

Segmentasi Wilayah Terdampak Pandemi Covid-19 Berdasarkan Suhu Dan Kelembaban Menggunakan Single Linkage Clustering

Lutfi Ali Muharom¹, Hardian Oktavianto^{1*}, Zainul Arifin¹

¹Teknik Informatika; Universitas Muhammadiyah Jember; Jl. Karimata 49 Jember, 0331-336728/0331-337957; e-mail: lutfi.muharom@unmujiember.ac.id hardian@unmujiember.ac.id zainul.arifin@unmujiember.ac.id

* Korespondensi: Hardian Oktavianto e-mail: hardian@unmujiember.ac.id

Diterima: 18 Maret 2023 ; Review:6 April 2023; Disetujui: 18 April 2023

Cara sitasi: Luffi Ali Muharom, Hardian Oktavianto, Zainul Arifin. 2021. Segmentasi Wilayah Terdampak Pandemi Covid-19 Berdasarkan Suhu Dan Kelembaban Menggunakan Single Linkage Clustering. Bina Insani ICT Journal. Vol 9(2): 162-173.

Abstrak: Tingginya proporsi antara jumlah orang yang terinfeksi dan jumlah kematian COVID-19 di Indonesia yang menempati urutan kedua setelah Italia, yaitu 1.285 infeksi dengan 114 kematian atau 11,2%, maka perlu dikaji karakteristik penyebaran COVID-19 di wilayah ini. Pada penelitian ini akan menerapkan algoritma single linkage clustering untuk menguji sejauh mana hasil cluster yang diperoleh untuk menunjukkan segmentasi wilayah terdampak pandemi yang dihubungkan dengan faktor suhu dan kelembaban udara. Hasil proses clustering yang telah dilakukan, dengan membatasi pembentukan cluster sejumlah 3 buah, maka diperoleh pada cluster pertama terdapat 15 propinsi, pada cluster kedua terdapat 16 propinsi, dan pada cluster ketiga mempunyai 3 propinsi. Untuk atribut Terkonfirmasi, cluster 1 mempunyai nilai sebesar 8830.6, pada cluster 2 sebesar 6463.2, dan cluster 3 sebesar 70179. Untuk atribut Suhu, cluster 1 mempunyai nilai 26.7, cluster 2 mempunyai nilai 27.5, dan cluster 3 mempunyai nilai 28.3. Untuk atribut Kecepatan Angin secara berurutan cluster 1, cluster 2, dan cluster 3 mempunyai nilai masing - masing adalah 2, 2.7, dan 2.8. Dan untuk atribut Kelembaban mempunyai nilai masing - masing 82.6, 77.9, dan 73.1. Berdasarkan hasil uji tersebut, maka profiling cluster yang dapat disimpulkan adalah, kejadian atau kasus covid-19 berbanding lurus dengan suhu dan kecepatan angin namun berbanding terbalik dengan kelembaban udara, semakin tinggi angka kasus covid-19 maka terjadi pada propinsi dengan suhu dan kecepatan angin tinggi, akan tetapi berbanding terbalik dengan tingkat kelembaban udara. Hasil evaluasi cluster menggunakan DBI menghasilkan nilai 0,31412426 yang menunjukkan bahwa hasil cluster dapat dikatakan optimal.

Kata kunci: clustering hierarki, covid-19, segmentasi

Abstract: Based on the proportion between the number of infected people and the number of COVID-19 deaths, Indonesia which ranks second after Italy has 1,285 infections with 114 deaths, namely 11.2%. In this study, a hierarchical algorithm will be applied to test the extent to which the cluster results obtained show segmentation of areas affected by a pandemic associated with temperature and humidity factors. By limiting the formation of clusters to 3, it is obtained that in the first cluster there are 15 provinces, in the second cluster there are 16 provinces, and in the third cluster there are 3 provinces. For the Confirmed attribute, cluster 1 has a value of 8830.6, in cluster 2 it is 6463.2, and cluster 3 is 70179. For the Temperature attribute, cluster 1 has a value of 26.7, cluster 2 has a value of 27.5, and cluster 3 has a value of 28.3. For the Wind Speed attribute sequentially cluster 1, cluster 2, and cluster 3 have respective values of 2, 2.7, and 2.8. And for the Humidity attribute it has a value of 82.6, 77.9 and 73.1 respectively. Based on the test results, the cluster profiling that can be concluded is, the incidence or case of covid-19 is directly proportional to temperature and wind speed but inversely proportional to air humidity, the higher the number of cases of covid-19, it occurs in provinces with high temperature and wind speed, but inversely proportional to the level of air humidity.

Keywords: hierarchical clustering, covid-19, segmentation

1. Pendahuluan

WHO (World Health Organization) pada 30 Januari 2020 mengeluarkan pernyataan resmi bahwa penyakit yang disebabkan virus SARS CoV-2 sebagai pandemi dan disebut COVID-19 [1]. Beberapa gejala virus ini antara lain adalah kesulitan bernapas, batuk kering, demam, kelelahan, dan infiltrasi paru pada kasus yang parah, mirip dengan yang disebabkan oleh infeksi SARS dan MERS [2]. Virus RNA tidak tersegmentasi milik keluarga Coronaviridae dan ordo Nidovirales dan tersebar luas pada manusia dan mamalia lainnya [3]. Analisis mendalam menemukan bahwa urutan penyakit yang disebabkan oleh COVID-19 yaitu 79,6% identik dengan SARS-CoV, dan 96% identik pada tingkat keseluruhan genom dengan virus corona kelelawar [4] [5].

Efektivitas pada jenis virus corona seperti SARS-CoV hilang setelah terpapar panas pada suhu 56 °C selama 15 menit [6]. Selain itu, kondisi sinar ultraviolet, basa ($\text{pH} > 12$), atau asam ($\text{pH} < 3$) membuat virus tidak aktif [7]. Virus corona manusia terbukti dapat bertahan hidup dalam Phosphate-buffered saline (PBS) atau media kultur dengan 5 - 10% Fetal Calf Serum (FCS) selama beberapa hari, tetapi hanya bertahan beberapa jam setelah pengeringan [8] [9] [10].

Beberapa penelitian melaporkan bahwa kondisi cuaca dan polusi udara terkait dengan wabah SARS [11] [12] [13]. Lingkungan pada kondisi suhu dan kelembaban yang berbeda signifikan terhadap kelangsungan hidup keluarga coronavirus seperti SARSCoV atau SCoV [6] [14]. Virus kering pada permukaan halus mempertahankan viabilitasnya selama lebih dari 5 hari pada suhu 22 - 25 °C dan kelembaban relatif 40 - 50%. Namun, virus dengan cepat menghilang pada suhu dan kelembaban relatif yang lebih tinggi (misalnya, 38 °C dan > 95% kelembaban relatif). Stabilitas virus corona pada suhu dan kelembaban rendah dapat memudahkan penularannya di daerah subtropis, dan dari hal tersebut di atas, beberapa negara Asia di kawasan tropis seperti Malaysia, Indonesia atau Thailand tidak memiliki wabah SARS komunitas yang besar [6].

Banyak penyakit menular memiliki pola musiman dalam kejadian dan penularannya secara global, contohnya influenza [15]. Influenza, SARS dan MERS menunjukkan fluktuasi musiman yang signifikan di daerah beriklim sedang di dunia, tetapi tetap menunjukkan musim yang lebih sedikit di daerah tropis [16] [17] [18]. Namun, berdasarkan proporsi antara jumlah orang yang terinfeksi dan jumlah kematian, Indonesia yang menempati urutan kedua setelah Italia memiliki 1.285 infeksi dengan 114 kematian, yaitu 11,2%. Ini sebenarnya bahkan lebih tinggi dari Italia, yang menyumbang 11,0%. Maka, mengingat jumlah penduduk Indonesia dan negara-negara Asia lainnya di daerah tropis sangat besar dan dengan fasilitas kesehatan yang minim, maka perlu dikaji karakteristik penyebaran COVID-19 di wilayah ini.

Clustering merupakan salah satu metode dari data mining yang dapat digunakan untuk mempelajari pola data berdasarkan sifat kemiripan diantara atribut data. Clustering pada intinya adalah proses pengelompokan sejumlah data ke dalam kelompok-kelompok yang memiliki ciri bahwa setiap anggota di dalam suatu kelompok memiliki kemiripan secara atribut data, sedangkan kelompok satu dengan kelompok lainnya akan memiliki kemiripan yang berbeda secara atribut data. Proses clustering juga memiliki beberapa teknik dalam penerapannya, diantaranya adalah teknik berbasis partisi dan teknik berbasis hierarki. Analisis cluster secara umum dibagi menjadi dua, yaitu hierarchical clustering dan nonhierarchical clustering. Hierarchical clustering adalah algoritma clustering dengan membentuk konstruksi hirarki atau berdasarkan tingkatan tertentu seperti struktur pohon atau struktur pertandingan. Dengan demikian proses clusteringnya bertingkat dan bertahap. Selanjutnya, hierarchical clustering terdiri dari dua macam, yakni agglomerative (penyatuan) dan divisive (penyebaran) [19].

Algoritma Agglomerative Hierarchical Clustering (AHC) adalah algoritma clustering yang dimulai dengan pengelompokan objek-objek yang individual. Banyaknya cluster sama dengan banyaknya objek. Algoritma AHC terdiri dari tiga macam yakni: AHC single linkage, AHC complete linkage, dan AHC average linkage. Single linkage adalah algoritma pengelompokan agglomerative berdasarkan jarak minimum (jarak terdekat) antar objek. Prosedur pengelompokan pada awalnya dipilih jarak terkecil dalam $D = \{d_{ij}\}$ dan menggabungkan objek-objek yang bersesuaian misalnya U dan V untuk mendapatkan cluster (UV). Langkah selanjutnya, jarak diantara (UV) dan cluster lainnya, misalnya W, $d(UV), W = \min(d(UW), d(VW))$ dengan $d(VW) =$ jarak antara tetangga terdekat dari cluster V dan W [20].

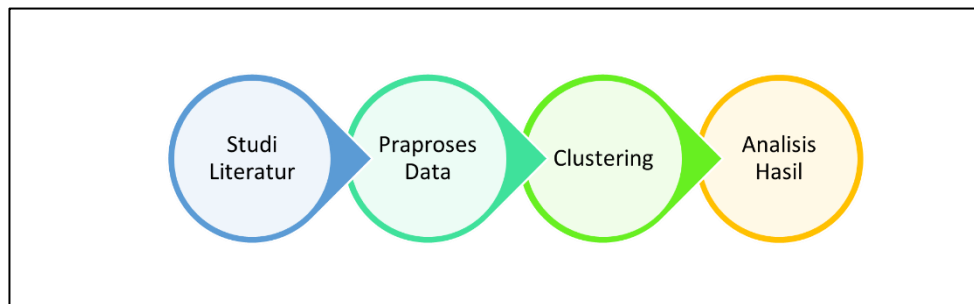
Pada penelitian ini akan menerapkan algoritma hierarki untuk menguji sejauh mana hasil cluster yang diperoleh untuk menunjukkan segmentasi wilayah terdampak pandemi yang dihubungkan dengan faktor suhu dan kelembaban udara.

2. Metode Penelitian

Metode penelitian yang dilakukan terdiri dari beberapa tahapan. Tahapan atau langkah – langkah penelitian ini secara umum dapat dilihat pada gambar 1. Dimulai dari studi literatur, praproses data, proses clustering, dan analisis hasil.

1. Studi Literatur

Studi literatur merupakan langkah yang dilakukan untuk mempelajari referensi berupa jurnal penelitian, paper, buku-buku referensi yang lain terkait dengan penelitian untuk melengkapi pengetahuan awal, guna memahami teori yang dapat digunakan untuk menunjang penelitian.



Gambar 1. Tahapan Penelitian

Sumber: Hasil Penelitian (2023)

2. Praproses Data

Dataset yang digunakan adalah dataset kasus covid-19, data kelembaban dan suhu per propinsi. Data kasus diperoleh dari Kaggle, sedangkan data kelembaban dan suhu diperoleh dari BPS. Dataset yang diperoleh pada umumnya belum siap untuk diolah ke dalam tahap clustering sehingga perlu dipersiapkan.

3. Clustering

Metode clustering yang dilakukan akan menggunakan algoritma agglomerative hierarchical yaitu Single Linkage. Proses pengklasteran pada Single Linkage didasarkan pada perhitungan jarak terdekat diantara dua objek. Dua objek yang terpisah oleh jarak yang pendek, maka otomatis kedua objek tersebut akan masuk dalam satu klaster yang sama. Metode Single Linkage ini diawali dengan penentuan objek – objek yang mempunyai jarak terpendek dalam matriks jarak $D = \{d_{ik}\}$. Objek – objek yang memiliki jarak terpendek akan dikelompokkan dalam satu klaster, misal U dan V adalah dua objek yang akan diklasterkan sehingga diperoleh klaster (UV). Untuk mencari jarak antara klaster (UV) dan klaster W atau klaster lainnya, dihitung berdasarkan rumus:

$$d_{(UV)W} = \min\{d_{UW}, d_{VW}\}$$

Keterangan:

d_{UW} adalah jarak terdekat dari klaster U dan W

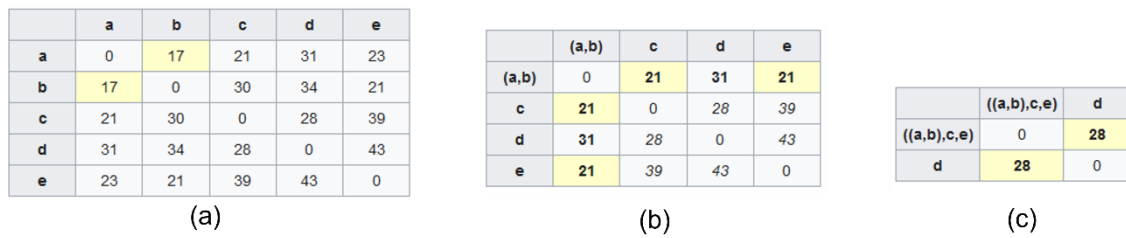
d_{VW} adalah jarak terdekat dari klaster V dan W

Adapun tahapan atau prosedur dari algoritma Single Linkage adalah sebagai berikut :

1. Mencari jarak minimum dari dua objek.
2. Gabungkan dua objek dengan jarak minimum menjadi satu cluster
3. Cari jarak antar cluster dengan menggunakan jarak minimum
4. Ulangi langkah 1-2 sampai semua objek bergabung menjadi satu cluster.

Sedangkan untuk jumlah cluster yang diujikan adalah membentuk sejumlah 3 cluster dengan harapan dapat menemukan kelompok propinsi berdasarkan fitur atau atribut yang dipakai, yaitu jumlah kasus, data kelembaban, dan data suhu.

Adapun ilustrasi contoh hasil proses single linkage clustering adalah ditunjukkan pada gambar 2.



Gambar 2. Matriks Jarak Berpasangan

Sumber: Wikipedia (2023)

Misal ada 5 elemen yang akan di-cluster, maka yang pertama dilakukan adalah menghitung jarak masing – masing elemen, kemudian dicari jarak antar elemen yang mempunyai nilai paling kecil, perhitungan jarak dilakukan dengan membentuk matriks berpasangan seperti pada gambar (a), elemen yang mempunyai jarak terkecil kemudian digabungkan menjadi sebuah cluster dan dilakukan pembaruan perhitungan jarak seperti yang ditunjukkan pada gambar (b). Sama dengan proses sebelumnya, elemen yang mempunyai jarak terkecil dimasukkan ke dalam sebuah cluster, proses berhenti ketika hanya tersisa 2 elemen saja.

4. Analisis Hasil

Tahap analisis hasil pada penelitian ini adalah melakukan observasi atau pengamatan terhadap hasil cluster yang dihasilkan dan analisis terhadap anggota pada masing – masing cluster terhadap fitur atau atribut yang digunakan. Pembatasan jumlah cluster yang dibentuk yaitu 3 buah cluster mempunyai tujuan untuk mengetahui mana cluster yang tinggi, sedang, dan rendah, berdasarkan atribut data yang dipakai. Adapun evaluasi hasil clustering yang akan dilakukan adalah menggunakan DBI (*Davies Bouldin Index*).

3. Hasil dan Pembahasan

Dataset yang digunakan adalah dataset kasus covid-19, data kelembaban dan suhu per propinsi. Data kasus diperoleh dari Kaggle, sedangkan data kelembaban dan suhu diperoleh dari BPS. Untuk data kelembaban udara, suhu udara, dan kecepatan angin, diambil data rata – rata periode 5 tahun terakhir per propinsi. Data kasus terkonfirmasi disajikan pada tabel 1 yang diperoleh dari website Kaggle dimana pada tabel menunjukkan jumlah kasus terkonfirmasi covid-19 pada tahun 2021 untuk tiap – tiap propinsi, sedangkan data suhu, kecepatan angin, dan kelembaban diperoleh dari BPS, dimana ketiga data ini adalah rata – rata data tahunan yang dikumpulkan mulai tahun 2017 sampai tahun 2021. Penggabungan data yang dilakukan adalah dengan menggabungkan atribut kasus terkonfirmasi dengan atribut suhu, atribut kecepatan angin, dan atribut kelembaban berdasarkan propinsi yang bersesuaian.

Tabel 1. Data Kasus Terkonfirmasi

Province_name	Confirmed_cases
Aceh	7.797
Sumatera Utara	13.277
Sumatera Barat	16.709
Riau	14.841
Jambi	1.399
Sumatera Selatan	8.534
Bengkulu	1.294
Kepulauan Bangka Belitung	777

Lampung	1.884
Kepulauan Riau	4.640
DKI Jakarta	115.174
Jawa Barat	42.572
Jawa Tengah	39.788
Daerah Istimewa Yogyakarta	4.399
Jawa Timur	55.575
Banten	10.554
Bali	12.519
Nusa Tenggara Barat	4.314
Nusa Tenggara Timur	694
Kalimantan Barat	2.035
Kalimantan Tengah	4.801
Kalimantan Selatan	12.365
Kalimantan Timur	16.384
Kalimantan Utara	928
Sulawesi Utara	5.452
Sulawesi Tengah	878
Sulawesi Selatan	18.372
Sulawesi Tenggara	5.073
Gorontalo	3.021
Sulawesi Barat	1.036
Maluku	3.863
Maluku Utara	2.232
Papua	9.037
Papua Barat	4.189

Sumber : Kaggle (2021)

Tabel 2. Data Suhu Kecepatan Angin Kelembaban

No	Propinsi	Suhu	Kecepatan Angin	Kelembaban
1	Aceh	27,10	2,71	80,00
2	Sumatera Utara	27,40	2,38	86,90
3	Sumatera Barat	26,50	2,88	84,00
4	Riau	27,20	2,97	80,50
5	Jambi	27,00	2,39	82,10
6	Sumatera Selatan	27,70	3,32	79,50
7	Bengkulu	27,00	2,12	83,20
8	Lampung	27,10	1,95	78,90
9	Kepulauan Bangka Belitung	27,30	4,05	79,90
10	Kepulauan Riau	27,00	3,16	84,10
11	DKI Jakarta	28,40	1,54	74,00
12	Jawa Barat	23,50	2,14	74,40
13	Jawa Tengah	28,50	2,83	70,00
14	DI Yogyakarta	26,10	0,07	82,80

15	Jawa Timur	28,00	3,88	75,20
16	Banten	27,30	0,99	79,30
17	Bali	27,30	3,27	79,10
18	Nusa Tenggara Barat	26,10	3,27	81,40
19	0Nusa Tenggara Timur	27,50	4,04	75,60
20	Kalimantan Barat	26,90	1,75	85,70
21	Kalimantan Tengah	27,70	2,17	80,50
22	Kalimantan Selatan	27,00	1,92	81,20
23	Kalimantan Timur	27,90	1,96	79,70
24	Kalimantan Utara	27,60	2,20	83,70
25	Sulawesi Utara	27,00	2,92	75,60
26	Sulawesi Tengah	28,40	2,32	72,60
27	Sulawesi Selatan	27,30	2,89	75,20
28	Sulawesi Tenggara	26,90	1,27	83,10
29	Gorontalo	27,30	1,95	77,50
30	Sulawesi Barat	27,90	1,93	77,20
31	Maluku	26,50	2,35	83,60
32	Maluku Utara	27,30	2,60	78,30
33	Papua Barat	27,40	1,54	83,60
34	Papua	27,80	2,60	75,50

Sumber : BPS

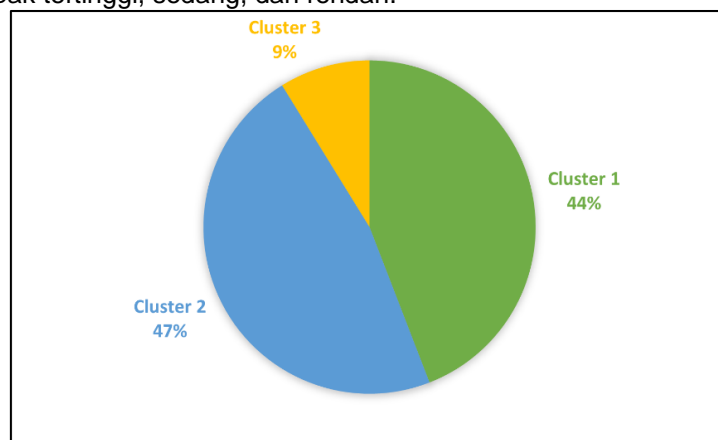
Tabel 3. Dataset yang digunakan

No	Provinsi	Terkonfirmasi	Suhu	Kecepatan Angin	Kelembaban
1	Aceh	7797	27.1	2.71	80
2	Sumatera Utara	13277	27.4	2.38	86.9
3	Sumatera Barat	16709	26.5	2.88	84
4	Riau	14841	27.2	2.97	80.5
5	Jambi	1399	27	2.39	82.1
6	Sumatera Selatan	8534	27.7	3.32	79.5
7	Bengkulu	1294	27	2.12	83.2
8	Kepulauan Bangka Belitung	777	27.1	1.95	78.9
9	Lampung	1884	27.3	4.05	79.9
10	Kepulauan Riau	4640	27	3.16	84.1
11	DKI Jakarta	115174	28.4	1.54	74
12	Jawa Barat	42572	23.5	2.14	74.4
13	Jawa Tengah	39788	28.5	2.83	70
14	Daerah Istimewa Yogyakarta	4399	26.1	0.07	82.8
15	Jawa Timur	55575	28	3.88	75.2
16	Banten	10554	27.3	0.99	79.3
17	Bali	12519	27.3	3.27	79.1
18	Nusa Tenggara Barat	4314	26.1	3.27	81.4
19	Nusa Tenggara Timur	694	27.5	4.04	75.6

20	Kalimantan Barat	2035	26.9	1.75	85.7
21	Kalimantan Tengah	4801	27.7	2.17	80.5
22	Kalimantan Selatan	12365	27	1.92	81.2
23	Kalimantan Timur	16384	27.9	1.96	79.7
24	Kalimantan Utara	928	27.6	2.2	83.7
25	Sulawesi Utara	5452	27	2.92	75.6
26	Sulawesi Tengah	878	28.4	2.32	72.6
27	Sulawesi Selatan	18372	27.3	2.89	75.2
28	Sulawesi Tenggara	5073	26.9	1.27	83.1
29	Gorontalo	3021	27.3	1.95	77.5
30	Sulawesi Barat	1036	27.9	1.93	77.2
31	Maluku	3863	26.5	2.35	83.6
32	Maluku Utara	2232	27.3	2.6	78.3
33	Papua	9037	27.4	1.54	83.6
34	Papua Barat	4189	27.8	2.6	75.5

Sumber: Biro Pusat Statistik (2017-2021)

Setelah melalui tahapan praproses data seperti yang disajikan pada tabel 3, maka dilakukan tahapan clustering, dengan bantuan perangkat lunak python dengan memanfaatkan library yang telah disediakan. Untuk jumlah cluster yang dibentuk adalah 3 cluster, dengan tujuan mengelompokkan propinsi – propinsi ke dalam 3 bagian, yaitu propinsi mana saja yang mengalami dampak tertinggi, sedang, dan rendah.



Gambar 3. Diagram Sebaran Persentase Anggota Cluster

Sumber: Hasil Penelitian (2023)

Dari proses clustering yang telah dilakukan, dengan membatasi pembentukan cluster sejumlah 3 buah dengan tujuan mendapatkan cluster atau kelompok mana yang mempunyai nilai tinggi, sedang, atau rendah berdasarkan atribut yang dipakai, maka diperoleh pada cluster pertama terdapat 15 propinsi, pada cluster kedua terdapat 16 propinsi, dan pada cluster ketiga mempunyai 3 propinsi. Adapun sebaran masing masing anggota yaitu nama – nama propinsi, ditunjukkan pada tabel 4 sampai tabel 6.

Tabel 2. Daftar Anggota Cluster 1

Provinsi	No	Terkonfirmasi	Suhu	Kecepatan Angin	Kelembaban	Klaster
Sumatera Utara	1	13277	27.4	2.38	86.9	1
Sumatera Barat	2	16709	26.5	2.88	84	1

Jambi	4	1399	27	2.39	82.1	1
Bengkulu	6	1294	27	2.12	83.2	1
Kepulauan Riau	9	4640	27	3.16	84.1	1
Jawa Barat	11	42572	23.5	2.14	74.4	1
Daerah Istimewa Yogyakarta	13	4399	26.1	0.07	82.8	1
Banten	15	10554	27.3	0.99	79.3	1
Nusa Tenggara Barat	17	4314	26.1	3.27	81.4	1
Kalimantan Barat	19	2035	26.9	1.75	85.7	1
Kalimantan Selatan	21	12365	27	1.92	81.2	1
Kalimantan Utara	23	928	27.6	2.2	83.7	1
Sulawesi Tenggara	27	5073	26.9	1.27	83.1	1
Maluku	30	3863	26.5	2.35	83.6	1
Papua	32	9037	27.4	1.54	83.6	1

Sumber: Hasil Penelitian (2023)

Tabel 3. Daftar Anggota Cluster 2

Provinsi	No	Terkonfirmasi	Suhu	Kecepatan Angin	Kelembaban	Klaster
Aceh	0	7797	27.1	2.71	80	2
Riau	3	14841	27.2	2.97	80.5	2
Sumatera Selatan	5	8534	27.7	3.32	79.5	2
Kepulauan Bangka Belitung	7	777	27.1	1.95	78.9	2
Lampung	8	1884	27.3	4.05	79.9	2
Bali	16	12519	27.3	3.27	79.1	2
Nusa Tenggara Timur	18	694	27.5	4.04	75.6	2
Kalimantan Tengah	20	4801	27.7	2.17	80.5	2
Kalimantan Timur	22	16384	27.9	1.96	79.7	2
Sulawesi Utara	24	5452	27	2.92	75.6	2
Sulawesi Tengah	25	878	28.4	2.32	72.6	2
Sulawesi Selatan	26	18372	27.3	2.89	75.2	2
Gorontalo	28	3021	27.3	1.95	77.5	2
Sulawesi Barat	29	1036	27.9	1.93	77.2	2
Maluku Utara	31	2232	27.3	2.6	78.3	2
Papua Barat	33	4189	27.8	2.6	75.5	2

Sumber: Hasil Penelitian (2023)

Tabel 4. Daftar Anggota Cluster 3

Provinsi	No	Terkonfirmasi	Suhu	Kecepatan Angin	Kelembaban	Klaster
DKI Jakarta	10	115174	28.4	1.54	74	2
Jawa Tengah	12	39788	28.5	2.83	70	2
Jawa Timur	14	55575	28	3.88	75.2	2

Sumber: Hasil Penelitian (2023)

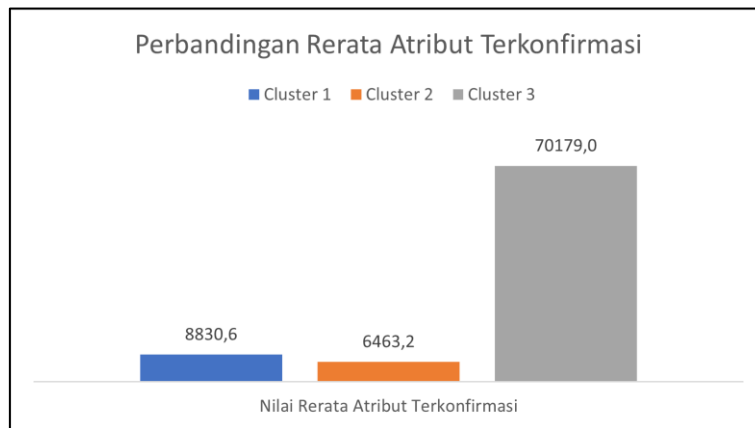
Dari masing – masing cluster yang terbentuk, maka bisa diambil suatu informasi yang dapat menggambarkan atau mendeskripsikan setiap kelompok cluster, tentunya berdasarkan atribut – atribut atau parameter yang dipakai, proses ini umumnya disebut profiling cluster. Adapun evaluasi cluster menggunakan DBI yang menghasilkan nilai 0,31412426, dimana dalam implementasinya nilai DBI yang mendekati 0 adalah yang optimal.

Dari masing masing kelompok cluster yang terbentuk, dapat diketahui nilai rata – rata dari masing – masing atribut, yang kemudian dapat digunakan sebagai acuan untuk melakukan profiling cluster. Nilai setiap atribut akan dijumlah kemudian dibagi dengan jumlah anggota pada masing – masing cluster, sehingga diperoleh rerata atribut per kelompok cluster. Perhitungan rerata ini untuk menunjukkan bahwa dari atribut atau fitur yang dipakai, cluster mana yang mempunyai rata – rata nilai tinggi, sedang, atau rendah.

Tabel 5. Nilai Rerata Atribut

	Terkonfirmasi	Suhu	Kecepatan Angin	Kelembaban
Cluster 0	8830.6	26.7	2.0	82.6
Cluster 1	6463.2	27.5	2.7	77.9
Cluster 2	70179.0	28.3	2.8	73.1

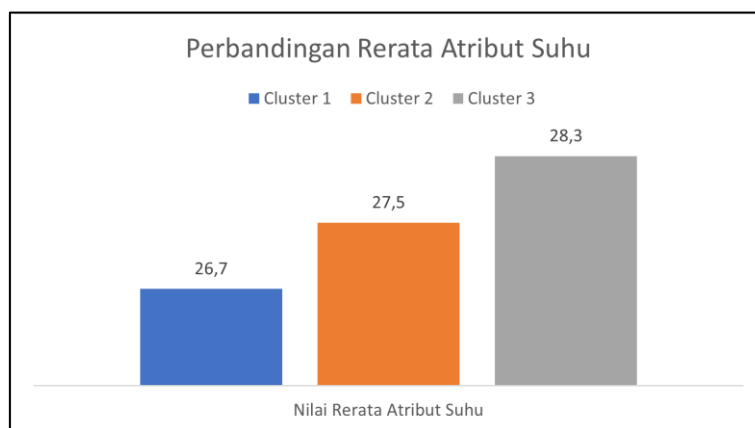
Sumber: Hasil Penelitian (2023)



Gambar 3. Grafik Rerata Atribut Terkonfirmasi

Sumber: Hasil Penelitian (2023)

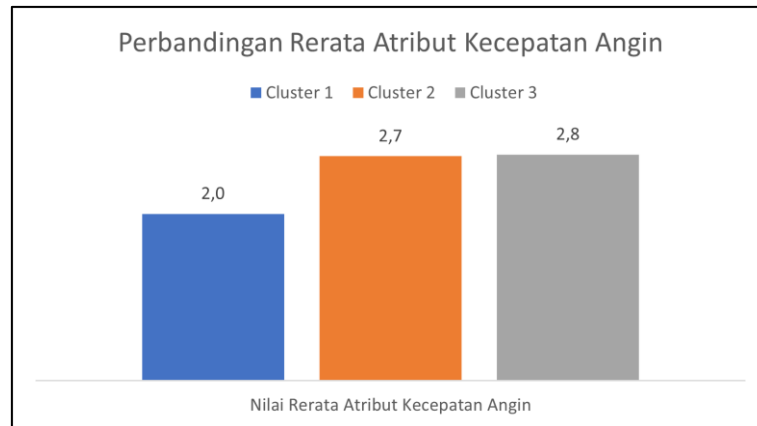
Pada atribut rerata atribut terkonfirmasi maka dapat dilihat bahwa pada cluster ketiga memiliki rerata tertinggi, sedangkan pada cluster kedua memiliki rerata terendah.



Gambar 4. Grafik Rerata Atribut Shuhu

Sumber: Hasil Penelitian (2023)

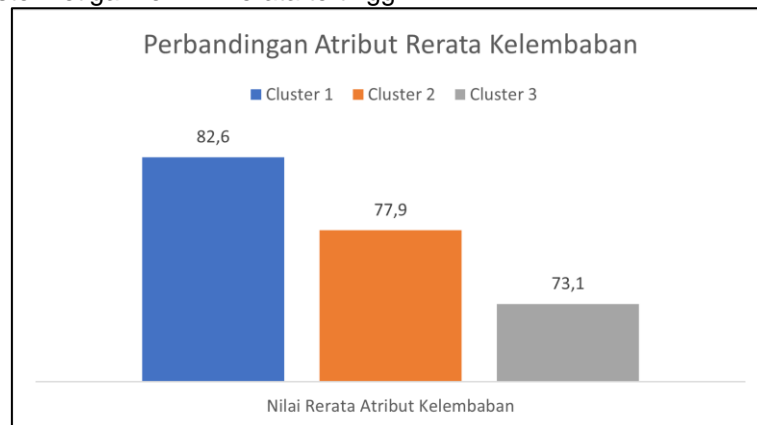
Pada rerata atribut suhu, dapat diketahui bahwa cluster ketiga memiliki rerata tertinggi, kemudian cluster pertama memiliki nilai terendah.



Gambar 5. Grafik Rerata Atribut Kecepatan Angin

Sumber: Hasil Penelitian (2023)

Pada rerata atribut kecepatan angin cluster pertama memiliki nilai rerata terendah, sedangkan cluster ketiga memiliki rerata tertinggi.



Gambar 6. Grafik Rerata Atribut Kelembaban

Sumber: Hasil Penelitian (2023)

Pada nilai rerata atribut kelembaban cluster pertama memiliki nilai tertinggi, sedangkan cluster ketiga memiliki nilai rerata terendah.

Berdasarkan profiling cluster yang telah dilakukan maka dapat diperoleh informasi bahwa, kejadian atau kasus covid-19 berbanding lurus dengan suhu dan kecepatan angin namun berbanding terbalik dengan kelembaban udara, semakin tinggi angka kasus covid-19 maka terjadi pada propinsi dengan suhu dan kecepatan angin tinggi, akan tetapi berbanding terbalik dengan tingkat kelembaban udara.

4. Kesimpulan

Kesimpulan dari penelitian ini adalah pada atribut rerata atribut terkonfirmasi maka dapat dilihat bahwa pada cluster ketiga memiliki rerata tertinggi, sedangkan pada cluster kedua memiliki rerata terendah. Pada rerata atribut suhu, dapat diketahui bahwa cluster ketiga memiliki rerata tertinggi, kemudian cluster pertama memiliki nilai terendah. Pada rerata atribut kecepatan angin cluster pertama memiliki nilai rerata terendah, sedangkan cluster ketiga memiliki rerata tertinggi. Pada nilai rerata atribut kelembaban cluster pertama memiliki nilai tertinggi, sedangkan cluster ketiga memiliki nilai rerata terendah. Semakin tinggi angka kasus covid-19 maka terjadi pada propinsi dengan suhu dan kecepatan angin tinggi, akan tetapi berbanding terbalik dengan tingkat kelembaban udara. Hasil evaluasi cluster menggunakan DBI menghasilkan nilai 0,31412426 yang menunjukkan bahwa hasil cluster dapat dikatakan optimal.

Ucapan Terima Kasih

Terima kasih kepada Lembaga Penelitian dan Pengabdian Masyarakat Universitas Muhammadiyah Jember yang telah mendukung dengan mendanai penelitian ini dan telah membantu pelaksanaan selama berjalannya kegiatan penelitian.

Referensi

- [1] WHO, "Coronavirus disease (COVID-19) weekly epidemiological update and weekly operational update," 2020. [Online]. Available: <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports>.
- [2] E. Gralinski and V. Menachery, Return of the coronavirus: 2019-ncov, 2020.
- [3] C. Huang, Y. Wang, X. Li, L. Ren, J. Zhao, Y. Hu, L. Zhang, G. Fan, J. Xu, X. Gu, Z. Cheng, T. Yu, J. Xia, Y. Wei, W. Wu, X. Xie, W. Yin, H. Li, M. Liu, Y. Xiao, H. Gao, L. Guo, J. Xie, G. Wang, R. Jiang, Z. Gao, Q. Jin, J. Wang and B. Cao, "Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China," *Lancet*, pp. 497-506, 2020.
- [4] P. Zhou, X. Yang, X. Wang, B. Hu, L. Zhang, W. Zhang, H. Si, Y. Zhu, B. Li, C. Huang, H. Chen, J. Chen, Y. Luo, H. Guo, R. Jiang, M. Liu, Y. Chen, X. Shen, X. Wang, X. Zheng, K. Zhao, Q. Chen, F. Deng, L. Liu, B. Yan, F. Zhan, Y. Wang, G. Xiao and Z. Shi, "A pneumonia outbreak associated with a new coronavirus of probable bat origin," *Nature*, p. 270, 2020.
- [5] D. Xu, Z. Zhang, F. Chu, Y. Li, L. Jin, L. Zhang, G. Gao and F. Wang, "Genetic Variation of SARS Coronavirus in Beijing Hospital," *Emerging Infectious Diseases*, pp. 789-794, 2004.
- [6] K. Chan, J. MalikPeiris, S. Lam, L. Poon, K. Yuen and W. Seto, "The effects of temperature and relative humidity on the viability of the SARS coronavirus," *Advances in Virology*, pp. 1-7, 2011.
- [7] M. Darnell, K. Subbarao, S. Feinstone and D. Taylor, "Inactivation of the coronavirus that induces severe acute respiratory syndrome, SARS-CoV," *Journal of Virological Methods*, pp. 85-91, 2004.
- [8] M. Ijaz, A. Brunner and S. Sattar, "Survival characteristics of airborne human coronavirus 229E," *Journal of General Virology*, pp. 2743-2748, 1985.
- [9] J. Sizun, M. Yu and P. Talbot, "Survival of human coronaviruses 229E and OC43 in suspension and after drying on surfaces: a possible source of hospital-acquired infections," *Journal of Hospital Infection*, pp. 55-60, 2000.
- [10] H. Rabenau, J. Cinatl, B. Morgenstern, G. Bauer, W. Preiser and H. Doerr, "Stability and inactivation of SARS coronavirus," *Medical Microbiology and Immunology*, pp. 1-6, 2005.
- [11] J. Tan, L. Mu, J. Huang, S. Yu, B. Chen and J. Yin, "An initial investigation of the association between the SARS outbreak and weather: With the view of the environmental temperature and its variation," *Journal of Epidemiology and Community Health*, pp. 186-192, 2005.
- [12] J. Yuan, H. Yun, W. Lan, W. Wang, S. Sullivan, S. Jia and A. Bittles, "A climatologic investigation of the SARS-CoV outbreak in Beijing, China," *American Journal of Infection Control*, pp. 234-236, 2006.
- [13] Q. Cai, J. Lu, Q. Xu, Q. Guo, D. Xu, Q. Sun, H. Yang, G. Zhao and Q. Jiang, "Influence of meteorological factors and air pollution on the outbreak of severe acute respiratory syndrome," *Public Health*, pp. 258-265, 2007.
- [14] L. Casanova, S. Jeon, W. Rutala, D. Weber and M. Sobsey, "Effects of air temperature and relative humidity on coronavirus survival on surfaces," *Applied and Environmental Microbiology*, pp. 2712-2717, 2010.
- [15] G. Collaborators, "Mortality, morbidity, and hospitalisations due to influenza lower respiratory tract infections, 2017: An analysis for the global burden of disease study 2017," *The Lancet Respiratory Medicine*, pp. 68-89, 2019.
- [16] K. Feshbach, W. Alonso, V. Charu, J. Tamerius, L. Simonsen, M. Miller and C. Viboud, "Latitudinal variations in seasonal activity of influenza and respiratory syncytial virus (RSV): A global comparative review," *PLOS Medicine*, 2013.

- [17] Y. Li, R. Reeves, X. Wang, Q. Bassat, W. Brooks, C. Cohen, D. Moore, M. Nunes, B. Rath, H. Campbell and H. Nair, "Global patterns in monthly activity of influenza virus, respiratory syncytial virus, parainfluenza virus, and metapneumovirus: A systematic analysis," *Lancet Global Health*, 2019.
- [18] C. Viboud, W. Alonso and L. Simonsen, "Influenza in tropical regions," *PLOS Medicine*, 2006.
- [19] X. Wu, V. Kumar, J. Ross Quinlan and e. al, "Top 10 algorithms in data mining," *Knowledge Information System*, no. 14, pp. 1-37, 2008.
- [20] J. Han, M. Kamber and J. Pei, *Data Mining: Concepts and Techniques*, Morgan Kaufmann, 2012.