

OPTIMASI PARAMETER DBSCAN DENGAN BAYESIAN OPTIMIZATION UNTUK PEMETAAN MITIGASI BENCANA ALAM

Chalifa Chazar^{1)*}, Muhammad Ghaza Azhar Lesmana²⁾

^{1,2)} Informatika; Institut Teknologi Nasional Bandung

*Corresponding Author: chalifa@itenas.ac.id;

ABSTRAK

Provinsi Jawa Barat memiliki tingkat kerawanan bencana alam yang tinggi, sehingga pemetaan wilayah rawan bencana hingga tingkat kecamatan penting untuk mendukung upaya mitigasi. Penelitian ini bertujuan mengelompokkan wilayah rawan bencana di Provinsi Jawa Barat menggunakan algoritma DBSCAN, yang kinerjanya sensitif terhadap parameter epsilon (ϵ) dan Minimum points (MinPts). Oleh karena itu, Bayesian Optimization digunakan untuk memperoleh parameter yang optimal. Data yang digunakan berupa jumlah kejadian bencana alam per kecamatan periode 2022–2024 dari Open Data Jawa Barat serta data koordinat kecamatan dari OpenStreetMap. Parameter awal DBSCAN ditentukan melalui analisis k-distance plot dan MinPts, kemudian dioptimalkan menggunakan Bayesian Optimization. Hasil evaluasi menggunakan Silhouette Coefficient menunjukkan bahwa DBSCAN dengan parameter baseline ($\epsilon = 0,18$ dan MinPts = 6) menghasilkan nilai Silhouette 0,5413, sedangkan setelah optimasi diperoleh parameter optimal $\epsilon = 0,34$ dan MinPts = 2 dengan peningkatan nilai Silhouette menjadi 0,7080. Hasil ini menunjukkan bahwa Bayesian Optimization efektif dalam meningkatkan kualitas klasterisasi wilayah rawan bencana.

Kata Kunci: *DBSCAN, Bayesian Optimization, Clustering, Mitigasi Bencana, Sistem Informasi Geografis.*

1 PENDAHULUAN

Indonesia memiliki risiko bencana alam yang tinggi karena letaknya di garis khatulistiwa dan terdiri dari banyak pulau. Salah satu bencana besar terjadi di Cianjur, Jawa Barat, pada 21 November 2022, yang mengakibatkan 268 korban jiwa dan kerusakan lebih dari 2.000 rumah (Supendi et al., 2022; Taufan Maulana & Andriansyah, 2024). Kondisi geografis yang kompleks, cuaca yang fluktuatif, serta dampak perubahan iklim menyebabkan Jawa Barat rentan terhadap berbagai bencana seperti gempa bumi, banjir, dan tanah longsor (Rohman et al., 2024).

Di sisi lain, situs resmi BNPB (<https://bnpb.go.id/>) pada tahun 2025 belum menyediakan sistem informasi berbasis data yang menampilkan klasterisasi wilayah rawan bencana secara spesifik di Provinsi Jawa Barat pada tingkat kecamatan di kabupaten. Pemetaan spasial berbasis data hingga level kecamatan penting untuk mendukung mitigasi dan kesiapsiagaan yang lebih tepat sasaran. Oleh karena itu, diperlukan penelitian untuk mengidentifikasi pola bencana dan menyajikannya dalam bentuk sistem informasi yang dapat diakses oleh berbagai pihak (Rohman et al., 2024).

Peraturan Daerah No. 2 Tahun 2013 tentang Penyelenggaraan Penanggulangan Bencana menjelaskan bahwa penanggulangan bencana terdiri atas tiga tahap, yaitu prabencana, tanggap darurat, dan pascabencana. Tahap prabencana mencakup kegiatan mitigasi yang bertujuan untuk mengurangi risiko serta dampak

bencana sebelum kejadian berlangsung. Mitigasi dibedakan menjadi dua bentuk, yaitu mitigasi struktural dan nonstruktural. Mitigasi struktural melibatkan upaya pembangunan fisik, seperti pembangunan tanggul dan pengaturan debit aliran sungai dalam penanggulangan banjir. Sementara itu, mitigasi nonstruktural mencakup kebijakan, perencanaan, dan edukasi, salah satunya melalui pemetaan wilayah rawan bencana (Ayu Sekar Ningrum & Kronika Br. Ginting, 2020; Rusfiana & Lestari, 2021).

Meskipun berbagai penelitian telah dilakukan terkait prediksi curah hujan, aliran sungai, dan fenomena hidrometeorologi lainnya, pemanfaatan dataset dunia nyata berskala besar dalam pengembangan model pengelompokan bencana berbasis pembelajaran mesin masih relatif terbatas. Oleh karena itu, diperlukan pengembangan model kluster yang lebih akurat dan adaptif guna mendukung upaya mitigasi bencana yang lebih efektif (Tin Tin et al., 2024).

Penelitian oleh Tin et al. (2024) menunjukkan bahwa algoritma DBSCAN memperoleh skor Silhouette tertinggi sebesar 0,9883 dibandingkan dengan algoritma clustering lainnya dalam mengelompokkan wilayah rawan bencana alam. Hal ini mencerminkan efektivitas DBSCAN dalam membentuk kluster berbasis kepadatan serta mengenali outlier tanpa memerlukan penentuan jumlah kluster di awal, sehingga cocok untuk data bencana alam (Fadhli Boer et al., 2023). Namun, seluruh algoritma klasterisasi berbasis kepadatan, termasuk DBSCAN, tetap memerlukan parameter input yang sulit ditentukan dan dapat memengaruhi hasil klasterisasi (Falahiazar et al., 2021).

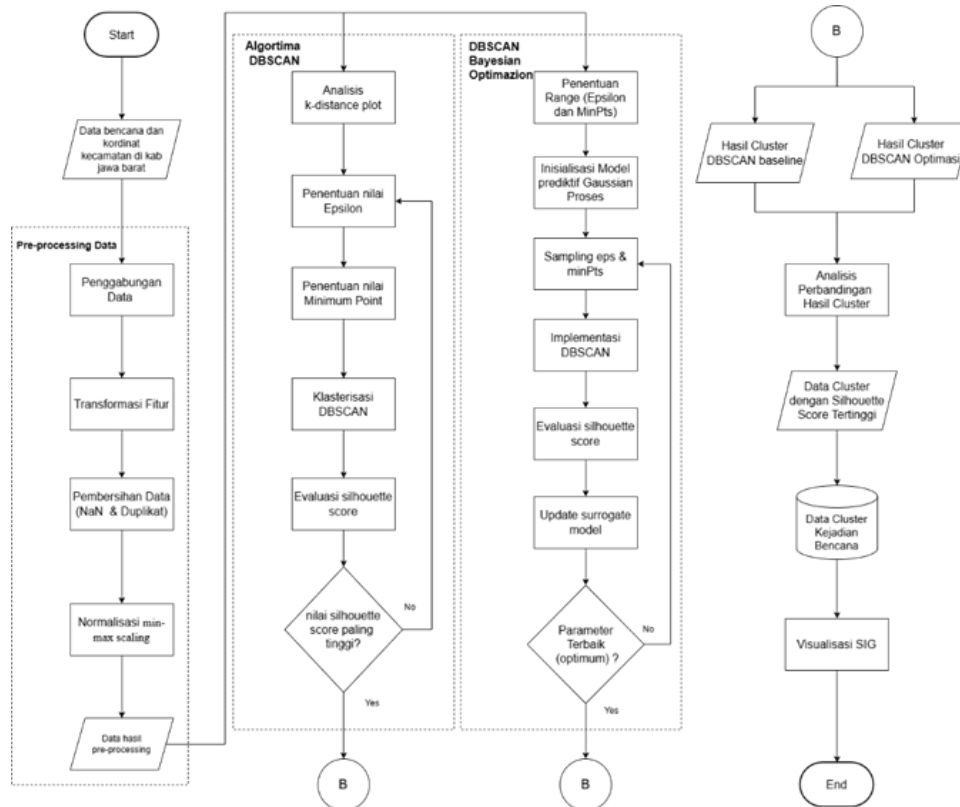
Untuk mengatasi tantangan tersebut, penelitian oleh Hammer et al. (2024) menunjukkan bahwa Bayesian Optimization merupakan pendekatan yang efisien dalam memaksimalkan hasil fungsi tanpa harus mengevaluasi seluruh kombinasi parameter (Hammer et al., 2024). Metode ini sangat sesuai digunakan pada kasus dengan pengetahuan terbatas terhadap parameter optimal, seperti pada penentuan parameter algoritma DBSCAN. Maka dari itu, tujuan utama penelitian ini adalah mengelompokkan wilayah rawan bencana alam di Provinsi Jawa Barat menggunakan algoritma DBSCAN dengan optimasi hyperparameter melalui Bayesian Optimization, serta menyajikan hasilnya dalam Sistem Informasi Geografis berbasis web..

2 METODE

Bagian ini menguraikan metode penelitian yang diterapkan, dimulai dari studi literatur terkait topik yang diangkat hingga alur kerja.

2.1 Flowchart

Untuk memberikan gambaran yang jelas mengenai tahapan penelitian, disusun sebuah flowchart yang memvisualisasikan urutan proses secara runtut dan terstruktur, sebagaimana ditampilkan pada Gambar 1



Gambar 1. Flowchart. Sumber: Hasil Penelitian (2026)

Pada Gambar 1 menunjukkan flowchart penelitian yang terdiri dari lima tahapan utama. Tahap pertama adalah pengumpulan data bencana, yang meliputi data jumlah kejadian bencana angin puting beliung, banjir, gempa bumi, kebakaran hutan, dan tanah longsor di kecamatan pada Kabupaten di Provinsi Jawa Barat periode 2022–2024 dari situs Open Data Jabar, di mana jumlah kejadian bencana pada tingkat kecamatan dihitung sebagai agregasi kejadian tahunan seluruh desa/kelurahan dalam kecamatan tersebut, serta data koordinat geografis dari OpenStreetMap.

Tahap kedua adalah pre-processing data yang diawali dengan penggabungan seluruh data bencana dan data koordinat ke dalam satu dataset. Selanjutnya dilakukan transformasi fitur numerik dengan mengelompokkan setiap jenis kejadian bencana ke dalam lima fitur utama, yaitu banjir (x1), angin puting beliung (x2), gempa bumi (x3), tanah longsor (x4), dan kebakaran hutan (x5). Proses pembersihan data kemudian dilakukan untuk menghilangkan nilai kosong (NaN) dan data duplikat. Terakhir, fitur numerik dinormalisasi menggunakan metode min–max scaling sehingga diperoleh dataset yang siap digunakan pada tahap klusterisasi.

Tahap ketiga merupakan penentuan parameter awal DBSCAN dengan Pendekatan bersifat heuristik, dimulai dari analisis nilai epsilon (ϵ) menggunakan k-distance plot dan penentuan minimum points (MinPts) berdasarkan jumlah fitur numerik. Nilai parameter awal ini digunakan untuk melakukan klusterisasi DBSCAN baseline, dan hasil klusterisasi dievaluasi menggunakan Silhouette Coefficient.

Tahap keempat optimasi hyperparameter DBSCAN menggunakan Bayesian Optimization. Proses ini meliputi penentuan rentang parameter, inisialisasi model prediktif berbasis Gaussian Process, sampling kombinasi ϵ dan MinPts, implementasi DBSCAN, evaluasi Silhouette Coefficient, dan pembaruan model hingga ditemukan kombinasi parameter terbaik.

Tahap akhir penelitian membandingkan hasil klasterisasi DBSCAN baseline dan hasil optimasi berdasarkan nilai Silhouette Coefficient, di mana model dengan nilai tertinggi dipilih sebagai hasil terbaik. Hasil klasterisasi kemudian divisualisasikan dalam Sistem Informasi Geografis (SIG) berbasis web menggunakan Streamlit untuk menampilkan sebaran wilayah rawan bencana, pola bencana dominan, dan informasi mitigasi pada tingkat kecamatan.

2.2 Pengumpulan Data

Data yang digunakan dalam penelitian ini bersumber dari dua sumber utama. Data pertama berupa jumlah kejadian bencana alam per kecamatan pada Kabupaten di Provinsi Jawa Barat periode 2022–2024 yang diperoleh dari Open Data Jawa Barat dan mencakup lima jenis bencana, yaitu banjir, angin puting beliung, gempa bumi, tanah longsor, dan kebakaran hutan. Jumlah kejadian bencana kecamatan dihitung sebagai agregasi kejadian tahunan seluruh desa/kelurahan dalam kecamatan. Kedua, data koordinat geografis (longitude dan latitude) setiap kecamatan yang diperoleh dari OpenStreetMap. Data ini digunakan sebagai atribut spasial untuk merepresentasikan posisi geografis masing-masing kecamatan.

2.3 Pre-processing Data

Tahap pre-processing dilakukan untuk memastikan data siap digunakan dalam proses klasterisasi. Langkah pertama penggabungan lima data jumlah kejadian bencana per kecamatan pada tingkat kabupaten di Provinsi Jawa Barat untuk periode 2022–2024 ke dalam satu dataset, di mana jumlah kejadian bencana kecamatan diperoleh melalui agregasi kejadian tahunan seluruh desa/kelurahan dalam kecamatan tersebut. Selanjutnya transformasi fitur numerik dengan melabeli setiap jenis kejadian bencana dikategorikan ke dalam lima fitur utama, yaitu banjir (x1), angin puting beliung (x2), gempa bumi (x3), tanah longsor (x4), dan kebakaran hutan (x5). Kemudian data koordinat digabungkan berdasarkan kecamatan, sehingga setiap data kecamatan memiliki informasi kejadian bencana sekaligus atribut spasial. Dan dilakukan proses pembersihan data untuk menghilangkan nilai kosong (NaN) dan data duplikat. Dataset hasil pengolahan ini terdiri dari 539 data.

Selanjutnya, data numerik pada fitur kejadian bencana dinormalisasi menggunakan metode min–max scaling. Pada tahap ini, terlebih dahulu dihitung nilai minimum, maksimum, dan rentang nilai (Max–Min) dari setiap fitur pada dataset. Hasil perhitungan tersebut disajikan pada Tabel 1.

Table 1. Hasil Nilai Min Max Max-Min

Fitur	Min	Max
banjir (x1)	0	301
angin topan (x2)	0	54
gempa bumi (x3)	0	694
tanah longsor (x4)	0	173

Sumber: Hasil Penelitian (2026)

Berdasarkan Tabel 1, seluruh fitur numerik memiliki nilai minimum 0, sementara nilai maksimum bervariasi pada tiap fitur. Nilai minimum, maksimum, dan rentang (Max–Min) digunakan sebagai dasar dalam proses normalisasi data jumlah kejadian

bencana per kecamatan. Hasil dari tahap pre-processing ini berupa dataset yang telah siap digunakan untuk proses klasterisasi, dengan contoh ditampilkan pada Tabel 2.

Table 2. Hasil Dataset Pre-processing

Kabupaten	Kecamatan	Longitude	Latitude	x1	x2	x3	x4	X5
Kabupaten Bandung	Arjasari	-7.0583258	107.6509845	0.305	1.000	0.714	1.000	1.000
Kabupaten Bandung	Cikancung	-7.0126297	107.8265136	0.356	0.111	0.428	0.187	0.571
Kabupaten Bandung	Cilengkrang	-6.9049243	107.6941319	0.000	0.333	1.000	0.312	0.000
Kabupaten Bandung	Cileunyi	-6.9411936	107.7434564	0.356	0.444	0.714	0.625	0.000
Kabupaten Bandung	Cimaung	-7.0869685	107.5627557	0.000	0.666	0.428	0.593	0.571
Kabupaten Bandung	Cimenyan	-6.8799677	107.6638481	0.000	0.555	0.000	0.593	0.000
Kabupaten Bandung	Ciwidey	-7.1027234	107.4641673	0.050	0.444	0.428	0.593	0.857
Kabupaten Bandung	Dayeuhkolot	-6.986415	107.626524	1.000	0.000	0.000	0.000	0.000

Sumber: Hasil Penelitian (2026)

Tabel 2 menampilkan struktur dataset hasil pre-processing yang terdiri dari 52lister52 fitur, yaitu kabupaten, kecamatan, longitude, latitude, serta jumlah kejadian bencana banjir (x1), angin putting beliung (x2), gempa bumi (x3), tanah longsor (x4), dan kebakaran hutan (x5). Fitur longitude dan latitude tidak dinormalisasi karena berfungsi sebagai atribut spasial, sedangkan fitur jumlah kejadian bencana telah dinormalisasi menggunakan metode min-max scaling, di mana nilai 0 menunjukkan tidak pernah terjadi kejadian bencana, nilai mendekati 0 menunjukkan kejadian yang jarang terjadi, sedangkan nilai mendekati 1 menunjukkan kejadian bencana yang sering terjadi, sehingga seluruh nilai berada pada rentang 0–1 sebelum proses klasterisasi.

2.4 Impelementasi DBSCAN

Penelitian ini menggunakan algoritma Density-Based Spatial Clustering of Applications with Noise (DBSCAN) untuk mengelompokkan data berdasarkan kepadatan. DBSCAN tidak memerlukan penentuan jumlah 52lister di awal, sehingga sesuai untuk data yang kompleks dan tidak beraturan (Bariklana & Fauzan, 2023). Dua parameter utama yang digunakan adalah epsilon (ϵ), Algoritma ini menggunakan dua parameter utama, yaitu epsilon (ϵ) sebagai radius maksimum kedekatan antar titik dan Minimum points (MinPts) sebagai jumlah minimum titik dalam radius ϵ untuk membentuk sebuah kluster (Monko & Kimura, 2023).

DBSCAN memiliki efektivitas tinggi dalam mengelompokkan data. Dalam konteks data bencana alam, DBSCAN dapat membentuk 52lister dari titik-titik yang berdekatan dan menandai titik-titik terisolasi sebagai noise (Tin Tin et al., 2024) Namun demikian, hasil klasterisasi DBSCAN sangat dipengaruhi oleh nilai parameter yang digunakan, sehingga perbedaan nilai ϵ dan MinPts dapat menghasilkan struktur 52lister yang bervariasi secara signifikan, serta menunjukkan keterbatasan DBSCAN dalam mendeteksi 52lister dengan kepadatan yang berbeda dan ketergantungannya terhadap parameter yang ditentukan oleh pengguna (Bushra & Yi, 2021).

Proses klusterisasi DBSCAN diawali dengan memilih satu titik data secara acak, kemudian menghitung jarak ke titik-titik lainnya menggunakan metrik Euclidean Distance. Jika jumlah tetangga dalam radius ϵ mencapai nilai MinPts, titik tersebut dikategorikan sebagai core point dan akan membentuk 53lister tetangga-tetangganya. Titik yang berada dalam radius ϵ dari core point namun tidak memenuhi jumlah tetangga minimum disebut border point, sedangkan titik yang tidak memenuhi kedua kriteria tersebut diklasifikasikan sebagai noise.

Dalam penelitian ini, parameter ϵ pada tahap awal ditentukan secara manual menggunakan analisis k-distance plot untuk menemukan nilai titik siku (elbow) sebagai radius maksimum antar titik, yang bersifat 53lister53e dan memerlukan interpretasi pengguna. Sementara itu, pemilihan nilai MinPts didasarkan pada pendekatan umum yang mempertimbangkan dimensi data numerik, meskipun literatur menyatakan bahwa tidak terdapat metode estimasi MinPts yang diterima secara universal dan pemilihannya (Monko & Kimura, 2023). Pada tahap ini akan menghasilkan 53lister awal (baseline) yang digunakan sebagai pembanding untuk mengevaluasi pengaruh optimasi hyperparameter pada tahap berikutnya

4.5 Optimasi Hyperparameter

Proses optimasi hyperparameter dilakukan untuk mencari kombinasi nilai epsilon (ϵ) dan MinPts yang menghasilkan kualitas 53lister terbaik pada algoritma DBSCAN. Nilai ϵ menentukan jarak maksimum antar titik dalam satu 53lister, sedangkan MinPts menetapkan jumlah minimum tetangga yang diperlukan untuk membentuk sebuah 53lister. Pemilihan kedua parameter ini sangat memengaruhi hasil klusterisasi.

Dalam penelitian ini, optimasi dilakukan menggunakan Bayesian Optimization (BO), yaitu metode optimasi berbasis 53lister53e53tic yang digunakan untuk mencari parameter optimal dari suatu fungsi objektif yang sulit dievaluasi secara langsung dan mahal secara komputasi. BO bekerja dengan membangun model 53lister53e53tic dari fungsi objektif dan memanfaatkan fungsi akuisisi untuk memutuskan titik parameter yang akan diuji berikutnya.

Secara umum, BO terdiri dari dua komponen utama, yaitu Model surrogate (umumnya Gaussian Process) dan fungsi akuisisi. Model surrogate memetakan distribusi posterior dari fungsi objektif berdasarkan titik evaluasi sebelumnya.

Setiap kombinasi parameter yang diusulkan oleh BO diterapkan pada algoritma DBSCAN untuk melakukan klusterisasi, kemudian dievaluasi menggunakan Silhouette Coefficient. Proses 53lister53e ini menghasilkan kombinasi parameter dengan skor Silhouette tertinggi yang digunakan sebagai parameter teroptimasi untuk membentuk 53lister akhir, yang kemudian dibandingkan dengan hasil baseline.

2.6 Metrix Evaluasi

Pada tahap evaluasi, penelitian ini menggunakan Silhouette Coefficient untuk menilai kualitas hasil klusterisasi. Metrik ini mengukur seberapa dekat suatu titik dengan anggota klasternya (intra-cluster distance) dibandingkan dengan jaraknya terhadap 53lister terdekat yang berbeda (nearest-cluster distance). Nilai koefisien dihitung dengan rumus yang di tunjukan pada Gambar 2 (Rahmawati et al., 2024).

$$s(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))} \quad [1]$$

Sumber: Rahmawati et al (2024)

Gambar 2. Rumus Silhouette Coefficient

Nilainya berada pada rentang -1 hingga 1, di mana mendekati 1 menunjukkan titik berada pada kluster yang tepat, mendekati 0 berarti berada di batas kluster, dan mendekati -1 menandakan kemungkinan tidak sesuai kluster. Metrik ini digunakan untuk membandingkan hasil klusterisasi (Aisyah Nur Fahira & Rani Nooraeni, 2023; Rahmawati et al., 2024) DBSCAN baseline dengan DBSCAN hasil optimasi hyperparameter menggunakan Bayesian Optimization.

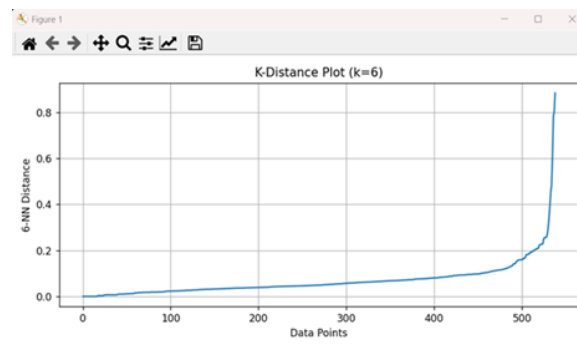
2.7 Visualisasi Kluster

Pada tahap akhir, hasil klusterisasi DBSCAN divisualisasikan dalam bentuk Sistem Informasi Geografis (SIG) berbasis web. Visualisasi ini dikembangkan menggunakan library Streamlit dan menampilkan sebaran spasial wilayah rawan bencana alam berdasarkan kluster yang terbentuk dengan peta dasar OpenStreetMap, sehingga memberikan representasi geografis yang informatif dan mudah dipahami. Selain itu, visualisasi juga menyajikan informasi pendukung berupa karakteristik kluster, jenis bencana yang dominan, serta informasi mitigasi bencana.

3 HASIL DAN PEMBAHASAN

3.1 Hasil DBSCAN Baseline

Pada tahap awal penelitian ini, penentuan parameter DBSCAN dilakukan menggunakan pendekatan heuristik. Dua parameter utama yang ditentukan adalah epsilon (ϵ) dan Minimum points (MinPts). Nilai MinPts ditetapkan sebesar 6, sedangkan nilai epsilon (ϵ) dianalisis menggunakan k-distance plot, sebagaimana divisualisasikan pada Gambar 3.



Gambar 3. *K-distance* Plot Sumber: Hasil Penelitian (2026)

Pada Gambar 3, penentuan nilai epsilon (ϵ) dilakukan menggunakan k-distance plot sebagai alat bantu visual, dengan nilai k disesuaikan dengan MinPts yang digunakan, yaitu $k = 6$. Grafik k-distance menggambarkan jarak ke tetangga terdekat ke-6 untuk setiap titik data yang telah diurutkan secara menaik. Berdasarkan grafik pada Gambar 2, rentang nilai epsilon yang dianalisis berada pada interval 0,08 hingga 0,20. Penentuan nilai epsilon optimal dilakukan dengan mengidentifikasi titik tekuk (knee atau elbow) pada grafik tersebut. Setelah nilai MinPts dan rentang epsilon (ϵ) ditentukan, diperoleh delapan pasangan kombinasi parameter ϵ dan MinPts yang diuji

untuk penentuan parameter awal dalam proses klusterisasi, sebagaimana ditampilkan pada Tabel 3.

Table 3. Hasil evaluasi 8 percobaan parameter awal

Percobaan	eps	MinPts	Silhouett Coefficient
1	0.08	6	-1.0000
2	0.09	6	0.3222
3	0.10	6	-1.0000
4	0.15	6	0.5337
5	0.16	6	0.5366
6	0.17	6	0.5334
7	0.18	6	0.5413
8	0.19	6	0.5353

Sumber: Hasil Penelitian (2026)

Pada Tabel 3 ditampilkan hasil evaluasi menggunakan Silhouette Coefficient dari delapan percobaan penerapan algoritma DBSCAN. Hasil evaluasi menunjukkan bahwa nilai Silhouette pada percobaan ke-1 dan ke-3 bernilai -1,000, sedangkan pada percobaan ke-2 diperoleh nilai 0,3222. Selanjutnya, nilai Silhouette meningkat secara signifikan pada percobaan ke-4 dan cenderung stabil dengan nilai di atas 0,53 pada percobaan ke-4 hingga ke-8.

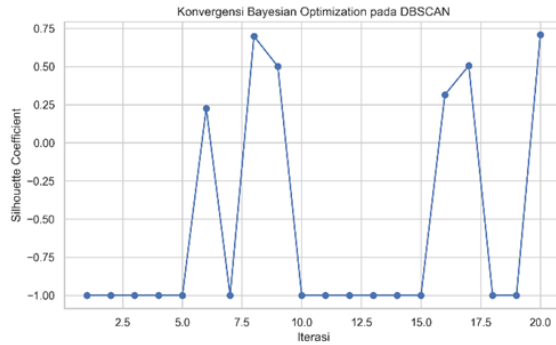
Dari delapan kombinasi parameter yang diuji, percobaan ke-7 dengan kombinasi $\epsilon = 0,18$ dan MinPts = 6 menghasilkan nilai Silhouette Coefficient tertinggi, yaitu sebesar 0,5413. Berdasarkan hasil tersebut, kombinasi parameter $\epsilon = 0,18$ dan MinPts = 6 ditetapkan sebagai parameter baseline, yang selanjutnya digunakan sebagai acuan pembandingan terhadap hasil klusterisasi DBSCAN setelah dilakukan optimasi parameter menggunakan pendekatan Bayesian Optimization.

3.2 Hasil Optimasi Hyperparameter

Pada tahap selanjutnya, dilakukan optimasi parameter DBSCAN menggunakan pendekatan Bayesian Optimization untuk memperoleh kombinasi nilai epsilon (ϵ) dan MinPts sehingga menghasilkan kualitas kluster terbaik. Ruang pencarian parameter ditetapkan dengan rentang epsilon sebesar 0,1 hingga 1,0 dan MinPts antara 2 hingga 10. Proses optimasi diawali dengan pengambilan lima sampel awal secara acak untuk membangun model awal fungsi objektif.

Fungsi objektif yang digunakan dalam proses optimasi adalah nilai Silhouette Coefficient dari hasil klusterisasi DBSCAN. Setiap kombinasi parameter yang diuji dijalankan pada algoritma DBSCAN, kemudian dievaluasi menggunakan metrik Silhouette. Nilai ini selanjutnya digunakan sebagai masukan untuk membangun Gaussian Process dengan kernel Matérn, yang berfungsi sebagai Model surrogate dalam Bayesian Optimization.

Setelah tahap inisialisasi, pemilihan parameter selanjutnya dilakukan menggunakan fungsi akuisisi Expected Improvement (EI). Fungsi ini mempertimbangkan peluang peningkatan nilai Silhouette dibandingkan nilai terbaik sebelumnya, sehingga proses pencarian parameter dapat berjalan lebih efisien dengan menyeimbangkan eksplorasi dan eksploitasi. Proses ini dilakukan secara iteratif hingga mencapai total 20 iterasi. Output grafik dari proses iterasi Bayesian Optimization diperlihatkan pada gambar 4.



Gambar 4. Grafik iterasi *Bayesian Optimization* Sumber: Hasil Penelitian (2026)

Berdasarkan Gambar 4, nilai Silhouette Coefficient pada beberapa iterasi awal masih berada pada $-1,0000$, yang menunjukkan kualitas klusterisasi yang rendah akibat pemilihan kombinasi parameter yang belum optimal. Seiring bertambahnya iterasi, Bayesian Optimization secara bertahap mengarahkan pencarian ke wilayah parameter yang lebih optimal sehingga kualitas klusterisasi meningkat. Rincian hasil pada setiap iterasi disajikan pada Tabel 4.

Table 4 Hasil parameter dan Silhouette Coefficient Bayesian Optimization

Iterasi	eps	MinPts	Silhouett Coefficient
1	0.4371	6	-1.0000
2	0.2651	9	-1.0000
3	0.6388	8	-1.0000
4	0.5012	8	-1.0000
5	0.1523	9	-1.0000
6	0.1000	2	0.2263
7	1.0000	2	-1.0000
8	0.2102	2	0.6989
9	0.1000	3	0.5007
10	0.1000	4	-1.0000
11	0.1000	5	-1.0000
12	0.1000	6	-1.0000
13	0.1000	7	-1.0000
14	0.1000	8	-1.0000
15	0.1000	9	-1.0000
16	0.1184	2	0.3150
17	0.1184	3	0.5062
18	0.1184	4	-1.0000
19	0.1184	5	-1.0000
20	0.3388	2	0.7080

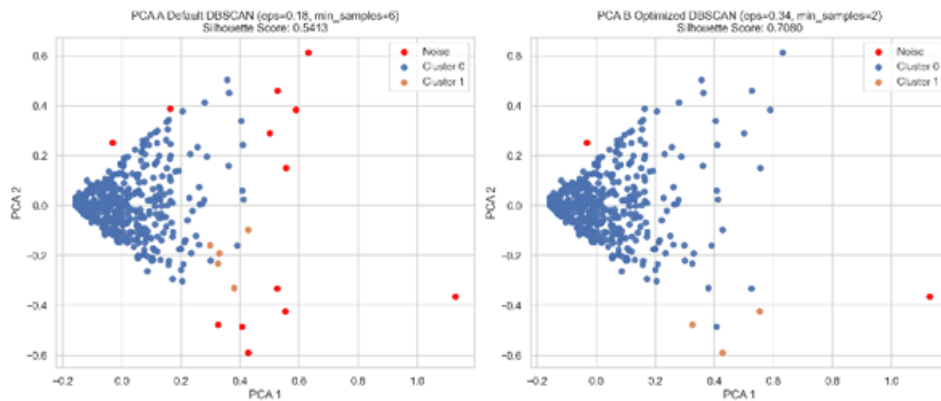
Sumber: Hasil Penelitian (2026)

Berdasarkan Tabel 4, pada iterasi ke-6 nilai Silhouette Coefficient mulai meningkat dan berada pada rentang positif, meskipun masih mengalami fluktuasi pada beberapa iterasi berikutnya. Pada iterasi ke-10 hingga ke-15, nilai Silhouette kembali berada pada $-1,0000$, yang menunjukkan bahwa kombinasi parameter pada rentang tersebut belum menghasilkan klaster yang optimal. Selanjutnya, pada iterasi ke-16 dan ke-17 terjadi peningkatan nilai Silhouette dengan fluktuasi yang lebih kecil, hingga mencapai nilai tertinggi pada iterasi ke-20 dengan $\epsilon = 0,3388$ dan MinPts = 2, yaitu sebesar 0,7080. Hal ini menunjukkan bahwa proses Bayesian Optimization telah mencapai konvergensi, di mana model Gaussian Process berhasil mempelajari

hubungan antara parameter DBSCAN dan kualitas klusterisasi, sehingga kombinasi parameter pada iterasi ke-20 ditetapkan sebagai hasil optimasi terbaik.

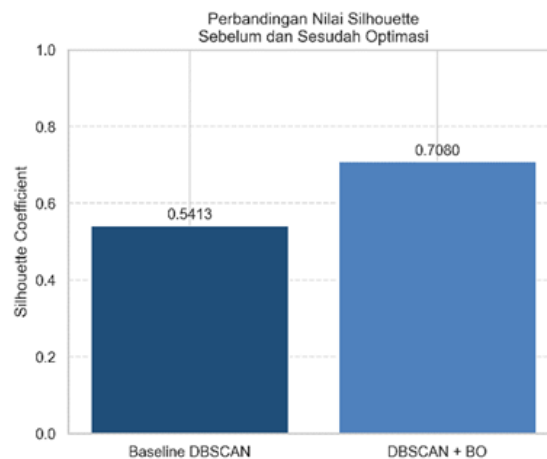
3.3 Analisis Perbandingan Klusterisasi

Selanjutnya dilakukan analisis perbandingan klusterisasi untuk mengevaluasi perbedaan kualitas dan karakteristik kluster antara DBSCAN baseline dan hasil optimasi Bayesian Optimization, yang divisualisasikan menggunakan Principal Component Analysis (PCA) pada Gambar 5 serta grafik batang pada Gambar 6.



Gambar 5. Visualisasi PCA (a) dan PCA (b) Sumber: Hasil Penelitian (2026)

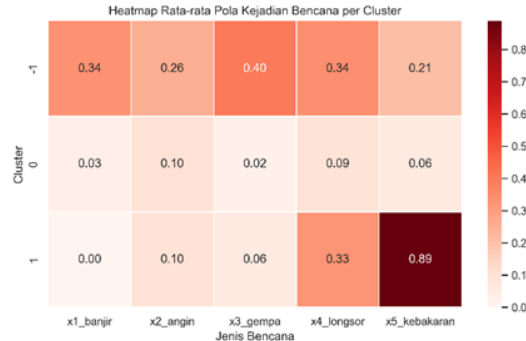
Berdasarkan visualisasi PCA (a) pada Gambar 5, DBSCAN dengan parameter baseline ($\epsilon = 0,18$ dan $\text{MinPts} = 6$) membentuk dua kluster utama dengan sejumlah titik berwarna merah yang masih teridentifikasi sebagai noise, serta pemisahan kluster yang belum optimal. Sebaliknya, visualisasi PCA (b) pada Gambar 6 menunjukkan bahwa DBSCAN dengan parameter hasil optimasi Bayesian menghasilkan pemisahan kluster yang lebih jelas dan jumlah noise yang relatif lebih sedikit.



Gambar 6. Perbandingan Nilai *Silhouette Coefficient* Sumber: Hasil Penelitian (2026)

Gambar 6 menunjukkan hasil klusterisasi sesuai dengan nilai *Silhouette Coefficient*. DBSCAN dengan parameter baseline ($\epsilon = 0,18$ dan $\text{MinPts} = 6$) menghasilkan nilai *Silhouette Coefficient* sebesar 0,5413, sedangkan DBSCAN

dengan parameter hasil optimasi Bayesian Optimization ($\epsilon = 0,34$ dan $\text{MinPts} = 2$) meningkatkan nilai Silhouette Coefficient menjadi 0,7080. Nilai yang lebih mendekati 1 ini menunjukkan struktur kluster yang lebih baik dan pemisahan antar kluster yang lebih jelas. Selanjutnya, dilakukan analisis karakteristik bencana pada masing-masing kluster yang ditampilkan dalam bentuk heatmap pada Gambar 7.



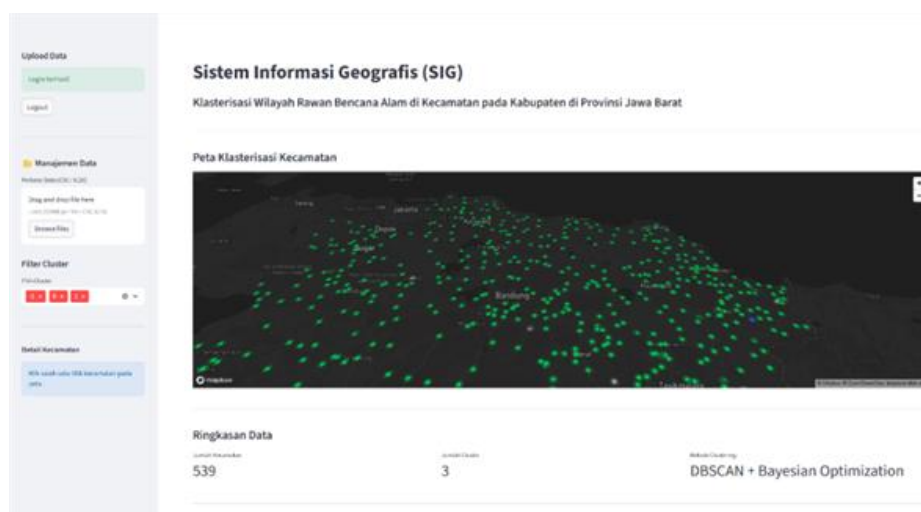
Gambar 7. *Heatmap* karakteristik bencana. Sumber: Hasil Penelitian (2026)

Berdasarkan hasil analisis pada gambar 6 tersebut, kluster 0 yang terdiri dari 534 kecamatan memiliki nilai rata-rata kejadian bencana yang relatif rendah dan cenderung merata pada seluruh jenis bencana. Seluruh nilai rata-rata berada pada kisaran di bawah 0,10, yang menunjukkan bahwa wilayah-wilayah dalam kluster ini tidak didominasi oleh jenis bencana tertentu. Kondisi tersebut mengindikasikan bahwa kluster 0 dapat dikategorikan sebagai wilayah dengan tingkat kerawanan bencana rendah hingga sedang.

Sementara itu, kluster 1 yang terdiri dari 2 kecamatan menunjukkan karakteristik yang kontras, dengan nilai rata-rata tinggi pada bencana kebakaran hutan (x5) dan tanah longsor (x4). Nilai kebakaran hutan yang mencapai 0,89 mengindikasikan dominasi risiko bencana tertentu, sehingga kluster ini dikategorikan sebagai wilayah dengan risiko bencana tinggi dan bersifat spesifik. Selain dua kluster utama tersebut, terdapat 3 kecamatan yang teridentifikasi sebagai noise.

3.4 Visualisasi Cluster

Hasil klusterisasi DBSCAN dengan optimasi Bayesian Optimization divisualisasikan melalui aplikasi web berbasis Streamlit, seperti ditunjukkan pada Gambar 8.



Gambar 7. Sistem Informasi Geografis Sumber: Hasil Penelitian (2026)

Pada Gambar 7 menampilkan antarmuka Sistem Informasi Geografis (SIG) yang terdiri dari header berisi judul dan tujuan sistem, body berupa peta Mapbox yang menampilkan sebaran klaster kecamatan (klaster 0 berwarna hijau, klaster 1 biru, dan noise abu-abu), serta footer yang menyajikan ringkasan hasil klasterisasi DBSCAN dengan optimasi Bayesian Optimization. Sidebar menyediakan fitur penyaringan klaster, pembaruan data, serta informasi detail kecamatan yang mencakup dominasi jenis bencana dan rekomendasi mitigasi. Visualisasi ini memudahkan pengguna dalam memahami sebaran dan tingkat kerawanan bencana hingga tingkat kecamatan.

4 KESIMPULAN

Hasil penelitian menunjukkan bahwa optimasi hyperparameter DBSCAN menggunakan Bayesian Optimization mampu meningkatkan kualitas klasterisasi wilayah rawan bencana, ditunjukkan oleh peningkatan nilai Silhouette Coefficient dari 0,5413 (parameter baseline) menjadi 0,7080 (parameter optimal). Hal ini membuktikan efektivitas Bayesian Optimization dalam menentukan parameter DBSCAN yang optimal. Untuk penelitian selanjutnya disarankan membandingkan DBSCAN dengan metode klasterisasi lain seperti K-Means, Hierarchical Clustering, dan HDBSCAN untuk mengevaluasi performa masing-masing metode dalam pemetaan wilayah rawan bencana.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih yang sebesar-besarnya kepada Ibu Chalifa Chazar, S.T., M.T. selaku dosen pembimbing yang telah dengan sabar memberikan bimbingan, arahan, serta masukan selama proses penyusunan penelitian ini. Penulis juga menyampaikan terima kasih kepada Ibu Lisa Kristiana, Dr. sc., S.T., M.T. dan Bapak Muhammad Ichwan, Ir., M.T. selaku dosen penguji yang telah memberikan saran, kritik, dan arahan yang sangat berharga guna penyempurnaan penelitian.

Selain itu, penulis mengucapkan terima kasih kepada keluarga tercinta atas doa, dukungan moral, dan motivasi yang tidak pernah putus. Tak lupa, penulis juga mengucapkan terima kasih kepada rekan-rekan yang telah memberikan dukungan, bantuan, serta semangat selama proses penyelesaian penelitian ini. Semoga segala kebaikan dan bantuan yang telah diberikan mendapatkan balasan yang setimpal.

DAFTAR PUSTAKA

- Aisyah Nur Fahira, & Rani Nooraeni. (2023). Optimasi Parameter St-DbSCAN Dengan Knn Dan Algoritma Genetika Studi Kasus: Data Bencana Alam Di Pulau Jawa 2021. *Jurnal Komputasi*, 11(1), 24–33.
- Ayu Sekar Ningrum, & Kronika Br. Ginting. (2020). Strategi Penanganan Banjir Berbasis Mitigasi Bencana Pada Kawasan Rawan Bencana Banjir Di Daerah Aliran Sungai Seulalah Kota Langsa. *Geography Science Education Journal (Geosee)*, 1(1), 6–13. [Http://Jurnal.Unsil.Ac.Id/Index.Php/Geosee/Index](http://Jurnal.Unsil.Ac.Id/Index.Php/Geosee/Index)
- Bariklana, M., & Fauzan, A. (2023). Implementation Of The DbSCAN Method For Cluster Mapping Of Earthquake Spread Location. *Barekeng*, 17(2), 867–878. <https://doi.org/10.30598/Barekengvol17iss2pp0867-0878>

- Bushra, A. A., & Yi, G. (2021). Comparative Analysis Review Of Pioneering DBSCAN And Successive Density-Based Clustering Algorithms. *IEEE Access*, 9, 87918–87935. <https://doi.org/10.1109/ACCESS.2021.3089036>
- Fadhli Boer, A., Al Haris, M., Wasono, R., Studi Statistika, P., & Muhammadiyah Semarang, U. (2023). Pengelompokan Daerah Rawan Bencana Di Pulau Sumatera Dengan Metode Density Based Spatial Clustering Of Applications With Noise (DBSCAN). *Prosiding Seminar Nasional Publikasi Hasil Penelitian Dan Pengabdian Masyarakat Universitas Muhammadiyah Semarang*, 6, 389–400.
- Falahiazar, Z., Bagheri, A., & Reshadi, M. (2021). Determining The Parameters Of DBSCAN Automatically Using The Multi-Objective Genetic Algorithm. *JOURNAL OF INFORMATION SCIENCE AND ENGINEERING*, 37(1), 157–183.
- Hammer, J. L., Devanny, A. J., & Kaufman, L. J. (2024). Density-Based Optimization For Unbiased, Reproducible Clustering Applied To Single Molecule Localization Microscopy (Biorxiv). <https://doi.org/10.1101/2024.11.01.621498>
- Monko, G. J., & Kimura, M. (2023). Optimized DBSCAN Parameter Selection: Stratified Sampling For Epsilon And Gridsearch For Minimum Samples. 43–61. <https://doi.org/10.5121/Csit.2023.132004>
- Oktaviani, A. N., Surya Atmaja, A., Putri, K., Aprianto, S. A., Kartiasih, F., Studi, P., Statistik, K., & Stis, S. (2025). ANALISIS PREDIKSI RADIASI MATAHARI DENGAN ALGORITMA MACHINE LEARNING DAN IMPLEMENTASI BAYESIAN OPTIMIZATION DI PROVINSI DKI JAKARTA. *Jurnal Manajemen Informatika & Sistem Informasi (MISI)*, 8(1), 45–58. <https://doi.org/10.36595/Misi.V5i2>
- R, A., Syafitri, U. D., & Erfiani, E. (2024). Effectiveness Of Machine Learning Models With Bayesian Optimization-Based Method To Identify Important Variables That Affect GPA. *JTAM (Jurnal Teori Dan Aplikasi Matematika)*, 8(3), 653. <https://doi.org/10.31764/Jtam.V8i3.21711>
- Rahmawati, T., Wilandari, Y., & Kartikasari, P. (2024). ANALISIS PERBANDINGAN SILHOUETTE COEFFICIENT DAN METODE ELBOW PADA PENGELOMPOKAN PROVINSI DI INDONESIA BERDASARKAN INDIKATOR IPM DENGAN K-MEDOIDS. *Jurnal Gaussian*, 13(1), 13–24. <https://doi.org/10.14710/J.Gauss.13.1.13-24>
- Rohman, D., Annisa, R., Efendi, D. I., & Solahudin, D. (2024). CLUSTERING BENCANA ALAM MENGGUNAKAN K-MEANS PADA WILAYAH JAWA BARAT. *Jurnal Mahasiswa Teknik Informatika*, 8(1), 493–500. <https://www.ncbi.nlm.nih.gov/books/NBK55>
- Rusfiana, Y., & Lestari, M. N. (2021). STRATEGI ANTISIPASI POTENSI BENCANA ALAM DI KABUPATEN BANDUNG. *JURNAL KONSTITUEN*, 3(1), 31–39.
- Supendi, P., Jatnika, J., Sianipar, D., Haidar Ali, Y., Heryandoko, N., Prayitno Adi, S., Karnawati, D., Dwi Anugerah, S., & Fatchurochman, I. (2022). Analisis Gempabumi Cianjur (Jawa Barat) Mw 5.6 Tanggal 21 November 2022.
- Taufan Maulana, A., & Andriansyah, A. (2024). Mitigasi Bencana Di Indonesia. *COMSERVA: Jurnal Penelitian Dan Pengabdian Masyarakat*, 3(10), 3996–4012. <https://doi.org/10.59141/Comserva.V3i10.1213>
- Tin Tin, T., Jia Hao, Y., Chang Yeou, Y., Siew Mooi, L., Ting Yew, G., Olumide Olugbade, T., & Aitizaz, A. (2024). Natural Disaster Clustering Using K-Means, DBSCAN, SOM, GMM, And Mean Shift: An Analysis Of Fema Disaster

Statistics. IJACSA) International Journal Of Advanced Computer Science And Applications, 15(9), 667–681. [Www.ijacsa.thesai.org](http://www.ijacsa.thesai.org)