

# KOMPARASI 2 METODE *CLUSTER* DALAM PENGELOMPOKAN INTENSITAS BENCANA ALAM DI INDONESIA

Fitri Khoirunnisa<sup>1\*</sup>, Yulianto Rahmawati<sup>2</sup>

<sup>1</sup>Universitas Adhirajasa Reswara Sanjaya; Jl. Terusan Sekolah No.1-2; (022) 7100124

<sup>2</sup>Departemen Magister Sistem Informasi, STMIK LIKMI; Ir. H. Juanda No. 96; (022) 2502121

## Riwayat artikel:

Received: 27 September 2022

Accepted: 29 Desember 2023

Published: 1 Januari 2024

## Keywords:

Natural Disasters:

Data Mining:

K-Means, K-Medoids.

## Correspondent Email:

fitrikhoirunnisa66@gmail.com

**Abstrak.** Data mining merupakan sebuah proses pengolahan informasi yang digunakan untuk keperluan tertentu dari suatu database. Salah satu metode dalam data mining adalah *Clustering*, yang berfungsi untuk mencari pola, titik, objek, atau dokumen yang dapat dikelompokkan. Algoritma K-Means *clustering* memiliki peran penting dalam bidang data mining karena mudah diimplementasikan dan dijalankan. Namun, terdapat variasi pengembangan dari metode K-Means *Clustering*, yaitu K-Medoids, yang diciptakan untuk mengatasi kelemahan Algoritma K-Means yang cenderung sensitif terhadap outlier. Dalam penelitian ini, kedua algoritma *clustering* dibandingkan dengan menggunakan dataset yang berisi informasi tentang bencana alam di Indonesia dari tahun 2013-2022, yang terdiri dari atribut banjir, gempa bumi, kekeringan, puting beliung, tanah longsor, dan tsunami. Hasil pengujian menunjukkan bahwa nilai DBI pada K-Means lebih rendah dibandingkan K-Medoids, yaitu masing-masing 0.425 dan 0.939. Berdasarkan data tersebut, terdapat tiga *cluster* yang terbentuk, yang dapat diinterpretasikan berdasarkan intensitas bencana alam, dengan masing-masing *cluster* memiliki karakteristik yang berbeda.

**Abstract.** Data mining is an information processing process that is used for certain purposes from a database. One of the methods in data mining is *Clustering*, which functions to look for patterns, points, objects or documents that can be grouped. The K-Means clustering algorithm has an important role in the field of data mining because it is easy to implement and run. However, there is a variation of the development of the K-Means Clustering method, namely K-Medoids, which was created to overcome the weakness of the K-Means Algorithm which tends to be sensitive to outliers. In this study, the two clustering algorithms are compared using a dataset that contains information about natural disasters in Indonesia from 2013-2022, consisting of attributes of floods, earthquakes, droughts, tornadoes, landslides, and tsunamis. The test results show that the DBI value on K-Means is lower than K-Medoids, namely 0.425 and 0.939 respectively. Based on these data, three clusters are formed, which can be interpreted based on the intensity of natural disasters, with each cluster having different characteristics.

## 1. PENDAHULUAN

Indonesia merupakan salah satu negara yang terdiri dari banyak pulau, negara ini memiliki kerentanan yang tinggi terhadap dampak dari perubahan iklim [1]. Dari perubahan iklim yang terjadi maka dapat

memicu terjadinya bencana hidrometeorologi seperti banjir, tanah longsor, angin topan, dan gelombang laut yang tinggi. [2]. Secara umum, ada dua jenis bencana, yakni bencana alam dan bencana non-alam. Bencana alam terjadi akibat peristiwa atau serangkaian peristiwa yang

berasal dari alam, seperti gempa bumi, tsunami, letusan gunung berapi, banjir, kekeringan, angin topan, dan longsor. Contoh dari bencana alam yang termasuk dalam kategori hidrometeorologi antara lain banjir, kekeringan, dan longsor [3]. Indonesia merupakan kepulauan yang terletak di tengah Semenanjung Malaya, terkenal dengan sering terjadi bencana alam maupun buatan manusia. Menurut *Emergency Events Database (EM-DAT)* dari *Center for Research on the Epidemiology of Disasters (CRED)*, sebanyak 516 bencana alam terjadi di Indonesia dari tahun 1955 hingga 2021, yang menyebabkan kerugian total sebesar \$47 miliar. Kepulauan Indonesia terletak di tiga lempeng tektonik, yaitu lempeng Indo-Australia, Eurasia, dan Pasifik. Akibat pergerakan relatif lempeng tektonik, Indonesia rentan terhadap gempa bumi yang dipicu oleh posisi geografisnya [4].

Menurut laporan dari Badan Nasional Penanggulangan Bencana (BNPB), terjadi 3.544 bencana alam di Indonesia pada tahun 2022. Dari jumlah tersebut, banjir merupakan bencana alam yang paling sering terjadi sepanjang tahun 2022, yakni sebanyak 1.531 peristiwa. Selain itu cuaca ekstrem juga mencecahi bencana alam yang paling banyak terjadi yaitu mencapai 1.068 peristiwa. Tanah longsor juga tercatat sebanyak 634 peristiwa. Kebakaran hutan dan lahan terjadi sebanyak 252 peristiwa, sementara gempa bumi mencapai 28 peristiwa. Selama tahun 2022, ada 26 kejadian gelombang pasang/abrasi. Erupsi gunung Merapi mengalami 1 kejadian dan kekeringan sendiri mengalami 4 kejadian. Bencana tersebut mengakibatkan 858 orang meninggal dunia, 8.733 orang luka-luka, dan 37 orang hilang. Lebih dari 6,14 juta orang mengalami kerugian dan terpaksa mengungsi karena bencana tersebut. Sebanyak 95.403 rumah serta 1.983 fasilitas seperti sekolah, tempat peribadahan, dan fasilitas kesehatan juga mengalami kerusakan akibat bencana alam yang terjadi di Indonesia[5].

Teknologi informasi kini memungkinkan akses mudah dan cepat ke informasi yang akurat tanpa batasan wilayah dan waktu. Informasi mengenai bencana alam masih belum terorganisir dengan baik dan keabsahannya diragukan. Analisis data bencana alam dapat membantu mengurangi dampak bencana. Oleh karena itu, dibutuhkan prediksi daerah yang

rawan bencana alam dengan metode data mining untuk pengolahan data yang mudah dipahami [6].

Data Mining merupakan proses ekstraksi informasi dari data yang besar untuk menemukan pengetahuan yang tersembunyi dan memungkinkan penggunaannya secara *real-time*. Proses data mining memiliki beberapa algoritma analisis data, dan beberapa teknik utama yang digunakan untuk analisis adalah *Clustering*, *Association*, dan *Klasifikasi*, serta teknik lainnya. *Clustering* merupakan teknik yang efektif untuk analisis data eksplorasi dan dapat diterapkan di berbagai bidang. Terdapat tiga kategori metode *clustering* utama, yaitu partisi, hierarki, berbasis grid, dan berbasis model [7].

Algoritma *K-Means* merupakan salah satu metode clustering dengan teknik pengelompokan data non-hirarkisasi dengan bertujuan untuk membagi data menjadi satu atau lebih kelompok/kluster. Metode ini membagi data ke dalam kelompok/kluster sehingga data dengan karakteristik yang serupa dikelompokkan ke dalam satu kelompok yang sama, sementara data dengan karakteristik yang berbeda dikelompokkan ke dalam kelompok yang berbeda[6].

Algoritma *K-Medoids* merupakan salah satu metode partisi klasik dalam *Clustering* yang digunakan untuk mengelompokkan dataset berdasarkan objek-objeknya ke dalam kelompok-kelompok yang telah ditentukan sebelumnya. Dalam perbandingannya dengan *K-Means*, *K-Medoids* terbukti lebih efektif dalam menangani kebisingan (*noise*) dan penciran (*outlier*) karena dapat meminimalkan beberapa dissimilarities yang berpasangan, bukan hanya jumlah kuadrat jarak Euclidean [8].

Penggunaan teknik data mining bukanlah suatu hal yang baru karena telah banyak penelitian sebelumnya yang menggunakannya. Salah satu contoh penelitian terdahulu adalah studi yang dilakukan oleh Anisa Fira, Chaerur Rozikin, dan Garno yang berjudul "Komparasi Algoritma K-Means dan K-Medoids untuk Pengelompokan Penyebaran Covid-19 di Indonesia". Tujuan dari penelitian ini adalah untuk mengelompokkan provinsi-provinsi di Indonesia berdasarkan tingkat keparahan Covid-19, serta membandingkan penggunaan

algoritma K-Means dan K-Medoids dalam proses pengelompokannya [9].

Tujuan dari penelitian yang dilakukan oleh Hafiz Yusuf Heraldi, Nabila Churin Aprilia, dan Hasih Pratiwi dalam jurnal yang berjudul "Analisis Cluster Intensitas Kebencanaan di Indonesia Menggunakan Metode K-Means" adalah untuk melakukan pengelompokan provinsi-provinsi di Indonesia berdasarkan tiga jenis bencana yang sering terjadi, yaitu banjir, tanah longsor, dan puting beliung [10].

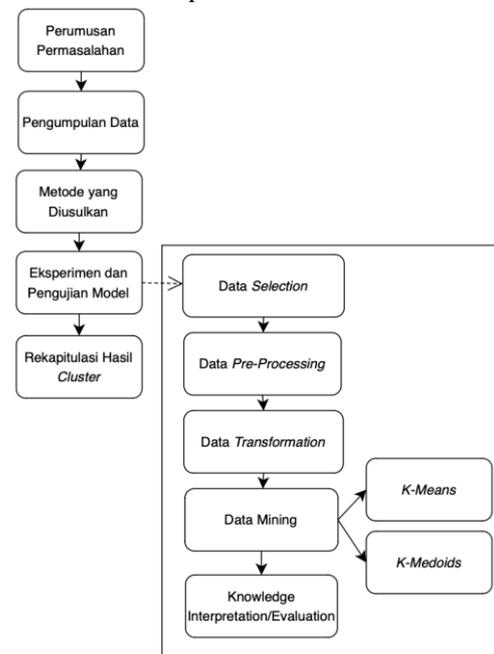
Penelitian yang dilakukan oleh Abdirahman Osman Hashi, Abdullahi Ahmed Abdirahman, Mohamed Abdirahman Elmi, Siti Zaiton Mohd Hashi, dan Octavio Ernesto Romo Rodriguez dalam jurnal yang berjudul "A Real-Time Flood Detection System Based on Machine Learning Algorithms with Emphasis on Deep Learning" bertujuan untuk mengusulkan sebuah model sistem deteksi banjir waktu nyata yang baru dan andal. Model ini menggunakan beberapa algoritma Machine-Learning dan Deep Learning seperti Random Forest, Naive Bayes, J48, dan Convolutional Neural Networks. Tujuan utama dari model ini adalah untuk dapat mendeteksi ketinggian air dan mengukur potensi banjir dengan tingkat akurasi yang tinggi dan memberikan peringatan dini tentang kemungkinan konsekuensi kemanusiaan yang dapat terjadi sebelum terjadinya banjir [11].

Dari hasil studi literatur yang dilakukan terhadap penggunaan data mining dengan menggunakan algoritma *K-Means* dan *K-Medoids* dapat di lihat bahwa penerapan algoritma *K-Means* dan *K-Medoids* dapat digunakan sebagai upaya dalam pengelompokan data, serta pada penelitian sebelumnya belum menerapkan metode komparasi sebagai salah satu pencegahan dalam bencana alam yang terjadi. Penelitian yang akan dilakukan bertujuan untuk menggunakan teknik data mining dengan algoritma *K-Means* dan *K-Medoids* untuk mengelompokkan data bencana alam di Indonesia berdasarkan intensitas rawan bencana. Penelitian ini diharapkan dapat membantu pemerintahan dalam merancang strategi penanggulangan bencana alam yang lebih fokus dan efektif serta dapat menjadi perhatian masyarakat dalam penanggulangan bencana alam sedini mungkin.

## 2. METODE PENELITIAN

### 2.1. Perancangan Penelitian

Dalam artikel [12] Dawson mengungkapkan bahwa terdapat empat metode penelitian yang sering digunakan, yaitu penelitian langsung, eksperimen, studi kasus, dan survey. Dalam penelitian yang sedang dilakukan, metode penelitian yang dipilih adalah eksperimen. Gambar dibawah ini dapat digunakan untuk melihat tahapan-tahapan yang dilakukan dalam penelitian ini.



Gambar 1. Tahapan Penelitian

#### 2.1.1. Perumusan Permasalahan

Perumusan permasalahan dalam penelitian yang dilakukan adalah belum ada informasi yang lengkap kepada masyarakat terhadap pengkategorian atau pengelompokan bencana alam yang terjadi di Indonesia maka dibuatlah penelitian ini untuk mengetahui bagaimana melakukan analisis *cluster* pada data bencana alam di Indonesia selama periode 2013-2022 dengan menggunakan algoritma *K-Means* dan *K-Medoids*. Perumusan masalah ini meliputi langkah-langkah pengelompokan data bencana alam ke dalam beberapa *cluster* berdasarkan kesamaan karakteristiknya serta menentukan jumlah *cluster* yang optimal. Selain itu, perumusan masalah juga mencakup analisis pola dan karakteristik setiap *cluster* untuk membantu pengambilan keputusan dalam pengelolaan bencana alam di Indonesia.

Dengan demikian, perumusan masalah tersebut memberikan gambaran tentang tujuan dari analisis *cluster* yang dilakukan serta manfaat dari hasil analisis tersebut untuk pengelolaan bencana alam di Indonesia.

### 2.1.2. Pengumpulan Data

Data yang diperoleh merupakan data sekunder yang berasal dari <https://gis.bnpp.go.id>, yang terdiri dari informasi mengenai bencana alam yang terjadi di Indonesia pada periode 2013-2022. Fokus data tersebut adalah bencana alam yang terjadi di Jawa Barat, yang mencakup 6 jenis bencana alam, yaitu banjir, gempa bumi, kekeringan, puting beliung, tanah longsor, dan tsunami. Data ini mencakup sekitar 29.234 kabupaten/kota yang berada di Indonesia.

### 2.1.3. Metode yang diusulkan

Pada tahap ini, dipilih model atau metode yang akan digunakan dalam penelitian. Salah satu metode yang diusulkan adalah pengelompokan *clustering* menggunakan algoritma K-Means dan K-Medoids. *Clustering* adalah proses pengelompokan kelas yang didasarkan pada kemiripan antar objek. Pengelompokan data sangat penting untuk mempercepat pengolahan data komputer, karena jumlah data yang sangat banyak. Oleh karena itu, *clustering* telah digunakan secara luas dan menjadi penting dalam pengelompokan data [13]

Metode K-means merupakan salah satu teknik yang sering digunakan dalam data mining untuk melakukan pengelompokan data tanpa menggunakan variabel kontrol. Metode ini efektif dalam mempartisi data ke dalam kelompok yang saling berhubungan [14]. Dalam penggunaan algoritma *k-means*, terdapat beberapa aturan, yaitu:

- a. Menentukan jumlah *cluster* yang dibutuhkan.
- b. Hanya menggunakan atribut yang bertipe numerik.

Algoritma k-means akan memilih beberapa komponen data secara acak untuk dijadikan pusat *cluster* awal, dan setiap komponen data akan dites dan dikelompokkan ke dalam salah satu pusat *cluster* tergantung pada jarak minimum dengan setiap pusat *cluster*. Kemudian, posisi pusat *cluster* akan dihitung ulang sampai semua komponen data telah

digolongkan ke dalam tiap-tiap *cluster* dan terbentuklah *cluster* baru [15].

Metode K-Medoids digunakan untuk mencari titik medoid yang menjadi pusat dari suatu kelompok atau *cluster*. Metode ini dianggap lebih superior dibandingkan dengan metode K-Means karena pada K-Medoids, kita mencari objek k yang mewakili kelompok untuk meminimalkan jumlah perbedaan antara objek data. Sementara pada metode K-Means, perbedaan antar objek data diukur dengan jarak Euclidean [16].

### 2.1.4. Eksperimen dan Pengujian Model

Langkah ini melibatkan proses pengujian model yang telah diajukan untuk mengevaluasi kinerjanya. Dataset yang telah melalui proses pemrosesan akan diuji menggunakan dua jenis algoritma, yaitu K-Means dan K-Medoids. Beberapa langkah yang dilakukan dalam tahap pengujian model ini antara lain sebagai berikut :

1. Data Selection
2. Data Preprocessing
3. Data Transformation
4. Data Mining
  - a. K-Means
  - b. K-Medoids

### 2.1.5. Rekapitulasi Hasil Cluster

Pada tahap ini, dilakukan validasi hasil dari model yang telah diajukan melalui eksperimen untuk mengukur kinerja dan akurasi model dalam menyelesaikan masalah yang dihadapi dalam penelitian. Evaluasi ini diharapkan dapat memberikan pemahaman yang lebih baik tentang keunggulan dan kelemahan dari model yang diusulkan, serta menilai efektivitas dan efisiensi model tersebut. Oleh karena itu, keberhasilan penelitian dan kontribusinya dalam pengembangan pengetahuan dan aplikasi di bidang yang terkait sangat bergantung pada model yang diusulkan.

## 3. HASIL DAN PEMBAHASAN

Di bawah ini adalah hasil dan analisis dari studi pengelompokan (*clustering*) bencana alam di Indonesia pada tahun 2013-2022 dengan menggunakan algoritma K-Means dan K-Medoids.

**A. Data Selection**

Data yang digunakan dalam penelitian ini terdiri dari jumlah kasus bencana alam yang terjadi di Indonesia. Atribut yang digunakan mencakup data bencana banjir, gempa bumi,

kekeringan, puting beliung, tanah longsor dan tsunami. Data tersebut didapatkan dari situs Geoportal Data Bencana Indonesia.

Tabel 1. Dataset Jumlah Bencana Alam Indonesia Tahun 2013-2022

No	ID Kabupaten	Tanggal Kejadian	Kejadian	Lokasi	Kabupaten	Provinsi	Kronologi & Dokumentasi
1	3319	14/07/2022	TANAH LONGSOR	Kec. Dawe Ds. Colo	KUDUS	JAWA TENGAH	Dokumentasi
..	.....	.....	.....	.....	.....	.....	.....
8120	1111	01/08/2020	PUTING BELIUNG	KEC CIBUNGBU LANG KEL CIARUTEN ILIR	BOGOR	JAWA BARAT	Dokumentasi
..	.....	.....	.....	.....	.....	.....	.....
29234	7322	01/01/2013	BANJIR	-	LUWU UTARA	Sulawesi Selatan	Dokumentasi

**B. Data Pre-Processing**

Selanjutnya, setelah proses pemilihan data (*data selection*), tahap *pre-processing* selanjutnya adalah membersihkan data dengan mengatasi masalah missing value, duplikasi data, dan data yang tidak konsisten. Namun

pada kedua dataset yang diolah, tidak ditemukan adanya masalah missing value, duplikasi data, maupun data yang tidak konsisten

Tabel 2. Dilakukan Tahap Data *Pre-Processing*

No	Kejadian	Kabupaten
1	TANAH LONGSOR	KUDUS
..	.....	.....
8120	PUTING BELIUNG	BOGOR
..	.....	.....
29234	BANJIR	LUWU UTARA

**C. Data Transformation**

Pada tahap data transformation ini maka dilakukanlah penggabungan dataset dengan atribut dan variabel yang sama. Dataset yang

telah digabungkan dapat dilihat pada tabel berikut.

Tabel 3. Dilakukan Tahap Data Transformation

No	Nama Kabupaten/Kota	Banjir	Gempa Bumi	Kekeringan	Puting Beliung	Tanah Longsor	Tsunami
1	ACEH BARAT	33	0	0	23	5	5
.....	.....	.....	.....	.....	.....	.....	.....
200	KOTA JAKARTA TIMUR	96	1	0	4	9	0
.....	.....	.....	.....	.....	.....	.....	.....
283	WONOSOBO	8	1	3	48	74	0

**D. Data Mining**

Pada proses data mining yang akan dilakukan yaitu dengan menggunakan algoritma *K-Means* dan *K-Medoids*. Dalam mengetahui hasil penentuan *cluster* yang dibentuk tergolong dalam kategori baik, maka

dengan itu dilakukanlah percobaan validasi dari pembentukan jumlah *k* sebanyak 2 sampai dengan 4 *k (cluster)*. Berikut merupakan table hasil percobaan validasi yang telah digunakan.

Tabel 4. Validasi Pembentukan Jumlah *k Cluster*

<i>k (Cluster)</i>	<i>Davies-Bouldin Index</i>	
	<i>K-Means</i>	<i>K-Medoids</i>
3	0.425	0.939
4	0.595	1.538
5	0.726	1.738

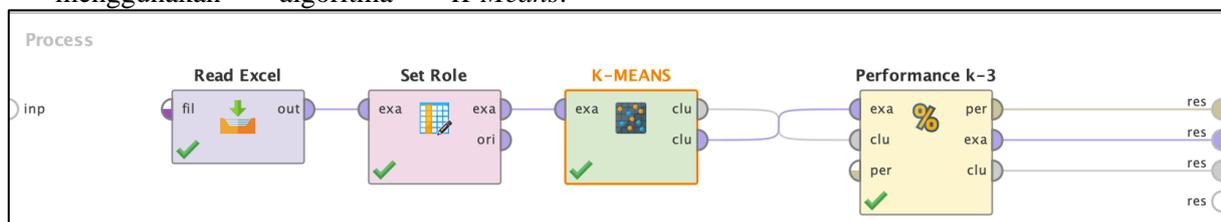
Nilai DBI yang paling terendah dari masing-masing algoritma yaitu dengan menggunakan 2 *k (cluster)*. Dari perhitungan yang telah dilakukan maka dibuatlah beberapa

skema pengujian menggunakan algoritma *K-Means* dan *K-Medoids* sesuai dengan nilai baik pengelompokan yaitu 2 *cluster*. berikut merupakan pengujian skema yang dilakukan.

**1) Skema 1 Pengujian dengan Metode K-Means**

Penggunaan dataset yang telah proses kemudian dilakukan pengujian dengan menggunakan algoritma *K-Means*.

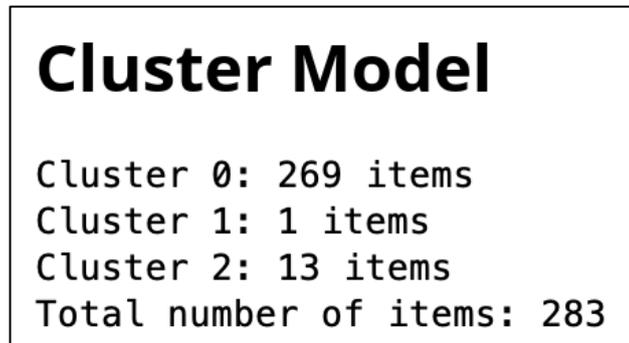
Operator yang digunakan dalam proses pengujian ini dapat dilihat dibawah ini.



Gambar 2. Operator Data Mining Pada Rapidminer Algoritma *K-Means*

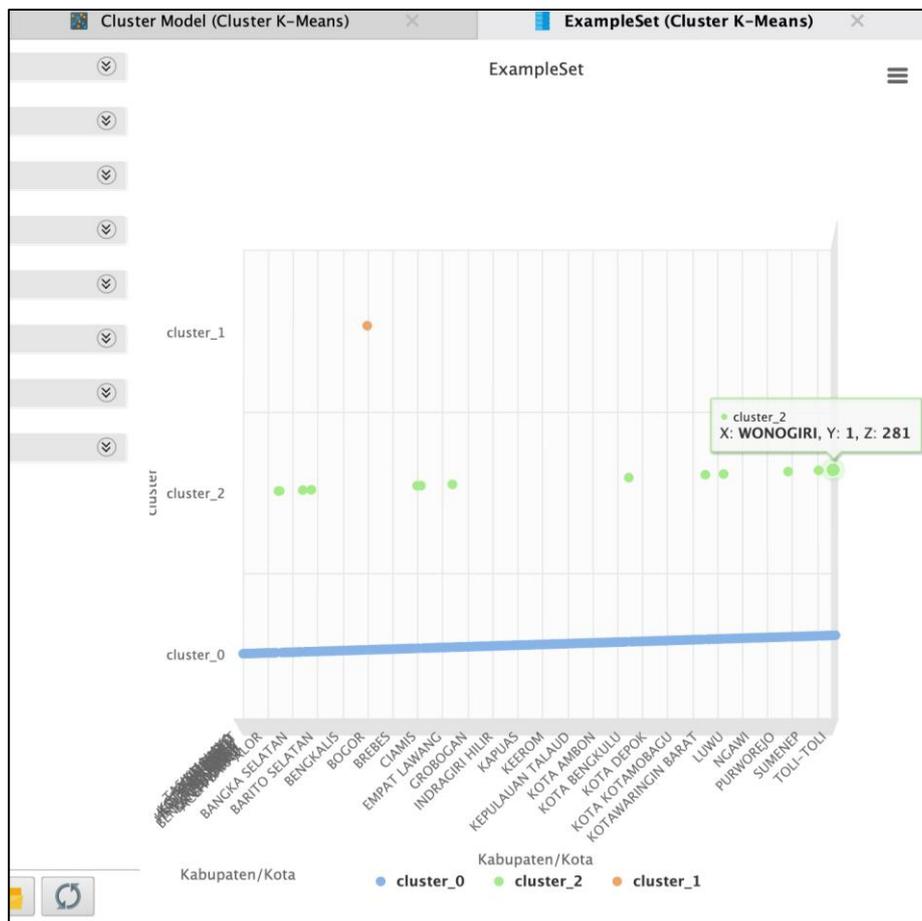
Dari proses pengujian yang telah dilakukan dengan menggunakan rapidminer

maka dapat dihasilkan 3 cluster dengan hasil yang dapat dilihat dibawah ini.



Gambar 3. Hasil Cluster Model Algoritma *K-Means*

Adapun gambar visualisasi dari sebaran titik cluster model yang sudah terbentuk adalah sebagai berikut,

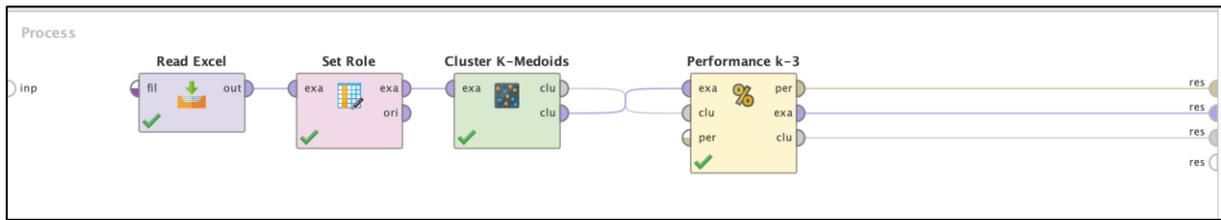


Gambar 4. Visualisasi sebaran titik cluster Algoritma *K-Means*

## 2) Skema 2 Pengujian dengan Metode *K-Medoids*

Penggunaan dataset yang telah proses kemudian dilakukan pengujian dengan menggunakan algoritma *K-Medoids*.

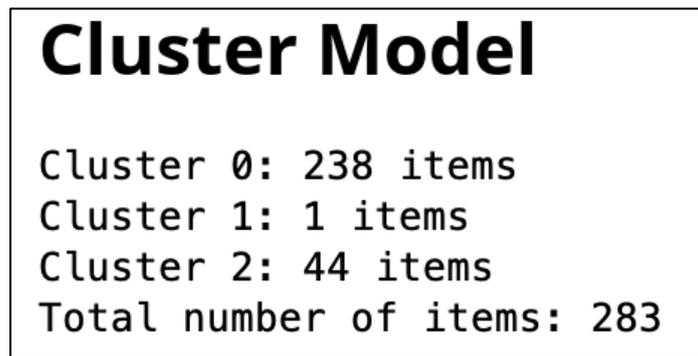
Operator yang digunakan dalam proses pengujian ini dapat dilihat dibawah ini.



Gambar 5. Operator Data Mining Pada Rapidminer Algoritma *K-Medoids*

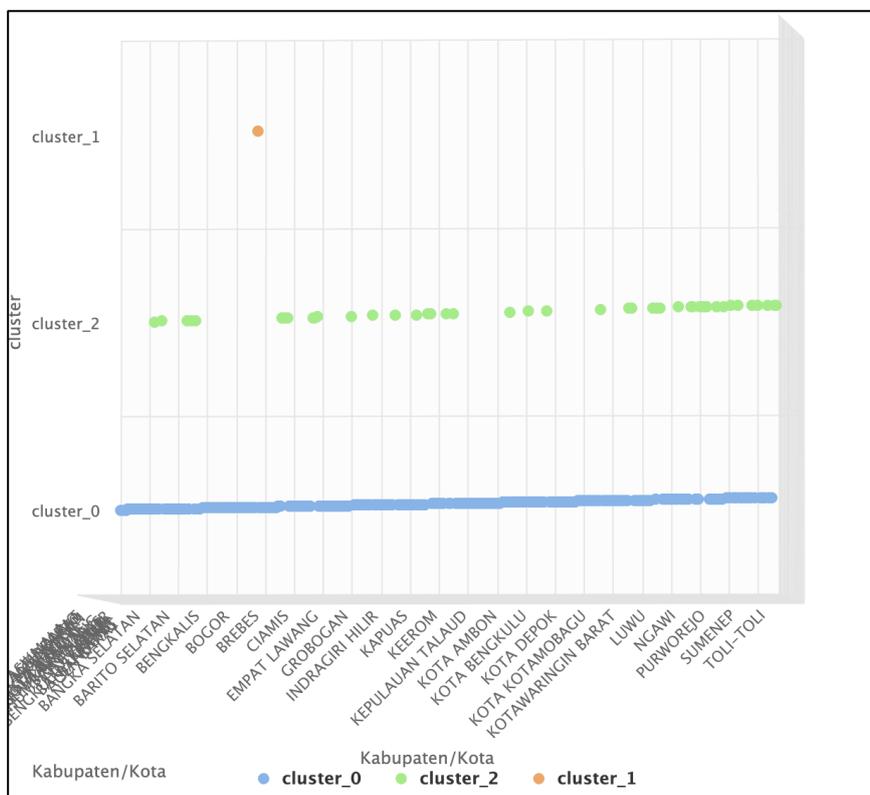
Dari proses pengujian yang telah dilakukan dengan menggunakan rapidminer maka dapat dihasilkan 3 cluster dengan

pengelompokan yang dapat dilihat pada gambar berikut .



Gambar 6. Hasil Cluster Model Algoritma *K-Medoids*

Adapun gambar visualisasi dari sebaran titik cluster model yang sudah terbentuk adalah sebagai berikut,



Gambar 7. Visualisasi sebaran titik cluster Algoritma *K-Medoids*

**E. Knowledge Interpretation/Evaluation**

Pada tahap ini, dilakukan evaluasi terhadap hasil pengelompokan data (*clustering*) untuk membandingkan kinerja yang lebih unggul. Dalam penelitian ini, Davies-Bouldin Index digunakan untuk mengevaluasi hasil pengelompokan. Kedua skema akan dibandingkan berdasarkan nilai DBInya, lalu akan disimpulkan skema mana yang memiliki kinerja yang lebih baik.

Tabel 5. Hasil Evaluasi Davies-Bouldin Index

Algoritma	Nilai DBI
K-Means	0.425
K-Medoids	0.939

Berdasarkan data tabel diatas maka dapat disimpulkan bahwa dengan menggunakan skema 1 yaitu menggunakan algoritma K-Means merupakan metode yang lebih baik, Skema 1 memiliki nilai DBI sebesar 0.425. Rincian dari kriteria hasil clustering dapat dilihat di bawah ini .

Tabel 6. Rincian Hasil Skema 1

<i>k (Cluster)</i>	Data Cluster	Nilai DBI
Cluster 0	269	0.425
Cluster 1	1	
Cluster 2	13	

**F. Rekapitulasi Hasil Cluster**

Berdasarkan hasil dari proses clustering dengan algoritma K-Means dan K-Medoids dengan menggunakan aplikasi rapidminer, serta algoritma yang tepat digunakan yaitu K-means karena memiliki nilai DBI terendah, maka dapat disimpulkan bahwa kategori cluster yang terbagi yaitu pada Cluster 0 memperoleh 269 data serta Cluster 1 memperoleh 1 data dan cluster 2 memperoleh 13 data. Dalam proses yang dilakukan maka metode yang baik diterapkan dalam pengkategorian intensitas bencana adalah dengan menggunakan algoritma K-Means dapat dilihat dalam perolehan nilai DBI sebesar 0.425. maka berikut merupakan rincian data dihasilkanlah pada setiap cluster yang terbentuk sebagai berikut :

Tabel 7. Kesimpulan Hasil Cluster

Cluster	Nama Kabupaten/Kota
0	Aceh Barat, Aceh Barat Daya, Aceh Besar, Aceh Jaya, Aceh Selatan, Aceh Singkil, Aceh Tamiang, Aceh Tengah, Aceh Tenggara, Aceh Timur, Aceh Utara, Agam, Alor, Asahan, Asmat, Badung, Balangan, Banggai, Banggai Kepulauan, Banggai Laut, Bangka, Bangka Barat, Bangka Selatan, Bangka Tengah, Bangkalan, Bangli, Banjar, Bantaeng, Bantul, Banyuasin, Banyuwangi, Barito Kuala, Barito Selatan, Barito Timur, Barito Utara, Barru, Batang, Batanghari, Batu Bara, Bekasi, Belitung, Belitung Timur, Belu, Bener Meriah, Bengkalis, Bengkayang, Bengkulu Selatan, Bengkulu Tengah, Bengkulu Utara, Berau, Bima, Bintan, Bireuen, Blitar, Blora, Boalemo, Bojonegoro, Bolaang Mongondow, Bolaang Mongondow Selatan, Bolaang Mongondow Timur, Bolaang Mongondow Utara, Bombana, Bondowoso, Bone, Bone Bolango, Boven Digoel, Boyolali, Brebes, Buleleng, Bulukumba, Bulungan, Bungo, Buol, Buru, Buru Selatan, Buton, Buton Selatan, Buton Tengah, Buton Utara, Cianjur, Cirebon, Dairi, Deiyai, Deli Serdang, Demak, Dharmasraya, Dogiyai, Dompu, Donggala, Empat Lawang, Ende, Enrekang, Fak-Fak, Flores Timur, Gayo Lues, Gianyar, Gorontalo, Gorontalo Utara, Gowa, Gresik, Grobogan, Gunung Mas, Gunungkidul, Halmahera Barat, Halmahera Selatan, Halmahera Tengah, Halmahera Timur, Halmahera Utara, Hulu Sungai Selatan, Hulu Sungai Tengah, Hulu Sungai Utara, Humbang Hasundutan, Indragiri Hilir, Indragiri Hulu, Indramayu, Intan Jaya, Jayapura, Jayawijaya,

	<p>Jember, Jembrana, Jenepono, Jepara, Jombang, Kampar, Kapuas, Kapuas Hulu, Karanganyar, Karangasem Karawang, Karimun, Karo, Katingan, Kaur, Kayong Utara, Kebumen, Kediri, Keerom, Kendal, Kepahiang, Kepulauan Anambas, Kepulauan Aru, Kepulauan Mentawai, Kepulauan Meranti, Kepulauan Sangihe, Kepulauan Selayar, Kepulauan Seribu, Kepulauan Siau Tagulandang Biaro, Kepulauan Sula, Kepulauan Talaud, Kepulauan Yapen, Kerinci, Ketapang, Klaten, Klungkung, Kolaka, Kolaka Timur, Kolaka Utara, Konawe, Konawe Selatan, Konawe Utara, Kota Ambon, Kota Balikpapan, Kota Banda Aceh, Kota Bandar Lampung, Kota Bandung, Kota Banjar, Kota Banjarbaru, Kota Banjarmasin, Kota Batam, Kota Batu, Kota Baubau, Kota Bekasi, Kota Bengkulu, Kota Bima, Kota Binjai, Kota Bitung, Kota Blitar, Kota Bontang, Kota Bukittinggi, Kota Cilegon, Kota Cimahi, Kota Cirebon, Kota Denpasar, Kota Depok, Kota Dumai, Kota Gorontalo, Kota Gunung Sitoli, Kota Jakarta Barat, Kota Jakarta Pusat, Kota Jakarta Selatan, Kota Jakarta Timur, Kota Jakarta Utara, Kota Jambi, Kota Jayapura, Kota Kendari, Kota Kotamobagu, Kota Kupang, Kota Surakarta, Kota Tangerang, Kota Tasikmalaya, Kota Tegal</p> <p>Kota Ternate, Kota Tidore Kepulauan, Kota Tomohon, Kota Tual, Kota Yogyakarta, Kotabaru, Kotawaringin Barat, Kotawaringin Timur, Kuantan Singingi, Kubu Raya, Kudus, Kupang, Lahat, Lamongan, Landak, Lebak, Lumajang, Luwu, Madiun, Magetan, Majalengka, Malang, Mamasa, Mamberamo Tengah, Manokwari, Manokwari Selatan, Mojokerto, Nganjuk, Ngawi, Pacitan, Pandeglang, Pangandaran,</p>
--	--

	<p>Pasuruan, Pati, Pekalongan, Pemalang, Penajam Paser Utara, Probolinggo, Purbalingga, Purwakarta, Purworejo, Semarang, Serang, Sidoarjo, Situbondo, Sleman, Sragen, Subang, Sukoharjo, Sumbawa, Sumedang, Sumenep, Tabanan, Tana Toraja, Tanah Bumbu, Tanah Datar Tanah Laut, Tangerang, Tapin, Tasikmalaya, Tegal, Tojo Una-Una, Toli-Toli, Toraja Utara, Trenggalek, Tuban, Wajo, Wonosobo.</p>
<b>1</b>	Bogor
<b>2</b>	Bandung, Bandung Barat, Banjarnegara, Banyumas, Ciamis, Cilacap, Garut, Kota Bogor, Kuningan, Magelang, Sukabumi, Temanggung, Wonogiri

Hasil dari penelitian pada klasterisasi data bencana alam di Indonesia pada tahun 2013-2022 dan telah dilakukan pengujian sehingga dapat menunjukkan bahwa :

1. Cluster 1 termasuk kedalam intensitas bencana alam tinggi dengan jumlah 1 data.
2. Cluster 2 termasuk kedalam intensitas bencana alam sedang dengan jumlah 13 data
3. Cluster 0 termasuk kedalam intensitas bencana alam rendah dengan jumlah 269 data.

#### 4. KESIMPULAN

Pada penelitian yang telah dilakukan maka dapat disimpulkan bahwa Hasil pengujian perbandingan antara *K-Means* dan *K-Medoids* untuk pengelompokan data menggunakan algoritma *clustering* menunjukkan bahwa tidak terdapat perbedaan yang signifikan. Namun, saat dihitung menggunakan Davies-Bouldin Index, nilai  $k=3$  pada *K-Means* lebih rendah dibandingkan dengan *K-Medoids* yaitu sebesar 0.425. Pengelompokan data dilakukan berdasarkan intensitas bencana alam yang terjadi di Indonesia dari tahun 2013-2022. Data tersebut menunjukkan bahwa terdapat tiga *cluster* yang terbentuk, yaitu *Cluster 1* dengan intensitas

bencana alam tinggi dan hanya memiliki 1 data, *Cluster 2* dengan intensitas bencana alam sedang dan memiliki 13 data, serta *Cluster 0* dengan intensitas bencana alam rendah dan memiliki 269 data.

### UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada pihak-pihak terkait yang telah memberi dukungan terhadap penelitian ini.

### DAFTAR PUSTAKA

- [1] I. Dewi Fitriani, W. Zulkarnaen, and A. Bagianto, "Analisis Manajemen Mitigasi Badan Penanggulangan Bencana Daerah (Bpbd) Terhadap Bencana Alam Erupsi Gunung Tangkuban Parahu Di Jawa Barat," *JIMEA / Jurnal Ilmiah MEA (Manajemen, Ekonomi, dan Akuntansi)*, vol. 5, no. 1, pp. 92–111, 2021.
- [2] A. Alijoyo and I. Sapuan, "Green Behavior and the Use of Eco-Product: Millennial Generation Perspective," *Eduvest-Journal of Universal Studies*, vol. 2, no. 1, pp. 121–130, 2022, [Online]. Available: <http://eduvest.greenvest.co.id>
- [3] I. J. A. Saragih, M. Sirait, and D. A. Sari, "Deskripsi Opini Publik tentang Bencana Alam untuk Rencana Studi Mitigasi di Indonesia (Studi kasus: Bencana Hidrometeorologi)," *MKGI: Jurnal Meteorologi, Klimatologi Geofisika dan Instrumentasi*, vol. 1, no. 1, pp. 33–39, 2021, [Online]. Available: <https://journal.physan.id/index.php/mkgip33> Journal homepage: <https://journal.physan.id/index.php/mkgi>
- [4] A. Dwi Maqfirah, N. Oktaviano Guntur, A. Burhanuddin, K. Atirah Monica Bubun, M. Daffa Rizqilah, and A. Apriliani, "Peran ASEAN Agreement on Disaster Management and Emergency Response (AADMER) dalam Penanggulangan Bencana Alam di Indonesia," Online, 2022.
- [5] BNPB, "BNPB (Badan Nasional Penanggulangan Bencana)," <https://www.bnpb.go.id/>, Mar. 2023. <https://www.bnpb.go.id/infografis/infografis-bencana-tahun-2022> (accessed Mar. 02, 2023).
- [6] M. Firman, A. Halik, and L. Septiana, "Analisa Data Untuk Prediksi Daerah Rawan Bencana Alam Di Jawa Barat Menggunakan Algoritma K-Means Clustering," *Journal of Information System, Applied, Management, Accounting and Research*, vol. 6, no. 4, pp. 856–870, 2022, doi: 10.52362/jisamar.v6i4.939.
- [7] M. Herviany, S. Putri Delima, T. Nurhidayah, and Kasini, "Comparison of K-Means and K-Medoids Algorithms for Grouping Landslide Prone Areas in West Java Province," *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 1, no. 1, pp. 34–40, 2021.
- [8] N. L. Anggraeni and S. Tresnawati, "Komparasi Algoritma K-Means Dan K-Medoids Untuk Menangani Strategi Promosi Di Politeknik Tedc Bandung," 2020.
- [9] A. Fira, C. Rozikin, and Garno, "Komparasi Algoritma K-Means dan K-Medoids Untuk Pengelompokan Penyebaran Covid-19 di Indonesia," *Journal of Applied Informatics and Computing (JAIC)*, vol. 5, no. 2, pp. 133–138, 2021.
- [10] H. Yusuf Heraldi, N. Churin Aprilia, and H. Pratiwi, "Analisis Cluster Intensitas Kebencanaan di Indonesia Menggunakan Metode K-Means," *Indonesian Journal of Applied Statistics*, vol. 2, no. 2, pp. 137–144, 2019.
- [11] A. O. Hashi, A. A. Abdirahman, M. A. Elmi, S. Z. M. Hashi, and O. E. R. Rodriguez, "A real-time flood detection system based on machine learning algorithms with emphasis on deep learning," *International Journal of Engineering Trends and Technology*, vol. 69, no. 5, pp. 249–256, May 2021, doi: 10.14445/22315381/IJETT-V69I5P232.
- [12] A. D. Andini and T. Arifin, "Implementasi Algoritma K-Medoids Untuk Klasterisasi Data Penyakit Pasien Di Rsud Kota Bandung," *JURNAL RESPONSIF*, vol. 2, no. 2, pp. 128–138,

- 2020, [Online]. Available: <http://ejurnal.ars.ac.id/index.php/jti>
- [13] D. Marlina, N. Fauzer Putri, A. Fernando, and A. Ramadhan, "Implementasi Algoritma K-Medoids dan K-Means untuk Pengelompokan Wilayah Sebaran Cacat pada Anak," *Jurnal CoreIT*, vol. 4, no. 2, 2018.
- [14] H. Firdaus and A. Sofro, "Analisa Cluster Menggunakan K-Means Dan Fuzzy C-Means Dalam Pengelompokan Provinsi Menurut Data Intesitas Bencana Alam Di Indonesia Tahun 2017-2021," *Jurnal Ilmiah Matematika*, vol. 10, no. 01, pp. 50–60, 2022.
- [15] M. Gading Sadewo, A. Perdana Windarto, and A. Wanto, "Penerapan Algoritma Clustering Dalam Mengelompokkan Banyaknya Desa/Kelurahan Menurut Upaya Antisipasi/ Mitigasi Bencana Alam Menurut Provinsi Dengan K-Means," *KOMIK (Konferensi Nasional Teknologi Informasi dan Komputer)*, vol. 2, no. 1, pp. 312–319, 2018, [Online]. Available: <http://ejurnal.stmik-budidarma.ac.id/index.php/komik>
- [16] S. Sindi *et al.*, "Analisis Algoritma K-Medoids Clustering Dalam Pengelompokan Penyebaran Covid-19 Di Indonesia," *Jurnal Teknologi Informasi*, vol. 4, no. 1, 2020.