

Analisis Gejala Covid-19 di RSUD Yogyakarta dengan Menggunakan Metode K-Means, Fuzzy C-Means dan K-Medoid Clustering

Kathina Deswi Aqsa^{1*}, Uswatun Hasanah², Urai Amanda¹

¹Program Studi Farmasi, Sekolah Tinggi Ilmu Kesehatan (STIKes) Sambas, Kalimantan Barat 79462, Indonesia

²Program Studi Farmasi, Fakultas Sains, Institut Teknologi Sumatera, Sumatera Selatan 35365, Indonesia

*email: kety.deswiagqa16@gmail.com

ABSTRACT

At the end of 2019, precisely in December, there was a discovery of a new disease originating from a virus, namely coronavirus. The purpose of this study was to group patients based on covid-19 symptoms using clustering with the K-Means, Fuzzy C-Means and K-Medoid algorithms at the Regional Public Hospital in the Special Region of Yogyakarta and see the comparison of the best algorithms. The data used are 5 (five) medical record data of patients from Yogyakarta Regional Public Hospitals, including inpatients at Yogyakarta City Hospital, inpatients at Sleman Hospital, inpatients at Wates Hospital, inpatients at Panembahan Senopati Hospital and inpatients at Wonosari Hospital for the 2020-2021 period using the proportional random sampling technique. The results obtained in this study, medical record data from 1066 patients selected into 622 patients have an optimal cluster of 2 clusters, using the K-Means algorithm, cluster 1 is moderate severity and cluster 2 is severe severity. Cluster evaluation was carried out with the Davies Bouldin Index (DBI) value on the K-Means algorithm 3.48. Based on the results of the study from the smallest DBI value, it can be concluded that the cluster uses the optimal K-Means algorithm.

Keywords: Cluster, Covid-19, Tingkat Keparahan, K-Means, Fuzzy C-Means, K-Medoid

Received: May 2025; Accepted: June 2025; Published: June 2025



©2025. Published by Institute for Research and Innovation Universitas Muhammadiyah Banjarmasin. This is Open Access article under the CC-BY-SA License (<http://creativecommons.org/licenses/by-sa/4.0/>).

LATAR BELAKANG

Pada akhir tahun 2019, kota Wuhan di China menjadi fokus perhatian dunia karena wabah penyakit pernapasan yaitu *Coronavirus*. (1). *Coronavirus* ini memiliki gejala demam (>90%), malaise, batuk kering (80%), sesak napas (20%) serta adanya gangguan pernapasan (15%). Tanda vital lainnya stabil namun terjadi leukopenia dan limfopenia (1). Menurut pedoman tatalaksana covid-19 edisi 4 di Indonesia gejala covid-19 dibagi menjadi 5 yaitu tanpa gejala, ringan, sedang, berat dan kritis (2)

Beberapa penelitian telah dilakukan di Indonesia terkait penyebaran covid-19 menggunakan algoritma *K-Means* dengan provinsi di Indonesia di kelompokkan berdasarkan data kasus positif, meninggal dan sembuh variabel yang digunakan adalah kasus positif

dan kasus meninggal per tanggal 19 April 2020 menghasilkan 3 *cluster* provinsi di Indonesia (3). Pengelompokan covid-19 dengan algoritma K-means dengan memasukkan data dari umur, jenis kelamin, penyakit penyerta yang mengelompokkan menjadi ringan, sedang dan berat (4). Tidak hanya itu terdapat algoritma lainnya seperti *Fuzzy Clustering Methode* (FCM) dengan mengetahui tingkat risiko Covid-19 berdasarkan beberapa variabel penentu di Provinsi Daerah Istimewa Yogyakarta (DIY) diperoleh 2 cluster terbaik yaitu "Risiko Tinggi" dan "Risiko Rendah" (5). Pada algoritma K-Medoid mengelompokkan aplikasi pendukung saat pembelajaran masa pandemi menggunakan menghasilkan 2 cluster yaitu aplikasi yang disukai dan tidak disukai. Pada penelitian ini menyebutkan *K-Medoid* lebih baik daripada *K-Means* karena pada K-Medoid menemukan k sebagai

representasi objek untuk meminimalkan jumlah ketidaksamaan objek data (6).

Senada dengan penelitian sebelumnya, penelitian ini bertujuan untuk mengelompokkan gejala covid-19 menggunakan clustering dengan algoritma *K-Means*, *Fuzzy C-Means* dan *K-Medoid*. di RSUD Provinsi Yogyakarta dengan melihat perbandingan algoritma terbaik. Adapun parameter perbandingan algoritma dengan menggunakan metode *Davies Bouldin Index* (DBI). Jumlah *cluster* yang dipilih adalah jumlah cluster dengan nilai DBI terkecil (7)

METODE

Metode yang digunakan adalah observasional analitik dengan desain kohort. Pengambilan data secara retrospektif menggunakan data rekam medis pasien covid-19 di RSUD Provinsi Daerah Istimewa Yogyakarta (DIY) (RSUD Kota Yogyakarta, RSUD Sleman, RSUD Wates, RSUD Panembahan Senopati dan RSUD Wonosari). Teknik pengambilan sampel *propotional random sampling* sebanyak 1066 pasien diseleksi menjadi 622 pasien berdasarkan kriteria inklusi yaitu pasien terkonfirmasi covid-19 usia ≥ 18 tahun, pasien tercatat menjalani rawat inap di 5 RSUD Provinsi DIY, pasien memiliki rekam medis lengkap. Kriteria eksklusi yaitu pasien yang memiliki data rekam medis tidak terbaca dengan jelas, pasien hamil dan menyusui.

Analisis menggunakan algoritma *Fuzzy C-Means* (FCM), *K-Means* dan *K-Medoid* dengan menggunakan bahasa pemrograman *Python*. Selanjutnya dianalisis menggunakan parameter perbandingan cluster terbaik dengan melihat nilai *Davies Bouldin Index* (DBI). Adapun langkah untuk clusterisasi adalah sebagai berikut:

1. Pengumpulan Data

Data yang digunakan ialah data pasien covid-19 dengan variabel usia (x1), jenis kelamin (x2), gejala (x3), komorbid (x4), komplikasi (x5), nilai *RR* (x6), suhu (x7), SpO₂ (x8), leukosit (x9), limfosit (x10), D-dimer (x11), vitamin C (x12), vitamin D (x13), antibiotik (x14), antivirus (x15), steroid (x16), antikoagulan (x17), terapi simptomatik (x18).

2. Pre-Processing Data

Mengubah data kategorik menjadi data numeric. Selanjutnya dilakukan pembersihan data dilakukan yang bertujuan menghilangkan data yang tidak valid atau tidak lengkap. Normalisasi data dilakukan menggunakan Z-Score yang bertujuan untuk membuat nilai dari tiap variabel menjadi bentuk data rasional yang setara. Menentukan jumlah cluster yang optimal. Penentuan jumlah *cluster* (k) membandingkan nilai DBI setiap pembagian *cluster*. Melihat nilai k terkecil dari DBI.

3. Melakukan Clusterisasi

Menggunakan metode *K-Means* dan selanjutnya dilakukan *clustering* menggunakan pemrograman *Python*.

4. Analisis Hasil Cluster

Melihat karakteristik atau distribusi tiap variabel dari masing-masing cluster yang diperoleh.

5. Evaluasi Nilai *Davies Bouldin Index* (DBI)

Evaluasi nilai DBI dengan menggunakan persamaan berikut :

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

Keterangan:

k merupakan jumlah cluster yang digunakan, nilai *R_{i,j}* adalah rasio perbandingan antara cluster ke-i dan *cluster* ke-j.

Semakin kecil nilai DBI yang diperoleh (non negative ≥ 0), maka semakin optimal cluster tersebut. Adapun rumus untuk mencari nilai DBI adalah sebagai berikut (8) :

a. Mencari SSW dengan rumus :

$$SSW_i = \sum_{j=1}^{m_i} D(x_j, c_i)^2$$

Keterangan :

- m_i : jumlah data dalam cluster ke- i
- X : data dalam cluster
- x_j : data pada cluster tersebut
- $D(x, c)$: jarak data ke centroid
- x_j : data pada cluster tersebut
- c_i : centroid cluster ke - i

b. Mencari SSB dengan rumus :

$$SSB_{ij} = d(c_i, c_j)$$

Keterangan :

- c_i : cluster 1
- c_j : cluster lainnya
- $d(c_i, c_j) ::$ jarak antara centroid sat dengan lainnya

c. Mencari rasio dengan rumus :

$$R_{ij} = \frac{SSW_i + SSW_j}{SSB_{ij}}$$

Keterangan :

- R_{ij} : rasio antar cluster
- SSW_i : cluster 1
- SSW_j : cluster 2
- SSB_{ij} : Separasi dari cluster 1 dan 2

4. Mencari DBI dengan rumus :

$$DBI = \frac{1}{K} \sum_{i=1}^K \max_{i \neq j} (R_{i,j})$$

Keterangan :

- K : cluster yang ada
- R_{ij} : Rasio antara cluster i dan j
- SSW_i : Dicari rasio antar cluster yang terbesar

HASIL DAN PEMBAHASAN

Analisis Cluster

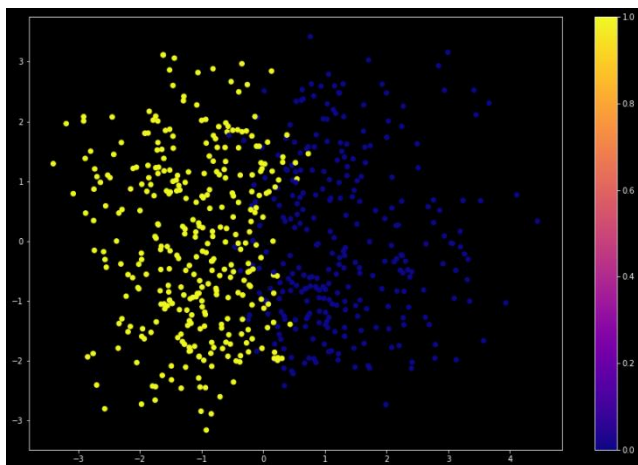
Berdasarkan algoritma *K-Means* dari 622 pasien dapat dikelompokkan 2 cluster berdasarkan nilai DBI terkecil (Tabel 1).

Tabel 1. Hasil Analisis Cluster Menggunakan Algoritma K-Means

Variabel	Jumlah	Cluster 1	Cluster 2	Hasil Uji Korelasi
Usia (tahun)				
• <50 tahun	220	147	77	0.27
• ≥ 50 tahun	402	143	255	
Jenis Kelamin				

• Laki-laki	320	132	188	-0.11
• Perempuan	302	158	144	
Diagnosa				
Komplikasi				0.20
• Tanpa Komplikasi	232	128	104	
• Ada Komplikasi 1	390	162	228	
Komorbid				
• Tanpa Komorbid	311	168	143	0.15
• Ada Komorbid	311	122	189	
Gejala				
• Tanpa Gejala	161	17	11	0.04
• 1-3 Gejala	339	165	192	
• 4-6 Gejala	122	108	129	
Tanda Vital				
• ΣRR				0.46
Normal	128	117	11	
Tidak Normal	494	173	321	
• Suhu				
Normal	452	249	203	0.28
Tidak Normal	170	41	129	
• SpO2				
Normal – Ringan	258	229	29	0.77
Sedang	120	53	67	
Berat	244	8	236	
• Leukosit				
Normal	473	250	223	0.22
Tidak Normal	149	40	109	
• Limfosit				
Normal	601	274	327	-0.11
Tidak Normal	21	16	5	
• D-dimer				
Normal	188	116	72	0.20
Tidak Normal	434	174	260	
• Vit C				
Ya	449	225	224	-0.11
Tidak	173	65	108	
• Vit D				
Ya	326	174	180	-0.14
Tidak	296	116	152	
• Antibiotik				
Ya	544	244	300	0.09
Tidak	78	46	32	
• Antivirus				
Ya	541	250	291	0.02
Tidak	81	40	41	
Steroid				
Ya	294	77	217	0.39
Tidak	328	213	115	
• Antikoagulan				
Ya	364	108	256	0.40
Tidak	258	182	76	
• Simptomatik				
Ya	537	227	310	0.22
Tidak	85	63	22	

Pada algoritma *K-Means* variabel yang lebih dominan yaitu gejala pada Gambar 1.



Gambar 1. Plot Hasil Sebaran Cluster Menggunakan Algoritma *K-Means*

Pada Gambar 2 menunjukkan plot sebaran data dari hasil cluster dengan menggunakan algoritma *K-Means*, terlihat pada gambar berwarna biru menunjukkan *cluster* 1 sedangkan warna kuning menunjukkan *cluster* 2. Sebaran plot gambar 4 menunjukkan bahwa *cluster* 1 dan *cluster* 2 terpisah dan pembentukan *cluster* jelas. Pada penelitian ini menggunakan evaluasi nilai *Davies Bouldin Index* (DBI). Dilihat dari nilai terkecil DBI pada algoritma *K-Means* yaitu 3,4839. Semakin kecil nilai DB Index mendekati nilai 0 menunjukkan cluster yang paling optimal (7). Selanjutnya Nilai *Silhouette Coefficient* berada pada rentan (-1) hingga 1. Semakin tinggi nilai *sillhoutte* maka akan berpengaruh ke hasil yang lebih baik. Pada penelitian ini menghasilkan urutan pertama nilai *sillhoutte coefficient* yang baik ialah *K-Means* sebesar 0.07. Pada algoritma Fuzzy C-Means menunjukkan 2 hasil cluster pada Tabel 2 di bawah ini.

Tabel 2. Hasil Analisis Cluster Menggunakan Algoritma Fuzzy C-Means

Variabel	Jumlah	Cluster 1	Cluster 2	Hasil Uji Korelasi
Usia (tahun)				0.33
• <50 tahun	220	158	62	
• ≥ 50 tahun	402	148	254	
Jenis Kelamin				-0.11
• Laki-laki	320	141	179	
• Perempuan	302	165	137	

Diagnosa				0.18
Komplikasi				
• Tanpa Komplikasi	232	132	100	
• Ada Komplikasi	380	174	222	
Komorbid				0.23
• Tanpa Komorbid	311	189	122	
• Ada Komorbid	311	117	194	
Gejala				-0.02
• Tanpa Gejala	161	16	12	
• 1-3 Gejala	339	167	190	
• 4-6 Gejala	122	123	114	
Tanda Vital				0.45
• RR				
Normal	128	119	9	
Tidak Normal	494	187	307	
• Suhu				0.25
Normal	452	257	195	
Tidak Normal	170	49	121	
• SpO2				0.70
Normal – Ringan	258	226	32	
Sedang	120	56	220	
Berat	244	24	64	
• Leukosit				0.18
Normal	473	257	216	
Tidak Normal	149	49	100	
• Limfosit				-0.08
Normal	601	291	310	
Tidak Normal	21	15	6	
• D-dimer				0.16
Normal	188	116	72	
Tidak Normal	434	190	244	
• Vitamin C				-0.19
Ya	449	247	202	
Tidak	173	59	114	
• Vitamin D				-0.18
Ya	326	189	179	
Tidak	296	117	137	
• Antibiotik				0.07
Ya	544	260	284	
Tidak	78	46	32	
• Antivirus				-0.03
Ya	541	269	272	
Tidak	81	37	44	
• Steroid				0.38
Ya	294	86	208	
Tidak	328	220	108	
• Antikoagulan				0.41
Ya	364	117	247	
Tidak	258	189	69	
• Simptomatik				0.17
Ya	537	246	291	
Tidak	85	60	25	

Pada tabel 2 dapat dijelaskan hal-hal sebagai berikut. *Cluster* 1 terdiri dari 306 pasien dan cluster 2 terdiri dari 316 pasien. Pada cluster 1 didominasi oleh

perempuan, dengan usia lebih dominan lansia yakni ≥ 50 tahun, tidak memiliki penyakit penyerta atau komorbid, gejala yang dialami dominan lebih sering 1-3 gejala serta nilai saturasi oksigen normal-ringan yaitu $\leq 95\%$. Pada cluster 1 untuk tanda vital seperti laju pernapasan mengalami takipnea/laju pernapasan tinggi dengan suhu normal (tidak demam). Pada uji hematologi seperti nilai leukosit cluster 1 normal, namun nilai limfosit menurun (limfositopenia), untuk uji D-dimer cenderung tinggi. Pada penggunaan terapi khusus cluster 1 pengobatan yang diberikan ialah vitamin C, vitamin D, antibiotik, antivirus, terapi simptomatik, tanpa steroid dan tanpa antikoagulan.

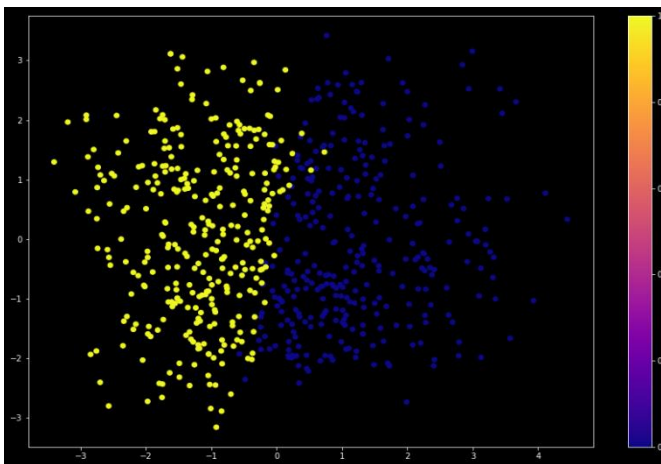
Pada cluster 2 total pasien sejumlah 316 orang dengan mayoritas berjenis kelamin laki-laki, usia paling dominan ≥ 50 tahun, memiliki komplikasi pneyakit, penyakit penyerta/komorbid, dengan ditemukannya 1-3 gejala Covid-19. Untuk tanda vital pada pasien cluster 2 ialah nilai *RR* lebih tinggi atau mengalami takipnea, saturasi oksigen sedang 93-95% namun tidak demam. Pada uji hematologi pasien seperti nilai leukosit pada cluster 2 normal tidak mnegalami kenaikan namun kadar limfosit menurun (limfositopenia) dan kadar D-dimer tinggi. Pada terapi farmakologi kelompok cluster 2 mengkonsumsi vitamin D, vitamin C, antibiotik, antivirus, steroid, antikoagulan serta terapi simptomatik. Berdasarkan hasil analisis XIX dapat dilihat bahwa pada cluster 1 lebih banyak pasien dengan gejala sedang.

Pada cluster 2 ialah kelompok pasien dengan gejala berat dilihat dari perbedaan saturasi oksigen yang lebih rendah, memiliki riwayat komorbid, usia diatas ≥ 50 tahun serta penggunaan terapi steroid serta antikoagulan yang mengindikasikan gejala Covid-19 berat sesuai dengan pedoman tatalaksana Covid-19 (2).

Berdasarkan hasil uji korelasi Pearson, variabel yang berkorelasi terhadap hasil cluster yaitu (1) saturasi oksigen (SpO_2) (0.70), (2) nilai *RR* (0.45), (3) terapi antikoagulan (0.41), (4) terapi steroid (0.38), (5) usia (0.33), (6) suhu (0.25), (7) komorbid (0.23), (8) komplikasi (0.18), (9) leukosit (0.18), (10) terapi simptomatik (0.17), (11) nilai D-dimer (0.16) dan (12) terapi antibiotik (0.07), (13) terapi vitamin C (-0,19), (14) terapi vitamin D (-0,18), (15) jenis kelamin (-0,11), (16) kadar limfosit (-0,08), (17) terapi antivirus (-0,03) dan (18) gejala (-0,02).

Nilai korelasi merupakan salah satu cara untuk mengukur tingkat korelasi dengan hasil cluster (cluster 1 dan cluster 2). Nilai korelasi berkisar dari -1 sampai 1, semakin negatif nilai korelasinya menunjukkan bahwa semakin besar variabel tersebut ditempatkan di cluster 1. Nilai korelasi positif menunjukkan bahwa semakin besar variabel tersebut ditempatkan pada cluster 2.

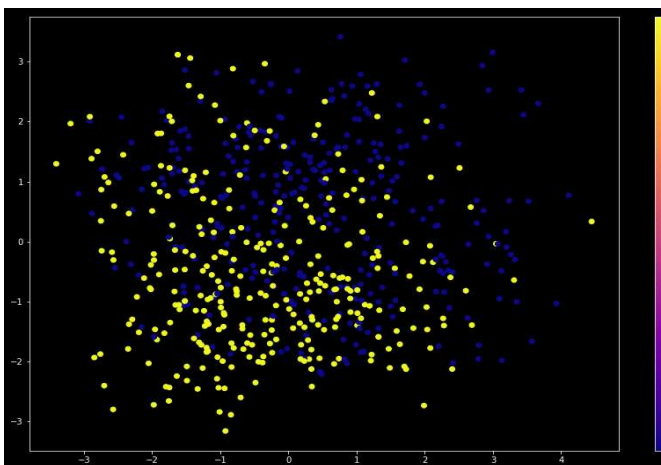
Pada penelitian ini juga mengukur kualitas cluster menggunakan *Davies Bouldin Index* (DBI) dimana evaluasi DBI ini bertujuan melihat kualitas cluster terbaik dengan cara melihat nilai DBI. Semakin kecil nilai DBI hingga mendekati 0 artinya pembagian jumlah cluster juga semakin baik. Pada algoritma *Fuzzy C-Means* didapat nilai terkecil DBI pada cluster 2 yaitu 3.54. Terlihat pada gambar warna biru menunjukkan klaster 1 sedangkan warna kuning menunjukkan klaster 2. Sebaran plot pasien covid-19 menggunakan algoritma *Fuzzy C-Means* dapat memisahkan klaster 1 dan klaster 2 dan pembentukan *cluster* jelas. Berdasarkan nilai evaluasi *Sillhouette Coefficient* dengan nilai tertinggi 0,07. Nilai koefisien siluet berkisar antara -1 hingga +1, di mana nilai yang lebih tinggi menunjukkan pengelompokan yang lebih baik (9).



Gambar 2. Plot Hasil Sebaran *Cluster* Menggunakan Algoritma *Fuzzy C-Means*

Pada Gambar 2 menunjukkan plot sebaran data hasil clusterisasi menggunakan algoritma *Fuzzy C-Means*. Terlihat pada gambar warna biru menunjukkan cluster 1 sedangkan warna kuning menunjukkan cluster 2. Sebaran plot pasien Covid-19 menggunakan algoritma *Fuzzy C-Means* dapat memisahkan cluster 1 dan cluster 2 dan pembentukan *cluster* jelas.

Pada penelitian ini juga menggunakan algoritma *K-Medoid*. Berikut gambar 3 plot hasil sebaran *cluster* menggunakan Algoritma *K-Medoid*.



Gambar 3. Plot Hasil Sebaran *Cluster* Menggunakan Algoritma *K-Medoid*

Berdasarkan hasil analisis cluster menggunakan *K-Medoid* diperoleh 2 cluster yang optimal dengan nilai *silhouette coefficient* sebesar 0.043. Pada algoritma *K-Medoid* didapat nilai terkecil DBI pada klaster 2 yaitu

4.604874 Terlihat pada gambar 3 tersebut warna biru menunjukkan cluster 1 dan warna kuning menunjukkan cluster 2. Namun pada sebaran plot menunjukkan bahwa cluster 1 dan cluster 2 tidak terpisah dengan baik. Hal ini menyebabkan hasil cluster pasien covid-19 dengan algoritma *K-Medoid* sulit diidentifikasi karakteristiknya.

Karakteristik Pasien Covid-19

Hasil penelitian ini didapatkan 2 cluster hasil analisis algoritma *K-Means* yang terdiri dari cluster 1 gejala sedang dan cluster 2 gejala berat. Pada variabel jenis kelamin penelitian ini dominan laki-laki, hal ini sejalan dengan penelitian meta-analisis yang mengumpulkan data dari 59.254 dari 11 negara yang berbeda menunjukkan jenis kelamin laki-laki lebih tinggi pada pasien Covid-19 (10). Banyaknya jenis kelamin laki-laki bisa dipengaruhi oleh aspek hormon seperti lebih tingginya ekspresi kadar reseptor ACE-2 yang dapat mempengaruhi meningkatkan viral load Covid-19 pada populasi laki-laki (11). Pada aspek sosial seperti laki-laki lebih cenderung beraktivitas di luar ketimbang perempuan sebagai contoh mencari nafkah hal ini memungkinkan tertularnya virus Covid-19 dan faktor lain seperti merokok menyebabkan peningkatan komorbiditas pada jenis kelamin laki-laki (12) Pada usia lanjut ≥ 50 tahun beresiko terjadi Covid-19 hal ini sejalan dengan penelitian yang menjelaskan usia >50 tahun memiliki risiko lebih tinggi mortalitas dikarenakan terdapat hubungan tingkat imunitas alami yang menurun memiliki kecenderungan untuk mudah terinfeksi (13). Sistem kekebalan tubuh melemah dengan bertambahnya usia, membuat lansia lebih sulit untuk melawan infeksi sehingga apabila peradangan terjadi bisa lebih hebat yang dapat menyebabkan kerusakan organ (14).

Adanya komorbid pada pasien Covid-19 dapat menyebabkan risiko lebih tinggi berkembangnya kejadian serius yaitu pasien masuk ruang ICU, intubasi

mekanis hingga kematian (15). Pada penelitian di Wuhan China, pasien Covid-19 banyak pasien yang ditemukan meninggal karena komorbid yaitu hipertensi (64%), DM (40%), gangguan jantung (32%) serta dilaporkan 2 atau lebih penyakit penyerta lebih sering terlihat pada kasus yang parah sebanyak 40% (16). Pada pasien Covid-19 dengan komplikasi memiliki pencetus pasien menjadi gejala kearah berat. Sistem pernapasan dan paru-paru adalah komplikasi tersering pada pasien Covid-19 dibandingkan komplikasi kardiovaskuler, hati, ginjal, gastrointestinal dan syaraf pusat (17). Jumlah gejala yang dimiliki pasien Covid-19 dapat mempengaruhi derajat keparahan Covid-19.

Pada tanda vital seperti laju pernapasan (nilai RR) merupakan skrining umum untuk mengidentifikasi terjadinya infeksi saluran pernapasan khususnya Covid-19 dengan tanda nilai $RR > 20x$ (18). Kedua suhu tubuh, faktor resiko terkait dengan Covid-19 yang paling banyak ialah demam (19). Hal ini sejalan dengan penelitian sistematik dan meta-analisis yang didapatkan secara keseluruhan pasien dewasa mengalami demam sebanyak 79,43% (20) Pada demam memiliki zat pirogen yang terbagi 2 yakni pirogen eksogen dari luar tubuh dan pirogen endogen seperti IL-1, IL-6, TNF-alpha dan IFN yang bersumber dari monosit, neutrophil dan limfosit(21). Tanda vital lainnya seperti nilai saturasi oksigen yang merupakan predictor kuat kematian pasien Covid-19 yang dapat dilihat dari kadar saturasi oksigen yang rendah pada tubuh. Saturasi oksigen rendah berhubungan dengan inflamasi, penurunan fungsi organ, lama rawat inap pasien terutama pada rawatan di ICU (22). Penelitian kohort retrospektif pasien Covid-19 yang mengalami kematian memiliki saturasi oksigen yang rendah dibawah 90% (23).

Pada pasien Covid-19, uji hematologi seperti leukosit pada gejala rendah hingga tinggi terjadi

perubahan nilai leukosit yang meningkat dikarenakan terjadinya aktivitas monosit, penurunan jumlah neutrophil dan sel *Natural Killer* (NK) hingga variasi produksi antibody dari sel B (24). Pada pasien Covid-19 terjadi penurunan limfosit, hal ini sejalan dengan penelitian meta-analisis pasien dengan hasil yang buruk memiliki nilai limfosit yang rendah dengan nilai $p\ value < 0,01$ (25). Penurunan yang menunjukkan bawah virus Covid-19 mempengaruhi sel imun dan menghambat fungsi sistem imun seluler, Melemahnya respon sel T dipengaruhi oleh sel Dendritik (DC) yang matang bermigrasi ke organ limfoid karena penting untuk pengaktifan sel T, hal ini yang menyebabkan terjadi limfositopenia (26).

Pada terapi pasien Covid-19 diberikan vitamin C, vitamin D, antibiotik, antivirus, antikoagulan, steroid dan terapi simptomatik sesuai dengan pedoman tatalaksana Covid-19 edisi 4 yaitu tahun 2022 (2)

Pada algoritma K-Means berdasarkan variabel gejala yang lebih banyak adalah 1-3 gejala. Pada penelitian ini pasien Covid-19 lebih banyak mengalami 1-3 gejala Covid-19 dikarenakan saat masuk rumah sakit gejala paling banyak yaitu demam, batuk dan sesak napas. Hal ini sejalan dengan penelitian di Korea Selatan sebanyak 5.628 pasien banyak mengalami demam 23,1% dan batuk 41,6%.

Penilaian kualitas cluster dievaluasi dari nilai DBI. Semakin kecil nilai DBI mendekati nilai 0 menunjukkan cluster yang optimal. Berdasarkan nilai DBI K-Means adalah 3,48. Nilai *Silhouette Coefficient* berada pada rentan (-1) hingga 1. Semakin tinggi nilai *sillhoutte* maka akan berpengaruh ke hasil yang lebih baik. Pada penelitian ini menghasilkan urutan pertama nilai *sillhoutte coefficient* yang baik ialah *K-Means* sebesar 0.07 selanjutnya *Fuzzy C-Means* sebesar 0.070 dan terakhir *K-Medoid* sebesar 0.04. Pada penelitian ini penggunaan metode K-means karena mampu mengelompokkan data besar dengan sangat cepat.

Pada data mining oleh IEEE algoritma K-means diakui sebagai salah dari 10 algoritma teratas (27). Selain itu *K-Means* juga mempunyai tingkat ketelitian yang cukup tinggi terhadap ukuran objek, sehingga algoritma ini efisien serta relatif lebih terukur untuk pengolahan objek dalam jumlah yang besar (28).

KESIMPULAN

Berdasarkan hasil penelitian dapat disimpulkan bahwa pengelompokan pasien Covid-19 berdasarkan gejala menggunakan algoritma K-Means menghasilkan 2 cluster yaitu sedang dan berat. Hasil evaluasi cluster menurut nilai DBI yaitu 3.48 dan berdasarkan *sillhouette coefficient* yang baik ialah *K-Means* sebesar 0.07.

SARAN

Bedasarkan hasil penelitian perlu ditambahkan variabel yang lebih mendukung seperti menilai koagulapati pasien dan marker inflamasi APTT, CT-Scan, LED, CRP untuk memudahkan pembagian tingkat keparahan Covid-19.

DAFTAR PUSTAKA

- Hui DS, I Azhar E, Madani TA, Ntoumi F, Kock R, Dar O, et al. The continuing 2019-nCoV epidemic threat of novel coronaviruses to global health The latest 2019 novel coronavirus outbreak in Wuhan, China. *Int J Infect Dis*. 2020;91:264–6.
- Burhan E, Susanto AD, Nasution SA, Eka G, Pitoyo ceva W, Susilo A, et al. Cedera miokardium pada infeksi COVID-19. *Pedoman tatalaksana COVID-19 edisi 4*. 2022. 79–85 p.
- Abdullah D, Susilo S, Ahmar AS, Rusli R, Hidayat R. The application of K-means clustering for province clustering in Indonesia of the risk of the COVID-19 pandemic based on COVID-19 data. *Qual Quant [Internet]*. 2021;(0123456789). Available from: <https://doi.org/10.1007/s11135-021-01176-w>
- Armstrong JJ, Zhu M, Hirdes JP, Stolee P. K-means cluster analysis of rehabilitation service users in the home health care system of Ontario: Examining the heterogeneity of a complex geriatric population. *Arch Phys Med Rehabil [Internet]*. 2012;93(12):2198–205. Available from: <http://dx.doi.org/10.1016/j.apmr.2012.05.026>
- Cristyani E, Tarigan B, Artika R, Febriawan MR, Muna KN. Covid-19 Risk Mapping with the Fuzzy C-Means Method in DIY Province. *Proc Internat Conf Sci Engin*. 2021;4(February):259–65.
- Samudi S, Widodo S, Brawijaya H. The K-Medoids Clustering Method for Learning Applications during the COVID-19 Pandemic. *SinkrOn*. 2020;5(1):116.
- Wijaya YA, Kurniady DA, Setyanto E, Tarihoran WS, Rusmana D, Rahim R. Davies Bouldin Index Algorithm for Optimizing Clustering Case Studies Mapping School Facilities. *TEM J*. 2021;10(3):1099–103.
- Di S, Utara S. Optimasi Jumlah Cluster Metode K-Medoids. 2022;(July).
- Kevin Herijanto C, Finsensia Riti Y. Perbandingan Algoritme K-Means Clustering Dan Fuzzy C-Means Untuk Penentuan Strategi Promosi Universitas Katolik Darma Cendika. *JATI (Jurnal Mhs Tek Inform*. 2024;8(3):3931–7.
- Do Nascimento IJB, Cacic N, Abdulazeem HM, von Groote TC, Jayarajah U, Weerasekara I, et al. Novel coronavirus infection (Covid-19) in humans: A scoping review and meta-analysis. *J Clin Med*. 2020;9(4).
- Alwani M, Yassin A, Al-Zoubi RM, Aboumarzouk OM, Nettleship J, Kelly D, et al. Sex-based differences in severity and mortality in COVID-19. *Rev Med Virol*. 2021;31(6).

12. Dalal J, Triulzi I, James A, Nguimbis B, Dri GG, Venkatasubramanian A, et al. COVID-19 mortality in women and men in sub-Saharan Africa: A cross-sectional study. *BMJ Glob Heal*. 2021;6(11):1–11.
13. Leng J, Goldstein DR. Impact of aging on viral infections. *Microbes and Infection*. 2010.
14. Lambrinou E, Hansen TB, Beulens JWJ. Lifestyle factors, self-management and patient empowerment in diabetes care. *Eur J Prev Cardiol*. 2019;
15. Nandy K, Salunke A, Pathak SK, Pandey A, Doctor C, Puj K, et al. Coronavirus disease (COVID-19): A systematic review and meta-analysis to evaluate the impact of various comorbidities on serious events. *Diabetes Metab Syndr Clin Res Rev*. 2020;14(5):1017–25.
16. Klein F. Risikofaktor Komorbiditäten bei COVID-19-Erkrankung. *Pneumologie*. 2020;74(10):640.
17. Kordzadeh-Kermani E, Khalili H, Karimzadeh I. Pathogenesis, clinical manifestations and complications of coronavirus disease 2019 (COVID-19). *Future Microbiol*. 2020;15(13):1287–305.
18. Massaroni C, Nicolò A, Schena E, Sacchetti M. Remote Respiratory Monitoring in the Time of COVID-19. *Front Physiol*. 2020;11(May):1–4.
19. World Health Organization. No Title. *Clin Manag COVID-19 interim Guid 27 May 2020 World Heal Organ*. 2020; <https://apps.who.int/iris/handle/10665/332196>.
20. Islam MA, Kundu S, Alam SS, Hossan T, Kamal MA HR. Prevalence and characteristics of fever in adult and paediatric patients with coronavirus disease 2019 (COVID-19): A Syst Rev meta-analysis 17515 patients. 2021; *PLoS ONE* 16(4): e0249788.
21. Guihur A, Rebeaud ME, Fauvet B, Tiwari S, Weiss YG, Goloubinoff P. Moderate Fever Cycles as a Potential Mechanism to Protect the Respiratory System in COVID-19 Patients. *Front Med*. 2020;7(September):1–8.
22. Docherty AB, Harrison EM, Green CA, Hardwick HE, Pius R, Norman L, et al. Features of 20 133 UK patients in hospital with covid-19 using the ISARIC WHO Clinical Characterisation Protocol: Prospective observational cohort study. *BMJ*. 2020;369(March):1–12.
23. Mejía F, Medina C, Cornejo E, Morello E, Vásquez S, Alave J, et al. Oxygen saturation as a predictor of mortality in hospitalized adult patients with COVID-19 in a public hospital in Lima, Peru. *PLoS One*. 2020;15(12 December):1–12.
24. Zhou F, Yu T, Du R, Fan G, Liu Y, Liu Z, et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. *Lancet [Internet]*. 2020;395(10229):1054–62. Available from: [http://dx.doi.org/10.1016/S0140-6736\(20\)30566-3](http://dx.doi.org/10.1016/S0140-6736(20)30566-3)
25. Huang I, Pranata R. Lymphopenia in severe coronavirus disease-2019 (COVID-19): Systematic review and meta-analysis. *J Intensive Care*. 2020;8(1):1–10.
26. Shi Y, Wang Y, Shao C, Huang J, Gan J, Huang X, et al. COVID-19 infection: the perspectives on immune responses. *Cell Death Differ [Internet]*. 2020;27(5):1451–4. Available from: <http://dx.doi.org/10.1038/s41418-020-0530-3>
27. Wu X, Kumar V, Quinlan JR, Ghosh J, Yang Q, Motoda H, et al. Top 10 algorithms in data mining. 2007.
28. Putu N, Merliana E, Santoso AJ. Analisa Penentuan Jumlah Cluster Terbaik pada Metode K-Means. :978–9.