

**IMPLEMENTASI GAK-MEANS CLUSTERING UNTUK
PENGELOMPOKAN DAERAH PENYEBARAN COVID-19 DI PROVINSI
JAWA TIMUR**

SKRIPSI



**UIN SUNAN AMPEL
S U R A B A Y A**

Disusun Oleh
ZULFATUL AIZZAH
H92218054

**PROGRAM STUDI MATEMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI SUNAN AMPEL
SURABAYA**

2022

PERNYATAAN KEASLIAN

Saya yang bertanda tangan di bawah ini,

Nama : ZULFATUL AIZZAH

NIM : H92218054

Program Studi : Matematika

Angkatan : 2018

Menyatakan bahwa saya tidak melakukan plagiat dalam penulisan skripsi saya yang berjudul "IMPLEMENTASI GA-K MEANS CLUSTERING UNTUK PENGELOMPOKAN DAERAH TERDAMPAK COVID-19 DI PROVINSI JAWA TIMUR". Apabila suatu saat nanti terbukti saya melakukan tindakan plagiat, maka saya bersedia menerima sanksi yang telah ditetapkan.

Demikian pernyataan keaslian ini saya buat dengan sebenar-benarnya.

Surabaya, 17 November 2021

Yang menyatakan,



ZULFATUL AIZZAH
NIM. H92218054

LEMBAR PERSETUJUAN PEMBIMBING

Skripsi oleh

Nama : ZULFATUL AIZZAH

NIM : H92218054

Judul Skripsi : IMPLEMENTASI GAK-MEANS CLUSTERING UNTUK
PENGELOMPOKAN DAERAH PENYEBARAN
COVID-19 DI PROVINSI JAWA TIMUR

telah diperiksa dan disetujui untuk diujikan.

Pembimbing I



Nurissaidah Ulinuha, M. Kom
NIP. 199007021014032004

Pembimbing II



Dr. Abdulloh Hamid, M. Pd
NIP. 198508282014031003

Mengetahui,
Ketua Program Studi Matematika
UIN Sunan Ampel Surabaya



Aris Fanani, M. Kom
NIP. 198701272014031002

PENGESAHAN TIM PENGUJI SKRIPSI

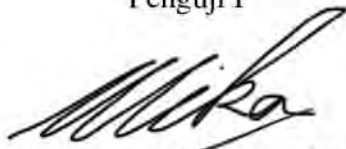
Skripsi oleh

Nama : ZULFATUL AIZZAH
NIM : H92218054
Judul Skripsi : IMPLEMENTASI GAK-MEANS CLUSTERING UNTUK
PENGELOMPOKAN DAERAH PENYEBARAN
COVID-19 DI PROVINSI JAWA TIMUR

Telah dipertahankan di depan Tim Penguji
pada tanggal 25 Januari 2022

Mengesahkan,
Tim Penguji

Penguji I



Wika Dianita Utami, M.Sc
NIP. 199206102018012003

Penguji II



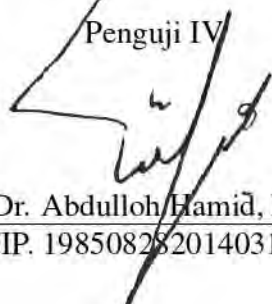
Putro Keumala Intan, M.Si
NIP. 198805282018012001

Penguji III



Nurissaidah Ulinnaha, M. Kom
NIP. 199011022014032004

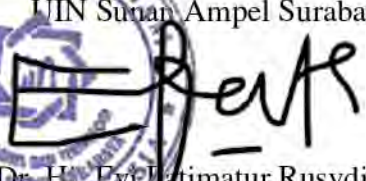
Penguji IV



Dr. Abdulloh Hamid, M. Pd
NIP. 198508282014031003



Mengetahui,
Dekan Fakultas Sains dan Teknologi
UIN Sunan Ampel Surabaya


Prof. Dr. H. Evi Fatmatur Rusydiyah, M.Ag
NIP. 197312272005012003



KEMENTERIAN AGAMA
UNIVERSITAS ISLAM NEGERI SUNAN AMPEL SURABAYA
PERPUSTAKAAN

Jl. Jend. A. Yani 117 Surabaya 60237 Telp. 031-8431972 Fax. 031-8413300
E-Mail: perpus@uinsby.ac.id

LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI
KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS

Sebagai sivitas akademika UIN Sunan Ampel Surabaya, yang bertanda tangan di bawah ini, saya:

Nama : ZULFATUL AIZZAH
NIM : H92218054
Fakultas/Jurusan : SAINTEK / MATEMATIKA
E-mail address : ZULFARINA09@GMAIL.COM

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Perpustakaan UIN Sunan Ampel Surabaya, Hak Bebas Royalti Non-Eksklusif atas karya ilmiah :

☒ Sekripsi ☐ Tesis ☐ Desertasi ☐ Lain-lain (.....)
yang berjudul :

IMPLEMENTASI GA-K MEANS CLUSTERING UNTUK

PENGELOMPOKAN DAERAH PENYEBARAN COVID-19

DI PROVINSI JAWA TIMUR

beserta perangkat yang diperlukan (bila ada). Dengan Hak Bebas Royalti Non-Eksklusif ini Perpustakaan UIN Sunan Ampel Surabaya berhak menyimpan, mengalih-media/format-kan, mengelolanya dalam bentuk pangkalan data (database), mendistribusikannya, dan menampilkan/mempublikasikannya di Internet atau media lain secara *fulltext* untuk kepentingan akademis tanpa perlu meminta ijin dari saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan atau penerbit yang bersangkutan.

Saya bersedia untuk menanggung secara pribadi, tanpa melibatkan pihak Perpustakaan UIN Sunan Ampel Surabaya, segala bentuk tuntutan hukum yang timbul atas pelanggaran Hak Cipta dalam karya ilmiah saya ini.

Demikian pernyataan ini yang saya buat dengan sebenarnya.

Surabaya, 25 JANUARI 2022

Penulis

(ZULFATUL AIZZAH)
nama terang dan tanda tangan

ABSTRAK

IMPLEMENTASI GAK-MEANS CLUSTERING UNTUK PENGELOMPOKAN DAERAH PENYEBARAN COVID-19 DI PROVINSI JAWA TIMUR

Saat ini COVID-19 masih menjadi topik terhangat di seluruh dunia tidak terkecuali Indonesia. Salah satu provinsi di Indonesia yang memiliki tren kasus COVID-19 tinggi adalah Provinsi Jawa Timur. Demikian maka penting adanya pengelompokan daerah di Provinsi Jawa Timur untuk memberikan informasi kepada masyarakat maupun pemerintah terkait penyebaran COVID-19 di Provinsi Jawa Timur. Metode yang digunakan adalah gabungan dari metode Algoritma Genetika (GA) dan k-Means. Metode k-Means dipilih karena prosesnya yang sederhana dengan komputasi yang ringan. Namun, hasil dari perhitungan metode k-Means memiliki kesensitifan karena pusat klaster awalnya ditentukan secara acak sehingga algoritma genetika digunakan untuk mengoptimasi nilai pusat klaster awal pada k-Means. Pengelompokan wilayah diterapkan pada data sebaran bulan Januari-November 2021. Pada implementasinya, perhitungan metode GAK-Means mendapatkan evaluasi yang lebih tinggi dibandingkan dengan perhitungan k-Means tanpa optimasi serta memiliki iterasi k-Means yang lebih cepat. Evaluasi yang digunakan adalah *Silhouette Coefficient* (SC). Berdasarkan uji coba parameter, hasil evaluasi tertinggi yang diperoleh yakni sebesar SC GAK-Means = 0.85875 dan SC k-Means = 0.52460. Pengelompokan dilakukan dengan jumlah klaster sebanyak $k = 2$ yakni klaster merah untuk daerah dengan tingkat kerawanan yang lebih tinggi daripada klaster hijau.

Kata kunci: COVID-19, Algoritma Genetika, Algoritma K-Means

ABSTRACT

IMPLEMENTATION OF GAK-MEANS CLUSTERING FOR GROUPING DISTRIBUTION AREAS OF COVID-19 IN EAST JAVA PROVINCE

Currently, COVID-19 is still the hottest topic around the world, don't expect Indonesia. One of the provinces in Indonesia that has a high trend of COVID-19 cases is East Java Province. Thus, it is important to group regions in East Java Province to provide information to the public and the government regarding the spread of COVID-19 in East Java Province. The method used is a combination of the Genetic Algorithm (GA) and k-Means methods. The k-Means method was chosen because the process is simple and computationally light. However, the results of the calculation of the k-Means method have sensitivity because the initial cluster center is determined randomly so that the algorithm used to optimize the initial cluster center value on k-Means. Regional grouping in the January-November 2021 distribution data. In its implementation, the calculation of the GAK-Means evaluation is higher than the calculation of k-Means without optimization and has a faster k-Means iteration. The evaluation used is Silhouette Coefficient (SC). Based on the parameter testing, the highest evaluation results obtained are SC GAK-Means = 0.85875 and SC k-Means = 0.52460. The grouping is done with the number of clusters as much as $k = 2$, namely red clusters for areas with a higher level of vulnerability than green clusters.

Keywords: COVID-19, Genetic Algorithm, K-Means Algorithm

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERNYATAAN KEASLIAN	ii
LEMBAR PERSETUJUAN PEMBIMBING	iii
PENGESAHAN TIM PENGUJI SKRIPSI	iv
PERSETUJUAN PUBLIKASI	v
ABSTRAK	vi
ABSTRACT	vii
DAFTAR ISI	1
DAFTAR TABEL	4
DAFTAR GAMBAR	6
I PENDAHULUAN	7
1.1. Latar Belakang Masalah	7
1.2. Rumusan Masalah	13
1.3. Tujuan Penelitian	13
1.4. Manfaat Penelitian	13
1.5. Batasan Masalah	14
1.6. Sistematika Penulisan	14
II TINJAUAN PUSTAKA	16
2.1. COVID-19 di Jawa Timur	16
2.2. Normalisasi Data	18
2.3. Uji Multikolinearitas	18
2.4. Analisis Klaster	20
2.5. Algoritma K-Means	20
2.6. Algoritma Genetika	22
2.6.1. Membangkitkan Populasi Acak	23
2.6.2. Fungsi Fitness	24
2.6.3. Seleksi Individu	24

2.6.4. Proses <i>Crossover</i>	25
2.6.5. Proses Mutasi Gen	27
2.6.6. Elitism	28
2.7. Algoritma Genetika dan K-Means	28
2.8. Evaluasi Klaster	29
2.9. Wabah Penyakit dan Integrasi Keislaman	31
III METODE PENELITIAN	36
3.1. Jenis Penelitian	36
3.2. Sumber Data	36
3.3. Variabel Penelitian	36
3.4. Kerangka Penelitian	37
3.4.1. <i>Pre-processing</i>	37
3.4.2. Implementasi GA dan K-Means	38
3.4.3. Uji Coba Parameter	39
3.4.4. Pemetaan Klaster	40
3.4.5. Analisis Hasil	40
IV HASIL DAN PEMBAHASAN	42
4.1. Deskripsi Data	42
4.2. <i>Pre-processing</i>	44
4.2.1. Normalisasi Data	44
4.2.2. Uji Multikolinearitas	45
4.3. Implementasi GA dan K-Means	50
4.3.1. Proses Algoritma Genetika	50
4.3.2. Proses K-Means	66
4.3.3. Proses Evaluasi Klaster	76
4.4. Uji Coba Parameter	85
4.4.1. Uji Coba Banyaknya Popsi	85
4.4.2. Uji Coba Banyaknya Generasi	86
4.4.3. Uji Coba Jumlah Klaster (GAK-Means)	87
4.4.4. Uji Coba Jumlah Klaster (K-Means)	87

4.5. Pemetaan Klaster	88
4.6. Analisis Hasil	98
4.7. Integrasi Keislaman	100
V PENUTUP	104
5.1. Kesimpulan	104
5.2. Saran	105
DAFTAR PUSTAKA	105
LAMPIRAN	113

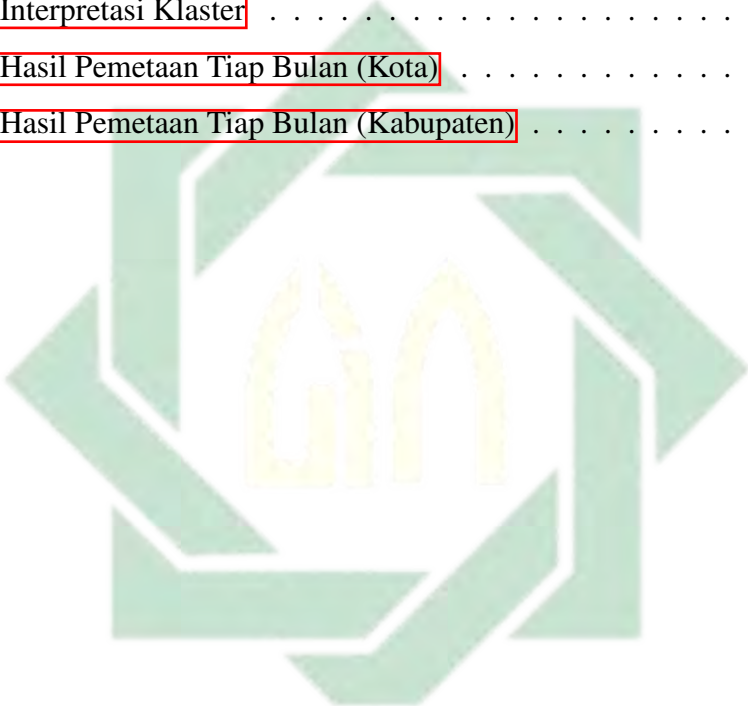


UIN SUNAN AMPEL
S U R A B A Y A

DAFTAR TABEL

2.1 Struktur Kategori <i>Silhouette Coefficient</i>	30
3.1 Variabel Penelitian	37
4.1 Perhitungan x_1 dan x_2 untuk Uji Multikolinearitas	46
4.2 Hasil Perhitungan Uji Multikolinearitas	48
4.3 Pembangkitan Individu dalam Populasi	50
4.6 Perhitungan Jarak <i>Manhattan Distance</i>	53
4.7 Nilai Fitness pada Semua Individu	55
4.8 Perhitungan Probabilitas dan Kumulatif	56
4.9 Proses Seleksi Individu	58
4.17 Populasi setelah Seleksi dan Reproduksi Gen	63
4.18 Populasi setelah Elitism	65
4.19 Pusat Klaster Terpilih	66
4.20 Pusat Klaster Awal (GAK-Means Iterasi = 1)	66
4.22 Perhitungan Algoritma GAK-Means (Iterasi = 1)	68
4.23 Pusat Klaster Baru (GAK-Means Iterasi = 2)	71
4.24 Perhitungan Algoritma GAK-Means (Iterasi = 2)	71
4.25 Pengelompokan Akhir GAK-Means	73
4.26 Pusat Klaster Awal (K-Means Iterasi = 1)	74
4.27 Pengelompokan K-Means (Iterasi = 1)	74
4.28 Pusat Klaster Baru (K-Means Iterasi = 2)	75
4.29 Pengelompokan K-Means (Iterasi = 2)	75
4.30 Pengelompokan Akhir K-Means	76
4.31 Perhitungan Jarak Euclidean (Kab. Tulungagung)	77
4.32 Nilai SC pada Tiap Data (GAK-Means)	79
4.33 Nilai SC pada Tiap Data (K-Means)	82
4.34 Uji Coba <i>Popsize</i>	85

4.35 Uji Coba Generasi (Popsize = 80)	86
4.36 Uji Coba Generasi (Popsize = 160)	86
4.37 Uji Coba Jumlah Klaster (GAK-Means)	87
4.38 Uji Coba Jumlah Klaster (K-Means)	88
4.39 Pusat Klaster Akhir sebelum Denormalisasi	89
4.40 Pusat Klaster Akhir setelah Denormalisasi	89
4.41 Interpretasi Klaster	89
4.42 Hasil Pemetaan Tiap Bulan (Kota)	90
4.43 Hasil Pemetaan Tiap Bulan (Kabupaten)	90



UIN SUNAN AMPEL
S U R A B A Y A

DAFTAR GAMBAR

2.1	Sebaran Kasus COVID-19 di Jawa Timur	17
2.2	Sebelum (kiri) dan Sesudah (kanan) Klastering	21
2.3	Representasi dalam Satu Populasi	23
2.4	Ilustrasi Mesin <i>roulette</i>	25
2.5	Proses <i>Crossover</i> 3-titik	26
2.6	Proses Elitism	28
3.1	Diagram Alir Implementasi GA dan K-Means	39
4.1	Penambahan Kasus COVID-19 di Provinsi Jawa Timur	42
4.2	Penyebaran COVID-19 di Provinsi Jawa Timur (Bulan November 2021)	43
4.3	Pemetaan Bulan Januari 2021	92
4.4	Pemetaan Bulan Februari 2021	92
4.5	Pemetaan Bulan Maret 2021	93
4.6	Pemetaan Bulan April 2021	93
4.7	Pemetaan Bulan Mei 2021	94
4.8	Pemetaan Bulan Juni 2021	95
4.9	Pemetaan Bulan Juli 2021	96
4.10	Pemetaan Bulan Agustus 2021	96
4.11	Pemetaan Bulan September 2021	97
4.12	Pemetaan Bulan Oktober 2021	97
4.13	Pemetaan Bulan November 2021	98

BAB I

PENDAHULUAN

1.1. Latar Belakang Masalah

Pada akhir Desember 2019 lalu, seluruh dunia digemparkan dengan ditemukannya pneumonia serius yang disebabkan oleh mutasi jenis baru dari *coronavirus* di China (Ulya, 2020). WHO secara resmi mengumumkan virus mutasi tersebut dengan sebutan *Corona Virus Diase 2019* (COVID-19). Virus mutasi baru ini menyerang sistem pernafasan manusia dan menyebar dengan sangat cepat (Harini, 2020). Hal ini dapat diamati pada penyebaran yang awalnya hanya berada di Provinsi Hubei, kemudian menuju provinsi di sekitar Hubei, sampai satu negara China (Wang et al., 2020). Tidak berhenti disana, virus COVID-19 menyebar menuju skala global dalam waktu yang singkat dan dinyatakan sebagai penyakit pandemik oleh WHO pada 11 Maret 2020. Pada akhir Agustus lalu dalam skala global, kasus terkonfirmasi dilaporkan melebihi 200 juta sedangkan kasus kematian mencapai angka 4 juta (WHO, 2021).

Berbagai negara telah memiliki aturan masing-masing untuk menurunkan laju kenaikan kasus COVID-19 yang begitu cepat tidak terkecuali Indonesia (Arswendi, 2021). Pemerintah Indonesia telah memaksimalkan berbagai upaya untuk menindaklanjuti wabah COVID-19. Program pemerintah meliputi 5M (menggunakan masker, rajin mencuci tangan, menghindari adanya kerumunan, menjaga jarak, mengurangi mobilitas bepergian), 3T (*testing, tracking, treatment*), pembatasan sosial berskala besar (PSBB), serta vaksinasi (Abidin et al., 2021). Hal

ini terus dilakukan untuk menekan penambahan kasus baru. Indonesia terdiri dari 34 provinsi dengan karakteristik yang berbeda-beda dan laju penambahan kasus yang berbeda pula. Terdapat beberapa provinsi dengan tren kasus rendah, sedang, maupun tinggi.

Wabah penyakit menular merupakan takdir serta musibah yang tidak dapat dihindari oleh manusia. Namun Allah Swt akan segera menolong dan memberikan petunjuk pada orang-orang beriman. Allah Swt berfirman dalam al-Qur'an surah At-Taghabun ayat 11 dengan redaksi sebagai berikut:

مَا أَصَابَ مِنْ مُصِيبَةٍ إِلَّا بِإِذْنِ اللَّهِ وَمَنْ يُؤْمِنْ بِاللَّهِ يَهْدِ اللَّهُ قَلْبَهُ وَاللَّهُ بِكُلِّ شَيْءٍ عَلِيمٌ ﴿١١﴾

Artinya: “Tidak ada suatu musibah yang menimpa (seseorang), kecuali atas izin Allah. Dan barangsiapa beriman kepada Allah, niscaya Allah akan memberi petunjuk melalui hatinya. Dan Allah Maha Mengetahui segala sesuatu.” (QS. At Taghabun : 11)

Wabah penyakit yang menular sudah pernah terjadi pada zaman dahulu. Rasulullah Saw memberikan anjuran apa yang harus dilakukan kaum muslim dalam menghadapi wabah penyakit. Hal ini dapat diamati melalui hadis Rasulullah Saw yang diriwayatkan oleh perawi masyhur yakni Bukhari, Nasa'I, dan Ahmad. Hadis tersebut adalah sebagai berikut:

عَنْ عَائِشَةَ، أَنَّهَا قَالَتْ: سَأَلْتُ رَسُولَ اللَّهِ صَلَّى اللَّهُ عَلَيْهِ وَسَلَّمَ عَنِ الطَّاعُونَ فَأَخْبَرَنِي رَسُولُ اللَّهِ صَلَّى اللَّهُ عَلَيْهِ وَسَلَّمَ: أَنَّهُ كَانَ عَذَابًا يَبْعَثُهُ اللَّهُ عَلَى مَنْ يَشَاءُ، فَيَجْعَلُهُ رَحْمَةً لِلْمُؤْمِنِينَ، فَلَيْسَ مِنْ رَجُلٍ يَقَعُ الطَّاعُونَ، فَيَمُوتُ فِي بَيْتِهِ صَابِرًا مُحْتَسِبًا يَعْلَمُ أَنَّهُ لَا يُصِيبُهُ إِلَّا مَا كَتَبَ اللَّهُ لَهُ إِلَّا كَانَ لَهُ مِثْلُ أَجْرِ الشَّهِيدِ (رواه البخاري و احمد و النسائي)

Artinya: “Dari Siti Aisyah ra, ia berkata, ‘Aku bertanya kepada Rasulullah Saw mengenai *tha’un*, lalu Rasulullah mengabarkan kepadaku, ‘Dahulu, *tha’un* adalah azab yang Allah kirimkan kepada siapa saja yang Dia kehendaki, tetapi Allah menjadikannya sebagai rahmat bagi orang beriman. Maka tiada seorang pun yang tertimpa *tha’un*, kemudian ia menahan diri di rumah dengan sabar serta mengharap ridha-Nya seraya menyadari bahwa *tha’un* tidak akan menimpanya kecuali telah menjadi ketentuan Allah untuknya, niscaya ia akan memperoleh ganjaran seperti pahala orang yang mati syahid.’” (HR. Bukhari, Nasa’I, dan Ahmad).

Pada hadis di atas, dijelaskan bahwa *tha’un* atau wabah penyakit menular merupakan azab atau peringatan dari Allah Swt kepada umat manusia yang dikehendakinya. Namun bagi orang-orang yang beriman, wabah penyakit tersebut merupakan rahmat dari Allah Swt. Saat menghadapi adanya wabah penyakit, Rasulullah Saw memerintahkan untuk bersabar dan menahan diri di rumah (Irham et al., 2020). Demikian sama halnya dengan program pemerintah yang telah dijalankan seperti pembatasan sosial di beberapa daerah, penutupan tempat ibadah, dan seruan untuk tetap di rumah saja. Dengan membatasi mobilitas serta menjaga jarak antar sesama, maka secara tidak langsung dapat mengurangi resiko menularkan maupun tertular COVID-19 (Mahmud, 2020).

Upaya-upaya tersebut merupakan beberapa upaya yang dilakukan untuk mengurangi laju penyebaran wabah COVID-19. Dalam hal ini bersesuaian dengan kaidah fikih mengenai kewajiban seorang muslim untuk mencegah bahaya lebih diutamakan daripada menarik kebaikan (Ridho, 2020).

دَرْءُ الْمَفَاسِدِ مُقَدَّمٌ عَلَى جَلْبِ الْمَصَالِحِ

Artinya: “Menolak kerusakan itu didahulukan daripada menarik kebaikan.”

Pada 23 Agustus 2021, Indonesia menjadi negara dengan kasus tertinggi urutan ke-4 di Asia dengan total lebih dari 3,9 juta kasus terkonfirmasi. Tingginya kasus terkonfirmasi di Indonesia merupakan akumulasi dari kasus terkonfirmasi di 34 provinsi yang mana salah satu provinsi dengan kasus yang tinggi adalah Provinsi Jawa Timur. Diakses melalui kawalcovid19.id pada 28 Agustus 2021 lalu, angka kenaikan kasus positif di Provinsi Jawa Timur menyalip Provinsi Jawa Barat yang sebelumnya berada di peringkat satu (KawalCovid, 2021).

Penyebaran COVID-19 yang cukup tinggi di daerah Provinsi Jawa Timur memerlukan adanya pengelompokan daerah-daerah penyebaran COVID-19. Dengan tujuan untuk memberikan informasi pada pemerintah maupun masyarakat di berbagai daerah terkait penyebaran COVID-19. Dengan demikian maka baik pemerintah ataupun masyarakat dapat mendapatkan informasi mengenai daerah mana yang menjadi rawan sehingga dapat dilakukan tindakan lanjutan untuk menurunkan laju penambahan kasus.

Beberapa peneliti sebelumnya telah melakukan penelitian mengenai pengelompokan daerah penyebaran COVID-19. Seperti penelitian yang ditulis oleh Nabila et al. (2021). Peneliti melakukan klustering menggunakan metode k-Means untuk mengelompokkan kasus spesifik di Provinsi Lampung. Daerah dibagi menjadi 4 kluster dengan keterangan warna meliputi zona merah, oranye, kuning, dan hijau. Validasi yang digunakan adalah *Davies-Bouldin Coefficient* (DBI). Berdasarkan evaluasi DBI, hasil yang didapatkan menunjukkan bahwa pengelompokan sudah sangat baik ditunjukkan dengan nilai yang didapatkan mendekati nol (Nabila et al., 2021).

Selain itu, pengelompokan daerah penyebaran COVID-19 pada provinsi lain dengan metode yang sama dilakukan oleh Mahmudan (2020). Peneliti menggunakan Provinsi Jawa Tengah sebagai studi kasus penelitian dalam mengelompokkan daerah kota atau kabupaten. Penelitian tersebut mengelompokkan daerah Jawa Tengah menjadi 3 klaster. Hasil yang didapatkan adalah dari 35 kota/kabupaten di Jawa Tengah, 33 daerah berada dalam klaster 1, 1 daerah berada dalam klaster 2 yakni Kota Semarang, dan 1 daerah sisanya berada dalam klaster 3 yakni Kabupaten Jepara (Mahmudan, 2020).

Pada kedua penelitian yang telah disebutkan sebelumnya, peneliti sama-sama menggunakan metode k-Means dan menghasilkan evaluasi yang baik dalam pengelompokan. Metode k-Means adalah metode yang sederhana, praktis, dan memiliki keakuratan yang baik dalam pengklasteran data. Namun, algoritma k-Means memiliki kekurangan yaitu inisialisasi nilai pusat klaster awal yang bersifat random dimana hal ini sangat berpengaruh pada hasil pengelompokan (Mursalim et al., 2021). Selain itu, iterasi yang dihasilkan pada metode k-Means cukup banyak seiring banyaknya data yang digunakan. Dengan demikian perlu dilakukan teknik optimasi dengan harapan dapat mempercepat proses pengklasteran namun hasil tetap maksimal (Zeebaree et al., 2017).

Dari berbagai metode optimasi, algoritma genetika (GA) adalah metode yang sering dijumpai (Santoso et al., 2017). Algoritma genetika merupakan algoritma optimasi yang prinsip dasar perhitungannya menggunakan seleksi alam melalui teori evolusi Darwin (Mursalim et al., 2021). Algoritma genetika sangat mumpuni dalam menyelesaikan permasalahan kompleks yang memiliki banyak variabel (Santoso et al., 2017). Selain itu, algoritma genetika mampu menangani bermacam-macam optimasi tergantung pada fungsi objektifnya (fitness) serta

memiliki kemampuan untuk mengeksplorasi ruang pencarian yang cukup besar (Prasetio et al., 2018).

Penelitian menggunakan gabungan metode GA dan k-Means telah dilakukan oleh Irfan et al. (2018). Pada penelitian tersebut ditemukan bahwa jumlah iterasi berkurang secara keseluruhan dan dengan demikian menghemat waktu dalam membentuk kluster untuk mencari informasi yang dibutuhkan (Irfan et al., 2018). Sedangkan pada penelitian Santoso et al. (2017) dilakukan pengelompokan kinerja akademik dosen dengan menggunakan metode yang sama. Berdasarkan nilai evaluasi *Silhouette Coefficient* (SC), gabungan metode GA dan k-Means mendapatkan nilai SC lebih tinggi daripada pengelompokan dengan metode k-Means saja. Nilai evaluasi yang didapatkan cukup baik yakni SC k-Means = 0.3803 dan SC GAK-Means = 0.4352 (Santoso et al., 2017).

Penelitian lain dengan menggunakan metode yang sama dilakukan oleh Kuntjoro et al. (2018). Peneliti melakukan pengelompokan pada data tsunami sebagai upaya mitigasi bencana. Berdasarkan perbandingan nilai *Silhouette Coefficient*, hasil pengujian pada klustering menggunakan gabungan algoritma GA dan k-Means yang telah dilakukan memiliki kualitas kluster lebih tinggi yaitu sebesar 0.9959 dibandingkan dengan algoritma k-Means tanpa algoritma genetika dengan nilai SC sebesar 0.8831 (Kuntjoro et al., 2018).

Berdasarkan penjelasan sebelumnya dengan mempertimbangkan kelebihan serta kekurangan metode k-Means, maka peneliti bermaksud untuk melakukan pengelompokan daerah-daerah penyebaran COVID-19 dengan mengimplementasikan gabungan metode GA dan k-Means di Provinsi Jawa Timur. Hal ini bertujuan untuk memberikan informasi pada pemerintah maupun masyarakat di berbagai daerah Provinsi Jawa Timur terkait penyebaran COVID-19

sehingga dapat dilakukan penindaklanjutan pada daerah dengan klaster tertentu.

1.2. Rumusan Masalah

Berdasarkan latar belakang permasalahan yang telah dijabarkan pada penelitian ini, diperoleh rumusan masalah sebagai berikut:

1. Bagaimana hasil evaluasi *Silhouette Coefficient* dalam pengklasteran pada data penyebaran COVID-19 di Jawa Timur?
2. Bagaimana hasil analisis klaster menggunakan metode gabungan GA dan k-Means pada data penyebaran COVID-19 di Jawa Timur?

1.3. Tujuan Penelitian

Dengan berdasarkan rumusan masalah yang telah didapatkan sebelumnya, maka pada penelitian ini diperoleh tujuan penelitian sebagai berikut:

1. Untuk mengetahui hasil evaluasi *Silhouette Coefficient* dalam pengklasteran pada data penyebaran COVID-19 di Jawa Timur.
2. Untuk mengetahui hasil analisis klaster menggunakan metode gabungan GA dan k-Means pada data penyebaran COVID-19 di Jawa Timur.

1.4. Manfaat Penelitian

Dengan berdasarkan pada tujuan penelitian yang telah dijabarkan sebelumnya, penelitian ini diharapkan dapat memberikan manfaat sebagai berikut:

1. Manfaat Teoritis

Berdasarkan teori, penelitian ini dapat menyalurkan pengetahuan mengenai pengelompokan daerah-daerah penyebaran COVID-19 menggunakan

gabungan dari metode GA dan k-Means serta dapat bermanfaat bagi para pembaca.

2. Manfaat Praktis

Sedangkan dalam manfaat praktisnya melalui penelitian ini, penulis berharap dapat menginformasikan kepada masyarakat umum serta pemerintah terkait pengelompokan daerah-daerah penyebaran COVID-19 sehingga dapat dilakukan tindakan lanjutan untuk menekan laju penambahan kasus. Selain itu, penelitian ini dapat menambah literatur untuk universitas serta dapat dijadikan acuan pada penelitian selanjutnya.

1.5. Batasan Masalah

Adapun batasan masalah dari penelitian ini adalah sebagai berikut:

1. Data yang digunakan adalah data penyebaran COVID-19 di Provinsi Jawa Timur pada bulan Januari – November 2021.
2. Uji coba parameter yang digunakan adalah uji coba banyaknya generasi, jumlah n populasi, serta inisialisasi jumlah klaster.

1.6. Sistematika Penulisan

Sistematika penulisan ini disusun untuk mempermudah dalam proses penulisan penelitian dengan membagi penelitian menjadi beberapa bab. Berikut adalah sistematika yang digunakan pada penelitian ini:

1. BAB I PENDAHULUAN

Pada bab pendahuluan akan diuraikan mengenai pendahuluan penelitian yang terdiri dari enam subbab meliputi latar belakang, rumusan masalah, tujuan penelitian, manfaat penelitian, batasan masalah, dan sistematika penulisan.

2. BAB II TINJAUAN PUSTAKA

Pada bab kajian pustaka akan diuraikan dengan rinci menggunakan literatur buku, jurnal ilmiah, dan sumber kepustakaan lain, serta hasil penelitian terdahulu yang relevan dengan penelitian ini.

3. BAB III METODE PENELITIAN

Pada bab metodologi penelitian akan diuraikan tahapan-tahapan penelitian yang meliputi jenis penelitian yang digunakan, sumber data yang digunakan, variabel-variabel yang digunakan, dan kerangka penelitian.

4. BAB IV HASIL DAN PEMBAHASAN

Pada bab hasil dan pembahasan akan diuraikan mengenai deskripsi data, implementasi GA dan k-Means, uji coba parameter, pemetaan tiap daerah, dan analisis hasil pemetaan.

5. BAB V PENUTUP

Pada bab penutup akan diuraikan mengenai kesimpulan yang didapatkan pada penelitian ini. Serta akan dituliskan saran untuk penelitian selanjutnya.

UIN SUNAN AMPEL
S U R A B A Y A

BAB II

TINJAUAN PUSTAKA

2.1. COVID-19 di Jawa Timur

Novel Coronavirus 2019 atau biasa dikenal COVID-19 merupakan virus yang diidentifikasi adanya di Provinsi Hubei China pada akhir Desember 2019 (Narin et al., 2021; Salman et al., 2020). Virus COVID-19 menyerang pernafasan dan pada umumnya menyebar melalui kontak dekat serta melalui droplet yang dikeluarkan ketika batuk atau bersin. Laju penyebaran virus COVID-19 sangat cepat, hal ini dapat diamati pada penyebaran yang meluas berawal dari negara China sampai memasuki negara-negara Asia termasuk negara Indonesia. COVID-19 secara resmi diumumkan sebagai penyakit pandemik oleh *World Health Organization* (WHO) yang bertepatan pada tanggal 11 Maret 2020 (Djalante et al., 2020).

Pemerintah Indonesia telah membuat beberapa kebijakan untuk menindaklanjuti wabah COVID-19 (Arswendi, 2021). Program dan seruan pemerintah meliputi 5M (menggunakan masker, mencuci tangan dengan rajin, menghindari adanya kerumunan, menjaga jarak, mengurangi mobilitas bepergian), 3T (*testing, tracking, treatment*), pembatasan sosial berskala besar (PSBB), serta vaksinasi (Abidin et al., 2021). Hal ini terus dilakukan untuk mengurangi penambahan kasus baru. Indonesia sendiri terdiri dari 34 provinsi dengan karakteristik yang berbeda-beda dan laju penambahan kasus yang berbeda pula. Terdapat beberapa provinsi dengan tren kasus rendah, sedang, maupun tinggi. Dan

Sampai akhir 2020 berdasarkan data Badan Pusat Statistik (BPS) Provinsi Jawa Timur, luas dari seluruh Provinsi Jawa Timur adalah $47.799,75 \text{ km}^2$. Daerah di Provinsi Jawa Timur terpecah menjadi 9 kota dan 29 kabupaten. kota/kabupaten dengan daerah terkecil adalah Kota Mojokerto ($16,47 \text{ km}^2$) sedangkan daerah terluas adalah Kabupaten Banyuwangi ($5.782,4 \text{ km}^2$). Berdasarkan hasil sensus penduduk pada tahun 2020, jumlah penduduk Provinsi Jawa Timur diperkirakan berjumlah sekitar 40,666 juta jiwa (BPS, 2021). Dengan artian lain, saat ini Provinsi Jawa Timur merupakan provinsi dengan jumlah penduduk yang sangat padat di Indonesia dan menempati posisi kedua setelah Provinsi Jawa Barat (48,274 juta penduduk).



Provinsi Jawa Timur pernah diklaim menduduki peringkat keempat dengan jumlah kasus konfirmasi COVID-19 paling banyak di Indonesia (Yustanti et al., 2020). Diakses melalui kawalcovid19.id pada 28 Agustus lalu, angka kenaikan kasus positif di Provinsi Jawa Timur menyalip Provinsi Jawa Barat yang

sebelumnya berada di peringkat satu dengan jumlah sebanyak 1.409 penambahan kasus baru. Penyebaran kasus COVID-19 yang tinggi di Provinsi Jawa Timur memerlukan adanya pengelompokan daerah-daerah penyebaran COVID-19 untuk memudahkan pemerintah menentukan strategi perencanaan dalam mengatasi percepatan laju penyebaran COVID-19 (KawalCovid, 2021).

2.2. Normalisasi Data

Dalam pemrosesan suatu data, adakalanya data memiliki masalah karena tidak meratanya suatu data. Demikian standarisasi data menggunakan normalisasi dapat dilakukan untuk mengatasi ketidakseimbangan pada data. Salah satu metode normalisasi yang cukup populer adalah normalisasi min-max. Pada normalisasi min-max data akan ditransformasikan secara linear sehingga menghasilkan keseimbangan dengan skala 0 sampai 1 (Nasution, 2019).

Normalisasi min-max digunakan untuk mengurangi nilai-nilai yang terlalu besar yang memungkinkan dapat memberi pengaruh besar pada proses perhitungan. Persamaan (2.1) adalah rumus untuk perhitungan menggunakan normalisasi min-max. Dimana w adalah nilai pada data asli dan w' adalah nilai pada data setelah dinormalisasi (Reynaldo et al., 2020).

$$w' = \frac{w - \min_x}{\max_x - \min_x} \quad (2.1)$$

2.3. Uji Multikolinearitas

Sebelum melakukan analisis klaster, diperlukan pengamatan terlebih dahulu pada masing-masing atribut yang akan digunakan. Salah satu uji statistik data yang dapat dilakukan adalah dengan melakukan uji multikolinieritas pada atribut. Multikolinearitas menemukan adanya hubungan linear yang sempurna

pada atribut-atribut yang digunakan atau dapat dikatakan memiliki korelasi yang tinggi (Aryanto et al., 2018). Dengan ditemukannya gejala multikolinearitas mengakibatkan hasil analisis yang diperoleh jauh dari akurat (Mu'afa and Ulinnuha, 2019).

Terdapat beberapa cara yang biasa digunakan untuk mengetahui adanya gejala multikolinearitas, salah satunya adalah dengan memperhatikan hasil perhitungan *Variances Inflation Factor* (VIF) (Mu'afa and Ulinnuha, 2019). Nilai VIF menyatakan faktor inflasi penyimpangan baku kuadrat. Jika nilai VIF tidak lebih dari 10, maka model dinyatakan bebas dari gejala multikolinieritas (Aryanto et al., 2018). Rumus untuk menghitung nilai VIF dapat diamati pada Persamaan (2.2).

$$VIF = \frac{1}{1 - R^2} \quad (2.2)$$

Dimana R^2 merupakan koefisien determinasi dengan Persamaan (2.3).

$$R^2 = \left[\frac{n \sum_{i=1}^n X_1 X_2 - \sum_{i=1}^n X_1 \sum_{i=1}^n X_2}{\sqrt{(n \sum_{i=1}^n (X_1)^2 - (\sum_{i=1}^n X_1)^2) - (n \sum_{i=1}^n (X_2)^2 - (\sum_{i=1}^n X_2)^2)}} \right]^2 \quad (2.3)$$

Dimana $i = 1, 2, \dots, n$; dan n = banyaknya data.

Selain menggunakan nilai VIF, dapat menggunakan nilai *Tolerance* (TOL) untuk mengetahui ada tidaknya multikolinieritas. Nilai TOL menyatakan besarnya tingkat kesalahan yang dibenarkan secara statistik. Jika nilai TOL kurang dari 0, 1 maka artinya tidak terdapat gejala multikolinieritas (Aryanto et al., 2018). Persamaan (2.4) merupakan hubungan antara TOL dan VIF.

$$TOL = \frac{1}{VIF} \quad (2.4)$$

2.4. Analisis Klaster

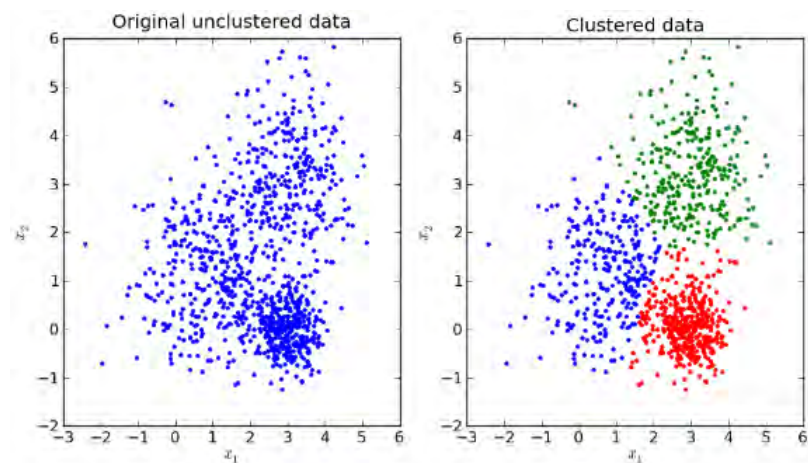
Analisis klaster adalah satu bagian dari beberapa metode *unsupervised learning* dikarenakan tidak adanya label awal yang ditentukan pada suatu data (Saxena et al., 2017). Analisis pengelompokan atau klaster merupakan suatu proses pembagian data dari suatu himpunan ke dalam kelompok atau klaster yang lebih banyak (Bastian et al., 2018). Tujuan dari analisis klaster yaitu mengelompokkan suatu data ke dalam klaster-klaster yang memiliki kemiripan karakteristiknya (Solichin and Khairunnisa, 2020). Kemiripan karakteristik data dapat diperoleh dari kedekatan nilai-nilai atributnya (Nabila et al., 2021).

Clustering terbagi menjadi dua macam yakni *hierarcial clustering* dan *non-hierarcial clustering* (Bastian et al., 2018). Metode *hierarcial clustering* mengelompokkan data dengan menggunakan konsep hirarki. Mulanya dua atau lebih data akan dikelompokkan terlebih dahulu dengan melihat kemiripan terdekat. Lalu akan dilanjutkan dengan mengelompokkan dengan data yang memiliki kedekatan kedua dan seterusnya. Hal ini berbeda dengan metode *non-hierarcial clustering* yang sudah ditentukan jumlah klasternya di awal. Setelah menginisialisasi jumlah klaster yang akan digunakan maka selanjutnya dilakukan proses klaster tanpa menggunakan konsep hirarki namun menggunakan sistem partisi (Annur, 2019).

2.5. Algoritma K-Means

Algoritma k-Means termasuk bagian dari algoritma *non-hierarcial clustering* yang digunakan untuk mengelompokkan data dari suatu himpunan menggunakan sistem partisi (Sampara, 2020). Algoritma k-Means merupakan metode dengan proses yang sederhana, praktis, dan memiliki keakuratan yang baik

dalam pengklasteran data. Selain itu kelebihan algoritma k-Means yang sangat populer adalah ketelitiannya saat menghadapi ukuran objek dalam jumlah besar serta tidak terpengaruh oleh urutan objek (Bastian et al., 2018). Gambar 2.2 merepresentasikan data sebelum dan sesudah dilakukan klastering.



Gambar 2.2 Sebelum (kiri) dan Sesudah (kanan) Klastering

Sumber: (Fauziyah Dewi, 2019)

Secara rinci, langkah-langkah klastering menggunakan algoritma k-Means adalah sebagai berikut (Nabila et al., 2021):

1. Menentukan jumlah klaster (k) yang diinginkan.
2. Menentukan pusat klaster. Untuk pusat klaster awal ditentukan secara acak sedangkan pada iterasi berikutnya pusat klaster ditentukan menggunakan Persamaan (2.5).

$$V_{ij} = \frac{1}{N_i} \sum_{k=0}^{N_i} X_{kj} \quad (2.5)$$

Dimana, N_i merupakan jumlah anggota klaster ke- i dan X_{kj} merupakan nilai data ke- k pada variabel ke- j .

3. Menghitung jarak data dengan pusat pusat kluster menggunakan metode ED (*Eucledian Distance*) dengan Persamaan (2.6).

$$ED = \sqrt{\sum_{i=1}^m (x_i - y_i)^2} \quad (2.6)$$

Dengan:

$i = 1, 2, \dots, m$; dan m = banyaknya variabel. Sedangkan x_i merupakan data pada variabel ke-i dan y_i merupakan nilai pusat kluster variabel ke-i.

4. Mengelompokkan data dengan berdasarkan jarak terdekat ke pusat kluster.
5. Lakukan iterasi dengan mengulangi langkah kedua hingga tidak terjadi perubahan pada pusat kluster (konvergen).

2.6. Algoritma Genetika

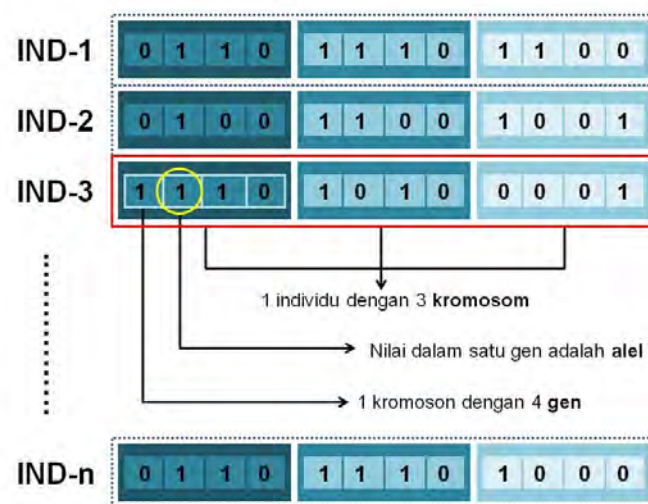
Salah satu metode adiptif yang cukup sering dijumpai untuk menyelesaikan masalah optimisasi adalah algoritma genetika (Ali and Sularto, 2019). Algoritma genetika menggunakan konsep teori evolusi yang diperkenalkan oleh Charles Darwin dengan prinsip dasar seleksi alam, dimana makhluk hidup dengan ketahanan atau fit terbaik yang akan bertahan hidup (Prasetio et al., 2018). Namun sebaliknya, makhluk hidup dengan fit yang rendah akan punah. Kelebihan yang dimiliki oleh algoritma genetika adalah kemampuannya dalam menyelesaikan masalah dengan banyak variabel (Santoso et al., 2017).

Dalam satu populasi biasanya terdiri dari puluhan, ratusan, bahkan ribuan solusi (individu). Akan dilakukan proses evaluasi fitness pada tiap-tiap individu menggunakan fungsi fitness (Suwirmayanti, 2018). Tujuan pada proses ini adalah untuk mendapatkan individu dengan fit terbaik yang kemudian dipilih menjadi

orang tua. Selanjutnya adalah proses reproduksi dengan *crossover* (perkawinan silang) dan mutasi. Hal ini dilakukan dengan harapan dari orang tua terbaik akan menghasilkan keturunan yang lebih baik sehingga populasi terus menjadi lebih baik. Langkah-langkah dalam menyelesaikan permasalahan menggunakan algoritma genetika secara terperinci adalah sebagai berikut (Setiawan et al., 2019):

2.6.1. Membangkitkan Populasi Acak

Langkah pertama yang dilakukan pada algoritma genetika adalah membangkitkan populasi yang berisikan individu-individu secara acak sebagai suatu solusi dari suatu permasalahan. Individu terdiri dari beberapa kromosom, lalu kromosom terdiri dari beberapa gen, dan nilai dari suatu gen disebut alel (Marimin, 2019). Gambar 2.3 adalah representasi antara populasi, individu, kromosom, gen, dan alel.



Gambar 2.3 Representasi dalam Satu Populasi

Sumber: (Roiha, 2017)

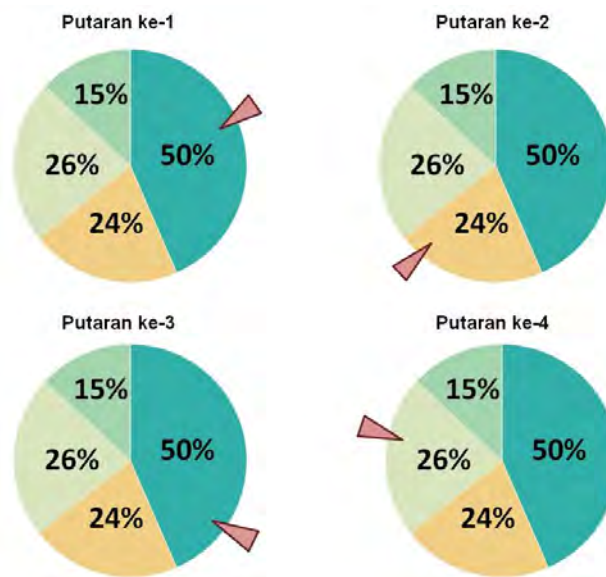
2.6.2. Fungsi Fitness

Pada setiap individu akan dihitung nilai fit-nya dengan menggunakan fungsi fitness. Fungsi fitness akan menjadi pengukuran yang menentukan seberapa fit suatu individu untuk dapat bertahan. Fungsi fitness mengeluarkan *fitness value* untuk tiap-tiap individu, dan individu terbaik memiliki fit yang paling tinggi. Fungsi fitness berbeda-beda tergantung terhadap permasalahan yang akan dioptimasi (Kuntjoro et al., 2018).

Selain ditentukan oleh jumlah generasi maksimal, kondisi iterasi dapat berhenti pada algoritma genetika dengan memperhatikan fungsi fitnessnya, yakni dengan cara menentukan batas maksimal nilai fitness yang ingin dicapai terlebih dahulu. Jika individu sudah mencapai batas tersebut, maka iterasi dapat berhenti (Handayani et al., 2020).

2.6.3. Seleksi Individu

Pada tahap seleksi individu akan dipilih individu-individu terbaik untuk dijadikan orang tua. Semakin besar nilai fitness pada tiap individu, maka probabilitas individu terpilih akan semakin besar (Handayani et al., 2020). Terdapat dua macam teknik yang sering dijumpai untuk melakukan seleksi individu yaitu turnamen dan mesin *roulette*. Pada penelitian ini akan digunakan metode mesin *roulette*. Gambar 2.4 menjelaskan bagaimana pemilihan individu terbaik menggunakan mesin *roulette*.



Gambar 2.4 Ilustrasi Mesin roulette

Sumber: (Roiha, 2017)

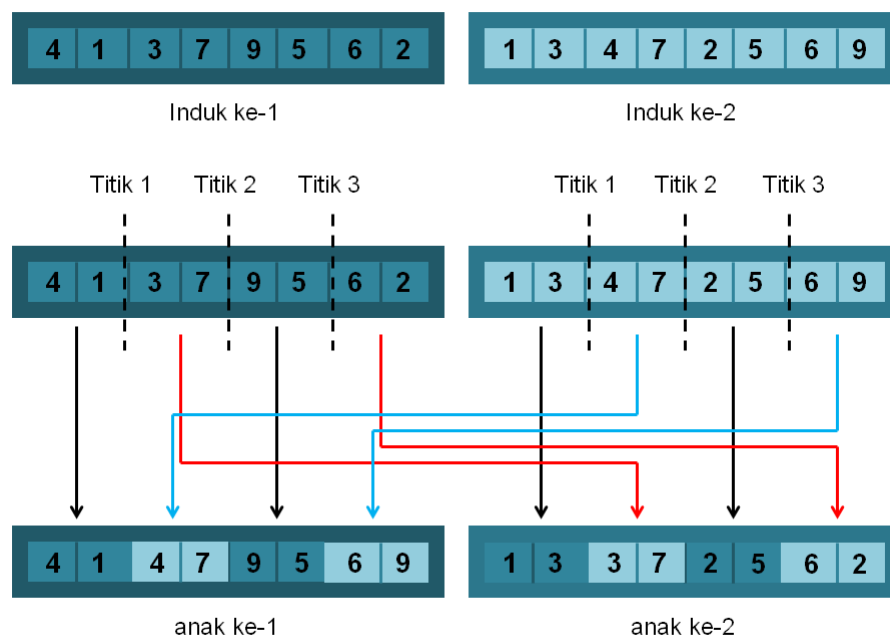
Nilai fitness tiap individu yang telah diketahui sebelumnya akan dijumlahkan secara keseluruhan untuk dihitung nilai probabilitas tiap-tiap individunya. Selanjutnya dibangkitkan nilai random antara 0 – 1 sebagai penentu individu manakah yang terpilih sebagai orang tua. Pembangkitan bilangan random dilakukan sebanyak n individu, namun dalam kasus lain dengan data yang cukup banyak dilakukan random sebanyak $n \times Cr$ dimana Cr adalah probabilitas *crossover*. Selanjutnya individu yang terpilih akan dilanjutkan ke proses *crossover* untuk dilakukan persilangan.

2.6.4. Proses Crossover

Setelah mendapatkan orang tua atau induk, maka dapat dilakukan *crossover*. *Crossover* adalah proses pertukaran dari sebagian solusi dengan bagian induk lain untuk menghasilkan individu yang berbeda (*offspring*) yang mana individu baru tersebut dapat menjadi solusi baru untuk menyelesaikan

permasalahan (Kuntjoro et al., 2018). Melalui perkawinan silang antara kedua individu diharapkan dapat melahirkan individu baru yang lebih baik namun tetap mewarisi sebagian sifat-sifat induknya. Hal ini sesuai dengan proses reproduksi yang terjadi di bumi. Persilangan kedua individu dilakukan secara acak dan akan mendapatkan dua anakan (*offspring*). Banyak *offspring* yang dihasilkan bergantung pada nilai *crossover rate* (Cr) yang ditentukan di awal (Iauiq et al., 2017).

Terdapat beberapa metode dalam melakukan *crossover* disesuaikan dengan representasi kromosom yang digunakan. Beberapa diantaranya adalah *crossover* satu titik, *crossover* n titik, *crossover* aritmatika, dan *crossover* kromosom permutasi. Dalam penelitian ini digunakan reproduksi *crossover* n titik ($n = 3$). Gambar 2.5 merupakan proses *crossover* 3 titik.



Gambar 2.5 Proses Crossover 3-titik

Sumber: (Hussain et al., 2018)

Pada Gambar 2.5 terdapat dua individu sebagai induk yang dipilih secara acak. Selanjutnya induk akan dibagi menjadi 4 bagian karena *crossover* yang dipilih menggunakan *crossover* 3 titik yang mana titik-titiknya ditentukan secara acak pula. Setelah ditentukan titik-titiknya, maka dilakukan persilangan dari kedua induk. Gen anak pertama diambil dari gen induk pertama bagian satu dan tiga, sisanya diambil dari induk kedua. Sedangkan gen anak kedua diambil dari gen induk kedua bagian satu dan tiga, sisanya diambil dari induk pertama.

2.6.5. Proses Mutasi Gen

Setelah melakukan reproduksi dengan *crossover* akan dilakukan reproduksi menggunakan mutasi gen. Mutasi merupakan operator genetika yang dilakukan untuk mencegah adanya konvergensi terlalu awal dikarenakan keberagaman tak terkendali dalam suatu populasi. Jika dalam *crossover* terdapat Cr maka dalam mutasi terdapat *mutation rate* (Mr). Proses mutasi pada gen dilakukan sebanyak Mr dikali jumlah gen (Kuntjoro et al., 2018). Misalkan $Mr = 0.2$ dan individu sebanyak 4 maka mutasi akan dilakukan sekali (pembulatan dari $0.2 \times 4 = 0.8$).

Sama halnya dengan reproduksi *crossover*, mutasi memiliki berbagai macam metode. Metode mutasi gen yang digunakan pada penelitian ini adalah *random mutation*. Pertama-tama akan dipilih gen dalam suatu kromosom secara acak lalu ditambahkan dengan nilai yang sangat kecil (r) di rentang $[-0.1, 0.1]$ yang didapatkan secara acak pula (Kuntjoro et al., 2018). Misalkan pada domain x_i terdapat $[max_i, min_i]$ dengan *offspring* yang dihasilkan adalah $S = [x_1, x_2, \dots, x_n]$, maka pembangkitan nilai gen *offspring* dapat menggunakan Persamaan (2.7) (Suryaningrum et al., 2017).

$$x_i' = x_i + r(max_i - min_i) \quad (2.7)$$

2.6.6. Elitism

Proses terakhir sebelum menuju iterasi selanjutnya adalah proses elitism. Pada bagian ini akan digunakan sistem ranking dengan memperhatikan nilai fitness pada tiap-tiap individu (Kuntjoro et al., 2018). Individu dengan fitness terbaik sebanyak n individu pada populasi awal akan dipilih untuk diteruskan menuju iterasi selanjutnya (Prasetio et al., 2018). Itulah mengapa jumlah populasi baru untuk iterasi selanjutnya tidak mengalami penambahan individu meskipun telah dilakukan reproduksi individu baru.



Gambar 2.6 Proses Elitism

Sumber: (Roiha, 2017)

2.7. Algoritma Genetika dan K-Means

Gabungan algoritma genetika dan k-Means merupakan metode yang telah dimodifikasi untuk mendapatkan hasil klastering yang optimal. Algoritma genetika akan mengoptimalkan pusat klaster awal sebagai masukan dari metode k-Means (Shi and Xu, 2018). Hal ini tentunya dapat mengatasi masalah pusat klaster awal pada k-Means yang dibangkitkan secara acak (Kuntjoro et al., 2018). Hasil pusat klaster dari optimasi GA menjadi pusat klaster awal yang diproses ke dalam metode k-Means. Algoritma genetika pada penelitian ini akan digunakan untuk mengoptimalkan pusat klaster (*centroid*) dengan mengambil jarak terdekat pada tiap anggota klaster. Perhitungan jarak antar data pada tiap klaster menggunakan

Manhattan Distance (MD) dengan Persamaan (2.8).

$$d(x_n, w_r) = \sum_{p=1}^m |x_n^p - w_r^p| \quad (2.8)$$

Dimana $p = 1, 2, ..m$; dan m = banyaknya variabel. Dengan x_n^p merupakan data ke- n pada variabel ke- p dan w_r^p merupakan pusat kluster ke- r pada variabel ke- p . Sedangkan Persamaan (2.9) merupakan rumus perhitungan jarak terdekat.

$$J = \sum_{n=1}^N \min_r d(x_n, w_r) \quad (2.9)$$

Lalu dalam permasalahan ini, untuk mendapatkan nilai fitness (F) pada tiap-tiap individu dalam populasi digunakan inversnya dengan Persamaan (2.10).

$$F = \frac{1}{J} \quad (2.10)$$

2.8. Evaluasi Kluster

Setelah melakukan analisis kluster perlu dilakukan evaluasi dari hasil yang didapatkan. Klustering yang baik dapat dilihat dari kemiripan (homogen) data pada kelasnya dan semakin tidak mirip (heterogen) dengan kelas lainnya (Solichin and Khairunnisa, 2020). Terdapat banyak metode evaluasi kluster untuk mengukur jarak antar data dan mengetahui seberapa dekat atau jauh data dari klasternya. Pada bagian ini metode yang digunakan oleh peneliti untuk evaluasi kluster adalah metode *Silhouette Coefficient* (SC) (Kuntjoro et al., 2018). Metode tersebut mengarah pada metode validasi suatu pengelompokan data. Metode ini memberikan informasi seberapa bagus setiap objek terletak dalam suatu kluster/kelompok (Mamat et al., 2018).

Nilai SC berkisar di rentang -1 dan 1 . Semakin mendekati 1 memperlihatkan bahwa objek tersebut semakin tepat berada pada kluster yang ditentukan (Siagian, 2021). Ketika nilai sama dengan atau mendekati 0 maka objek berada di antara perbatasan dua kluster. Sedangkan nilai SC sama dengan atau mendekati -1 mengindikasikan bahwa objek menjauh dari kluster yang ditentukan. Tabel 2.1 menyatakan kriteria subjektif pengukuran dengan menggunakan SC.

Tabel 2.1 Struktur Kategori Silhouette Coefficient

Nilai SC	Kategori
0.71-1.00	Struktur kluster dinyatakan kuat
0.51-0.70	Struktur kluster dinyatakan baik
0.26-0.50	Struktur kluster dinyatakan cukup
-1.00-0.26	Struktur kluster dinyatakan buruk

Sumber: (Sari, 2016)

Untuk mendapatkan nilai SC pertama-tama dihitung jarak rata-rata dari suatu data- i di satu kluster (a_i) dengan menggunakan Persamaan (2.11) dengan memisalkan i terhadap seluruh data lain pada satu kluster.

$$a_i = \frac{1}{m-1} \sum_{r=1, r \neq i}^m d(x_i, x_r), \quad i = 1, 2, \dots, m \quad (2.11)$$

Setelah itu, dihitung jarak rata-rata data ke- i pada seluruh data di kluster lain (b_i) lalu diambil nilai minimumnya dengan Persamaan (2.12).

$$b_i = \min \frac{1}{m} \sum_{r=1, r \neq i}^m d(x_i, x_r), \quad i = 1, 2, \dots, m \quad (2.12)$$

Langkah terakhir adalah menghitung nilai SC pada data ke- i (SC_i), nilai SC pada suatu klaster (SCK), dan nilai SC pada seluruh pengklasteran (SC) menggunakan Persamaan (2.13), (2.14), dan (2.15).

$$SC_i = \frac{b_i - a_i}{\max(a_i, b_i)} \quad (2.13)$$

$$SCK = \frac{1}{m} \sum_{i=1}^m SC_i \quad (2.14)$$

$$SC = \frac{1}{k} \sum_{j=1}^k SCK \quad (2.15)$$

2.9. Wabah Penyakit dan Integrasi Keislaman

Sebelum mengetahui relevansi antara pandemi COVID-19 dan wabah dalam Islam maka alangkah baiknya untuk mengetahui apa itu wabah. Wabah berasal dari bahasa arab (*awbaa'*, *yubiu'*) dengan arti terserang wabah. Selain itu dalam kamus al-Munawwir kata *al-wabaa* memiliki arti penyakit pes atau infeksi terhadap bakteri *Yersinia Pestis* (Ridho, 2020). Sedangkan definisi wabah yang disepakati oleh para ahli adalah suatu penyakit yang dapat menyebar dengan cepat dikarenakan adanya bakteri atau virus yang menjangkit individu dan kemudian menularkan pada individu lain (Mahmud, 2020).

Hal ini sangat relevan dengan adanya kemunculan virus COVID-19 yang kini menjadi permasalahan karena sudah menyebar ke seluruh dunia serta menyebabkan banyak kerugian baik secara ekonomi maupun pengurangan sumber daya manusia. *Maqashid al-syariah* memiliki beberapa kaidah salah satunya adalah *hifdzun nafs* (menjaga jiwa). Sebelum menjaga agama, menjaga nyawa atau jiwa penting dilakukan oleh setiap muslim. Itulah mengapa dalam Islam sangat

diwajibkan untuk menghindari setiap sesuatu yang berpotensi mengancam eksistensi nyawa manusia (Mukharom and Aravik, 2020). Saat terjadi wabah COVID-19, manusia tidak bisa hanya berdiam diri dan menunggu wabah pergi dengan sendirinya namun diperlukan upaya untuk menekan angka kenaikan kasus COVID-19 (Ginting et al., 2021).

Saat wabah COVID-19 masuk ke Indonesia, pemerintah turut andil untuk membuat segala macam aturan serta strategi untuk menekan laju penambahan kasus COVID-19 yang terus meningkat setiap harinya. Aturan yang dibuat oleh pemerintah didasarkan pada kemaslahatan masyarakat. Demikian itu maka kita diwajibkan untuk melakukan upaya dengan berpedoman kepada pemerintah. Dalam al-Qur'an surah An-Nisa ayat 59 Allah Swt memerintahkan orang-orang yang beriman untuk menaati Allah, Rasulullah, beserta *ulil amri* (penguasa).

يَا أَيُّهَا الَّذِينَ آمَنُوا أَطِيعُوا اللَّهَ وَأَطِيعُوا الرَّسُولَ وَأُولِيَ الْأَمْرِ مِنْكُمْ فَإِنْ تَنَازَعْتُمْ فِي شَيْءٍ فَرُدُّوهُ إِلَى اللَّهِ وَالرَّسُولِ إِنْ كُنْتُمْ تُؤْمِنُونَ بِاللَّهِ وَالْيَوْمِ الْآخِرِ ذَلِكَ خَيْرٌ وَأَحْسَنُ تَأْوِيلًا ﴿٥٩﴾

Artinya: "Wahai orang-orang yang beriman! Taatilah Allah dan taatilah Rasul (Muhammad), serta *ulil amri* (pemegang kekuasaan) di antara kamu. Kemudian, jika kamu berbeda pendapat tentang sesuatu, maka kembalikanlah kepada Allah (al-Qur'an) dan Rasul (sunnahnya), jika kamu beriman kepada Allah dan hari akhir. Yang demikian itu lebih utama (bagimu) dan lebih baik akibatnya." (QS. An-Nisa: 59).

Dalam ayat tersebut, Allah Swt memerintahkan manusia untuk menaati-Nya, lalu kemudian menaati Rasulullah Saw, dan kemudian menaati *ulil amri*. Kata *ulil amri* yang dimaksud merupakan penguasa atau pemimpin di suatu tempat seperti presiden maupun pemerintah daerah. Saat pemerintah sudah

menetapkan suatu aturan maka wajib hukumnya bagi setiap muslim untuk mematuhi aturan yang dibuat oleh pemerintah (Rosanti, 2021).

Pemerintah memiliki peran yang sangat besar untuk menurunkan laju penyebaran COVID-19 di Indonesia. Karena penyebarannya yang cepat, maka langkah yang diambil juga harus cepat dan tepat (Arswendi, 2021; Ulya, 2020). Seperti yang disebutkan sebelumnya bahwa penyebaran virus tersebut menyebar dengan sangat cepat sehingga pemerintah menghimbau untuk selalu menjaga jarak antar sesama individu. Hal ini dilakukan untuk mengurangi resiko menularkan maupun tertular dari individu lain (Mukharom and Aravik, 2020). Rasulullah Saw menganjurkan hal tersebut melalui hadis yang diriwayatkan oleh Imam Ahmad sebagai berikut:

حَدَّثَنَا وَكِيعٌ قَالَ حَدَّثَنَا النَّهَّاسُ عَنْ شَيْخِ بَمَكَةَ عَنْ أَبِي هُرَيْرَةَ قَالَ سَمِعْتُ رَسُولَ اللَّهِ صَلَّى اللَّهُ عَلَيْهِ وَسَلَّمَ يَقُولُ فِرٌّ مِنَ الْمَجْدُومِ فِرَارَكَ مِنَ الْأَسَدِ (رواه احمد)

Artinya: "Telah menceritakan kepada kami Waki' berkata; telah menceritakan kepada kami An Nahas dari seorang syaikh di Makkah, dari Abu Hurairah berkata; Aku mendengar Rasulullah Saw bersabda, "Jauhilah penyakit kusta sebagaimana engkau lari dari kejaran singa."" (HR. Ahmad).

Pada hadis di atas disebutkan bahwa Rasulullah Saw memerintahkan kita untuk menjauhi penderita kusta atau lepra. Dimana penyakit tersebut merupakan suatu penyakit menular yang disebabkan oleh bakteri *Mycobacterium Leprae*. Sama-sama merupakan penyakit yang mudah menular, hal ini juga berlaku pada larangan untuk menjauhi penderita COVID-19 agar tidak tertular (Nurhidayati and Rofiq, 2021).

Menjauhi yang dimaksud bukan berarti tidak boleh menjalin hubungan sama sekali dengan penderita COVID-19 ataupun orang sekitar. Namun menjauhi yang dimaksud adalah dengan menjaga jarak antar sesama. Sehingga hal demikian dapat mengurangi resiko tertular maupun menularkan penyakit pada orang lain (Solichin and Khairunnisa, 2020).

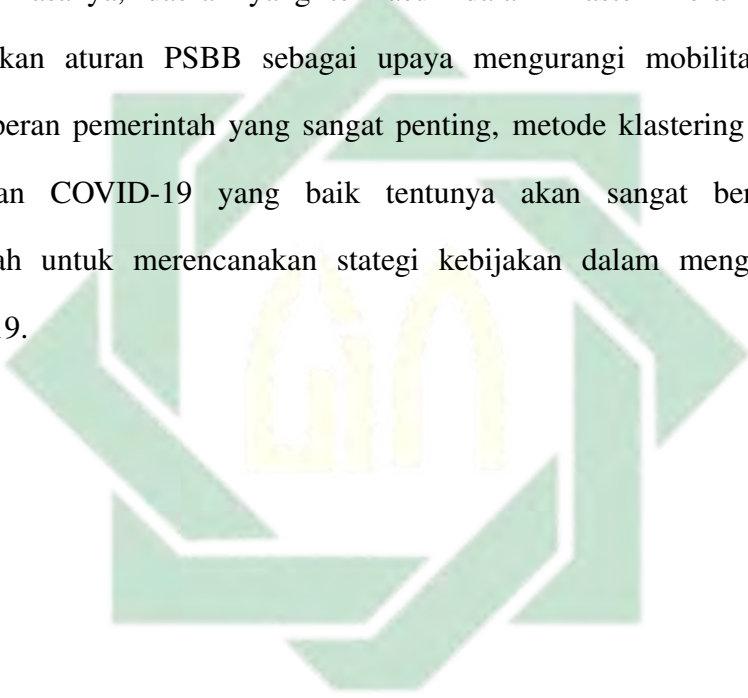
Majelis Ulama Indonesia (MUI) mengeluarkan fatwa mengenai ikhtiar yang dapat dilakukan untuk menghadapi wabah COVID-19. Dalam fatwanya (No. 14/2020) seluruh masyarakat Indonesia diwajibkan untuk menjaga kesehatan diri beserta menjauhi hal-hal yang dapat menyebabkan terpaparnya penyakit. Beberapa hal yang biasanya dikerjakan di luar rumah semisal belajar di sekolah, bekerja, salat berjamaah di masjid, dan lain sebagainya harus dilakukan pembatasan aktivitas (Rosanti, 2021). Misalkan dengan bersekolah dan bekerja yang dapat dilakukan secara daring ataupun salat berjamaah yang dapat dilakukan di rumah. Fatwa MUI tersebut digunakan oleh pemerintah untuk menentukan kebijakan berupa pembatasan dalam melakukan interaksi sosial dengan lebih spesifik. Hal ini bersesuaian dengan kaidah fikih di bawah ini:

UIN SUNAN / مَا لَا يَتِمُّ الْوَاجِبُ إِلَّا بِهِ فَهُوَ وَاجِبٌ
S U R A B A Y A

Artinya: "Sesuatu yang menjadi syarat bagi sebuah kewajiban, maka hukumnya juga menjadi wajib."

Dalam kaidah tersebut dimaksudkan pada hal yang sebelumnya tidak wajib dilakukan dapat berubah menjadi wajib dalam kondisi tertentu. Pada kasus ini, kondisi yang dimaksudkan adalah kondisi dalam pentingnya menjaga jiwa sehingga dibuat keputusan pemerintah yang bersifat wajib untuk menjaga jiwa dalam ikhtiar menghadapi wabah COVID-19.

Selain pembatasan dalam melakukan aktivitas dalam skala tertentu, pemerintah juga menerapkan Pembatasan Sosial Berskala Besar (PSBB). Pada aturan PSBB, dilakukan pembatasan berdasarkan daerah-daerah dengan tingkat kerawanan yang tinggi. Daerah-daerah di Indonesia akan dikelompokkan ke dalam beberapa kelompok daerah zona dengan 4 warna (merah, oranye, kuning, dan hijau). Biasanya, daerah yang termasuk dalam klaster merah (rawan) akan diberlakukan aturan PSBB sebagai upaya mengurangi mobilitas masyarakat. Melihat peran pemerintah yang sangat penting, metode klastering daerah-daerah penyebaran COVID-19 yang baik tentunya akan sangat bermanfaat bagi pemerintah untuk merencanakan strategi kebijakan dalam menghadapi wabah COVID-19.



UIN SUNAN AMPEL
S U R A B A Y A

BAB III

METODE PENELITIAN

3.1. Jenis Penelitian

Penelitian ini memiliki tujuan untuk memberikan informasi kepada masyarakat maupun pemerintah terkait pengelompokan daerah-daerah terdampak COVID-19 di Provinsi Jawa Timur. Hal ini dilakukan agar dapat membantu pemerintah dalam membuat strategi perencanaan di masa pandemi dengan lebih optimal. Oleh karena itu, jenis penelitian ini adalah penelitian terapan yang mana hasil penelitiannya merupakan jawaban atas adanya suatu permasalahan untuk kemudian dapat diterapkan.

3.2. Sumber Data

Penelitian ini menggunakan data sekunder yang diambil secara daring di halaman *website* resmi Jawa Timur Tanggap COVID-19. Data yang digunakan adalah data sebaran pada bulan Januari - November 2021 yang dapat diamati pada halaman lampiran A-K. Data yang digunakan sebagai contoh perhitungan adalah data bulan November 2021 yang dapat diamati pada halaman lampiran K.

3.3. Variabel Penelitian

Variabel atau atribut yang digunakan pada penelitian ini berdasarkan data sebaran pada halaman *website* Jawa Timur Tanggap Covid-19 (Jatim Pemprov, 2021).

Tabel 3.1 Variabel Penelitian

Variabel	Keterangan
x_1	Jumlah kasus terkonfirmasi COVID-19.
x_2	Jumlah kasus pasien sembuh dari COVID-19.
x_3	Jumlah kasus pasien meninggal saat menderita COVID-19.
x_4	Jumlah kasus COVID-19 dengan riwayat bergejala (<i>suspect</i>).
x_5	Jumlah kasus COVID-19 tanpa riwayat bergejala (<i>non-suspect</i>).
x_6	Jumlah kasus COVID-19 dengan riwayat kontak erat.
x_7	Jumlah kasus COVID-19 tanpa riwayat kontak erat.
x_8	Jumlah kasus COVID-19 dengan riwayat perjalanan.

3.4. Kerangka Penelitian

Secara garis besar penelitian ini memiliki beberapa tahapan utama yakni *pre-processing*, implementasi GA dan k-Means, uji coba parameter kluster terbaik menggunakan data bulan November 2021, pemetaan kluster setiap bulan menggunakan parameter terbaik yang diperoleh dari hasil uji coba, dan terakhir adalah analisis perubahan hasil kluster dari bulan Januari sampai November 2021.

3.4.1. *Pre-processing*

Langkah pertama pada tahapan *pre-processing* adalah data akan dilakukan normalisasi min-max menggunakan Persamaan (2.1). Lalu setelah itu akan dilakukan uji multikolinearitas menggunakan Persamaan (2.2) dan (2.4). Jika terdapat gejala multikolinearitas, maka salah satu variabel yang memiliki gejala multikolinearitas harus direduksi. Jika sudah tidak ada variabel yang memiliki gejala multikolinearitas, maka dapat dilanjutkan pada langkah berikutnya.

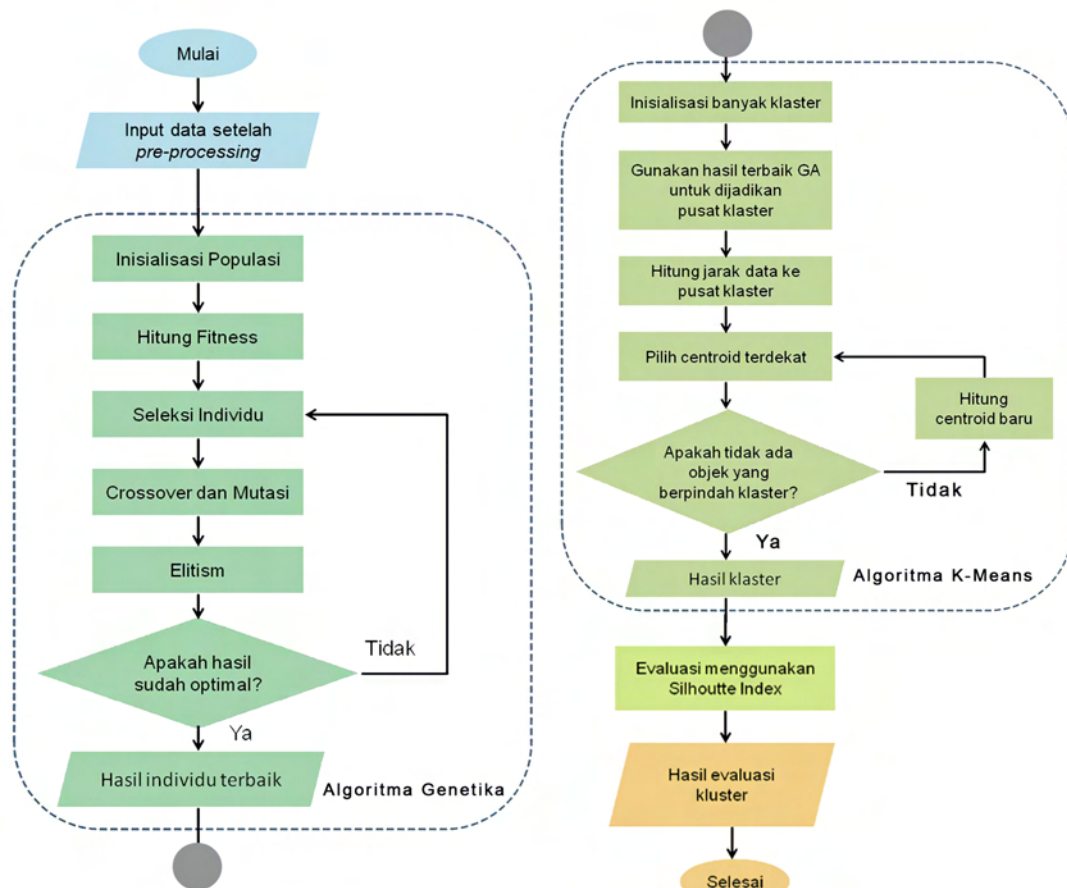
3.4.2. Implementasi GA dan K-Means

Penelitian ini menggunakan gabungan metode GA dan k-Means. Dalam implementasinya terdapat beberapa bagian penting yakni tahap optimasi pusat klaster menggunakan algoritma genetika, tahap klastering menggunakan k-Means, dan tahap evaluasi hasil klastering. Alur penelitian dapat dirincikan sebagai berikut:

1. Masukkan data yang sudah dilakukan *pre-processing* sebagai inputan.
2. Tahap optimasi pusat klaster: metode yang digunakan untuk optimasi pusat klaster adalah algoritma genetika. Pada algoritma genetika dilakukan inisialisasi populasi terlebih dahulu, lalu hitung nilai fitness tiap individu menggunakan Persamaan (2.6), (2.7) dan (2.8), seleksi menggunakan mesin *roulette*, dilakukan persilangan (*crossover*) menggunakan *crossover* di 3-titik, setelah itu dimutasi secara acak dengan pembangkitan gen menggunakan Persamaan (2.9), terakhir proses seleksi perankingan populasi selanjutnya (*elitism*). Hal ini dilakukan secara berulang sampai mencapai kondisi batas yang digunakan (maksimal generasi).
3. Tahap klastering: metode klastering yang digunakan adalah k-Means. Pertama inisialisasi jumlah klaster yang diinginkan, lalu ambil pusat klaster awal dari hasil optimasi algoritma genetika, hitung jarak antara data dengan pusat klaster menggunakan Persamaan (2.10), kelompokkan data dengan berdasarkan jarak terdekat ke pusat klaster, ulangi langkah-langkah tersebut sampai tidak ada pengelompokan data yang berpindah klaster.
4. Tahap evaluasi klaster: pada bagian ini hasil analisis klaster akan dievaluasi menggunakan metode *Silhouette Coefficient*. Perhitungan nilai SC dapat

menggunakan Persamaan (2.13), (2.14), dan (2.15).

5. Keluaran yang dihasilkan merupakan hasil evaluasi klaster.



Gambar 3.1 Diagram Alir Implementasi GA dan K-Means

3.4.3. Uji Coba Parameter

Pada penelitian ini dilakukan beberapa uji coba parameter untuk menentukan parameter terbaik. Uji coba pada tahap GA adalah uji coba banyaknya individu dalam suatu populasi ($Popsiz = 10, 20, 40, 80, 160$) dan banyaknya generasi ($I = 100, 200, 300, 400, 500$). Sedangkan pada tahap k-Means, dilakukan uji coba jumlah klaster ($k = 2, 3$, dan 4). Perhitungan uji coba parameter menggunakan data bulan November 2021.

Parameter terbaik dinilai dari validasi SC yang memiliki nilai paling tinggi. Hasil uji coba parameter nantinya akan menjadi model klaster terbaik yang akan diterapkan sebagai inisialisasi parameter pada implementasi klastering data di bulan lain untuk menganalisis perubahan klaster dalam beberapa bulan.

3.4.4. Pemetaan Klaster

Selanjutnya, parameter klaster terbaik yang didapat akan diimplementasikan pada pengelompokan data dari bulan Januari sampai November 2021. Artinya, pengelompokan tersebut menggunakan parameter klaster yang sama yang didapat dari hasil ujicoba parameter sebelumnya. Hasil klastering tiap bulannya kemudian dianalisis kabupaten/kota mana saja yang masuk pada daerah dengan kerawanan yang tinggi/rendah dengan mengamati nilai pusat klasternya.

Setelah diketahui label klaster kelompoknya, dilakukan pemetaan dengan memberi warna sesuai klasternya untuk melihat karakteristik klasternya. Warna merah akan digunakan untuk menunjukkan daerah dengan tingkat kerawanan COVID-19 yang tinggi. Warna hijau akan digunakan untuk menunjukkan daerah dengan tingkat kerawanan COVID-19 yang lebih rendah. Sedangkan jika jumlah klaster lebih dari dua, maka akan diambil warna transisi merah ke hijau.

3.4.5. Analisis Hasil

Tahap terakhir adalah tahap analisis perubahan hasil klaster pada pemetaan sebelumnya. Hal ini bertujuan untuk mengamati perubahan anggota klaster. Data yang diamati dari bulan Januari sampai November 2021. Identifikasi yang dilakukan diantaranya yaitu identifikasi kabupaten/kota mana saja yang mengalami peningkatan dan penurunan level kerawanan klaster serta kabupaten/kota mana saja yang tidak mengalami perubahan level.

Diharapkan dengan adanya analisis perubahan ini dapat menjadi pertimbangan bagi pemerintah untuk mengambil kebijakan penanganan khususnya pada kabupaten/kota yang mengalami peningkatan level kerawanan klaster.

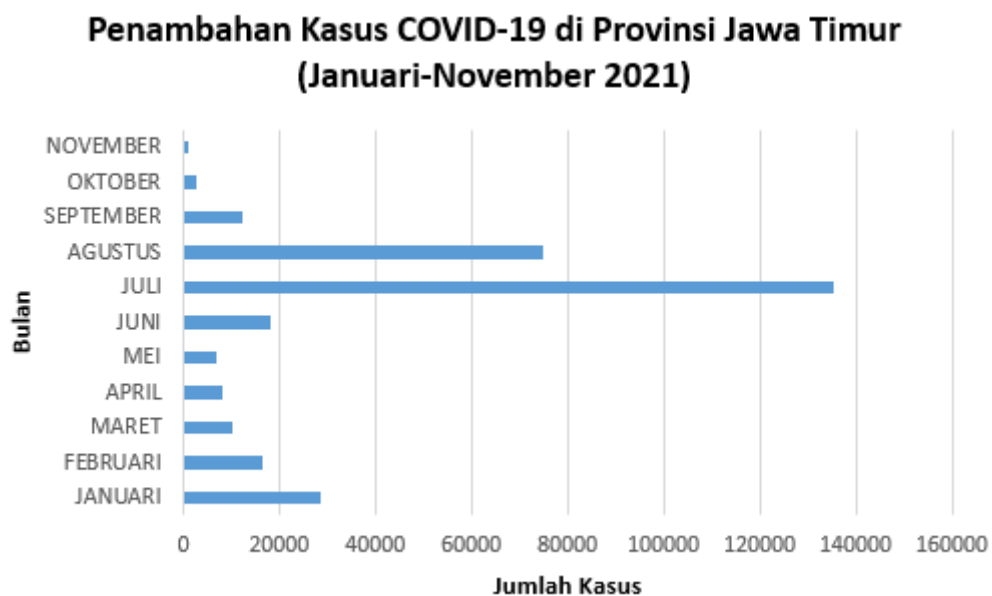


BAB IV

HASIL DAN PEMBAHASAN

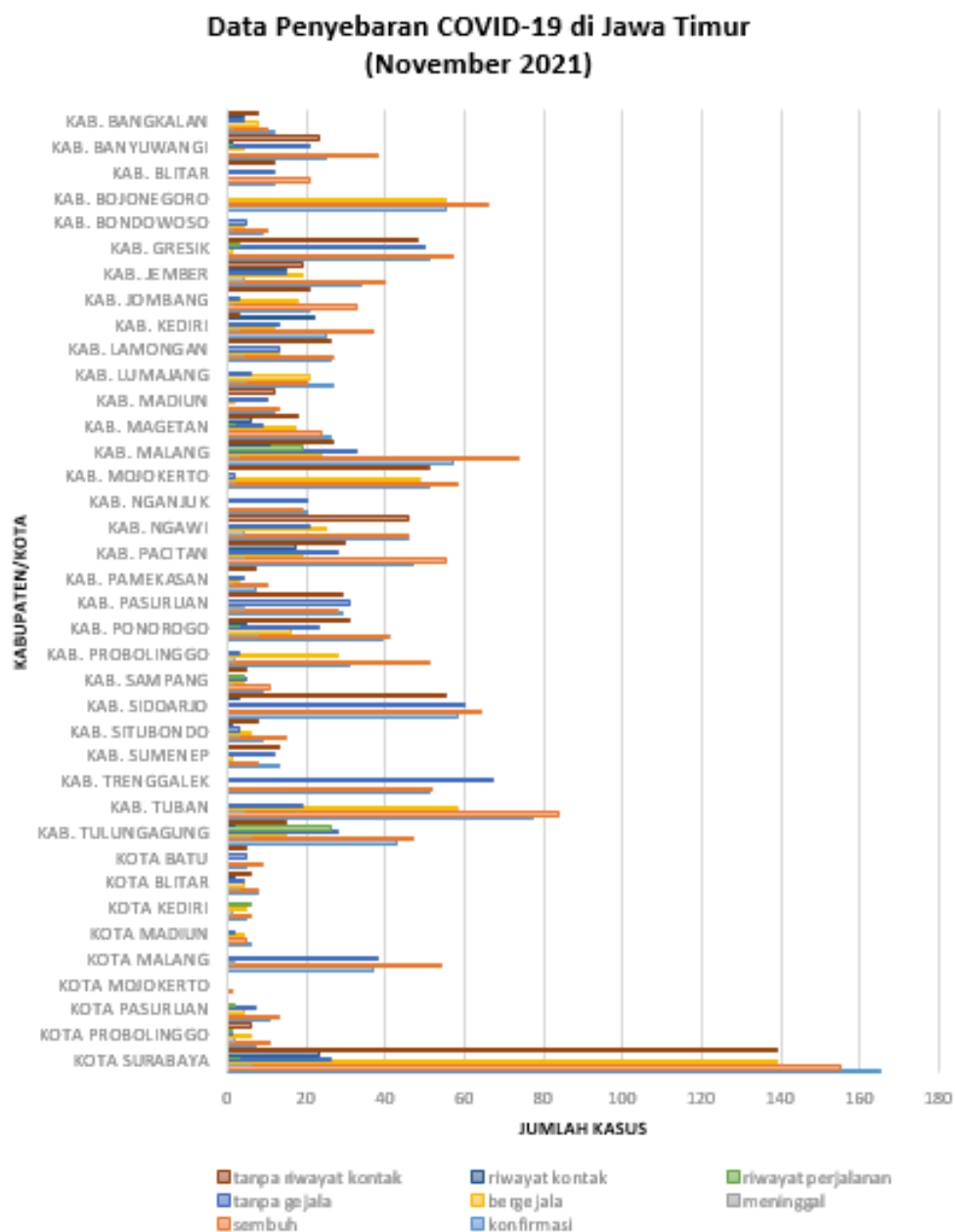
4.1. Deskripsi Data

Penelitian ini menggunakan penelitian jenis terapan dengan data sekunder yang diambil dari situs resmi pemerintah Provinsi Jawa Timur yang dapat diakses langsung secara daring. Situs tersebut memberikan informasi terkait data-data penambahan serta perkembangan kasus COVID-19 di berbagai daerah Jawa Timur. Terdapat delapan variabel yang akan digunakan dalam penelitian ini dan dapat dilihat pada Tabel 3.1. Pada penelitian ini akan dilakukan analisis kluster tiap bulan dalam periode Januari sampai November 2021. Berikut adalah infografis terkait penambahan kasus COVID-19 di Provinsi Jawa Timur pada tiap bulannya.



Gambar 4.1 Penambahan Kasus COVID-19 di Provinsi Jawa Timur

Dalam Gambar 4.1 dapat dilihat bahwa puncak penambahan kasus COVID-19 berada di bulan Juli 2021 yang kemudian mengalami penurunan hingga saat ini. Namun meskipun dalam kondisi laju penyebaran yang sudah menurun, penambahan kasus COVID-19 di Provinsi Jawa Timur masih terbilang cukup banyak jika dibandingkan dengan provinsi-provinsi lain di Indonesia.



Gambar 4.2 Penyebaran COVID-19 di Provinsi Jawa Timur (Bulan November 2021)

Dalam implementasinya, perhitungan dilakukan dengan menggunakan data bulan November 2021. Gambar 4.2 merepresentasikan penambahan kasus COVID-19 di berbagai daerah Provinsi Jawa Timur pada bulan November 2021. Pada Gambar 4.2 dapat dilihat bahwa setiap provinsi memiliki karakteristik yang berbeda-beda terkait banyaknya laju penambahan kasus COVID-19. Dengan memperhatikan Gambar 4.2, Kota Surabaya merupakan daerah yang memiliki penambahan kasus konfirmasi tertinggi. Selain itu Kabupaten Tuban, Sidoarjo, dan Malang menempati posisi kasus konfirmasi terbanyak setelah Kota Surabaya. Jika diperhatikan, kasus pasien sembuh cukup baik karena tidak banyak selisinya dengan kasus konfirmasi. Begitu juga dengan kasus pasien meninggal yang hanya berkisar di rentang satuan.

4.2. *Pre-processing*

4.2.1. Normalisasi Data

Data yang digunakan tidak selalu bersifat ideal untuk diproses ke dalam suatu algoritma. Dilakukan *pre-processing* terlebih dahulu pada data untuk mengatasi permasalahan tidak idealnya suatu data yang mana hal tersebut dapat mempengaruhi hasil dari proses algoritma. Langkah pertama tahap *pre-processing* data adalah dengan cara menormalisasi data. Normalisasi data dilakukan untuk menyeimbangkan data saat terdapat beberapa data yang memiliki jangkauan sangat besar dengan data lain. Dalam hal ini digunakan normalisasi min-max dengan mengubah data berkisar di rentang 0 – 1. Pertama-tama dicari nilai maksimal dan minimal pada semua data yang akan dinormalisasi, setelah itu akan dilakukan perhitungan. Berikut adalah contoh perhitungan normalisasi data menggunakan Persamaan (2.1). Pada manualisasi perhitungan akan ditunjukkan normalisasi data pada 3 provinsi dengan variabel x_1 (jumlah kasus konfirmasi).

Nilai maksimal (max_{x_1}) = 165

Nilai minimal (min_{x_1}) = 0

Maka perhitungannya adalah,

$$\begin{aligned}
 w'_{Tuban} &= \frac{w - min_{x_1}}{max_{x_1} - min_{x_1}} \\
 &= \frac{77 - 0}{165 - 0} = 0.46667 \\
 w'_{Banyuwangi} &= \frac{w - min_{x_1}}{max_{x_1} - min_{x_1}} \\
 &= \frac{25 - 0}{165 - 0} = 0.15151 \\
 w'_{Surabaya} &= \frac{w - min_{x_1}}{max_{x_1} - min_{x_1}} \\
 &= \frac{165 - 0}{165 - 0} = 1
 \end{aligned}$$

Normalisasi data dilakukan pada setiap data di 38 provinsi dan 8 variabel sehingga menghasilkan data baru yang telah dinormalisasi seperti pada lampiran L.

4.2.2. Uji Multikolinearitas

Setelah data dilakukan normalisasi, selanjutnya akan dilakukan uji multikolinearitas pada data. Model yang baik seharusnya tidak memiliki gejala multikolinearitas pada variabel-variabel yang digunakannya. Dengan kata lain, korelasi antar variabel-variabelnya tidak terlalu tinggi karena korelasi yang tinggi dapat mengakibatkan hasil perhitungan yang menjauhi akurat. Berikut adalah contoh perhitungan uji multikolinearitas pada variabel x_1 dan x_2 dengan menggunakan Persamaan (2.2), (2.3), dan (2.4). Tabel 4.1 merupakan perhitungan keseluruhan data pada x_1 dan x_2 untuk menghitung koefisien determinasi.

Tabel 4.1 Perhitungan x_1 dan x_2 untuk Uji Multikolinearitas

Kabupaten/Kota	x_1	x_2	x_1x_2	x_1^2	x_2^2
Kota Surabaya	1.00000	1.00000	1.00000	1.00000	1.00000
Kota Probolinggo	0.04242	0.06494	0.00275	0.00180	0.00422
Kota Pasuruan	0.06667	0.07792	0.00519	0.00444	0.00607
Kota Mojokerto	0.00000	0.00000	0.00000	0.00000	0.00000
Kota Malang	0.22424	0.34416	0.07717	0.05028	0.11844
Kota Madiun	0.03636	0.02597	0.00094	0.00132	0.00067
Kota Kediri	0.03030	0.03247	0.00098	0.00092	0.00105
Kota Blitar	0.04848	0.04545	0.00220	0.00235	0.00207
Kota Batu	0.03030	0.05195	0.00157	0.00092	0.00270
Kab. Tulungagung	0.26061	0.29870	0.07784	0.06792	0.08922
Kab. Tuban	0.46667	0.53896	0.25152	0.21778	0.29048
Kab. Trenggalek	0.30909	0.33117	0.10236	0.09554	0.10967
Kab. Sumenep	0.07879	0.04545	0.00358	0.00621	0.00207
Kab. Situbondo	0.05455	0.09091	0.00496	0.00298	0.00826
Kab. Sidoarjo	0.35152	0.40909	0.14380	0.12356	0.16736
Kab. Sampang	0.05455	0.06494	0.00354	0.00298	0.00422
Kab. Probolinggo	0.18788	0.32468	0.06100	0.03530	0.10541
Kab. Ponorogo	0.23636	0.25974	0.06139	0.05587	0.06747
Kab. Pasuruan	0.17576	0.17532	0.03081	0.03089	0.03074
Kab. Pamekasan	0.04242	0.05844	0.00248	0.00180	0.00342
Kab. Pacitan	0.28485	0.35065	0.09988	0.08114	0.12295
Kab. Ngawi	0.27879	0.29221	0.08146	0.07772	0.08539

Kab. Nganjuk	0.12121	0.11688	0.01417	0.01469	0.01366
Kab. Mojokerto	0.30909	0.37013	0.11440	0.09554	0.13700
Kab. Malang	0.34545	0.47403	0.16375	0.11934	0.22470
Kab. Magetan	0.15758	0.14935	0.02353	0.02483	0.02231
Kab. Madiun	0.07273	0.07792	0.00567	0.00529	0.00607
Kab. Lumajang	0.16364	0.12338	0.02019	0.02678	0.01522
Kab. Lamongan	0.15758	0.16883	0.02660	0.02483	0.02850
Kab. Kediri	0.15152	0.23377	0.03542	0.02296	0.05465
Kab. Jombang	0.12727	0.20779	0.02645	0.01620	0.04318
Kab. Jember	0.20606	0.25325	0.05218	0.04246	0.06413
Kab. Gresik	0.30909	0.36364	0.11240	0.09554	0.13223
Kab. Bondowoso	0.05455	0.05844	0.00319	0.00298	0.00342
Kab. Bojonegoro	0.33333	0.42208	0.14069	0.11111	0.17815
Kab. Blitar	0.07273	0.12987	0.00945	0.00529	0.01687
Kab. Banyuwangi	0.15152	0.24026	0.03640	0.02296	0.05772
Kab. Bangkalan	0.07273	0.05844	0.00425	0.00529	0.00342
Jumlah	7.06667	8.33117	2.80421	2.49778	3.22310

Dalam perhitungan pada Tabel 4.1 dapat dituliskan ringkasan di bawah ini sehingga dapat dihitung koefisien determinasinya (R^2).

$$\begin{aligned}\sum x_1 &= 7.06667, \sum x_2 = 8.33117, \sum x_1 x_2 = 2.80421, \\ \sum x_1^2 &= 2.49778, \sum x_2^2 = 3.22310\end{aligned}$$

$$\begin{aligned}
 R^2 &= \left[\frac{n \sum_{i=1}^n x_1 x_2 - \sum_{i=1}^n x_1 \sum_{i=1}^n x_2}{\sqrt{(n \sum_{i=1}^n (x_1)^2 - (\sum_{i=1}^n x_1)^2) - (n \sum_{i=1}^n (x_2)^2 - (\sum_{i=1}^n x_2)^2)}} \right]^2 \\
 &= \left[\frac{(38)(2.80421) - (7.06667)(8.33117)}{\sqrt{((38)(2.49778) - (7.06667)^2)((38)(3.22310) - (8.33117)^2)}} \right]^2 \\
 &= (0.97605)^2 = 0.95268
 \end{aligned}$$

Setelah mendapatkan nilai koefisien determinasi, selanjutnya dihitung nilai *VIF* dan *tolerance (TOL)* sebagai berikut.

$$\begin{aligned}
 VIF &= \frac{1}{1 - R^2} = \frac{1}{1 - 0.95268} = 21.133 \\
 TOL &= \frac{1}{VIF} = \frac{1}{21.133} = 0.04732
 \end{aligned}$$

Dari perhitungan nilai *VIF* dan *TOL* variabel x_1 dan x_2 , didapati bahwa terdapat gejala multikolinearitas antar variabelnya. Hal ini dapat dilihat pada nilai *VIF* sebesar $21.133 > 10$ dan *TOL* sebesar $0.04732 < 0.1$. Karena variabelnya memiliki gejala multikolinearitas, maka akan dilakukan reduksi pada salah satu variabelnya. Dalam perhitungan ini, variabel yang direduksi adalah variabel x_2 (jumlah kasus sembuh). Setelah dilakukan reduksi variabel selanjutnya akan dilakukan perhitungan ulang pada data yang sudah direduksi. Hal tersebut dilakukan pada semua variabel yang digunakan. Sehingga menghasilkan nilai *VIF* dan *TOL* sebagai berikut.

Tabel 4.2 Hasil Perhitungan Uji Multikolinearitas

Variabel	R^2	<i>TOL</i>	<i>VIF</i>
X1X3	0.20167	0.79833	1.25262
X1X4	0.73494	0.26506	3.77271

X1X5	0.23462	0.76538	1.30655
X1X6	0.03051	0.96949	1.03147
X1X7	0.27713	0.72287	1.38337
X1X8	0.62887	0.37113	2.69447
X3X4	0.19636	0.80364	1.24433
X3X5	0.00900	0.99100	1.00909
X3X6	0.09971	0.90029	1.11076
X3X7	0.18493	0.81507	1.22689
X3X8	0.13738	0.86262	1.15926
X4X5	0.00088	0.99912	1.00088
X4X6	0.00473	0.99527	1.00475
X4X7	0.24375	0.75625	1.32231
X4X8	0.46437	0.53563	1.86696
X5X6	0.03719	0.96281	1.03863
X5X7	0.02696	0.97304	1.02771
X5X8	0.12418	0.87582	1.14179
X6X7	0.01632	0.98368	1.01659
X6X8	0.00723	0.99277	1.00728
X7X8	0.24797	0.75203	1.32973

Berdasarkan Tabel 4.2 semua variabel sudah tidak memiliki gejala multikolinearitas sehingga data siap diproses ke dalam algoritma genetika untuk mencari pusat kluster awal algoritma k-Means. Demikian maka variabel yang diteruskan adalah 7 variabel yakni jumlah kasus konfirmasi (x_1), jumlah kasus meninggal (x_3), jumlah positif dengan riwayat bergejala (x_4), jumlah positif tanpa

riwayat gejala (x_5), jumlah kasus dengan riwayat kontak (x_6), jumlah kasus tanpa riwayat kontak (x_7), dan jumlah kasus dengan riwayat perjalanan (x_8).

4.3. Implementasi GA dan K-Means

4.3.1. Proses Algoritma Genetika

Pembangkitan Populasi Acak

Tahap awal dari algoritma genetika adalah pembangkitan individu-individu dalam suatu populasi. Individu-individu tersebut merupakan kumpulan solusi yang dibangkitkan secara acak. Pada kasus ini, dalam satu individu terdiri dari beberapa gen yang mana gen tersebut berjumlah banyaknya variabel dikali banyaknya kelas (klaster). Sehingga jika menginisialisasi $k = 3$, maka jumlah gen dalam satu individu adalah 21 ($7 \times 3 = 21$). Berikut adalah contoh pembangkitan populasi sebanyak 10 individu.

Tabel 4.3 Pembangkitan Individu dalam Populasi

IND-1	0.48643	0.13533	0.84778	0.04342	0.62092	0.06526	0.10416
	0.22736	0.72473	0.54375	0.62910	0.90425	0.44209	0.14925
	0.92019	0.84938	0.32533	0.47713	0.70284	0.39695	0.57521
IND-2	0.62728	0.08777	0.74430	0.21833	0.48187	0.35687	0.24883
	0.11062	0.08397	0.05793	0.35922	0.19832	0.85703	0.76839
	0.75726	0.66102	0.01478	0.59463	0.92607	0.76314	0.38582
IND-3	0.30299	0.79662	0.30195	0.01893	0.35354	0.34501	0.06761
	0.28822	0.12949	0.77426	0.62423	0.60511	0.26175	0.93414
	0.41328	0.78743	0.26814	0.97143	0.10704	0.19196	0.36811
IND-4	0.48601	0.88481	0.71721	0.29978	0.66065	0.94416	0.64157
	0.96155	0.57801	0.48870	0.46958	0.06495	0.07456	0.69036

	0.21663	0.78721	0.29787	0.07639	0.40062	0.71540	0.47373
IND-5	0.60470	0.44081	0.37169	0.37962	0.00718	0.46885	0.93726
	0.50343	0.12259	0.56854	0.50884	0.59995	0.16967	0.09925
	0.55880	0.27102	0.91366	0.70021	0.45009	0.10156	0.80032
IND-6	0.55445	0.59951	0.65114	0.25786	0.27004	0.28995	0.31731
	0.82463	0.04164	0.47160	0.38681	0.56299	0.57903	0.50508
	0.48419	0.33506	0.29460	0.47805	0.93195	0.96784	0.87624
IND-7	0.04716	0.19512	0.08925	0.32067	0.01378	0.32696	0.02918
	0.20779	0.80640	0.28735	0.57881	0.37337	0.31387	0.62187
	0.99958	0.70463	0.61969	0.87575	0.65900	0.93958	0.81301
IND-8	0.49997	0.04687	0.90614	0.77583	0.73953	0.99528	0.79122
	0.94187	0.18159	0.17951	0.84630	0.43503	0.67854	0.12008
	0.36874	0.93802	0.67830	0.87329	0.72655	0.44328	0.83464
IND-9	0.92033	0.82761	0.15118	0.73132	0.84019	0.33305	0.52521
	0.92806	0.65553	0.70541	0.65432	0.25747	0.16139	0.10373
	0.15539	0.72528	0.36683	0.72570	0.29051	0.20278	0.09046
IND-10	0.88498	0.46318	0.92565	0.42037	0.91339	0.97846	0.45595
	0.46468	0.35678	0.30807	0.93994	0.91888	0.52269	0.82920
	0.26683	0.99695	0.30635	0.95591	0.10475	0.92236	0.10520

Perhitungan Nilai Fitness

Setelah dilakukan pembangkitan individu, selanjutnya adalah perhitungan nilai fitness untuk melihat ketahanan tiap-tiap individu. Sesuai dengan konsep evolusi, individu dengan fit (ketahanan) yang baik akan bertahan dengan baik pula. Langkah pertama yang dilakukan adalah dengan menghitung jarak dari tiap solusi

ke data menggunakan Persamaan (2.8). Berikut adalah contoh perhitungan jarak solusi (IND-1) dengan data ke-1.

IND-1						
0.48643	0.13533	0.84778	0.04342	0.62092	0.06526	0.10416
0.22736	0.72473	0.54375	0.62910	0.90425	0.44209	0.14925
0.92019	0.84938	0.32533	0.47713	0.70284	0.39695	0.57521

Data-1						
1.00000	0.75000	1.00000	0.38806	0.11538	1.00000	1.00000

$$\begin{aligned}
 J1 &= \sum_p |x_n^p - w_r^p| \\
 &= |x_1^1 - w_1^1| + |x_1^2 - w_1^2| + \dots + |x_1^7 - w_1^7| \\
 &= |1.00000 - 0.48643| + |0.75000 - 0.13533| + \dots + |1.00000 - 0.10416| \\
 &= 3.96122
 \end{aligned}$$

$$\begin{aligned}
 J2 &= \sum_p |x_n^p - w_r^p| \\
 &= |x_1^1 - w_2^1| + |x_1^2 - w_2^2| + \dots + |x_1^7 - w_2^7| \\
 &= |1.00000 - 0.22736| + |0.75000 - 0.72473| + \dots + |1.00000 - 0.14925| \\
 &= 3.69271
 \end{aligned}$$

$$\begin{aligned}
 J3 &= \sum_p |x_n^p - w_r^p| \\
 &= |x_1^1 - w_3^1| + |x_1^2 - w_3^2| + \dots + |x_1^7 - w_3^7| \\
 &= |1.00000 - 0.92019| + |0.75000 - 0.84938| + \dots + |1.00000 - 0.57521| \\
 &= 2.55822
 \end{aligned}$$

Setelah mendapatkan jarak pada tiap-tiap klaster, selanjutnya akan diambil nilai minimumnya. Proses perhitungan jarak dilakukan pada seluruh data dengan IND-1. Sehingga didapatkan Tabel 4.6 dan kemudian diambil jarak terdekat antara data pada klasternya menggunakan Persamaan (2.9).

Tabel 4.6 Perhitungan Jarak *Manhattan Distance*

Data-i	J1	J2	J3	Minimal
1	3.96122	3.69271	2.55822	2.55822
2	2.10050	3.18840	3.81488	2.10050
3	2.14856	3.34369	3.97018	2.14856
4	2.30329	3.62054	4.24702	2.30329
5	2.53871	2.57913	3.38569	2.53871
6	2.20830	3.52555	4.15203	2.20830
7	1.88125	3.19849	3.82498	1.88125
8	2.21650	2.97845	3.60494	2.21650
9	2.22480	3.47964	4.10612	2.22480
10	2.35937	1.19777	2.11004	1.19777
11	1.84544	2.43163	2.76338	1.84544
12	2.90736	3.21671	3.98368	2.90736
13	2.21605	3.26193	3.88841	2.21605

14	2.16682	3.00202	3.62850	2.16682
15	2.94774	3.12140	3.31066	2.94774
16	1.89294	3.14777	3.77426	1.89294
17	1.85125	2.93644	3.56293	1.85125
18	2.92380	2.08606	2.29771	2.08606
19	2.73304	2.39222	2.89995	2.39222
20	2.03678	3.32147	3.94795	2.03678
21	3.05828	2.16833	2.63698	2.16833
22	2.42320	2.48374	2.64401	2.42320
23	2.39374	3.20082	3.82730	2.39374
24	1.74977	2.91094	2.99304	1.74977
25	2.11781	1.23012	1.70741	1.23012
26	2.04551	2.73904	3.36553	2.04551
27	2.19226	3.29784	3.92432	2.19226
28	2.34563	2.59127	3.21776	2.34563
29	2.36745	2.56395	3.11484	2.36745
30	3.08139	2.86442	3.58119	2.86442
31	1.82222	3.04657	3.66940	1.82222
32	2.77695	2.18520	2.90198	2.18520
33	2.66806	2.88722	3.26205	2.66806
34	2.08275	3.33759	3.96407	2.08275
35	1.57428	3.10346	3.65872	1.57428
36	2.23650	3.28238	3.90886	2.23650
37	2.22480	2.91184	3.50589	2.22480
38	2.00671	3.07409	3.70057	2.00671

Jumlah (J)	82.30150
------------	----------

Setelah didapatkan jumlah dari keseluruhan nilai minimum, selanjutnya dicari *fitness value* pada IND-1 dengan menghitung nilai *invers*-nya menggunakan Persamaan (2.10).

$$F = \frac{1}{J}$$

$$= \frac{1}{82.30150} = 0.01215$$

Perhitungan seperti di atas diterapkan terhadap semua individu yang ada, sehingga didapatkan nilai fitness pada tiap-tiap individu yang dapat diamati pada Tabel 4.7.

Tabel 4.7 Nilai Fitness pada Semua Individu

Individu	Nilai Fitness
IND-1	0.01215
IND-2	0.01214
IND-3	0.01428
IND-4	0.01215
IND-5	0.01238
IND-6	0.01136
IND-7	0.02374
IND-8	0.00933
IND-9	0.01310
IND-10	0.00926

Proses Seleksi Individu

Selanjutnya masuk pada tahap seleksi. Pada tahap ini dilakukan seleksi individu menggunakan mesin *roulette*. Dengan harapan individu yang memiliki nilai fitness tinggi akan sering terpilih untuk kemudian dijadikan *parent* (induk). Pada tahap ini seluruh individu akan diberikan probabilitas dengan rentang 0 – 1. Kemudian akan dibangkitkan p secara acak ($p = [0, 1]$) sebanyak n populasi untuk menentukan individu mana yang akan terpilih. Berikut adalah contoh perhitungan dalam proses seleksi individu.

Tabel 4.8 Perhitungan Probabilitas dan Kumulatif

Individu	Nilai Fitness	Probabilitas	Kumulatif	Rentang
IND-1	0.01215	0.09355	0.09355	0.00000-0.09355
IND-2	0.01214	0.09351	0.18706	0.09356-0.18706
IND-3	0.01428	0.10998	0.29704	0.18707-0.29704
IND-4	0.01215	0.09351	0.39055	0.29705-0.39055
IND-5	0.01238	0.09532	0.48587	0.39056-0.48587
IND-6	0.01136	0.08743	0.57330	0.48588-0.57330
IND-7	0.02374	0.18275	0.75605	0.57331-0.75605
IND-8	0.00933	0.07186	0.82791	0.75606-0.82791
IND-9	0.01310	0.10083	0.92874	0.82792-0.92874
IND-10	0.00926	0.07126	1.00000	0.92875-1.00000

Pada Tabel 4.8 kolom pertama merupakan individu ke- i , kolom kedua adalah nilai fitness yang didapatkan pada perhitungan sebelumnya, kolom ketiga merupakan probabilitas pada masing-masing individu, dan kolom keempat adalah kumulatif dari individu 1 sampai 10. Pada perhitungan nilai probabilitas pertama-tama adalah dengan menjumlahkan semua nilai fitness individu ($\sum F$). Selanjutnya nilai fitness tiap individu dibagi dengan total keseluruhan fitness. Sedangkan untuk menghitung nilai kumulatif individu- i , dilakukan penjumlahan probabilitas individu- i dengan nilai kumulatif- $(i - 1)$.

$$\sum F = 0.12988$$

$$\begin{aligned} Prob_1 &= \frac{F_1}{\sum F} \\ &= \frac{0.01215}{0.12988} = 0.09355 \end{aligned}$$

$$\begin{aligned} Prob_2 &= \frac{F_2}{\sum F} \\ &= \frac{0.01214}{0.12988} = 0.09351 \end{aligned}$$

$$\begin{aligned} Prob_3 &= \frac{F_3}{\sum F} \\ &= \frac{0.01428}{0.12988} = 0.10998 \end{aligned}$$

$$Kum_1 = 0 + 0.09355 = 0.09355$$

$$Kum_2 = 0.09355 + 0.09351 = 0.18706$$

$$Kum_3 = 0.18706 + 0.10998 = 0.29704$$

Tujuan dilakukan perhitungan kumulatif adalah sebagai rentang nilai probabilitas pada setiap individu. Pada individu pertama berada di rentang $0 - 0.09355$, pada individu kedua berada di rentang $0.09356 - 0.18706$, dan seterusnya. Kemudian akan dilakukan pembangkitan bilangan acak p sebanyak n

individu (10 kali) untuk menentukan individu mana saja yang akan terseleksi. Berikut adalah Pembangkitan bilangan acak p dan individu yang terseleksi.

Tabel 4.9 Proses Seleksi Individu

Acak (p)	Individu Terseleksi
0.69837	IND-7
0.81093	IND-8
0.42609	IND-5
0.88551	IND-9
0.49658	IND-6
0.10206	IND-2
0.39937	IND-5
0.72947	IND-7
0.14364	IND-2
0.18592	IND-2

Pada seleksi pertama, dengan memperhatikan Tabel 4.9 IND-7 terpilih karena nilai $p = 0.69837$ berada di rentang $0.57331 - 0.75605$. Pada seleksi kedua, IND-8 terpilih karena nilai $p = 0.81093$ berada di rentang $0.75606 - 0.82791$. Hal serupa diperlakukan sama sampai individu ke-10. Individu-individu yang telah terpilih akan menjadi calon *parent* untuk menghasilkan *offspring* atau solusi baru pada proses reproduksi gen.

Proses Reproduksi *Crossover*

Proses reproduksi gen yang pertama adalah proses *crossover*. Individu yang terpilih pada tahap sebelumnya akan dipilih secara acak sebanyak dua individu lalu

dilakukan *crossover*. Banyaknya *offspring* yang dihasilkan bergantung pada banyaknya individu dan *crossover rate* yang digunakan ($Cr \times Popsiz$ e). Jumlah *offspring* pada proses *crossover* selalu bersifat genap dengan pembulatan ke atas. Pada penelitian ini jumlah *offspring* yang didapatkan pada proses *crossover* adalah 2 karena Cr yang digunakan adalah 0.2 ($0.2 \times 10 = 2$). Metode *crossover* yang digunakan adalah *crossover n point* ($n = 3$). Pertama dua individu dipilih secara acak, pada iterasi pertama individu yang terpilih adalah IND-7 dan IND-8.

IND-7						
0.04716	0.19512	0.08925	0.32067	0.01378	0.32696	0.02918
0.20779	0.80640	0.28735	0.57881	0.37337	0.31387	0.62187
0.99958	0.70463	0.61969	0.87575	0.65900	0.93958	0.81301
IND-8						
0.49997	0.04687	0.90614	0.77583	0.73953	0.99528	0.79122
0.94187	0.18159	0.17951	0.84630	0.43503	0.67854	0.12008
0.36874	0.93802	0.67830	0.87329	0.72655	0.44328	0.83464

Karena titik yang digunakan adalah $n = 3$, maka tiap individu akan terbagi menjadi 4 bagian. Penentuan ketiga titik akan dilakukan secara acak, pada iterasi pertama didapatkan titik yang berturut-turut 3, 7, dan 12. Berikut adalah bentuk pembagian individu menjadi 4 bagian.

IND-7 (Parent-1)						
0.04716	0.19512	0.08925	0.32067	0.01378	0.32696	0.02918
0.20779	0.80640	0.28735	0.57881	0.37337	0.31387	0.62187
0.99958	0.70463	0.61969	0.87575	0.65900	0.93958	0.81301
IND-8 (Parent-2)						

0.49997	0.04687	0.90614	0.77583	0.73953	0.99528	0.79122
0.94187	0.18159	0.17951	0.84630	0.43503	0.67854	0.12008
0.36874	0.93802	0.67830	0.87329	0.72655	0.44328	0.83464

Setelah dibagi menjadi 4 bagian, selanjutnya adalah proses pertukaran gen-gen pada masing-masing individu. Gen *offspring-1* diambil dari gen *parent-1* bagian satu dan tiga, sisanya diambil dari *parent-2*. Sedangkan gen *offspring-2* diambil dari gen *parent-2* bagian satu dan tiga, sisanya diambil dari *parent-1*. Sehingga dihasilkan dua *offspring* sebagai berikut.

OFF-1						
0.04716	0.19512	0.08925	0.77583	0.73953	0.99528	0.79122
0.20779	0.80640	0.28735	0.57881	0.37337	0.67854	0.12008
0.36874	0.93802	0.67830	0.87329	0.72655	0.44328	0.83464
OFF-2						
0.49997	0.04687	0.08925	0.32067	0.01378	0.32696	0.02918
0.99958	0.70463	0.61969	0.87575	0.65900	0.31387	0.62187
0.99958	0.70463	0.61969	0.87575	0.65900	0.93958	0.81301

Proses Reproduksi Mutasi Gen

Setelah dilakukan reproduksi gen menggunakan operator *crossover*, selanjutnya adalah reproduksi gen menggunakan operator mutasi. Satu individu akan dipilih secara acak untuk kemudian dilakukan mutasi pada gennya. Banyaknya *offspring* yang dihasilkan bergantung pada banyaknya individu dan *mutation rate* yang digunakan ($Mr \times Popsiz$ e).

Pada penelitian ini, *offspring* yang dihasilkan adalah 2 karena Mr yang digunakan adalah 0.2 ($0.2 \times 10 = 2$) dengan kata lain untuk menghasilkan dua *offspring*, mutasi gen dilakukan sebanyak dua kali. Pada iterasi pertama, individu yang terpilih secara acak untuk mutasi pertama adalah IND-5. Sedangkan individu yang terpilih secara acak untuk mutasi kedua adalah IND-2.

IND-5 (<i>Parent-3</i>)						
0.60470	0.44081	0.37169	0.37962	0.00718	0.46885	0.93726
0.50343	0.12259	0.56854	0.50884	0.59995	0.16967	0.09925
0.55880	0.27102	0.91366	0.70021	0.45009	0.10156	0.80032

IND-2 (<i>Parent-4</i>)						
0.62728	0.08777	0.74430	0.21833	0.48187	0.35687	0.24883
0.11062	0.08397	0.05793	0.35922	0.19832	0.85703	0.76839
0.75726	0.66102	0.01478	0.59463	0.92607	0.76314	0.38582

Pertama-tama akan dibangkitkan nilai berupa bilangan bulat secara acak untuk menentukan gen mana yang akan dimutasi. Didapatkan bilangan bulat berturut-turut 10 dan 7, maka gen-10 pada *parent-3* ($x_i = 0.56854$) dan gen-7 pada *parent-4* ($x_i = 0.24883$) akan dilakukan mutasi. Setelah itu ditambahkan dengan nilai r dengan interval $[-0.1, 0.1]$ yang didapatkan secara acak pula sehingga didapatkan $r = -0.0521$. Berikut adalah pembangkitan nilai *gen offspring* yang baru pada iterasi pertama menggunakan Persamaan (2.7).

$$\begin{aligned}
 x_i'(off-3) &= x_i + r(max_i - min_i) \\
 &= 0.56854 + (-0.0521)(0.93726 - 0.00718) = 0.52008
 \end{aligned}$$

$$\begin{aligned}
 x_i'(off-4) &= x_i + r(max_i - min_i) \\
 &= 0.24883 + (-0.0521)(0.92607 - 0.01478) = 0.20135
 \end{aligned}$$

OFF-3						
0.60470	0.44081	0.37169	0.37962	0.00718	0.46885	0.93726
0.50343	0.12259	0.52008	0.50884	0.59995	0.16967	0.09925
0.55880	0.27102	0.91366	0.70021	0.45009	0.10156	0.80032

OFF-4						
0.62728	0.08777	0.74430	0.21833	0.48187	0.35687	0.20135
0.11062	0.08397	0.05793	0.35922	0.19832	0.85703	0.76839
0.75726	0.66102	0.01478	0.59463	0.92607	0.76314	0.38582

Proses Elitism

Setelah melalui proses reproduksi gen, tentunya jumlah individu akan semakin bertambah banyak. Pada awalnya hanya terdapat 10 individu kemudian setelah dilakukan reproduksi menjadi 14 individu dengan 4 individu tambahan (*offspring*). Karena individu menjadi semakin banyak pada tiap penambahan iterasi, maka dilakukan perankingan menggunakan seleksi elitism. Seluruh individu akan dihitung nilai fitnessnya untuk kemudian dapat dilakukan perankingan. Tabel 4.17 merepresentasikan populasi beserta nilai fitness individu setelah proses seleksi dan reproduksi gen.

Tabel 4.17 Populasi setelah Seleksi dan Reproduksi Gen

IND-7							Fitness
0.04716	0.19512	0.08925	0.32067	0.01378	0.32696	0.02918	0.02374
0.20779	0.80640	0.28735	0.57881	0.37337	0.31387	0.62187	
0.99958	0.70463	0.61969	0.87575	0.65900	0.93958	0.81301	
IND-8							Fitness
0.49997	0.04687	0.90614	0.77583	0.73953	0.99528	0.79122	0.00933
0.94187	0.18159	0.17951	0.84630	0.43503	0.67854	0.12008	
0.36874	0.93802	0.67830	0.87329	0.72655	0.44328	0.83464	
IND-5							Fitness
0.60470	0.44081	0.37169	0.37962	0.00718	0.46885	0.93726	0.01238
0.50343	0.12259	0.56854	0.50884	0.59995	0.16967	0.09925	
0.55880	0.27102	0.91366	0.70021	0.45009	0.10156	0.80032	
IND-9							Fitness
0.92033	0.82761	0.15118	0.73132	0.84019	0.33305	0.52521	0.01310
0.92806	0.65553	0.70541	0.65432	0.25747	0.16139	0.10373	
0.15539	0.72528	0.36683	0.72570	0.29051	0.20278	0.09046	
IND-6							Fitness
0.55445	0.59951	0.65114	0.25786	0.27004	0.28995	0.31731	0.01136
0.82463	0.04164	0.47160	0.38681	0.56299	0.57903	0.50508	
0.48419	0.33506	0.29460	0.47805	0.93195	0.96784	0.87624	
IND-2							Fitness
0.62728	0.08777	0.74430	0.21833	0.48187	0.35687	0.24883	0.01214
0.11062	0.08397	0.05793	0.35922	0.19832	0.85703	0.76839	
0.75726	0.66102	0.01478	0.59463	0.92607	0.76314	0.38582	

IND-5							Fitness
0.60470	0.44081	0.37169	0.37962	0.00718	0.46885	0.93726	0.01238
0.50343	0.12259	0.56854	0.50884	0.59995	0.16967	0.09925	
0.55880	0.27102	0.91366	0.70021	0.45009	0.10156	0.80032	
IND-7							Fitness
0.04716	0.19512	0.08925	0.32067	0.01378	0.32696	0.02918	0.02374
0.20779	0.80640	0.28735	0.57881	0.37337	0.31387	0.62187	
0.99958	0.70463	0.61969	0.87575	0.65900	0.93958	0.81301	
IND-2							Fitness
0.62728	0.08777	0.74430	0.21833	0.48187	0.35687	0.24883	0.01214
0.11062	0.08397	0.05793	0.35922	0.19832	0.85703	0.76839	
0.75726	0.66102	0.01478	0.59463	0.92607	0.76314	0.38582	
IND-2							Fitness
0.62728	0.08777	0.74430	0.21833	0.48187	0.35687	0.24883	0.01214
0.11062	0.08397	0.05793	0.35922	0.19832	0.85703	0.76839	
0.75726	0.66102	0.01478	0.59463	0.92607	0.76314	0.38582	
OFF-1							Fitness
0.04716	0.19512	0.08925	0.77583	0.73953	0.99528	0.79122	0.00970
0.20779	0.80640	0.28735	0.57881	0.37337	0.67854	0.12008	
0.36874	0.93802	0.67830	0.87329	0.72655	0.44328	0.83464	
OFF-2							Fitness
0.04716	0.19512	0.08925	0.77583	0.73953	0.99528	0.79122	0.01321
0.20779	0.80640	0.28735	0.57881	0.37337	0.67854	0.12008	
0.36874	0.93802	0.67830	0.87329	0.72655	0.44328	0.83464	
OFF-3							Fitness

0.60470	0.44081	0.37169	0.37962	0.00718	0.46885	0.93726	0.01260
0.50343	0.12259	0.52008	0.50884	0.59995	0.16967	0.09925	
0.55880	0.27102	0.91366	0.70021	0.45009	0.10156	0.80032	
OFF-4							Fitness
0.62728	0.08777	0.74430	0.21833	0.48187	0.35687	0.20135	0.01216
0.11062	0.08397	0.05793	0.35922	0.19832	0.85703	0.76839	
0.75726	0.66102	0.01478	0.59463	0.92607	0.76314	0.38582	

Selanjutnya adalah dilakukan perankingan atau pengurutan individu dari nilai fitness yang tertinggi yang kemudian diambil 10 individu untuk diteruskan pada iterasi selanjutnya. Sedangkan 4 individu lainnya yang berada di posisi paling bawah akan direduksi. Pada Tabel 4.17 dapat dilihat bahwa 4 nilai fitness terkecil adalah satu IND-2, IND-6, OFF-1, dan IND-8. Sehingga 4 individu tersebut harus direduksi. Berikut adalah 10 individu dengan nilai fitness terbaik yang akan diteruskan ke dalam iterasi selanjutnya.

Tabel 4.18 Populasi setelah Elitism

Individu	Nilai Fitness
IND-7	0.02374
IND-7	0.02374
OFF-2	0.01321
IND-9	0.01310
OFF-3	0.01260
IND-5	0.01238
IND-5	0.01238
OFF-4	0.01216

IND-2	0.01214
IND-2	0.01214

Individu-individu pada Tabel 4.18 akan diteruskan ke dalam iterasi selanjutnya. Pada penelitian ini, pemberhentian iterasi dilihat dari jumlah generasi yang ditetapkan di awal yakni sebesar $I = 10$. Menggunakan bantuan komputer, maka didapatkan pusat klaster terpilih dengan fitness sebesar 0.02397. Berikut adalah pusat klaster terpilih dari generasi ke-10 yang kemudian akan menjadi pusat klaster awal untuk optimasi algoritma k-Means pada tahap selanjutnya.

Tabel 4.19 Pusat Klaster Terpilih

0.10706	0.19512	0.09925	0.32067	0.01378	0.32696	0.02811
0.20779	0.80640	0.28735	0.57881	0.37337	0.35387	0.62187
0.99958	0.70463	0.61969	0.87575	0.65900	0.93958	0.81301

4.3.2. Proses K-Means

Perhitungan Algoritma GAK-Means (Iterasi = 1)

Selanjutnya adalah pengklasteran menggunakan algoritma k-Means. Pertama akan ditentukan nilai k terlebih dahulu. Inisialisasi nilai k yang digunakan adalah sebanyak tiga klaster. Selanjutnya adalah mengambil pusat klaster awal yang telah dioptimasi menggunakan GA melalui tahap sebelumnya. Lalu akan dihitung jarak tiap-tiap data ke pusat klaster menggunakan Persamaan (2.6). Berikut adalah contoh perhitungan jarak data ke-1 dengan pusat klaster.

Tabel 4.20 Pusat Klaster Awal (GAK-Means Iterasi = 1)

Pusat Klaster Awal

0.10706	0.19512	0.09925	0.32067	0.01378	0.32696	0.02811
0.20779	0.80640	0.28735	0.57881	0.37337	0.35387	0.62187
0.99958	0.70463	0.61969	0.87575	0.65900	0.93958	0.81301

Data-1						
1.00000	0.75000	1.00000	0.38806	0.11538	1.00000	1.00000

$$\begin{aligned}
 ED_1 &= \sqrt{\sum_{i=1}^m (x_i - y_i)^2} \\
 &= \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_7 - y_7)^2} \\
 &= \sqrt{(1.00000 - 0.10706)^2 + (0.75000 - 0.19512)^2 + \dots + (1.00000 - 0.02811)^2} \\
 &= 1.82455 \\
 ED_2 &= \sqrt{\sum_{i=1}^m (x_i - y_i)^2} \\
 &= \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_7 - y_7)^2} \\
 &= \sqrt{(1.00000 - 0.20779)^2 + (0.75000 - 0.80640)^2 + \dots + (1.00000 - 0.62187)^2} \\
 &= 1.34240 \\
 ED_3 &= \sqrt{\sum_{i=1}^m (x_i - y_i)^2} \\
 &= \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_7 - y_7)^2} \\
 &= \sqrt{(1.00000 - 0.99958)^2 + (0.75000 - 0.70463)^2 + \dots + (1.00000 - 0.81301)^2} \\
 &= 0.84775
 \end{aligned}$$

Setelah dilakukan perhitungan jarak, selanjutnya diambil nilai minimum yang merepresentasikan jarak terdekat antara data dengan klasternya. Perhitungan

jarak data ke pusat kluster diterapkan pada seluruh data. Tabel 4.22 merupakan perhitungan jarak data pada pusat kluster di 38 data.

Tabel 4.22 Perhitungan Algoritma GAK-Means (Iterasi = 1)

Kabupaten/Kota	C1	C2	C3	Klaster
Kota Surabaya	1.82455	1.34240	0.84775	C3
Kota Probolinggo	0.45995	1.13438	2.01415	C1
Kota Pasuruan	0.45066	1.24975	2.04780	C1
Kota Mojokerto	0.51971	1.32756	2.15051	C1
Kota Malang	0.44186	1.02172	1.80698	C1
Kota Madiun	0.49041	1.30355	2.11355	C1
Kota Kediri	0.52194	1.19428	2.02739	C1
Kota Blitar	0.40851	1.04564	1.94481	C1
Kota Batu	0.47065	1.27533	2.09349	C1
Kab. Tulungagung	1.17371	0.89001	1.53626	C2
Kab. Tuban	0.65777	0.95723	1.64183	C1
Kab. Trenggalek	0.81124	1.25374	1.83052	C1
Kab. Sumenep	0.42280	1.20373	2.00987	C1
Kab. Situbondo	0.44249	1.05291	1.95772	C1
Kab. Sidoarjo	0.78234	1.04609	1.60127	C1
Kab. Sampang	0.44695	1.15132	1.99043	C1
Kab. Probolinggo	0.45165	1.11978	1.92590	C1
Kab. Ponorogo	0.85184	0.60638	1.54047	C2
Kab. Pasuruan	0.51708	0.79143	1.71976	C1
Kab. Pamekasan	0.43683	1.19290	2.04316	C1
Kab. Pacitan	0.58371	0.77531	1.35156	C1

Kab. Ngawi	0.57252	0.72813	1.61882	C1
Kab. Nganjuk	0.39560	1.21255	1.99081	C1
Kab. Mojokerto	0.64375	0.97728	1.72815	C1
Kab. Malang	0.82987	0.74278	1.21822	C2
Kab. Magetan	0.30743	1.10298	1.80810	C1
Kab. Madiun	0.43165	1.21614	2.02359	C1
Kab. Lumajang	0.59328	0.97167	1.87642	C1
Kab. Lamongan	0.49389	0.85774	1.78924	C1
Kab. Kediri	0.66866	1.11386	1.62420	C1
Kab. Jombang	0.45230	1.12569	1.94498	C1
Kab. Jember	0.48089	0.84045	1.51792	C1
Kab. Gresik	0.69739	1.01751	1.66200	C1
Kab. Bondowoso	0.42552	1.20839	2.04893	C1
Kab. Bojonegoro	0.62278	1.29002	1.93277	C1
Kab. Blitar	0.42383	1.20925	2.01744	C1
Kab. Banyuwangi	0.38067	1.09888	1.86775	C1
Kab. Bangkalan	0.31689	1.13801	1.94119	C1

Dengan melihat Tabel 4.22 maka dapat disimpulkan bahwa data yang berada di klaster 1 berjumlah 34 data, klaster 2 berjumlah 3 data, dan klaster 3 berjumlah 1 data. Setelah mendapatkan hasil pada iterasi-1, selanjutnya akan dihitung kembali pada iterasi-2 dengan memperbarui pusat klaster (*centroid*) terlebih dahulu.

Perhitungan Algoritma GAK-Means (Iterasi = 2)

Pada iterasi selanjutnya, dilakukan perhitungan pusat kluster terlebih dahulu menggunakan Persamaan 2.5. Dimana, N_i merupakan jumlah anggota kluster ke- i dan $\sum X_{kj}$ merupakan penjumlahan seluruh data pada variabel ke- j . Berikut adalah contoh perhitungan nilai pertama pada tiap kluster.

$$N_1 = 34$$

$$N_2 = 3$$

$$N_3 = 1$$

$$\begin{aligned} V_{11} &= \frac{1}{N_1} \sum_{k=0}^{N_1} X_{k1} \\ &= \frac{1}{34} (0.04242 + 0.06667 + \dots + 0.07273) \\ &= 0.15365 \end{aligned}$$

$$\begin{aligned} V_{21} &= \frac{1}{N_2} \sum_{k=0}^{N_2} X_{k1} \\ &= \frac{1}{3} (0.26061 + 0.23636 + 0.34545) \\ &= 0.28081 \end{aligned}$$

$$\begin{aligned} V_{31} &= \frac{1}{N_3} \sum_{k=0}^{N_3} X_{k1} \\ &= \frac{1}{1} (1.00000) \\ &= 1.00000 \end{aligned}$$

Perhitungan seperti contoh di atas dilakukan pada data di semua variabel. Sehingga didapatkan pusat kluster baru yang dapat diamati pada Tabel 4.23.

Tabel 4.23 Pusat Klaster Baru (GAK-Means Iterasi = 2)

0.15365	0.19118	0.08252	0.21642	0.02149	0.09079	0.09543
0.28081	0.70833	0.13189	0.41791	0.61538	0.26087	0.17506
1.00000	0.75000	1.00000	0.38806	0.11538	1.00000	1.00000

Setelah didapatkan pusat klaster baru, selanjutnya adalah menghitung ulang jarak terdekat data dengan pusat klaster yang terbaru menggunakan Persamaan (2.6). Sehingga didapatkan hasil pengelompokan baru yang dapat dilihat dalam Tabel 4.24.

Tabel 4.24 Perhitungan Algoritma GAK-Means (Iterasi = 2)

Kabupaten/Kota	C1	C2	C3	Klaster
Kota Surabaya	1.88508	1.65841	0.00000	C3
Kota Probolinggo	0.26314	0.92491	2.03548	C1
Kota Pasuruan	0.28277	1.02230	2.11158	C1
Kota Mojokerto	0.36279	1.11805	2.17403	C1
Kota Malang	0.39508	0.85460	1.97413	C1
Kota Madiun	0.32529	1.09550	2.13916	C1
Kota Kediri	0.36020	0.91297	2.10338	C1
Kota Blitar	0.27479	0.85462	1.96458	C1
Kota Batu	0.30153	1.07973	2.13160	C1
Kab. Tulungagung	1.15005	0.43058	1.93789	C2
Kab. Tuban	0.57247	0.80921	1.64644	C1
Kab. Trenggalek	0.83626	1.15592	2.10420	C1
Kab. Sumenep	0.24068	1.03370	2.06772	C1

Kab. Situbondo	0.28064	0.86555	1.97087	C1
Kab. Sidoarjo	0.79742	1.09391	1.83717	C1
Kab. Sampang	0.25786	0.90574	2.06321	C1
Kab. Probolinggo	0.25699	0.91661	1.91792	C1
Kab. Ponorogo	0.84812	0.58914	1.62691	C2
Kab. Pasuruan	0.42993	0.72230	1.84039	C1
Kab. Pamekasan	0.23614	0.99994	2.07071	C1
Kab. Pacitan	0.76904	0.80780	1.42001	C1
Kab. Ngawi	0.44055	0.72641	1.64979	C1
Kab. Nganjuk	0.26261	1.01797	2.08713	C1
Kab. Mojokerto	0.46662	0.94535	1.63982	C1
Kab. Malang	0.90439	0.42832	1.60144	C2
Kab. Magetan	0.27936	0.94311	1.84729	C1
Kab. Madiun	0.24706	1.04194	2.07338	C1
Kab. Lumajang	0.47635	0.77863	1.88101	C1
Kab. Lamongan	0.33632	0.74638	1.81816	C1
Kab. Kediri	0.88866	1.03258	1.64484	C1
Kab. Jombang	0.22033	0.97489	1.93956	C1
Kab. Jember	0.64686	0.78737	1.53124	C1
Kab. Gresik	0.65317	0.98638	1.89273	C1
Kab. Bondowoso	0.23454	0.99867	2.08311	C1
Kab. Bojonegoro	0.48120	1.10731	1.88044	C1
Kab. Blitar	0.24513	1.03638	2.07704	C1
Kab. Banyuwangi	0.23725	0.95927	1.96178	C1
Kab. Bangkalan	0.21194	0.95785	1.95715	C1

Dengan melihat Tabel 4.24 maka dapat disimpulkan kembali bahwa data yang berada di klaster 1 berjumlah 34 data, klaster 2 berjumlah 3 data, dan klaster 3 berjumlah 1 data. Jika dibandingkan dengan hasil pengelompokan pada iterasi-1, pengelompokan sudah tidak mengalami perpindahan sehingga iterasi dapat dihentikan. Berikut adalah hasil pengelompokan menggunakan algoritma GAK-Means dengan $k = 3$.

Tabel 4.25 Pengelompokan Akhir GAK-Means

Klaster	Kab/Kota
Klaster 1	Kota Blitar, Kota Pasuruan, Kota Malang, Kota Mojokerto, Kota Probolinggo, Kota Madiun, Kota Batu, Kota Kediri, Kab. Tuban, Kab. Trenggalek, Kab. Sumenep, Kab. Situbondo, Kab. Sidoarjo, Kab. Sampang, Kab. Probolinggo, Kab. Pasuruan, Kab. Pamekasan, Kab. Pacitan, Kab. Ngawi, Kab. Nganjuk, Kab. Mojokerto, Kab. Magetan, Kab. Madiun, Kab. Lumajang, Kab. Lamongan, Kab. Kediri, Kab. Jombang, Kab. Jember, Kab. Gresik, Kab. Bondowoso, Kab. Bojonegoro, Kab. Blitar, Kab. Banyuwangi, dan Kab. Bangkalan.
Klaster 2	Kab. Tulungagung, Kab. Ponorogo, dan Kab. Malang
Klaster 3	Kota Surabaya

Perhitungan Algoritma K-Means tanpa Optimasi GA

Selain dilakukan perhitungan k-Means dengan optimasi GA, pada penelitian ini juga dilakukan pengelompokan dengan metode k-Means saja dengan tujuan sebagai perbandingan. Tahapan-tahapannya sama seperti sebelumnya yakni dengan inisialisasi jumlah k , menentukan secara acak pusat klaster awal, menghitung jarak

terdekat dari data ke pusat klaster, dilakukan pengulangan sampai tidak ada klaster yang berpindah. Jumlah klaster yang digunakan adalah $k = 3$ sedangkan Tabel 4.26 adalah pusat klaster yang didapatkan secara acak.

Tabel 4.26 Pusat Klaster Awal (K-Means Iterasi = 1)

0.09091	0.00808	0.04621	0.04470	0.00404	0.01288	0.04848
0.52727	0.01818	0.45606	0.07121	0.00455	0.03485	0.28788
0.27758	0.01879	0.06061	0.22970	0.03091	0.02303	0.17030

Setelah itu dihitung jarak menggunakan Persamaan (2.6) dan diambil jarak terkecil dari ketiga klaster sehingga menghasilkan pengelompokan sebagai berikut:

Tabel 4.27 Pengelompokan K-Means (Iterasi = 1)

Klaster	Kab/Kota
Klaster 1	Kota Mojokerto, Kota Probolinggo, Kota Pasuruan, Kota Kediri, Kota Madiun, Kota Batu, Kota Blitar, Kab. Sumenep, Kab. Situbondo, Kab. Sampang, Kab. Probolinggo, Kab. Pamekasan, Kab. Madiun, Kab. Lumajang, Kab. Jombang, Kab. Bondowoso, Kab. Blitar, dan Kab. Bangkalan.
Klaster 2	Kota Surabaya, Kab. Tuban, Kab. Mojokerto, dan Kab. Bojonegoro.
Klaster 3	Kab. Trenggalek, Kota Malang, Kab. Tulungagung, Kab. Sidoarjo, Kab. Ponorogo, Kab. Pasuruan, Kab. Pacitan, Kab. Ngawi, Kab. Nganjuk, Kab. Malang, Kab. Magetan, Kab. Lamongan, Kab. Kediri, Kab. Jember, Kab. Gresik, dan Kab. Banyuwangi.

Langkah selanjutnya adalah memperbarui pusat klaster dengan Persamaan (2.5). Dan didapatkan pusat klaster yang baru pada Tabel 4.28.

Tabel 4.28 Pusat Klaster Baru (K-Means Iterasi = 2)

0.06869	0.14583	0.04716	0.07131	0.02778	0.01691	0.04117
0.52727	0.37500	0.54137	0.17537	0.02885	0.25000	0.34173
0.23258	0.32813	0.07419	0.43843	0.12981	0.22283	0.16637

Setelah itu dilakukan perhitungan jarak kembali menggunakan Persamaan (2.6) dan diambil jarak terkecil dari ketiga klaster sehingga menghasilkan pengelompokan sebagai berikut:

Tabel 4.29 Pengelompokan K-Means (Iterasi = 2)

Klaster	Kab/Kota
Klaster 1	Kota Pasuruan, Kota Blitar, Kota Probolinggo, Kota Mojokerto, Kota Batu, Kota Madiun, Kota Kediri, Kab. Sumenep, Kab. Situbondo, Kab. Sampang, Kab. Madiun, Kab. Probolinggo, Kab. Pamekasan, Kab. Nganjuk, Kab. Magetan, Kab. Lumajang, Kab. Jombang, Kab. Bondowoso, Kab. Blitar, Kab. Bojonegoro, Kab. Banyuwangi, dan Kab. Bangkalan.
Klaster 2	Kota Surabaya, Kab. Tuban, dan Kab. Mojokerto.
Klaster 3	Kota Malang, Kab. Lamongan, Kab. Trenggalek, Kab. Sidoarjo, Kab. Tulungagung, Kab. Ponorogo, Kab. Pasuruan, Kab. Pacitan, Kab. Ngawi, Kab. Malang, Kab. Kediri, Kab. Jember, dan Kab. Gresik.

Karena data masih terus berpindah klaster, maka iterasi dilanjutkan sampai data sudah tidak berpindah klaster lagi. Dengan menggunakan bantuan komputer, perhitungan menggunakan k-Means tanpa optimasi GA baru berhenti di iterasi-4. Berikut adalah hasil akhir pengelompokan menggunakan metode k-Means tanpa optimasi GA.

Tabel 4.30 Pengelompokan Akhir K-Means

Klaster	Kab/Kota
Klaster 1	Kota Probolinggo, Kota Pasuruan, Kota Mojokerto, Kota Madiun, Kota Kediri, Kota Blitar, Kota Batu, Kab. Sumenep, Kab. Magetan, Kab. Situbondo, Kab. Sampang, Kab. Probolinggo, Kab. Pamekasan, Kab. Mojokerto, Kab. Nganjuk, Kab. Madiun, Kab. Lumajang, Kab. Lamongan, Kab. Jombang, Kab. Bondowoso, Kab. Blitar, Kab. Bojonegoro, Kab. Banyuwangi, dan Kab. Bangkalan.
Klaster 2	Kota Surabaya.
Klaster 3	Kota Malang, Kab. Tulungagung, Kab. Tuban, Kab. Trenggalek, Kab. Malang, Kab. Sidoarjo, Kab. Ponorogo, Kab. Pasuruan, Kab. Pacitan, Kab. Ngawi, Kab. Kediri, Kab. Jember, dan Kab. Gresik.

4.3.3. Proses Evaluasi Klaster

Langkah terakhir setelah dilakukan pengelompokan adalah evaluasi hasil klaster. Pengelompokan yang baik adalah ketika data dapat dikelompokkan ke dalam klaster yang memiliki karakteristik yang sama. Evaluasi klaster yang digunakan adalah metode *Silhouette Coefficient* (SC).

Perhitungan Evaluasi Kluster GAK-Means

Berikut akan dicontohkan perhitungan mencari nilai SC pengelompokan GAK-Means pada data yang berada di kluster 2 yakni data-10 (Kab. Tulungagung). Untuk mendapatkan nilai SC, akan dicari jarak data-10 ke data lain terlebih dahulu menggunakan *Euclidean Distance* dengan data-10 sebagai pusat klasternya.

Tabel 4.31 Perhitungan Jarak Euclidean (Kab. Tulungagung)

$d_{1,10}$	1.93789	$d_{11,10}$	1.11261	$d_{21,10}$	1.22511	$d_{31,10}$	1.24801
$d_{2,10}$	1.18343	$d_{12,10}$	1.39087	$d_{22,10}$	1.06594	$d_{32,10}$	1.19342
$d_{3,10}$	1.25535	$d_{13,10}$	1.29248	$d_{23,10}$	1.27555	$d_{33,10}$	1.23664
$d_{4,10}$	1.35496	$d_{14,10}$	1.15366	$d_{24,10}$	1.23992	$d_{34,10}$	1.25555
$d_{5,10}$	1.14213	$d_{15,10}$	1.37667	$d_{25,10}$	0.62501	$d_{35,10}$	1.35811
$d_{6,10}$	1.33748	$d_{16,10}$	1.13398	$d_{26,10}$	1.23957	$d_{36,10}$	1.29402
$d_{7,10}$	1.11104	$d_{17,10}$	1.19267	$d_{27,10}$	1.29875	$d_{37,10}$	1.23343
$d_{8,10}$	1.15082	$d_{18,10}$	0.93889	$d_{28,10}$	1.07421	$d_{38,10}$	1.25175
$d_{9,10}$	1.32581	$d_{19,10}$	1.04931	$d_{29,10}$	1.06643		
$d_{10,10}$	0.00000	$d_{20,10}$	1.25892	$d_{30,10}$	1.40239		

Setelah didapatkan nilai ED dari data ke-10 ke data yang lain, selanjutnya akan diambil jarak data-10 ke data yang berada dalam satu kluster (a_i) lalu akan dilakukan perhitungan menggunakan Persamaan (2.11). Karena data-10 merupakan data yang berada dalam kluster 2 dengan banyaknya jumlah anggota kluster 2 sebanyak 3 data ($m = 3$), maka dilakukan perhitungan berikut:

$$\begin{aligned}
 a_i &= \frac{1}{m-1} \sum_{r=1, r \neq i}^m d(x_i, x_r) \\
 a_{10} &= \frac{1}{3-1} (d_{18,10} + d_{25,10}) \\
 &= \frac{1}{2} (0.93889 + 0.62501) = 0.78195
 \end{aligned}$$

Setelah mendapatkan a_i , selanjutnya akan dicari jarak data-10 ke data yang berada dalam kluster lain (b_i) yakni kluster 1 ($b_{10,1}$) dan kluster 3 ($b_{10,3}$). Kemudian akan diambil nilai minimalnya menggunakan Persamaan (2.12).

$$\begin{aligned}
 b_i &= \frac{1}{m} \sum_{r=1, r \neq i}^m d(x_i, x_r) \\
 b_{10,1} &= \frac{1}{34} (d_{2,10} + d_{3,10} + \dots + d_{38,10}) \\
 &= \frac{1}{34} (1.18342 + 1.25534 + \dots + 1.25174) = 1.22885 \\
 b_i &= \frac{1}{m} \sum_{r=1, r \neq i}^m d(x_i, x_r) \\
 b_{10,3} &= \frac{1}{1} (d_{1,10}) \\
 &= 1.93789 \\
 b_{10} &= \min(b_{10,1}, b_{10,3}) \\
 &= \min(1.22885, 1.93789) = 1.22885
 \end{aligned}$$

Setelah didapatkan nilai a_i dan b_i , maka dapat dihitung nilai SC pada Kab. Tulungagung menggunakan Persamaan (2.13).

$$\begin{aligned}
 SC_i &= \frac{b_i - a_i}{\max(a_i, b_i)} \\
 SC_{10} &= \frac{b_{10} - a_{10}}{\max(a_{10}, b_{10})} \\
 &= \frac{1.22885 - 0.78195}{1.22885} = 0.36367
 \end{aligned}$$

Perhitungan seperti contoh di atas, dilakukan pada seluruh data sehingga didapatkan nilai SC sebagai berikut:

Tabel 4.32 Nilai SC pada Tiap Data (GAK-Means)

Data-i	SC_i	Data-i	SC_i	Data-i	SC_i	Data-i	SC_i
1	1.00000	11	0.22590	21	0.03613	31	0.60036
2	0.57035	12	0.24657	22	0.27750	32	0.13654
3	0.60536	13	0.62078	23	0.58245	33	0.28654
4	0.59693	14	0.51624	24	0.40356	34	0.62009
5	0.39757	15	0.23827	25	0.22276	35	0.47684
6	0.61170	16	0.56951	26	0.52690	36	0.62035
7	0.50650	17	0.54825	27	0.62488	37	0.56168
8	0.51034	18	0.00116	28	0.29540	38	0.59128
9	0.61857	19	0.28182	29	0.38673		
10	0.36367	20	0.62119	30	0.11857		

Setelah mendapatkan semua nilai SC pada masing-masing data, selanjutnya akan dihitung nilai SC pada masing-masing kluster (SCK) untuk mendapatkan nilai SC menggunakan Persamaan (2.14) dan (2.15).

$$\begin{aligned}
SCK &= \frac{1}{m} \sum_{i=1}^m SC_i \\
SC1 &= \frac{1}{34}(SC_2 + SC_3 + \dots + SC_{38}) \\
&= \frac{1}{34}(0.57035 + 0.60536 + \dots + 0.59128) = 0.45387 \\
SCK &= \frac{1}{m} \sum_{i=1}^m SC_i \\
SC2 &= \frac{1}{2}(SC_{10} + SC_{18} + SC_{25}) \\
&= \frac{1}{3}(0.36367 + 0.00116 + 0.22276) = 0.19586 \\
SCK &= \frac{1}{m} \sum_{i=1}^m SC_i \\
SC3 &= \frac{1}{1}(SC_1) \\
&= 1 \\
SC &= \frac{1}{k} \sum_{j=1}^k SCK \\
&= \frac{1}{3}(SC1 + SC2 + SC3) \\
&= \frac{1}{3}(0.45387 + 0.19586 + 1) = 0.54991
\end{aligned}$$

Perhitungan sebelumnya merupakan hasil evaluasi dari pengelompokan menggunakan GAK-Means dengan mendapatkan evaluasi $SC1 = 0.45387$, $SC2 = 0.19586$, dan $SC3 = 1$. Demikian SC rata-rata yang didapatkan yakni sebesar $SC = 0.54991$ dan struktur klasternya termasuk dalam kategori baik.

Perhitungan Evaluasi Kluster K-Means tanpa Optimasi

Berikut juga dicontohkan perhitungan mencari nilai SC pengelompokan k-Means pada data yang berada di kluster 3 yakni data-10 (Kab. Tulungagung).

Untuk mendapatkan nilai SC, akan dicari jarak data-10 ke data lain terlebih dahulu menggunakan *Euclidean Distance* dengan data-10 sebagai pusat klasternya. Tabel 4.31 merupakan perhitungan jarak Euclidean data-10 ke data-data lainnya. Selanjutnya diambil jarak data-10 ke data yang berada dalam satu klaster (a_i) lalu akan dilakukan perhitungan menggunakan Persamaan (2.11). Karena data-10 merupakan data yang berada dalam klaster 3 dengan banyaknya jumlah anggota klaster 3 sebanyak 13 data ($m = 13$), maka dilakukan perhitungan berikut:

$$\begin{aligned}
 a_i &= \frac{1}{m-1} \sum_{r=1, r \neq i}^m d(x_i, x_r) \\
 a_{10} &= \frac{1}{13-1} (d_{5,10} + d_{11,10} + \dots + d_{33,10}) \\
 &= \frac{1}{12} (1.14213 + 1.11261 + \dots + 1.23664) = 1.14658
 \end{aligned}$$

Setelah mendapatkan a_i , selanjutnya akan dicari jarak data-10 ke data yang berada dalam klaster lain (b_i) yakni klaster 1 ($b_{10,1}$) dan klaster 2 ($b_{10,2}$). Kemudian akan diambil nilai minimalnya menggunakan Persamaan (2.12).

$$\begin{aligned}
 b_i &= \frac{1}{m} \sum_{r=1, r \neq i}^m d(x_i, x_r) \\
 b_{10,1} &= \frac{1}{24} (d_{2,10} + d_{3,10} + \dots + d_{38,10}) \\
 &= \frac{1}{24} (1.18342 + 1.25534 + \dots + 1.25174) = 1.23274
 \end{aligned}$$

$$\begin{aligned}
 b_i &= \frac{1}{m} \sum_{r=1, r \neq i}^m d(x_i, x_r) \\
 b_{10,2} &= \frac{1}{1} (d_{1,10}) \\
 &= 1.93789 \\
 b_{10} &= \min(b_{10,1}, b_{10,2}) \\
 &= \min(1.23274, 1.93789) = 1.23274
 \end{aligned}$$

Setelah didapatkan nilai a_i dan b_i , maka dapat dihitung nilai SC pada Kab. Tulungagung menggunakan Persamaan (2.13).

$$\begin{aligned}
 SC_i &= \frac{b_i - a_i}{\max(a_i, b_i)} \\
 SC_{10} &= \frac{b_{10} - a_{10}}{\max(a_{10}, b_{10})} \\
 &= \frac{1.23274 - 1.14658}{1.14658} = 0.06989
 \end{aligned}$$

Perhitungan seperti contoh di atas, dilakukan pada seluruh data sehingga didapatkan nilai SC sebagai berikut:

Tabel 4.33 Nilai SC pada Tiap Data (K-Means)

Data-i	SC	Data-i	SC	Data-i	SC	Data-i	SC
1	1.00000	11	-0.16998	21	0.13718	31	0.66463
2	0.65976	12	0.02869	22	-0.18967	32	-0.00316
3	0.68501	13	0.66638	23	0.56346	33	-0.02913
4	0.67876	14	0.56855	24	0.36287	34	0.70501
5	-0.21257	15	0.07585	25	0.11896	35	0.47014

6	0.69476	16	0.66758	26	0.51570	36	0.66679
7	0.61543	17	0.59251	27	0.68252	37	0.51167
8	0.55358	18	0.08790	28	0.30290	38	0.64289
9	0.69610	19	-0.14363	29	0.33115		
10	0.06989	20	0.70917	30	-0.00642		

Setelah mendapatkan semua nilai SC pada masing-masing data, selanjutnya akan dihitung nilai SC pada masing-masing klaster (SCK) untuk mendapatkan nilai SC menggunakan Persamaan (2.14) dan (2.15).

$$SCK = \frac{1}{m} \sum_{i=1}^m SC_i$$

$$\begin{aligned} SC1 &= \frac{1}{24}(SC_2 + SC_3 + \dots + SC_{38}) \\ &= \frac{1}{24}(0.65976 + 0.68501 + \dots + 0.64289) = 0.59197 \end{aligned}$$

$$SCK = \frac{1}{m} \sum_{i=1}^m SC_i$$

$$SC2 = \frac{1}{1}(SC_1)$$

$$SCK = \frac{1}{m} \sum_{i=1}^m SC_i$$

$$\begin{aligned} SC3 &= \frac{1}{13}(SC_5 + SC_{11} + \dots + SC_{33}) \\ &= \frac{1}{13}(-0.21257 + (-0.16998) + \dots + (-0.02913)) = -0.01816 \end{aligned}$$

$$\begin{aligned}
 SC &= \frac{1}{k} \sum_{j=1}^k SC_j \\
 &= \frac{1}{3}(SC1 + SC2 + SC3) \\
 &= \frac{1}{3}(0.59197 + 1 + (-0.01816)) = 0.52460
 \end{aligned}$$

Perhitungan sebelumnya merupakan hasil evaluasi dari pengelompokan menggunakan k-Means tanpa optimasi GA dengan mendapatkan evaluasi $SC1 = 0.59197$, $SC2 = 1$, dan $SC3 = -0.01816$. Demikian SC rata-rata yang didapatkan yakni sebesar $SC = 0.52460$ dan struktur klasternya termasuk dalam kategori baik.

Perbandingan Evaluasi K-Means dan GAK-Means

Berdasarkan perhitungan manual yang telah dilakukan, jika memperhatikan Tabel 2.1 maka nilai evaluasi menggunakan GAK-Means maupun menggunakan k-Means saja sama-sama menghasilkan evaluasi yang baik ($0.5 < SC < 0.7$). Selain itu, dapat dilihat bahwa perhitungan menggunakan metode gabungan GA dan k-Means mendapatkan nilai evaluasi yang lebih tinggi dari perhitungan menggunakan metode k-Means tanpa optimasi GA. Lalu pada perhitungan evaluasi menggunakan metode GAK-Means, nilai SC di setiap klasternya tidak memiliki nilai yang negatif (*overleaping*) sehingga struktur pengelompokan dapat dikatakan baik.

Perbedaan lain juga dapat diamati melalui hasil perhitungan manual antara kedua metode tersebut. Pada perhitungan menggunakan GAK-Means, proses k-Means menjadi sangat cepat karena sudah konvergen di iterasi-2. Sedangkan perhitungan menggunakan k-Means tanpa optimasi GA baru mendapatkan pengelompokan secara konvergen di iterasi-4. Demikian dikatakan bahwa metode

GA dapat diimplementasikan dengan baik untuk mengoptimasi pusat kluster awal pada algoritma k-Means sehingga proses iterasi k-Means menjadi lebih optimal.

4.4. Uji Coba Parameter

Selanjutnya adalah tahap uji coba parameter. Data yang digunakan pada tahap uji coba parameter ini adalah data sebaran COVID-19 pada bulan November 2021. Parameter yang dilakukan uji coba yakni banyaknya individu dalam suatu populasi (*Popsi*), banyaknya generasi (*I*), dan jumlah kluster (*k*). Berikut adalah serangkaian uji coba parameter yang dilakukan:

4.4.1. Uji Coba Banyaknya Popsi

Pada tahap pengujian jumlah individu dalam populasi yang digunakan (*Popsi*), akan dilakukan uji coba sebanyak 5 kali ($Popsi = 10, 20, 40, 80, 160$). Sedangkan untuk parameter lainnya akan diinisialisasi terlebih dahulu yakni $I = 100$, $Cr = 0.3$, $Mr = 0.2$, dan $k = 3$. Tabel 4.34 merupakan hasil uji coba parameter *Popsi*.

Tabel 4.34 Uji Coba *Popsi*

<i>Popsi</i>	SC1	SC2	SC3	SC
10	-0.02150	0.59700	1.00000	0.52517
20	-0.01820	0.59200	1.00000	0.52460
40	0.59200	-0.01820	1.00000	0.52460
80	0.46380	1.00000	0.43210	0.63197
160	0.46380	1.00000	0.43210	0.63197

Pada Tabel 4.34 dapat diketahui bahwa nilai SC terbaik diperoleh saat $Popsiz = 80$ dan $Popsiz = 160$. Semakin banyak individu dalam suatu populasi semakin banyak pula solusi yang ditawarkan. Sehingga dapat memungkinkan untuk mendapatkan solusi dengan nilai SC yang lebih baik. $Popsiz$ terbaik akan dilanjutkan pada proses uji coba banyaknya generasi.

4.4.2. Uji Coba Banyaknya Generasi

Pada tahap pengujian banyaknya generasi (I), akan dilakukan uji coba sebanyak 5 kali juga ($I = 100, 200, 3000, 400, 500$). Tabel 4.35 merupakan hasil uji coba parameter I dengan $Popsiz = 80$ dan Tabel 4.36 dengan $Popsiz = 160$.

Tabel 4.35 Uji Coba Generasi ($Popsiz = 80$)

Generasi	SC1	SC2	SC3	SC
100	0.46380	1.00000	0.43210	0.63197
200	0.46380	1.00000	0.43210	0.63197
300	0.46380	1.00000	0.43210	0.63197
400	-0.01820	0.59200	1.00000	0.52460
500	1.00000	-0.09900	0.58720	0.49607

Tabel 4.36 Uji Coba Generasi ($Popsiz = 160$)

Generasi	SC1	SC2	SC3	SC
100	0.46380	1.00000	0.43210	0.63197
200	0.46380	1.00000	0.43210	0.63197
300	0.46380	1.00000	0.43210	0.63197
400	0.46380	1.00000	0.43210	0.63197
500	0.46380	1.00000	0.43210	0.63197

Tabel 4.35 memberikan informasi bahwa nilai SC terbaik diperoleh saat generasi 100, 200, dan 300. Sedangkan pada Tabel 4.36 dari semua generasi yang diuji cobakan mendapatkan hasil terbaik. Demikian generasi yang dipilih untuk diteruskan ke proses selanjutnya adalah $I = 100$ karena dari kedua uji coba didapatkan hasil yang sama dan tentunya dapat mempercepat iterasi pada GA.

4.4.3. Uji Coba Jumlah Klaster (GAK-Means)

Uji coba parameter yang terakhir adalah uji coba jumlah klaster. Pengujian ini dilakukan sebanyak 3 kali ($k = 2, 3, 4$). Parameter yang digunakan diambil dari hasil pengujian sebelumnya yakni $Popsiz$ = 160, $I = 100$, $Cr = 0.3$, dan $Mr = 0.2$. Tabel 4.37 merupakan hasil uji coba jumlah klaster (k).

Tabel 4.37 Uji Coba Jumlah Klaster (GAK-Means)

Jumlah K	SC1	SC2	SC3	SC4	SC
2	0.71750	1.00000	-	-	0.85875
3	0.46380	1.00000	0.43210	-	0.63197
4	0.45750	1.00000	0.38160	0.38090	0.55500

Tabel 4.37 memberikan kesimpulan bahwa nilai SC terbaik adalah saat $k = 2$ yakni sebesar $SC = 0.85875$. Sehingga $k = 2$ akan terpilih sebagai parameter terbaik untuk masuk ke tahap pemetaan klaster.

4.4.4. Uji Coba Jumlah Klaster (K-Means)

Selain dilakukan uji coba parameter menggunakan metode GAK-Means, pada penelitian ini juga dilakukan uji coba parameter terhadap perhitungan k-Means tanpa optimasi GA. Tabel 4.38 merupakan hasil uji coba parameter menggunakan algoritma k-Means tanpa optimasi GA.

Tabel 4.38 Uji Coba Jumlah Klaster (K-Means)

Jumlah K	SC1	SC2	SC3	SC4	SC
2	0.55539	0.05588	-	-	0.30564
3	0.59197	1.00000	-0.01816	-	0.52460
4	0.59613	1.00000	-0.01644	0.34369	0.48084

Pada Tabel 4.38 dapat dilihat uji coba jumlah klaster k-Means tanpa optimasi mendapatkan nilai SC terbaik saat $k = 3$ yakni sebesar 0.52460. Sedangkan pada uji coba GAK-Means, nilai SC terbaik didapatkan saat $k = 2$ dengan perolehan nilai SC yang lebih tinggi yakni 0.85875. Berdasarkan struktur kategori SC pada Tabel 2.1, hasil evaluasi GAK-Means masuk dalam kategori struktur klaster kuat. Demikian dapat membuktikan bahwa klastering menggunakan metode GAK-Means mampu mengelompokkan data dengan lebih baik daripada metode k-Means tanpa optimasi GA.

Selain itu dapat diperhatikan pula pada Tabel 4.38 terdapat nilai negatif pada klaster ketiga ($k = 3$ dan 4), yang mengartikan bahwa data dalam klaster ketiga mengalami *overlapping* atau data tidak dikelompokkan dengan baik. Sedangkan pada Tabel 4.37, tidak ditemukan evaluasi yang bernilai negatif sehingga pengelompokan dapat dikatakan baik. Hal ini kemungkinan disebabkan oleh sensitifnya algoritma k-Means terhadap pemilihan pusat klaster awal.

4.5. Pemetaan Klaster

Berdasarkan uji coba yang telah dilakukan sebelumnya, maka parameter terbaik akan diterapkan untuk memetakan daerah Provinsi Jawa Timur pada setiap bulan periode Januari-November 2021. Parameter yang terpilih adalah banyaknya

individu dalam populasi ($Popsize = 160$), banyaknya generasi ($I = 100$), dan jumlah klaster ($k = 2$).

Sebelum dilakukan pemetaan akan dilakukan interpretasi klaster terlebih dahulu dengan melihat nilai pusat klaster akhirnya. Tabel 4.39 adalah hasil perhitungan pusat klaster pada data November 2021 dengan $k = 2$. Nilai pusat klaster yang diperoleh akan dilakukan denormalisasi terlebih dahulu karena sebelumnya telah dilakukan normalisasi. Tabel 4.40 merupakan hasil denormalisasi.

Tabel 4.39 Pusat Klaster Akhir sebelum Denormalisasi

klaster1	0.16400	0.23310	0.08650	0.23280	0.06960	0.10460	0.10190
klaster2	1.00000	0.75000	1.00000	0.38810	0.11540	1.00000	1.00000

Tabel 4.40 Pusat Klaster Akhir setelah Denormalisasi

klaster1	27.06	1.86	12.02	15.60	1.81	2.41	14.16
klaster2	165.00	6.00	139.00	26.00	3.00	23.00	139.00

Tabel 4.41 Interpretasi Klaster

klaster1	rendah	rendah	rendah	rendah	rendah	rendah	rendah
klaster2	tinggi	tinggi	tinggi	tinggi	tinggi	tinggi	tinggi

Berdasarkan Tabel 4.41 maka dapat dikatakan bahwa klaster 1 merupakan klaster dengan kategori lebih rendah yang akan diberi warna hijau (H) sedangkan klaster 2 merupakan klaster dengan kategori tinggi yang akan diberi warna merah (M). Dengan artian lain berdasarkan interpretasi di atas, dapat dinyatakan bahwa daerah yang berada pada klaster 2 merupakan daerah yang masuk dalam kriteria lebih rawan dari pada daerah yang masuk pada klaster 1.

Setelah dilakukan interpretasi klaster, selanjutnya akan dilakukan pemetaan pada data Januari-November 2021. Tabel 4.42 adalah hasil pemetaan tiap bulan dengan menggunakan parameter terbaik pada kota di daerah Jawa Timur sedangkan pemetaan pada kabupaten di Jawa Timur terdapat pada Tabel 4.43. Pada Tabel 4.42 dan 4.43 akan dilakukan penyingkatan nama kota/kabupaten untuk memudahkan penulisan.

Tabel 4.42 Hasil Pemetaan Tiap Bulan (Kota)

Kota	Jan	Feb	Mar	Apr	Mei	Jun	Jul	Agu	Sep	Okt	Nov
Sura	M	M	M	M	M	H	M	M	M	M	M
Prob	H	H	H	H	H	H	H	H	H	H	H
Pasu	H	H	H	H	H	H	H	H	H	H	H
Mojo	H	H	H	H	H	H	H	H	H	H	H
Mala	M	H	H	H	M	H	H	H	H	H	H
Madi	H	H	H	H	M	H	H	H	H	H	H
Kedi	H	H	H	H	H	H	H	H	H	H	H
Blit	H	H	H	H	H	H	H	H	H	H	H
Batu	H	H	H	H	H	H	H	H	H	H	H

Tabel 4.43 Hasil Pemetaan Tiap Bulan (Kabupaten)

Kab	Jan	Feb	Mar	Apr	Mei	Jun	Jul	Agu	Sep	Okt	Nov
Tulu	H	H	H	H	H	H	H	H	H	H	H
Tuba	H	H	H	H	H	H	H	H	H	H	H
Tren	H	H	H	H	M	H	H	H	H	H	H
Sume	H	H	H	H	H	H	H	H	H	H	H
Situ	H	H	H	H	H	H	H	H	H	H	H

Sido	H	H	H	H	H	H	H	H	H	H	H
Samp	H	H	H	H	H	H	H	H	H	H	H
Prob	H	H	H	H	H	H	H	H	H	H	H
Pono	H	H	H	H	M	H	H	H	H	H	H
Pasu	H	H	H	H	H	H	H	H	H	H	H
Pame	H	H	H	H	H	H	H	H	H	H	H
Paci	H	H	H	H	H	H	H	H	H	H	H
Ngaw	H	H	H	H	M	H	H	H	H	H	H
Ngan	M	H	H	H	H	H	H	H	H	H	H
Mojo	H	H	H	H	H	H	H	H	H	H	H
Mala	H	H	H	M	H	H	H	H	H	H	H
Mage	H	H	H	M	H	H	H	H	H	H	H
Madi	H	H	H	M	H	H	H	H	H	H	H
Luma	H	H	H	H	H	H	H	H	H	H	H
Lamo	H	H	H	H	H	H	H	H	H	H	H
Kedi	H	H	H	H	H	H	H	H	H	H	H
Jomb	H	H	H	H	H	H	H	H	H	H	H
Jemb	M	H	H	H	H	H	H	H	H	H	H
Gres	H	H	H	H	H	H	H	H	H	H	H
Bond	H	H	H	H	H	H	H	H	H	H	H
Bojo	H	H	H	H	H	H	H	H	H	H	H
Blit	M	H	H	H	M	H	H	H	H	H	H
Bany	H	H	H	H	M	H	H	H	H	H	H
Bang	H	H	H	H	H	M	H	H	H	H	H



Gambar 4.3 Pemetaan Bulan Januari 2021

Berdasarkan Gambar 4.3 dapat diamati bahwa pengelompokan pada data bulan Januari 2021 memperlihatkan 5 daerah dengan kluster merah yang menandakan 5 daerah tersebut memiliki tingkat kerawanan yang lebih tinggi dibandingkan daerah lainnya yang berada di kluster hijau.



Gambar 4.4 Pemetaan Bulan Februari 2021



Gambar 4.5 Pemetaan Bulan Maret 2021

Pada penambahan kasus di bulan Februari dan Maret 2021, penyebaran COVID-19 terlihat lebih terkendali padahal kasus penambahan pasien positif masih terus bertambah di setiap harinya. Hal ini dikarenakan penyebaran Kota Surabaya memiliki jumlah kasus yang cukup besar dibandingkan daerah lain.



Gambar 4.6 Pemetaan Bulan April 2021

Berbeda dengan bulan Februari dan Maret sebelumnya, pengelompokan pada bulan April 2021 yang dapat diamati melalui Gambar 4.6 menyatakan bahwa selain Kota Surabaya, terdapat pula daerah yang masuk dalam kluster merah yakni Kabupaten Madiun.



Gambar 4.7 Pemetaan Bulan Mei 2021

Masuk pada bulan Mei 2021 daerah yang berada dalam kluster merah bertambah cukup banyak menjadi 11 daerah. Terlebih lagi daerah di sekitar Kabupaten Madiun. 11 daerah tersebut adalah Kota Surabaya, Kota Malang, Kota Madiun, Kabupaten Banyuwangi, Kabupaten Ponorogo, Kabupaten Madiun, Kabupaten Blitar, Kabupaten Ngawi, Kabupaten Trenggalek, Kabupaten Magetan, dan Kabupaten Malang.



Gambar 4.8 Pemetaan Bulan Juni 2021

Pada bulan Juni 2021, terdapat 1 daerah yang masuk pada klaster merah yakni Kabupaten Bangkalan. Hal tersebut menandakan bahwa terdapat kenaikan drastis pada penambahan kasus di Kabupaten Bangkalan. Kenaikan drastis tersebut membuat daerah Kabupaten Bangkalan menjadi satu-satunya daerah yang masuk pada klaster merah.

UIN SUNAN AMPEL
SURABAYA



Gambar 4.13 Pemetaan Bulan November 2021

Dalam lima bulan berturut-turut dimulai Juli-November 2021, klaster merah hanya beranggotakan satu daerah yakni Kota Surabaya. Hal tersebut menandakan bahwa Kota Surabaya dalam lima bulan terakhir menjadi daerah dengan tingkat kerawanan yang tinggi jika dibandingkan dengan daerah lain.

4.6. Analisis Hasil

Berdasarkan pemetaan tiap bulan yang telah dilakukan dalam subbab sebelumnya, dapat dilihat penyebaran COVID-19 pada tiap bulan di Provinsi Jawa Timur berhasil dikelompokkan dengan baik. Hal tersebut dapat diamati melalui data sebaran serta pemetaan yang telah dilakukan. Pada bulan Januari 2021, penambahan kasus berada di rentang ratusan sampai ribuan kasus. Lima daerah masuk dalam klaster merah yang menandakan lima daerah tersebut menjadi daerah dengan penyebaran yang cukup tinggi dibandingkan daerah lainnya.

Sedangkan pada bulan Februari dan Maret 2021, Kota Surabaya menjadi satu-satunya daerah dengan penyebaran yang tinggi. Pada April 2021, Kabupaten

Madiun menyusul Kota Surabaya yang berada di klaster merah. Pada data sebaran, keduanya sama-sama menjadi penyumbang kasus tertinggi dengan selisih yang sedikit. Pada Mei 2021, penyebaran kasus Kabupaten Madiun menyalip Kota Surabaya dan menjadi daerah dengan penyebaran tertinggi. Selain itu jika diamati, daerah yang berada di sekitar Kabupaten Madiun juga masuk pada klaster merah.

Pada bulan Juni 2021, daerah yang berada dalam klaster merah adalah Kabupaten Bangkalan. Penambahan kasus secara drastis terjadi di Kabupaten Bangkalan. Hal ini disebabkan kenaikan jumlah pemudik di pertengahan bulan Mei dimana Kabupaten Bangkalan menjadi salah satu daerah yang dituju para pemudik untuk merayakan hari raya idul fitri. Dilansir melalui artikel pada laman *webiste* Dewan Perwakilan Daerah (DPD), Penambahan kasus secara drastis tersebut menjadi salah satu penyebab dilakukannya penyekatan serta *testing* secara massal pada seluruh pengendara yang melintasi jembatan Suramadu (DPD, 2021).

Bersamaan dengan itu masih di bulan Juni 2021, di Provinsi Jawa Timur telah terdeteksi virus COVID-19 dengan mutasi varian Delta. Dimana virus COVID-19 varian Delta memiliki penyebaran yang sangat cepat. Varian Delta pada awalnya ditemukan menginfeksi tiga warga Jawa Timur. Ketiganya ditemukan reaktif saat dilakukan penyekatan suramadu usai menghadiri acara di Kabupaten Bangkalan (Farid, 2021). Satu orang berasal dari Kabupaten Bojonegoro sedangkan dua orang yang lainnya berasal dari Kabupaten Bangkalan (Khusaini, 2021).

Pada bulan Juli 2021 tepatnya saat kasus COVID-19 di Indonesia berada pada puncaknya, penambahan kasus melonjak drastis di berbagai daerah. Namun, Kota Surabaya menjadi sorotan karena penambahan kasusnya yang jauh lebih tinggi dibandingkan daerah lain hingga mencapai angka 27.385 jiwa. Itulah kenapa pada

saat itu pemerintah menyatakan bahwa Kota Surabaya menjadi daerah dengan zona merah tua (Ghinan Salman, 2021). Pada bulan selanjutnya sampai November 2021 penyebaran kasus bulanan mengalami penurunan. Namun Kota Surabaya masih menjadi daerah dengan penyebaran yang cukup tinggi jika dibandingkan dengan daerah lain di Provinsi Jawa Timur.

Jika diperhatikan, Kota Surabaya menjadi satu-satunya daerah dengan tingkat kerawanan yang tinggi dalam periode Januari-November 2021 selain pada bulan Juni 2021. Hal ini dapat terjadi karena Kota Surabaya merupakan ibukota Provinsi Jawa Timur dengan populasi yang padat serta aktivitas masyarakat yang cukup tinggi (Khikmah and Sofro, 2021). Selain itu kota Surabaya juga menjadi satu-satunya kota yang menghubungkan pulau jawa dengan pulau madura sehingga semakin besar resiko penularan yang disebarkan oleh masyarakat dari daerah lain.

4.7. Integrasi Keislaman

Berdasarkan analisis pemetaan di subbab sebelumnya, pada tiga bulan terakhir penyebaran COVID-19 sudah mengalami penurunan laju penyebaran. Namun masyarakat tetap tidak boleh lengah terhadap hal tersebut. Hal ini dikarenakan belum ada kepastian terkait kapan pandemi COVID-19 akan mereda. Tentunya kesehatan adalah salah satu hal yang paling penting saat ini. Dengan menjaga imunitas tetap kuat serta mematuhi protokol kesehatan masih harus terus diperhatikan. Menjaga kesehatan adalah kewajiban bagi setiap muslim yang mana hal tersebut telah disebutkan dalam firman Allah Swt dengan redaksi sebagai berikut:

وَأَنْفَقُوا فِي سَبِيلِ اللَّهِ وَلَا تُلْقُوا بِأَيْدِيكُمْ إِلَى التَّهْلُكَةِ وَأَحْسِنُوا إِنَّ اللَّهَ يُحِبُّ الْمُحْسِنِينَ ﴿١٤٥﴾

Artinya: “Dan infakkanlah (hartamu) di jalan Allah Swt, dan janganlah

kamu jatuhkan (diri sendiri) ke dalam kebinasaan dengan tangan sendiri, dan berbuat baiklah. Sungguh, Allah menyukai orang-orang yang berbuat baik.” (QS. Al-Baqarah: 195).

Selain memerintahkan kepada manusia untuk menginfakkan sebagian hartanya di jalan Allah Swt dan berbuat baik kepada sesama, dalam al-Qur'an surah al-Baqarah ayat 195 diperintahkan pula untuk tidak menjatuhkan diri ke dalam kebinasaan. Artinya, kita dituntut untuk tidak melakukan sesuatu yang dapat menyakiti atau menjatuhkan diri sendiri ke dalam hal-hal yang tidak baik. Salah satu contohnya adalah dengan menjaga kesehatan agar tidak sakit akibat kelalaian kita sendiri. Di masa pandemi kita diwajibkan untuk mematuhi protokol kesehatan sehingga kesehatan kita dan orang lain juga terjaga misalkan dengan menerapkan protokol *physical distancing*.

Selain itu, protokol mengurangi mobilitas bepergian juga penting untuk mengurangi resiko penularan yang tinggi. Hal ini dapat dilakukan dengan memperhatikan daerah-daerah dengan kerawanan yang cukup tinggi melalui proses klastering daerah. Pada penelitian ini telah dilakukan pemetaan pada tiap bulan dengan membagi daerah-daerah menjadi dua klaster. Klaster merah mengindikasikan bahwa daerah tersebut memiliki laju penyebaran yang cukup tinggi dibandingkan dengan daerah-daerah yang berada dalam klaster hijau. Masyarakat yang berada dalam lingkup klaster hijau diharapkan tetap menahan diri untuk melakukan aktivitas bepergian ke daerah-daerah dengan klaster merah. Hal tersebut sejatinya sudah pernah diajarkan dalam Islam melalui sabda Rasulullah Saw.

عَنْ عَبْدِ اللَّهِ بْنِ عَامِرٍ بْنِ رَبِيعَةَ أَنَّ عُمَرَ خَرَجَ إِلَى الشَّامِ فَلَمَّا جَاءَ سَرَعَ بَلَّغَهُ أَنَّ الْوَبَاءَ قَدْ وَقَعَ بِالشَّامِ فَأَخْبَرَهُ عَبْدُ الرَّحْمَنِ بْنُ عَوْفٍ أَنَّ رَسُولَ اللَّهِ صَلَّى اللَّهُ عَلَيْهِ وَسَلَّمَ قَالَ إِذَا سَمِعْتُمْ بِهِ بِأَرْضٍ فَلَا تَقْدَمُوا عَلَيْهِ وَإِذَا وَقَعَ بِأَرْضٍ وَأَنْتُمْ بِهَا فَلَا تَخْرُجُوا فِرَارًا مِنْهُ فَرَجَعَ عُمَرُ بْنُ الْخَطَّابِ مِنْ سَرَعٍ (رواه البخاري ومسلم)

Artinya: “Dari Abdullah bin Amir bin Rabi’ah, Sesungguhnya Umar bin Khattab ra menempuh perjalanan menuju Syam. Ketika sampai di Sargh, Umar mendapat kabar bahwa wabah sedang menimpa daerah Syam. Abdurrahman bin Auf mengatakan kepada Umar bahwa Rasulullah Saw pernah bersabda, ‘Bila kamu mendengar wabah di suatu daerah, maka kalian jangan memasukinya. Tetapi jika wabah terjadi wabah di daerah kamu berada, maka jangan tinggalkan tempat itu.’ Lalu Umar bin Khattab berbalik arah meninggalkan Sargh.” (HR. Bukhari dan Muslim).

Pada hadis di atas, Rasulullah Saw secara gamblang memerintahkan untuk menjauhi daerah yang terkena wabah penyakit dikarenakan dapat beresiko untuk ikut tertular penyakit juga (Rosanti, 2021). Lalu Rasulullah Saw juga memerintahkan kepada orang-orang yang telah berada di daerah wabah untuk tetap di daerah tersebut agar tidak menularkan pada orang lain yang belum tertular. Hal ini relevan dengan anjuran untuk tidak mengunjungi atau bepergian ke daerah yang memiliki tingkat kerawanan tinggi (Mukharom and Aravik, 2020).

Aturan atau kebijakan yang disebutkan sebelumnya telah ditetapkan oleh pemerintah sebagai bentuk cepat tanggap menghadapi wabah COVID-19. Dalam islam, saat pemerintah menetapkan kebijakan terdapat syarat penting yang harus dipenuhi. Hal ini dapat ditemukan dalam kaidah fikih mengenai aturan dalam membuat kebijakan oleh kepala Negara sebagai berikut:

تصرف الإمام على الراعية منوط بالمصلحة

Artinya: “Tindakan kepala Negara/imam terhadap rakyat itu harus didasarkan atas pertimbangan kemaslahatan (masyarakat).”

Berdasarkan kaidah di atas, kebijakan yang diatur serta diterapkan oleh pemerintah haruslah berdasarkan kemaslahatan masyarakat dengan berbagai pertimbangan. Sebagai warga Negara yang baik, upaya-upaya tersebut harus dipatuhi karena dibuat demi kemaslahatan masyarakat serta tidak menyimpang dari apa yang telah diajarkan oleh islam (Mahargiani et al., 2021).

Dalam menentukan kebijakan suatu daerah tentunya diperlukan strategi yang tepat sesuai dengan kebutuhan masyarakat. Penelitian ini telah melakukan klastering pada tiap bulan dengan harapan dapat menjadi salah satu refrensi yang dapat digunakan pemerintah dalam mengambil kebijakan lanjutan. Daerah-daerah yang masuk ke dalam kategori dengan kerawanan yang tinggi akan lebih diperhatikan untuk kemudian dilakukan strategi yang lebih optimal.

UIN SUNAN AMPEL
S U R A B A Y A

BAB V

PENUTUP

5.1. Kesimpulan

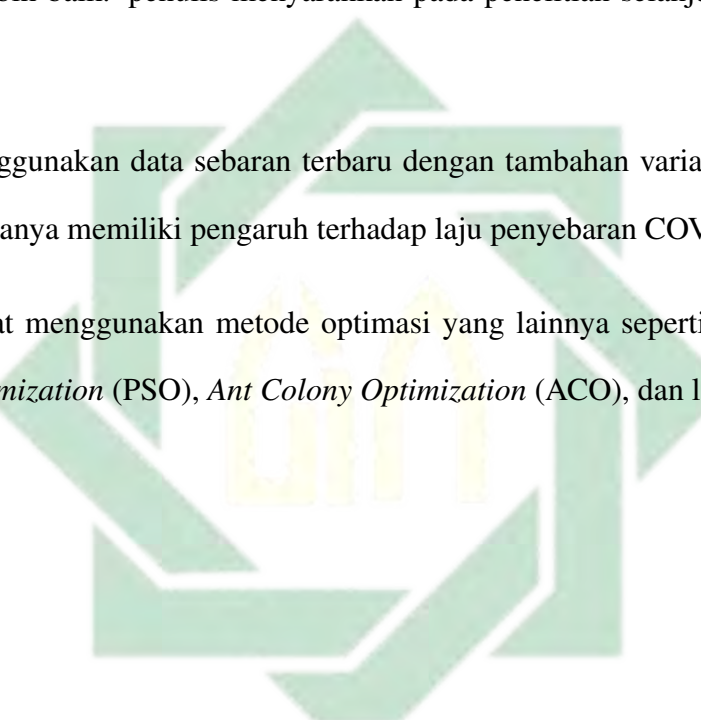
Berdasarkan hasil yang didapatkan dari pengelompokan daerah penyebaran COVID-19 di Provinsi Jawa Timur menggunakan gabungan metode GA dan k-Means, dapat disimpulkan bahwa:

1. Evaluasi yang dihasilkan pada perhitungan implementasi GAK-Means mendapatkan hasil yang lebih baik dibandingkan dengan perhitungan menggunakan k-Means tanpa optimasi. Berdasarkan uji coba parameter, hasil evaluasi tertinggi yang diperoleh yakni sebesar SC GAK-Means = 0.85875 dengan $k = 2$ dan SC k-Means = 0.52460 dengan $k = 3$. Selain mendapatkan hasil evaluasi yang lebih baik, metode GA berhasil melakukan optimasi dalam menentukan pusat kluster awal pada k-Means dibuktikan dengan proses iterasi yang lebih cepat.
2. Hasil analisis kluster pada pemetaan daerah yang telah dilakukan adalah Kota Surabaya menjadi satu-satunya kota yang masuk dalam kluster dengan kerawanan yang lebih tinggi (kluster merah) pada bulan Januari-November 2021 selain bulan Juni 2021. Sedangkan pada Juni 2021, satu-satunya wilayah yang masuk pada kluster merah adalah Kabupaten Bangkalan dan 37 kab/kota yang lain masuk dalam kluster hijau.

5.2. Saran

Penelitian mengenai implementasi GAK-Means clustering untuk pengelompokan COVID-19 di Provinsi Jawa Timur memiliki banyak kekurangan. Perbaikan dan inovasi sangat diperlukan untuk menjadikan sistem pengelompokan menjadi lebih baik. penulis menyarankan pada penelitian selanjutnya diharapkan untuk:

1. Menggunakan data sebaran terbaru dengan tambahan variabel lainnya yang sekiranya memiliki pengaruh terhadap laju penyebaran COVID-19.
2. Dapat menggunakan metode optimasi yang lainnya seperti *Particle Swarm Optimization* (PSO), *Ant Colony Optimization* (ACO), dan lain sebagainya.



UIN SUNAN AMPEL
S U R A B A Y A

DAFTAR PUSTAKA

- Abidin, Z., Yanti, A., Zulfayanto, I., and Rasid, M. (2021). Peningkatan Kualitas Kesehatan Masyarakat Untuk Pencegahan Covid-19. *Jurnal Pembelajaran Pemberdayaan Masyarakat (JP2M)*, 2(1):41–46.
- Ali, I. and Sularto, L. (2019). Optimasi Parameter Artificial Neural Network Menggunakan Algoritma Genetika Untuk Prediksi Kelulusan Mahasiswa. *Jurnal ICT : Information Communication dan Technology*, 18(1):54–59.
- Annur, H. (2019). Penerapan Data Mining Menentukan Strategi Penjualan Variasi Mobil Menggunakan Metode K-Means Clustering. *Jurnal Informatika Upgris*, 5(1).
- Arswendi (2021). Analisis Penerapan 5M dan Vaksinasi Dalam Meningkatkan Kesadaran Diri Masyarakat pada saat Pandemi Covid 19 di Desa Limbung. *Jurnal Abdimas Bina Bangsa*, 02(01):63–67.
- Aryanto, U. R., Titisari, K. H., and Nurlaela, S. (2018). Pengaruh Likuiditas, Leverage, Profitabilitas, Dan Aktivitas Terhadap Pertumbuhan Laba (Studi Empiris : Perusahaan Food And Beverages Yang Terdaftar Di Bursa Efek Indonesia Periode 2011-2015) Ulfinabella. *Seminar Nasional IENACO*, pages 625–631.
- Bastian, A., Sujadi, H., and Febrianto, G. (2018). Penerapan Algoritma K-Means Clustering Analysis Pada Penyakit Menular Manusia (Studi Kasus Kabupaten Majalengka). *Jurnal Sistem Informasi (Journal of Information System)*, 14(1):26–32.

BPS (2021). *Provinsi Jawa Timur dalam Angka 2021*. BPS Provinsi Jawa Timur, Surabaya.

Djalante, R., Lassa, J., Setiamarga, D., Sudjatma, A., Indrawan, M., Haryanto, B., Mahfud, C., Sinapoy, M. S., Djalante, S., Rafliana, I., Gunawan, L. A., Surtiari, G. A. K., and Warsilah, H. (2020). Review and analysis of current responses to COVID-19 in Indonesia: Period of January to March 2020. *Progress in Disaster Science*, 6.

DPD (2021). Polemik Penyekatan Suramadu, LaNyalla Minta Pemda Pakai Pendekatan Persuasi.

Farid (2021). Pasien Covid Delta di Jatim Berawal dari Penyekatan Suramadu.

Fauziyah Dewi (2019). K-Means Clustering with Scikit-Learn in Python.

Ghinan Salman (2021). Pemprov Jatim: Surabaya Masuk Zona Merah Tua, Bukan Hitam.

Ginting, N., Pradesyah, R., Amini, A., and Panggabean, H. S. (2021). Memperkuat Nalar Teologi Islam Moderat Dalam Menyikapi Pandemi Covid-19 Di Pimpinan Ranting Pemuda Muhammadiyah Bandar Pulau Pekan. *Martabe : Jurnal Pengabdian Kepada Masyarakat*, 4(1):30.

Handayani, N. S., Kusumadewi, S., and Fitriyanto, E. (2020). Rekomendasi Makanan untuk Ibu Hamil Menggunakan Algoritma Genetika. *JUITA: Jurnal Informatika*, 8(1):45.

Harini, S. (2020). Identification COVID-19 Cases in Indonesia with The Double Exponential Smoothing Method. *Jurnal Matematika "MANTIK"*, 6(1):66–75.

- Hussain, A., Muhammad, Y. S., and Nawaz, A. (2018). Optimization through Genetic Algorithm with a New and Efficient Crossover Operator. *International Journal of Advances in Mathematics*, 2018(1):1–14.
- Irfan, S., Dwivedi, G., and Ghosh, S. (2018). Optimization of K-means clustering using genetic algorithm. *2017 International Conference on Computing and Communication Technologies for Smart Nation, IC3TSN 2017*, 2017-Octob:156–161.
- Irham, R. A. W., Siregar, A. K., and Jubba, H. (2020). Dampak Wabah Corona Virus Disease Terhadap Budaya Ibadah Umat Muslim Dan Kebijakan Pemerintah. *POROS ONIM: Jurnal Sosial Keagamaan*, 1(2):131–143.
- Jatim Pemprov (2021). Jawa Timur Tanggap Covid-19.
- KawalCovid (2021). Informasi Terkini Covid di Indonesia.
- Khikmah, K. N. and Sofro, A. (2021). Clustering Regency in East Java Based on Population Density and Cumulative Confirmed COVID-19 Cases. *ComTech: Computer, Mathematics and Engineering Applications*, 12(2):111–121.
- Khusaini, A. (2021). Covid-19 Varian Delta di Bangkalan.
- Kuntjoro, D. A., Setiawan, B. D., and Perdana, R. S. (2018). Algoritme Genetika Untuk Optimasi K-Means Clustering Dalam Pengelompokan Data Tsunami. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 2(10):3865–3872.
- Mahargiani, E., Afnan, A. N., and Sumarjoko, S. (2021). Covid-19 dalam Perspektif Teologis, Fiqh dan Sains. *Syariat: Jurnal Studi Al-Qur'an dan Hukum*, 7(1):43–56.

- Mahmud, M. (2020). Pola Penyikapan Terhadap Penyakit Menular dan Wabah Berdasarkan Perspektif Fiqh dan Islam. *Jurnal Al-Maqasid: Jurnal Ilmu-Ilmu Kesyariahan dan Keperdataan*, 6(1):141–151.
- Mahmudan, A. (2020). Clustering of District or City in Central Java Based COVID-19 Case Using K-Means Clustering. *Jurnal Matematika, Statistika dan Komputasi*, 17(1):1–13.
- Mamat, A. R., Mohamed, F. S., Mohamed, M. A., Rawi, N. M., and Awang, M. I. (2018). Silhouette index for determining optimal k-means clustering on images in different color models. *International Journal of Engineering and Technology(UAE)*, 7(April):105–109.
- Marimin (2019). *Sistem Pendukung Pengambilan Keputusan dan Sistem Pakar*. PT. Penerbit IPB Press, Bandung.
- Mu'afa, S. F. and Ulinnuha, N. (2019). Perbandingan Metode Single Linkage, Complete Linkage Dan Average Linkage dalam Pengelompokan Kecamatan Berdasarkan Variabel Jenis Ternak Kabupaten. *Jurnal Ilmiah Bidang Teknologi Informasi dan Komunikasi*, 4(2).
- Mukharom, M. and Aravik, H. (2020). Kebijakan Nabi Muhammad Saw Menangani Wabah Penyakit Menular dan Implementasinya dalam Konteks Penanggulangan Coronavirus Covid-19. *SALAM: Jurnal Sosial dan Budaya Syar-i*, 7(3).
- Mursalim, M., Purwanto, P., and Soeleman, M. A. (2021). Penentuan Centroid Awal Pada Algoritma K-Means Dengan Dynamic Artificial Chromosomes Genetic Algorithm Untuk Tuberculosis Dataset. *Techno.Com*, 20(1):97–108.

- Nabila, Z., Rahman Isnain, A., and 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.
- Narin, A., Kaya, C., and Pamuk, Z. (2021). Automatic detection of coronavirus disease (COVID-19) using X-ray images and deep convolutional neural networks. *Pattern Analysis and Applications*, 24(3):1207–1220.
- Nasution, D. L. (2019). Perbaikan Kualitas Citra Maps Menggunakan Metode Contrast Limited Adaptive Histogram Equalization (Clahe). *KOMIK (Konferensi Nasional Teknologi Informasi dan Komputer)*, 3(1):49–56.
- Nurhidayati, T. and Rofiq, M. (2021). Pandemi COVID-19 dalam Kacamata Islam. *Falasifa*, 12(September):6.
- Prasetio, R. T., Rismayadi, A. A., and Anshori, I. F. (2018). Implementasi Algoritma Genetika pada k-nearest neighbours untuk Klasifikasi Kerusakan Tulang Belakang. *Jurnal Informatika*, 5(2):186–194.
- Reynaldo, J., Adikara, P. P., and Wihandika, R. C. (2020). Analisis Sentimen Mengenai Produk Toyota Avanza Menggunakan Metode Learning Vector Quantization Versi 3 (LVQ 3) dengan Seleksi Fitur Chi Square, Lexicon-Based Features serta Normalisasi Min-Max. 4(3):830–839.
- Ridho, M. R. (2020). Wabah Penyakit Menular dalam Sejarah Islam dan Relevansinya dengan Covid-19. *JUSPI (Jurnal Sejarah Peradaban Islam)*, 4(1):24–33.
- Roiha, N. U. (2017). *Segmentasi Pengguna Web Menggunakan Metode Genetic K-Means Algorithm*. PhD thesis, Institut Teknologi Sepuluh Nopember.

- Rosanti, C. (2021). Tinjauan Hukum Islam pada Edaran Pemerintah dan MUI dalam Menyikapi Wabah Covid-19 Setelah Pemberlakuan New Normal. *Jurnal Ilmiah Ekonomi Islam*, 7(1):393–402.
- Salman, F. M., Abu-Naser, S. S., Alajrami, E., Abu-Nasser, B. S., and Ashqar, B. A. M. (2020). COVID-19 Detection using Artificial Intelligence. *International Journal of Academic Engineering Research*, 4(3):18–25.
- Sampara (2020). Penerapan Algoritma K-Means Clustering Dalam Perancangan Planogram Pada Sang Surya Mart Kota Makassar. *Paulus Informatics Journal*, 1(2):19–25.
- Santoso, B., Cholissodin, I., and Setiawan, B. D. (2017). Optimasi K-Means untuk Clustering Kinerja Akademik Dosen Menggunakan Algoritme Genetika. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 1(12):1652–1659.
- Sari, B. N. (2016). Identification of Tuberculosis Patient Characteristics Using K-Means Clustering. *Scientific Journal of Informatics*, 3(2):129–138.
- Saxena, A., Prasad, M., Gupta, A., Bharill, N., Patel, O. P., Tiwari, A., Er, M. J., Ding, W., and Lin, C. T. (2017). A review of clustering techniques and developments. *Neurocomputing*, 267:664–681.
- Setiawan, D., Putri, R., Dan, R. S. R. J. T., and 2019, U. (2019). Implementasi algoritma genetika untuk prediksi penyakit autoimun. *jurnal.univrab.ac.id*, 4(1):8–19.
- Shi, H. and Xu, M. (2018). A Data Classification Method Using Genetic Algorithm and K-Means Algorithm with Optimizing Initial Cluster Center. 2018

IEEE International Conference on Computer and Communication Engineering Technology, CCET 2018, pages 224–228.

Siagian, R. (2021). E-Commerce Customer Segmentation Using K-Means Algorithm and Length, Recency, Frequency, Monetary Model. *JITE (Journal of Informatics and Telecommunication Engineering)*, 5(July):152–159.

Solichin, A. and Khairunnisa, K. (2020). Klasterisasi Persebaran Virus Corona (Covid-19) Di DKI Jakarta Menggunakan Metode K-Means. *Fountain of Informatics Journal*, 5(2):52.

Suryaningrum, D. A., Ratnawati, D. E., and Setiawan, B. D. (2017). Prediksi Waktu Panen Tebu Menggunakan Gabungan Metode Backpropagation dan Algoritma Genetika. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer (J-PTIIK) Universitas Brawijaya*, 1(11):1443–1450.

Suwirmayanti, N. L. G. P. (2018). Optimasi Pusat Cluster Fuzzy C-Means Untuk Pengelompokan Data. *Prosiding SINTAK*, pages 142–149.

Taufiq, M., Dewi, C., and Mahmudy, W. (2017). Optimasi Komposisi Pakan Untuk Penggemukkan Sapi Potong Menggunakan Algoritma Genetika. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 1(7):571–582.

Ulya, H. N. (2020). Alternatif Strategi Penanganan Dampak Ekonomi Covid-19 Pemerintah Daerah Jawa Timur Pada Kawasan Agropolitan. *El-Barka: Journal of Islamic Economics and Business*, 3(1):80–109.

Wang, D., Hu, B., Hu, C., Zhu, F., Liu, X., Zhang, J., Wang, B., Xiang, H., Cheng, Z., Xiong, Y., Zhao, Y., Li, Y., Wang, X., and Peng, Z. (2020). Clinical

Characteristics of 138 Hospitalized Patients with 2019 Novel Coronavirus-Infected Pneumonia in Wuhan, China. *JAMA - Journal of the American Medical Association*, 323(11):1061–1069.

WHO (2021). Coronavirus disease (COVID-19).

Yustanti, W., Rahmawati, N., and Yamasari, Y. (2020). Klastering Wilayah Kota/Kabupaten Berdasarkan Data Persebaran Covid-19 Di Propinsi Jawa Timur dengan Metode K-Means. *JIEET (Journal of Information Engineering and Educational Technology)*, 4(1):1–9.

Zeebaree, D. Q., Haron, H., Abdulazeez, A. M., and Zeebaree, S. R. (2017). Combination of k-means clustering with genetic algorithm: A review. *International Journal of Applied Engineering Research*, 12(24):14238–14245.

UIN SUNAN AMPEL
S U R A B A Y A