

PEMETAAN PENYEBARAN COVID-19 PADA TINGKAT KABUPATEN/KOTA DI PULAU JAWA MENGGUNAKAN ALGORITMA K-MEANS CLUSTERING

Lidia Gayatri¹⁾ dan Hendry²⁾

^{1,2} Teknik Informatika, Fakultas Teknologi Informasi, Universitas Kristen Satya Wacana

^{1,2} Jl. Dr. O. Notohamidjojo No. 1, Salatiga, 50715

E-mail : lidiagayatri123@gmail.com¹⁾, hendry@uksw.edu²⁾

ABSTRAK

Coronavirus atau Covid-19 adalah virus yang ditemukan di Wuhan, China pada Desember 2019, memiliki kemampuan penyebaran yang cukup cepat diseluruh dunia termasuk di Indonesia. Kasus Covid-19 mulai masuk di Indonesia pada 2 Maret 2020 dan kasus Covid-19 terus meningkat setiap hari, tercatat 1.492.002 kasus Covid-19 pada 27 Maret 2021 di Indonesia. Kasus Covid-19 yang terus meningkat menyebabkan perlunya pemetaan tingkat kerawanan penyebaran Covid-19 khususnya di Pulau Jawa menggunakan data dari *website* resmi pemerintah pada tingkat provinsi dengan menggunakan 3 parameter, yaitu jumlah kasus dirawat, sembuh, dan meninggal. Untuk menentukan banyaknya *cluster* digunakan algoritma K-Means dan metode *Davies Bouldin Index* (DBI). Algoritma K-Means adalah salah satu metode *clustering* yang dapat membagi data ke dalam beberapa kelompok atau *cluster*, sedangkan DBI digunakan untuk menghitung kemiripan setiap *cluster*. Hasil pengujian menunjukkan *cluster* terbaik pada ukuran *cluster* 3 dengan nilai DBI 0.609. Terdapat 3 tingkat kerawanan, yaitu kerawanan rendah terdapat pada *cluster* 0 yang memiliki 105 kabupaten/kota, kerawanan sedang terdapat pada *cluster* 2 yang memiliki 7 kabupaten/kota, dan kerawanan tinggi terdapat pada *cluster* 1 yang memiliki 7 kabupaten/kota. Hasil pemetaan menunjukkan kabupaten/kota yang berada pada tingkat kerawanan tinggi berada di Kota Jakarta Utara, Kota Jakarta Barat, Kota Jakarta Pusat, Kota Jakarta Selatan, Kota Jakarta Timur, Kota Surabaya, dan Kota Semarang. Hasil dari penelitian diharapkan dapat digunakan sebagai acuan oleh masyarakat maupun pemerintah dalam melihat setiap daerah yang memiliki tingkat kerawanan rendah, sedang, dan tinggi dalam penyebaran Covid-19 di Pulau Jawa, agar pemerintah memfokuskan perhatian lebih kepada kabupaten/kota yang memiliki tingkat kerawanan tinggi.

Kata Kunci: Covid-19, DBI, Data Mining, K-Means, Clustering, Pemetaan.

1. PENDAHULUAN

Coronavirus adalah virus yang menyebabkan penyakit pada manusia dan hewan. Ditemukan di Wuhan China, pada Desember 2019, kemudian virus ini diberi nama *Severe Acute Respiratory Syndrome Coronavirus 2* (SARS-COV2) dan menyebabkan penyakit Corona Disease-2019 (Covid-19). Covid-19 pada manusia biasanya menginfeksi saluran pernapasan, seperti flu sampai penyakit yang serius seperti *Middle East Respiratory Syndrome* dan *Severe Acute Respiratory Syndrome*. Indonesia mulai mengkonfirmasi penyebaran Covid-19 pada 2 Maret 2020 (Kementerian Kesehatan RI, 2020). Hampir di seluruh wilayah Indonesia sudah terdapat kasus Covid-19, tercatat pada tanggal 27 Maret 2021 jumlah kasus Covid-19 di Indonesia mencapai 1.492.002 kasus (Purnomo, 2021). Menurut (Solichin & Khairunnisa, 2020) Pulau Jawa merupakan pulau dengan provinsi yang menjadi titik awal persebaran Covid-19 di Indonesia dengan 1.005.292 kasus Covid-19 di Pulau Jawa pada 27 Maret 2021 (Tim Detikcom, 2021).

Usaha pemerintah untuk mencegah penularan Covid-19 serta memutus rantai penularan yaitu dengan cara melakukan isolasi diri bagi mereka yang terkena virus

atau yang melakukan perjalanan diluar kota maupun luar pulau, tetap berada di rumah dan selalu cuci tangan dan menggunakan masker apabila di luar rumah (Komite Penanganan Covid-19, 2020). Namun, walaupun sudah diberlakukannya anjuran atau Pembatasan Sosial Berskala Besar (PSBB) kasus Covid-19 terus bertambah hari demi hari. Mengelompokkan daerah atau kabupaten/kota yang mempunyai tingkat kerawanan penyebaran Covid-19 lebih tinggi merupakan salah satu cara untuk membatasi aktivitas penyebaran Covid-19 semakin meluas sehingga dapat mengurangi peningkatan kasus Covid-19.

Clustering merupakan metode dalam *data mining* yang digunakan untuk mengelompokkan data atau memetakan data berdasarkan kemiripannya (Ridlo, Defiyanti & Primajaya, 2017) (Heraldi, Aprilia & Pratiwi, 2019). Algoritma yang dapat digunakan untuk *clustering* yaitu K-Means. K-Means adalah algoritma *non-hierarchical* yang cukup populer dan sudah banyak digunakan di beberapa penelitian.

Penelitian terdahulu berjudul Klasterisasi Persebaran Virus Corona (Covid-19) di DKI Jakarta Menggunakan Metode K-Means, bertujuan untuk melakukan *clustering*

wilayah Provinsi DKI Jakarta per kelurahan dengan menggunakan 5 parameter, yaitu jumlah ODP, PDP, kasus positif, pasien sembuh dan pasien meninggal. Hasil perhitungan dengan metode K-Means berdasarkan nilai SSE mendapatkan 9 *cluster* (Solichin & Khairunnisa, 2020).

Penelitian terdahulu lainnya berjudul Penentuan Jumlah *Cluster* Ideal SMK di Jawa Tengah Dengan Metode X-Means *Clustering* dan K-Means *Clustering*, membahas tentang pengembangan revitalisasi SMK dengan pengelompokan SMK berdasarkan data kemendikbud di Jawa Tengah. X-means dan K-Means adalah metode yang digunakan untuk menghitung nilai *Davies Bouldin Index* (DBI) menggunakan data pokok Kemendikbud untuk mencari pembagian *cluster* terbaik. Hasil perhitungan X-means dan K-Means membagi data ke dalam 4 kelompok yaitu kurang, cukup, baik, dan unggul (Adhitama, 2020).

Penelitian terdahulu lainnya berjudul Pemetaan Kasus Covid-19 di Sulawesi Tenggara Tahun 2020, memiliki tujuan untuk mengetahui hubungan antara kepadatan penduduk dengan kejadian Covid-19 di Sulawesi Tenggara tahun 2020. Dengan menggunakan desain penelitian deskriptif analitik, teknik *sampling* mengambil sampel penelitian dari data sekunder yaitu data kasus Covid-19 tanggal 11 Mei 2020. Perangkat lunak yang digunakan untuk memetakan yaitu GIS dan memperoleh hasil bahwa kepadatan penduduk dengan kejadian Covid-19 di Sulawesi Tenggara tidak memiliki hubungan yang signifikan (Hanafi et al., 2020).

Penelitian terdahulu lainnya berjudul Analisis *Cluster* Untuk Pemetaan Data Kasus Covid-19 di Indonesia Menggunakan K-Means, membahas tentang *clustering* wilayah provinsi di Indonesia dengan data Covid-19 untuk menghindari wilayah dengan kasus persebaran Covid-19 yang tinggi menggunakan algoritma K-Means untuk membagi wilayah ke dalam beberapa kelompok. Penelitian ini mendapatkan hasil *cluster* optimal pada 2 kelompok dengan nilai SC 0,74, yaitu kelompok rawan sebanyak 6 provinsi dan kelompok aman sebanyak 28 provinsi (Sholihah, 2021)

Penelitian ini bertujuan melakukan *clustering* data Covid-19 di Pulau Jawa dari tahun 2020-2021 menggunakan algoritma K-Means dan DBI. K-Means digunakan untuk pengelompokan berdasarkan kemiripan data, sehingga dapat digunakan untuk mengelompokkan data Covid-19 di Pulau Jawa. Setelah mendapatkan jumlah *cluster* data Covid-19 dilakukan penilaian menggunakan DBI untuk mendapatkan jumlah *cluster* yang optimal. Untuk memudahkan membaca data *clustering* Covid-19 hasil DBI divisualisasikan ke dalam bentuk peta.

2. RUANG LINGKUP

Penelitian ini memiliki permasalahan yang dapat dilihat sebagai berikut:

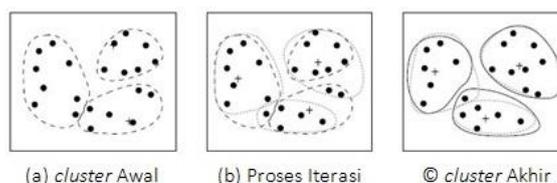
1. Bagaimana cara mengelompokkan data Covid-19 berdasarkan tingkat kabupaten/kota. Diharapkan data yang sudah dikelompokkan kemudian dapat dilakukan pemetaan. Data yang digunakan adalah data Covid-19 dari portal website milik pemerintah provinsi dan kabupaten/kota.
2. Menerapkan algoritma K-Means untuk mengelompokkan data Covid-19 di Pulau Jawa.
3. Menggunakan data Covid-19 di 6 provinsi yang ada di Pulau Jawa pada tingkat kabupaten/kota. Analisis ini diterapkan dengan mengelompokkan tingkat kerawanan Covid-19 di Pulau Jawa.
4. Pengujian menggunakan DBI dari hasil algoritma K-Means untuk mendapatkan jumlah *cluster* kemudian dilakukan visualisasi dalam bentuk peta.

3. BAHAN DAN METODE

Bagian ini akan membahas metode atau langkah-langkah yang digunakan untuk melakukan pemetaan.

3.1 *Clustering*

Clustering merupakan metode yang digunakan untuk mengelompokkan data ke dalam beberapa kelompok sehingga satu kelompok memiliki tingkat kemiripan maksimum setiap datanya, sedangkan antar *cluster* memiliki kemiripan minimum setiap datanya. Beberapa manfaat *clustering* adalah prediksi dan analisa masalah bisnis tertentu karena *clustering* merupakan metode segmentasi data serta dapat mengidentifikasi objek dalam berbagai bidang (Irwansyah & Faisal, 2015). Proses *clustering* berlangsung berdasarkan data yang ada, kemudian *cluster* yang sudah terbentuk akan dilakukan iterasi hingga *cluster* tersebut tidak berubah. Proses *clustering* dapat dilihat pada Gambar (Han, Kamber & Pei, 2012) dapat dilihat pada gambar 1.



Gambar 1. Proses *clustering*

3.2 K-Means

Algoritma K-Means adalah algoritma yang digunakan untuk *clustering* dan berbentuk *non-hierarchical* yang memiliki waktu komputasi relatif cepat (Ridlo, Defiyanti & Primajaya, 2017). Menentukan jumlah berapa *cluster* yang akan digunakan adalah awal yang dilakukan algoritma K-Means, kemudian menentukan nilai awal untuk masing-masing *cluster*. Selanjutnya menghitung jarak dari tiap data dengan nilai awal yang telah ditentukan Data tersebut kemudian ditempatkan dalam *cluster* terdekat. *Cluster* yang sudah memiliki data kemudian dihitung nilai rata-rata dari tiap

cluster yang nantinya nilai tersebut digunakan untuk nilai awal yang baru dalam menghitung jarak dari tiap data. Iterasi terus berlanjut sampai nilai *cluster* yang baru sama dengan nilai *cluster* sebelumnya atau tidak mengalami perubahan (Heraldi, Aprilia & Pratiwi, 2019). Untuk menghitung jarak ke pusat *cluster*, menggunakan rumus pada persamaan (1).

$$D(i, j) = \sqrt{(X_{1i} - X_{1j})^2 + \dots + (X_{ki} - X_{kj})^2} \quad (1)$$

Di mana $D(i, j)$ adalah jarak ke i ke pusat *cluster* j , X_{ki} adalah data ke i pada atribut data ke k , dan X_{kj} adalah titik pusat ke j pada atribut ke k (Santosa, 2007).

3.3 Davies Bouldin Index

Davies Bouldin Index (DBI) adalah metode menghitung kesamaan antar anggota *cluster* atau yang memiliki kemiripan (Nawrin, Rahatur & Akhter, 2017). DBI melakukan perhitungan dengan membandingkan rasio *cluster* ke i dan ke j . Nilai DBI yang semakin kecil menunjukkan *cluster* yang dihasilkan semakin baik (Adhitama, Burhanuddin & Ananda, 2020). Untuk menghitung nilai DBI dapat menggunakan persamaan (2):

$$DBI = \frac{1}{k} \sum_{i=1}^k \max_{i \neq j} (R_{i,j}) \quad (2)$$

Di mana k adalah banyak *cluster*, I dan j adalah *cluster* ke i dan ke j .

3.4 Pemetaan

Peta adalah gambar permukaan bumi dua dimensi yang diperkecil dan termuat dalam sebuah kertas. Dengan menggunakan sebuah peta, memudahkan dalam pengamatan permukaan bumi secara luas (Banyumanik, 2018). Pemetaan adalah seluruh proses yang dilakukan untuk membuat sebuah peta (Yusuf & Halim, 2014). Proses pemetaan dapat dilakukan dengan banyak cara, salah satunya secara digital. Cara ini memanfaatkan teknologi digital yang dapat mengurangi sumber-sumber kesalahan seperti faktor manusia ataupun karena distorsi media (Armijon, 2019).

3.5 Metodologi Penelitian

Metode penelitian yang digunakan dalam penelitian ini adalah kuantitatif yang menggunakan data sekunder. Penelitian ini terdiri dari tahapan yang saling berhubungan satu sama lain. Tahap ini dapat dilihat pada Gambar 2.

1. Studi Literatur

Tahapan ini melakukan pencarian informasi baik melalui jurnal, artikel, buku, skripsi maupun tesis untuk mendapatkan informasi yang sesuai dengan penelitian yang sedang dilakukan.

2. Pengumpulan Data

Tahapan ini dilakukan untuk mengumpulkan data yang akan digunakan dalam penelitian dengan menggunakan data sekunder.

3. Preprocessing

Tahapan ini merupakan tahapan untuk membersihkan data yang tidak digunakan, memiliki *missing value*, dan normalisasi data.

4. Implementasi Algoritma K-Means Clustering

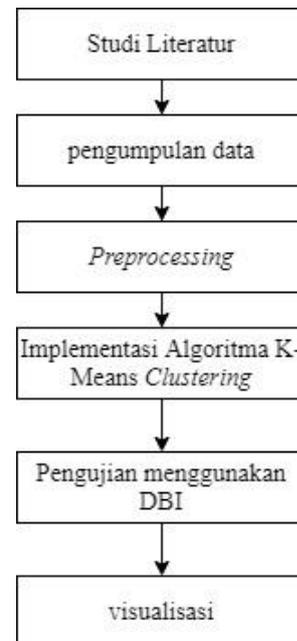
Tahapan ini merupakan tahapan terpenting dalam penelitian ini. Algoritma K-Means digunakan untuk mengelompokkan data Covid-19 ke dalam beberapa kelompok dari karakteristik setiap data.

5. Pengujian Menggunakan DBI

Tahapan ini dilakukan untuk mencari ketepatan hasil penelitian, apakah hasil penelitian baik atau buruk dilihat dari nilai DBI yang dihasilkan setiap ukuran *cluster*, semakin kecil nilai DBI maka semakin baik hasil penelitian tersebut.

6. Visualisasi

Tahapan ini bertujuan untuk membuat data menjadi peta sehingga mudah untuk dibaca dapat dilihat pada gambar 2.



Gambar 2. Metodologi Penelitian

4. PEMBAHASAN

Bagian pembahasan akan menguraikan metode yang digunakan dalam melakukan analisis.

4.1 Pengumpulan Data

Data yang digunakan dalam penelitian adalah data Covid-19 pada tingkat kabupaten/kota pada 6 provinsi, yaitu DKI Jakarta (Pemerintah Provinsi DKI Jakarta, 2021), Banten (Pemerintah Provinsi Banten, 2021), Jawa Barat (Pemerintah Provinsi Jawa Barat, 2021), Jawa Timur (Pemerintah Provinsi Jawa Timur, 2021), Jawa Tengah (Pemerintah Provinsi Jawa Tengah, 2021), dan DI Yogyakarta (Pemerintah Provinsi DI Yogyakarta,

2021). Data yang digunakan mulai tanggal 2 Maret 2020 hingga 27 Maret 2021 melalui *website* resmi pemerintah pada tingkat provinsi.

4.2 Preprocessing

Preprocessing adalah tahapan yang digunakan dalam membersihkan data dari *noise*, memperkecil ukuran dimensi, dan membuat data lebih terstruktur sehingga data dapat diolah pada tahapan selanjutnya. Salah satunya adalah menghapus data yang tidak sesuai atau tidak diperlukan terdapat dalam *dataset* yang diperoleh melalui *website* resmi pemerintah. Proses pembersihan data dan analisis algoritma K-Means *clustering* menggunakan RapidMiner Studio. Data Covid-19 dapat dilihat pada Tabel 1.

Tabel 1. Data Covid-19

No	Kabupaten /Kota	Konfirmasi	Dirawat	Sembuh	Meninggal
1	Kota Jakarta Barat	61213	1268	58818	1127
2	Kota Jakarta Pusat	37676	768	36202	706
3	Kota Jakarta Selatan	73337	1608	70490	1239
⋮	⋮	⋮	⋮	⋮	⋮
123	Kab. Pamekasan	1143	14	1043	86
124	Kab. Tuban	3360	40	2951	369
125	Kab. Pasuruan	3381	50	3070	261

Tabel 1 menunjukkan kolom konfirmasi yang tidak diperlukan sehingga kolom tersebut perlu dihapus dari *dataset*. Selain itu, terdapat 6 baris data yang tidak bisa digunakan karena data tersebut tidak termasuk ke dalam kabupaten/kota yang dianalisis. Hasil dari pembersihan data adalah terdapat 119 data dan 3 atribut.

Tahap selanjutnya melakukan normalisasi data. Karena menggunakan data yang mempunyai nilai bervariasi, membuat dimensi data menjadi tidak seimbang. Normalisasi data menggunakan metode *Min-Max* (Nishom & Fathoni, 2018), yaitu merubah nilai data antara 0 – 1 sehingga membuat data menjadi seimbang. Data setelah normalisasi dapat dilihat pada Tabel 2.

Tabel 2. Data Setelah Normalisasi

No	Kabupaten/Kota	Dirawat	Sembuh	Meninggal
1	Kota Jakarta Barat	0.232	0.674	0.634
2	Kota Jakarta Pusat	0.141	0.413	0.397
3	Kota Jakarta Selatan	0.295	0.809	0.697
⋮	⋮	⋮	⋮	⋮
117	Kab. Pamekasan	0.003	0.006	0.047
118	Kab. Tuban	0.007	0.028	0.207
119	Kab. Pasuruan	0.009	0.029	0.146

4.3 Implementasi Algoritma K-Means Clustering

Penentuan jumlah *cluster* dalam proses implementasi menggunakan 2 cara. Pertama, *cluster* akan dibagi menjadi 2 ukuran yaitu 3 dan 4. Hasil dari nilai *centroid* setiap *cluster* menentukan tingkat kerawanan persebaran Covid-19, semakin tinggi nilai *centroid* suatu *cluster* maka tingkat kerawanan semakin tinggi. Kedua, *cluster* yang sudah dibagi ke beberapa ukuran akan ditarik menjadi *cluster* yang baik menggunakan DBI untuk melihat seberapa baik *cluster* tersebut.

4.3.1 Implementasi 3 Cluster

Implementasi menggunakan 3 *cluster* menghasilkan jumlah wilayah bervariasi pada setiap *cluster*. Pada *cluster* 0 berjumlah 105 wilayah, *cluster* 1 berjumlah 7 wilayah, dan *cluster* 2 berjumlah 7 wilayah. Perbandingan nilai *centroid* dapat dilihat pada Tabel 3.

Tabel 3. Nilai Centroid Pada 3 Cluster

Cluster	Centroid
0	0.013
1	0.135
2	0.061

Tabel 3 menunjukkan *cluster* 0 adalah wilayah yang mengalami tingkat kerawanan Covid-19 terendah, *cluster* 2 adalah wilayah yang mengalami tingkat kerawanan Covid-19 sedang, dan *cluster* 1 adalah wilayah yang mengalami tingkat kerawanan Covid-19 tinggi.

4.3.2 Implementasi 4 Cluster

Implementasi menggunakan 4 *cluster* menghasilkan jumlah wilayah bervariasi pada setiap *cluster*. Pada *cluster* 0 berjumlah 105 wilayah, *cluster* 1 berjumlah 3. Perbandingan nilai *centroid* dapat dilihat pada Tabel 4.

Tabel 4. Nilai Centroid Pada 4 Cluster

Cluster	Centroid
0	0.013
1	0.034
2	0.061
3	0.082

Tabel 4 menunjukkan *cluster* 0 adalah wilayah yang mengalami tingkat kerawanan Covid-19 terendah, *cluster* 1 adalah wilayah yang mengalami tingkat kerawanan Covid-19 sedang, *cluster* 2 adalah wilayah yang mengalami tingkat kerawanan Covid-19 tinggi, dan *cluster* 3 adalah wilayah yang mengalami tingkat kerawanan Covid-19 sangat tinggi.

4.4 Pengujian Menggunakan DBI

Setelah mendapatkan hasil dari implementasi K-Means pada 3 *cluster* dan 4 *cluster* yang terjadi pada iterasi ke 5 dengan nilai yang tidak berubah, maka selanjutnya melakukan pengujian dan perbandingan nilai menggunakan DBI. Pengujian DBI dapat dilihat pada Tabel 5.

Tabel 5. Nilai DBI

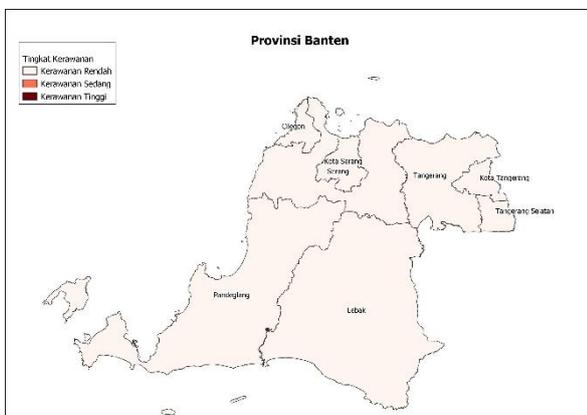
Cluster	Nilai DBI
3	0.609
4	0.698

Tabel 5 menunjukkan perbandingan nilai DBI pada 3 *cluster* dan 4 *cluster* di mana nilai DBI yang semakin kecil menunjukkan *cluster* yang dihasilkan semakin baik sehingga jumlah *cluster* terbaik dimiliki 3 *cluster* dengan nilai DBI 0.609.

4.5 Visualisasi

Hasil pengujian menggunakan DBI dibuat dalam bentuk peta per provinsi di Pulau Jawa menggunakan aplikasi QGIS (*Quantum Geographic Information System*) mendapatkan hasil seperti pada Gambar 3 – 8.

Kabupaten/kota yang berada di Provinsi Banten mempunyai tingkat kerawanan rendah, dapat dilihat pada Gambar 3.

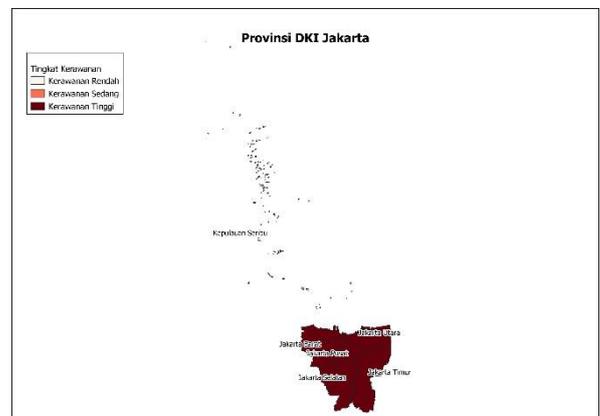


Gambar 3. Persebaran Covid-19 di Provinsi Banten

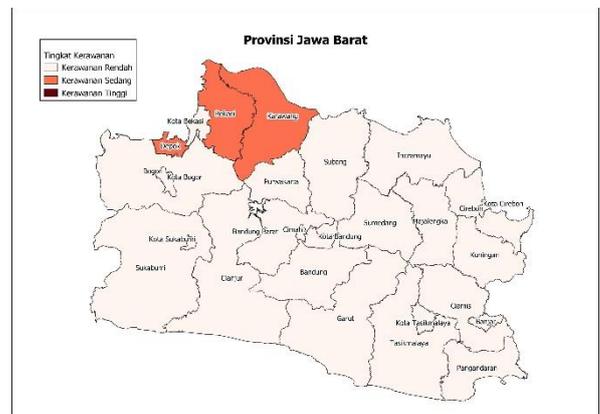
Kabupaten/kota yang berada di Provinsi DKI Jakarta mempunyai tingkat kerawanan rendah pada Kabupaten Kepulauan Seribu, dan tingkat kerawanan tinggi pada Jakarta barat, Jakarta Timur, Jakarta utara, Jakarta Pusat dan Jakarta Selatan, dapat dilihat pada Gambar 4.

Kabupaten/kota di Provinsi Jawa Barat yang mengalami tingkat kerawanan sedang berada di Karawang, Bekasi, dan Depok. Sedangkan kabupaten/kota lainnya mengalami tingkat kerawanan rendah, dapat dilihat pada Gambar 5.

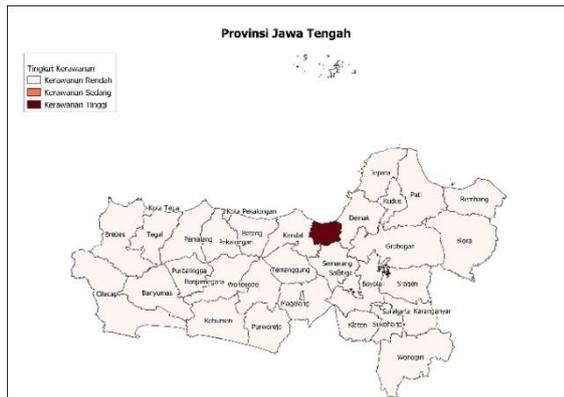
Kabupaten/kota di Provinsi Jawa Tengah yang mengalami tingkat kerawanan tinggi berada pada Kota Semarang, sedangkan kabupaten/kota lainnya berada pada tingkat kerawanan rendah, dapat dilihat pada Gambar 6.



Gambar 4. Persebaran Covid-19 di Provinsi DKI Jakarta

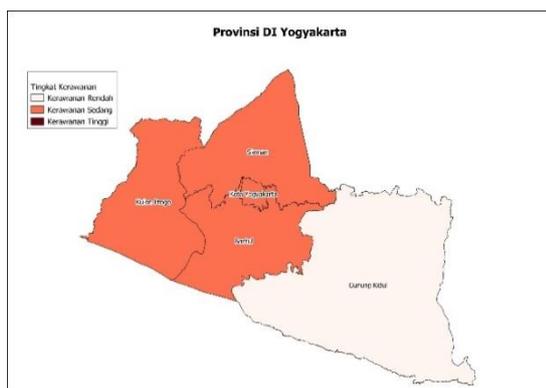


Gambar 5. Persebaran Covid-19 di Provinsi Jawa Barat



Gambar 6. Persebaran Covid-19 di Provinsi Jawa Tengah

Kabupaten/kota di Provinsi DI Yogyakarta yang mengalami tingkat kerawanan sedang berada di Yogyakarta, Sleman, Bantul dan Kulon Progo. Sedangkan kabupaten/kota yang mengalami tingkat kerawanan rendah berada di Gunung Kidul, dapat dilihat pada Gambar 7.



Gambar 7. Persebaran Covid-19 di Provinsi DI Yogyakarta

Kabupaten/kota yang berada di Provinsi Jawa Timur yang mengalami tingkat kerawanan tinggi berada di Kota Surabaya, sedangkan kabupaten/kota lainnya mengalami tingkat kerawanan rendah, dapat dilihat pada Gambar 8.



Gambar 8. Persebaran Covid-19 di Provinsi Jawa Timur

Gambar 3 – 8 merupakan data hasil implementasi K-Means yang dibuat ke dalam bentuk peta, terbagi menjadi 3 tingkat kerawanan. Pada tingkat kerawanan rendah terdapat 105 kabupaten/kota, tingkat kerawanan sedang terdapat 7 kabupaten/kota, dan tingkat kerawanan tinggi terdapat 7 kabupaten/kota.

5 KESIMPULAN

Hasil implementasi persebaran kasus Covid-19 di Pulau Jawa menggunakan algoritma K-Means clustering mendapatkan 3 cluster dari hasil pengujian menggunakan DBI dengan nilai 0,609. 3 cluster tersebut dibagi ke dalam 3 tingkat kerawanan, yaitu kerawanan rendah terdapat pada cluster 0, kerawanan sedang terdapat pada cluster 2, dan kerawanan tinggi terdapat pada cluster 1. Pada cluster 0 memiliki 105 kabupaten/kota, cluster 1 memiliki 7 kabupaten/kota, dan cluster 2 memiliki 7 kabupaten/kota. Dari peta, dapat diketahui bahwa kabupaten/kota yang memiliki tingkat kerawanan tinggi berada di Kota Jakarta Utara, Kota Jakarta Barat, Kota Jakarta Pusat, Kota Jakarta Selatan, Kota Jakarta Timur, Kota Surabaya, dan Kota Semarang.

Hasil dari penelitian diharapkan dapat digunakan sebagai acuan oleh masyarakat maupun pemerintah dalam melihat setiap daerah yang memiliki tingkat kerawanan rendah, sedang, dan tinggi dalam penyebaran Covid-19 di Pulau Jawa, agar pemerintah memfokuskan perhatian lebih kepada kabupaten/kota yang memiliki tingkat kerawanan tinggi.

6 SARAN

Penelitian selanjutnya diharapkan dalam melakukan pengelompokan data Covid-19 dapat memperluas daerah, membuat peta pada tingkat desa/kelurahan, menambahkan parameter seperti kepadatan penduduk, perpindahan penduduk dan usia, dan menambah beberapa ataupun menggunakan algoritma clustering selain K-Means untuk mendapatkan hasil yang lebih optimal.

7 DAFTAR PUSTAKA

- Adhitama, R., Burhanuddin, A. & Ananda, R. 2020. Penentuan Jumlah Cluster Ideal Smk di Jawa Tengah Dengan Metode X-Means Clustering dan K-Means Clustering. *JIKO (Jurnal Informatika dan Komputer)*. 3(1):1–5. DOI: 10.33387/jiko.
- Armijon. 2019. *PEMETAAN DIGITAL PRAKTIS*. Armijon, Ed. Bandar Lampung: Anugrah Utama Raharja.
- Banyumanik, K. 2018. Analisis Potensi Desa Berbasis Sistem Informasi Geografis (Studi Kasus: Kelurahan Sumurboto, Kecamatan Banyumanik, Kabupaten Semarang). *Jurnal Geodesi Undip*. 7(4):1–7.
- Han, J., Kamber, M. & Pei, J. 2012. *Data mining: Data mining concepts and techniques*. DOI:



- 10.1109/ICMIRA.2013.45.
- Hanafi, M.A., Ode, L., Sety, M. & Lestari, H. 2020. PEMETAAN KASUS COVID-19 DI SULAWESI TENGGARA TAHUN 2020. 1(1):11–20.
- Heraldi, H.Y., Aprilia, N.C. & Pratiwi, H. 2019. Analisis Cluster Intensitas Kebencanaan di Indonesia Menggunakan Metode K-Means. *Indonesian Journal of Applied Statistics*. 2(2):137. DOI: 10.13057/ijas.v2i2.34911.
- Irwansyah, E. & Faisal, M. 2015. *Advanced clustering teori dan aplikasi*. Yogyakarta: Deepublish.
- Kementrian Kesehatan RI. 2020. *Apa itu Coronavirus? Informasi Tentang Virus Corona / Stoppneumonia.id*. <https://stoppneumonia.id/informasi-tentang-virus-corona-novel-coronavirus/> [diakses tanggal 02 Juli 2020].
- Komite Penanganan Covid-19. 2020. *Jubir Pemerintah: Atasi COVID-19 Dengan Putus Rantai Penularan - Berita Terkini | Satgas Penanganan COVID-19*. <https://covid19.go.id/p/berita/jubir-pemerintah-atasi-covid-19-dengan-putus-rantai-penularan> [diakses tanggal 13 Oktober 2020].
- Nawrin, S., Rahatur, M. & Akhter, S. 2017. Exploreing K-Means with Internal Validity Indexes for Data Clustering in Traffic Management System. *International Journal of Advanced Computer Science and Applications*. 8(3). DOI: 10.14569/ijacsa.2017.080337.
- Nishom, M. & Fathoni, M.Y. 2018. Implementasi Pendekatan Rule-Of-Thumb untuk Optimasi Algoritma K-Means Clustering. *Jurnal Informatika: Jurnal Pengembangan IT*. 3(2):237–241. DOI: 10.30591/jpit.v3i2.909.
- Pemerintah Provinsi Banten. 2021. *Corona Provinsi Banten*. <https://infocorona.bantenprov.go.id/> [diakses tanggal 27 Maret 2021].
- Pemerintah Provinsi DI Yogyakarta. 2021. *Yogyakarta Tanggap COVID-19*. <https://corona.jogjaprov.go.id/data-statistik> [diakses tanggal 27 Maret 2021].
- Pemerintah Provinsi DKI Jakarta. 2021. *Covid-19*. <https://corona.jakarta.go.id/id/peta-persebaran> [diakses tanggal 27 Maret 2021].
- Pemerintah Provinsi Jawa Barat. 2021. *Sebaran Kasus - Pikobar [Pusat Informasi dan Koordinasi COVID-19 Jawa Barat]*. <https://pikobar.jabarprov.go.id/distribution-case> [diakses tanggal 27 Maret 2021].
- Pemerintah Provinsi Jawa Tengah. 2021. *Jateng Tanggap COVID-19*. <https://corona.jatengprov.go.id/data> [diakses tanggal 27 Maret 2021].
- Pemerintah Provinsi Jawa Timur. 2021. *JATIM TANGGAP COVID-19*. Available: <https://www.infocovid19.jatimprov.go.id/> [diakses tanggal 27 Maret 2021].
- Purnomo, K. 2021. *UPDATE 27 Maret 2021: Bertambah 4.461 Orang, Kasus Covid-19 di Indonesia Capai 1.492.002*. <https://nasional.kompas.com/read/2021/03/27/1644-0421/update-27-maret-2021-bertambah-4461-orang-kasus-covid-19-di-indonesia-capai> [diakses tanggal 04 April 2021].
- Ridlo, M.R., Defiyanti, S. & Primajaya, A. 2017. Implementasi Algoritme K-Means Untuk Pemetaan Produktivitas Panen Padi Di Kabupaten Karawang. *Citee 2017*. 426–433.
- Santosa, B. 2007. *Data Mining: Teknik Pemanfaatan Data untuk Keperluan Bisnis*. Yogyakarta: Graha Ilmu.
- Sholihah, S.A. 2021. Analisis Cluster untuk Pemetaan Data Kasus Covid-19 di Indonesia Menggunakan K-Means.
- 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. DOI: 10.21111/fij.v5i2.4905.
- Tim Detikcom. 2021. *Sebaran 4.461 Kasus Baru Corona di RI 27 Maret, DKI Tertinggi*. <https://news.detik.com/berita/d-5510377/sebaran-4461-kasus-baru-corona-di-ri-27-maret-dki-tertinggi> [diakses tanggal 27 Maret 2021].
- Yusuf, H. & Halim, H. 2014. *Survey Dan Pemetaan*. 1st ed. Yogyakarta: Deepublish.