



Clustering Daerah Penyebaran Covid-19 di Indonesia Menggunakan Algoritma K-Medoids

Husdi ^{1*}, Muis Nanja ²

^{1,2} Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Ichsan Gorontalo.

article info

Article history:

Received 25 February 2022

Received in revised form

23 March 2022

Accepted 27 April 2022

Available online October 2022

DOI:

<https://doi.org/10.35870/jtik.v6i4.608>

Keywords:

Clustering, Covid-19, K-Medoids.

Kata Kunci:

Clustering, Covid-19, K-Medoids.

abstract

The year 2019 was the beginning of the outbreak of a disease that hit the world today called the Covid-19 Virus. The virus originated from Wuhan and is now designated as a pandemic. The main problem raised in this study is how to determine the spread of covid-19 in Indonesia based on provincial data. Because there are several areas that have high prevalence cases, it is necessary to improve health services and protocols in these areas compared to areas with rather low distribution areas. So that in this study a clustering method is needed to be able to classify data on the spread of COVID-19 in Indonesia. The algorithm used for clustering is the K-Medoids algorithm. Based on the results of the research conducted, the K-Medoids Method is able to group the Covid 19 Spreading Areas in Indonesia.

abstrak

Tahun 2019 adalah Awal Mulanya muncul Wabah Penyakit yang Melanda Dunia Saat Ini yang disebut dengan Virus Covid-19. Virus Tersebut bermula dari Wuhan dan kini ditetapkan sebagai pandemi. Permasalahan Utama yang di angkat dalam penelitian ini adalah bagaimana mengetahui penyebaran covid-19 di Indonesia berdasarkan data provinsi. Karena terdapat beberapa daerah yang mempunyai kasus penyebaran yang tinggi sehingga dibutuhkan peningkatan pelayanan dan protokol kesehatan pada daerah tersebut dibandingkan dengan wilayah yang daerah penyebaran agak rendah. Sehingga dalam dalam penelitian ini dibutuhkan metode clustering untuk dapat mengelompokkan data penyebaran covid-19 di Indonesia. Adapun Algoritma yang digunakan untuk clustering adalah algoritma K-Medoids. Berdasarkan hasil penelitian yang dilakukan, maka Metode K-Medoids ini dapat Mengelompokkan Daerah Penyebaran Covid 19 Di Indonesia.

* Corresponding author. Email: mr.husdi@unisan.ac.id ^{1*}.

1. Latar Belakang

Tahun 2019 adalah Awal Mulanya muncul Wabah Penyakit yang Melanda Dunia Saat Ini yang disebut dengan Virus *Covid-19*. Virus Tersebut bermula dari Wuhan dan kini ditetapkan sebagai pandemi. Terlepas dari dampak mewabahnya virus *Covid-19* yang menghebohkan dunia sekarang, ini bukan pertama kalinya dunia dilanda wabah mematikan. Selama dua dekade terakhir, sayangnya asia seringkali menjadi episentrumnya [1]. Pada kasus wabah *Covid-19* sekarang ini, berbagai upaya telah dilakukan untuk penanggulangan wabah penyakit tersebut di Indonesia dengan penerapan PSBB (Pembatasan Sosial Skala Besar-Besaran) antara lain berupa meliburkan sekolah untuk sementara waktu, menutup fasilitas umum untuk sementara waktu, melakukan pengamatan secara *intensif/surveilans* selama terjadi KLB serta melakukan evaluasi terhadap upaya penanggulangan secara keseluruhan. dilakukan secara terpadu oleh Pemerintah, pemerintah daerah dan masyarakat [2].

Tak terkecuali di Provinsi Gorontalo sangat merasakan dampak dari *covid-19* dari berbagai kalangan masyarakat. Di Gorontalo sendiri PSBB dimulai pada tanggal 4 mei dan di perpanjang sebanyak 3 kali. Kebijakan Pemerintah yang terbaru dengan meminta masyarakat untuk "berdamai" dengan *Covid-19* dengan menggaungkan apa yang disebut *New Normal* atau Pola Hidup Baru tentunya tidak terlepas dari upaya yang telah dilakukan oleh pemerintah selama ini dalam menangani penyebaran *Covid-19* [3]. Permasalahan Utama yang di angkat dalam penelitian ini adalah bagaimana mengetahui penyebaran *Covid-19* di Indonesia berdasarkan data provinsi. Karena terdapat beberapa daerah yang mempunyai kasus penyebaran yang tinggi sehingga dibutuhkan peningkatan pelayanan dan protokol kesehatan pada daerah tersebut dibandingkan dengan wilayah yang daerah penyebaran agak rendah. Sehingga dalam penelitian ini dibutuhkan metode *clustering* untuk dapat mengelompokkan data penyebaran *Covid-19* di Indonesia.

Tujuan dari Penelitian ini adalah menghasilkan data pengelompokan wilayah penyebaran *Covid-19* di Indonesia berdasarkan provinsi. Sehingga data tersebut dapat dijadikan acuan dalam peningkatan

protokol kesehatan dan memupus mata rantai penyebaran *Covid-19*. Beberapa Penelitian tentang clustering wilayah penyebaran penyakit di Indonesia telah banyak dilakukan dengan menggunakan metode/ algoritma clustering yang ada seperti pada penelitian berikut:

1. *K-Means Clustering Data Covid-19* [4],
2. Clustering Penyakit DBD Pada Rumah Sakit Dharma Kerti Menggunakan Algoritma *K-Means*[5],
3. Pengelompokan kasus penyakit *AIDS* Berdasarkan provinsi dengan *Data Mining K-Medoids Clustering* [6],
4. Penerapan Metode *K-Medoids Clustering* Pada Penanganan Kasus Diare *Di* Indonesia dan lain-lain. [7].

Adapun kebaruan dalam penelitian ini adalah mencluster daerah penyebaran covid-19 di Indonesia berdasarkan provinsi dengan dataset yang digunakan bersumber dari dataset *Covid 19* di Indonesia pada laman <http://sinta.ristekbrin.go.id/>[8]

a) *Clustering*

Clustering adalah salah satu bagian dari Konsep Data Mining. Yakni bertujuan untuk menggali informasi atau menemukan pola-pola yang bermanfaat dari sekumpulan data. Selain *Clustering* dalam data mining Cuma terdapat. Klasifikasi, Prediksi, Asosiasi, Regresi dan lain-lain [9,10]. *Clustering* merupakan proses partisi satu set objek data ke dalam himpunan bagian yang disebut dengan cluster. Objek yang di dalam *cluster* memiliki kemiripan karakteristik antar satu sama lainnya dan berbeda dengan cluster yang lain. Partisi tidak dilakukan secara manual melainkan dengan suatu algoritma *clustering* [11].

Metode *clustering* terdiri dari dua jenis, yaitu *Partitioning clustering* dan *Hierarchical clustering*. *Partitioning clustering* merupakan metode pengelompokan data yang dimulai dengan menentukan terlebih dahulu jumlah *cluster* yang diinginkan (dua *cluster*, tiga *cluster*, atau lain sebagainya). Setelah jumlah *cluster* diketahui, kemudian dilakukan proses cluster tanpa mengikuti proses *hierarki Hierarchical* [12].

a) K-Medoids

Menurut Listiyanti, Youanda Alim Syahbana & Henim [13] "Algoritma K-Medoids adalah algoritma pengelompokan yang berkaitan dengan algoritma K-Means dan algoritma *medoidshif*". Algoritma K-Medoids dikembangkan oleh Leonard Kaufman dan Peter J. Rousseeuw pada tahun 1987. Algoritma K-Medoids sering disebut juga algoritma *Partitioning Around Medoids (PAM)* [14]. Algoritma *PAM (Partitioning Around Medoid)* adalah sebuah algoritma yang merepresentasikan *cluster* yang dibentuk menggunakan *medoids cluster* di bangun dengan menghitung kedekatan yang dimiliki antara *medoid* dengan objek *non-medoid* menggunakan *distance measure* [6].

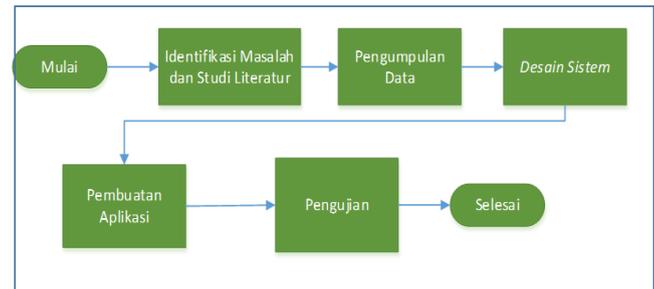
K-Medoids dan bisa disebut algoritma *PAM* yang memiliki kemiripan dengan K-Means karena keduanya tujuannya membagi data menjadi kelompok/*cluster*. Algoritma K-Means dan algoritma K-Medoids memiliki perbedaan pada penentuan pusat cluster, yang mana algoritma K-Means ditentukan dari nilai rata-rata (means) pada setiap cluster, sedangkan K-Medoids menggunakan objek data sebagai perwakilan (*medoids*) [15]. Berikut langkah - langkah yang digunakan dalam perhitungan algoritma K-Medoids [15]:

1. Menentukan pusat *cluster* sebanyak k
2. Mengalokasikan data untuk *cluster* terdekat dengan rumus ukuran jarak *euclidean distance*
3. Memilih acak objek pada setiap cluster yang digunakan untuk medoid baru
4. Menghitung jarak masing – masing objek berada pada setiap cluster.
5. kandidat *medoids* baru dengan menggunakan rumus *euclidian distance*
6. Menghitung total simpangan (S) dengan menghitung nilai total *distance* baru sampai total *distance* lama. Jika $S < 0$, maka tukar objek dengan hasil data clustering untuk membentuk sekumpulan k objek baru untuk medoids.
7. Ulangi urutan 3 sampai 5 hingga tidak ada perubahan medoids, sampai mendapatkan cluster beserta masing-masing anggota cluster.

2. Metode Penelitian

Tahapan Penelitian :

Metode pelaksanaan dalam penelitian ini terdapat beberapa tahapan untuk mencapai solusi dari permasalahan yang ada. Berikut ini adalah tahapan-tahapan dalam pelaksanaan penelitian ini:



Gambar 1. Tahapan Penelitian

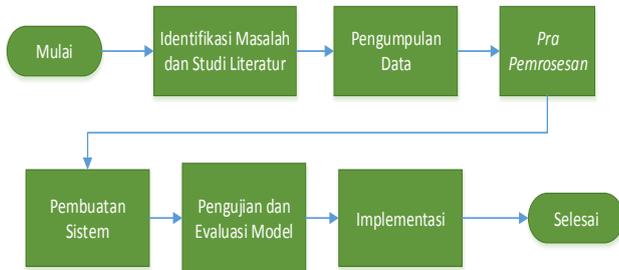
Seluruh dataset yang digunakan dalam penelitian ini di ambil dari laman <https://sinta.ristekbrin.go.id> pada menu *Covid-19*[8]. Yang mana menu tersebut terdapat dataset berisi data nasional *Covid-19* secara real time yang diperuntukkan untuk peneliti yang akan melakukan riset tentang *Covid-19*. Adapun model data yang disajikan dan menjadi data dalam penelitian ini adalah sebagai berikut:

Tabel 1. Sampel Data

No	Provinsi Asal	Kasus	Sembuh	Kematian
1	Aceh	4.554	2.325	176
2	Bali	8.878	7.365	275
3	Banten	5.643	3.964	171
4	Bangka Belitung	361	303	4
5	Bengkulu	693	441	34
6	DI Yogyakarta	2.643	1.885	67
7	DKI Jakarta	73.736	60.231	1.729
8	Jambi	513	265	11
9	Jawa Barat	22.205	13.650	420
10	Jawa Tengah	22.435	16.034	1.414
11	Jawa Timur	43.744	37.058	3.192
12	Kalimantan Barat	982	810	9
13	Kalimantan Timur	8.651	5.772	332
14	Kalimantan Tengah	3.645	2.816	133
15	Kalimantan Selatan	10.348	8.784	421
16	Kalimantan Utara	577	495	4
17	Kepulauan Riau	2.234	1.418	56
18	Nusa Tenggara Barat	3.316	2.638	196
19	Sumatera Selatan	6.095	4.357	348
20	Sumatera Barat	6.281	3.185	127
21	Sulawesi Utara	4.487	3.596	176

22	Sumatera Utara	10.313	6.974	429
23	Sulawesi Tenggara	2.862	1.934	60
24	Sulawesi Selatan	15.579	12.340	420
25	Sulawesi Tengah	410	267	15
26	Lampung	894	639	33
27	Riau	7.622	3.849	159
28	Maluku Utara	2.065	1.802	74
29	Maluku	2.815	1.800	41
30	Papua Barat	2.135	1.153	35
31	Papua	6.321	3.835	86
32	Sulawesi Barat	794	425	11
33	Nusa Tenggara Timur	424	274	6
34	Gorontalo	2.753	2.263	76
35	Dalam proses investigasi	0		
Total		287.008	214.947	10.740

Berikut merupakan diagram alir dari tahapan penelitian.



Gambar 2. Tahapan penelitian

a) *Tahap Pengumpulan Data :*

Dalam penelitian ini, data yang digunakan adalah data berita hoax, baik itu dari sosial media ataupun dari situs web yang khusus menyajikan berita untuk khalayak umum.

Tahap Pembuatan Aplikasi

Merupakan tahapan di mana kita melakukan pengembangan, melakukan tahap pembangunan sesuai dari hasil analisa dan desain sistem yang sebelumnya, termasuk didalamnya membangun sebuah aplikasi, menulis code program dan membuatnya dalam bentuk sebuah antarmuka dan integrasi dari input, proses dan, output yang meruakan bagian-bagian dari system program. Adapun untuk pembuatan website menggunakan PHP dan MYSQL.

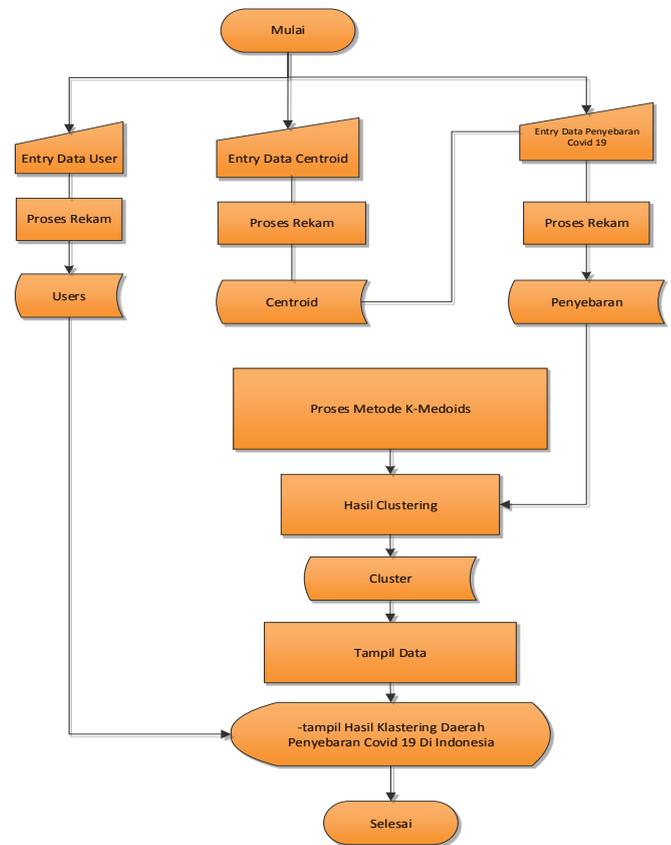
b) *Tahap Pengujian:*

Tahap pengujian di lakukan setelah semua model selesai di buat, dan program dapat berjalan, di mana seluruh perangkat lunak, program tambahan, dan

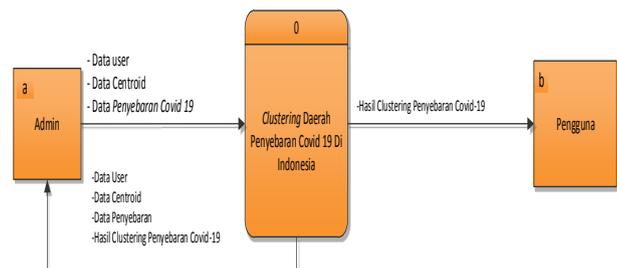
semua program yang terlibat dalam pembangunan sistem diuji untuk memastikan apakah sistem yang dirancang dapat berjalan dengan baik maka di lakukan pengujian whitebox untuk menguji prosesur-prosedur pada proses Selanjutnya pengujian *blackbox* untuk menguji *event-event* pada aplikasi yang di bangun.

3. Hasil dan Pembahasan

Sistem Usulan



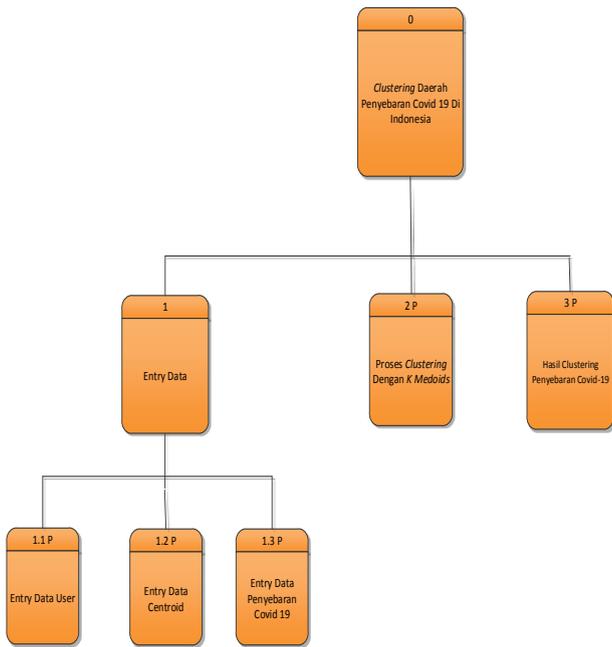
Gambar 3. Sistem Usulan



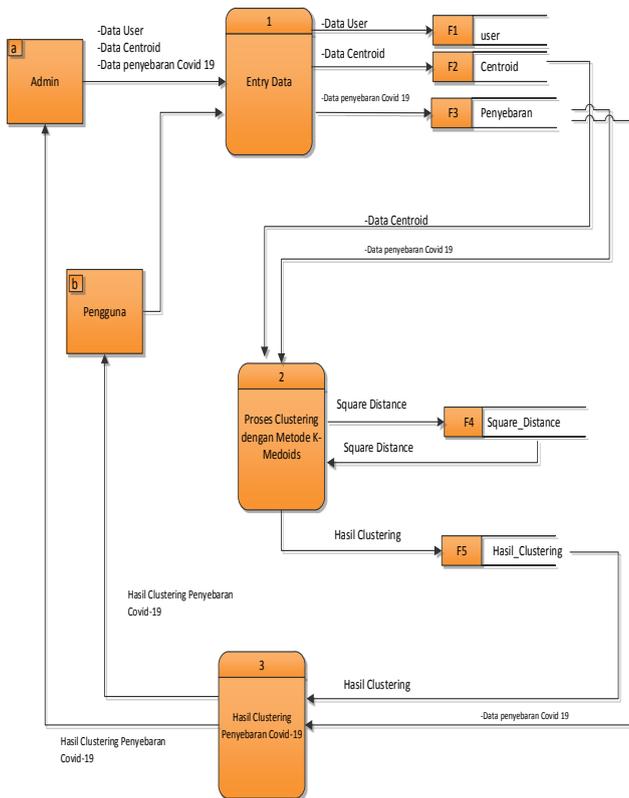
Gambar 4. Diagram Konteks

Diagram Berjenjang yaitu suatu diagram yang digunakan untuk menggambarkan proses-proses

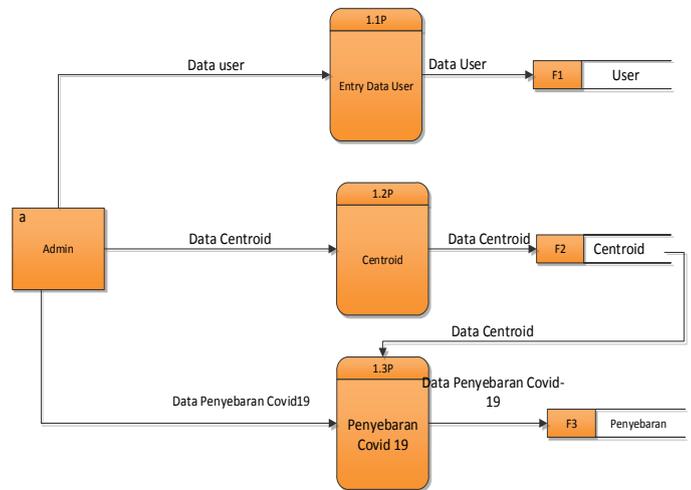
yang terdapat dalam sistem, yakni menggambarkan input, proses, output yang dibutuhkan dalam sistem.



Gambar 5. Diagram Berjenjang



Gambar 6. DAD Level 0



Gambar 7. DAD Level 1 Proses 1

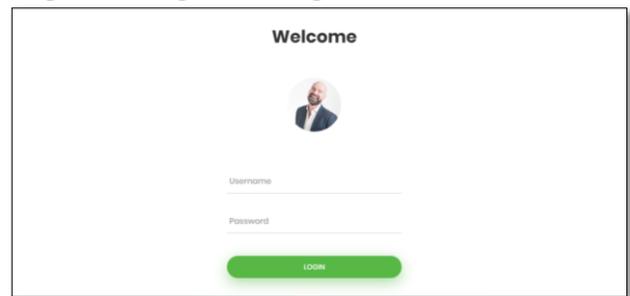
Hasil Perancangan Sistem

Berikut ini merupakan tampilan dari aplikasi yang sudah dirancang dengan menggunakan bahasa pemrograman *php*



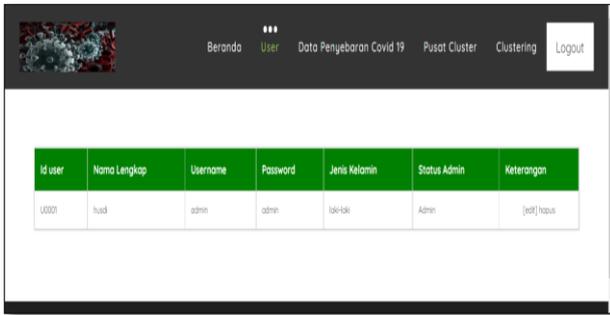
Gambar 8. Halaman Awal Aplikasi

Halaman Awal Aplikasi pada gambar 8. merupakan tampilan awal pada saat aplikasi ini akan di buka

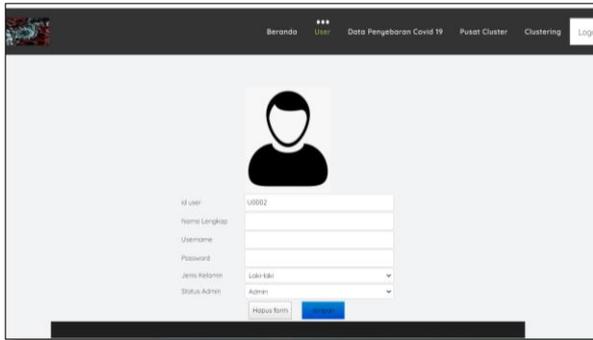


Gambar 9. Tampilan Menu Login

Halaman ini berfungsi untuk menginput username dan password untuk masuk ke halaman aplikasi Clustering Penyebaran Covid 19.



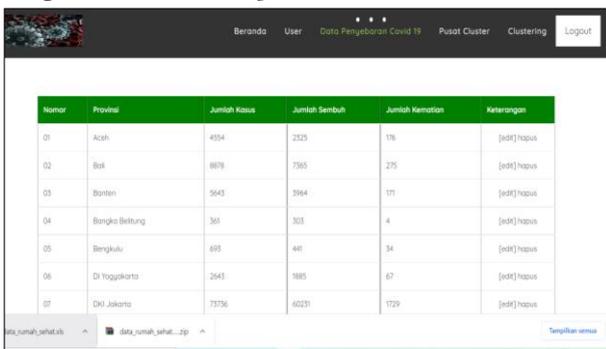
Gambar 10. Tampilan Menu User



Gambar 11. Tampilan Halaman Tambah User

Halaman ini menampilkan username dan password yang tersimpan dan tersedia menu input user baru untuk buat user baru.

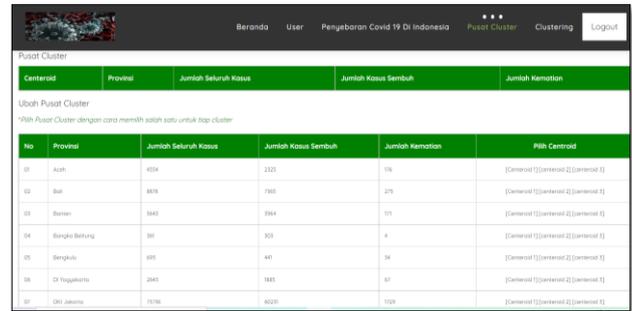
Tampilan Halaman Penyebaran Covid 19



Gambar 12. Tampilan Halaman Tabel Penyebaran

Halaman ini menampilkan tabel data penyebaran Covid 19 DI Indonesia.

Tampilan Halaman centroid



Gambar 13. Tampilan Halaman centroid

Centroid adalah Pusat Cluster yang akan digunakan untuk menghitung jarak data dalam proses clustering.

Tampilan Halaman Tabel Hasil Cluster

Halaman ini menampilkan hasil cluster Pertama Penyebaran Covid 19 di Indonesia



Gambar 14. Tampilan Hasil Cluster 1

Halaman ini menampilkan hasil cluster Kedua Penyebaran Covid 19 di Indonesia



Gambar 15. Tampilan Hasil Cluster 2

Halaman ini menampilkan hasil cluster Ketiga Penyebaran Covid 19 di Indonesia.

- [7] Hardiyanti, F., Tambunan, H.S. and Saragih, I.S., 2019. Penerapan Metode K-Medoids Clustering pada Penanganan Kasus Diare di Indonesia. *KOMIK (Konferensi Nasional Teknologi Informasi dan Komputer)*, 3(1). doi: 10.30865/komik.v3i1.1666.
- [8] "Dataset Covid 19 Di Indonesia, 2021. [Online]. Available: <http://sinta.ristekbrin.go.id/covid/>.
- [9] Husdi, H., 2020. Sistem Cerdas dalam Mengidentifikasi Kematangan Buah Naga Berdasarkan Fitur Tekstur dengan Metode K-Nearest Neighbor. *ILKOM Jurnal Ilmiah*, 12(3), pp.225-232. doi: 10.33096/ilkom.v12i3.665.225-232.
- [10] H. Dalai. 2017. Klasifikasi Blogger Menggunakan Algoritma K-Nearest Neighbor (k-NN) Berbasis Backward Elimination, *Multek*.
- [11] Zuhail, N.K., 2022, February. Study Comparison K-Means Clustering Dengan Algoritma Hierarchical Clustering: AHC, K-Means Clustering, Study Comparison. In *STAINS (SEMINAR NASIONAL TEKNOLOGI & SAINS)* (Vol. 1, No. 1, pp. 200-205).
- [12] Putri, H.N. and Saputro, D.R.S., 2022, February. Clustering Data Campuran Numerik dan Kategorik Menggunakan Algoritme Ensemble Quick ROBust Clustering using linKs (QROCK). In *PRISMA, Prosiding Seminar Nasional Matematika* (Vol. 5, pp. 716-720).
- [13] Syahbana, Y.A., 2017, January. Perancangan dan Implementasi Aplikasi Android Penentu Salient Area pada Video dengan Algoritma K-Medoids. In *Annual Research Seminar (ARS)* (Vol. 2, No. 1, pp. 96-101). [Online]. Available: <http://ars.ilkom.unsri.ac.id>.
- [14] Ashari, A., 2020. *Analisa K-Medoids Pada Tingkat Pengangguran Berdasarkan Tingkat Pendidikan Di Indonesia* (Doctoral dissertation, STMIK ROYAL KISARAN). [Online]. Available: http://repository.stmikroyal.ac.id/182/%0Ahttp://repository.stmikroyal.ac.id/182/1/SKR_IPSI.pdf.
- [15] Wira, B., Budianto, A.E. and Wiguna, A.S., 2019. Implementasi Metode K-Medoids Clustering Untuk Mengetahui Pola Pemilihan Program Studi Mahasiswa Baru Tahun 2018 Di Universitas Kanjuruhan Malang. *Rainstek: Jurnal Terapan Sains dan Teknologi*, 1(3), pp.53-68. doi: 10.21067/jtst.v1i3.3046.