

4409-9985-1-PB.pdf

by

Submission date: 01-Feb-2023 06:30AM (UTC+0700)

Submission ID: 2003662121

File name: 4409-9985-1-PB.pdf (489.64K)

Word count: 3372

Character count: 19068

Klastering Penyebaran Covid-19 Di Indonesia Tahun 2020 Menggunakan *K-Medoids Clustering*

Ayu Khairunnisa^{1*}, Alwis Nazir², Novriyanto³, Yusra⁴, Lola Oktavia⁵, Amany Akhyar⁶

¹ Sains dan Teknologi, Teknik Informatika, UIN SUSKA Riau, Pekanbaru, Indonesia

Jl. H.R. Soebrantas No. 155 Km 15, Simpang Baru, Tampan, Pekanbaru

Corresponding author's e-mail: 11651203467@students.uin-suska.ac.id

Abstrak - Badan Nasional Penanggulangan Bencana (BNPB) telah menetapkan status darurat bencana terkait virus Covid-19. Penyebaran pada kasus Covid-19 yang sudah merata di seluruh provinsi di Indonesia sudah merupakan penyebaran yang cukup cepat dan berdampak negative pada seluruh bidang. Dari banyaknya data Covid-19 di Indonesia yang ada, diperlukan pengelompokan penyebaran kasus Covid-19 sehingga dapat dimanfaatkan oleh berbagai pihak, salah satunya yaitu Dinas Kesehatan mengenai pengelompokan penyebaran Covid-19 pada 34 provinsi di Indonesia agar dapat diketahui pola penyebarannya menggunakan data mining metode clustering dan algoritma yang digunakan yaitu K-Medoids. Berdasarkan penelitian yang telah dilakukan dapat disimpulkan bahwa penerapan data mining dengan algoritma K-Medoids dapat diterapkan. penelitian ini mendapatkan cluster sebanyak 3 pengelompokan dan mendapatkan hasil dengan yaitu cluster tinggi berjumlah 3 provinsi, cluster sedang berjumlah 9 provinsi dan cluster rendah berjumlah 22 provinsi. Pengujian cluster dengan algoritma K-Medoids ini menggunakan Silhouette Coefficient dan menunjukkan bahwa cluster optimal untuk pengelompokan penyebaran Covid-19 di Indonesia adalah cluster 3 dengan nilai 0.262463.

Kata kunci: Covid-19, Data mining, Clustering, K-Medoids.

Abstract - Badan Nasional Penanggulangan Bencana (BNPB) has declared a disaster emergency status related to the Covid-19 virus. The spread of Covid-19 cases that have been evenly distributed in all provinces in Indonesia is already a fairly rapid spread and has negative impact on all fields. From the large number of existing Covid-19 data in Indonesia, it is necessary to classify the spread of Covid-19 cases so that it can be utilized by various parties, one of which is the Health Office regarding the grouping of the spread of Covid-19 in 34 provinces in Indonesia so that the distribution pattern can be known using the data mining method. clustering and the algorithm used is K-Medoids. Based on the research that has been done, it can be concluded that the application of data mining with the K-Medoids algorithm can be applied. This study obtained 3 clusters and obtained results with 3 provinces in the high cluster, 9 provinces in the medium cluster and 22 in the low cluster. The cluster testing of K-Medoids algorithm uses the Silhouette Coefficient and shows that the optimal cluster for clustering the spread of Covid-19 in Indonesia is cluster 3 with a value of 0.262463. **Keywords:** Covid-19, Data mining, Clustering, K-Medoids.

1. Pendahuluan

Pandemi Covid-19 masuk di Indonesia pada awal maret 2020. Badan Nasional Penanggulangan Bencana (BNPB) telah menetapkan status darurat bencana terkait virus Covid-19 ini [1][2]. Gejala klinis dari penyakit ini yaitu demam, myalgia, pneumonia, batuk kering dan dapat menyebabkan kegagalan pernafasan karena kerusakan pada alveolar dan dapat berujung kematian. Menurut Komisi Kesehatan Nasional China, penularan Covid-19 terjadi akibat adanya kontak langsung dari orang ke orang dan tetesan cairan yang berasal dari saluran pernafasan [3]. Kasus Covid-19 di provinsi Indonesia yang sudah merata dan penyebarannya yang cukup cepat berdampak negatif pada seluruh bidang [4]. Sekarang ini jumlah kasus Covid-19 di dunia pada tanggal 16 september 2021, berjumlah 227,429,294 kasus. Di Indonesia sendiri kasus Covid-19 pada tanggal 16 september 2021 sudah mencapai 41,309 kasus (worldometer, 2021). Dari banyaknya data Covid-19 yang ada di perlukan pengelompokan penyebaran kasus Covid-19 berdasarkan 34 provinsi di Indonesia untuk mendapatkan informasi pengelompokan penyebaran Covid-19 dengan melakukan pengelompokan penyebaran data menggunakan data mining metode clustering. Dalam penelitian ini algoritma yang digunakan yaitu K-Medoids. K-Medoids juga dikenal dengan *partitioning around medoids* dan merupakan salah satu varian metode K-Means. Algoritma ini melakukan pengelompokan berdasarkan penggunaan medoids bukan dari pengamatan means yang dimiliki oleh setiap cluster, tujuannya yaitu mengurangi sensitivitas terhadap partisi yang dihasilkan oleh nilai ekstrim yang terdapat dalam dataset [5].

Penelitian terkait yang menjadi landasan dari pengelompokan penyebaran Covid-19 di Indonesia yang dilakukan oleh [4], hasil perhitungan pusat cluster yang didapatkan yaitu cluster pertama pada provinsi DKI Jakarta Pusat. Cluster kedua pada provinsi Jawa Barat, dan pusat cluster ketiga yaitu pada provinsi Banten.

Daerah penyebaran Covid-19 terbesar di Indonesia dengan 3 kasus positif dan kasus negative dan kasus meninggal 3 ling besar terdapat pada pusat cluster pertama, penyebaran kasus Covid-19 yang cukup potensial yaitu pada daerah-daerah pada cluster kedua dan ketiga, dari hasil yang di dapatkan harus menjadi perhatian pemerintah untuk menanganinya. Penelitian terkait selanjutnya yaitu oleh [6]. Hasilnya yaitu data yang di uji sebanyak 34 provinsi yang menggunakan dua cluster. 31 provinsi termasuk Cluster 1 (rendah). 3 provinsi termasuk Cluster 2 (nilai tinggi). Dari hasil yang didapatkan, cluster tertinggi dapat menjadi masukan kepada pihak pemerintah Indonesia agar pemerintah lebih memperhatikan provinsi di Indonesia yang memiliki cluster tertinggi agar memberikan penanganan untuk kasus diare.

Dalam penelitian ini penulis melakukan penelitian tentang penyebaran Covid-19 di Indonesia. Penelitian ini bertujuan untuk mengetahui pengelompokan penyebaran Covid-19 tahun 2020 di Indonesia menggunakan algoritma *K-Medoids* dan untuk memberikan informasi dan pengetahuan yang dapat dimanfaatkan oleh berbagai pihak, salah satunya yaitu Dinas Kesehatan tentang pengelompokan penyebaran Covid-19 pada 34 provinsi di Indonesia dengan menerapkan algoritma *K-medoids* untuk dapat mengetahui pola penyebarannya. Objek penelitian berasal dari situs <https://www.kaggle.com/hendratno/covid19-indonesia>.

2. Tinjauan Pustaka

2.1. Data Mining

Data mining adalah pengolahan data untuk menggali pengetahuan dari data yang belum diketahui untuk mendapatkan informasi yang berguna berupa pola atau aturan dan relasi [7]. Berikut beberapa pendapat para ahli tentang data mining:

Menurut Dermawan *data mining* adalah sebuah proses untuk menentukan hubungan kecenderungan dan pola untuk memeriksa kumpulan data yang besar. Penyimpanan data dengan menggunakan pola seperti teknik matematika dan statistik [8]. Menurut [9] *Data mining* yaitu teknik yang menggabungkan teknik analisis data dan merupakan pola-pola penting didalam data. Sederhananya data mining adalah proses eksplorasi, seleksi dan pemodelan beberapa data untuk mencari pola-pola yang tidak diketahui kebenarannya.

2.2. K-Medoids Clustering

Clustering adalah metode yang cukup banyak di pakai. *Clustering* dapat membagikan sebuah data yang tidak mempunyai label agar menjadi beberapa kelomok [10]. Atribut yang digunakan pada penelitian ini ada 4 yaitu: *Location*, *New Cases*, *New Deaths*, dan *New Recovered*. Data Covid-19 2020. Data tersebut akan diolah menggunakan *clustering* yaitu cluster rendah, cluster sedang, dan cluster tinggi.



Gambar 1 Alur Proses Perhitungan Algoritma K-Medoids

Flowchart diatas menggambarkan tahapan-tahapan bagaimana penerapan algoritma *K-Medoids* pada proses *data mining*. Merode *K-Medoid* memiliki langkah-langkah sebagai berikut:

- Inisialisasi pusat cluster (*k*)
- Data dipilih secara acak untuk digunakan menjadi *medoid* awal dan sebanyak jumlah cluster yang ditentukan.
- Melakukan perhitungan pada jarak terdekat data terhadap *medoid* awal menggunakan persamaan jarak *Euclidian Distance* dengan persamaan:

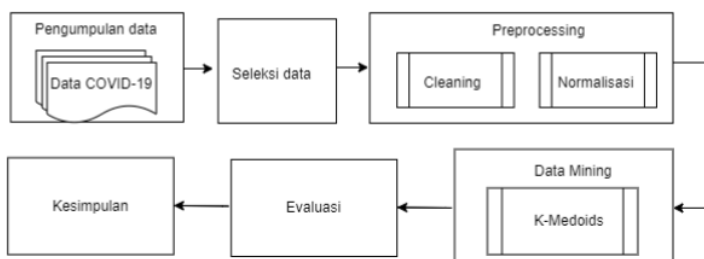
$$d(x, y) = \sqrt{\sum (x_i - y_i)^2}$$

- Mengelompokkan data menggunakan *medoid* yang berjarak paling dekat, lalu hitung jumlah *cost*.

- e. Hitung simpangan (S) menggunakan perhitung nilai total iterasi ke-1 dan iterasi ke-2, maka tukar objek dengan data *cluster* untuk membentuk sekumpulan k objek baru sebagai *medoid*.
- f. Ulangi proses c – e sampai semua data telah menjadi *medoid*.

3. Metode Penelitian

Tahapan yang dilakukan untuk melakukan clustering data Covid-19 terdiri dari 6 tahapan utama yaitu: pengumpulan data, seleksi data, *preprocessing*, transformasi, *data mining* dan kesimpulan seperti ditunjukkan dalam gambar 2:



Gambar 2 Tahapan Metodologi Penelitian

3.1 Pengumpulan Data

Tahapan pengumpulan data pada penelitian ini dilakukan dengan mengambil informasi mengenai penelitian terkait dari studi literatur. Dataset yang dipakai yaitu data Covid-19 di Indonesia yang berasal dari <https://www.kaggle.com/hendratno/covid19-indonesia> yang diakses pada tanggal 11 april 2021 dan pada penelitian ini menggunakan data pada tahun 2020 yang berjumlah 9959 *records* data. Data bisa dilihat pada table berikut:

Tabel 1 Data Covid-19 Tahun 2020

No	Date	Location iso code	Location	New case	New Date	New Recovery	Growth Factor of New Deaths
1	03/01/2020	ID-JK	DKI Jakarta	2	0	0	
2	03/02/2020	ID-JK	DKI Jakarta	2	0	0	1.00
3	03/02/2020	IDN	Indonesia	2	0	0	
...
9957	12/31/2020	ID-SB	Sumatra Barat	3	26	88	3.00
9958	12/31/2020	ID-SS	Sumatra Selatan	1	50	41	0.50
9959	12/31/2020	ID-SU	Sumatra Utara	0	82	1	0.00

1. Data Selection

Data selection atau seleksi data adalah tahapan pemilihan data yang akan digunakan di dalam proses *data mining*. Pada data Covid-19 tahun 2020 yang di ambil dari www.kaggle.com terdapat 37 atribut dan dari atribut yang ada atribut yang dipakai ada 4 yaitu: *Location*, *New Cases*, *New Deaths*, dan *New Recovered*.

2. Pre-processing / Cleaning

Tahapan *pre-processing* ini bertujuan untuk memeriksa dan membersihkan kesalahan pada data. Pada tahapan ini dilakukan normalisasi data yang memiliki tujuan yaitu mendapatkan jarak yang lebih kecil antara nilai di setiap atribut, dengan menentukan nilai tertinggi (max) dan nilai terendah (min) dengan rumus:

$$X_{new} = \frac{X_{old} - X_{min}}{X_{max} - X_{min}}$$

Berikut adalah hasil dari data yang sudah di normalisasikan.

Tabel 2 Data Covid-19 yang Sudah di Normalisasi

No	Location	New Cases	New Deaths	New Recovered
1	Aceh	0,0375	0,0574	0,0363
2	Bali	0,0863	0,0846	0,0707
3	Banten	0,0892	0,0652	0,0551
4	Bengkulu	0,0092	0,0150	0,0083
5	DKI Jakarta	1,0000	0,5559	1,0000
6	Daerah Istimewa Yogyakarta	0,0563	0,0403	0,0427
7	Gorontalo	0,0105	0,0136	0,0129
8	Jambi	0,0071	0,0052	0,0068
9	Jawa Barat	0,4502	0,1969	0,4255
10	Jawa Tengah	0,4399	0,5739	0,3249
11	Jawa Timur	0,4532	1,0000	0,4338
12	Kalimantan Barat	0,0065	0,0000	0,0094
13	Kalimantan Selatan	0,0737	0,0965	0,0771
14	Kalimantan Tengah	0,0430	0,0300	0,0218
15	Kalimantan Timur	0,1386	0,1236	0,1298
16	Kalimantan Utara	0,0102	0,0053	0,0041
17	Kepulauan Bangka Belitung	0,0022	0,0017	0,0032
18	Kepulauan Riau	0,0278	0,0252	0,0298
19	Lampung	0,0239	0,0429	0,0190
20	Maluku	0,0209	0,0079	0,0204
21	Maluku Utara	0,0046	0,0110	0,0070
22	Nusa Tenggara Barat	0,0205	0,0405	0,0165
23	Nusa Tenggara Timur	0,0012	0,0043	0,0000
24	Papua	0,0622	0,0210	0,0358
25	Papua Barat	0,0223	0,0126	0,0260
26	Riau	0,1270	0,0960	0,1340
27	Sulawesi Barat	0,0000	0,0016	0,0021
28	Sulawesi Selatan	0,1605	0,0979	0,1504
29	Sulawesi Tengah	0,0089	0,0140	0,0041
30	Sulawesi Tenggara	0,0329	0,0214	0,0336
31	Sulawesi Utara	0,0426	0,0491	0,0359
32	Sumatera Barat	0,1187	0,0858	0,1029
33	Sumatera Selatan	0,0545	0,0986	0,0499
34	Sumatera Utara	0,0894	0,1127	0,0869

4. Hasil dan Pembahasan

4.1 Penerapan Metode *K-Medoids Clustering*

Berikut adalah perhitungan manual metode *K-Medoids clustering*.

- Menetapkan jumlah dari *cluster* (*k*) yang diinginkan. Inisialisasi pusat *cluster* sebanyak 3 *cluster*.
- Untuk mendapatkan medoid awal sebanyak jumlah *cluster* yang ditentukan, digunakan data yang di pilih secara acak

Tabel 3 Table Medoid Awal

Location	New Cases	New Deaths	New Recovered
Riau(C1)	0,1270	0,0960	0,1340
Jawa Barat(C2)	0,4502	0,1969	0,4255
Sulawesi Selatan(C3)	0,0545	0,0986	0,0499

c) Selanjutnya dilakukan perhitungan jarak terdekat dengan data *medoid* awal menggunakan persamaan ukuran jarak *Euclidian Distance*

$$d_{Aceh,C1} = \sqrt{(0,1270 - 0,0375)^2 + (0,0960 - 0,0574)^2 + (0,1340 - 0,0363)^2} = 0,137942$$

$$d_{Aceh,C2} = \sqrt{(0,4502 - 0,0375)^2 + (0,1969 - 0,0574)^2 + (0,4255 - 0,0363)^2} = 0,584168$$

$$d_{Aceh,C3} = \sqrt{(0,0545 - 0,0375)^2 + (0,0986 - 0,0574)^2 + (0,0499 - 0,0363)^2} = 0,172596$$

Hasil dari keseluruhan dapat dilihat pada table 4 berikut:

Tabel 4 Hasil Perhitungan Algoritma *K-Medoids* Iterasi ke-1

No	Location	Cost1	Cost2	Cost3	Kedekatan	Cluster
1	Aceh	0,137942	0,584168	0,172596	0,137941773	1
2	Bali	0,076092	0,520541	0,109756	0,076092269	1
3	Banten	0,092744	0,533777	0,12349	0,092743755	1
4	Bengkulu	0,190365	0,633795	0,223577	0,190365265	1
5	DKI Jakarta	1,312891	0,872463	1,279166	0,872463328	2
6	Daerah Istimewa Yogyakarta	0,128222	0,571235	0,160632	0,128222157	1
7	Gorontalo	0,187153	0,63027	0,220318	0,187152602	1
8	Jambi	0,196964	0,639091	0,229712	0,196963711	1
9	Jawa Barat	0,446838	0	0,411582	0	2
.
.
.
34	Sumatera Selatan	0,111008	0,554373	0,146086	0,11100839	1
35	Sumatera Utara	0,062562	0,501995	0,096566	0,06256156	1
Total Cost					6,221375028	

Setelah mendapatkan hasil jarak dari setiap objek (*cost*) pada iterasi ke-1 maka lanjut ke iterasi ke-2. Inisialisasi *medoid* baru (*non-medoid*) pada iterasi ke-2 dapat dilihat di table 5 berikut:

Tabel 5 Tabel *Medoid* Baru

Location	New Cases	New Deaths	New Recovered
Aceh (C1)	0,0375	0,0574	0,0363
Jambi (C2)	0,0071	0,0052	0,0068
Kalimantan Barat (C3)	0,0065	0,0000	0,0094

Gunakan medoid baru untuk menghitung jarak objek pada iterasi ke-2. Selanjutnya keseluruhan hasil iterasi ke-2 diperoleh yang ditunjukkan tabel 6 sebagai berikut :

Tabel 6 Hasil Perhitungan Algoritma *K-Medoids* Iterasi ke-2

No	Location	Cost1	Cost2	Cost3	Kedekatan	Cluster
1	Aceh	0	0,06727	0,070597	0	1
2	Bali	0,065571	0,129097	0,131485	0,065570852	1
3	Banten	0,055509	0,112561	0,114782	0,055508695	1
4	Bengkulu	0,058209	0,010151	0,015278	0,010151124	2
5	DKI Jakarta	1,450362	1,508513	1,509112	1,450361896	1
6	Daerah Istimewa Yogyakarta	0,026112	0,070271	0,072183	0,026112034	1
7	Gorontalo	0,056556	0,010933	0,014612	0,010933056	2
8	Jambi	0,06727	0	0,005815	0	2
9	Jawa Barat	0,584168	0,639091	0,639398	0,584167857	1
.
.
.
35	Sumatera Utara	0,091119	0,15731	0,159935	0,091119372	1
Total Cost					5,103008055	

d) Menghitung jumlah simpangan

Sesudah nilai jarak dari iterasi ke-1 dan iterasi k-2 didapatkan, kemudian jumlah simpangan dihitung untuk mendapatkan nilai selisih dengan cara mengurangi total *cost* baru – total *cost* lama. Menggunakan ketentuan $S < 0$, dan tentukan medoid yang baru dengan menukar nilai objek.

$$\begin{aligned}
 S &= \text{Total } cost \text{ baru} - \text{Total } cost \text{ lama} \\
 &= 5,103008055 - 6,221375028 \\
 &= -1,118366973
 \end{aligned}$$

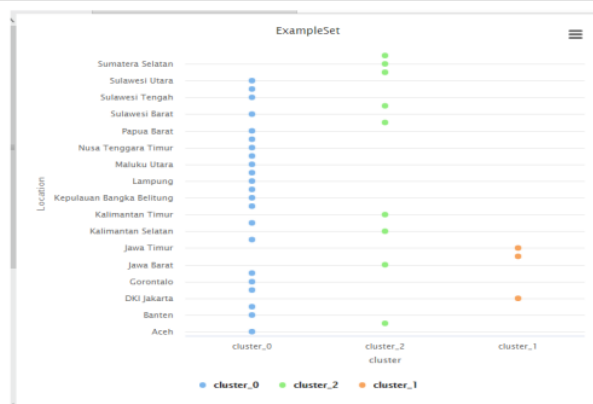
e) Ulangi tahapan sebelumnya sampai $S > 0$, sehingga hasil masing- masing *cluster* mendapat anggota *cluster*. Kemudian hitung jumlah simpangan dengan cara mengurangi nilai total *cost* baru – total *cost* lama setelah jarak iterasi ke-3 didapatkan.

$$\begin{aligned}
 S &= \text{Total dari } cost \text{ baru} - \text{Total dari } cost \text{ lama} \\
 &= 21,36510265 - 6,870660836 \\
 &= 15,494441819
 \end{aligned}$$

Dengan nilai simpangan $S > 0$ maka proses *clustering* di hentikan, sehingga dapat diperoleh anggota tiap *cluster*.

4.1 Implementasi Data Mining Menggunakan Rapidminer

Tahap selanjutnya yaitu dilakukan pengelompokan data menggunakan Tool *Rapidminer*. Hasil akhir yang didapatkan berupa pengelompokan hasil dari pengujian data akan terlihat *cluster* dengan masing-masing anggota. Bisa dilihat pada gambar berikut:



Gambar 3 Tampilan Scatter

Berdasarkan gambar diatas terdapat 3 warna pada masing masing provinsi yaitu warna biru, merah dan hijau. Warna biru menunjukkan *cluster* 0, warna merah menunjukkan *cluster* 1 dan warna hijau menunjukkan *cluster* 2. Hasil gambar diatas memperoleh 22 provinsi (Sulawesi Utara, Maluku Utara, Sulawesi Tenggara, Sulawesi Barat, Papua Barat, Papua, Nusa Tenggara Timur, Sulawesi Tengah, Nusa Tenggara Barat, Maluku, Lampung, Kepulauan Riau, Kepulauan Bangka Belitung, Kalimantan Utara, Kalimantan Tengah, Jambi, Gorontalo, Kalimantan Barat, Daerah Istimewa Yogyakarta, Bengkulu, Banten, Aceh) pada *cluster* rendah, 9 provinsi (Sumatra Utara, Sumatra Selatan, Kalimantan Selatan, Sumatra Barat, Sulawesi Selatan, Kalimantan Timur, Riau, Jawa Barat, dan Bali) pada *cluster* sedang dan 3 provinsi (Jawa Timur, Jawa Tengah, dan DKI Jakarta) pada *cluster* tinggi.

Cluster Model

```
Cluster 0: 22 items
Cluster 1: 3 items
Cluster 2: 9 items
Total number of items: 34
```

Gambar 4 Tampilan Text View Clustering Model

4.2 Pengujian menggunakan *Silhouette Coefficient*

Pengujian hasil *cluster* dari algoritma *K-medoids* menggunakan *Silhouette Coefficient* bertujuan untuk melihat kualitas dan kekuatan dari *cluster* yang dibentuk dari proses Analisa pengelompokan Covid-19 pada 34 provinsi di Indonesia. Hasil yang di dapat dari pengujian menggunakan *Silhouette Coefficient* yaitu *cluster* 3 dengan nilai 0.262463 dengan menggunakan *library python Scikit-learn*. Hal ini dapat menunjukkan bahwa pengelompokan untuk kasus penyebaran covid-19 yang optimal yaitu 3 *cluster*. Dengan demikian, hasil clustering penelitian ini tergolong pada *intermediate case*[11].

5. Kesimpulan

Berdasarkan penelitian yang telah dilakukan dapat disimpulkan bahwa penerapan data mining dengan algoritma *K-medoids* dapat diterapkan dengan baik. Data yang digunakan bersumber dari <https://www.kaggle.com/hendratno/covid19-indonesia>. Data yang digunakan adalah data covid-19 tahun 2020. Data yang digunakan sebanyak 34 provinsi. Penelitian ini menetapkan *cluster* sebanyak 3 pengelompokan dan mendapatkan hasil dengan *cluster* tinggi berjumlah 3 provinsi, *cluster* sedang berjumlah 9 provinsi dan *cluster* rendah berjumlah 22 provinsi. Pola penyebaran yang di dapatkan dari penelitian ini yaitu, *cluster* rendah terdapat 22 provinsi (Sulawesi Utara, Sulawesi Tenggara, Sulawesi Tengah, Sulawesi Barat, Papua Barat, Papua, Nusa Tenggara Timur, Nusa Tenggara Barat, Maluku Utara, Maluku, Lampung, Kepulauan Riau, Kepulauan Bangka Belitung, Kalimantan Tengah, Kalimantan Barat, Jambi, Gorontalo, Daerah Istimewa Yogyakarta, Bengkulu, Banten, Aceh). Untuk *cluster* sedang terdapat 9 provinsi (Sumatra Utara, Sumatra

Selatan, Sumatra Barat, Kalimantan Utara, Sulawesi Selatan, Riau, Kalimantan Timur, Kalimantan Selatan, Jawa Barat, dan Bali). Pada *cluster* tinggi terdapat 3 provinsi (Jawa Timur, Jawa Tengah, dan DKI Jakarta). Hasil pengujian cluster dari algoritma K-Medoids menggunakan *Silhouette Coefficient* menunjukkan bahwa *cluster* optimal untuk pengelompokan penyebaran Covid-19 di Indonesia adalah cluster 3 dengan nilai 0.262463.

Daftar Pustaka

- [1] S. Hikmawan, A. Pardamean, and S. N. Khasanah, "Sentimen Analisis Publik Terhadap Joko Widodo Terhadap Wabah Covid-19 Menggunakan Metode Machine Learning," vol. 20, no. 2, pp. 167–176, 2020.
- [2] Z. Zulfan *et al.*, "Perancangan Storyboard Konten Animasi 3 Dimensi Untuk Edukasi Anak Usia Sekolah Dasar Tentang Mitigasi Penyebaran COVID-19," *J. Serambi Eng.*, vol. 7, no. 1, 2022.
- [3] N. Nurhaliza, "IJIRSE : Indonesian Journal of Informatic Research and Software Engineering," no. April, pp. 1–8, 2020.
- [4] N. Dwitri, J. A. Tampubolon, S. Prayoga, F. Ilmi Zer, and D. Hartama, "Penerapan Algoritma K-Means Dalam Menentukan Tingkat Kepuasan Pembelajaran Online Pada Masa Pandemi Covid-19 di Indonesia," *Jti (Jurnal Teknol. Informasi)*, vol. 4, no. 1, pp. 101–105, 2020.
- [5] D. A. S. Simamora, M. T. Furqon, and B. Priyambadha, "Clustering Data Kejadian Tsunami Yang Disebabkan Oleh Gempa Bumi Dengan Menggunakan Algoritma K-Medoids," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 1, no. 8, pp. 635–640, 2017.
- [6] F. Hardiyanti, H. S. Tambunan, and I. S. Saragih, "Penerapan Metode K-Medoids Clustering Pada Penanganan Kasus Diare Di Indonesia," *KOMIK (Konferensi Nas. Teknol. Inf. dan Komputer)*, vol. 3, no. 1, pp. 598–603, 2019.
- [7] S. Al Syahdan and A. Sindar, "Data Mining Penjualan Produk Dengan Metode Apriori Pada Indomaret Galang Kota," *J. Nas. Komputasi dan Teknol. Inf.*, vol. 1, no. 2, 2018.
- [8] A. Darmawan, N. Kustian, and W. Rahayu, "Implementasi Data Mining Menggunakan Model SVM untuk Prediksi Kepuasan Pengunjung Taman Tabebuya," *STRING (Satuan Tulisan Ris. dan Inov. Teknol.*, vol. 2, no. 3, p. 299, 2018.
- [9] N. A. Hasibuan *et al.*, "Implementasi Data Mining Untuk Pengaturan Layout," vol. 4, no. 4, pp. 6–11, 2017.
- [10] H. A. Negara, A. R. Putra, and U. Enri, "Clustering Clustering Data Ekspor Buah-Buahan Berdasarkan Negera Tujuan Menggunakan Algoritma K-Means," *Bina Insa. Ict J.*, vol. 8, no. 1, p. 73, 2021.
- [11] P. J. Rousseeuw, "Silhouettes: A graphical aid to the interpretation and validation of cluster analysis," *J. Comput. Appl. Math.*, vol. 20, no. C, pp. 53–65, 1987.

ORIGINALITY REPORT

16%

SIMILARITY INDEX

11%

INTERNET SOURCES

11%

PUBLICATIONS

6%

STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to Higher Education Commission Pakistan Student Paper	6%
2	Sutrisno, Razali Haron. "INCREASING THE ROLE OF ZAKAT INSTITUTIONS IN POVERTY REDUCTION THROUGH PRODUCTIVE ZAKAT PROGRAMS IN INDONESIA", Humanities & Social Sciences Reviews, 2020 Publication	4%
3	jurnal.una.ac.id Internet Source	3%
4	jurnal.unsil.ac.id Internet Source	2%
5	repository.usd.ac.id Internet Source	2%

Exclude quotes On

Exclude matches < 2%

Exclude bibliography On