

PENERAPAN METODE CLUSTERING K-MENS UNTUK MEMBANTU MENTUKAN TINGKATAN STATUS DAERAH DAMPAK COVID 19

Eki Juliana¹, Yuliana², dan Vivi Nur Aleyda³

^{1,2,3}Jurusan Sistem Informasi Fakultas Ilmu Komputer UNDA University
Ekijuliana6@gmail.com¹, yuliana08016@gmail.com², vnurlyda@gmail.com³

Abstrak

Penyakit menular yang dapat menyerang daerah pernafasan telah dideteksi masuk di Indonesia sekitar awal tahun 2020 yaitu Corona Virus Disease-2019 atau Covid-19. Penularan virus ini cukup ringan dengan cairan yang dikeluarkan seseorang dari hidung, bersin, dan batuk dari pasien yang terjangkit, maka dari itu penularannya sangat cepat. Perkembangan klaster covid-19 ini dapat diakibatkan dari faktor manusianya yang tidak mematuhi peraturan protokol kesehatan. Peristiwa ini dapat diketahui dari jumlah kasus Covid-19 tertinggi antar per provinsinya yaitu Jakarta dengan kasus positif 9.128 jiwa perhari dengan kesembuhan 11.857 jiwa perhari, sedangkan penularan Covid-19 dengan kasus yang terendah penyebarannya adalah Maluku terdapat 69 jiwa kasus positif perhari dan kesembuhannya 23 jiwa perhari. Penelitian ini menerapkan algoritma k-means clustering dengan batuan tool weka agar dapat melakukan pengklasifikasian pengelolaan kasus Covid-19 untuk menentukan pengelompokan status zona pada setiap provinsinya dengan menggunakan 3 cluster yaitu rendah, Sedang dan Tinggi. Sehingga diperoleh hasil data 4 provinsi zona Merah, 11 Provinsi zona Kuning, dan 19 provinsi zona hijau.

Kata Kunci: *Algoritma K-means, Clustering, Covid-19, Status Zona, tool weka*

Abstract

Infectious diseases that can attack the respiratory area have been detected in Indonesia around the beginning of 2020, namely Corona Virus Disease-2019 or Covid-19. The transmission of this virus is quite light, with fluids released by a person from the nose, sneezing, and coughing from infected patients, therefore transmission is very fast. The development of this covid-19 cluster can be caused by human factors that do not comply with health protocol regulations. This event can be seen from the highest number of Covid-19 cases per province, namely Jakarta with 9,128 positive cases per day with healing 11,857 people per day, while the transmission of Covid-19 with the lowest case spread was Maluku, there were 69 positive cases per day and 23 people recovered. per day. This study applies the k-means clustering algorithm with the Weka tool in order to classify the management of Covid-19 cases to determine the grouping of zone status in each province using 3 clusters, namely positive patients, recovered and dead patients from each province in Indonesia. So that the data obtained from 4 provinces in the red zone, 11 provinces in the yellow zone, and 19 provinces in the green zone.

Keywords: *Algoritma K-means, Clustering, Covid-19, Status Zona, tool weka*

1. Pendahuluan

Pada awal tahun 2020 terdapat sebuah virus

yang membuat semua orang didunia khususnya di Wuhan, China menjadi heboh dan ketakutan Virus tersebut yaitu

CoronaVirus Disease-2019 atau Covid-19. Virus tersebut baru diidentifikasi sebagai SARS-CoV-2, virus ini pertama kali terdeteksi di daerah wuhan, China pada Desember 2019 lalu kemudian mulai menyebar keseluruh negara-negara lain. Penyebaran virus ini sangat cepat sehingga masyarakat Indonesia menimbulkan ketakutan dan keresahan pada virus tersebut karena penyebarannya dapat melalui cairan yang dikeluarkan seseorang dari hidung, bersin, dan batuk dari pasien yang terjangkit. Kemudian pemerintah langsung melakukan pemberlakuan kebijakan-kebijakan yaitu salah satunya Pemberlakuan Pembatasan Kegiatan Masyarakat (PPKM) yang saat ini sedang diterapkan.

Semenjak Indonesia terdeteksi Covid-19 pada awal tahun 2020, kasusnya semakin meningkat dengan cepat. Provinsi yang pertama kali terdampak adalah Jakarta dan yang paling tinggi tingkat kasusnya diantara provinsi lainnya yang ada di Indonesia sampai saat ini. Pemantauan data pada tanggal 18 Juli 2021 dengan kasus positif 9.128 jiwa perhari dengan kesembuhan 11.857 jiwa perhari, sedangkan penularan Covid-19 dengan kasus yang terendah penyebarannya adalah Maluku terdapat 69 jiwa kasus positif perhari dan kesembuhannya 23 jiwa perhari. Pemantauan data pada tanggal 18 Juli 2021 di seluruh provinsi Indonesia dengan kasus Positif mencapai 44.721 jiwa, Sembuh mencapai 29.264 jiwa dan Kematian mencapai 1.093 jiwa atau sebanding 78,6 %. Sehingga masyarakat harus membentuk imun tubuh yang baik agar kita dianjurkan untuk mengonsumsi vitamin C, buah-buahan dan sayuran yang cukup serta menjaga pola tidur yang baik dan berjemur dipagi hari. Karena sampai saat ini banyak sekali orang yang terjangkit virus ini tetapi tidak memiliki gejala seperti batuk ataupun panas yang tinggi, dan sekarang pun belum ditemukan obat untuk virus ini, walaupun dunia sudah membuat sebuah vaksin untuk virus ini tetapi belum menutup kemungkinan seseorang yang sudah divaksin tersebut tidak dapat lagi terjangkit virus ini.

Dalam penelitian ini bertujuan untuk

melakukan clustering pada data penularan Covid-19 di Provinsi seluruh Indonesia. Sehingga dapat membantu untuk membentuk cluster zona penyebaran covid 19. Clustering ini adalah metode data mining yang dapat digunakan dalam melakukan pengategorian data yang serupa atau sama sehingga menjadi cluster. Proses penggunaan algoritma clustering yaitu dengan mencari data pada seluruh set/bagian dan dijadikan subkelompok atau bagian kelompok. Adapun harapan dari penelitian ini agar dapat membantu pemerintah dalam mengelompokkan zona penyebaran covid-19 dan memberikan keterangan status zona dari angka penyebaran covid tertinggi, sedang dan rendah.

2. Metode

A. Tahap Pengumpulan Data

Data yang digunakan dalam penelitian ini di peroleh dari kasus penyebaran covid-19 pada tanggal 18 Juli 2021 dari 34 provinsi yang ada di Indonesia yang di ambil langsung dari situs web resmi Data Riset dan Teknologi Covid-19 Indonesia (ristekbrin.go.id)

B. Pembagian Zona di Indonesia pada pandemi covid-19

Menurut Gugus Tugas Percepatan Penanganan Covid-19, kriteria zonasi wilayah berdasarkan warna tersebut merupakan indikator kategori risiko Covid-19 yang dilihat dari tingkat penularannya ataupun penyebarannya

Berikut ini Beberapa pembagian zona di Indonesia :

- a) Zona Hijau (Tidak Terdampak atau sudah tidak terinfeksi virus covid 19)
- b) Zona Kuning (Resiko penyebaran covid rendah)
- c) Zona Orange (Resiko penyebaran covid sedang)
- d) Zona Merah (Resiko penyebaran covid Tinggi)

C. Metode Clustering

Clustering adalah memisahkan data ke dalam kelompok-kelompok yang memiliki obyek dengan karakteristik yang serupa. Clustering ini merupakan awalan dari sebuah data mining.[3]

D. Algoritma K-Means

Salah satu algoritma Clustering adalah K-Means yang digunakan untuk mengelompokkan data menjadi beberapa kelompok dengan beberapa Cluster.[1] Data – data dipilih menjadi beberapa kelompok dengan kriteria yang telah ditentukan lalu dikumpulkan menjadi satu dalam sebuah Cluster. Dimana setiap Cluster memiliki titik pusat yang disebut Centroid. Berikut adalah tahapan-tahapan untuk melakukan optimasi menggunakan algoritma K-Means[1]

- Pilihlah jumlah Cluster (k) yang diinginkan pada dataset
- Tentukan titik pusat (Centroid) secara acak/random pada tahap awal
- Hitunglah jarak terdekat setiap data dengan Centroid. Untuk menghitung jarak terdekat dengan Centroid adalah Euclidean distance
- Hitung kembali pusat Cluster dengan keanggotaan Cluster yang sekarang. Pusat Cluster adalah rata – rata dari semua data dalam sebuah Cluster.
- Kelompokan hasil cluster dan berikan tanda warna

E. Waikato Environment for Knowledge Analysis (WEKA)

WEKA adalah sebuah paket tools machine learning praktis. WEKA merupakan singkatan dari Waikato Environment for Knowledge Analysis, yang dibuat di Universitas Waikato, New Zealand untuk penelitian, pendidikan dan berbagai aplikasi. WEKA mampu menyelesaikan masalah-masalah data mining di dunia nyata, khususnya klasifikasi yang mendasari pendekatan-pendekatan machine learning. Perangkat lunak ini ditulis dalam hirarki class Java dengan metode berorientasi objek dan dapat berjalan hampir di semua platform.[4]

3. Hasil dan Pembahasan

A. Data Pengujian

Penggunaan data yang di gunakan dalam penelitian ini berupa dataset kasus penularan covid 19 yang terdiri dari 7 atribut yaitu nama provinsi, kasus positif, kasus positif hari ini, kasus sembuh, kasus sembuh hari ini, kasus meninggal, kasus meninggal hari ini dan 34

data Provinsi Indonesia.

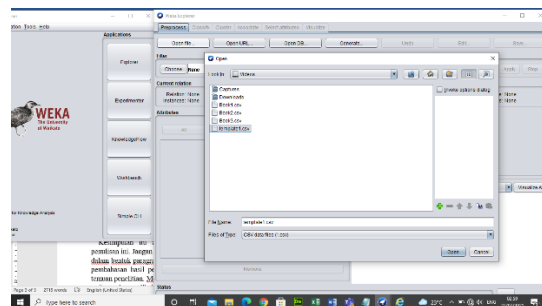
Table 1
Dataset Kasus pasien covid 19 disetiap provinsi pada 18 juli 2021

Provinsi Asal	Kasus	Kasus Hari Ini	Sembuh	Sembuh Hari Ini	Kem atian	Kematian Hari ini
Aceh	20949	63	16193	86	923	8
Bali	61179	944	52293	465	1769	20
Banten	87718	3139	55381	310	1554	9
Bangka Belitung	26777	309	22371	271	455	10
Bengkulu	14003	468	10569	232	256	1
DI Yogyakarta	92084	2119	62428	780	2382	48
DKI Jakarta	746312	9128	627462	11857	10167	158

Jambi	16441	442	12809	133	321	6
Jawa Barat	516591	7777	390203	3895	7340	91
Jawa Tengah	321763	4699	245593	3420	14397	179
Jawa Timur	239168	5726	178766	3053	15836	328
Kalimantan Barat	21149	358	15987	249	501	12
Kalimantan Timur	96564	1514	78479	708	2392	51
Kalimantan Tengah	30347	305	21950	160	584	11
Kalimantan Selatan	40029	299	35531	103	1141	6
Kalimantan Utara	16286	284	13123	172	243	3
Kepulauan Riau	36581	928	27657	347	826	31
Nusa Tenggara Barat	16331	333	13337	210	533	2
Sumatera Selatan	36945	556	29430	300	1707	15
Sumatera Barat	61350	641	51508	316	1347	0
Sulawesi Utara	19720	306	15957	89	607	6
Sumatera Utara	44535	814	35379	217	1319	14
Sulawesi Tenggara	13817	159	11336	126	284	6
Sulawesi Selatan	72239	434	64778	349	1094	17
Sulawesi Tengah	16792	251	13677	72	473	7
Lampung	28009	677	20791	176	1253	11
Riau	82107	847	72081	639	2190	28
Maluku Utara	8204	84	5270	90	179	6
Maluku	12127	69	7913	23	187	0
Papua Barat	15902	156	11626	147	235	3
Papua	23385	135	11846	67	216	3
Sulawesi Barat	6919	90	5841	30	137	1
Nusa Tenggara Timur	28363	565	18276	141	537	2
Gorontalo	6790	102	5817	31	197	0

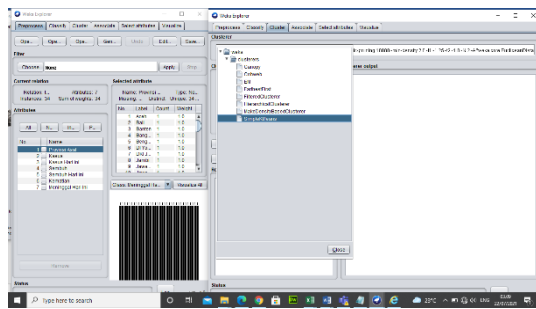
B. Implemtasi Pada Tool Weka

- Pertama –tama Explorer Dataset kedalam tool Weka dalam format csv atau xls



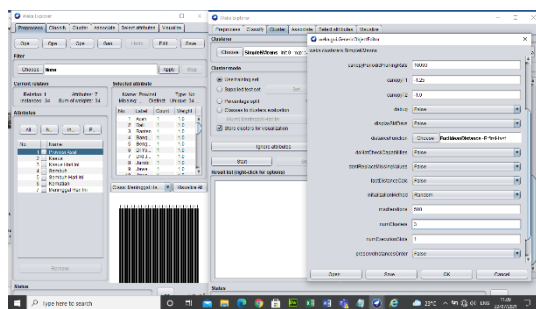
Gambar 1. Ekspor Dataset

2. Pilih Algoritma cluster kemudian tentukan metodenya yaitu SimpleKmeans



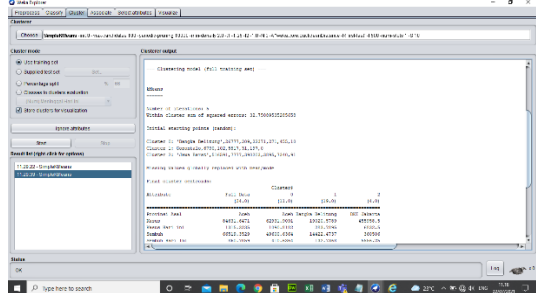
Gambar 2. Pemilihan Metode

3. Selanjutnya kita masukan jumlah cluster yang sudah kita tentukan sebelumnya dengan cara double klik pada simpleKMean



Gambar 3. Penentuan Jumlah Kluster

4. akan diperoleh hasil seperti berikut ini



Gambar 4. Tampilan Hasil

C. Hasil

Dengan Menggunakan Tool Weka hasil dari Perhitungannya Secara otomatis langsung di proses oleh komputer tanpa harus menghitung secara manual lagi, berikut ini adalah hasil dari perhitungan metode Kmeans

Final cluster centroids:				
Attribute	Full Data (34.0)	Cluster# 0 (11.0)	1 (19.0)	2 (4.0)
Provinsi Asal	Aceh	Aceh	Bangka Belitung	DKI Jakarta
Kerus	84631.6471	62931.9091	19020.5789	455958.5
Kerus Hari Ini	1315.3235	1090.8182	203.7895	6832.5
Sembuh	46819.3529	49600.6364	14422.4737	360806
Sembuh Hari Ini	840.7059	410.6264	152.7368	5556.25
Kematian	2164.1765	1591.1818	438.8947	11935
Meninggal Hari Ini	32.1471	21.9091	5.0526	189

Gambar 5. Hasil Penjumlahan

- Initial starting poin (Random) merupakan titik awal centroids yang di mana apa bila kita menghitung dengan cara manual kita menentukannya sendiri secara acak karena kita menggunakan tool weka maka secara otomatis akan terpilih dengan sendirinya
 - Final Cluster Centroids adalah hitungan pusat rata rata Cluster dengan keanggotaan Cluster yang sekarang.
- Terdapat 3 pembagian hasil klastering dengan rincian Cluster 0 terdiri dari 11hasil data, cluster 1 terdiri dari 19 hasil data, cluster 2 terdiri dari 4 hasil data.

D. Pengkelompokan zona

Zona di kelompokan berdasarkan hasil pembagian dari perhitungan clustering. Terdapat 3 cluster untuk menentukan status zona yang telah ditentukan dan di peroleh hasil cluster 0, cluster 1 dan cluster 2. Berdasarkan hasil yang kita peroleh melalui perhitungan dengan batuan tool weka. Cluster 2 dikategorikan sebagai zona merah karena cluster tersebut termasuk status zona penyebaran covid 19 tingkat tinggi, cluster 0 bewarna orane melambangkan bawah cluster tersebut termasuk ke status zona penyebaran covid-19 tingkat sedang dan cluster 1 bewarna kuning melambangkan penyebaran covid tingkat rendah.

Provinsi Asal	Cluster
---------------	---------

Aceh	cluster0
Bali	cluster0
Banten	cluster0
'Bangka Belitung'	cluster1
Bengkulu	cluster1
'DI Yogyakarta'	cluster0
DKI Jakarta'	cluster2
Jambi	cluster1
'Jawa Barat'	cluster2
'Jawa Tengah'	cluster2
'Jawa Timur'	cluster2
'Kalimantan Barat'	cluster1
'Kalimantan Timur'	cluster0
'Kalimantan Tengah'	cluster1
'Kalimantan Selatan'	cluster1
'Kalimantan Utara'	cluster1
'Kepulauan Riau'	cluster0
'Nusa Tenggara Barat'	cluster1
'Sumatera Selatan'	cluster0
'Sumatera Barat'	cluster0
'Sulawesi Utara'	cluster1
'Sumatera Utara'	cluster0
'Sulawesi Tenggara'	cluster1
'Sulawesi Selatan'	cluster0
'Sulawesi Tengah'	cluster1
Lampung	cluster1
Riau	cluster0
'Maluku Utara'	cluster1
Maluku	cluster1
'Papua Barat'	cluster1
Papua	cluster1
'Sulawesi Barat'	cluster1
'Nusa Tenggara Timur'	cluster1
Gorontalo	cluster1

dikaji sehingga dapat di kembangkan dalam membatu penanganan kasus data covid-19

Daftar Pustaka

- N. Dwitri, J. A. Tampubolon, S. Prayoga, P.P.P.A.N.W. F. Ilmi R.H Zer, D Hartama (2020) "Penerapan Algoritma K-Means Dalam Menentukan Tingkat Penyebaran Pandemi Covid-19 Di Indonesia,"
- E. Bu'ulolo, B. Purba (2021) "Algoritma Clustering Untuk Membentuk Cluster Zona Penyebaran Covid-19" Digital Zone : Jurnal Teknologi Informasi & Komunikasi
- A. H. Ardiansyah, W. Nugroho, N. H. Alfiyah, R. A. Handoko, M. A. Bakhtiar (2020) "Penerapan Data Mining Menggunakan Metode Clustering Untuk Menentukan Status Provinsi Di Indonesia 2020" Seminar Nasional Inovasi Teknologi.
- S. Pujiono, A. Amborowati, M. Suyanto (2013) "Analisis Kepuasan Publik Menggunakan Weka Dalam Mewujudkan Good Governance Di Kota Yogyakarta" Jurnal Dasi

5. KESIMPULAN DAN SARAN

Berdasarkan penelitian yang sudah kita lakukan di atas maka dapat di Tarik kesimpulan bahwasanya algoritma clustering KMeans dapat di terapkan dalam pengeolaan kasus data covid-19 dalam menentukan Tingkatan Status zona penyebaran covid 19 di provinsi indonesia, dengan batuan tool weka maka proses perhitungannya lebih mudah dalam membatu mendapatkan hasil di penelitian ini. Terdapat 4 provinsi yang termasuk ke zona merah , 11 provinsi zona orange dan 19 provinsi zona kuning.

Tentunya dalam penelitian ini masi banyak kekurangan sehingga penelitian ini masi perlu