

Pengelompokan Daerah Rawan Gempa Bumi di Indonesia menggunakan Algoritma *Clustering Large Applications based on Randomized Search* (CLARANS)

Vina Afiah Hasna^{*}, Fitri Amanah

Prodi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Islam Bandung, Indonesia.

vinaafiyahh24@gmail.com, fitriamanah@unisba.ac.id

Abstract. Indonesia is a country prone to earthquakes because it is located at the intersection of three active tectonic plates. To improve disaster mitigation, efforts are needed to identify areas with potential for earthquakes. This research aims to classify earthquake-prone areas in Indonesia using the CLARANS (Clustering Large Applications based on Randomized Search) method. This method is an improvement over the previous medoid-based methods, namely PAM and CLARA, with the advantage of being more efficient and optimal for large-scale data. The data used is secondary data on earthquake events in Indonesia from 2005 to 2024, with the variables of depth and magnitude. Based on the research results, the silhouette coefficient value obtained was 0,6023 with the formation of 2 clusters. Cluster 1 is categorized as an area with moderate vulnerability, encompassing a total of 138 earthquake events, while cluster 2 falls into the high vulnerability category with a total of 698 earthquake events. The provinces of Papua and Sulawesi Tengah had the highest total number of earthquakes in cluster 2, with 228 and 71 events respectively. Meanwhile, provinces like Sumatra Utara and Maluku Utara had a relatively balanced total number of earthquakes in both clusters, indicating that both clusters were equally dominant in those areas.

Keywords: *Clustering, CLARANS, Earthquake.*

Abstrak. Indonesia merupakan negara yang rawan terhadap bencana gempa bumi karena terletak di pertemuan tiga lempeng tektonik aktif. Untuk meningkatkan mitigasi bencana, diperlukan upaya identifikasi daerah yang memiliki potensi rawan gempa bumi. Penelitian ini bertujuan untuk mengelompokkan daerah rawan gempa bumi di Indonesia menggunakan metode CLARANS (*Clustering Large Applications based on Randomized Search*). Metode ini merupakan pengembangan dari metode berbasis medoid sebelumnya, yaitu PAM dan CLARA, dengan keunggulan lebih efisien dan lebih optimal untuk data berskala besar. Data yang digunakan merupakan data sekunder kejadian gempa bumi di Indonesia tahun 2005–2024, dengan variabel kedalaman (*depth*) dan magnitudo (*magnitude*). Berdasarkan hasil penelitian diperoleh nilai *silhouette coefficient* sebesar 0,6023 dengan pembentukan 2 *cluster*. *Cluster* 1 dikategorikan sebagai daerah dengan tingkat kerawanan sedang, mencakup total 138 kejadian gempa bumi, sementara *cluster* 2 termasuk dalam kategori dengan tingkat kerawanan tinggi dengan total 698 kejadian gempa bumi. Provinsi Papua, dan Sulawesi Tengah menempati total kejadian gempa bumi tertinggi pada *cluster* 2, masing-masing dengan 228 dan 71 kejadian. Sementara itu, provinsi seperti Sumatera Utara dan Maluku Utara memiliki total kejadian gempa bumi yang relative seimbang di kedua *cluster*, sehingga mengindikasikan bahwa kedua *cluster* tersebut sama-sama dominan di daerah tersebut.

Kata Kunci: *Clustering, CLARANS, Gempa Bumi.*

A. Pendahuluan

Negara Indonesia terletak di wilayah cincin api pasifik, yang dikenal sebagai kawasan dengan aktivitas dan seismik vulkanik yang tinggi. Posisi geografis Indonesia yang berada di pertemuan tiga lempeng tektonik, yaitu lempeng Indo-Australia, lempeng Eurasia, dan lempeng Pasifik menyebabkan Indonesia memiliki aktivitas seismik yang tinggi sehingga rentan terhadap bencana alam seperti gempa bumi. Dampak dari gempa tidak hanya mencakup kerusakan fisik pada infrastruktur, tetapi juga korban jiwa serta gangguan terhadap stabilitas sosial dan ekonomi (Susanta et al., 2019). Pada 21 November 2022, terjadi gempa dengan magnitudo 5,6 dan kedalaman 11 km di Kabupaten Cianjur, Jawa Barat. Kejadian gempa bumi ini mengakibatkan 268 orang meninggal dunia, 1083 orang luka-luka, sejumlah 6.570 bangunan dalam kondisi rusak berat atau hancur dan 2.071 bangunan mengalami rusak sedang. Oleh karena itu, pemahaman terhadap pola kejadian gempa bumi menjadi sangat penting untuk meningkatkan upaya mitigasi dan kesiapsiagaan dalam menghadapi bencana gempa bumi ini.

Salah satu langkah dalam meningkatkan kesiapsiagaan terhadap terjadinya gempa bumi adalah dengan mengidentifikasi daerah rawan gempa bumi, yaitu daerah yang berpotensi mengalami kejadian gempa bumi serta berisiko menimbulkan dampak signifikan. Hal ini dapat dilakukan dengan mengelompokkan daerah rawan gempa bumi berdasarkan data kejadian gempa bumi. Dalam proses pengelompokan daerah rawan gempa bumi, digunakan data gempa bumi dengan magnitudo $> 4,99$ dan depth < 300 km. Gempa dengan magnitudo $> 4,99$ umumnya memiliki potensi merusak area kecil hingga kerusakan area luas (Irawan & Hermawan Hasibuan, n.d.). Sedangkan gempa dengan depth (kedalaman) < 300 km (gempa dangkal dan menengah) berpotensi menimbulkan kerusakan yang signifikan seperti kerusakan infrastruktur (Kharis Pratama et al., 2023).

Dalam pengelompokan daerah rawan gempa bumi digunakan metode clustering berbasis medoid, dimana medoid merupakan objek atau data asli yang mewakili setiap cluster. Setiap cluster memiliki satu medoid yang paling mendekati objek-objek lain dalam cluster tersebut (Sindi et al., 2020). Metode clustering berbasis medoid dipilih karena lebih tahan terhadap outlier yang sering ditemukan dalam data gempa bumi. Selain itu, karena medoid berasal dari data asli hasil pengelompokan menjadi lebih mudah diinterpretasikan. Salah satu metode clustering berbasis medoid adalah PAM (Partitioning Around Medoids) yaitu metode yang memilih beberapa objek sebagai pusat cluster (medoid), kemudian mencoba semua kombinasi untuk menemukan medoid yang menghasilkan cluster dengan jarak antar objek yang paling kecil. Metode PAM umumnya efektif untuk data kecil, namun kurang efektif dalam menangani data besar (Data Mining: Concepts and Techniques Second Edition, n.d.).

Untuk mengatasi keterbatasan algoritma PAM, (Finding Groups in Data, n.d.) mengembangkan algoritma baru yaitu CLARA (Clustering LARge Application), yang dirancang untuk menangani dataset berukuran besar. Algoritma ini bekerja dengan mengambil sampel dari data. Kemudian menerapkan algoritma PAM pada sampel tersebut untuk menentukan k medoid. Namun, kelemahan utama dari algoritma ini adalah bahwa k medoid yang diperoleh dari sampel tidak selalu mewakili keseluruhan dataset, sehingga kualitas cluster yang dihasilkan sering kali lebih rendah dibandingkan dengan yang diperoleh menggunakan PAM.

Ng & Han mengembangkan algoritma CLARANS (Clustering Large Applications based on RANdomized Search) untuk mengatasi keterbatasan metode CLARA. CLARANS menggunakan pencarian acak untuk mencari pusat cluster (medoid). Dalam prosesnya, CLARANS akan mencoba menukar medoid yang ada dengan objek lain secara acak, lalu menghitung apakah pertukaran tersebut membuat total jarak antar objek dalam cluster menjadi lebih kecil, jika iya maka pertukaran dilakukan. Proses ini diulang beberapa kali sampai batas percobaan atau iterasi yang ditentukan tercapai. CLARANS dapat memberikan hasil pengelompokan yang lebih efisien. CLARANS dapat menghasilkan pengelompokan yang lebih optimal dan efisien dibandingkan metode PAM dan CLARA, khususnya ketika diterapkan pada data besar (Ng & Han, n.d.). Oleh karena itu, dalam pengelompokan daerah rawan gempa bumi digunakan metode CLARANS karena lebih efisien dalam menangani data berukuran besar. Data berukuran besar disini

Penelitian dengan menggunakan metode CLARANS sebelumnya pernah dilakukan oleh (Fitri et al., 2024) yang mengelompokan daerah berdasarkan sebaran titik panas di Indonesia. Hasil yang diperoleh, yaitu memiliki jumlah klaster optimal 2 dan nilai silhouette coefficient sebesar 0,896. Terdapat juga penelitian yang dilakukan oleh (Arif, 2016) dengan menggunakan CLARANS untuk menentukan lokasi tempat pengungsian bencana banjir di Bandung Selatan. Hasil menunjukkan bahwa klasterisasi dengan 2 klaster memberikan nilai silhouette coefficient yaitu 0,9. Hal ini menjadi pertimbangan dalam memilih algoritma CLARANS pada penelitian ini karena efektivitasnya telah terbukti pada studi sebelumnya yang menunjukkan hasil evaluasi cluster yang kuat.

Berdasarkan latar belakang yang telah diuraikan, maka tujuan dalam penelitian ini adalah untuk mengetahui hasil pengelompokan daerah rawan gempa bumi di Indonesia menggunakan algoritma CLARANS.

B. Metode

Penelitian ini menggunakan metode *clustering* CLARANS, di mana data yang digunakan merupakan data kejadian gempa bumi di Indonesia tahun 2005 – 2024. Data tersebut merupakan data sekunder yang didapatkan dari *website* USGS (*United States Geological Survey*). Adapun parameter atau variabel yang digunakan, yaitu kedalaman gempa bumi (*depth*) dan kekuatan gempa bumi (*magnitude*). Penelitian ini menggunakan data yang mencakup 836 kejadian gempa bumi dengan *magnitude* > 4,99 serta kedalaman (*depth*) ≤ 300 Km.

Preprocessing Data

Preprocessing data merupakan salah satu tahapan penting dalam data mining yang mencakup proses persiapan dan transformasi data ke dalam bentuk yang sesuai untuk dianalisis lebih lanjut (Hambardzumyan, n.d.). Pada *Preprocessing* data ini terdapat dua kategori utama yang sering digunakan yaitu *cleaning* data dan normalisasi data.

Clustering

Clustering atau pengelompokan merupakan suatu proses untuk membentuk kelompok-kelompok data, di mana objek-objek yang memiliki kesamaan karakteristik akan dikelompokkan ke dalam satu klaster yang sama. Sebaliknya, objek yang memiliki perbedaan karakteristik akan ditempatkan pada klaster yang berbeda. Tujuan utama dari teknik clustering ini adalah untuk meminimalkan jarak antar objek di dalam klaster dan memaksimalkan jarak antar klaster (Hidayatur Rifa et al., n.d.).

CLARANS

CLARANS (Clustering Large Applications based on Randomized Search) merupakan metode clustering yang dirancang untuk menangani data skala besar dengan efisien (Ng & Han, n.d.). CLARANS menggunakan pendekatan berbasis pencarian acak (*randomized search*) untuk menemukan medoid yang optimal tanpa harus memeriksa seluruh kemungkinan yang ada. Metode ini merupakan pengembangan dari dua metode sebelumnya, yaitu PAM (Partitioning Around Medoids) dan CLARA (Clustering Large Applications).

CLARANS memiliki dua parameter utama, yaitu *maxneighbor* dan *numlocal*. *Maxneighbor* merupakan jumlah maksimum tetangga (*neighbor*) yang diperiksa dari suatu node yang dipilih, sedangkan *numlocal* menentukan berapa kali pencarian lokal dilakukan yang berfungsi sebagai acuan seberapa sering node diperiksa atau jumlah iterasi yang dilakukan. Node adalah kombinasi dari beberapa objek yang dipilih sebagai medoid.

Berdasarkan penelitian yang dilakukan oleh (Ng & Han, n.d.) nilai *maxneighbor* ditentukan berdasarkan perbandingan antara $p\% \times k(n - k)$ dengan *minmaxneighbor*, di mana jika $k(n - k) \leq \text{minmaxneighbor}$ maka *maxneighbor* = $k(n - k)$, sedangkan jika lebih besar, maka *maxneighbor* diambil dari nilai terbesar antara $p\% \times k(n - k)$ dan *minmaxneighbor*, nilai *p* adalah 1,25% dari total kemungkinan pertukaran antara medoid dan non-medoid, yang dihitung menggunakan rumus $k(n - k)$, dengan *k* menyatakan jumlah klaster yang terbentuk, sedangkan *n* adalah total objek dalam dataset. Nilai *numlocal* yang digunakan yaitu 2. Nilai-nilai tersebut dipilih guna menjaga keseimbangan antara runtime (waktu pengoperasian) dan kualitas klaster yang terbentuk. Semakin tinggi nilai *maxneighbor*

dan numlocal maka semakin lama runtime yang diperlukan, tetapi semakin tinggi kualitas kluster yang dihasilkan.

Berikut ini merupakan langkah - langkah algoritma CLARANS (Ng & Han, n.d.):

1. Menentukam parameter numlocal dan maxneighbor. Setelah itu, set nilai awal $i = 1$. Nilai i menunjukkan jumlah pencarian lokal (numlocal).
2. Pilih secara acak sebuah node dari $G_{n,k}$ sebagai current node
3. Tetapkan $j = 1$. Nilai j menunjukkan jumlah tetangga yang telah diperiksa dalam satu pencarian lokal.
4. Pilih secara acak node tetangga (S) dari node yang dipilih sebelumnya (current node). Node tetangga didapat dengan mengganti salah satu medoid dari node awal dengan objek lain yang belum menjadi medoid.
5. Menghitung tingkat kemiripan kedua node tersebut (S dan current node) dengan menggunakan ukuran jarak Euclidean berdasarkan persamaan berikut.

$$d_{ij} = \sqrt{\sum_{k=1}^p (x_{ik} - x_{jk})^2} \quad (1)$$

Dimana:

d_{ij} = jarak Euclidean antar objek ke- i dengan objek ke- j

x_{ik} = objek ke- i pada variabel ke- k

x_{jk} = objek ke- j pada variabel ke- k

p = jumlah variabel

6. Kemudian hitung cost dari kedua node tersebut (S dan current node) berdasarkan persamaan berikut.

$$E = \sum_{i=1}^k \sum_{p \in C_i} d(p, O_i) \quad (2)$$

Dimana:

E = total cost

p = objek non-medoid dari kluster ke- i

O_i = medoid dari kluster ke- i

7. Jika S mempunyai cost lebih rendah dibanding current node, maka tukar current node dengan S dan kembali ke langkah 4.
8. Sebaliknya, jika S mempunyai cost lebih tinggi dibanding current node maka naikan nilai j menjadi $j = j + 1$. Jika $j \leq \text{maxneighbor}$ maka kembali ke langkah 4.
9. Jika $j > \text{maxneighbor}$ naikan nilai i menjadi $i = i + 1$. Jika $i < \text{numlocal}$ maka kembali ke langkah 2. Namun, jika $i > \text{numlocal}$ maka proses berhenti dan outputnya adalah node dengan cost paling rendah.

Silhouette Coefficient

Silhouette coefficient merupakan suatu metode yang digunakan untuk mengevaluasi kualitas dan kekuatan hasil *clustering*, yaitu dengan menilai seberapa baik suatu objek ditempatkan dalam suatu cluster (Sains dan Teknologi -Universitas Kanjuruhan Malang Bagus Wira et al., 2019). Metode ini merupakan kombinasi dari metode cohesion dan metode separation. Metode cohesion digunakan untuk mengukur tingkat kedekatan antar data dalam satu cluster yang sama, sedangkan metode separation berfungsi untuk mengukur kedekatan antar cluster yang terbentuk (Utami et al., n.d.). Berikut adalah langkah-langkah metode *silhouette coefficient* (Sains dan Teknologi -Universitas Kanjuruhan Malang Bagus Wira et al., 2019):

1. Menghitung rata-rata jarak objek ke- i dengan semua objek yang berada di dalam satu kelompok A dengan persamaan berikut.

$$a(i) = \frac{1}{|A| - 1} \sum_{j \in A, j \neq i} d(i, j) \quad (3)$$

Dimana:

$a(i)$ = nilai rata-rata objek i dengan semua objek lain dalam kelompok A

A = jumlah anggota kelompok A

$d(i, j)$ = jarak antara objek i dan j

- Menghitung rata-rata jarak objek ke- i dengan semua objek yang berada pada kelompok lain dengan persamaan berikut.

$$d(i, C) = \frac{1}{|C|} \sum_{j \in C} d(i, j) \quad (4)$$

Dimana:

$d(i, C)$ = rata-rata jarak objek i dengan semua objek lain dalam kelompok C dimana $A \neq C$
 C = jumlah anggota kelompok C

- Menentukan nilai minimum yaitu $b(i)$ yang menunjukkan rata-rata jarak objek i dengan objek lain di luar kelompok dapat dituliskan dengan persamaan berikut.

$$b(i) = \min d(i, C) \quad (5)$$

- Menghitung nilai silhouette dengan persamaan berikut.

$$s(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))} \quad (6)$$

Hasil perhitungan $s(i)$ berada pada kisaran -1 hingga 1.

- Menghitung silhouette coefficient yang didefinisikan sebagai rata-rata $s(i)$.

$$SC = \frac{1}{n} \sum_{i=1}^n s(i) \quad (7)$$

Dimana:

n = banyak pengamatan

C. Hasil Penelitian dan Pembahasan

Preprocessing Data

Sebelum melakukan analisis data, perlu dilakukan *preprocessing* data terlebih dahulu. Beberapa tahapan yang perlu dilakukan adalah *cleaning* dan normalisasi data. Pada tahap *cleaning* data, tidak ditemukan adanya missing value dan data duplikat. Selanjutnya, dilakukan normalisasi data untuk variabel *magnitude* dan *depth* karena terdapat perbedaan satuan pada kedua variabel tersebut. Metode normalisasi data yang digunakan adalah metode normalisasi *min-max* yang dapat mengubah skala data berada pada rentang 0-1.

Pengelompokan dengan CLARANS

Untuk memperoleh cluster terbaik, dilakukan beberapa percobaan dengan variasi jumlah *cluster* sebanyak 2 – 4. Penentuan jumlah *cluster* ini didasarkan pada kombinasi karakteristik gempa bumi, seperti *magnitude* dan *depth* (kedalaman) gempa, untuk mengelompokkan wilayah berdasarkan tingkat kerawannya. Masing-masing percobaan diawali dengan penentuan parameter *maxneighbor* dan *numlocal* melalui kombinasi beberapa nilai *minmaxneighbor* dan *numlocal* untuk memperoleh nilai *cost* terkecil.

Pada pengelompokan 2 klaster, kombinasi optimal diperoleh pada *minmaxneighbor* 250 dan *numlocal* 3, dengan nilai *cost* sebesar 94,417. Pengelompokan 3 klaster menghasilkan nilai *cost* terkecil sebesar 86,213 dengan *minmaxneighbor* 500 dan *numlocal* 3. Sementara itu, pada 4 klaster, nilai *cost* terbaik sebesar 88,764 diperoleh dengan *minmaxneighbor* 750 dan *numlocal* 4.

Evaluasi Cluster

Evaluasi cluster dilakukan dengan menggunakan *silhouette coefficient*. Dalam penelitian ini, perhitungan *silhouette coefficient* dilakukan dengan menggunakan 2, 3, dan 4 cluster untuk mengevaluasi kualitas hasil pengelompokan. Nilai *silhouette coefficient* yang lebih tinggi menunjukkan kualitas pengelompokan yang lebih baik. Berikut merupakan tabel yang menyajikan nilai *silhouette coefficient* untuk masing-masing jumlah cluster.

Tabel 1. Nilai *Silhouette Coefficient* pada Setiap Cluster

Jumlah Cluster	<i>Silhouette Coefficient</i>
2	0,6023
3	0,4722
4	0,4014

Berdasarkan tabel diatas, nilai *silhouette coefficient* terbesar diperoleh saat jumlah *cluster* adalah 2, yaitu sebesar 0,6023. Oleh karena itu, jumlah *cluster* yang paling baik yaitu 2 *cluster*.

Interpretasi Hasil Pengelompokan

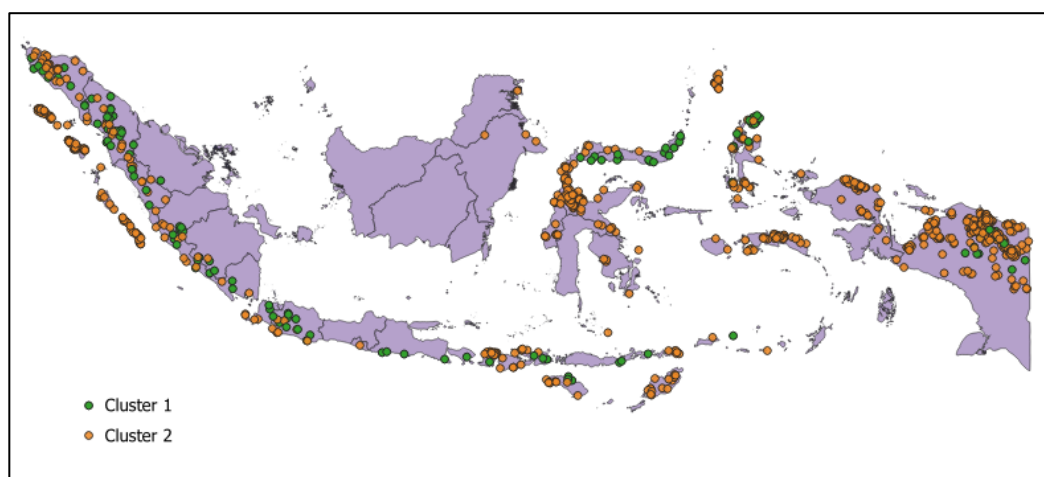
Berikut merupakan tabel rata-rata variabel dan total kejadian gempa bumi untuk masing-masing *cluster*.

Tabel 2. Rata-rata Variabel dan Total dari Setiap Cluster

Cluster	Rata-rata Depth	Rata-rata Magnitude	Total
1	126,0905	5,2	138
2	27,15732	5,3	698

Berdasarkan Tabel 4.21 Cluster 1 terdiri dari 138 kejadian gempa bumi dengan rata-rata depth 126,0905 dan rata-rata magnitude 5,2 yang termasuk sebagai gempa bumi menengah. Jenis gempa ini umumnya menimbulkan getaran kuat dengan potensi kerusakan ringan. Sementara itu, cluster 2 mencakup lebih banyak kejadian gempa bumi, yaitu 698 kejadian dengan rata-rata depth 27,15732 dan rata-rata magnitude 5,3 yang termasuk dalam kategori gempa dangkal. Jenis gempa ini umumnya berpotensi menyebabkan kerusakan yang signifikan. Selain itu, magnitude pada kedua cluster > 4,99 yang berarti keduanya berpotensi menyebabkan kerusakan, namun dampaknya lebih besar pada cluster 2 karena total kejadian yang lebih besar dan depth yang lebih dangkal. Dapat disimpulkan bahwa cluster 1 dikategorikan sebagai daerah dengan kerawanan sedang dan cluster 2 dikategorikan sebagai daerah dengan kerawanan tinggi. Berikut merupakan hasil cluster untuk tingkat kerawanan sedang dan tinggi.

Berikut merupakan visualisasi pada peta berdasarkan pemetaan gempa bumi pada masing-masing cluster dengan keterangan bahwa simbol titik berwarna hijau merupakan cluster dengan tingkat kerawanan sedang dan simbol titik berwarna oranye merupakan cluster dengan tingkat kerawanan tinggi.

**Gambar 1.** Pemetaan Daerah Rawan Gempa Bumi

Berdasarkan Gambar 3 diperoleh informasi bahwa titik berwarna oranye memiliki jumlah yang lebih banyak dibandingkan titik berwarna hijau. Hal ini menunjukkan bahwa potensi terjadinya gempa bumi di Indonesia, khususnya pada wilayah dengan titik oranye tergolong tinggi. Titik oranye tersebar

cukup banyak di wilayah Papua, Maluku, Sulawesi, dan Sumatera. Ini menunjukkan bahwa wilayah tersebut memiliki tingkat kerawanan gempa bumi yang tinggi. Untuk mengetahui sebaran daerah berdasarkan hasil pengelompokan, dilakukan rekapitulasi total kejadian gempa bumi di masing-masing provinsi untuk setiap cluster. Langkah ini bertujuan untuk mengidentifikasi provinsi-provinsi yang termasuk ke dalam cluster dengan tingkat kerawanan yang berbeda. Berikut merupakan tabel yang menyajikan total kejadian gempa bumi berdasarkan provinsi untuk masing-masing cluster.

D. Kesimpulan

Berdasarkan hasil pengelompokan daerah rawan gempa bumi di Indonesia menggunakan algoritma CLARANS diperoleh 2 *cluster* optimal dengan nilai *silhouette coefficient* sebesar 0,6023. *Cluster* 1 dikategorikan sebagai daerah dengan tingkat kerawanan sedang, mencakup total 138 kejadian gempa bumi, sementara *cluster* 2 termasuk dalam kategori dengan tingkat kerawanan tinggi dengan total 698 kejadian gempa bumi. Provinsi Papua dan Sulawesi Tengah menempati total kejadian gempa bumi tertinggi pada *cluster* 2, masing-masing dengan 228 dan 71 kejadian. Sementara itu, provinsi seperti Sumatera Utara dan Maluku Utara memiliki total kejadian gempa bumi yang relatif seimbang di kedua cluster, sehingga mengindikasikan bahwa kedua cluster tersebut sama-sama dominan di daerah tersebut.

Ucapan Terimakasih

Peneliti mengucapkan terima kasih kepada semua pihak yang telah membantu dan mendukung penulis dalam menyelesaikan penelitian ini.

Daftar Pustaka

- Arif, Salmawaty. (2016). *2016 12th International Conference on Mathematics, Statistics, and Their Applications (ICMSA) : proceedings : 4-6 October 2016, Syiah Kuala University, Banda Aceh, Indonesia*. IEEE.
- Data Mining: Concepts and Techniques Second Edition*. (n.d.).
- Fazrilillah, U. N., & Rohaeni, O. (2024). Penerapan Metode Holt's Exponential Smoothing Dalam Memprediksi Jumlah Nasabah Kredit. *DataMath: Journal of Statistics and Mathematics*, 2(1), 11–16.
- Finding Groups in Data*. (n.d.).
- Fitri, S. wisa, Martha, Z., Kurniawati, Y., & Zilrahmi. (2024). Pengelompokan Potensi Kebakaran Hutan/Lahan di Indonesia Berdasarkan Sebaran Titik Panas Menggunakan Metode CLARANS. *UNP Journal of Statistics and Data Science*, 2(3), 273–278. <https://doi.org/10.24036/ujsds/vol2-iss3/182>
- Hambardzumyan, K. (n.d.). *Data Preprocessing in Real-time Education Management System*.
- Hidayatur Rifa, I., Pratiwi, H., Studi Statistika, P., & Matematika dan Ilmu Pengetahuan, F. (n.d.). *IMPLEMENTASI ALGORITMA CLARA UNTUK DATA GEMPA BUMI DI INDONESIA*.
- Irawan, L., & Hermawan Hasibuan, L. (n.d.). *ANALISA PREDIKSI EFEK KERUSAKAN GEMPA DARI MAGNITUDO (SKALA RICHTER) DENGAN METODE ALGORITMA*

ID3 MENGGUNAKAN APLIKASI DATA MINING ORANGE.
<https://doi.org/10.47111/JTI>

- Kharis Pratama, A., Ashaury, H., Rakhmat Umbara Informatika, F., Jenderal Achmad Yani Cimahi Jl Terusan Jend Sudirman, U., Cimahi Sel, K., Cimahi, K., & Barat, J. (2023). KLASIFIKASI DATA GEMPA BUMI DI PULAU JAWA MENGGUNAKAN ALGORITMA EXTREME GRADIENT BOOSTING. In *Jurnal Mahasiswa Teknik Informatika* (Vol. 7, Issue 4). <https://repogempa.bmkg.go.id>,
- Luhung Mustika Budiharti, & Sunendiari, S. (2021). Pemodelan dan Pemetaan Jumlah Penderita Kusta di Jawa Barat dengan Regresi Binomial Negatif dan Flexibly Shaped Spatial Scan Statistic. *Jurnal Riset Statistika*, 1(2), 99–106. <https://doi.org/10.29313/jrs.v1i2.409>
- Ng, R. T., & Han, J. (n.d.). *CLARANS: A Method for Clustering Objects for Spatial Data Mining*.
- Raisa Filmi Suryahadikusumah, & Sutawanir Darwis. (2024). Penentuan Premi Asuransi Gempa Berdasarkan Declustering Katalog Jawa Barat. *Jurnal Riset Statistika*, 127–136. <https://doi.org/10.29313/jrs.v4i2.5029>
- Sains dan Teknologi -Universitas Kanjuruhan Malang Bagus Wira, F., Wira, B., Endy Budianto, A., & Sartika Wiguna, A. (2019). IMPLEMENTASI METODE K-MEDOIDS CLUSTERING UNTUK MENGETAHUI POLA PEMILIHAN PROGRAM STUDI MAHASIWA BARU TAHUN 2018 DI UNIVERSITAS KANJURUHAN MALANG. In *Jurnal Terapan Sains & Teknologi* (Vol. 1, Issue 3).
- Sindi, S., Ratnasari, W., Ningse, O., Sihombing, I. A., Fikrul, P. P. P. A. N. W., Zer, I. R. H., & Hartama, D. (2020). ANALISIS ALGORITMA K-MEDOIDS CLUSTERING DALAM PENGELOMPOKAN PENYEBARAN COVID-19 DI INDONESIA. *Jurnal Teknologi Informasi*, 4(1).
- Susanta, F. F., Pratama, C., Aditya, T., Khomaini, A. F., & Abdillah, H. W. K. (2019). Geovisual Analytics of Spatio-Temporal Earthquake Data in Indonesia. *JGISE: Journal of Geospatial Information Science and Engineering*, 2(2). <https://doi.org/10.22146/jgise.51131>
- Utami, D. S., Retno, D., & Saputro, S. (n.d.). *Program Studi Pendidikan Matematika FKIP UMS PENGELOMPOKAN DATA YANG MEMUAT PENCILANDENGAN KRITERIA ELBOWDAN KOEFISIEN SILHOUETTE (ALGORITME K-MEDOIDS)*.
- Wildan, & Karyana, Y. (2021). Evaluasi Kesalahan Proyeksi Penduduk Tahun 2020 untuk Memproyeksikan Penduduk Tahun 2025 Provinsi Jawa Barat. *Jurnal Riset Statistika*, 1(2), 92–98. <https://doi.org/10.29313/jrs.v1i2.407>