

## Analisis Pengaruh Kardiovaskular Dalam Kasus Covid-19 Terhadap Obesitas Menggunakan Metode K-Medoid

Hadi Ramdan<sup>1</sup>, A.Gunawan<sup>2</sup>, Gunawan<sup>3</sup>

Universitas Bina Sarana Informatika

haramvolly@gmail.com<sup>1</sup>, a.gunawan.agn@bsi.ac.id<sup>2</sup>, gunawan.gnz@bsi.ac.id<sup>3</sup>

Diterima (30-09-2023)	Direvisi (18-04-2024)	Disetujui (25-04-2024)
--------------------------	--------------------------	---------------------------

**Abstrak** - Penelitian ini bertujuan untuk menganalisis pasien COVID-19 yang menderita obesitas, non-obesitas dan menderita obesitas sekaligus kardiovaskular melalui melihat perbandingan sekaligus pengaruh dari antar atribut yang ada dalam dataset serta mengetahui jumlah persentase dari pasien yang terjangkit COVID-19 dan tidak terjangkit COVID-19, dengan menggunakan metode K-Medoid. Hasil dari penelitian ini adalah penelitian dapat menggunakan K-medoid sebagai metode klustering untuk menemukan data, dalam dataset mentah menunjukkan persentase tidak terjangkit covid lebih banyak dibandingkan dengan pasien terjangkit COVID-19 dengan 62,62% untuk pasien yang tidak terjangkit covid-19 dan terjangkit covid-19 sebanyak 37,38%. Dari populasi dan sampel yang di uji sampel non-obesitas memiliki persentase sebesar 74,54%, obesitas memiliki persentase sebesar 25,46%, Sampel pasien yang memiliki obesitas sekaligus kardiovaskular hanya memiliki persentase 0,57%. Hasil pemodelan K-medoid mendapatkan tingkat persentase pengaruh pada setiap klaster, pada dataset obesitas memiliki perbedaan atribut pneumonia dengan tingkat pengaruh sebesar 150,15% dan hipertension sebesar 172,04%, sedangkan untuk dataset non-obesitas memiliki perbedaan atribut sex dengan tingkat pengaruh sebesar 39,50% dan hipertension sebesar 106,61%. Serta penderita obesitas sekaligus menderita kardiovaskular terdapat perbedaan atribut sex dan pneumonia dengan tingkat pengaruh sebesar 159,07% dan pneumonia sebesar 300%.

Kata Kunci : Covid-19, K-medoid, Obesitas, Kardiovaskular

**Abstract** - This research aims to analyze COVID-19 patients with obesity, non-obesity, and those with both obesity and cardiovascular conditions by examining the comparison and influence of attributes within the dataset. It also aims to determine the percentage of patients affected by COVID-19 and those not affected, using the K-Medoid method. The results of this study indicate that K-Medoid can be utilized as a clustering method to discover data patterns. The raw dataset shows a higher percentage of patients not affected by COVID-19 compared to those affected, with 62.62% not affected and 37.38% affected by COVID-19. From the tested population and sample, non-obese patients account for 74.54%, obese patients account for 25.46%, while patients with both obesity and cardiovascular conditions only constitute 0.57%. The K-Medoid modeling results reveal the percentage of influence in each cluster. In the obesity dataset, there are differences in the attributes of pneumonia with an influence rate of 150.15%, and hypertension with an influence rate of 172.04%. Meanwhile, in the non-obesity dataset, there are differences in the attribute of sex with an influence rate of 39.50%, and hypertension with an influence rate of 106.61%. Additionally, among obese patients with cardiovascular conditions, there are differences in the attributes of sex with an influence rate of 159.07% and pneumonia with an influence rate of 300%.

Keywords: Covid-19, K-medoid, Obesity, Cardiovascular

### I. PENDAHULUAN

COVID-19 adalah penyakit pernapasan yang disebabkan oleh virus SARS-CoV-2 yang menyebar dengan cepat ke seluruh dunia. Pandemi ini telah menyebabkan kelelahan tenaga kesehatan, angka kematian yang tinggi, dan dampak sosial dan ekonomi yang luas (Nursofwa et al., 2020)

COVID-19 dapat menyebabkan gejala ringan hingga berat, termasuk pneumonia, sesak napas, dan kegagalan organ. Komplikasi

jangka panjang juga dapat terjadi (Hasanah et al., 2020)

Obesitas dan COVID-19 telah menjadi dua masalah kesehatan global yang saling berhubungan. Menurut banyak ahli, obesitas dapat meningkatkan risiko tertular COVID-19 dan mengalami komplikasi penyakit yang lebih serius (Rahayu et al., 2021).

Obesitas dan COVID-19 saling terkait. Obesitas meningkatkan risiko tertular COVID-19 dan mengalami komplikasi yang lebih serius. Obesitas dapat melemahkan sistem kekebalan

tubuh, menyebabkan peradangan kronis, dan menurunkan fungsi paru-paru. Obesitas juga dapat meningkatkan risiko penyakit kardiovaskular, diabetes tipe 2, dan masalah pernapasan. Obesitas dapat memengaruhi respons terhadap vaksin COVID-19. Menjaga berat badan yang sehat penting untuk melindungi dari COVID-19 (Febrianti et al., 2020).

Ahli berpendapat bahwa pendekatan holistik diperlukan untuk mengatasi obesitas dan COVID-19. Selain vaksinasi dan pengendalian penyebaran virus, perubahan gaya hidup sehat, dukungan masyarakat, pendekatan berbasis bukti, dan akses setara ke layanan kesehatan juga penting (S.Kwok et al., 2020)

Obesitas dan COVID-19 saling terkait dan menjadi perhatian serius. Pendekatan komprehensif diperlukan untuk mengatasi tantangan ini, termasuk langkah-langkah pengendalian virus, vaksinasi, dan manajemen obesitas. Menjaga berat badan yang sehat dapat membantu melindungi diri dari COVID-19 dan meningkatkan kesehatan secara keseluruhan.

Namun dengan berbagai masalah yang muncul, diperlukannya sebuah Analisa untuk menjadikan sebuah acuan dalam penyelesaian masalah. Salah satu caranya menggunakan proses data mining.

*Clustering* adalah metode data mining yang mengelompokkan data ke dalam kelompok berdasarkan kesamaan fitur. Salah satu algoritma *clustering* yang umum digunakan adalah K-medoid, yang menggunakan titik data aktual sebagai pusat *cluster*. Algoritma K-medoids bekerja dengan mengidentifikasi medoid pada setiap *cluster* dan memperbaharui medoid pada setiap iterasi hingga tercapai konvergensi (Dinata et al., 2021)

Penelitian menggunakan metode data mining dengan menggunakan algoritma K-medoid penting dilakukan karena dapat memberikan informasi yang berharga bagi pengambilan keputusan. Algoritma K-medoid merupakan salah satu teknik clustering yang digunakan untuk mengelompokkan data menjadi beberapa kelompok. Keunggulan algoritma ini adalah tidak sensitif terhadap outlier, sehingga dapat menghasilkan pengelompokan data yang lebih akurat.

### 1. Tujuan Penelitian

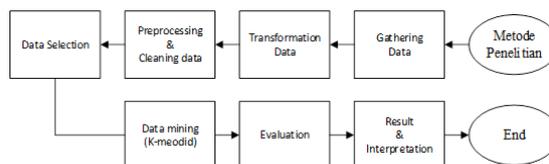
Tujuan dalam penelitian ini adalah sebagai berikut:

- Melakukan Analisa data set COVID-19 serta penerapan algoritma K-medoids.
- Mengetahui hasil perbandingan antara orang yang terkena COVID-19 antara orang yang mempunyai penyakit obesitas dan

tidak punya penyakit obesitas serta tingkat persentase pengaruhnya.

- Mengetahui hasil perbandingan kepada pasien COVID-19 yang menderita obesitas dan kardiovaskular serta tingkat persentase pengaruhnya.
- Mengetahui persentase yang terjangkit covid-19 dan tidak terjangkit covid-19 dari dataset mentah serta populasi dan sampel yang akan di uji.
- Mengetahui nilai evaluasi dari DBI dan AVG. *Within centroid distance* dari hasil penelitian.

### 2. Gambaran Metode Usulan



Sumber: Olahan Peneliti (2023)

Gambar 1. Gambaran Metode usulan

- Gathering data* adalah tahapan untuk mengumpulkan data dari berbagai sumber, baik internal maupun eksternal. Data yang dikumpulkan dapat berupa data teks, data numerik, atau data gabungan.
- Data cleaning* adalah tahapan untuk membersihkan data dari kesalahan, duplikasi, atau data yang tidak relevan. Data yang bersih akan memastikan hasil data mining yang akurat dan valid.
- Data transformation* adalah tahapan untuk mengubah format data agar sesuai dengan kebutuhan proses data mining. Data yang ditransformasi akan memudahkan proses ekstraksi pola dari data.
- Data selection* adalah tahapan untuk memilih data yang relevan dengan tujuan proses data mining. Data yang dipilih akan digunakan untuk proses data mining selanjutnya.
- Data mining* adalah tahapan untuk mengekstrak pola dari data. Pola yang diekstrak dapat berupa pola hubungan antar data, pola tren data, atau pola lainnya menggunakan algoritma k-medoid.
- Evaluation* adalah tahapan untuk mengevaluasi hasil data mining. Hasil data mining dievaluasi untuk menilai akurasi, relevansi, dan interpretabilitasnya.
- Result & interpretation* adalah tahapan untuk menyajikan hasil data mining dalam bentuk yang mudah dipahami. Hasil data mining dapat disajikan dalam bentuk tabel, grafik, atau laporan.

## II. METODOLOGI PENELITIAN

Penelitian ini menggunakan algoritma k-medoid untuk menganalisis hubungan antara obesitas, kardiovaskular, dan covid-19. Penelitian ini menggunakan metode kuantitatif dan akan dievaluasi berdasarkan hasil pengujian.

Penelitian kuantitatif menggunakan angka dan data numerik untuk mengumpulkan dan menganalisis informasi. Penelitian ini biasanya dilakukan melalui survei, eksperimen, atau sumber data lainnya (I Made Laut Mertha Jaya, 2020)

Penelitian kuantitatif menggunakan teknik statistik untuk menguji hipotesis, mengidentifikasi pola, dan membuat generalisasi. Data dikumpulkan dengan kuesioner atau instrumen yang memberikan angka (Mukhid, 2021)

Mengutip dari buku Metode Penelitian Kuantitatif dan Kualitatif (2020): Ciri-ciri penelitian kuantitatif adalah sebagai berikut:

1. Model pemikiran: Penelitian kuantitatif menggunakan model pemikiran deduktif, yaitu dimulai dari teori umum untuk menjelaskan fenomena spesifik.
2. Sasaran: Penelitian kuantitatif bertujuan untuk mengembangkan ilmu nomotetik, yaitu ilmu yang bertujuan membuat hukum dari generalisasinya.
3. Pengumpulan data: Penelitian kuantitatif mengumpulkan data melalui pengukuran dengan menggunakan alat yang obyektif dan terstandar.
4. Analisis data: Penelitian kuantitatif menggunakan analisis statistik untuk menggeneralisasi hasil penelitian.
5. Peneliti: Peneliti kuantitatif harus objektif dan tidak terlibat secara emosional dengan objek penelitian.

Data yang diambil dan diolah adalah data yang bersumber dari internet. Khususnya dari forum kaggle, dataset yang digunakan adalah daataset sekunder. Data sekunder adalah sumber yang tidak langsung memberikan data kepada pengumpul data kepada pengumpulan data. Data sekunder didapatkan dari sumber yang dapat mendukung penelitian antara lain dari dokumentasi dan literatur (Novi Yona Sidratul Munti & Dwi Asril Syaifuddin, 2020).

Dataset yang akan digunakan dalam penelitian ini adalah data yang dimiliki oleh Meir Nizri sebagai kolabolator sekaligus *owner* dari dataset itu sendiri. Dataset tersebut memiliki beberapa atribut di antaranya, *sex*, *age*, *classification*, *patient type*, *pneumonia*, *pregnancy*, *diabetes*, *copd*, *asthma*, *inmsupr*, *hypertension*, *kardiovaskular*, *renal chronic*, *other deseas*, *obesity*, *tobacco*, *usmr*, *medical unit*, *intubed*, *icu*, *date died*. Dengan definisi

atribut dijelaskan pada tabel berikut.

Tabel 1. Definisi Atribut

Atribut	Definisi
<i>Sex</i>	1 untuk perempuan dan 2 untuk laki-laki.
<i>Age</i>	Umur dari pasien
<i>Classification</i>	Temuan tes covid. Nilai 1-3 berarti pasien didiagnosis dengan covid berbeda derajat. 4 atau lebih tinggi berarti pasien bukan pembawa covid atau tesnya tidak meyakinkan.
<i>Patient Type</i>	jenis perawatan yang diterima pasien di unit. 1 untuk kembali ke rumah dan 2 untuk rawat inap.
<i>Pneumonia</i>	apakah pasien sudah mengalami radang kantung udara atau belum.
<i>Pregnancy</i>	apakah pasien hamil atau tidak.
<i>Diabetes</i>	apakah pasien menderita diabetes atau tidak.
<i>Copd</i>	Menandakan apakah pasien mengidap penyakit paru obstruktif kronik atau tidak.
<i>Asthma</i>	apakah pasien menderita asma atau tidak.
<i>Inmsupr</i>	apakah pasien immunosupresi atau tidak.
<i>Hypertension</i>	apakah pasien menderita hipertensi atau tidak.
<i>Cardiovascular</i>	apakah pasien memiliki penyakit terkait jantung atau pembuluh darah.
<i>Renal chronic</i>	apakah pasien memiliki penyakit ginjal kronis atau tidak.
<i>Other disease</i>	apakah pasien memiliki penyakit lain atau tidak.
<i>Obesity</i>	apakah pasien obesitas atau tidak.
<i>Tobacco</i>	apakah pasien adalah pengguna tembakau.
<i>Usmr</i>	Menunjukkan apakah pasien dirawat unit medis tingkat pertama, kedua atau ketiga.
<i>Medical unit</i>	jenis lembaga Sistem Kesehatan Nasional yang menyediakan perawatan.
<i>Intubed</i>	apakah pasien terhubung ke ventilator.
<i>Icu</i>	Menunjukkan apakah pasien telah dirawat di Unit Perawatan Intensif.
<i>Date died</i>	Jika pasien meninggal diindikasikan dengan tanggal kematiannya, untuk tanggal 9999-99-99 adalah kebalikan nya.

Sumber: kaggle.com (2023)

### 1. K-medoid

K-Medoid adalah algoritma pengelompokan data yang digunakan untuk mengelompokkan data menjadi beberapa kelompok atau *cluster*. Algoritma ini merupakan varian dari algoritma K-Means tetapi menggunakan medoid sebagai pusat *cluster* sedangkan K-Means menggunakan mean (rata-rata) sebagai pusat

*cluster*. Medoid adalah titik data aktual yang mewakili sebuah *cluster* (Supriyadi et al., 2021) Dalam setiap *cluster*, medoid adalah titik data dengan jarak total terkecil ke semua titik data lain dalam *cluster*. Sebuah medoid bukanlah titik data yang berada di pusat geometris sebuah *cluster*, seperti di K-Means. langkah-langkah K-Medoid (Dewi & Pramita, 2019) :

- a. Inisialisasi: Pilih K titik data awal sebagai medoid acak dari dataset.
- b. *Assignment*: Setiap titik data dalam dataset diberi label sesuai dengan medoid terdekat. Jarak antara titik data dan medoid bisa dihitung dengan menggunakan metrik jarak seperti jarak Euclidean atau Manhattan.
- c. *Update*: Untuk setiap medoid dalam setiap klaster, hitung total jarak dari medoid ke semua titik data lain dalam klaster. Pilih titik data baru sebagai medoid yang memiliki total jarak terkecil.
- d. Iterasi: Ulangi langkah 2 dan 3 hingga tidak ada perubahan dalam penugasan medoid.
- e. Output: Setelah konvergensi, hasil akhir adalah partisi klaster dengan medoid dan titik data yang terkait.

Langkah-langkah ini mengoptimalkan fungsi objektif yang disebut total jarak dalam klaster. Tujuan akhir dari algoritma K-Medoid adalah untuk meminimalkan total jarak dalam klaster dan memaksimalkan total jarak antara klaster. Penting untuk dicatat bahwa K-Medoid adalah algoritma iteratif yang memerlukan inisialisasi medoid awal dan berulang kali menghitung jarak dan memperbarui medoid hingga konvergensi (Sindi et al., 2020).

**2. Clustering**

*Clustering*, juga dikenal sebagai analisis *cluster*, adalah proses pengelompokan objek atau data ke dalam kelompok atau *cluster* yang memiliki kesamaan hal tertentu. Pengelompokan bertujuan untuk mengelompokkan objek-objek yang mirip menjadi satu dan memisahkan objek-objek yang berbeda satu sama lain (Anjelita et al., 2020).

Dalam konteks analisis data, *clustering* digunakan untuk mengidentifikasi pola atau struktur dalam data yang sebelumnya tidak diketahui. Menggunakan algoritma pengelompokan, data dapat diatur ke dalam kelompok dengan kesamaan internal yang tinggi dan perbedaan eksternal yang tinggi. Pengelompokan dapat membantu Anda memahami data, mendapatkan wawasan baru, dan membuat keputusan yang lebih baik (Dogan & Birant, 2021)

**3. Rapidminer**

RapidMiner adalah aplikasi data mining yang powerful dan serbaguna, cocok untuk berbagai pengguna. Dapat digunakan untuk berbagai jenis data dan aplikasi, mulai dari bisnis hingga ilmiah. User interface intuitif, komunitas aktif, dan sangat direkomendasikan. Rapidminer yang akan digunakan adalah 9.10.

**III. HASIL DAN PEMBAHASAN**

**1. Pemilihan Data Sampel dan Populasi**

Data yang akan diolah berasal dari kaggle, berjumlah 1.048.576 data. Data perlu dibersihkan dan ditransformasi sebelum diolah dengan algoritma K-medoids. Data mentah perlu ditransformasi dan dibersihkan sebelum diolah. Transformasi meliputi eksklusi data. Atribut yang digunakan seperti *medical\_unit*, *sex*, *patient\_type*, *intubed*, *pneumonia*, *age*, *diabetes*, *copd*, *asthma*, *hipertensi*, *cardiovaskular*, *obesity*, *renal\_chronic*, *tobacco*, *classification\_final*, *icu*. Dengan atribut yang sudah disiapkan, peneliti mencoba mengerucutkan kembali dataset dengan filter dan menyaring data sesuai dengan usia (*age*) dan klasifikasi final (*classification\_final*).

Sumber: Dataset (2023)

Gambar 2. Dataset Covid-19

memilih data pada rentang usia 12-45 yang memiliki kriteria dari masa remaja awal hingga masa dewasa akhir, rentang usia dijelaskan dalam jurnal “Urgensi Revisi Undang-undang tentang Kesejahteraan Lanjut Usia” (Hakim, 2020) bahwa ada 9 kriteria umur:

- a. Masa Balita (usia 0-5 tahun)
- b. Masa Kanak-kanak (usia 5-11 tahun)
- c. Masa Remaja Awal (usia 12-16 tahun)
- d. Masa Remaja Akhir (usia 17-25 tahun)
- e. Masa Dewasa Awal (usia 26-35 tahun)
- f. Masa Dewasa Akhir (usia 36-45 tahun)
- g. Masa Lansia Awal (usia 46-55 tahun)
- h. Masa Lansia Akhir (usia 56-65 tahun)
- i. Masa Manula (usia 65 tahun ke atas)

Data difilter berdasarkan nilai atribut *classification\_final*. Nilai 1-3 berarti pasien terkena COVID-19. Setelah filter, jumlah data menjadi 25.970 (sample data) dari 1.048.576 (populasi data). Dataset kembali dipisah menjadi 3 bagian seperti:

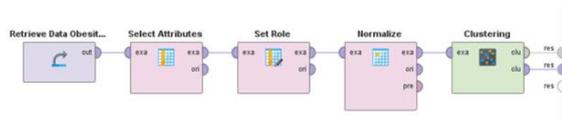
- a. Dataset 1: Pasien COVID-19 dengan obesitas: 6.615 data
- b. Dataset 2: Pasien COVID-19 tanpa obesitas: 19.355 data
- c. Dataset 3: Pasien COVID-19 dengan obesitas dan kardiovaskular: 147 data

**2. Model dan Pengujian**

Pengujian klustering K-medoids menggunakan rapidminer dengan operator retrieve data, select attributes, set role, normalize, dan clustering k-medoids. Operator *retrieve* data memiliki fungsi untuk mengambil dataset yang akan diolah dari *repository* kedalam proses modeling, dalam pengujian akan mencoba memasukan dataset yang non-obesitas, obesitas, serta dataset obesitas dan kardiovaskular.

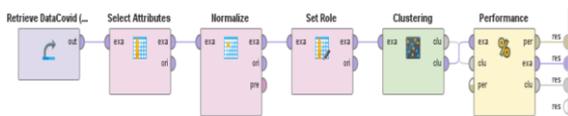
*Select attributes* adalah operator untuk memilih atribut yang akan dimasukkan atau yang tidak akan dimasukkan pada proses nanti, dalam pengujian ini akan tidak memasukkan atribut *age* yang dinilai tidak relevan setelah masuk pada tahapan proses.

*Set role* merupakan atribut yang menjelaskan bagaimana operator lain menangani Atribut yang akan dipilih pada pengaturannya. Pada pengujian ini akan menetapkan atribut *classification\_final* sebagai Id. untuk contoh penggunaan operator seperti berikut:



Sumber: Rapidminer (2023)  
Gambar 3. Operator rapidminer

Pada pengujian berikutnya akan menggunakan operator yang sama seperti sebelumnya namun ditambahkan operator performance. Operator ini digunakan untuk evaluasi kinerja metode klustering berbasis centroid, dalam pengujian akan menggunakan K-medoids seperti berikut:



Sumber: Rapidminer (2023)  
Gambar 4. Evaluasi

**3. Hasil Akhir**

Peneliti menunjukkan berbagai bentuk visualisasi dan hasil serta hasil dari proses perhitungan yang dapat diperoleh untuk menyelesaikan penelitian dengan mencapai tujuan akhir. seperti *centroid table*, *cluster model*, *example data clustering*, dan *plot chart*.

- a. Obesitas
- 1) *Centroid table* obesitas

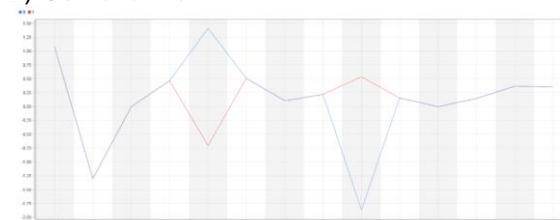
Tabel 2. *Centroid Table* obesitas

Attribute	cluster_0	cluster_1
Medical_Unit	1,072	1,072
Sex	-1,305	-1,305
Patient_Type	0	0
Intubed	0,469	0,469
Pneumonia	1,413	-0,708
Diabetes	0,508	0,508
Copd	0,106	0,106
Asthma	0,215	0,215
Hipertension	-1,872	0,534
Cardiovascular	0,151	0,151
Obesity	0	0
Renal_Chronic	0,149	0,149
Tobacco	0,368	0,368
Icu	0,352	0,352

Sumber: Rapidminer (2023)

Klaster 1 memiliki nilai hipertensi yang lebih rendah dan nilai pneumonia yang lebih tinggi daripada klaster 2. Hal ini menunjukkan bahwa pneumonia lebih memengaruhi klaster 1, sedangkan hipertensi lebih memengaruhi klaster 2.

- 2) *Centroid Plot*



Sumber: Rapidminer (2023)

Gambar 5. *Centroid Plot* obesitas

Centroid plot menunjukkan bahwa klaster 1 memiliki nilai hipertensi yang lebih rendah dan nilai pneumonia yang lebih tinggi daripada klaster 2.

- 3) *Cluster model*

Tabel 3. *Cluster model* obesitas

Cluster Model	
Cluster 0	1468
Cluster 1	5147
Total Numbes of items	6615

Sumber: Rapidminer (2023)

Algoritma K-medoid membagi data pasien COVID-19 dengan obesitas menjadi 2 klaster, yaitu klaster 1 dengan 1.468 data dan klaster 2 dengan 5.147 data.

- b. Non-obesitas
- 1) *Centroid table* Non-obesitas

Tabel 3. *Centroid Table* non-obesitas

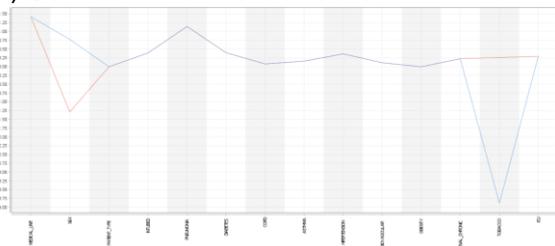
Attribute	cluster_0	cluster_1
Medical_Unit	1,421	1,421
Sex	0,776	-1,288
Patient_Type	0	0
Intubed	0,383	0,383

<i>Pneumonia</i>	1,141	1,141
<i>Diabetes</i>	0,395	0,395
<i>Copd</i>	0,076	0,076
<i>Asthma</i>	0,151	0,151
<i>Hipertension</i>	0,367	0,367
<i>Cardiovascular</i>	0,108	0,108
<i>Obesity</i>	0	0
<i>Renal_Chronic</i>	0,291	0,291
<i>Tobacco</i>	-3,886	0,257
<i>Icu</i>	0,292	0,292

Sumber: Rapidminer (2023)

Klaster 1 memiliki nilai *sex* yang lebih tinggi dan nilai *tobacco* yang lebih rendah daripada klaster 2. Hal ini menunjukkan bahwa *sex* lebih memengaruhi klaster 1, sedangkan *tobacco* lebih memengaruhi klaster 2.

2) *Centroid Plot*



Sumber: Rapidminer (2023)

Gambar 6. *Centroid Plot* non-obesitas

*Centroid plot* menunjukkan bahwa klaster 1 memiliki nilai *sex* yang lebih tinggi dan nilai *tobacco* yang lebih rendah daripada klaster 2.

3) *Cluster model*

Tabel 3. *Cluster model* non-obesitas

<i>Cluster Model</i>	
<i>Cluster 0</i>	1202
<i>Cluster 1</i>	18153
Total Numbes of items	19355

Sumber: Rapidminer (2023)

Algoritma K-medoid membagi data pasien COVID-19 tanpa obesitas menjadi 2 klaster, yaitu klaster 1 dengan 1.202 data dan klaster 2 dengan 18.153 data. Jumlah data yang diproses adalah 19.355 data.

c. Obesitas dan kardiovaskular

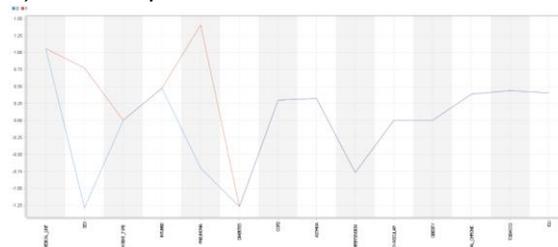
1) *Centroid table* obesitas dan kardiovaskular

Attribute	<i>cluster_0</i>	<i>cluster_1</i>
<i>Medical_Unit</i>	1,060	1,060
<i>Sex</i>	-1,289	0,771
<i>Patient_Type</i>	0	0
<i>Intubed</i>	0,473	0,473
<i>Pneumonia</i>	-0,705	1,409
<i>Diabetes</i>	-1,270	-1,270
<i>Copd</i>	0,297	0,297
<i>Asthma</i>	0,323	0,323
<i>Hipertension</i>	-0,771	-0,771
<i>Cardiovascular</i>	0	0
<i>Obesity</i>	0	0
<i>Renal_Chronic</i>	0,384	0,384
<i>Tobacco</i>	0,440	0,440
<i>Icu</i>	0,407	0,407

Sumber: Rapidminer (2023)

Klaster 1 memiliki nilai *sex* yang lebih rendah dan nilai *pneumonia* yang lebih tinggi daripada klaster 2. Hal ini menunjukkan bahwa *sex* lebih memengaruhi klaster 2, sedangkan *pneumonia* lebih memengaruhi klaster 1.

2) *Centroid plot*



Sumber: Rapidminer (2023)

Gambar 7. *Centroid Plot* obesitas & kardiovaskular

*Centroid plot* menunjukkan bahwa klaster 1 memiliki nilai *sex* yang lebih rendah dan nilai *pneumonia* yang lebih tinggi daripada klaster 2.

3) *Cluster model*

Tabel 4. *Cluster model* non-obesitas

<i>Cluster Model</i>	
<i>Cluster 0</i>	58
<i>Cluster 1</i>	49
Total Numbes of items	147

Sumber: Rapidminer (2023)

Algoritma K-medoid membagi data pasien COVID-19 dengan obesitas dan kardiovaskular menjadi 2 klaster, yaitu klaster 1 dengan 58 data dan klaster 2 dengan 49 data. Jumlah data yang diproses adalah 147 data.

d. *Performance cluster*

Dataset utama dipilih untuk mengevaluasi hasil clustering dengan nilai rata-rata jarak centroid dan DBI (*Davies Bouldin Indeks*) dari proses yang telah dilakukan dalam aplikasi rapidminer, hasilnya seperti berikut:

Tabel 5. *Cluster model* non-obesitas

<i>PerformanceVector</i>	
AVG. Within centroid distance	18,064
AVG. Within centroid distance_ <i>cluster_0</i>	26,506
AVG. Within centroid distance_ <i>cluster_1</i>	17,715
Davies Bouldin	1,367

Sumber: Rapidminer (2023)

Pada tabel *cluster performance* terdapat beberapa hasil yang didapatkan, ada *AVG. Within centroid distance* yang memiliki *value* 18,064 yang merupakan nilai rata-rata jarak titik centroid secara keseluruhan, untuk perhitungan pada klaster 1 memiliki *value* 26,506 dan pada klaster 2 memiliki *value* 17,715, untuk hasil yang terakhir yaitu evaluasi dari *davies bouldin index* memiliki hasil 1,367.

Hasil analisa yang didapat pada dataset ini akan menampilkan beberapa tabel yang berkaitan dengan pasien covid-19 yang terjangkit dan tidak terjangkit, dari dataset asli, yang sudah difilter terhadap pasien penderita obesitas, pasien non-obesitas, serta pasien yang menderita obesitas sekaligus kardiovaskular dalam bentuk persentasenya, berikut tabelnya:

Tabel 6. Covid dan Non-Covid dari dataset

Atribut	Kalkulasi	Jumlah	Persentase
Raw data	total	1048575	100%
	covid-19	391979	37,38%
	non-covid-19	656596	62,62%
Obesitas	total	162848	100%
	covid-19	74127	45,62%
	non-covid-19	88721	54,48%
Non obesitas	total	888759	100%
	covid-19	319205	35,91%
	non-covid-19	569554	64,09%
Obesitas & kardiovaskular	total	8932	100%
	Covid-19	4030	45,11%
	non-covid-19	4902	54,89%

Sumber: Olahan Peneliti (2023)

Dari hasil analisis, pasien yang tidak terjangkit COVID-19 lebih banyak daripada pasien yang terjangkit, baik untuk pasien obesitas maupun non-obesitas. Tabel berikutnya memperlihatkan hasil dari persentase dalam jumlah populasi dan sampel dalam penelitian ini, berikut tabelnya:

Tabel 7. Sampel dan populasi (persentase)

Nama	Jumlah Data	Persentase
Populasi	25970	100%
Sampel (Obesitas)	6615	25,46%
Sampel (Non-Obesitas)	19355	74,54%
Sampel (Obesitas dan Kardiovaskular)	147	0,57%

Sumber: Olahan Peneliti (2023)

Sampel non-obesitas lebih banyak daripada sampel obesitas dan sampel obesitas sekaligus kardiovaskular. Tabel terakhir menunjukkan tingkat pengaruh setiap atribut pada klaster pasien COVID-19 dengan obesitas, non-obesitas, dan obesitas kardiovaskular:

Tabel 8. Pengaruh antar atribut

Nama Dataset	Atribut	Persentase
obesitas	<i>pneumonia</i>	150,15%
	<i>hipertension</i>	172,04%

non obesitas	<i>sex</i>	39,50%
	<i>tobacco</i>	106,61%
obesitas & kardiovaskular	<i>sex</i>	159,07%
	<i>pneumonia</i>	300%

Sumber: Olahan Peneliti (2023)

Persentase yang didapatkan dari hasil perhitungan k-medoid memperlihatkan persentase yang bervariasi dan dipengaruhi dengan 2 atribut, pada sampel obesitas dipengaruhi oleh pneumonia dengan tingkat pengaruh sebesar 150,15% dan hipertension sebesar 172,04%, pada sampel non-obesitas dipengaruhi oleh sex dengan tingkat pengaruh sebesar 39,50% dan hipertension sebesar 106,61% serta sampel terakhir sampel obesitas dan kardiovaskular dipengaruhi oleh sex dengan tingkat pengaruh sebesar 159,07% dan pneumonia sebesar 300%.

#### IV. KESIMPULAN

1. Penelitian dengan menggunakan algoritma k-medoid dapat diterapkan pada penelitian ini dengan menggunakan aplikasi rapidminer untuk penerapan dan menganalisa hasil.
2. Dataset yang menderita obesitas dan non-obesitas memiliki perbedaan, pada dataset obesitas memiliki perbedaan pada atribut pneumonia dengan tingkat pengaruh sebesar 150,15% dan hipertension sebesar 172,04%, sedangkan untuk dataset non-obesitas memiliki perbedaan pada atribut sex dengan tingkat pengaruh sebesar 39,50% dan hipertension sebesar 106,61%.
3. Dataset penderita obesitas sekaligus menderita kardiovaskular terdapat perbedaan pada atribut sex dan pneumonia dengan tingkat pengaruh sebesar 159,07% dan pneumonia sebesar 300%.
4. Dari dataset mentah mendapatkan hasil 62,62% untuk pasien yang tidak terjangkit covid-19 dan terjangkit covid-19 sebanyak 37,38%. , Obesitas yang tidak terjangkit covid-19 lebih banyak dengan persentase 54,48% dan terjangkit covid-19 sebanyak 45,62%, non-obesitas yang tidak terjangkit covid-19 lebih banyak dengan persentase 64,09% dan terjangkit covid-19 sebanyak 35,91%, dan obesitas sekaligus kardiovaskular yang tidak terjangkit covid-19 lebih banyak dengan persentase 54,89% dan terjangkit covid-19 sebanyak 45,11%.
5. Dari populasi dan sampel yang di uji sampel non-obesitas memiliki persentase sebesar 74,54% , obesitas memiliki persentase sebesar 25,46% , Untuk sampel pasien yang memiliki obesitas sekaligus

kardiovaskular hanya memiliki persentase 0,57%.

6. Hasil performa pada klaster pengujian menggunakan algoritma k-medoid mendapatkan AVG. Within centroid distance yang memiliki value 18,064, pada klaster 1 memiliki nilai 26,506 dan pada klaster 2 memiliki nilai 17,715, serta evaluasi DBI memiliki value 1,367.

Untuk saran dari peneliti, Dataset COVID-19 dapat diolah dengan berbagai metode, seperti K-means, klasifikasi, dan prediksi. Dalam penelitian ini, evaluasi hanya dilakukan dengan menghitung *performance* pada 2 kaster. Bisa dicoba untuk melakukan komparasi hasil dengan jumlah kaster 2-5 dan memperluas data. Namun, hal ini tidak dilakukan karena keterbatasan perangkat.

## V. REFERENSI

- Anjelita, M., Perdana Windarto, A., Wanto, A., & Sudahri, I. (2020). Pengembangan Datamining Klustering Pada Kasus Pencemaran Lingkungan Hidup. *Seminar Nasional Teknologi Komputer & Sains (SAINTEKS)*, 1(1), 309–313.
- Dewi, D. A. I. C., & Pramita, D. A. K. (2019). Analisis Perbandingan Metode Elbow dan Silhouette pada Algoritma Clustering K-Medoids dalam Pengelompokan Produksi Kerajinan Bali. *Matrix : Jurnal Manajemen Teknologi Dan Informatika*, 9(3), 102–109. <https://doi.org/10.31940/matrix.v9i3.1662>
- Dinata, R. K., Retno, S., & Hasdyna, N. (2021). Minimization of the Number of Iterations in K-Medoids Clustering with Purity Algorithm. *Revue d'Intelligence Artificielle*, 35(3), 193–199. <https://doi.org/10.18280/ria.350302>
- Dogan, A., & Birant, D. (2021). Machine learning and data mining in manufacturing. *Expert Systems with Applications*, 166, 114060. <https://doi.org/10.1016/j.eswa.2020.114060>
- Febrianti, C. D., Candi, C., & Sulistiadi, W. (2020). *Effect of Obesity on Covid-19 Patients*. 30(1chd), 315–320.
- Hakim, L. N. (2020). The Urgency of The Elderly Welfare Law Revision. *Aspirasi: Jurnal Masalah-Masalah Sosial*, 11(1), 43–55. <https://doi.org/10.22212/aspirasi.v11i1.1589>
- Hasanah, D. Y., Nauli, S. E., Prima Putri, V. K., Arifianto, H., Suryana, N. M., Suryani, L. D., Aditya, W., & Probodewi, P. (2020). Gangguan Kardiovaskular pada infeksi COVID 19. *Indonesian Journal of Cardiology*, 41(2), 59–68. <https://doi.org/10.30701/ijc.994>
- I Made Laut Mertha Jaya. (2020). *Metode Penelitian Kuantitatif dan Kualitatif: Teori, Penerapan, dan Riset Nyata* (p. 232). Fira Husaini. [https://books.google.co.id/books/about/Method\\_Penelitian\\_Kuantitatif\\_dan\\_Kualitatif?id=yz8KEAAAQBAJ&redir\\_esc=y](https://books.google.co.id/books/about/Method_Penelitian_Kuantitatif_dan_Kualitatif?id=yz8KEAAAQBAJ&redir_esc=y)
- Mukhid, A. (2021). Metodologi Penelitian Pendekatan Kuantitatif. In *Syria Studies* (Vol. 7, Issue 1). [https://www.researchgate.net/publication/269107473\\_What\\_is\\_governance/link/548173090cf22525dcb61443/download%0Ahttps://www.econ.upf.edu/~reynal/Civilwars\\_12December2010.pdf%0Ahttps://think-asia.org/handle/11540/8282%0Ahttps://www.jstor.org/stable/41857625](https://www.researchgate.net/publication/269107473_What_is_governance/link/548173090cf22525dcb61443/download%0Ahttps://www.econ.upf.edu/~reynal/Civilwars_12December2010.pdf%0Ahttps://think-asia.org/handle/11540/8282%0Ahttps://www.jstor.org/stable/41857625)
- Nisri, M. (2020). COVID-19 dataset. Kaggle. Retrieved from <https://www.kaggle.com/datasets/meirnizri/covid19-dataset?rvi=1>
- Novi Yona Sidratul Munti, & Dwi Asril Syaifuddin. (2020). Analisa Dampak Perkembangan Teknologi Informasi Dan Komunikasi Dalam Bidang Pendidikan. *Jurnal Pendidikan Tambusai*, 4(2), 1799–1805. <https://jptam.org/index.php/jptam/article/view/655>
- Nursofwa, R. F., Sukur, M. H., Kurniadi, B. K., & . H. (2020). Handling of Health Services During the Covid-19 Pandemic in the Perspective of Health Law. *Inicio Legis*, 1(1), 1–17. <https://doi.org/10.21107/il.v1i1.8822>
- Rahayu, L. A. D., Admianty, J. C., Khalda, Y. I., Ahda, F. R., Agistany, N. F. F., Setiawati, S., Shofiyanti, N. I., & Warnaini, C. (2021). Hipertensi, Diabetes Mellitus, Dan Obesitas Sebagai Faktor Komorbiditas Utama Terhadap Mortalitas Pasien Covid-19: Sebuah Studi Literatur. *JIMKI: Jurnal Ilmiah Mahasiswa Kedokteran Indonesia*, 9(1), 90–97. <https://doi.org/10.53366/jimki.v9i1.342>
- S.Kwok, S.Adam, J.H.Ho, Z.Iqbal, P.turkington, S.Razvi, C.W.Le roux, H.Soran, A. A. S. (n.d.). *Clinical Obesity - 2020 - Kwok - Obesity A critical risk factor in the COVID-19 pandemic.pdf*.
- Sindi, S., Ningse, W. R. O., Sihombing, I. A., R.H.Zer, F. I., & Hartama, D. (2020). Analisis Algoritma K-Medoids Clustering Dalam Pengelompokan Penyebaran Covid-19 Di Indonesia. *Jurnal Teknologi Informasi*, 4(1), 166–173. <https://doi.org/10.36294/jurti.v4i1.1296>

Supriyadi, A., Triayudi, A., & Sholihati, I. D. (2021). Perbandingan Algoritma K-Means Dengan K-Medoids Pada Pengelompokan Armada Kendaraan Truk Berdasarkan Produktivitas. *JIPi (Jurnal Ilmiah Penelitian Dan Pembelajaran Informatika)*,

6(2), 229–240.  
<https://doi.org/10.29100/jipi.v6i2.2008>