

# Klasterisasi Sebaran Kasus Covid-19 Di Kota Denpasar Menggunakan Algoritme *K-Means*

Ida Bagus Gede Sarasvananda<sup>1\*</sup>, I Gusti Made Ngurah Desnanjaya<sup>2</sup>, Yunita Dewi<sup>3</sup>  
<sup>1,2,3</sup> STMIK STIKOM Indonesia, Jl. Tukad Pakerisan No. 97 Panjer Denpasar – Bali  
<sup>1</sup>sarasvananda@stiki-indonesia.ac.id, <sup>2</sup>ngurah.desnanjaya@stiki-indonesia.ac.id,  
<sup>3</sup>yunita.dewi@gmail.com

## Abstract

Various efforts have been made by the government to tackle the spread of Covid-19 in Denpasar City. Starting from creating a data collection system for Covid-19 patients and a map of the distribution of Covid-19 patients based on villages, which are updated every day and open for access to the general public. In the observations made by researchers, the Denpasar City Government needs to add efforts to organize the distribution of Covid-19 cases, by knowing the characteristics of Covid-19 patients and grouping villages based on similar characteristics so that the distinctive characteristics of each urban village group can be identified in Denpasar City. The problem that will be discussed in this research is how to cluster the distribution of Covid-19 cases using the K-means algorithm. The purpose of this proposed study is to cluster Covid-19 data using the K-Means algorithm. The clustering of the distribution of Covid-19 cases in Denpasar City was successfully carried out using the K-Means algorithm and the number of clusters used was 3, with cluster 0 (Low Cluster) of 16 villages / wards, cluster 2 (Regular Cluster) of 21 villages / wards, and cluster 1 (High Cluster) of 6 villages / wards. Cluster validity was tested using the Davies Bouldin method with the resulting value of -0.522.

**Keywords:** clustering, covid-19, k-means

## Abstrak

Berbagai upaya telah dilakukan oleh pemerintah untuk menanggulangi penyebaran Covid-19 di Kota Denpasar. Mulai dari membuat sistem informasi pendataan pasien Covid-19 dan peta sebaran pasien Covid-19 berdasarkan desa/kelurahan yang diupdate setiap hari dan terbuka untuk diakses oleh masyarakat umum. Dalam pengamatan yang dilakukan oleh peneliti, Pemerintah Kota Denpasar perlu menambahkan upaya untuk mengorganisir sebaran kasus Covid-19 khususnya di Kota Denpasar, dengan mengetahui karakteristik pasien Covid-19 dan mengelompokkan kelurahan-kelurahan berdasarkan kesamaan karakteristik sehingga dapat diketahui ciri khas dari masing masing kelompok kelurahan di Kota Denpasar. Masalah yang akan dibahas dalam penelitian ini yaitu, bagaimana melakukan klasterisasi sebaran kasus Covid-19 algoritme K-means. Tujuan dari penelitian yang diusulkan ini adalah untuk melakukan klasterisasi data Covid-19 menggunakan algoritme K-Means. Klasterisasi sebaran kasus Covid-19 di Kota Denpasar berhasil dilakukan menggunakan algoritme K-Means dan jumlah klaster yang digunakan yaitu 3, dengan klaster 0 (Low Cluster) sebanyak 16 desa/kelurahan, klaster 2 (Regular Cluster) sebanyak 21 desa/kelurahan, dan klaster 1 (High Cluster) sebanyak 6 desa/kelurahan. Validitas klaster diuji menggunakan metode Davies Bouldin dengan nilai yang dihasilkan yaitu -0.522.

**Kata kunci:** klastering, covid-19, k-means

## 1. PENDAHULUAN

Sudah lebih dari setahun pandemi COVID-19 melanda Indonesia semenjak kasus terkonfirmasi positif COVID-19 diumumkan pertama kali pada Senin 2 Maret 2020. Tidak hanya Indonesia, hampir seluruh negara di dunia mengalami pandemi COVID-19. Pandemi COVID-19 adalah pandemi

berskala global yang disebabkan oleh virus SARS-CoV-2 (Severe acute respiratory syndrome coronavirus 2) [1]. Virus SARS-CoV-2 tergolong dalam kategori virus berbahaya karena dapat menyebabkan kematian.

Hingga saat ini kasus terkonfirmasi positif di Indonesia masih terus bertambah setiap harinya. Berdasarkan data yang diperoleh dari laman website [infocorona.baliprov.go.id](http://infocorona.baliprov.go.id) pada tanggal 27 Maret 2021 jumlah kasus terkonfirmasi positif di Indonesia sebanyak 1.492.002, sembuh sebanyak 1.327.121, dalam perawatan sebanyak 124.517 dan meninggal sebanyak 40.364 orang. Sedangkan di Provinsi Bali, jumlah kasus terkonfirmasi positif sebanyak 39.323, sembuh sebanyak 36.357, dalam perawatan sebanyak 1.640 dan meninggal sebanyak 1.111 orang. Dari 9 Kabupaten/Kota yang ada di Provinsi Bali, Kota Denpasar merupakan penyumbang kasus terkonfirmasi positif terbanyak yaitu sebanyak 12.244 orang.

	Kabupaten	Kasus Terkonfirmasi	Dalam Perawatan	Sembuh	Meninggal
1.	JEMBRANA	2180	38	2070	72
2.	TABANAN	4141	137	3862	142
3.	BADUNG	7551	335	7026	190
4.	DENPASAR	12.244	641	11355	248
5.	GIANYAR	4470	3	4338	129
6.	BANGLI	2057	53	1925	79
7.	KLUNGKUNG	1584	93	1444	47
8.	KARANGASEM	1654	56	1514	84
9.	BULELENG	2894	188	2589	117
10.	KABUPATEN LAINNYA	462	78	380	4
11.	WARGA NEGARA ASING	86	18	64	4

Sumber: [infocorona.baliprov.go.id](http://infocorona.baliprov.go.id)

**Gambar 1.** Kasus Covid-19 Per Kabupaten di Provinsi Bali

Berbagai upaya telah dilakukan oleh pemerintah untuk menanggulangi penyebaran Covid-19 di Kota Denpasar. Mulai dari membuat sistem informasi pendataan pasien Covid-19 dan peta sebaran pasien Covid-19 berdasarkan kelurahan yang diupdate setiap hari dan terbuka untuk diakses oleh masyarakat umum melalui laman <https://safecity.denpasarkota.go.id/id/covid19>. Selain membuat sistem informasi pendataan pasien Covid-19, Pemerintah Kota Denpasar juga membuat Peraturan Walikota Denpasar untuk menanggulangi penyebaran Covid-19 Nomor 32 Tahun 2020 tentang Pembatasan Kegiatan Masyarakat di Desa, Kelurahan dan Desa Adat dalam percepatan penanganan Covid-19. Berbasis Desa/Kelurahan dan bahkan Desa Adat dimaksudkan karena



penyebaran kasus Covid-19 telah berdampak secara multidimensional, sehingga dibutuhkan kolaborasi dari seluruh [2].

Dalam pengamatan yang dilakukan oleh peneliti, Pemerintah Kota Denpasar perlu menambahkan upaya untuk mengorganisir sebaran kasus Covid-19 khususnya di Kota Denpasar, dengan mengetahui karakteristik pasien Covid-19 dan mengelompokkan kelurahan-kelurahan berdasarkan kesamaan karakteristik sehingga dapat diketahui ciri khas dari masing-masing kelompok kelurahan di Kota Denpasar.

Beberapa penelitian yang berkaitan dengan klusterisasi kasus Covid-19 telah dilakukan, diantaranya yaitu: memanfaatkan klustering untuk melakukan analisis pengelompokan penyebaran dan perkembangan kasus Covid-19 di Indonesia [3],[4],[5]. Yustanti, W dkk [6] melakukan klustering wilayah Kota/Kabupaten berdasarkan data harian yang disajikan dalam situs resmi Pemerintah Jawa Timur menggunakan metode *K-Means*. Beberapa peneliti yang lain seperti [1],[7] juga menggunakan klustering *K-Means* untuk melakukan pengelompokan penyebaran kasus Covid-19. Pemanfaatan klustering *K-Means* dalam pengelompokan data berdasarkan wilayah telah terbukti membantu dalam pengambilan keputusan yang tepat dalam mengurangi sebaran Covid-19. Sarasvananda, dkk [8] menggunakan *K-Means* dengan *semantic similarity* untuk melakukan klusterisasi data pasien. Pengelompokan kasus dan kematian akibat Covid-19 di Asia Tenggara juga pernah dilakukan oleh Hutagalung, dkk [9].

Berdasarkan latar belakang yang telah dipaparkan di atas, dapat dirumuskan masalah yang akan dibahas dalam penelitian ini yaitu, bagaimana melakukan klusterisasi sebaran kasus Covid-19 berdasarkan Desa/Kelurahan di Kota Denpasar menggunakan algoritme *K-means*. Tujuan dari penelitian yang diusulkan ini adalah untuk melakukan klusterisasi data Covid-19 berdasarkan data harian yang disajikan resmi oleh Pemerintah Kota Denpasar menggunakan algoritme *K-Means*, di mana data dikategorikan berdasarkan *mean* terdekat, sehingga Desa/Kelurahan dapat dikategorikan menjadi kluster-kluster. Diharapkan juga dengan klusterisasi Covid-19 berdasarkan Desa/Kelurahan dapat membantu Pemerintah Kota Denpasar dalam pengambilan keputusan yang tepat dalam mengoptimalkan penanganan Covid-19 di Kota Denpasar.

## 2. METODOLOGI PENELITIAN

Penelitianawali dengan melakukan literatur terhadap data sebaran kasus covid-19 setiap Desa / Kelurahan di Kota Denpasar melalui situs resmi Pemerintah Kota Denpasar, kemudian dilanjutkan literatur *review* dari jurnal yang berkaitan dengan masalah dan metode yang digunakan dalam melakukan kluster data.

### 2.1. Sumber Data

Data yang digunakan dalam penelitian ini merupakan data harian yang disajikan resmi oleh Pemerintah Kota Denpasar pada laman

<https://safecity.denpasarkota.go.id/id/covid19> mulai dari tanggal 2 Maret 2020 sampai dengan tanggal 27 Maret 2021.

## 2.2. Variabel Penelitian

Variabel penelitian yang digunakan pada praktikum ini akan diuraikan pada Tabel 1 berikut.

**Tabel 1.** Variabel Penelitian

Variabel	Keterangan
X <sub>1</sub>	Sembuh
X <sub>2</sub>	Meninggal
X <sub>3</sub>	Masih Dirawat
X <sub>4</sub>	Kasus Terkonfirmasi positif Covid-19

## 2.3. Struktur Data

Struktur data yang digunakan pada praktikum ini ditunjukkan pada tabel sebagai berikut.

**Tabel 2.** Struktur Data

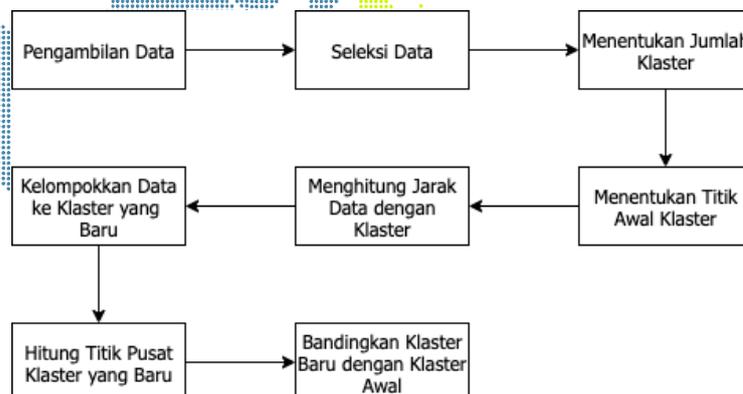
Desa/Kelurahan	X1	X2	X3	X4
1	X1,1	X1,2	X1,3	X1,4
2	X2,1	X2,2	X2,3	
3	X3,1	X3,2	X3,3	X3,4
....	....	....	....	....
44	X44,1	X44,2	X44,3	X44,4

Langkah analisis yang dilakukan pada praktikum ini adalah sebagai berikut:

- 1) Melakukan identifikasi dan merumuskan masalah.
- 2) Mengumpulkan data sekunder yang diperoleh dari website resmi Safecity Pemerintah Kota Denpasar.
- 3) Melakukan pengolahan dan analisis karakteristik data COVID-19.
- 4) Menentukan kluster optimum berdasarkan metode *silhouette coefficient*.
- 5) Melakukan analisis kluster dengan metode non hierarki menggunakan *K-Means*.
- 6) Menarik kesimpulan dan saran.

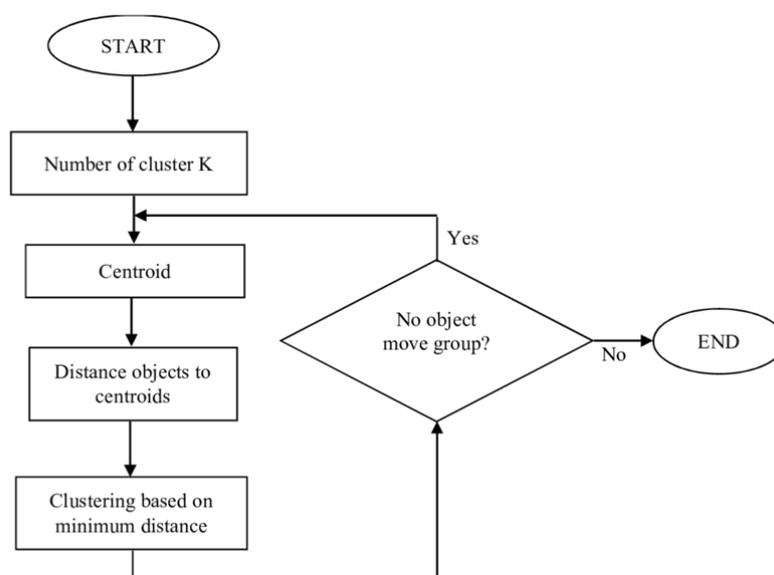
## 2.4. Penerapan Metode

Penerapan metode *K-means* pada penelitian Klasterisasi Persebaran Virus Corona di Kota Denpasar dilakukan dalam beberapa tahap seperti disajikan pada Gambar 2.



**Gambar 2.** Penerapan Metode

Langkah pertama yang dilakukan yaitu melakukan pengambilan data berdasarkan data harian yang disajikan pada situs resmi Pemerintah Kota Denpasar [safecity.denpasarkota.go.id](http://safecity.denpasarkota.go.id), dilanjutkan dengan melakukan pemilihan data yang akan diproses sesuai dengan variabel dan struktur data penelitian yang telah ditetapkan. Kemudian dilanjutkan dengan melakukan klusterisasi dengan algoritme *K-Means*. Langkah pertama yang dilakukan dalam algoritme *K-Means* yaitu menentukan banyaknya jumlah kluster yang akan dibentuk. Penentuan banyaknya jumlah kluster berpengaruh terhadap penentuan jumlah *centroid*. Jika jumlah kluster yang akan dibentuk sebanyak tiga, maka akan memilih sebanyak tiga data sebagai *centroid*. Nilai *initial centroid* pada iterasi yang pertama diberikan secara acak. Ketika *centroid* awal sudah dipilih, selanjutnya adalah menghitung kemiripan untuk setiap data ke masing-masing *centroid*.



**Gambar 3.** Flowchart algoritme *K-Means*



Data yang memiliki nilai jarak yang dekat dengan *centroid* pada suatu kluster tertentu maka data tersebut dikategorikan dan dialokasikan pada kluster tersebut. Proses untuk menghitung nilai kemiripan data dengan *centroid* menggunakan pendekatan *euclidean distance*. Setelah semua data dialokasikan pada suatu kluster, langkah selanjutnya yaitu dilakukan pengecekan konvergensi hasil kluster dengan membandingkan hasil kluster pada iterasi sebelumnya dengan hasil kluster pada iterasi yang sedang berjalan atau menggunakan nilai fungsi objektif yang ditentukan. Jika hasilnya sama atau apabila perubahan nilai fungsi objektif yang digunakan berada di bawah nilai *threshold* yang ditentukan maka data hasil klustering telah konvergen, tetapi jika berbeda atau apabila perubahan nilai fungsi objektif yang digunakan berada di atas nilai *threshold* yang ditentukan maka belum dikatakan konvergen sehingga perlu dilakukan iterasi selanjutnya dan menentukan kembali *centroid* baru berdasarkan data dari masing-masing kluster. Penentuan *centroid* baru dilakukan dengan mencari nilai jarak rata-rata dari seluruh anggota di setiap kluster.

Langkah tersebut akan berulang kembali hingga tidak ada lagi perubahan keanggotaan dari setiap kluster sehingga dapat dikatakan data telah konvergen. Sehingga hasil akhir dari metode ini merupakan pengelompokan desa / kelurahan yang memiliki jarak yang dekat antar data pada suatu kluster.

### 3. HASIL DAN PEMBAHASAN

#### 3.1. Pengelohan dan Analisis Karakteristik data COVID-19

Data yang digunakan yaitu data harian sebaran kasus Covid-19 per Desa / Kelurahan yang bersumber dari website resmi Pemerintah Kota Denpasar. Tabel 3 merupakan data yang telah diolah dan akan digunakan untuk melakukan klusterisasi.

**Tabel 3.** Data yang digunakan untuk klustering

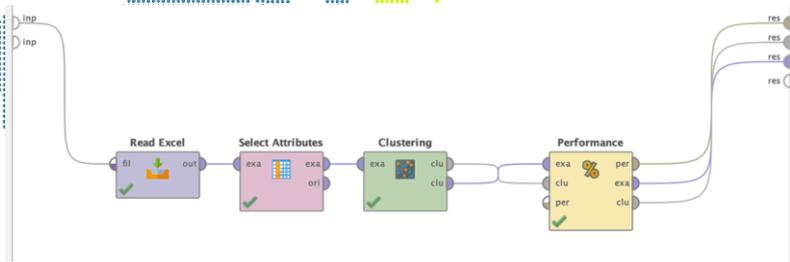
Desa/Kelurahan	Sembuh	Meninggal	Masih Dirawat	Kasus konfirmasi positif
Pemecutan Kelod	442	12	62	516
Padangsambian Kelod	314	9	36	359
Dauh Puri	235	4	34	273
Pemogan	492	12	33	537
Dauh Puri Kelod	226	5	28	259
Pedungan	399	2	26	427
Sesetan	709	12	25	746
Ubung Kaja	490	8	23	521
Sidakarya	330	5	21	356
Tonja	315	5	20	340
Kesiman Kertalangu	296	9	18	323
Penatih	215	5	18	238
Peguyangan	288	9	17	314
Panjer	523	6	16	545

Desa/Kelurahan	Sembuh	Meninggal	Masih Dirawat	Kasus konfirmasi positif
Sumerta Kelod	283	3	15	301
Dauh Puri Kauh	162	1	14	177
Padangsambian Kaja	364	9	11	384
Dauh Puri Kangin	47	3	9	59
Peguyangan Kangin	313	7	9	329
Renon	277	6	9	292
Sanur Kauh	213	9	9	231
Sumerta Kauh	134	5	9	148
Kesiman	224	6	7	237
Pemecutan	250	12	7	269
Pemecutan Kaja	414	7	7	428
Tegal Harum	253	6	7	266
Tegal Kertha	278	2	7	287
Kesiman Petilan	160	5	6	171
Dangin Puri Kangin	256	5	5	266
Peguyangan Kaja	169	1	5	175
Sanur Kaja	141	8	5	154
Dangin Puri	115	8	4	127
Dangin Puri Kaja	230	6	4	240
Dauh Puri Kaja	153	7	4	164
Padangsambian	628	7	4	639
Penatih Dangin Puri	146	2	4	152
Ubung	172	5	4	181
Sumerta	179	3	3	185
Dangin Puri Kelod	185	3	2	190
Sumerta Kaja	141	6	2	149
Dangin Puri Kauh	49	2	1	52
Sanur	183	6	1	190
Serangan	37	1	1	39

Data variabel yang digunakan dalam klusterisasi sebaran kasus Covid-19 di Kota Denpasar meliputi  $X_1$  = sembuh,  $X_2$  = Meninggal,  $X_3$  = Masih dirawat, dan  $X_4$  = Kasus terkonfirmasi positif. Tipe data yang digunakan dari keseluruhan variabel yaitu integer, sehingga untuk menghitung jarak terdekat dengan pusat kluster bisa menggunakan metode *euclidean distance*.

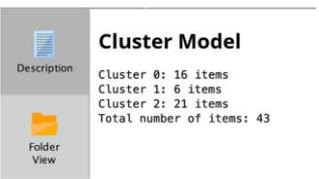
### 3.2. Klusterisasi Data dengan *K-Means*

Berdasarkan data pada Tabel 3, akan dilakukan klusterisasi dengan menggunakan algoritme *K-Means* untuk mengetahui kluster sebaran Covid-19 di Kota Denpasar. Proses klusterisasi dilakukan dengan menggunakan aplikasi *Rapid Miner*. Jumlah kluster yang digunakan dalam penelitian ini yaitu  $K=3$ .



**Gambar 4.** Proses Klasterisasi dengan menggunakan algoritme *K-Means*

Gambar 4 merupakan proses klasterisasi dengan menggunakan algoritme *K-Means* pada *Rapid Miner*. Data pada Tabel 3 akan disimpan dalam format *Microsoft excel* sebagai input klaster dan dilakukan pemilihan atribut sesuai dengan variabel penelitian. Setelah dilakukan pemilihan atribut, data akan diproses menggunakan algoritme *K-Means* dan untuk menghitung jarak terdekat dengan pusat klaster bisa menggunakan metode *euclidean distance*. Hasil akhir dari proses klasterisasi data dengan menggunakan algoritme *K mean* ditunjukkan pada Gambar 5.



**Gambar 5** Hasil proses klasterisasi dengan menggunakan algoritme *K-Means*

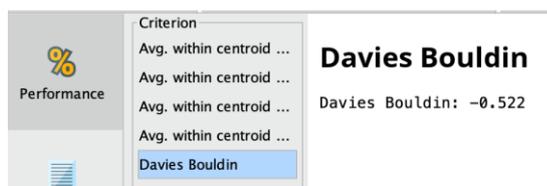
Data diklaster menjadi tiga klaster, dengan klaster 0 (*Low Cluster*) sebanyak 16 desa/kelurahan, klaster 2 (*Regular Cluster*) sebanyak 21 desa/kelurahan, dan klaster 1 (*High Clutser*) sebanyak 6 desa/kelurahan. Gambar 6 merupakan data detail dari anggota masing-masing klaster.

- cluster\_0
  - Dauh Puri Kauh
  - Dauh Puri Kangin
  - Sumerta Kauh
  - Kesiman Petilan
  - Peguyangan Kaja
  - Sanur Kaja
  - Dangin Puri
  - Dauh Puri Kaja
  - Penatih Dangin Puri
  - Ubung
  - Sumerta
  - Dangin Puri Kelod
  - Sumerta Kaja
  - Dangin Puri Kauh
  - Sanur
  - Serangan
- cluster\_1
  - Pemecutan Kelod
  - Pemogan
  - Sesetan
  - Ubung Kaja
  - Panjer
  - Padangsambian
- cluster\_2
  - Padangsambian Kelod
  - Dauh Puri
  - Dauh Puri Kelod
  - Pedungan
  - Sidakarya
  - Tonja
  - Kesiman Kertalangu
  - Penatih
  - Peguyangan
  - Sumerta Kelod
  - Padangsambian Kaja
  - Peguyangan Kangin
  - Renon
  - Sanur Kauh
  - Kesiman
  - Pemecutan
  - Pemecutan Kaja
  - Tegal Harum
  - Tegal Kertha
  - Dangin Puri Kangin
  - Dangin Puri Kaja

**Gambar 6.** Anggota dari masing-masing klaster

### 3.3. Validitas Kluster

Proses validitas kluster dilakukan dengan menggunakan metode validitas *Davies Bouldin*. Indeks validitas *Davies-Bouldin* diperoleh dengan menghitung rata-rata dari setiap titik pada himpunan data. Nilai setiap titik diperoleh dengan menghitung jumlah nilai *compactness* yang dibagi dengan jarak antara kedua titik pusat kluster sebagai *separation* [10]. Hasil validitas kluster dengan menggunakan metode *Davies Bouldin* yaitu sebesar -0.522.



Gambar 7. Validitas kluster dengan *Davies Bouldin*

## 4. SIMPULAN

Klusterisasi sebaran kasus Covid-19 di Kota Denpasar berhasil dilakukan menggunakan algoritme *K-Means* dan jumlah kluster yang digunakan yaitu 3, dengan kluster 0 (*Low Cluster*) sebanyak 16 desa/kelurahan, kluster 2 (*Regular Cluster*) sebanyak 21 desa/kelurahan, dan kluster 1 (*High Clutser*) sebanyak 6 desa/kelurahan. Validitas kluster diuji menggunakan metode *Davies Bouldin* dengan nilai yang dihasilkan yaitu -0.522.

## DAFTAR PUSTAKA

- [1] R. A. Indraputra and R. Fitriana, "K-Means Clustering Data COVID-19," vol. 10, no. 3, pp. 275–282, 2020.
- [2] P. L. O. Permadhi and I. M. Sudirga, "Problematika Penerapan Pembatasan Kegiatan Masyarakat (Pkm) Kota Denpasar Berbasis Adat Dalam Upaya Penanganan Covid-19," *Pros. Webinar Nas. ...*, 2020, [Online]. Available: <http://e-journal.unmas.ac.id/index.php/webinaradat/article/view/1169>.
- [3] T. Khotimah *et al.*, "Clustering The Development Of The Covid-19 Case In Indonesia," vol. 1, no. 1, pp. 23–26, 2020.
- [4] S. Sindi, W. R. O. Ningse, I. A. Sihombing, F. Ilmi R.H.Zer, and D. Hartama, "Analisis Algoritma K-Medoids Clustering Dalam Pengelompokan Penyebaran Covid-19 Di Indonesia," *Jti (Jurnal Teknol. Informasi)*, vol. 4, no. 1, pp. 166–173, 2020.
- [5] 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.
- [6] W. Yustanti, N. Rahmawati, and Y. Yamasari, "Klastering Wilayah Kota/Kabupaten Berdasarkan Data Persebaran Covid-19 di Propinsi Jawa Timur dengan Metode K-Means," *Journal Information Engineering*



- and Educational Technology*, vol. 04, no. 01. pp. 1–9, 2020.
- [7] 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.
- [8] I. B. G. Sarasvananda, R. Wardoyo, and A. K. Sari, "The K-Means Clustering Algorithm With Semantic Similarity To Estimate The Cost of Hospitalization," *IJCCS (Indonesian J. Comput. Cybern. Syst.)*, vol. 13, no. 4, p. 313, 2019, doi: 10.22146/ijccs.45093.
- [9] J. Hutagalung, N. L. W. S. R. Ginantra, G. W. Bhawika, W. G. S. Parwita, A. Wanto, and P. D. Panjaitan, "COVID-19 Cases and Deaths in Southeast Asia Clustering using K-Means Algorithm," *J. Phys. Conf. Ser.*, vol. 1783, no. 1, p. 012027, 2021, doi: 10.1088/1742-6596/1783/1/012027.
- [10] A. F. Khairati, A. A. Adlina, G. F. Hertono, and B. D. Handari, "Kajian Indeks Validitas pada Algoritma K-Means Enhanced dan K-Means MMCA," *Pros. Semin. Nas. Mat.*, vol. 2, pp. 161–170, 2019.