018, doi: 10.24014/coreit.v4i2.4498.
- [9] R. K. Dinata, H. Novriando, N. Hasdyna, and S. Retno, “Reduksi Atribut Menggunakan Information Gain untuk Optimasi Cluster Algoritma K-Means,” *J. Edukasi dan Penelit. Inform.*, vol. 6, no. 1, p. 48, 2020, doi: 10.26418/jp.v6i1.37606.
- [10] B. Surarso and R. Gernowo, “IMPLEMENTATION OF K-MEDOIDS CLUSTERING FOR HIGH EDUCATION ACCREDITATION DATA,” vol. 10, no. 3, pp. 119–128, 2020.
- [11] N. Butsianto, Sufajar;Saepudin, “PENERAPAN DATA MINING TERHADAP MINAT SISWA DALAM MATA PELAJARAN MATEMATIKA DENGAN METODE K-MEANS,” *SIGMA – J. Teknol. Pelita Bangsa*, vol. 10, pp. 113–121, 2019.