



Implementasi Algoritma *K-Means* untuk Pengelompokan Lama Sembuh Pasien Covid-19

Nadya Octavianna Lompoliuw^{1*}, Hindriyanto Dwi Purnomo²

^{1,2} Program Studi Teknik Informatika, Fakultas Teknologi Informasi, Universitas Kristen Satya Wacana, Kota Salatiga, Provinsi Jawa Tengah, Indonesia.

article info

Article history:

Received 8 July 2022

Received in revised form

22 November 2022

Accepted 12 February 2023

Available online April 2023

DOI:

<https://doi.org/10.35870/jtik.v7i2.706>

Keywords:

COVID-19; Clustering; K-Means; RapidMiner.

Kata Kunci:

COVID-19; Clustering; K-Means; RapidMiner.

abstract

COVID-19 (coronavirus disease 2019) is a disease caused by infection with the severe acute respiratory syndrome Coronavirus 2 (SARS-CoV-2). COVID-19 can cause disturbances in the respiratory system, ranging from mild symptoms such as flu, cough to lung infections. The spread of this virus is still happening, therefore the purpose of this study is so that the public can find out more information about COVID-19, one of which is the recovery time for COVID-19 patients consisting of 4 clusters, which are very long, long, fast, and very fast. Judging by the length of the recovery. Grouping the recovery time of COVID-19 patients based on age, positive date and negative date using the K-Means Algorithm. The K-Means algorithm is an iterative grouping that partitions the data set into a number of K clusters that have been determined in the initial data. Of the 4 clusters that have been determined, the results obtained are cluster 1 (very long) which is the cluster that has the longest recovery time, which is 41-48 days with ages 31 and 48 years. It can also be seen based on the symptoms experienced at each different age, on average all experience symptoms of fever and cough.

abstract

COVID-19 (coronavirus disease 2019) adalah penyakit akibat infeksi virus severe acute respiratory syndrome Coronavirus 2(SARS-CoV-2). COVID-19 dapat mengakibatkan gangguan pada sistem pernapasan, dimulai dari gejala ringan seperti flu, batuk hingga infeksi paru-paru. Penyebaran virus ini sampai sekarang masih terjadi, maka dari itu tujuan dalam penelitian ini adalah agar masyarakat bisa lebih mengetahui informasi tentang COVID-19 salah satunya yaitu lama sembuhnya pasien COVID-19 yang terdiri dalam 4 cluster sangat lama, lama, cepat, dan sangat cepat dilihat dari lama sembuhnya. Pengelompokan lama sembuh pasien COVID-19 berdasarkan umur, tanggal positif dan tanggal negatif menggunakan Algoritma K-Means. Algoritma K-Means adalah pengelompokan iteratif yang melakukan partisi set data ke dalam sejumlah K cluster yang telah ditentukan di data awal. Dari 4 cluster yang telah ditentukan hasil yang di dapatkan yaitu cluster 1 (sangat lama) merupakan cluster yang memiliki lama sembuh yang paling lama yaitu 41 – 48 hari dengan umur 31 dan 48 tahun. Bisa dilihat juga berdasarkan gejala yang dialami di setiap umur berbeda – beda, rata – rata semua mengalami gejala demam dan batuk.

Corresponding author. Email: 672018029@student.uksw.edu ^{1}.

1. Latar Belakang

Dalam era kemajuan teknologi yang terus berkembang hingga saat ini, penggunaan teknologi telah meluas ke berbagai sektor, termasuk di bidang ekonomi dan bidang lainnya [1]. Dalam bidang kesehatan, teknologi juga telah terbukti sangat membantu dalam mengelompokkan data pasien COVID-19, khususnya dalam hal lama waktu penyembuhan pasien. Pentingnya kesehatan tidak dapat dipungkiri, karena kesehatan yang baik sangat menentukan produktivitas dan kualitas hidup manusia [2]. Oleh karena itu, teknologi yang digunakan dalam bidang kesehatan sangat penting untuk memastikan bahwa pasien mendapatkan perawatan yang tepat dan waktu penyembuhan yang optimal.

Pada akhir tahun 2019 tepatnya awal bulan Desember, seluruh negara digemparkan dengan sebuah kejadian yang membuat banyak masyarakat gelisah yaitu COVID-19. Kejadian itu bermula di Wuhan, Provinsi Hubei, China. Awalnya virus COVID-19 ini diduga karena paparan pasar grosir makanan laut Huanan yang menjual banyak spesies hewan hidup [3]. Virus ini begitu cepat menyebar di wilayah lain di China. Munculnya Covid-19 sudah menarik perhatian dunia, pada 30 Januari 2020 WHO (*World Health Organization*) telah mengumumkan COVID-19 sebagai darurat kesehatan masyarakat yang menjadi perhatian nasional [4]. Kenaikan jumlah kasus virus ini bisa dibilang cukup cepat dan telah terjadi penyebaran antar negara. Pada tanggal 25 Maret 2020, diberitahukan total kasus konfirmasi 414.179 dengan 18.440 kematian (CR 4.4%) dimana kasus yang dilaporkan 192 wilayah/negara. Dari beberapa kasus yang terjadi, terkonfirmasi ada sebagian petugas kesehatan yang telah dilaporkan terinfeksi. Virus COVID-19 ini adalah zoonosis yang artinya yaitu penyakit yang dapat ditularkan hewan ke manusia. Orang yang berisiko terjangkit virus tersebut yaitu mempunyai kontak erat dengan orang yang terpapar COVID-19 dan orang yang merawat pasien positif tersebut [5]. Infeksi ini menular melalui droplet yang berasal dari batuk ataupun bersin dengan gejala, tetapi juga bisa menular dengan tanpa gejala. Sebab, Droplet bisa menebar hingga 1–2 meter dan virus yang terdapat didalam itu bisa bertahan dalam beberapa hari, tetapi virus ini bisa dimusnahkan dengan disinfektan. Penyebab

COVID-19 ini terjadi karena menghirup udara sekitar, dan menyentuh permukaan benda atau tempat yang terkontaminasi. Penyebaran COVID-19 ini bisa melalui mukosa hidung, mulut dan mata. Penyebaran virus dapat terjadi pada semua kelompok umur [6].

Berdasarkan penelitian terdahulu yang pertama ini untuk analisis survival pada pasien COVID-19 di Kabupaten Timur Tengah Utara (TTU), dimana pengolahan data menggunakan uji *Kaple Meier* dan uji *Log-Rank*. Tujuannya adalah untuk mengetahui lama waktu sembuh dan perbedaan lama waktu sembuh pada pasien COVID-19 yang memiliki usia kurang dari 40 tahun dan di atas 40 tahun. Data penelitian ini menggunakan data pasien COVID-19 di Timor Tengah Utara yang didapatkan dari Dinas Kesehatan Kabupaten TTU pada April 2020 hingga 31 Agustus 2021. Hasil analisis dalam penelitian ini adalah berdasarkan hasil uji *Kaple Meier* peluang lama waktu sembuh pada pasien COVID-19 secara keseluruhan memiliki median 10 hari perawatan dan perbedaan lama waktu sembuh usia kurang dari 40 tahun memiliki median 9 hari perawatan dan di atas 40 tahun memiliki median 10 hari perawatan, dan berdasarkan hasil uji *Log-Rank* perbedaan waktu perawatan yang signifikan, pasien usia kurang dari 40 tahun sembuh lebih cepat dibandingkan pasien di atas 40 tahun [7].

Penelitian terdahulu yang kedua ini untuk mengetahui gambaran usia pada kejadian COVID-19, dimana pengolahan data nya menggunakan desain *case control*. Tujuannya adalah untuk melakukan analisis terhadap gambaran usia terhadap pasien COVID-19. Data dari penelitian ini menggunakan data sekunder dari pasien yang dirawat inap di ruang isolasi COVID-19 di RSMH Palembang pada tanggal 1 Maret 2020 s/d 31 Juli 2020. Penelitian ini memiliki 2 kriteria yaitu usia berisiko (>65 tahun) dan usia tidak berisiko (<65 tahun). Hasil analisis dalam penelitian ini adalah usia berisiko memiliki persentasi 40,2% dan usia tidak berisiko presentasi nya 47,45% [8].

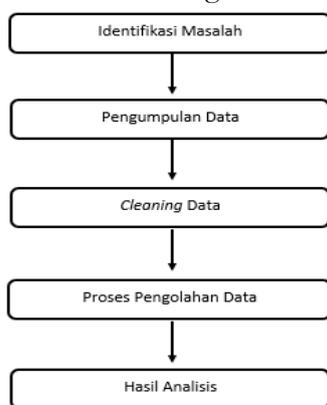
Dengan begitu, bukan hanya gejala saja yang penting, umur juga sangat penting untuk kita kenali dan memahami berapa lama sembuhnya seseorang yang positif COVID-19. Mengenali rentan umur mana saja yang proses penyembuhan paling lama sejak dini. Pada dasarnya seseorang yang positif COVID-19 disarankan untuk melakukan Swab Test paling lambat setelah 14 hari terkonfirmasi tertular COVID-19.

Pada saat ini kasus COVID-19 di Indonesia bisa dibilang cukup menurun. Walaupun kasus COVID-19 menurun kita harus tetap waspada, karena tidak menutup kemungkinan virus ini masih ada di sekitar kita dan berisiko kita bisa tertular virus ini. Infeksi virus ini bisa terinfeksi kepada siapa saja, tidak memandang umur dan kondisi kesehatan seseorang. Apalagi masih ada banyak orang yang meremehkan virus ini, dengan tidak mematuhi protokol kesehatan salah satunya tidak menggunakan masker.

Maka dari itu, dalam penelitian ini akan dilakukan pengelompokan berdasarkan umur, tanggal positif, dan tanggal negatif pada orang yang positif COVID-19 di Kota Salatiga menggunakan metode algoritma *K-Means*. Data yang akan digunakan diambil dari Dinas Kesehatan Kota Salatiga. Dari hasil penelitian ini, masyarakat bisa lebih mengetahui informasi tentang COVID-19 salah satunya lama sembuhnya pasien COVID-19 dalam 4 kategori sangat lama, lama, cepat, dan sangat cepat dilihat dari lama sembuhnya.

2. Metode Penelitian

Penelitian ini menggunakan algoritma *K-Means* untuk mengelompokkan lama sembuhnya pasien COVID-19 berdasarkan umur, tanggal positif, dan tanggal negatif. Dataset yang akan digunakan dalam penelitian ini diambil dari Data Pasien COVID-19 di Dinas Kesehatan Kota Salatiga.



Gambar 1. Tahapan Penelitian

Identifikasi Masalah

Tahap yang pertama ini adalah identifikasi masalah. Identifikasi masalah ini dilakukan guna untuk mengetahui tujuan dan manfaat dari penelitian yang dilakukan. Tujuan dari penelitian ini adalah untuk mengetahui pengelompokan lama sembuhnya

pasien COVID-19. Manfaat dari penelitian ini adalah agar masyarakat bisa lebih mengetahui informasi tentang COVID-19 salah satunya yaitu lama sembuhnya pasien COVID-19.

COVID-19

COVID-19 (*coronavirus disease 2019*) adalah penyakit yang berasal dari infeksi virus *severe acute respiratory syndrome Coronavirus 2* (SARS-CoV-2). Virus ini bisa mengakibatkan gangguan pada sistem pernapasan, dimulai dari gejala ringan seperti flu sampai yang lebih serius, seperti *Middle East Respiratory Syndrome* (MERS) dan *Severe Acute Respiratory Syndrome* (SARS) atau sindrom pernapasan akut berat [9]. Virus ini adalah sejenis penyakit baru yang belum pernah ditemukan sebelumnya pada manusia [10]. Ada beberapa faktor yang harus diperhatikan terhadap pasien COVID-19 agar bisa cepat sembuh yaitu, faktor yang pertama harus berpikiran positif karena kesehatan mental sangat berpengaruh terhadap kondisi fisik seseorang, faktor selanjutnya konsumsi nutrisi yang tepat, minum air putih yang banyak, makan makanan sehat, makan buah, sayur dan diusahakan hindari makanan yang berlemak tinggi dan juga istirahat yang cukup [11].

Pengumpulan Data

Tahap yang kedua adalah melakukan pengumpulan data yang orang yang positif COVID-19 pada bulan Januari – Maret 2022 yang didapatkan dari data Dinas Kesehatan Kota Salatiga. Dari data yang didapatkan terdapat beberapa atribut. Atribut tersebut yang akan digunakan pada penelitian ini yang selanjutnya akan diproses menggunakan metode *K-Means*.

Cleaning Data

Tahap yang ketiga adalah *cleaning* data yang bertujuan untuk mengambil data yang akan digunakan dalam penelitian ini yaitu umur, tanggal positif, tanggal negatif. Dan juga mengambil gejala untuk informasi tambahan pada penelitian ini. Dalam menghitung hari sembuhnya pasien COVID-19 akan menggunakan *Microsoft Office Excel*.

Microsoft Office Excel

Microsoft Office Excel adalah salah satu produksi dari *Microsoft Corporation* dalam paket *Microsoft Office*. Fungsi *Microsoft Excel* ini untuk memproses data seperti angka atau bilangan dengan menggunakan *spreadsheet* yang terdapat baris dan kolom [12]. Selain cara penggunaannya mudah, *Microsoft Excel* ini

mendapatkan banyak pengakuan dari berbagai segala jenis profesi akan ketangguhannya dalam menangani sekaligus menyelesaikan berbagai persoalan. Dari penggunaan yang ringan sampai penggunaan yang berat yang berhubungan dengan pengolahan perhitungan dan data diperoleh dengan hasil yang maksimal dan otomatis [13].

Proses Pengolahan Data

Tahap yang keempat adalah dalam *clustering* data akan diproses dengan metode Algoritma *K-Means* menggunakan aplikasi *RapidMiner*. Terdapat 4 *cluster* yang akan digunakan yaitu sangat lama, lama, cepat, dan sangat cepat. Untuk variabel yang akan digunakan dalam proses ini yaitu umur dan lama sembuh.

Algoritma K-Means

Algoritma *K-Means* adalah pengelompokan iteratif yang melakukan partisi set data ke dalam sejumlah *K cluster* yang telah ditentukan di data awal. Algoritma *K-Means* biasa jika untuk diimplementasikan dan dijalankan, termasuk relatif cepat, dapat mudah beradaptasi, penggunaannya umum dalam praktek. *K-Means* bisa digunakan pada data yang dipresentasikan dalam *r*-dimensi ruang tempat. *K-means* mengelompokkan set data *r*-dimensi, $X = \{x_i | i=1, \dots, N\}$. Algoritma *KMeans* mengelompokkan semua titik data dalam *X* sehingga setiap titik x_i hanya jatuh dalam satu *K* partisi [14].

Clustering

Clustering adalah proses dimana pengelompokkan benda serupa dalam sebuah kelompok yang berbeda, atau tepatnya partisi dari sebuah data set kedalam subset, hingga data di dalam setiap subset memiliki arti yang berguna [15]. Dalam pengelompokkan *clustering* menggunakan teknik *unsupervised learning* dimana tidak dibutuhkan *fase learning* dan juga tidak memakai pelabelan pada setiap kelompok [16].

RapidMiner Studio

RapidMiner adalah platform perangkat lunak ilmu data yang dikembangkan oleh perusahaan bernama sama dengan yang menyuplai lingkungan terintegrasi untuk persiapan data, pembelajaran dan analisis prediktif. Hal tersebut dipakai untuk bisnis dan komersial, dan untuk penelitian, pendidikan, pelatihan, *rapid prototyping*, dan juga pengembangan aplikasi [17]. *RapidMiner* dikenal dengan sebutan

YALE (*Yet Another Learning Environment*), dikemukakan oleh Ralf Klinkenberg, Ingo Mierswa, dan Simon Fischer merupakan Unit Kecerdasan Buatan Universitas Teknik Dortmund tahun 2001 [18].

3. Hasil dan Pembahasan

Sumber data dalam penelitian ini adalah data awal yang diambil dari Dinas Kesehatan Kota Salatiga, dimana data awal ini merupakan data COVID-19 di Kota Salatiga pada bulan Januari - Maret tahun 2022 terdiri dari 1.600 data yang terdiri dari berbagai variabel. Untuk penelitian ini diambil data yang dibutuhkan yaitu umur, tanggal positif, dan tanggal negatif. Kemudian nanti akan di proses untuk menentukan lama sembuhnya berdasarkan tanggal.

No.	Umur (tahun)	Tanggal Positif	Tanggal Negatif
1.	37	04/02/2022	16/02/2022
2.	55	04/02/2022	16/02/2022
3.	66	05/02/2022	17/02/2022
4.	28	06/02/2022	20/02/2022
5.	15	07/02/2022	20/02/2022
6.	4	02/02/2022	17/02/2022
7.	51	06/02/2022	20/02/2022
8.	16	07/02/2022	18/02/2022
9.	9	07/02/2022	18/02/2022
10.	25	07/02/2022	17/02/2022
11.	17	09/02/2022	22/02/2022
12.	29	10/02/2022	22/02/2022
13.	21	10/02/2022	23/02/2022
14.	16	10/02/2022	23/02/2022
15.	14	10/02/2022	23/02/2022
16.	3	10/02/2022	23/02/2022
17.	43	10/02/2022	21/02/2022
18.	6	04/02/2022	08/02/2022
19.	2	10/02/2022	23/02/2022
20.	6	07/02/2022	20/02/2022
21.	10	07/02/2022	20/02/2022
22.	35	12/02/2022	28/02/2022
23.	11	11/02/2022	22/02/2022
24.	12	13/02/2022	01/03/2022
25.	31	14/02/2022	03/04/2022

Gambar 2. Dataset COVID-19

Pada gambar 2 merupakan sebagian data yang telah di cleaning dari data awalnya yang berjumlah 1.600 data. Dari data awal diambil 150 data, karena banyak data yang ada tanggal positifnya tapi tidak ada tanggal negatifnya, begitupun sebaliknya.

No.	Umur	Gejala COVID-19
1.	37	Batuk, gatal tenggorokan
2.	55	Demam
3.	66	Demam, batuk, pilek
4.	28	Demam, batuk, tenggorokan panas
5.	15	Batuk
6.	4	Batuk
7.	51	Batuk, demam, pegal linu
8.	16	Batuk, gatal tenggorokan
9.	9	Batuk
10.	25	Batuk, gatal tenggorokan
11.	17	Demam, batuk
12.	29	Demam
13.	21	Demam
14.	16	Demam
15.	14	Demam, batuk
16.	3	Demam
17.	43	Demam
18.	6	Demam, batuk
19.	2	Batuk, pilek
20.	6	Demam, batuk
21.	10	Demam, batuk, gatal tenggorokan, mengigil
22.	35	Batuk, gatal tenggorokan
23.	11	Batuk, gatal tenggorokan
24.	12	Demam, batuk
25.	31	Demam, batuk, pilek

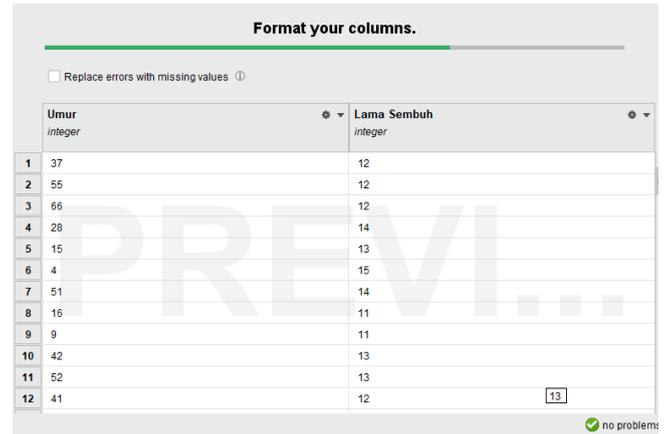
Gambar 3. Data Umur dan Gejala

Pada gambar 3 merupakan sebagian data gejala yang dialami oleh pasien COVID-19 sebagai informasi tambahan dalam penelitian ini. Rata – rata gejala yang dialami yaitu demam, batuk, dan pilek.

No.	Umur	Lama Sembuh (hari)
1.	37	12
2.	55	12
3.	66	12
4.	28	14
5.	15	13
6.	4	15
7.	51	14
8.	16	11
9.	9	11
10.	25	10
11.	17	13
12.	29	12
13.	21	13
14.	16	13
15.	14	13
16.	3	13
17.	43	11
18.	6	4
19.	2	13
20.	6	13
21.	10	13
22.	35	16
23.	11	11
24.	12	16
25.	31	48
26.	32	14
27.	35	20
28.	22	17
29.	74	16

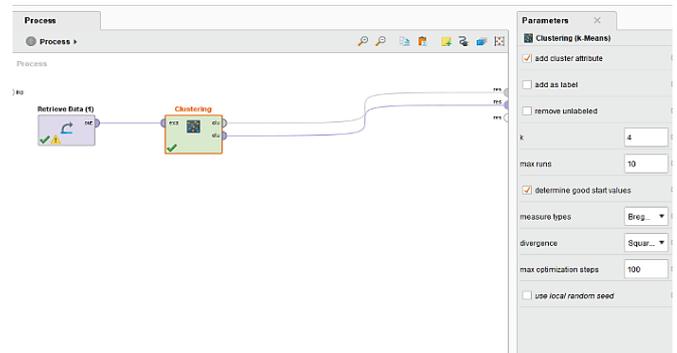
Gambar 4. Umur dan Lama Sembuh

Pada gambar 4 merupakan data umur beserta lama sembuh seseorang pasien COVID-19. Lama sembuh dihitung dalam hari. Dari dataset awal tidak ada data lama sembuh. Maka dari itu, untuk menentukan lama sembuh menggunakan *Microsoft Office Excel* dengan formula `=DATEDIF(C2,D2,"d")` sebagai contoh tinggal menyesuaikan dengan *column* dan *row* yang ada pada data berdasarkan tanggal negatif dan tanggal positifnya.



Gambar 5. Memberikan Label Pada Dataset

Pada gambar 5 merupakan data yang telah di *cleaning* yang akan diproses di *RapidMiner* yang sebelumnya diberikan label terdahulu. Dalam pengolahan dataset ini menggunakan algoritma *K-Means*, terdapat 4 *cluster* yaitu sangat lama, lama, cepat, dan sangat cepat.



Gambar 6. Hasil Akhir Design Pada *RapidMiner*

Pada gambar 6 merupakan hasil akhir *design* pada aplikasi. Proses ini mulai dari import data jarak sembuh lalu dihubungkan dengan *operators Clustering K-Means* lalu mengisi berapa “k” yang akan di *cluster* yaitu 4.

Row No.	id	cluster	Umur	Lama Sembuh
1	1	cluster_0	37	12
2	2	cluster_3	55	12
3	3	cluster_3	66	12
4	4	cluster_0	28	14
5	5	cluster_2	15	13
6	6	cluster_2	4	15
7	7	cluster_3	51	14
8	8	cluster_2	16	11
9	9	cluster_2	9	11
10	10	cluster_0	42	13
11	11	cluster_3	52	13
12	12	cluster_0	41	12
13	13	cluster_0	39	10
14	14	cluster_0	25	10
15	15	cluster_2	17	13

ExampleSet (150 examples, 2 special attributes, 2 regular attributes)

Gambar 7. Data View

Pada gambar 7 merupakan sheet di Result untuk menampilkan data yang telah diproses secara keseluruhan dengan cluster dari example set. Hasil nya terdapat 4 cluster yaitu cluster 0, cluster 1, cluster 2.

Cluster Model

Cluster 0: 48 items
 Cluster 1: 2 items
 Cluster 2: 63 items
 Cluster 3: 37 items
 Total number of items: 150

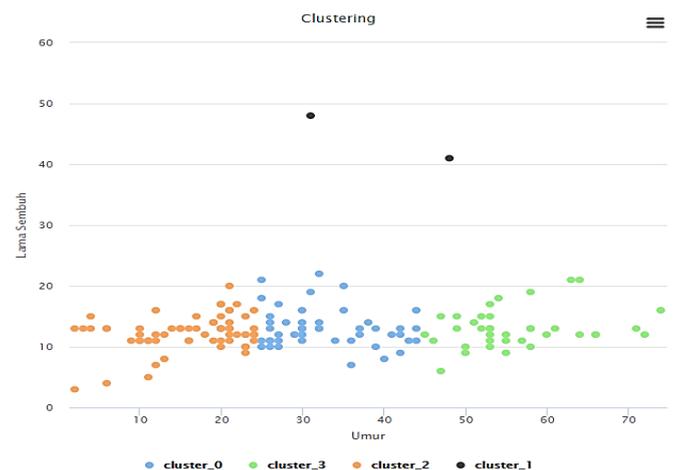


Gambar 8. Cluster Model

Pada gambar 8 merupakan hasil dari Cluster Model terdapat jumlah items pada setiap cluster yaitu cluster 1 (sangat lama) terdapat 2 items, cluster 0 (lama) terdapat 37 items, cluster 3 (cepat) terdapat 48 items, dan cluster 2

(sangat cepat) terdapat 63 items. Berikut uraian dari hasil cluster tersebut :

- 1) Cluster 1 (sangat lama) terdiri dari umur 31 tahun dan 48 tahun.
- 2) Cluster 0 (lama) terdiri dari umur 55 tahun, 66 tahun, 51 tahun, 52 tahun, 47 tahun, 74 tahun, 54 tahun, 60 tahun, 50 tahun, 45 tahun, 66 tahun, 71 tahun, 53 tahun, 61 tahun, 58 tahun, 49 tahun, 64 tahun, 72 tahun, 58 tahun, 46 tahun, 63 tahun, 64 tahun, dan 57 tahun.
- 3) Cluster 3 (cepat) terdiri dari umur 37 tahun, 28 tahun, 42 tahun, 41 tahun, 39 tahun, 25 tahun, 29 tahun, 43 tahun, 35 tahun, 36 tahun, 30 tahun, 32 tahun, 26 tahun, 27 tahun, 34 tahun, 38 tahun, 44 tahun, 40 tahun, dan 31 tahun.
- 4) Cluster 2 (sangat cepat) terdiri dari umur 15 tahun, 4 tahun, 16 tahun, 9 tahun, 17 tahun, 21 tahun, 14 tahun, 3 tahun, 6 tahun, 2 tahun, 10 tahun, 11 tahun, 20 tahun, 12 tahun, 23 tahun, 22 tahun, 19 tahun, 24 tahun, 13 tahun, 18 tahun, dan 4 tahun.



Gambar 9. Visualizations

Pada gambar 8 merupakan sheet di Result untuk menampilkan data yang telah diproses secara keseluruhan dengan cluster dari example set dalam bentuk diagram Scatter/Bubble. Dari hasil clustering nya dilihat lama sembuh nya cluster 1 = 'sangat lama' dengan lama sembuh 41 – 48 hari, cluster 0 = 'lama' dengan lama sembuh 8 – 22 hari, cluster 3 = 'cepat' dengan lama sembuh 6 – 21 hari, dan cluster 2 'sangat cepat' dengan lama sembuh 4 – 20 hari.

Perbedaan dari hasil yang ada dapat dilihat dari penelitian terdahulu yang pertama hanya menentukan lama sembuh pasien COVID-19 dengan kriteria peluang dan perbedaan lama waktu sembuh dan untuk

penelitian terdahulu yang kedua dengan kriteria usia berisiko dan usia tidak berisiko. Sedangkan dalam penelitian dengan kriteria sangat lama, lama, cepat dan sangat cepat dalam pengelompokkan lama sembuh pasien COVID-19 masih belum ada penelitiannya, maka dari itu penelitian ini dilakukan.

4. Kesimpulan

Berdasarkan dari hasil penelitian ini terdapat kesimpulan dalam penerapan metode Algoritma *K-Means* untuk pengelompokkan lama sembuh COVID-19 berdasarkan umur, tanggal positif dan tanggal negatif menggunakan aplikasi *RapidMiner*. Dari 150 data yang ada terdapat 4 *cluster* yang telah ditentukan, dapat dilihat *cluster* 1 (sangat lama) adalah *cluster* yang memiliki lama sembuh yang paling lama yaitu 41 – 48 hari dengan umur 31 dan 48 tahun. Dari dataset yang ada bisa dilihat juga berdasarkan gejala yang dialami di setiap umur berbeda-beda, rata-rata semua mengalami gejala demam dan batuk. Hasil penelitian ini kiranya dapat dijadikan informasi tambahan bagi masyarakat agar dapat lebih mengetahui informasi tentang COVID-19 yaitu lama sembuh pasien COVID-19 karena penyebaran virus ini masih terjadi hingga saat ini, maka dari itu masyarakat harus tetap mewaspadai virus ini dan menjaga kesehatan. Dari hasil penelitian dalam pengelompokkan lama sembuh pasien COVID-19, mendapatkan hasil yang cukup baik karena hasil analisis sesuai dengan data yang ada. Dalam penelitian ini tentunya tidak terlepas dari kekurangan dan kelemahan. Maka dari itu, terdapat saran perlu dilakukan proses penelitian lebih lanjut terhadap *clustering* lainnya agar mendapat perbandingan hasil yang lebih baik dan lebih maksimal.

5. Daftar Pustaka

- [1] Kidi, O. and Sos, S., 2018. Teknologi dan Aktivitas Dalam Kehidupan Manusia. *Jurnal Media*.
- [2] Lifiyasari, M. 2019. Universitas Ngudi Waluyo, *Lifiyasar*, (186), hal. 6925408.
- [3] Putri, N.A., Putra, A.E. and Mariko, R., 2021. Hubungan usia, Jenis kelamin dan gejala dengan kejadian COVID-19 di Sumatera Barat. *Majalah Kedokteran Andalas*, 44(2), pp.104-111. DOI: <https://doi.org/10.25077/mka.v44.i2.p104-111.2021>.
- [4] Susilo, A., Rumende, C.M., Pitoyo, C.W., Santoso, W.D., Yulianti, M., Herikurniawan, H., Sinto, R., Singh, G., Nainggolan, L., Nelwan, E.J. and Chen, L.K., 2020. Coronavirus disease 2019: Tinjauan literatur terkini. *Jurnal penyakit dalam Indonesia*, 7(1), pp.45-67. DOI: 10.7454/jpdi.v7i1.415.
- [5] Yuliana, L.W., 2020. Karakteristik gejala klinis kehamilan dengan Coronavirus disease (COVID-19). *Jurnal Ilmiah Kesehatan Sandi Husada*, 9(2), pp.726-734. DOI: 10.35816/jiskh.v10i2.397.
- [6] Putri, R.N., 2020. Indonesia dalam menghadapi pandemi Covid-19. *Jurnal Ilmiah Universitas Batanghari Jambi*, 20(2), pp.705-709. DOI: 10.33087/jiubj.v20i2.1010.
- [7] Matematika, P. S. *et al.* 2022. ANALISIS SURVIVAL PADA PASIEN COVID-19 DI KABUPATEN TIMOR TENGAH UTARA (TTU). 4(1), hal. 87–95.
- [8] Elviani, R., Anwar, C. and Sitorus, R.J., 2021. Gambaran usia pada kejadian Covid-19. *JAMBI MEDICAL JOURNAL" Jurnal Kedokteran dan Kesehatan"*, 9(2), pp.204-209. DOI: 10.22437/jmj.v9i1.11263.
- [9] Nasution, N.H. and Hidayah, A., 2021. Gambaran pengetahuan masyarakat tentang pencegahan covid-19 di kecamatan padangsidiempuan batunadua, kota padangsidiempuan. *Jurnal Kesehatan Ilmiah Indonesia/Indonesian Health Scientific Journal*, 6(1), pp.107-114. DOI: <http://dx.doi.org/10.51933/health.v6i1.419>.

- [10] Levani, Y., Prastya, A.D. and Mawaddatunnadila, S., 2021. Coronavirus disease 2019 (COVID-19): patogenesis, manifestasi klinis dan pilihan terapi. *Jurnal Kedokteran dan kesehatan*, 17(1), pp.44-57. DOI: <https://doi.org/10.24853/jkk.17.1.44-57>.
- [11] Azania, D. and Naan, N., 2021. Peran Spiritual Bagi Kesehatan Mental Mahasiswa di Tengah Pandemi Covid-19. *HUMANISTIKA: Jurnal Keislaman*, 7(1), pp.26-45. DOI: <https://doi.org/10.55210/humanistika.v7i1.384>.
- [12] Andriyani, D., Harahap, E., Badruzzaman, F.H., Fajar, M.Y. and Darmawan, D., 2019. Aplikasi Microsoft Excel Dalam Penyelesaian Masalah Rata-rata Data Berkelompok. *Matematika: Jurnal Teori dan Terapan Matematika*, 18(1). DOI: <https://doi.org/10.29313/jmtm.v18i1.5078>.
- [13] Idayat, R. *et al.* 2022. Pemanfaatan Penggunaan Microsoft Excel Secara Efektif Terhadap Pembelajaran Siswa-Siswi SMP Muhammadiyah 44 Pamulang. *JATIMIKA: Jurnal ...*, 2, hal. 331–336. Tersedia pada: <http://openjournal.unpam.ac.id/index.php/JATIMIKA/article/view/15996>.
- [14] Auliasari, K. and Kertaningtyas, M., 2019. Penerapan Algoritma K-Means untuk Segmentasi Konsumen Menggunakan R. *JTMI Jurnal Teknologi Dan Manajemen Informatika*, 5(2). DOI: 10.26905/jtmi.v5i2.3644.
- [15] Indraputra, R.A. and Fitriana, R., 2020. K-Means clustering data COVID-19. *Jurnal Teknik Industri*, 10(3), pp.275-282.
- [16] Priyatman, H., Sajid, F. and Haldivany, D., 2019. Klasterisasi Menggunakan Algoritma K-Means Clustering untuk Memprediksi Waktu Kelulusan Mahasiswa. *Jurnal Edukasi Dan Penelitian Informatika (JEPIN)*, 5(1), p.62. DOI: 10.26418/jp.v5i1.29611.
- [17] Prasetyo, V.R., Lazuardi, H., Mulyono, A.A. and Lauw, C., 2021. Penerapan Aplikasi RapidMiner Untuk Prediksi Nilai Tukar Rupiah Terhadap US Dollar Dengan Metode Regresi Linier. *Jurnal Nasional Teknologi dan Sistem Informasi (TEKNOSI)*, 7(1), pp.8-17. DOI: 10.25077/teknosi.v7i1.2021.8-17.
- [18] Nasir, J., 2020. Penerapan Data Mining Clustering Dalam Mengelompokan Buku Dengan Metode K-Means. *Jurnal Simetris*, 11(2). DOI: 10.24176/simet.v11i2.5482.