Implementasi Algoritma *K-Means Clustering* Untuk Tingkat Rawan Bencana Di Wilayah Kabupaten Bandung

Tiara Nurhaliza Anggraeni[1], Ignatius Wiseto Prasetyo Agung[2], Rizal Rachman[3]
[1,3]Program Studi Sistem Informasi, Universitas Adhirajasa Reswara Sanjaya
[2]Program Studi Teknik Informatika,, Universitas Adhirajasa Reswara Sanjaya
e-mail: [1]itskraa@gmail.com, [2]wiseto.agung@ars.ac.id, [3]rizalrachman@ars.ac.id

Abstrak

Bencana alam adalah kejadian yang secara signifikan mempengaruhi populasi manusia. Tanah longsor, gempa bumi, banjir, kebakaran, kekeringan dan bencana alam lainnya sering terjadi di Provinsi Jawa Barat. Salah satu wilayah di Jawa Barat dengan risiko bencana tertinggi adalah Kabupaten Bandung. Saat ini, terdapat beberapa metode pengembangan untuk menganalisa data bencana alam termasuk menggunakan keahlian teknologi informasi. *Data Mining* adalah metode populer untuk menganalisis data bencana. Penelitian ini bertujuan ntuk mengelompokan wilayah yang rentan terhadap bencana alam di 31 kecamatan di Kabupaten Bandung ke dalam dua kelompok, yaitu kelompok berisiko rendah dan kelompok berisiko tinggi, berdasarkan tingkat kerentanan bencananya. Perhitungan *K-Means Clustering* secara manual serta menggunakan aplikasi RapidMiner menghasilkan kelompok rentan bencana rendah (C1) yang terdiri dari 28 kecamatan, sedangkan kelompok rentan bencana tinggi (C2) terdiri dari 3 kecamatan. Nilai *DBI* (K=2) sebesar 0,702 adalah nilai yang optimal untuk pembentukan kelompok dari K=2 hingga K=5. Hal ini menunjukkan bahwa pembentukan dua kelompok cukup baik karena nilai *DBI* mendekati nol. Diharapkan temuan penelitian ini dapat membantu pemerintah dalam menetapkan prioritas pengelolaan dan mitigasi bencana. Untuk mengurangi risiko dan biaya di masa depan.
Kata kunci—*K-Means*, *Clustering*, *Data Mining*, *RapidMiner*, Bencana

Abstract

Natural disasters are occurrences that greatly impact human communities. In West Java Province, events such as landslides, earthquakes, floods, fires, droughts, and various other natural calamities frequently take place. A region in West Java that faces a significant level of disaster threat is Bandung Regency. At present, there are various approaches for examining data related to natural disasters, which involve the application of skills in information technology. One common technique for analyzing disaster data is data mining. The purpose of this study is to classify 31 subdistricts in Bandung Regency that are susceptible to natural disasters into low-risk and high-risk groups according to the degree of disaster susceptibility. A low disaster vulnerability group (C1) with 28 sub-districts and a high disaster vulnerability group (C2) with 3 sub-districts were produced via manual K-Means clustering calculations and the RapidMiner application. The ideal DBI value for group formation between K=2 and K=5 is 0.702 (K=2). The DBI score is around zero, which suggests that the establishment of two groups is quite good. In order to lower future risks and expenses, it is believed that the results of this study would help the government establish priorities for disaster management and mitigation.
Keywords— *K-Means*, *Clustering*, *Data Mining*, *RapidMiner*, *Disaster*

Corresponding Author:
Ignatius Wiseto Prasetyo Agung,
Email: wiseto.agung@ars.ac.id

1. PENDAHULUAN

Karena berbagai karakteristik geografisnya, seperti kondisi tektonik, meteorologi, dan klimatologi, Kepulauan Indonesia rentan terhadap bencana alam [1]. Lokasi geografis dan geologis Indonesia meningkatkan risiko bencana, menurut data dari Badan Nasional Penanggulangan Bencana (BNPB). Lempeng Eurasia, Indo-Australia, Filipina, dan Pasifik adalah empat lempeng tektonik utama yang secara geografis berada di antara wilayah Indonesia. Karena letaknya yang berada di daerah tropis, yang diapit oleh dua benua dan dua samudra, Indonesia rentan terhadap sejumlah bencana alam, termasuk kekeringan, banjir, tanah longsor, banjir bandang, cuaca ekstrem, gelombang pantai, dan erosi [2]. Forest and land fires (2,051 incidents), extreme weather (1,261 incidents), floods (1,255 incidents), landslides (591 incidents), droughts (174 incidents), earthquakes (31 incidents), tidal waves and coastal erosion (33 incidents), and volcanic eruptions (4 incidents) were among the 5,400 disaster incidents reported in 2023. As of December 31, 2023, the Indonesian Disaster Data Geoportal reported that 8,491,288 people had been affected and displaced, 275 had died, 33 were missing, and 5,795 had been injured [2].

Peristiwa alam yang berdampak signifikan terhadap penduduk dikenal sebagai bencana alam. Provinsi Jawa Barat sering dilanda tanah longsor, gempa bumi, banjir, kebakaran, kekeringan, dan bencana alam lainnya [3]. Salah satu wilayah di Jawa Barat dengan risiko bencana tertinggi adalah Kabupaten Bandung Kabupaten Bandung merupakan salah satu wilayah di Jawa Barat yang paling rentan terhadap bencana alam [4]. Bencana yang sering terjadi di wilayah ini meliputi banjir, tanah longsor, dan angin puting beliung, yang disebabkan oleh berbagai faktor baik alamiah maupun non-alamiah serta kondisi geografis Kabupaten Bandung, yang merupakan dataran tinggi dengan banyak alih fungsi lahan, berkontribusi terhadap tingginya risiko bencana. Misalnya, curah hujan yang tinggi dan perilaku masyarakat seperti penanganan lahan yang tidak tepat dapat memperburuk situasi[5]. Data dari BPBD Kabupaten Bandung menunjukkan bahwa pada tahun 2021 saja, terjadi 111 kejadian banjir yang berdampak pada lebih dari 214.000 jiwa, serta 92 kejadian longsor [6]. Saat ini, terdapat beberapa metode pengembangan untuk menganalisa data bencana alam termasuk menggunakan keahlian teknologi informasi. Data bencana alam tersebut dapat dianalisa menggunakan beberapa metode dibawah ini.

Saat ini, teknologi dan informasi berkembang dengan pesat. Berkat teknologi modern, setiap orang kini memiliki akses tak terbatas terhadap informasi. Pengetahuan sangat penting dalam segala aspek kehidupan [3] Informasi mengenai bencana alam adalah salah satu contohnya, karena pengelolaan bencana memerlukan data semacam ini [3]. Karena *Data Mining* dianggap sebagai solusi potensial untuk mengatasi tantangan yang terkait dengan pengelolaan data bencana, teknik ini banyak digunakan untuk menganalisis data bencana [7].

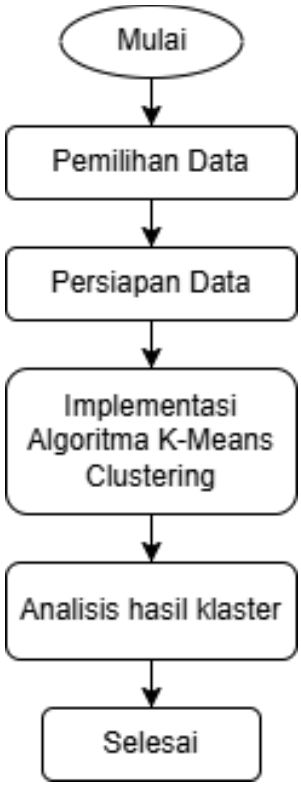
Data Mining adalah proses pengambilan informasi berharga dari basis data untuk keperluan tertentu. Pengelompokan (clustering) adalah salah satu pendekatan dalam penambangan data yang bertujuan untuk mencari pola, titik, objek, atau dokumen yang dapat dikelompokkan ke dalam klaster [8]. Algoritma Pengelompokan *K-Means* memainkan peran penting dalam *Data Mining* karena kesederhanaan dan kemudahan implementasinya[9]. Sebuah teknik analisis data yang disebut algoritma K-Means membagi data menjadi beberapa kelompok berdasarkan tingkat kemiripannya, dengan terlebih dahulu menentukan nilai k [10].

Penelitian ini bertujuan untuk mengelompokan daerah rawan bencana alam yang ada di Kabupaten Bandung menjadi 2 klaster yaitu klaster rendah dan tinggi. Mengelompokkan wilayah Kabupaten Bandung berdasarkan tingkat kerawanan bencana menggunakan algoritma *K-Means Clustering* untuk membantu pemerintah dalam menetapkan prioritas mitigasi dan penanggulangan bencana. Untuk mengurangi risiko dan biaya di masa mendatang, lokasi-lokasi yang rawan bencana harus diidentifikasi secara akurat.

2. METODE PENELITIAN

2.1. Tahapan Penelitian

Prosedur yang digunakan dalam penelitian ini ditunjukkan pada Gambar 1.



Gambar 1. Alur Penelitian

Berikut adalah alur penelitian yang dibagi menjadi beberapa tahap sebagai berikut:

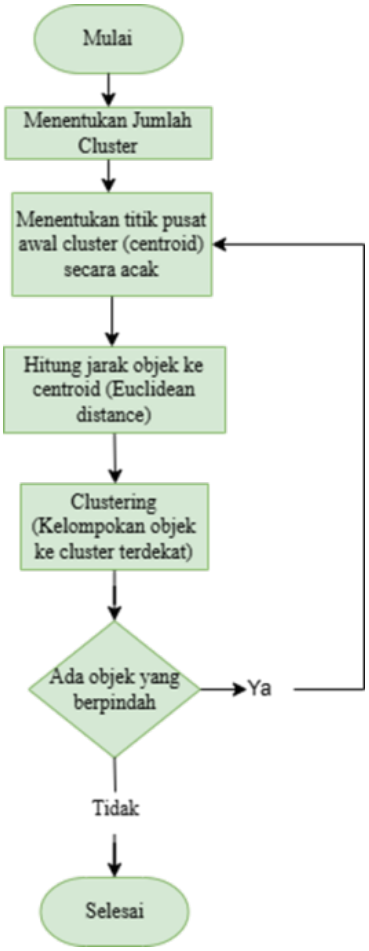
1. Pemilihan Data: Situs web resmi Pemerintah Kabupaten Bandung menyediakan kumpulan data mengenai bencana alam di 31 kecamatan untuk periode tahun 2014–2024 [11].
2. Persiapan Data: Proses pengelompokan hanya menggunakan data numerik selama tahap pemilihan data. Data mengenai bencana alam dikumpulkan dari 31 kecamatan di Kabupaten Bandung. Tabel 1 di bawah ini mencantumkan tujuh variabel yang dipertimbangkan dalam analisis ini untuk periode 2014 hingga 2024: kecamatan, banjir, badai, tanah longsor, gempa bumi, kebakaran, dan kekeringan.

Tabel 1. Dataset Bencana

No	Kecamatan	Banjir	Badai	Longsor	Gempa	Kebakaran	Kekeringan
1	Arjasari	7	11	25	3	17	35
2	Baleendah	106	18	12	1	52	42
3	Banjaran	26	15	23	3	10	25
4	Bojongsoang	66	17	1	1	8	2
5	Cangkuang	7	8	13	1	10	25
6	Cicalengka	22	15	42	2	44	5

No	Kecamatan	Banjir	Badai	Longsor	Gempa	Kebakaran	Kekeringan
7	Cikancung	5	10	5	1	26	25
8	Cilengkrang	3	2	13	1	11	1
9	Cileunyi	31	9	13	1	22	1
10	Cimaung	3	8	22	2	14	4
11	Cimenyan	2	9	24	1	18	10
12	Ciparay	19	19	9	1	24	43
13	Ciwidey	8	5	47	1	21	10
14	Dayeuhkolot	113	6	2	3	9	3
15	Ibun	5	3	40	1	11	9
16	Katapang	17	9	1	1	19	1
17	Kertasari	11	3	25	6	4	13
18	Kutawaringin	5	8	56	1	39	11
19	Majalaya	27	7	8	1	15	10
20	Margaasih	9	4	3	1	14	1
21	Margahayu	4	1	1	1	8	1
22	Nagreg	6	3	8	1	16	15
23	Pacet	3	5	57	2	8	15
24	Pameungpeuk	13	12	5	2	15	22
25	Pangalengan	10	13	65	8	17	3
26	Paseh	8	5	13	1	17	5
27	Pasirjambu	4	6	50	1	21	18
28	Rancabali	1	4	32	1	10	3
29	Rancaekek	42	21	3	1	10	36
30	Solokanjeruk	12	5	2	1	6	1
31	Soreang	17	6	42	1	36	12

3. Algoritma Pengelompokan K-Means: Algoritma adalah serangkaian instruksi sistematis untuk memecahkan suatu masalah. Matematika, ilmu komputer, dan kehidupan sehari-hari hanyalah beberapa bidang di mana algoritma dapat diterapkan. Secara umum, algoritma terdiri dari tahapan-tahapan yang logis dan terstruktur yang dirancang untuk mencapai tujuan tertentu [12]. Gambaran alir pengelompokan menggunakan algoritma *K-Means* ditampilkan pada Gambar 2.



Gambar 2. Alir *K-Means Clustering*

Penjelasan berikut ini akan menguraikan langkah-langkah metode K-Means [13]:

a. Ditentukannya jumlah klaster (*k*)
b. Tentukan titik pusat (*Ci*) secara acak
c. Gunakan rumus Jarak *Euclid* untuk menentukan jarak terdekat dari setiap data. Dibawah ini merupakan persamaan Jarak *Euclidean*:

$$d_{Euclidean}(x, c)\sqrt{\sum_i (x_i - c_i)}^{\,2} \quad (1)$$

Jarak yang dihitung dengan mengurangkan X_i (nilai data) dan C_i (centroid) merupakan hasil perhitungan Jarak.

d. Mengelompokkan semua data berdasarkan jarak terpendek atau pusat kelompok terdekat. Untuk memperbarui nilai pusat klaster, gunakan rumus berikut untuk menghitung nilai rata-rata setiap klaster:

$$ClusterCenter = \sum \frac{x_i}{n} \quad (2)$$

Keterangan:

x_i = titik data ke-i dalam klaster

n = jumlah titik data dalam klaster tersebut

e. Ulangi langkah 3–5 hingga anggota setiap kelompok tetap sama. Setelah proses selesai, pusat-pusat kelompok yang ditemukan pada iterasi terakhir akan digunakan sebagai parameter untuk menentukan cara pengelompokan data.

4. Analisis hasil pengelompokan: Pada tahap ini, Indeks Davies-Bouldin (DBI) digunakan sebagai metode penilaian untuk menganalisis hasil pengelompokan. Hal ini membantu menentukan apakah hasil tersebut dapat digunakan serta mengevaluasi kualitas kelompok-kelompok yang terbentuk [14]. Klaster yang lebih baik ditandai dengan nilai DBI yang mendekati nol. Hal ini menunjukkan bahwa peningkatan validitas klaster berkorelasi dengan skor DBI yang lebih rendah. [15].

2.2. Metode Analisa Data

Dengan menerapkan teknik *Data Mining*, khususnya algoritma pengelompokan *K-means*, yang terkenal karena keefektifannya dalam mengklasifikasikan data berdasarkan atribut tertent [8]. Pada peneilitian ini juga digunakan aplikasi RapidMiner untuk memvisualisasikan pengolahan data dengan cara yang intuitif. Dengan menggunakan Indeks *Davies-Bouldin* (DBI), yang mengevaluasi kualitas kluster dan membantu dalam menentukan apakah temuan penambangan data dapat digunakan. Nilai *DBI* yang mendekati 0 (non-negatif) menunjukkan klaster yang unggul [16]. Diharapkan bahwa metode yang digunakan pada penelitian ini dapat memberikan pemetaan risiko yang lebih akurat dan bermanfaat untuk menangani bencana di wilayah tersebut.

2.3. RapidMiner

RapidMiner adalah alat perangkat lunak sumber terbuka berlisensi di bawah Lisensi Publik GNU, yang dikembangkan oleh Dr. Markus Hofmann dan Ralf Klinkenberg. Program berbasis Java ini dirancang khusus untuk pemrosesan dan penambangan data. RapidMiner memanfaatkan basis data, kecerdasan buatan, dan statistik untuk mengidentifikasi pola dalam kumpulan data berskala besar. Program ini mendukung berbagai metode analisis data, seperti klasifikasi, pengelompokan, analisis asosiasi, model Bayesian, pohon keputusan, dan jaringan saraf tiruan, serta dilengkapi dengan antarmuka pengguna grafis (GUI) yang intuitif [17][18].

3. HASIL DAN PEMBAHASAN

3.1. K-Means Clustering

1. Menentukan titik pusat (Ci) secara acak
   Karena K-Means tidak memiliki pengetahuan awal mengenai struktur data atau lokasi kluster yang ideal, diketahui bahwa pemilihan ini dilakukan secara acak [17]. Selanjutnya, klasifikasi tingkat bawah (C1) diinisialisasi menggunakan data kecamatan Cikancung, sedangkan klasifikasi tingkat atas (C2) diinisialisasi menggunakan data kecamatan Baleendah. Titik pusat awal dapat dilihat pada Tabel 2 di bawah ini:

Tabel 2. Iterasi Centroid Awal

Centroid	Banjir	Badai	Longsor	Gempa	Kebakaran	Kekeringan
C1	5	10	5	1	26	25
C2	106	18	12	1	52	42

2. Menggunakan rumus Jarak Euclid untuk menentukan jarak terdekat dari setiap data. Berikut adalah persamaan Jarak Euclidean:

$$d_{Euclidean}(x, c)\sqrt{\sum_i (x_i - c_i)}^{\,2} \quad (1)$$

Contoh perhitungan di bawah ini adalah data untuk kecamatan Arjasari yang tercantum dalam tabel 3:

Tabel 3. Data Kecamatan Arjasari

Kecamatan	Banjir	Badai	Longsor	Gempa	Kebakaran	Kekeringan
Arjasari	7	11	25	3	17	35

Maka:

Data 1,1 = $\sqrt{(7-5)^2 + (11-10)^2 + (25-5)^2 + (3-1)^2 + (17-26)^2 + (35-25)^2}$

Data 1,1 = 24,2899156

Data 1,2 = $\sqrt{(7-106)^2 + (11-18)^2 + (25-13)^2 + (3-1)^2 + (17-52)^2 + (35-42)^2}$

Data 1,2 = 106,2873464

Data 1,1 adalah data pertama dengan pusat massa pertama, dan data 1,2 adalah data pertama dengan pusat massa kedua. Lakukan hal ini hingga semua data dihitung[18]. Berikut adalah hasil perhitungan Tabel 4 pada iterasi 1:

Tabel 4. Jarak Centroid Iterasi 1

Kecamatan	Centroid 1	Centroid 2
Arjasari	24,2899156	106,2873464
Baleendah	106,2026365	0
Banjaran	32,40370349	92,66606714
Bojongsoang	68,11020482	72,51206796
Cangkuang	18,11077028	109,3389226
Cicalengka	49,07137659	96,9484399
Cikancung	0	106,2026365
Cilengkrang	30,5450487	119,2811804
Cileunyi	36,51027253	91,03845341
Cimaung	29,71531592	117,0384552
Cimenyan	25,69046516	114,9826074
Ciparay	24,91987159	91,45490692
Ciwidey	45,254834	113,9429682
Dayeuhkolot	111,6512427	60,55575943
Ibun	41,89272013	118,23705
Katapang	28,03569154	104,3695358
Kertasari	33,73425559	112,200713
Kutawaringin	54,49770637	115,6157429

Kecamatan	Centroid 1	Centroid 2
Majalaya	29,12043956	93,65361712
Margaasih	27,85677655	113,1856881
Margahayu	31,591138	120,1290972
Nagreg	16,09347694	110,7519752
Pacet	56,19608527	124,3744347
Pameungpeuk	14,10673598	102,4890238
Pangalengan	65,17668295	121,8400591
Paseh	24,06241883	111,2115102
Pasirjambu	46	116,3142296
Rancabali	39	122,0901306
Rancaekek	43,25505751	77,36924454
Solokanjeruk	32,54228019	113,5869711
Soreang	42,40283009	100,6031809

Dapat dilihat bahwa hasil data 1,1, jarak data-1 ke pusat klaster-1, memiliki nilai terkecil sebesar 24.2899156 dibandingkan dengan hasil data 1,2. Oleh karena itu, data-1 termasuk dalam kategori klaster 1. Pengelompokan dilakukan pada semua dataset.

3. Cari nilai centroid terbaru. Nilai ini diperoleh dari rata-rata setiap klaster menggunakan persamaan berikut:

$$ClusterCenter = \sum \frac{x_i}{n} \quad (2)$$

Contoh:

Data yang termasuk dalam kategori klaster 1 adalah 29 dari 31 data, dan jumlah total nilai data x_i yang termasuk dalam klaster 1 adalah 393. Oleh karena itu, temukan pusat klaster C_i terbaru yang termasuk dalam klaster 1 dengan menghitung rata-ratanya.

$$Cx_1 = \frac{393}{29} = 13{,}55172414$$

Hasil data pusat baru dari Iterasi 1 ditampilkan dalam Tabel 5 di bawah ini:

Tabel 5. Centroid Baru Iterasi 2

Centroid	Banjir	Badai	Longsor	Gempa	Kebakaran	Kekeringan
C1	13,55172414	8,379310345	22,34482759	1,689655172	16,93103448	12,48275862
C2	109,5	12	7	2	30,5	22,5

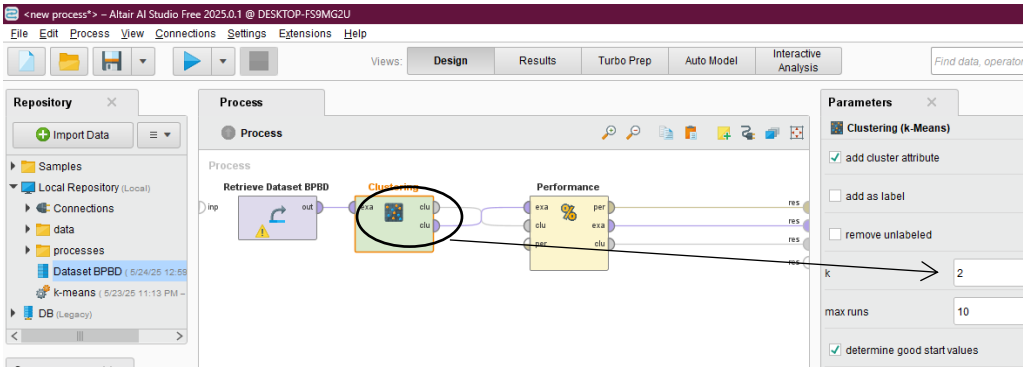
4. Ulangi langkah 3-5 hingga anggota setiap kelompok tetap tidak berubah
   Hasil perhitungan iterasi 3 tidak lagi memiliki anggota klaster yang berpindah dan dianggap stabil jika hasil iterasi baru sama dengan hasil iterasi sebelumnya[13]. Klaster 1 masih memiliki 28 anggota yang sama seperti perhitungan iterasi 2, dan ketiga anggota klaster 2 tetap sama seperti pada iterasi 2. Oleh karena itu, proses algoritma *K-Means* dihentikan pada iterasi 3. Hasil pengelompokan berikut ditampilkan dalam Tabel 6 di bawah ini:

Tabel 6. Hasil *Clustering*

Klaster	Jumlah Anggota
C1	28
C2	3

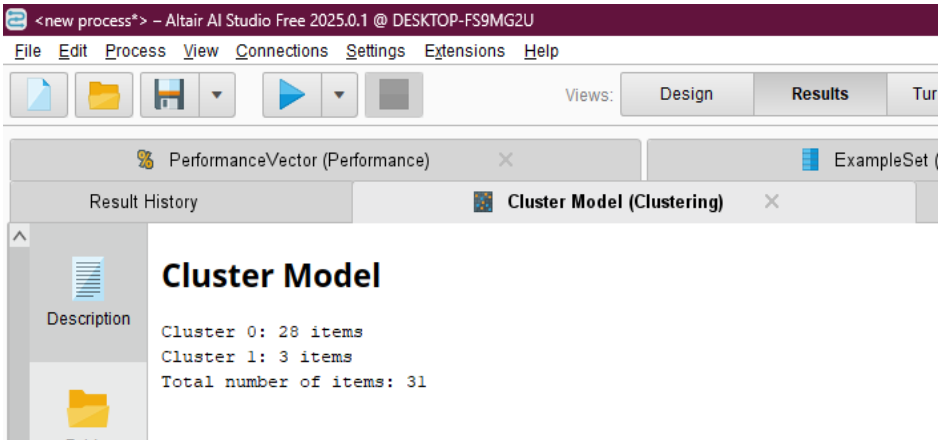
3.2. K-Means Clustering Menggunakan RapidMiner

1. Pada kolom pencarian operator, ketik *K-Means* dan kinerja jarak sebagai operator hasil pengelompokan, lalu seret ke lembar kerja dan hubungkan operator dataset dengan operator lain. Ubah parameter *K-Means* menjadi 2 klaster, lalu klik Jalankan seperti yang ditunjukkan pada Gambar 3 di bawah ini:



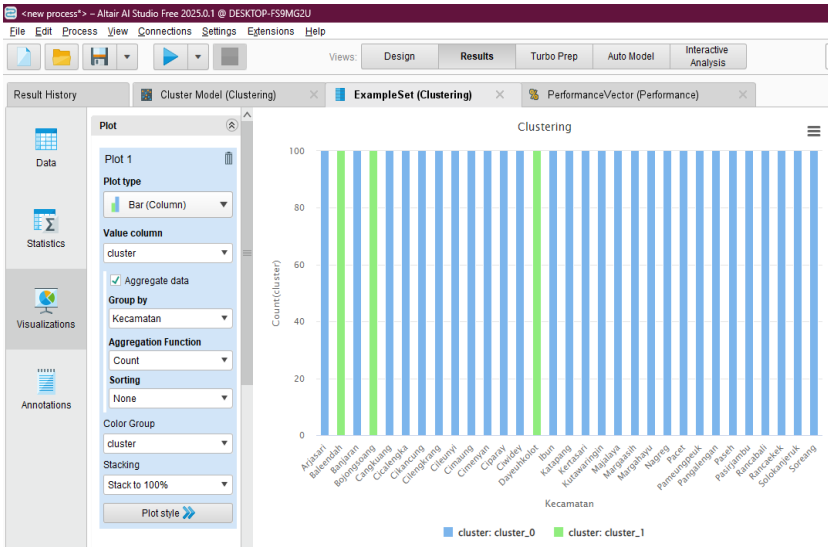
Gambar 3. Pilih Operator *K-Means* dan *Distance Performance*

2. Hasil dari Algoritma Pengelompokan *K-Means* pada Gambar 4 di bawah ini:



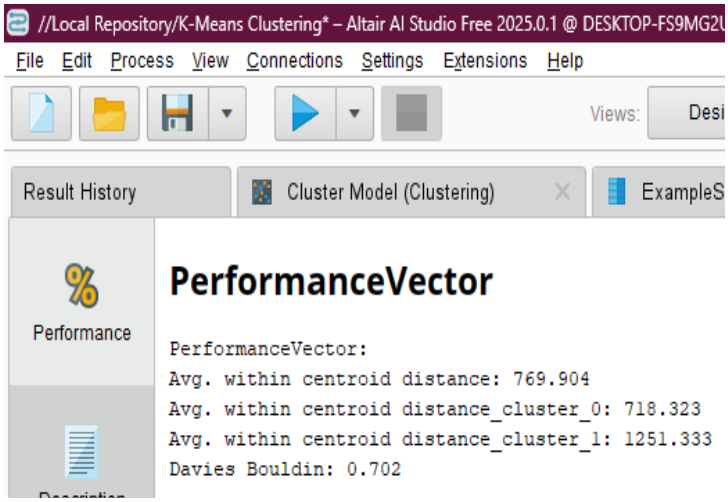
Gambar 4. Hasil *Clustering*

3. Pada operator *ExampleSet*, klik menu visualisasi lalu pilih diagram yang akan digunakan seperti pada Gambar 5 di bawah ini. Dalam studi ini, digunakan diagram batang:



Gambar 5. Diagram Hasil *Clustering*

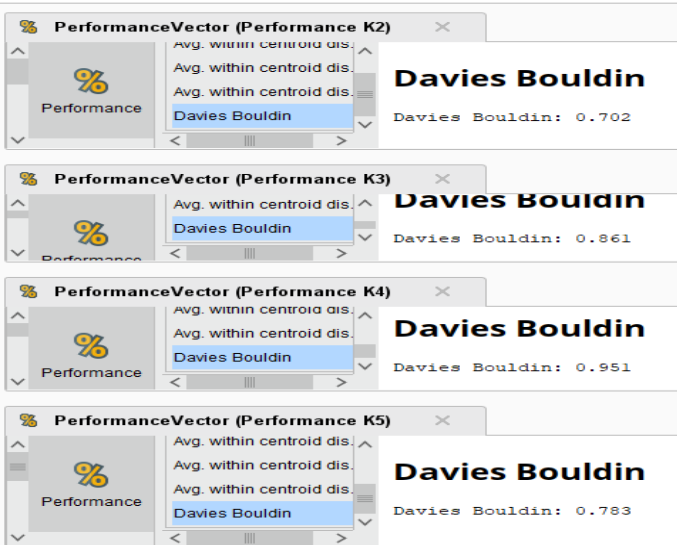
4. Karakterisrtik Klaster
Hasil perhitungan dalam RapidMiner ditampilkan pada Gambar 6 di bawah ini:



Gambar 6. Karakteristik Klaster

Berdasarkan hasil pengelompokan *K-Means* menggunakan aplikasi Rapidminer, klaster 0 memiliki nilai 789.904 sedangkan klaster 1 memiliki nilai 1.251.333. Klaster 0 adalah klaster dengan nilai rata-rata terendah dari semua variabel dibandingkan dengan klaster 1. Akibatnya, Kelompok 1 (C1) diklasifikasikan sebagai berisiko rendah dan Kelompok 2 (C2) sebagai berisiko tinggi.

5. Analisis hasil klaster
Untuk perbandingan, para peneliti membandingkan pembentukan klaster K2 hingga K5 dalam aplikasi RapidMiner menggunakan operator Multiply. Nilai *DBI* untuk K2 hingga K5 ditampilkan pada Gambar 7 di bawah ini:



Gambar 7. Perbandingkan DBI K=2 - K=5

Berdasarkan perhitungan *Davies Bouldin* Index (*DBI*) di RapidMiner, nilai *DBI* dari setiap pembentukan klaster dari K2 hingga K5 adalah sebagai berikut: (K2) sebesar 0.702, (K3) sebesar 0.861, (K4) sebesar 0.951, dan (K5) sebesar 0.783. Hal ini berarti pembentukan dua klaster (K2) dapat menjadi jumlah klaster terbaik karena nilai *DBI*-nya merupakan nilai terdekat dengan 0 di antara klaster lainnya.

4. KESIMPULAN

Hasil penelitian pengelompokan tingkat kerentanan bencana di 31 kecamatan di Kabupaten Bandung dari tahun 2014 hingga 2024 menggunakan Algoritma *K-Means Clustering* dengan metode Data Mining. 28 kecamatan, termasuk Ajasari, Banjaran, Cangkuang, Cicalengka, Cikancung, Cilengkrang, Cileunyi, Cimaung, Cimenyan, Ciparay, Ciwidey, Ibun, Katapang, Kertasari, Kutawaringin, Majalaya, Margaasih, Margahayu, Nagreg, Pacet, Pameungpeuk, Pangalengan, Paseh, Pasirjambu, Rancabali, Rancaekek, dan Soreang, diidentifikasi sebagai klaster risiko rendah (C1). Kemudian, klaster tingkat bencana tinggi (C2) terdiri dari 3 kecamatan, yaitu: Baleendah, Bojongsoang, dan Dayeuhkolot. Hasil evaluasi klastering dengan perhitungan *Davies Bouldin* Index (*DBI*) di RapidMiner adalah 0.702. Hal ini berarti pembentukan dua klaster sudah cukup baik karena nilai *DBI* mendekati nol.

DAFTAR PUSTAKA

[1] T. Qodrifuddin *et al.*, "Peningkatan Pemahaman Masyarakat terhadap Bahaya dan Dampak Bencana Alam Serta Penanggulangannya," *J. Pengabdi. Magister Pendidik. IPA*, vol. 5, no. 1, pp. 173–177, 2022, doi: 10.29303/jpmpi.v5i1.1400.

[2] A. Adi *et al.*, *IRBI Indeks Risiko Bencana Indonesia Tahun 2023*, vol. 02. 2024. [Online]. Available: https://inarisk.bnpb.go.id/IRBI-2023/mobile/index.html

[3] I. Rinjani, S. Anwar, and R. Herdiana, "Pengelompokan Daerah Bencana Alam Menggunakan Algoritma K-Means *Clustering*," *J. Ilm. Sist. Inf. DAN ILMU Komput.*, vol. 3, no. 1, pp. 35–51, 2023, doi: 10.55606/juisik.v3i1.417.

[4] A. Ardiansyah, F. I. Budi, Z. Hiya, R. Fadillah, and I. Ibnu, "Analisa Kerawanan Banjir Kabupaten Bandung Dengan Software ARGIS," *Konstr. Publ. Ilmu Tek. Perenc. Tata Ruang dan Tek. Sipil*, vol. 2, no. 2, pp. 198–210, 2024, doi:

10.61132/konstruksi.v2i2.273.

[5] L. H. Baihaqi and A. Kurniadi, "Implementasi Kebijakan Pengurangan Risiko Bencana Banjir Pemerintah Kabupaten Bandung Dalam Rangka Mendukung Ketahanan Wilayah," *J. PAPATUNG*, vol. 6, no. 2, pp. 35–49, 2023, doi: doi.org/10.54783/japp.v6i2.874.

[6] S. P. Valentina, "Evaluasi Pelaksanaan Kesiapsiagaan Bencana Banjir Di Kabupaten Bandung Provinsi Jawa Barat," *Sustain.*, vol. 11, no. 1, pp. 1–14, 2020, [Online]. Available: http://eprints.ipdn.ac.id/17075/1/Evaluasi Pelaksanaan Kesiapsiagaan Bencana Banjir di Kabupaten Bandung Provinsi Jawa Barat.pdf

[7] M. F. Al Halik and L. Septiana, "Analisa Data Untuk Prediksi Daerah Rawan Bencana Alam Di Jawa Barat Menggunakan Algoritma *K-Means Clustering*," *J. Inf. Syst. Applied, Manag. Account. Res.*, vol. 6, no. 4, pp. 856–870, 2022, [Online]. Available: https://journal.stmikjayakarta.ac.id/index.php/jisamar/article/view/939

[8] P. Rahayu *et al.*, *Buku Ajar Data Mining*, vol. 1. PT. Sonpedia Publishing Indonesia, 2024. [Online]. Available: https://www.researchgate.net/publication/377415198_BUKU_AJAR_DATA_MINING

[9] F. Khoirunnisa and Y. Rahmawati, "Komparasi 2 Metode Cluster Dalam Pengelompokan Intensitas Bencana Alam Di Indonesia," *J. Inform. dan Tek. Elektro Terap.*, vol. 12, no. 1, pp. 68–79, 2024, doi: 10.23960/jitet.v12i1.3619.

[10] D. Rohman, R. Annisa, D. Indriyana Efendi, and D. Solahudin, "*Clustering* Bencana Alam Menggunakan *K-Means* Pada Wilayah Jawa Barat," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 8, no. 1, pp. 493–500, 2024, doi: 10.36040/jati.v8i1.8409.

[11] BPBD, "Jumlah kasus bencana yang terjadi," 20 Februari. [Online]. Available: https://satudata.bandungkab.go.id/dataset/jumlah-kasus-bencana-yang-terjadi

[12] M. Herviany, S. Putri Delima, T. Nurhidayah, and K. Kasini, "Perbandingan Algoritma *K-Means* dan K-Medoids untuk Pengelompokkan Daerah Rawan Tanah Longsor Pada Provinsi Jawa Barat," *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 1, no. 1, pp. 34–40, 2021, doi: 10.57152/malcom.v1i1.60.

[13] I. F. Rozi, M. A. Hendrawan, and T. A. Rifda, "Tingkat Kerawanan Desa Berdasarkan Dampak Bencana Kab . Bojonegoro Dengan Metode *Clustering* Algoritma *K-Means*," *MULTINETICS (jurnal Multimed. Netw. informatics)*, vol. 10, no. 2, pp. 147–155, 2024, doi: 10.32722/multinetics.v10i2.6686.

[14] A. A. Putrie and R. Sanjaya, "Pengelompokan Kabupaten/Kota Berdasarkan Indikator Tingkat Pengangguran Menggunakan Algoritma *K-Means Clustering* (Studi Kasus: Provinsi Jawa Barat)," *eProsiding Sist. Inf.*, vol. 2, no. 2, pp. 111–121, 2021, [Online]. Available: https://eprosiding.ars.ac.id/index.php/psi/article/view/595

[15] D. N. Sari, Bambang, and B. Agus, "Mengidentifikasi Resiko Banjir Di Kabupaten Cirebon Dalam Menghadapi Curah Hujan Yang Ekstrem Menggunakan Algoritma *K-Means*," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 8, no. 3, pp. 2982–2987, 2024, doi: 10.36040/jati.v8i3.9610.

[16] E. Tasia and M. Afdal, "Perbandingan Algoritma *K-Means* Dan K-Medoids Untuk *Clustering* Daerah Rawan Banjir Di Kabupaten Rokan Hilir," *Indones. J. Inform. Res. Softw. Eng.*, vol. 3, no. 1, pp. 65–73, 2023, doi: 10.57152/ijirse.v3i1.523.

[17] B. G. Sudarsono, M. I. Leo, A. Santoso, and F. Hendrawan, "Analisis Data Mining Data Netflix Menggunakan Aplikasi Rapid Miner," *JBASE - J. Bus. Audit Inf. Syst.*, vol. 4, no. 1, pp. 13–21, 2021, doi: 10.30813/jbase.v4i1.2729.

[18] N. Dwitiyanti, S. Ayu Kumala, and S. Dwi Handayani, "Penerapan Metode K-Means Pada Klasterisasi Wilayah Rawan Gempa Di Indonesia Implementation Of *K-Means* Method In Classterization Of Earthquake Prone Areas In Indonesia," *Pros. Semin. Nas. UNIMUS*, vol. 6, pp. 1029–1037, 2023.