

PENERAPAN K-MEANS DALAM PENGELOMPOKKAN PEMANTAUAN KASUS TERKONFIRMASI COVID-19

Desy Winarty^{1*}; Yumi Novita Dewi²

Sistem Informasi^{1,2}
Universitas Nusa Mandiri^{1,2}
<https://nusamandiri.ac.id/>^{1,2}
desytrjmn@gmail.com^{1*}; yumi.ymd@nusamandiri.ac.id²



Abstract— Covid-19 is a large group of viruses that can cause disease in humans and animals. A new type of corona virus that was discovered in humans since an extraordinary incident in December 2019 in Wuhan, China was later named Severe Acute Respiratory Syndrome coronavirus 2 (SARS-COV2) and causes Coronavirus disease 2019 (COVID-19). Cases of the spread of Covid-19 in Indonesia continue to increase, so people are asked to maintain their distance to break the chain of transmission of the virus. In this study, calculations will be carried out using Rapidminer with the K-Means algorithm, in monitoring confirmed cases of Covid-19 in five Sub-Districts namely; Ex. Cipinang Melayu, Kel. Prime Minister. Ex. Nutmeg Gardens, Kel. Makassar, and Ex. Pinang Ranti in Makassar District. By looking at the data, it will be known which groups should receive special attention, such as; independent isolation, treated, recovered, and died. The grouping is divided into 4 clusters based on the data criteria for confirmed cases of Covid-19 totaling 265 cases. The cluster model leads to 225 cases of cluster0, 6 cases of cluster1, 12 cases of cluster2 and 22 cases of cluster3 with a total of 265 cases. This study resulted in grouping monitoring data for Covid-19 patients in the Makassar Village which required special monitoring in cluster 0.

Keywords: Covid-19, K-Means, Monitoring

Abstrak— Covid-19 adalah sekelompok besar virus yang dapat menyebabkan penyakit pada manusia dan hewan. Virus corona jenis baru yang ditemukan pada manusia sejak kejadian luar biasa pada Desember 2019 di Wuhan, China kemudian diberi nama Severe Acute Respiratory Syndrome coronavirus 2 (SARS-COV2) dan menyebabkan penyakit Coronavirus 2019 (COVID-19). Kasus penyebaran Covid-19 di Indonesia terus meningkat, sehingga masyarakat diminta menjaga jarak untuk memutus mata rantai penyebaran virus tersebut. Dalam penelitian ini akan dilakukan perhitungan menggunakan Rapidminer dengan algoritma K-Means, dalam memantau kasus terkonfirmasi positif Covid-19 pada lima Kelurahan yaitu; Kel. Cipinang Melayu, Kel. Halim Perdana. Kel. Kebon Pala, Kel. Makasar, dan Kel. Pinang Ranti di Kecamatan Makasar. Dengan melihat data tersebut akan diketahui kelompok mana yang harus mendapatkan perhatian khusus seperti; isolasi mandiri, dirawat, sembuh, dan meninggal. Pengelompokan dibagi menjadi 4 Cluster berdasarkan kriteria data kasus yang terkonfirmasi Covid-19 sebanyak 265 kasus. Cluster model mengarah kepada *cluster0* sebanyak 225 kasus, *cluster1* sebanyak 6 kasus, *cluster2* sebanyak 12 kasus dan *cluster3* sebanyak 22 kasus dengan total 265 kasus. Penelitian ini menghasilkan pengelompokan data pemantauan pasien Covid-19 pada Kelurahan Makasar yang memerlukan pemantauan khusus pada *cluster0*.

Kata kunci: Covid-19, K-Means, Pemantauan

PENDAHULUAN

Keluarga coronavirus adalah sekelompok besar virus yang dapat menyebabkan penyakit pada manusia dan hewan. Pada manusia biasanya menyebabkan infeksi saluran pernapasan mulai dari flu biasa hingga penyakit serius seperti Middle East Respiratory Syndrome (MERS) dan Severe Acute Respiratory Syndrome (SARS). Virus corona

jenis baru yang ditemukan pada manusia sejak kejadian luar biasa pada Desember 2019 di Wuhan, China kemudian diberi nama Severe Acute Respiratory Syndrome coronavirus 2 (SARS-COV2) dan menyebabkan penyakit Coronavirus 2019 (COVID-19) (WHO, n.d.-b).

Menurut (WHO, n.d.-b) seperti penyakit pernapasan lainnya, COVID-19 dapat menyebabkan gejala ringan, seperti pilek, sakit tenggorokan,

batuk, dan demam. Sekitar 80% kasus sembuh tanpa perawatan khusus. Sekitar 1 dari 6 orang dapat mengembangkan penyakit parah, seperti pneumonia atau kesulitan bernapas, yang biasanya berkembang secara bertahap, dan meskipun angka kematian untuk penyakit ini tetap rendah (sekitar 3%), orang tua dan orang dengan kondisi kondisi medis yang sudah ada sebelumnya (seperti diabetes, tekanan darah tinggi, penyakit jantung) biasanya lebih mudah untuk mengalami gejala kritis.

Kasus penyebaran Covid-19 di Indonesia terus meningkat, sehingga masyarakat diminta menjaga jarak untuk memutus mata rantai penyebaran virus tersebut. Kasus Covid-19 pertama dilaporkan di Indonesia pada 2 Maret 2020. Wilayah yang luas ini memungkinkan pengelompokan kasus berdasarkan wilayah di Indonesia (Nabila et al., 2021).

Menurut Kementerian Suku Dinas Kesehatan kasus jumlah terpapar di Indonesia pada bulan Maret 2022 mencapai hingga 6.012.828 kasus konfirmasi meningkat sebanyak 3.332 kasus (WHO, n.d.-a). Untuk di wilayah Kecamatan Makasar dengan jumlah pasien masih aktif 74 kasus, 70 pasien isolasi mandiri, 4 pasien di rawat dan 196 pasien meninggal akibat Covid-19.

Dinas Kesehatan sebelumnya membagikan peta penyebaran Covid-19 dengan merinci tiga kategori secara nasional. Belum ada perhitungan khusus yang dapat dibuat berdasarkan data yang dikumpulkan di setiap Wilayah Kecamatan untuk memantau pengelompokan kasus perkembangan penyebaran Covid-19.

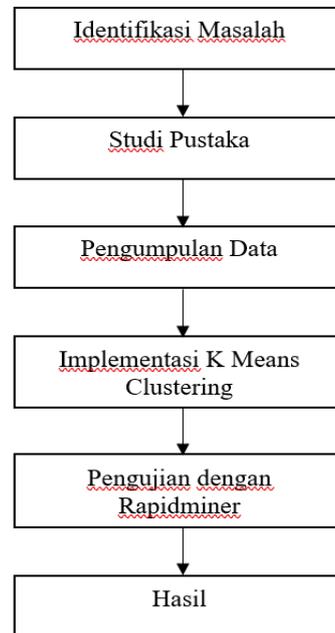
Dalam penelitian ini akan dilakukan perhitungan menggunakan Rapidminer dengan algoritma K-Means, dalam memantau kasus terkonfirmasi positif Covid-19 pada lima Kelurahan yaitu; Kel. Cipinang Melayu, Kel. Halim Perdana. Kel. Kebon Pala, Kel. Makasar, dan Kel. Pinang Ranti di Kecamatan Makasar.

Melalui penelitian terdahulu, dalam jurnal yang ditulis oleh (Solichin & Khairunnisa, 2020) Data Mining Clustering dapat digunakan untuk pengelompokan data Covid 19 di DKI Jakarta. Berdasarkan riset yang dilakukan, diperoleh kesimpulan bahwa data mining mampu mengelompokkan kasus Covid-19 dengan baik.

Dengan melihat data tersebut akan diketahui kelompok mana yang harus mendapatkan perhatian khusus seperti; isolasi mandiri, dirawat, sembuh, dan meninggal. Pengelompokan dibagi menjadi 4 Cluster berdasarkan kriteria data kasus yang terkonfirmasi Covid-19 sebanyak 265 kasus. Dengan demikian penelitian ini penulis ingin mengangkat judul "Penerapan K-Means Untuk Pengelompokan Pemantauan Kasus Terkonfirmasi Covid-19 Pada Puskesmas Kecamatan Makasar."

BAHAN DAN METODE

Tahapan penelitian merupakan langkah-langkah yang dilakukan oleh peneliti, berikut adalah langkah-langkah yang dilakukan oleh peneliti:



Gambar 1. Tahapan Penelitian

Metode yang digunakan adalah penelitian studi kasus. Adapun jenis data yang digunakan adalah data sekunder. Dimana data tersebut diambil dari Puskesmas Kecamatan Makasar yang berada di jalan Pusdiklat Depnaker No.4 RT 8/6 Kecamatan Makasar Jakarta Timur.

Data yang di dapatkan adalah data sekunder dalam bentuk excel yang terdiri dari 265 field. Field tersebut berisikan Kelurahan, RW, Pasien Isolasi Mandiri, Pasien Rawat, Pasien Sembuh, Pasien Meninggal dan Total Pasien. Data tersebut kemudian di olah menjadi dataset seperti berikut:

	KELURAHAN	RW	ISOLASI MANDIRI	DIRAWAT	SEMBUH	MENINGGAL	TOTAL PASIEN
3	CIPINANG MELAYU	1	0	0	102	1	0
4	CIPINANG MELAYU	2	3	2	211	5	221
5	CIPINANG MELAYU	3	1	0	274	2	277
6	CIPINANG MELAYU	4	0	0	147	2	149
7	CIPINANG MELAYU	5	8	1	504	5	518
8	CIPINANG MELAYU	6	5	0	482	5	492
9	CIPINANG MELAYU	7	11	0	234	4	249
10	CIPINANG MELAYU	8	3	0	226	2	231
11	CIPINANG MELAYU	9	2	0	286	4	292
12	CIPINANG MELAYU	10	5	0	304	2	311
13	CIPINANG MELAYU	11	0	0	184	2	186
14	CIPINANG MELAYU	12	1	0	345	4	350
15	CIPINANG MELAYU	13	0	0	241	6	247
16	CIPINANG MELAYU	1	0	0	102	1	103
17	CIPINANG MELAYU	2	0	0	216	5	221
18	CIPINANG MELAYU	3	0	0	275	2	277
19	CIPINANG MELAYU	4	0	0	147	2	149
20	CIPINANG MELAYU	5	7	1	505	5	518
21	CIPINANG MELAYU	6	4	0	485	5	494
22	CIPINANG MELAYU	7	1	0	244	4	249

Gambar 2. Tabel Sampel Dataset

1. Algoritma K-Means

Algoritma k-means merupakan metode untuk mengelompokkan data ke dalam cluster-cluster

berdasarkan kemiripan. Hal ini memungkinkan data yang terkait satu sama lain untuk dikelompokkan bersama, sedangkan data yang berbeda ditempatkan ke dalam cluster yang berbeda (Nurul Rohmawati, Sofi Defiyanti, 2015).

Menurut (Sari & Sukestiyarno, 2021) Analisis cluster merupakan teknik multivariat yang mempunyai tujuan utama untuk mengelompokkan objek-objek berdasarkan karakteristik yang dimilikinya.

K-Means adalah metode clustering yang mencoba mengelompokkan data ke dalam kelompok-kelompok dengan membaginya menjadi bagian-bagian yang lebih kecil. Cluster adalah kumpulan data serupa, berbeda satu sama lain dan berbeda dari data di cluster lain. Metode K-Means melakukan analisis grup yang terkait dengan partisi n objek k (Cluster) berdasarkan mean (Mean) terdekat (Fibriany, 2022).

K-Means merupakan salah satu algoritma Clustering yang masuk dalam kelompok Unsupervised Literacy yang digunakan untuk mengelompokkan data kedalam beberapa kelompok dengan sistem partisi. Algoritma ini menerima masukan berupa data tanpa marker kelas. Pada algoritma K-Means, komputer menerima data- data yang tidak diketahui kelasnya terlebih dahulu lalu mengelompokkannya. Input yang diterima ialah data dan jumlah kelompok (Cluster) yang diinginkan (Dwitri et al., 2020).

Algoritma K- Means menjelaskan bagaimana suatu algoritma dapat menyelesaikan suatu masalah dalam data yang dikelompokkan dengan tujuan akhir untuk dapat mengurangi kegagalan atau kecacatan secara berulang- ulang (Sianipar et al., 2020).

K-Means adalah metode pengelompokan data yang mengelompokkan data yang serupa menjadi satu. Metode ini membagi data ke dalam kelompok-kelompok berdasarkan kesamaan, sehingga data yang memiliki kesamaan karakteristik dikelompokkan menjadi satu dan data dengan karakteristik berbeda dikelompokkan secara terpisah (Yudi Agusta, 2007).

2. Data Mining

Data mining adalah bidang ilmiah yang membahas masalah pengambilan informasi dari database besar dengan menggabungkan statistik, pembelajaran mesin, visualisasi data, pengenalan pola, dan teknik database (Werdiningsih et al., 2020).

Data mining adalah analisis pada data untuk mencari hubungan yang jelas dan menyimpulkan belum diketahui sebelumnya dengan cara terkini agar mudah dipahami dan bermanfaat (Ardiansyah et al., 2020).

Menurut (Ni Luh Wiwik Sri Rahayu Ginantra, 2021) data mining merupakan proses pengambilan informasi yang tersembunyi dan informasi tersebut sebelumnya tidak dikenal. Proses data mining meliputi proses pendekatan secara teknis seperti clustering, klasifikasi dengan meliputi metode yang merupakan irisan Artificial Intelligence (AI), machine learning (ML) dan statistik. Data mining dapat menangani data dalam skala besar serta dapat memanfaatkan data pengalaman masa lampau untuk meningkatkan proses model pembelajarannya seperti pada penerapan klasifikasi.

3. Covid-19

Covid-19 adalah penyakit infeksi saluran pernapasan yang disebabkan oleh virus. Hal ini paling sering terlihat pada anak kecil dan bisa sangat serius. Meski terlihat biasa saja dengan gejala yang umumnya terlihat seperti flu, namun penyakit ini sangat mengkhawatirkan dan membutuhkan perawatan yang cukup intensif. Beberapa gejala yang ditemui yang disebabkan oleh virus ini antara lain kehilangan rasa, kehilangan penciuman, hidung tersumbat, diare, sakit kepala, sakit tenggorokan dan sebagainya. Gejala-gejala ini adalah tanda ringan yang dapat berkembang secara bertahap. Bahkan beberapa orang tidak mengalami gejala awal dan diumumkan terkena virus corona (Puryono & Handayani, 2021).

Wabah Covid-19 pertama kali diidentifikasi pada Desember 2019 di Kota Wuhan, Provinsi Hubei, China. Dibandingkan dengan virus pernapasan lainnya seperti SARS-CoV dan MERS-CoV, Covid-19 menyebar dengan cepat. Virus SARS membutuhkan waktu sekitar empat bulan untuk menginfeksi 1000 orang, sedangkan MERS membutuhkan waktu sekitar dua setengah tahun untuk menginfeksi orang sebanyak ini. Hanya dalam

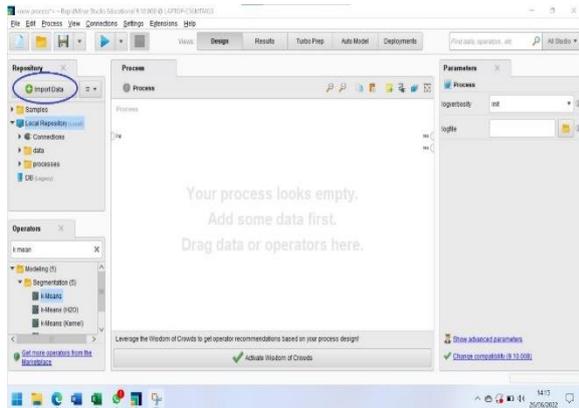
48 hari, Covid-19 bisa mencapai angka itu (Isfahani et al., 2020).

Virus Corona dapat menular dari manusia ke manusia melalui droplet (tetesan sangat kecil) yang dihasilkan ketika seseorang batuk, bersin, atau berbicara. Beberapa penelitian juga membuktikan bahwa virus corona berlangsung beberapa jam di benda yang disentuh pasien. Jika ada seseorang yang menyentuh benda yang terkontaminasi dengan tetesan dan sentuhan mata, hidung atau mulut, maka orang tersebut bisa terinfeksi Covid19. Salah satu cara untuk membantu mencegah penularan virus ini adalah dengan menjaga jarak fisik, memakai masker, dan menjaga kesehatan fisik dan non fisik (Solichin & Khairunnisa, 2020).

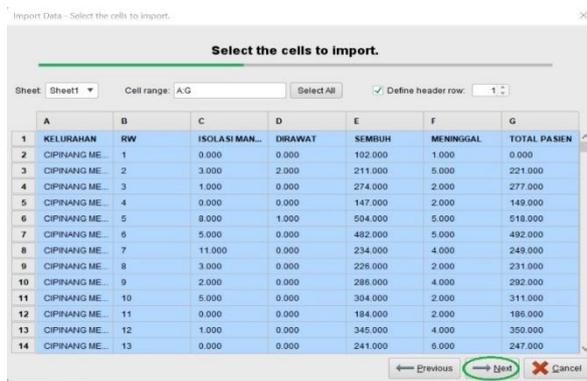
HASIL DAN PEMBAHASAN

1. Import Data

Sebelum menerapkan algoritma K-Means maka peneliti melakukan import dataset terlebih dahulu. Adapun dataset yang digunakan oleh peneliti yaitu dalam ormat excel.



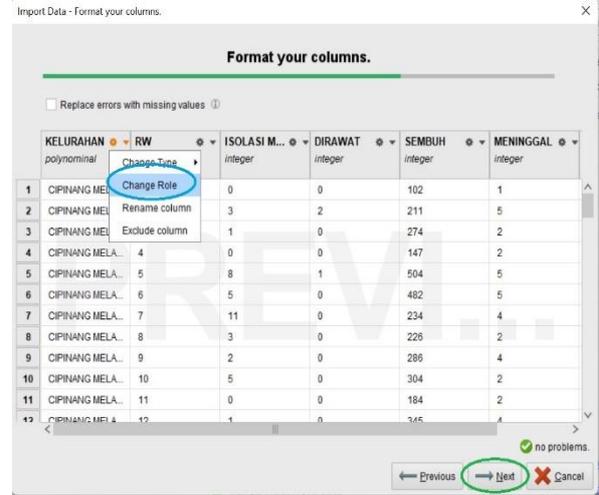
Gambar 3. Import Data



Gambar 4. Kolom dataset yang akan di import ke program RapidMiner

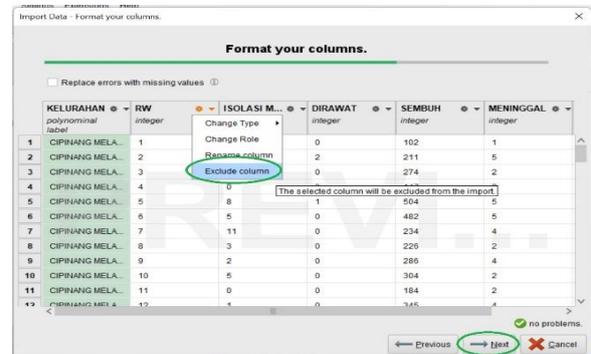
Pada tahap import dataset, terdapat beberapa atribut yang digunakan oleh peneliti,

adapun atribut yang digunakan sebagai label ialah atribut kelurahan.

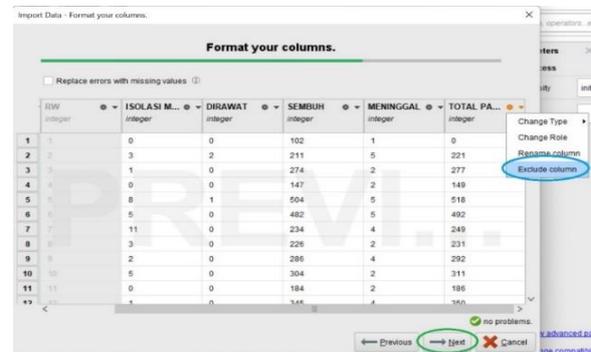


Gambar 5. Penetapan atribut label

Gambar 5 menunjukkan proses terkait pemilihan label pada salah satu atribut yang ada pada dataset, pemilihan label dilakukan dengan mengubah type field. Pada tahapan import dataset dan penggunaan dataset sebagai bahan penelitian, peneliti juga melakukan proses pemilihan data atau filter terhadap beberapa atribut yang tidak digunakan dalam penelitian ini diantaranya ialah atribut RW dan Total Pasien.



Gambar 6. Exclude column RW



Gambar 7. Exclude column Total Pasien

Adapun dari hasil proses penyeleksian atribut yang sudah dilakukan sebelumnya serta proses filterisasi

terkait atribut yang tidak digunakan oleh peneliti, maka akan ditampilkan pada Gambar 8 berikut:

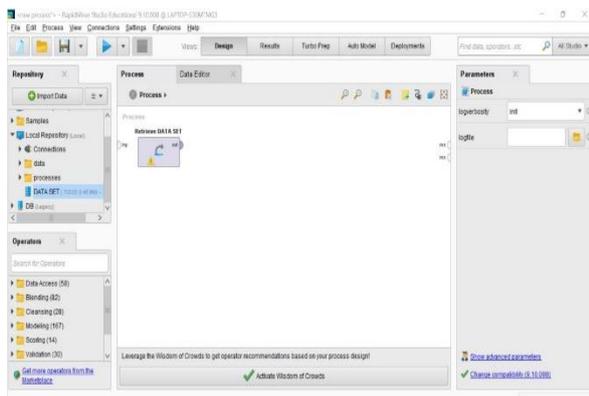
Row No.	KELURAHAN	ISLAMSAL	DIRAWAT	SEMBUN	MENINGGAL
1	OPPIWANG M.	0	0	702	1
2	OPPIWANG M.	3	2	211	5
3	OPPIWANG M.	1	0	274	2
4	OPPIWANG M.	0	0	147	2
5	OPPIWANG M.	0	1	504	5
6	OPPIWANG M.	5	0	482	5
7	OPPIWANG M.	11	0	234	4
8	OPPIWANG M.	3	0	255	2
9	OPPIWANG M.	2	0	269	4
10	OPPIWANG M.	5	0	304	2
11	OPPIWANG M.	0	0	194	2
12	OPPIWANG M.	1	0	345	4
13	OPPIWANG M.	0	0	241	6

Gambar 8. Hasil dataset yang telah tersimpan dalam RapidMiner

Dari Gambar 8 menunjukkan hasil dataset yang telah selesai proses setting input kedalam RapidMiner, terdapat lima atribut yang digunakan oleh peneliti, adapun salah satunya merupakan label.

1. Proses Clustering pada RapidMiner

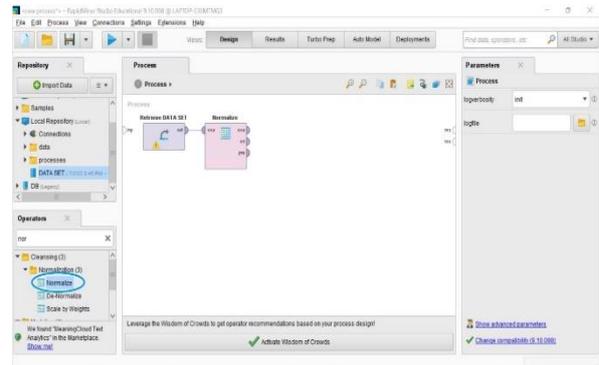
Setelah dataset sudah tersimpan. peneliti memulai tahapan clustering pada program RapidMiner. Adapun clustering dilakukan menggunakan dataset yang telah diimport sebelumnya sebagai berikut:



Gambar 9. Memasukkan dataset ke dalam process layout untuk proses clustering

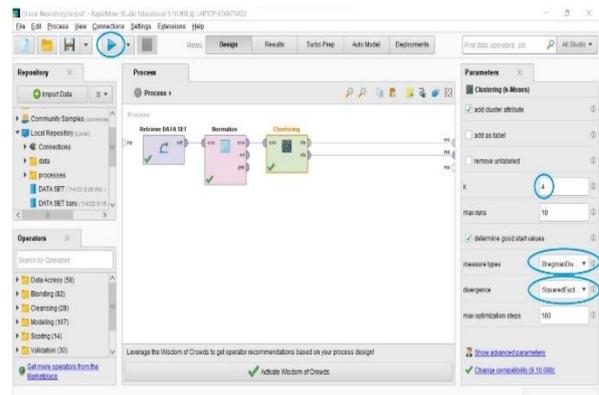
Dari gambar 9 diatas, dataset telah masuk kedalam layout proses clustering dan selanjutnya dilakukannya normalize pada kolom operators dan drag pada process layout untuk menormalisasi dataset yang telah disimpan. Adapun Fungsi dari normalisasi ini adalah untuk mendapatkan nilai akurasi yang lebih tinggi, mengurangi hal-hal yang

tidak diperlukan dalam data set, dan mengurangi kesamaan data dalam data set.



Gambar 10. Normalisasi pada dataset

Data yang telah di normalisasi selanjutnya akan mulai di proses menggunakan k-means. Pada process layout untuk memulai proses perhitungan menggunakan k-means. Proses modelling dengan K-Means ini menggunakan jumlah K 4 karena pada penelitian ini akan ada 4 cluster yang digunakan.



Gambar 11. Clustering pada dataset

2. Hasil Clustering

Setelah data berhasil di proses, maka diperoleh hasil clustering data seperti berikut.

Row No.	KELURAHAN	cluster	ISLAMSAL	DIRAWAT	SEMBUN	MENINGGAL
1	OPPIWANG M.	cluster_0	-0.804	-0.284	-1.335	-0.815
2	OPPIWANG M.	cluster_1	0.501	3.783	-0.635	0.391
3	OPPIWANG M.	cluster_0	-0.358	-0.284	-0.231	-0.513
4	OPPIWANG M.	cluster_0	-0.804	-0.284	-1.546	-0.513
5	OPPIWANG M.	cluster_0	2.875	1.700	1.245	0.391
6	OPPIWANG M.	cluster_0	1.370	-0.284	1.165	0.391
7	OPPIWANG M.	cluster_0	3.980	-0.284	-0.488	0.090
8	OPPIWANG M.	cluster_0	0.501	-0.284	-0.539	-0.513
9	OPPIWANG M.	cluster_0	0.956	0.284	-0.154	0.090
10	OPPIWANG M.	cluster_0	1.370	-0.284	-0.838	-0.513
11	OPPIWANG M.	cluster_0	-0.804	-0.284	-0.809	-0.513
12	OPPIWANG M.	cluster_0	-0.358	-0.284	0.225	0.090
13	OPPIWANG M.	cluster_0	-0.804	-0.284	-0.443	0.693

Gambar 12. Hasil Clustering

Terkait Gambar 12 diatas, berikut adalah cluster model yang diperoleh adalah :

Tabel 1. Hasil Clusterisasi

Cluster Model		
No	Cluster	Item
1	Cluster 0	225
2	Cluster 1	6
3	Cluster 2	12
4	Cluster 3	22
Total Number of item		265

Tabel 1 Menjelaskan untuk cluster 0 sebanyak 225 kasus, cluster 1 sebanyak 6 kasus, cluster 2 sebanyak 12 kasus dan cluster 3 sebanyak 22 kasus dengan total 265 kasus.

KESIMPULAN

Kesimpulan yang dapat diperoleh dari penelitian ini adalah Cluster model mengarah kepada *cluster0* sebanyak 225 kasus, *cluster1* sebanyak 6 kasus, *cluster2* sebanyak 12 kasus dan *cluster3* sebanyak 22 kasus dengan total 265 kasus. Penelitian ini menghasilkan pengelompokan data pemantauan pasien Covid-19 pada Kelurahan Makasar yang memerlukan pemantauan khusus pada *cluster 0*.

REFERENSI

Ardiansyah, A. H., Nugroho, W., Alfiyah, N. H., Handoko, R. A., & Bakhtiar, M. A. (2020). Penerapan Data Mining Menggunakan Metode Clustering untuk Menentukan Status Provinsi di Indonesia 2020. *Seminar Nasional Inovasi Teknologi*, 4(3), 329-333.

Dwitri, N., Tampubolon, J. A., Prayoga, S., R.H Zer, F. I., & Hartama, D. (2020). Penerapan Algoritma K-Means Dalam Menentukan Tingkat Penyebaran Pandemi Covid-19 Di Indonesia. *Jurnal Teknologi Informasi*, 4(1), 128-132. <https://doi.org/10.36294/jurti.v4i1.1266>

Fibriany, F. W. (2022). *Penerapan Metode K-Means Dalam Ujian Nasional*. 6(1), 26-31. <https://doi.org/10.52362/jisamar.v6i1.670>

Isfahani, F. Al, Nugraha, F., Mubarak, R., & Rahmatulloh, A. (2020). Implementasi Web Service untuk Aplikasi Pemantau Coronavirus Disease 2019 (COVID-19). *Innovation in Research of Informatics (INNOVATICS)*, 2(1), 33-39. <https://doi.org/10.37058/innovatics.v2i1.1696>

Nabila, Z., Rahman Isnain, A., & Abidin, Z. (2021). Analisis Data Mining Untuk Clustering Kasus Covid-19 Di Provinsi Lampung Dengan Algoritma K-Means. *Jurnal Teknologi Dan Sistem Informasi (JTSI)*, 2(2), 100. <http://jim.teknokrat.ac.id/index.php/JTSI>

Ni Luh Wiwik Sri Rahayu Ginantra, D. (2021). *Data Mining dan Penerapan Algoritma* (R. Watrianthos & J. Simarmata (eds.)). Yayasan Kita Menulis. https://books.google.co.id/books?hl=id&lr=&id=v0gtEAAAQBAJ&oi=fnd&pg=PA5&dq=pengertian+data+mining&ots=RiY3vvekkI&sig=GNMKFdf_yTuKk57ip6IkhSBysOc&redir_esc=y#v=onepage&q=pengertian+data+mining&f=false

Nurul Rohmawati, Sofi Defiyanti, M. J. (2015). Implementasi Algoritma K-Means Dalam Pengklasteran Mahasiswa Pelamar Beasiswa. *Jitter 2015, I(2)*, 62-68.

Puryono, D. A., & Handayani, D. (2021). Sistem Informasi Pelacakan dan Pemantauan Covid-19 Berbasis Internet Of Things (IoT). *INOVTEK Polbeng - Seri Informatika*, 6(2), 306. <https://doi.org/10.35314/isi.v6i2.2150>

Sari, D. N. P., & Sukestiyarno, Y. L. (2021). Analisis Cluster dengan Metode K-Means pada Persebaran Kasus Covid-19 Berdasarkan Provinsi di Indonesia. *PRISMA, Prosiding Seminar Nasional Matematika*, 4, 602-610. <https://journal.unnes.ac.id/sju/index.php/prisma/>

Sianipar, K. D. R., Siahaan, S. W., Siregar, M., R.H Zer, F. I., & Hartama, D. (2020). Penerapan Algoritma K-Means Dalam Menentukan Tingkat Kepuasan Pembelajaran Online Pada Masa Pandemi Covid-19. *Jurnal Teknologi Informasi*, 4(1), 101-105. <https://doi.org/10.36294/jurti.v4i1.1258>

Solichin, A., & Khairunnisa, K. (2020). Klasterisasi Persebaran Virus Corona (Covid-19) Di DKI Jakarta Menggunakan Metode K-Means. *Fountain of Informatics Journal*, 5(2), 52. <https://doi.org/10.21111/fij.v5i2.4905>

Werdiningsih, I., Nuqoba, B., & Muhammadun. (2020). *Data Mining Menggunakan Android, Weka dan SPSS*. Airlangga University Press. https://books.google.co.id/books?hl=id&lr=&id=xEwCEAAAQBAJ&oi=fnd&pg=PP1&dq=pengertian+data+mining&ots=z3TXs0F2oK&sig=WlvGc4UTYOhNLMxqmvJ3n0cifpc&redir_esc=y#v=onepage&q=pengertian+data+mining&f=false

WHO. (n.d.-a). *Situasi Covid-19*. <https://covid19.go.id/artikel/2022/03/31/situasi-covid-19-di-indonesia-update-31-maret-2022>

WHO. (n.d.-b). *Tentang Covid-19*. <https://covid19.go.id/tentang-covid-19>

Yudi Agusta. (2007). K-Means - Penerapan, Permasalahan dan Metode Terkait. *Jurnal Sistem Dan Informatika*, 3(Februari), 47-60.