

Article Identity

Volume 1 Issue 1 Pages 39–46

October 2026

E-ISSN xxxx-xxxx

Article History

Received 04 August 2026

Revised 30 August 2026

Accepted 08 September 2026

Published 14 September 2026

Keywords

Disaster Mitigation; Indonesian Provinces; Fuzzy C-Means; Clustering; Silhouette Coefficient

Copyright © 2026 Author(s).

Editorial office: Journal of Statistics and Data Analytics, PT. Gammabeta Edu and Tech.

To Cite this Article: Aning Saidatul Abadiyah et al. Analisis Kluster Provinsi di Indonesia Berdasarkan Indikator Mitigasi Bencana dan Karakteristik Wilayah Menggunakan Metode Fuzzy C-Means. *Journal of Statistics and Data Analytics*. 2026;1(1):39–46.

RESEARCH ARTICLE

Analisis Kluster Provinsi di Indonesia Berdasarkan Indikator Mitigasi Bencana dan Karakteristik Wilayah Menggunakan Metode Fuzzy C-Means

Aning Saidatul Abadiyah¹, Nadiyah Syahri Rehand², Nur Roudlotur Rosidah³, Dian Candra Rini Novitasari^{4*}, Adelia Rova Chumairo⁵

^{1,2,3,4,5}Program Studi Matematika, Fakultas Sains dan Teknologi, UIN Sunan Ampel Surabaya, Indonesia

Corresponding author: diancrini@uinsa.ac.id

Abstract. *The high frequency of natural disasters across Indonesia underscores the urgent need for a structured, data-driven mitigation system. Situated at the convergence of active tectonic plates, each region in Indonesia exhibits varying levels of disaster preparedness and response capacity. This study applies the Fuzzy C-Means method to cluster Indonesia's 38 provinces based on their natural disaster mitigation capacity. Variables used include tsunami early warning systems, evacuation routes, number of disasters, number of sub-districts, land area, and population, all sourced from the Central Statistics Agency in 2024. Prior to clustering, data underwent correlation analysis and normalization through the Min-Max Scaler method. Two clusters emerged: the first representing provinces with high mitigation capacity comprising 5 provinces (13.16%), and the second representing those with low mitigation capacity comprising 33 provinces (86.84%). Cluster quality was measured using the silhouette coefficient, yielding a score of 0.744, which reflects a very strong cluster structure. These findings are intended to support the government in formulating more targeted disaster mitigation policy priorities.*

Abstrak. *Tingginya frekuensi bencana alam di Indonesia menjadi latar belakang pentingnya sistem mitigasi yang terstruktur dan berbasis data. Sebagai negara yang secara geografis berada di jalur pertemuan lempeng tektonik aktif, setiap wilayah di Indonesia memiliki tingkat kesiapsiagaan yang berbeda-beda. Penelitian ini menerapkan metode Fuzzy C-Means untuk membentuk pengelompokan 38 provinsi berdasarkan kapasitas mitigasi bencana alamnya. Variabel yang digunakan meliputi sistem peringatan dini tsunami, jalur evakuasi, jumlah bencana, jumlah kelurahan, luas wilayah, dan jumlah penduduk, yang seluruhnya bersumber dari Badan Pusat Statistik tahun 2024. Data diproses melalui tahap analisis korelasi dan normalisasi dengan Min-Max Scaler sebelum dilakukan klusterisasi. Dua kelompok terbentuk dari proses ini: kelompok pertama dengan kapasitas mitigasi tinggi mencakup 5 provinsi (13,16%), sedangkan kelompok kedua dengan kapasitas mitigasi rendah mencakup 33 provinsi (86,84%). Kualitas pengelompokan diukur menggunakan silhouette coefficient dengan nilai 0,744, yang mencerminkan struktur kluster sangat kuat. Temuan ini diharapkan menjadi landasan bagi pemerintah dalam menyusun prioritas kebijakan mitigasi bencana secara lebih terarah.*

1. PENDAHULUAN

Kerugian akibat bencana alam mencakup dimensi yang sangat luas, mulai dari hilangnya nyawa, rusaknya infrastruktur, hingga terguncangnya stabilitas ekonomi dan sosial suatu wilayah [1]. Interaksi antara potensi bahaya alam dan kerentanan wilayah dalam manajemen risiko menjadi faktor penentu besarnya dampak yang ditimbulkan [2]. Mitigasi bencana dipandang sebagai pendekatan yang lebih strategis dalam pengurangan risiko dibandingkan penanganan yang dilakukan setelah bencana terjadi. Karakteristik geografis Indonesia yang berada pada pertemuan tiga lempeng tektonik utama serta termasuk dalam kawasan Cincin Api Pasifik menempatkan Indonesia sebagai salah satu negara dengan tingkat kerawanan bencana yang tinggi di dunia [3]. Badan Nasional Penanggulangan Bencana mencatat sedikitnya 5.400 kejadian bencana sepanjang tahun 2023 [4], menegaskan bahwa sistem mitigasi yang terstruktur dan berbasis data valid merupakan kebutuhan yang tidak dapat ditunda, terutama mengingat kondisi tiap wilayah di Indonesia tidaklah sama.

Pola kebencanaan di Indonesia tidak seragam; setiap provinsi memiliki profil risiko yang unik, bergantung pada kondisi alam, kepadatan penduduk, dan kapasitas daerah [5]. Data BPS tahun 2024 menunjukkan bahwa provinsi seperti Aceh dan Sumatera Utara sudah memiliki sistem peringatan dini tsunami, sementara provinsi seperti Riau, Jambi, dan sebagian wilayah Papua belum memilikinya sama sekali [6]. Ketidakmerataan ini berdampak pada sektor ekonomi, pertanian, dan ketahanan sosial yang responsnya berbeda di tiap wilayah [7], sehingga diperlukan langkah konkret untuk memetakan kesenjangan tersebut secara menyeluruh.

Pengelompokan provinsi berdasarkan kapasitas mitigasi memiliki urgensi yang tinggi karena memberikan kontribusi secara teoritis maupun praktis. Kontribusi teoritis penelitian ini terletak pada penerapan metode *Fuzzy C-Means* dengan memanfaatkan indikator kesiapsiagaan sebagai dasar pengelompokan, bukan semata-mata frekuensi kejadian bencana. Kontribusi praktisnya berupa penyediaan informasi yang dapat digunakan pemerintah sebagai dasar penyusunan kebijakan mitigasi yang lebih tepat sasaran. Ketiadaan pengelompokan berdasarkan kapasitas mitigasi berpotensi menghasilkan kebijakan yang bersifat seragam meskipun tingkat risiko antarprovinsi berbeda, sehingga provinsi dengan kapasitas mitigasi yang rendah dapat kurang memperoleh perhatian dalam proses pengambilan keputusan strategis di masa mendatang.

Penelitian klasterisasi dalam bidang kebencanaan telah banyak dilakukan sebelumnya. Penelitian [8] menggunakan *Fuzzy C-Means* pada data BPS tahun 2021 untuk mengelompokkan provinsi rawan bencana di Indonesia dan memperoleh dua klaster dengan nilai *silhouette* tertinggi 0,744, di mana klaster pertama mencakup Aceh, Sumatera Utara, Jawa Barat, Jawa Tengah, dan Jawa Timur. Penelitian [9] menerapkan metode *K-Means* untuk mengelompokkan wilayah bencana banjir di Indonesia dan memperoleh dua klaster terbaik dengan nilai *silhouette coefficient* sebesar 0,8461. Hasil pengelompokan menunjukkan bahwa 35 wilayah termasuk dalam kategori rawan banjir tinggi, sedangkan 463 wilayah lainnya berada pada kategori rawan banjir rendah. Penelitian [10] menerapkan FCM untuk mengelompokkan 34 provinsi berdasarkan enam indikator kualitas pendidikan dari data BPS periode 2022–2024, dengan dua klaster optimal yang diidentifikasi melalui nilai *partition coefficient index* tertinggi. Penelitian [11] mengaplikasikan FCM pada data penyakit menular dan berhasil membentuk tiga klaster optimal dengan nilai *silhouette coefficient* sebesar 0,9151 dan 0,91899 untuk tahun 2019 dan 2020, sekaligus membuktikan keandalan FCM dalam mengevaluasi pemerataan fasilitas kesehatan antardaerah. Penelitian [12] menggunakan metode *Fuzzy C-Means* untuk mengelompokkan data penerima Bantuan Pangan Non Tunai (BPNT) di Kota Pekanbaru dan memperoleh dua klaster optimal dengan nilai *silhouette coefficient* sebesar 0,4267. Klaster pertama beranggotakan 331 data, sedangkan klaster kedua terdiri atas 351 data. Hasil pengelompokan menunjukkan bahwa klaster pertama memiliki karakteristik yang mencerminkan kondisi ekonomi yang relatif lebih rendah dibandingkan klaster kedua, sehingga diprioritaskan sebagai penerima bantuan, sementara klaster kedua menunjukkan kondisi ekonomi yang relatif lebih baik [12]. Berbagai penelitian tersebut menunjukkan bahwa metode klasterisasi, khususnya *Fuzzy C-Means*, mampu menghasilkan pengelompokan yang baik pada berbagai jenis data. Sebagian besar penelitian masih berfokus pada tingkat kerawanan, intensitas, atau kejadian bencana, serta pada bidang selain kebencanaan seperti pendidikan, kesehatan, dan bantuan sosial. Penelitian yang secara khusus mengelompokkan provinsi berdasarkan kapasitas mitigasi bencana alam menggunakan indikator kesiapsiagaan dan kemampuan mitigasi daerah masih relatif terbatas.

Penelitian ini mengangkat kapasitas mitigasi bencana alam sebagai objek kajian utama. Berbeda dengan penelitian terdahulu yang umumnya menggunakan indikator kerawanan atau kejadian bencana, penelitian ini memanfaatkan indikator yang mencerminkan kesiapsiagaan dan kemampuan mitigasi setiap provinsi sehingga mampu memberikan perspektif yang lebih berorientasi pada kapasitas penanggulangan risiko. Pengelompokan dilakukan menggunakan metode *Fuzzy C-Means*. Pendekatan tersebut diharapkan mampu memberikan gambaran yang lebih komprehensif mengenai kondisi mitigasi bencana di Indonesia serta memperlihatkan karakteristik masing-masing kelompok provinsi secara lebih jelas.

Fuzzy C-Means merupakan metode klasterisasi yang memungkinkan setiap objek memiliki derajat keanggotaan pada lebih dari satu klaster. Proses pengelompokan dilakukan melalui pembentukan pusat klaster dan pembaruan nilai keanggotaan secara berulang hingga mencapai kondisi yang stabil [13]. Karakteristik kapasitas mitigasi bencana yang bersifat kompleks dan dipengaruhi oleh berbagai indikator menyebabkan batas antarkelompok tidak selalu dapat dibedakan secara tegas. Kemampuan FCM dalam memberikan derajat keanggotaan pada setiap objek menjadikan metode ini sesuai untuk digunakan dalam mengelompokkan provinsi berdasarkan kapasitas mitigasi bencana.

Berdasarkan uraian yang telah dipaparkan, penelitian ini bertujuan untuk mengelompokkan provinsi-provinsi di Indonesia berdasarkan kapasitas mitigasi bencana alam menggunakan metode *Fuzzy C-Means* (FCM). Indikator yang digunakan mencerminkan tingkat kesiapsiagaan dan kemampuan mitigasi masing-masing provinsi. Melalui pengelompokan ini, diharapkan perbedaan profil mitigasi antarprovinsi dapat dipetakan secara lebih jelas, sehingga hasil penelitian dapat dijadikan acuan dalam pengambilan keputusan strategis terkait pengurangan risiko bencana dan penguatan kapasitas masyarakat secara nasional.

2. KAJIAN TEORI

2.1. Bencana Alam

Bencana alam merujuk pada peristiwa yang berasal dari aktivitas alam, antara lain gempa bumi, tsunami, erupsi gunung berapi, banjir, kekeringan, angin puting beliung, serta tanah longsor [14]. Kejadian semacam ini pada umumnya berlangsung secara tiba-tiba dan sangat sulit untuk diprediksi kapan terjadinya [11]. Kerugian yang diakibatkan sangat luas, mulai dari hilangnya nyawa manusia, rusaknya ekosistem dan lingkungan hidup, hancurnya berbagai infrastruktur, hingga timbulnya gangguan psikologis yang berdampak panjang terhadap kehidupan masyarakat. Minimnya pemahaman warga terhadap potensi dan risiko bencana di wilayah tempat tinggal mereka menjadi salah satu faktor yang memperparah dampak kerugian tersebut, sehingga upaya mitigasi bencana menjadi sangat mendesak untuk dilaksanakan [15].

Mitigasi bencana pada dasarnya mencakup serangkaian upaya yang dilakukan sebagai bentuk antisipasi terhadap tingkat kerawanan bencana sekaligus menekan risiko yang dapat ditimbulkan [16]. Penerapan mitigasi bencana difokuskan pada penurunan tingkat risiko serta pengendalian dampak buruk yang mungkin terjadi, termasuk kehilangan nyawa dan kerusakan harta benda, sekaligus memperkuat kapasitas masyarakat dalam melakukan tindakan kewaspadaan dan respons terhadap ancaman bencana [17]. Berdasarkan bentuk pelaksanaannya, mitigasi bencana dibedakan menjadi dua jenis, yaitu mitigasi struktural dan mitigasi nonstruktural. Mitigasi struktural diwujudkan melalui pembangunan fisik seperti tanggul penahan banjir, jalur evakuasi, serta konstruksi bangunan tahan gempa. Mitigasi nonstruktural dilaksanakan melalui penyusunan kebijakan, kegiatan sosialisasi dan edukasi, serta penyebaran informasi kepada masyarakat luas guna meningkatkan kesiapsiagaan dalam menghadapi ancaman bencana [18].

2.2. Analisis Korelasi

Tingkat hubungan antarvariabel dapat diidentifikasi melalui analisis korelasi, yaitu metode statistika yang digunakan untuk mengukur keterkaitan antara dua variabel atau lebih [19]. Variabel dalam konteks ini dapat dimaknai sebagai karakteristik atau atribut yang menjadi objek pengamatan dalam suatu penelitian. Teknik ini memusatkan perhatian pada hubungan antarvariabel tanpa mempermasalahkan hubungan sebab-akibat di antara keduanya. Koefisien korelasi digunakan sebagai ukuran hubungan antarvariabel dengan nilai yang berada pada kisaran -1 hingga 1 . Koefisien positif mencerminkan hubungan yang bergerak searah, sedangkan koefisien negatif mencerminkan hubungan yang bergerak berlawanan arah [20].

2.3. Normalisasi

Normalisasi merupakan suatu proses pengorganisasian struktur tabel dalam basis data dengan tujuan untuk memperoleh struktur yang lebih baik, meminimalkan pengulangan data, serta menjaga konsistensi data secara keseluruhan [21]. Proses ini dilakukan dengan cara membagi dan mengelompokkan data ke dalam tabel-tabel yang sesuai berdasarkan keterkaitan antardata. Tujuan utama normalisasi adalah menjaga konsistensi data sekaligus meminimalkan kemungkinan terjadinya kesalahan dalam pengelolaan basis data. Struktur data yang telah dinormalisasi akan memudahkan proses penambahan, pembaruan, maupun penghapusan data karena setiap data tersimpan pada tempat yang tepat. Normalisasi juga berkontribusi dalam meningkatkan efisiensi penggunaan ruang penyimpanan. Data yang tertata dengan baik akan memudahkan proses pengelolaan dan pemeliharaan basis data dalam jangka panjang, sehingga menghasilkan struktur tabel yang lebih terorganisasi dan mudah diakses sesuai kebutuhan sistem [22].

Pada penelitian ini, data dinormalisasi terlebih dahulu menggunakan metode Min-Max sebelum tahap klusterisasi. Metode ini berfungsi untuk menyeragamkan skala setiap variabel ke dalam rentang $[0, 1]$ dengan rumus sebagai berikut [23]:

$$x' = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (1)$$

dengan x' merupakan hasil proses normalisasi, x merupakan data sebelum normalisasi, $x_{\min}$ merupakan batas bawah nilai variabel, dan $x_{\max}$ merupakan batas atas nilai variabel.

2.4. Fuzzy C-Means (FCM)

Fuzzy C-Means (FCM) merupakan teknik klusterisasi nonhierarkis yang memberikan nilai keanggotaan pada setiap objek terhadap seluruh kluster berdasarkan tingkat kedekatannya dengan pusat kluster [24]. Berbeda dengan metode klusterisasi konvensional, FCM memungkinkan setiap data memiliki derajat keanggotaan pada lebih dari satu kluster secara bersamaan. Proses FCM diawali dengan penentuan pusat kluster awal yang kemudian diperbarui secara bertahap melalui serangkaian iterasi. Pada setiap iterasi, nilai derajat keanggotaan dan pusat kluster dihitung ulang sampai kondisi stabil tercapai. Pembaruan ini bertujuan untuk mendapatkan pusat kluster yang paling representatif terhadap karakteristik data. Hasil pengelompokan akhir ditentukan berdasarkan nilai minimum fungsi objektif yang mencerminkan tingkat kedekatan data terhadap pusat kluster dan derajat keanggotaannya.

Tahapan algoritma *Fuzzy C-Means* mengikuti prosedur pada [25] sebagai berikut.

1. Masukkan data dalam bentuk matriks berukuran $n \times m$, dengan n sebagai jumlah data, m sebagai jumlah variabel, dan X_{ij} sebagai data ke- i pada variabel ke- j .
2. Tentukan parameter yang diperlukan, yaitu banyaknya kluster c , nilai eksponen pembobot w , batas akhir iterasi MaxIter , nilai toleransi kesalahan $\varepsilon = 0.00001$, nilai awal fungsi objektif $P_0 = 0$, dan iterasi pertama $t = 1$.
3. Tetapkan angka acak μ_{ik} sebagai elemen penyusun matriks partisi awal U . Jumlah setiap baris matriks dihitung dengan

$$Q_i = \sum_{k=1}^c \mu_{ik} \quad (2)$$

4. Selanjutnya, nilai keanggotaan dinormalisasi menggunakan

$$\mu_{ik} \leftarrow \frac{\mu_{ik}}{Q_i} \quad (3)$$

5. Pusat kluster ke- k , V_{kj} , untuk $k = 1, 2, \dots, c$ dan $j = 1, 2, \dots, m$, dihitung menggunakan

$$V_{kj} = \frac{\sum_{i=1}^n (\mu_{ik})^w X_{ij}}{\sum_{i=1}^n (\mu_{ik})^w} \quad (4)$$

6. Fungsi objektif pada iterasi ke- t dihitung menggunakan

$$P_t = \sum_{i=1}^n \sum_{k=1}^c \left[\sum_{j=1}^m (X_{ij} - V_{kj})^2 \right] (\mu_{ik})^w \quad (5)$$

7. Matriks partisi diperbarui menggunakan

$$\mu_{ik} = \frac{\left[\sum_{j=1}^m (X_{ij} - V_{kj})^2 \right]^{-1/(w-1)}}{\sum_{\ell=1}^c \left[\sum_{j=1}^m (X_{ij} - V_{\ell j})^2 \right]^{-1/(w-1)}} \quad (6)$$

8. Periksa kriteria penghentian iterasi. Iterasi berhenti ketika $|P_t - P_{t-1}| < \varepsilon$ atau $t > \text{MaxIter}$. Jika kriteria belum terpenuhi, nilai iterasi diperbarui menjadi $t = t + 1$ dan proses kembali ke langkah ke-5.

2.5. Denormalisasi

Denormalisasi merupakan proses kebalikan dari normalisasi [26]. Proses ini dilaksanakan dengan cara menggabungkan kembali data yang sebelumnya telah dipisahkan ke dalam beberapa tabel guna meningkatkan kecepatan akses dan kinerja pengambilan data. Tujuan utama denormalisasi adalah mempercepat proses pengambilan informasi dari basis data. Penerapan denormalisasi dapat mengurangi kebutuhan penggabungan tabel pada saat pencarian data dilakukan. Struktur basis data hasil denormalisasi umumnya mengandung data yang tersimpan

secara berulang, namun kondisi tersebut justru membantu meningkatkan kecepatan akses terutama pada sistem basis data yang mengelola data dalam jumlah besar [22].

Pada penelitian ini, proses denormalisasi dilakukan untuk mengembalikan data ke skala aslinya setelah melalui proses normalisasi. Proses tersebut dihitung menggunakan persamaan berikut [23]:

$$x_i = x' (x_{\max} - x_{\min}) + x_{\min}. \quad (7)$$

dengan x_i merupakan data aktual, x' merupakan data hasil normalisasi, $x_{\max}$ merupakan batas atas data empiris, dan $x_{\min}$ merupakan batas bawah data empiris.

2.6. Silhouette Coefficient

Silhouette coefficient adalah metode evaluasi yang digunakan untuk mengukur seberapa baik kualitas hasil pengelompokan data [27]. Penilaian dilakukan dengan mempertimbangkan tingkat kedekatan suatu data terhadap anggota dalam klasternya sendiri serta tingkat perbedaannya dengan data yang berada pada klaster lain. Rentang nilai *silhouette coefficient* berada antara -1 hingga 1 . Kualitas pengelompokan dinilai semakin optimal apabila nilai *silhouette coefficient* mendekati angka 1 . Nilai kurang dari nol dapat diartikan bahwa objek memiliki tingkat kesamaan yang lebih tinggi terhadap anggota klaster lain [28]. Nilai *silhouette coefficient* untuk elemen ke- i dihitung dengan

$$S_i = \frac{b_i - a_i}{\max(a_i, b_i)}, \quad (8)$$

dengan a_i merupakan rata-rata jarak objek ke- i terhadap anggota klaster yang sama dan b_i merupakan rata-rata jarak objek ke- i terhadap anggota klaster lain terdekat.

Penilaian kualitas hasil klasterisasi berdasarkan nilai koefisien *silhouette* pada kriteria Kaufman dan Rousseeuw ditunjukkan pada Tabel 1 [28].

Tabel 1. Kriteria Silhouette Coefficient

Rentang Nilai Silhouette Coefficient	Interpretasi
0.71–1.00	Klaster sangat kuat
0.51–0.70	Klaster baik
0.26–0.50	Klaster kurang kuat
≤ 0.25	Klaster buruk

3. METODE PENELITIAN

Data yang digunakan dalam penelitian ini merupakan data sekunder yang diperoleh langsung dari laman resmi BPS [6]. Data dapat dilihat pada Tabel 2.

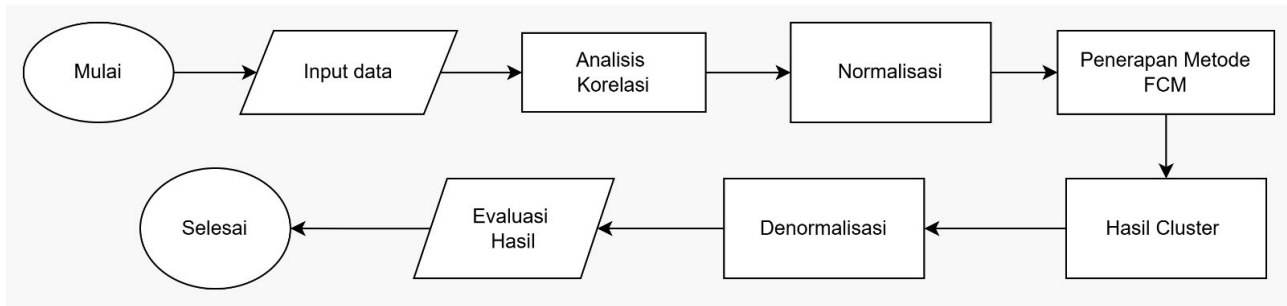
Tabel 2. Data Sampel Penelitian

Provinsi	Sistem Peringatan Dini Tsunami	Jalur Evakuasi	Jumlah Bencana	Jumlah Kelurahan	Luas Wilayah	Jumlah Penduduk
Aceh	63	448	103	6516	56834	5570
Sumatera Utara	16	231	164	6113	72460	2640
Sumatera Barat	90	288	67	1265	42119	5790
Riau	0	238	66	1867	89935	6970
Jambi	0	49	21	1585	49026	3800
⋮	⋮	⋮	⋮	⋮	⋮	⋮
Papua Barat Daya	18	40	4	1018	39122	616
Papua	15	97	7	1029	82680	1090
Papua Selatan	0	0	3	690	117849	545
Papua Tengah	0	37	5	1208	61072	1360
Papua Pegunungan	0	5	3	2634	51213	1470

Dataset yang digunakan dalam penelitian ini bersumber dari BPS tahun 2024, mencakup informasi Banyaknya Desa atau Kelurahan dengan Upaya Antisipasi atau Mitigasi Bencana Alam di 38 provinsi di Indonesia. Variabel-variabel yang digunakan meliputi sistem peringatan dini tsunami, jalur evakuasi, jumlah kejadian bencana, jumlah kelurahan, luas wilayah, serta jumlah penduduk. Keseluruhan variabel tersebut dijadikan sebagai indikator dalam proses pengelompokan provinsi berdasarkan kondisi mitigasi bencana alam melalui pendekatan *Fuzzy C-Means* (FCM).

Pada proses pengelompokan data, diperlukan sebuah *flowchart* yang menggambarkan alur tahapan perhitungan secara sistematis. *Flowchart* ini berfungsi sebagai panduan dalam pelaksanaan algoritma sehingga proses pembentukan kluster dapat tergambar secara sistematis dan jelas. Alur kerja algoritma *Fuzzy C-Means* dalam penelitian ini ditunjukkan pada Gambar 1.

Tahapan penelitian diawali dengan pengumpulan dan penyusunan data mitigasi bencana yang bersumber dari laman resmi BPS. Data yang telah terkumpul selanjutnya dianalisis menggunakan korelasi Pearson guna mengidentifikasi tingkat keterkaitan antarvariabel penelitian. Data kemudian dinormalisasi menggunakan metode Min-Max sehingga seluruh variabel memiliki skala yang seragam. Data hasil normalisasi diolah dengan algoritma *Fuzzy C-Means* guna membentuk kluster provinsi berdasarkan kesamaan karakteristik mitigasi bencana. Hasil kluster yang terbentuk selanjutnya dikembalikan ke skala semula melalui proses denormalisasi agar interpretasi hasil dapat dilakukan dengan lebih mudah. Tahap akhir adalah evaluasi kualitas kluster menggunakan *silhouette coefficient* guna mengetahui tingkat ketepatan pengelompokan yang telah terbentuk.

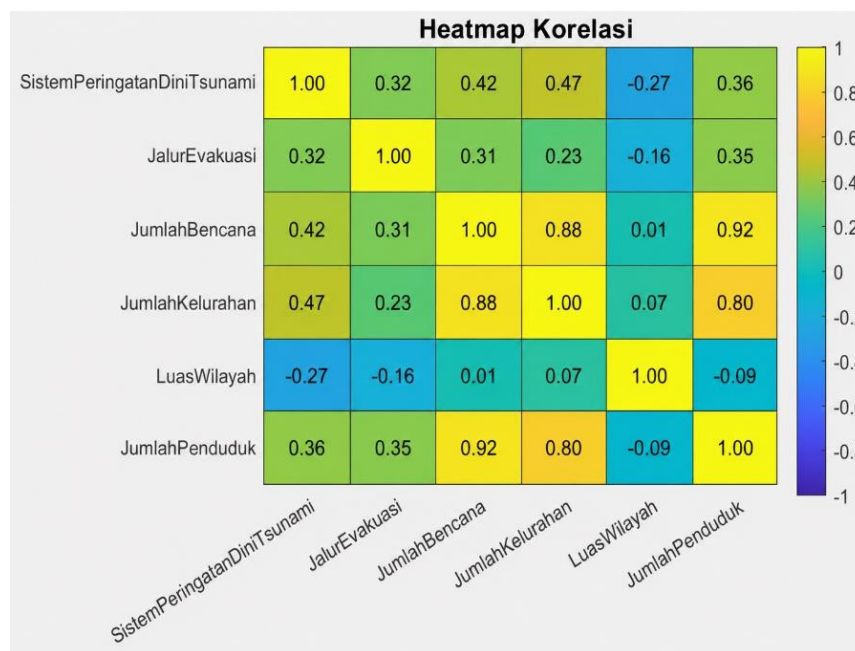


Gambar 1. Flowchart Fuzzy C-Means

4. HASIL DAN PEMBAHASAN

4.1. Analisis Korelasi

Langkah awal dalam penelitian ini dimulai dengan melakukan analisis hubungan korelasi antarvariabel menggunakan visualisasi *heatmap*. Matriks korelasi Pearson dihitung untuk keenam variabel yang mencakup Sistem Peringatan Dini Tsunami, Jalur Evakuasi, Jumlah Bencana, Jumlah Kelurahan, Luas Wilayah, serta Jumlah Penduduk pada seluruh wilayah kajian. Hasil analisis ini memberikan gambaran mengenai sejauh mana keterkaitan linier di antara indikator-indikator kesiapan dan kerentanan bencana tersebut.



Gambar 2. Matriks Korelasi Antarvariabel

Berdasarkan hasil analisis korelasi, terdapat variasi tingkat hubungan antarvariabel penelitian. Hubungan paling kuat teridentifikasi antara variabel Jumlah Bencana dan Jumlah Penduduk dengan koefisien korelasi sebesar 0,92 yang mencerminkan korelasi positif berkategori tinggi. Variabel Jumlah Bencana dan Jumlah Kelurahan menunjukkan hubungan positif yang kuat berdasarkan koefisien korelasi sebesar 0,88. Hubungan yang kuat antara kedua variabel ditunjukkan oleh nilai korelasi yang tinggi.

Variabel Luas Wilayah memperlihatkan hubungan yang relatif lemah dengan sebagian besar variabel lainnya. Korelasi antara Luas Wilayah dan Jumlah Bencana hanya sebesar 0,01, sedangkan korelasinya dengan Jumlah Penduduk bernilai -0,09. Nilai-nilai tersebut menunjukkan bahwa luas wilayah tidak memiliki hubungan yang signifikan terhadap kedua variabel tersebut. Berdasarkan hasil korelasi yang diperoleh, dilakukan seleksi variabel dengan mempertimbangkan tingkat hubungan antarvariabel, yakni hanya variabel dengan nilai korelasi dalam rentang 0,2 hingga 0,8 yang diikutsertakan dalam analisis lanjutan. Hal ini dilakukan untuk menghindari masalah multikolinearitas maupun penggunaan variabel yang memiliki keterkaitan terlalu lemah, sehingga informasi yang digunakan dalam proses klusterisasi tidak bersifat terlalu mirip maupun terlalu lemah.

4.2. Metode Algoritma Fuzzy C-Means (FCM)

Fuzzy C-Means (FCM) diterapkan sebagai metode inti dalam membentuk pengelompokan data pada penelitian ini. Sebelum proses klusterisasi dijalankan, seluruh data terlebih dahulu dinormalisasi dengan metode Min-Max guna menyetarakan skala nilai antarvariabel ke dalam interval [0, 1].

Tabel 3. Pusat Klaster Terakhir

Klaster	Sistem Peringatan Dini Tsunami	Jalur Evakuasi	Jumlah Bencana	Jumlah Kelurahan	Luas Wilayah	Jumlah Penduduk
1	0.1479	0.1567	0.0965	0.1427	0.3047	0.0681
2	0.5068	0.2919	0.5628	0.6780	0.3396	0.4716

Nilai pusat klaster pada Tabel 3 memperlihatkan adanya perbedaan karakteristik yang cukup jelas antara kedua kelompok yang terbentuk.

Klaster 1 secara konsisten mencatat angka pusat yang lebih kecil pada sebagian besar variabel, sedangkan Klaster 2 menunjukkan angka pusat yang secara umum lebih besar. Perbedaan tersebut mengindikasikan bahwa provinsi-provinsi yang masuk ke dalam Klaster 2 mempunyai indikator mitigasi bencana yang lebih baik dibandingkan provinsi-provinsi pada Klaster 1. Kondisi ini menunjukkan bahwa kedua klaster memiliki karakteristik yang tidak sama sehingga proses pengelompokan mampu memisahkan provinsi berdasarkan tingkat kemiripan datanya. Pembagian sampel provinsi berdasarkan klaster ditunjukkan pada Tabel 4.

Tabel 4. Klaster Mitigasi Bencana

Provinsi	Tingkat Mitigasi Rendah	Tingkat Mitigasi Tinggi
Aceh	–	2
Sumatera Utara	–	2
Sumatera Barat	1	–
Riau	1	–
Jambi	1	–
⋮	⋮	⋮
Papua Barat Daya	1	–
Papua	1	–
Papua Selatan	1	–
Papua Tengah	1	–
Papua Pegunungan	1	–

Pembagian provinsi ke dalam masing-masing klaster diperoleh melalui proses klasterisasi *Fuzzy C-Means* yang bekerja dengan menghitung derajat keanggotaan tiap provinsi secara iteratif hingga konvergen. Nilai derajat keanggotaan tersebut mencerminkan seberapa dekat suatu provinsi dengan pusat klaster yang terbentuk. Penempatan provinsi pada klaster tertentu ditentukan berdasarkan nilai keanggotaan tertinggi yang dimilikinya, sehingga setiap provinsi hanya menjadi anggota dari satu kelompok pada hasil akhir klasterisasi. Penentuan Klaster 1 dan Klaster 2 dilakukan berdasarkan hasil perhitungan tersebut dengan mempertimbangkan seluruh variabel penelitian yang digunakan. Hasil data pengelompokan provinsi untuk setiap klaster ditampilkan dalam Tabel 5.

Tabel 5. Hasil Klasterisasi Kesiapsiagaan di Seluruh Provinsi di Indonesia

Kategori	Persentase	Nama Provinsi
Tingkat Kesiapsiagaan Rendah	86.84%	Sumatera Barat, Riau, Jambi, Sumatera Selatan, Bengkulu, Kep. Bangka Belitung, Kep. Riau, DKI Jakarta, Jawa Barat, Jawa Tengah, Jawa Timur, Banten, Nusa Tenggara Barat, Nusa Tenggara Timur, Kalimantan Barat, Kalimantan Tengah, Kalimantan Selatan, Kalimantan Timur, Kalimantan Utara, Sulawesi Utara, Sulawesi Tengah, Sulawesi Selatan, Sulawesi Tenggara, Gorontalo, Sulawesi Barat, Maluku, Maluku Utara, Papua Barat, Papua Barat Daya, Papua, Papua Selatan, Papua Tengah, Papua Pegunungan
Tingkat Kesiapsiagaan Tinggi	13.16%	Aceh, Sumatera Utara, Lampung, DI Yogyakarta, Bali

Hasil pengelompokan memperlihatkan sebagian besar provinsi di Indonesia masuk pada Klaster 1 dengan kategori tingkat mitigasi rendah, yakni sebanyak 33 provinsi atau setara dengan 86,84% dari total keseluruhan. Sementara itu, Klaster 2 dengan kategori tingkat mitigasi tinggi hanya mencakup 5 provinsi atau sebesar 13,16%. Komposisi tersebut menunjukkan bahwa karakteristik mitigasi bencana pada sebagian besar provinsi di Indonesia masih lebih mendekati profil kelompok mitigasi rendah. Provinsi-provinsi yang termasuk dalam klaster mitigasi tinggi cenderung memiliki kondisi yang lebih baik pada indikator-indikator yang dipakai dalam penelitian ini, antara lain keberadaan sistem peringatan dini tsunami, jalur evakuasi, jumlah kejadian bencana, jumlah kelurahan, luas wilayah, dan jumlah penduduk. Hasil pengelompokan ini mengungkapkan adanya kesenjangan yang cukup signifikan dalam hal karakteristik mitigasi bencana antarprovinsi di Indonesia.

4.3. Evaluasi Hasil Klasterisasi (*Silhouette Coefficient*)

Kualitas pengelompokan yang dihasilkan dalam penelitian ini diukur melalui *silhouette coefficient* sebagai indikator ketepatan penempatan objek pada klasternya. Nilai *silhouette coefficient* berada pada rentang -1 hingga $+1$, di mana nilai yang semakin mendekati $+1$ mencerminkan kekohesifan yang tinggi dalam klaster yang sama sekaligus pemisahan yang semakin tegas terhadap klaster lainnya. Nilai yang mendekati 0 mengindikasikan objek di wilayah batas antara dua klaster, sedangkan nilai negatif mengisyaratkan adanya ketidaktepatan dalam penempatan objek tersebut.

Dari hasil perhitungan, rata-rata *silhouette coefficient* yang diperoleh sebesar $0,74396$. Berpedoman pada kriteria evaluasi Kaufman dan Rousseeuw, angka tersebut masuk ke dalam kategori struktur klaster yang sangat kuat. Setiap anggota dalam satu klaster mempunyai kemiripan yang besar antara satu dengan lainnya, sekaligus perbedaan yang cukup signifikan terhadap anggota di klaster yang berbeda. Hal ini menunjukkan bahwa hasil klasterisasi yang diperoleh berkualitas baik dan mampu merepresentasikan perbedaan karakteristik mitigasi bencana antarprovinsi di Indonesia secara akurat.

4.4. Visualisasi Hasil Klaster

Hasil klaster yang diperoleh kemudian divisualisasikan melalui peta pada Gambar 3.

Hasil pemetaan memperlihatkan bahwa mayoritas wilayah Indonesia masih berada dalam kategori tingkat mitigasi rendah, dengan 33 provinsi (86,84%) ditandai dengan warna biru. Lima provinsi (13,16%), yaitu Aceh, Sumatera Utara, Lampung, DI Yogyakarta, dan Bali, masuk ke dalam kategori mitigasi tinggi dan ditandai dengan warna merah. Persebaran tersebut mencerminkan adanya ketimpangan kapasitas penanggulangan bencana antarwilayah yang dipengaruhi oleh perbedaan kapasitas kelembagaan daerah, keterbatasan sumber daya manusia terlatih di bidang kebencanaan, serta belum meratanya sistem peringatan dini di berbagai wilayah. Provinsi-provinsi yang termasuk kategori mitigasi tinggi pada umumnya telah memiliki infrastruktur kebencanaan yang lebih memadai, mencakup sistem koordinasi tanggap darurat, jalur evakuasi, serta regulasi penanggulangan bencana yang lebih baik.



Gambar 3. Pemetaan Tingkat Mitigasi Bencana di Indonesia

Tingginya nilai mitigasi tidak selalu menjamin respons bencana yang optimal mengingat masih terdapat banyak faktor operasional yang perlu terus ditingkatkan. Ketidakseimbangan antara kebutuhan dan kapasitas mitigasi ini menjadi tantangan besar dalam mewujudkan perlindungan masyarakat yang merata di seluruh wilayah Indonesia, sehingga diperlukan intervensi kebijakan yang lebih terfokus khususnya bagi provinsi-provinsi yang masih berada dalam kategori mitigasi rendah. Penelitian selanjutnya disarankan menggunakan indikator yang lebih beragam dan sesuai dengan kondisi aktual di setiap wilayah, serta melakukan perbandingan metode klusterisasi *Fuzzy C-Means* dengan metode klusterisasi lainnya guna memperoleh hasil pengelompokan yang lebih optimal.

Berdasarkan peta Indeks Risiko Bencana Indonesia (IRBI) tahun 2024, beberapa wilayah seperti Papua, Maluku, dan Sulawesi masih memiliki tingkat risiko bencana yang relatif tinggi [29]. Hasil klusterisasi penelitian ini menunjukkan bahwa sebagian besar provinsi di wilayah tersebut masih termasuk dalam kelompok mitigasi rendah. Provinsi lainnya seperti Aceh, Sumatera Utara, DI Yogyakarta, Lampung, dan Bali tergolong dalam kelompok mitigasi tinggi yang mengindikasikan adanya kapasitas kesiapsiagaan dan infrastruktur mitigasi yang lebih baik. Temuan ini menunjukkan bahwa kapasitas mitigasi bencana antarprovinsi di Indonesia masih belum merata dan perlu terus ditingkatkan, terutama pada wilayah dengan tingkat risiko bencana yang tinggi.

5. KESIMPULAN

Penerapan metode *Fuzzy C-Means* (FCM) dalam penelitian ini terbukti efektif untuk mengklasifikasikan 38 provinsi di Indonesia ke dalam dua kelompok berdasarkan kondisi mitigasi bencana dan karakteristik wilayahnya. Sebanyak 5 provinsi (13,16%), yakni Aceh, Sumatera Utara, Lampung, DI Yogyakarta, dan Bali, teridentifikasi memiliki kapasitas mitigasi yang lebih tinggi, sementara 33 provinsi sisanya (86,84%) masih tergolong dalam kelompok dengan kapasitas mitigasi yang perlu ditingkatkan. Nilai *silhouette coefficient* sebesar 0,74396 mengonfirmasi bahwa struktur kluster yang dihasilkan sangat kuat dan representatif, sehingga temuan ini layak dijadikan landasan bagi pemangku kebijakan dalam menyusun strategi prioritas penguatan sistem penanggulangan bencana di seluruh wilayah Indonesia.

DAFTAR PUSTAKA

- [1] Y. Subiakto and B. Kustiawan, "Manajemen Mitigasi Bencana pada Peserta Didik untuk Mengurangi Risiko Bencana Gempa Bumi," vol. 6, no. 2, pp. 609–615, 2022, doi: 10.33369/pendipa.6.2.609-615.
- [2] D. Trimaily and A. A. Said, *Mitigasi Bencana Penyelamat Jiwa Manusia*, 1st ed. Tasikmalaya: Perkumpulan Rumah Cemerlang Indonesia, 2024.
- [3] Suniah and N. Rochayati, "Interaksi Tiga Lempeng di Kawasan Indonesia Timur," *Educ. Soc. Sci. Linguist. Conf. Ser.*, vol. 1, no. 2, pp. 127–132, 2025.
- [4] V. A. Zulfa, H. Widyasamratri, and J. Kautsary, "Bencana Tanah Longsor Studi Kasus: Lereng Gunung Wilis Kabupaten Nganjuk, Desa Sendangrejo Kecamatan Sambeng Kabupaten Lamongan dan Desa Sriharjo Kecamatan Imogiri Kabupaten Bantul," vol. 1, no. 2, pp. 154–169, 2022.
- [5] D. D. Utomo, F. Yul, and D. Marta, "Dampak Bencana Alam terhadap Perekonomian Masyarakat di Kabupaten Tanah Datar," vol. 2, no. 1, pp. 92–97, 2022.
- [6] Badan Pusat Statistik, "Banyaknya Desa/Kelurahan Menurut Upaya Antisipasi/Mitigasi Bencana Alam, 2024," 2024. [Online]. Available: <https://www.bps.go.id/id/statistics-table/2/OTcxIzI=/banyaknya-desa-kelurahan-menurut-upaya-antisipasi-mitigasi-bencana-alam--desa-.html>
- [7] F. Khoirunnisa and Y. Rahmawati, "Komparasi 2 Metode Cluster dalam Pengelompokan Intensitas Bencana Alam di Indonesia," vol. 12, no. 1, pp. 68–79, 2024.
- [8] A. Baldah, A. V. Duarisah, and R. A. Maulana, "Clustering Daerah Rawan Bencana Alam di Indonesia Berdasarkan Provinsi dengan Metode K-Means," vol. 14, pp. 31–36, 2023.

- [9] W. T. Oktaviany, "Pengelompokan Wilayah Bencana Banjir di Indonesia Menggunakan Algoritma K-Means," *Bull. Comput. Sci. Res.*, vol. 5, no. 4, pp. 542–552, 2025.
- [10] G. F. Anwar, U. Khaira, and P. E. P. U. Utomo, "Clustering Wilayah di Indonesia Berdasarkan Kualitas Pendidikan Menggunakan Algoritma Fuzzy C-Means," vol. 8, no. 2, pp. 167–178, 2025, doi: 10.24076/infosjournal.2025v8i02.2442.
- [11] F. Novianti, Y. R. A. Yasmin, and D. C. R. Novitasari, "Penerapan Algoritma Fuzzy C-Means (FCM) dalam Pengelompokan Provinsi di Indonesia berdasarkan Indikator Penyakit Menular Manusia," vol. 6, no. 1, pp. 23–33, 2022.
- [12] S. Huddin, "Penerapan Fuzzy C-Means pada Klasterisasi Penerima Bantuan Pangan Non Tunai," vol. 4, no. 1, pp. 453–461, 2023, doi: 10.30865/klik.v4i1.988.
- [13] S. Nurdianti and D. N. Yunianti, "Penentuan Banyak Kluster Optimal Hasil Fuzzy C-Means dengan Metode Elbow pada Klasifikasi Kecemasan Belajar Siswa," vol. 11, no. 3, pp. 401–413, 2023, doi: 10.26740/mathunesa.v11n3.p401-413.
- [14] T. Yulianto, A. F. Rahmah, Faisol, and R. Amalia, "Clustering Daerah Bencana Alam di Indonesia Menggunakan Metode Fuzzy C-Means," *Unisda J. Math. Comput. Sci.*, vol. 9, no. 2, pp. 29–39, 2023, doi: 10.52166/ujmc.v9i2.
- [15] H. Firdaus and A. S. Sofro, "Analisa Cluster Menggunakan K-Means dan Fuzzy C-Means dalam Pengelompokan Provinsi Menurut Data Intensitas Bencana Alam di Indonesia Tahun 2017–2021," vol. 10, no. 1, pp. 50–60, 2022, doi: 10.26740/mathunesa.v10n1.p50-60.
- [16] G. Ramadhan, A. S. Gutama, and M. Fedryansyah, "Tahap Pra Pengorganisasian Komunitas Lokal dalam Mitigasi Bencana di Desa Karangjaladri, Kecamatan Parigi, Kabupaten Pangandaran," *Sosio-Didaktika Soc. Sci. Educ. J.*, vol. 11, no. 2, pp. 149–160, 2024, doi: 10.15408/sd.v11i2.39847.
- [17] A. R. Susmita, E. Darmawi, and E. W. Suri, "Analisis Pendekatan Mitigasi Bencana Alam di Badan Nasional Penanggulangan Bencana Kota Bengkulu," *Masy. Demokr. J. Ilm. Adm. Publik*, vol. 1, no. 2, pp. 9–18, 2023, doi: 10.32663/md.v1i2.3812.
- [18] N. L. V. Anggaini, A. T. Widiyati, R. Adiono, F. Amalia, and N. N. Islami, *Mitigasi Bencana dan Emergency Management Arsip pada Organisasi*, 1st ed. Tulungagung, 2023.
- [19] T. S. Prayoga and Suliadi, "Korelasi Rank-Spearman pada Hubungan Beberapa Variabel Produk Domestik Regional Bruto," *J. Ris. Stat.*, vol. 4, no. 2, pp. 137–144, 2024, doi: 10.29313/jrs.v4i2.5162.
- [20] P. G. Subhaktiyasa, I. W. Numertayasa, N. P. Sumaryani, and S. A. Ketut, "Uji Korelasi dalam Penelitian Kuantitatif: Kajian Konseptual, Asumsi Statistik dan Implikasi Praktis," vol. 10, pp. 3297–3308, 2025, doi: 10.29303/jipp.v10i4.3952.
- [21] K. A. K-nn and R. G. Whendasmoro, "Analisis Penerapan Normalisasi Data dengan Menggunakan Z-Score pada," vol. 9, no. 4, pp. 872–876, 2022, doi: 10.30865/jurikom.v9i4.4526.
- [22] R. A. Eryadi, A. Maulana, S. Padilaturmah, and Z. Z. Shibghatullah, "Penerapan Normalisasi dan Denormalisasi untuk Optimasi Performa Database," *Kohesi J. Sains dan Teknol.*, vol. 10, no. 8, 2026, doi: 10.8734/Kohesi.v1i2.365.
- [23] P. P. Alloserung, A. Erna, M. Bagussahrir, and S. Alam, "Analisis Performa Normalisasi Data untuk Klasifikasi K-Nearest Neighbor pada Dataset Penyakit," *J. Inform. Sunan Kalijaga*, vol. 9, no. 3, pp. 178–191, 2024, doi: 10.14421/jjska.2024.9.3.178-191.
- [24] I. N. Audy, T. N. Padilah, and B. N. Sari, "Pengelompokan Daerah Rawan Bencana Alam di Jawa Barat Menggunakan Algoritma Fuzzy C-Means," vol. 7, no. 4, pp. 2799–2803, 2023, doi: 10.36040/jati.v7i4.7205.
- [25] A. R. Maulana, N. Ulinuha, and D. C. R. Novitasari, "Penerapan Fuzzy C-Means untuk Pengelompokan Tingkat Kualitas Pendidikan di Jawa Timur," *J. SIMETRIS*, vol. 14, no. 2, pp. 419–425, 2023, doi: 10.24176/simet.v14i2.9442.
- [26] Z. Sidik and A. R. Ismail, "Perbandingan Teknik Normalisasi dan Denormalisasi dalam Pengelolaan Data Skala Besar," *J. Tecnoscienza*, vol. 10, no. 1, pp. 42–53, 2025, doi: 10.51158/pqfg7z82.
- [27] A. Rizal, D. Candra, R. Novitasari, and M. Hafiyusholeh, "Pengelompokan Karyawan Berdasarkan Kesalehan Menggunakan Perbandingan Fuzzy C-Means, K-Means, dan Probabilistic Distance Clustering," vol. 11, no. 2, pp. 69–77, 2022, doi: 10.14421/fourier.2022.112.69-77.
- [28] C. Ifadah, C. D. Ratnasri, and D. C. R. Novitasari, "Clustering Daerah Banjir di Jawa Timur dengan Algoritma Fuzzy C-Means," *Din. Inform.*, vol. 14, no. 2, pp. 99–104, 2022, doi: 10.35315/informatika.v14i2.8885.
- [29] AgroIndonesia, "Cegah Bencana Hidrometeorologi dengan UU Daerah Aliran Sungai." [Online]. Available: <https://agroindonesia.co.id/cegah-bencana-hidrometeorologi-dengan-uu-daerah-aliran-sungai/>