

Analisa Data Untuk Prediksi Daerah Rawan Bencana Alam Di Jawa Barat Menggunakan Algoritma K-Means Clustering

Muhamad Firman Al Halik¹, Laila Septiana²

Program Studi Sistem Informasi
Fakultas Teknologi Informasi
Universitas Nusa Mandiri

firmanalkhalik89@gmail.com¹, laila@nusamandiri.ac.id²

Received: September 18, 2022. **Revised:** October 15, 2022. **Accepted:** October 26, 2022.

Issue Period: Vol.6 No.4 (2022), Pp.856-870

Abstrak: Bencana alam adalah suatu peristiwa yang mengakibatkan dampak dan pengaruh besar bagi populasi manusia. Jawa Barat merupakan provinsi yang memiliki kejadian bencana alam cukup banyak, diantaranya bencana alam tanah longsor, gempa bumi, banjir, angin puting beliung, kebakaran, dan sebagainya. Saat ini pengetahuan teknologi dan informasi berkembang sangat pesat. Canggihnya teknologi membuat setiap orang mampu mengakses dan mendapat informasi tanpa batasan. Hal ini membuat informasi sangat dibutuhkan di setiap lini kehidupan. Salah satunya ialah informasi tentang bencana alam, dimana informasi tentang bencana alam di butuhkan untuk penanggulangan bencana. Data mining merupakan teknik yang umum dilakukan untuk pengolahan data bencana alam, sebab teknik tersebut dianggap mampu menjadi sebuah solusi atas permasalahan penanggulangan bencana alam. Oleh karena itu, dalam penelitian ini membahas tentang pengelompokan data bencana alam untuk prediksi daerah rawan bencana alam di Jawa Barat dengan teknik data mining menggunakan algoritma k-means clustering. Hasil penelitian didapatkan 3 cluster diantaranya cluster rendah, cluster sedang, dan cluster tinggi..

Kata kunci: Data Mining, Bencana Alam, K-Means, Clustering

Abstract: A natural disaster is an event that has a major impact and impact on the human population. West Java is a province that has quite a lot of natural disasters, including landslides, earthquakes, floods, hurricanes, fires, and so on. Currently, knowledge of technology and information is growing very rapidly. Sophisticated technology makes everyone able to access and obtain information without restrictions. This makes information indispensable in every line of life. One of them is information about natural disasters, where information about natural disasters is needed for disaster management. Data mining is a commonly used technique for processing natural disaster data, because the technique is considered capable of being a solution to natural disaster management problems. Therefore, this study discusses the grouping of natural disaster data for prediction of natural disaster-prone areas in West Java with data mining techniques using the k-means clustering algorithm. The results of the study obtained 3 clusters including low cluster, medium cluster, and high cluster.

Keywords: Data Mining, Natural Disasters, K-Means, Clustering

I. PENDAHULUAN

Indonesia memiliki resiko bencana yang tinggi sebagai konsekuensi letak negara dari segi geologis dan geografis. Secara geologis, Indonesia berada pada pertemuan empat lempeng utama yaitu Eurasia, Indo



DOI: 10.52362/jisamar.v6i4.939

Ciptaan disebarluaskan di bawah [Lisensi Creative Commons Atribusi 4.0 Internasional](https://creativecommons.org/licenses/by/4.0/).

Australia, Filipina, dan Pasifik yang menjadikan Indonesia rawan bencana gempa bumi, tsunami, dan letusan gunung berapi. Sedangkan dari segi geografis Indonesia berada di daerah tropis dan pada pertemuan dua samudera serta dua benua membuat wilayah ini rawan bencana banjir, tanah longsor, banjir bandang, kekeringan, cuaca ekstrim dan abrasi yang juga dapat memicu kebakaran hutan dan lahan [1]. Daerah dengan tingkat rawan bencana hampir semuanya berada pada daerah dengan tingkat penduduk yang tinggi [2].

BNPB mencatat total kejadian bencana alam sejak 1 Januari hingga 31 Desember 2021 sebanyak 3.092 kasus, dengan wilayah paling banyak mengalami peristiwa bencana alam ialah provinsi Jawa Barat dan Jawa Timur. Sebaran kejadian bencana alam antara lain Jawa Barat dengan 768 peristiwa, Jawa Timur 398, Jawa Tengah 313, Aceh 216, Kalimantan Selatan 110, Sumatera Utara 104, Sulawesi Selatan 102, dan Kalimantan Tengah 100 kejadian [3].

Jawa Barat memiliki jenis bencana alam yang beragam. Diantaranya bencana Tanah Longsor, Banjir, Gempa Bumi, Angin Puting Beliung, dan Kebakaran Hutan serta Bangunan. Pada Januari sampai Desember tahun 2016, Tanah Longsor menjadi bencana yang sering terjadi di Jawa Barat, dengan 480 kali kejadian, kemudian bencana Kebakaran sebanyak 257 kejadian. Disusul bencana Banjir dengan 215 kali kejadian, bencana Angin Puting Beliung terjadi sebanyak 185 kali dan bencana Gempa Bumi sebanyak 48 kali. Pada tahun 2017 di bulan Januari sampai Maret, tercatat kejadian Kebakaran sebanyak 81 kali, kemudian Banjir sebanyak 62 kali, Tanah Longsor sebanyak 170 kali, Angin Puting Beliung 102 kali, dan Gempa Bumi 59 kali kejadian. Data di atas merupakan data bencana dengan intensitas yang cukup besar yang laporannya di terima oleh BPBD Provinsi Jawa Barat [4].

Perkembangan teknologi dan informasi saat ini mengalami perkembangan yang sangat pesat. Teknologi yang canggih membuat setiap orang dapat mengakses dan mendapatkan informasi secara cepat dan akurat, tanpa mengenal batasan wilayah dan waktu. Hal ini menjadikan informasi itu sangat berharga dan dibutuhkan untuk pengambilan keputusan. Salah satu informasi yang dibutuhkan oleh warga saat ini adalah informasi mengenai kejadian bencana alam. Namun informasi mengenai bencana alam yang tersaji saat ini masih bersifat acak, sulit dipahami dan belum terbukti keabsahannya. Meskipun bencana alam merupakan kejadian yang tidak bisa dihindari, akan tetapi dampak bencana dapat dikurangi dengan mengenali penyebab dan mempelajari kejadian bencana yang telah terjadi dengan menganalisa data bencana yang ada. Perlu dilakukan prediksi daerah rawan bencana alam supaya dapat menjadi indikator penanggulangan bencana alam. Berbagai metode pengolahan data menjadi sebuah informasi yang mudah dipahami sudah banyak dilakukan. Data mining merupakan teknik yang umum dilakukan untuk pengolahan data bencana alam [2].

Teknik data mining dapat diterapkan untuk menggali nilai tambah dari suatu kumpulan data berupa pengetahuan yang selama ini tidak diketahui secara manual [5]. Dari polanya data mining di kelompokkan ke dalam beberapa Teknik yaitu deskripsi, kemudian estimasi, prediksi, klasifikasi, clustering, dan asosiasi [2]. Data mining menjadi sebuah inti dari proses KDD (*knowledge discovery in database*) dimana KDD merupakan sebuah proses terorganisir untuk menemukan pengetahuan dan pola-pola yang berlaku dalam sebuah data yang besar. *Data mining* perlu dipelajari dan dipahami, karena manusia menghasilkan banyak data yang besar dalam berbagai bidang [6].

Pendekatan data mining bukanlah hal baru, terbukti dengan banyaknya penelitian terdahulu yang memanfaatkan [7]. Beberapa penelitian terdahulu diantaranya, penelitian yang dilakukan oleh Muhamad Iqbal Ramadan dan Prihandoko dalam jurnal yang berjudul “Penerapan *Data Mining* Untuk Analisa Data Bencana Milik BNPB Menggunakan Algoritma *K-means* dan *Linear Regression*” menggunakan algoritma *k-means* untuk pengelompokan data dan *linear regression* untuk melakukan prediksi data bencana milik BNPB selama 5 tahun kedepan [2]. Dalam jurnal berjudul “Implementasi *K-means Clustering* Untuk Pengelompokan Daerah Rawan Bencana Kebakaran Menggunakan Model *Crisp-dm*” yang disusun oleh F N Dheawiyanti dkk, mengelompokkan daerah rawan bencana kebakaran dengan algoritma *k-means clustering* dengan model *crisp-dm* bertujuan menghasilkan kelompok daerah yang memiliki potensi tinggi terjadinya bencana kebakaran [8]. Dalam penelitian yang dilakukan Risa Halilintar dan Intan Nur Farida menggunakan algoritma *k-means clustering* untuk melakukan prediksi prestasi nilai akademik mahasiswa yang terbagi menjadi 4 kelompok atau *cluster* yaitu sangat baik, baik, cukup, kurang yang berguna untuk mengetahui mahasiswa yang mempunyai predikat kelulusan yang sesuai dengan pengelompokan tersebut [9].

Berdasarkan latar belakang masalah, pada penelitian ini menggunakan teknik data mining untuk menggali informasi data bencana alam dengan menggunakan algoritma *K-Means Clustering* untuk mengelompokkan daerah rawan bencana alam yang ada di 27 Kabupaten/Kota di Provinsi Jawa Barat menjadi 3 *cluster* yaitu



cluster rendah, sedang, dan tinggi. Kemudian akan dilakukan pengelompokan daerah rawan bencana alam dengan intensitas kerawanan diatas 50% pada tiap variable bencana alam.

II. METODE DAN MATERI

2.1 Data Mining

Data mining merupakan sebuah proses dengan menggunakan Teknik dan alat analisis data untuk mencari pola dan hubungan yang tersembunyi dari sebuah data. Sebuah pendekatan yang mendasar dalam data mining ialah untuk meringkas dan mengekstrak informasi yang berguna serta masuk akal yang sebelumnya belum diketahui [6].

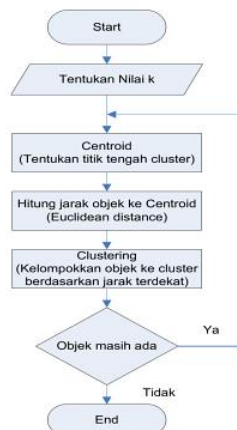
2.2 Clustering

Clustering adalah metode pengelompokan data ke dalam beberapa *cluster* maupun kelompok dimana data dalam *cluster* memiliki kemiripan yang maksimum dan data antar *cluster* memiliki kemiripan yang minimum [10]. *Clustering* digunakan untuk pengelompokan maupun mengidentifikasi data atau informasi yang mempunyai karakteristik tertentu [6].

2.3 Algoritma K-Means Clustering

K-means clustering adalah metode data clustering non-hirarki memperbaiki data yang ada ke dalam bentuk satu atau lebih cluster atau kelompok, sehingga data yang memiliki karakteristik yang sama di kelompokkan ke dalam satu cluster yang sama dan data yang mempunyai karakteristik yang berbeda dikelompokkan ke dalam kelompok yang lain [5].

K-means Clustering adalah, K sebagai konstanta jumlah *cluster* yang diinginkan, means yang berarti nilai suatu rata-rata dari suatu grup data yang dalam hal ini didefinisikan sebagai cluster, sehingga *k-means clustering* adalah suatu metode analisa data atau metode data mining yang melakukan proses pemodelan tanpa supervise dan merupakan salah satu metode yang melakukan pengelompokan data dengan system partisi. Metode *k-means* berguna untuk mengelompokkan data ke dalam satu kelompok, dengan karakteristik yang berbeda antara kelompok satu dengan kelompok yang lain [9]. Diagram alir algoritma *K-Means Clustering* :



Gambar 1. Diagram alir algoritma *K-Means Clustering* [6]

Pada algoritma *K-Means Clustering* perhitungan Ecludian Distance didapat dari jarak hasil pengurangan xi (nilai data) dan y (nilai pusat).

2.4 Metode Penelitian

2.4.1 Pengumpulan Data



DOI: 10.52362/jisamar.v6i4.939

Ciptaan disebarluaskan di bawah [Lisensi Creative Commons Atribusi 4.0 Internasional](https://creativecommons.org/licenses/by/4.0/).

Data merupakan kumpulan kejadian atau peristiwa yang berupa angka, symbol, kode dan lainnya. Pengumpulan data adalah pencatatan peristiwa atau kejadian yang dilakukan untuk mendapatkan informasi yang dibutuhkan dalam mencapai tujuan penelitian.

1. Kebutuhan Data

Data yang dibutuhkan berupa data bencana alam provinsi Jawa Barat menurut Kabupaten/Kota yang dihimpun dalam website resmi open data provinsi Jawa Barat.

2. Sumber Data

Pengumpulan data berdasarkan dokumen-dokumen dan keterangan yang dihasilkan Dinas Pemberdayaan Masyarakat dan Desa serta Dinas Lingkungan Hidup Provinsi Jawa Barat, yang terangkum dalam Open Data Jabar (<https://opendata.jabarprov.go.id>) sebagai portal data terbuka yang menyediakan data akurat dari seluruh organisasi perangkat daerah di Jawa Barat.

3. Cara Memperoleh Data

Data diperoleh dari mendownload atau mengunduh melalui website <https://opendata.jabarprov.go.id/> yang sudah disediakan oleh pemerintah Provinsi Jawa Barat.

4. Jumlah Data

Data yang diambil oleh penulis yaitu data bencana alam berdasarkan kabupaten/kota di Jawa Barat pada rentan tahun 2012 sampai 2021.

2.4.2 Populasi

Populasi merupakan area generalisasi yang tersusun dari beberapa objek maupun subjek yang memiliki kualitas serta karakter tertentu yang ditetapkan oleh peneliti sebagai bahan untuk dipelajari kemudian di tarik kesimpulan [12]. Dalam penelitian ini penulis menetapkan populasi diantaranya data bencana alam yang terjadi di 27 Kabupaten/Kota di Jawa Barat pada rentan waktu tahun 2012 sampai tahun 2021.

2.4.3 Sampel Penelitian

Sampel diartikan sebagai bagian dari jumlah serta karakter yang dimiliki oleh populasi. Pengambilan sampel subyek yang jumlahnya kurang dari 100, baiknya diambil semuanya sehingga menjadi penelitian populasi [12].

2.4.4 Analisis Data

Pada tahap ini dilakukan analisis data jumlah dan jenis bencana alam yang terjadi di Kabupaten/Kota di Jawa Barat. Data kemudian diolah dengan perhitungan bobot dari tiap indeks. Selanjutnya data akan dicluster ke dalam 3 *cluster* yakni *cluster* tingkat bencana alam rendah, *cluster* tingkat bencana alam sedang, dan *cluster* tingkat bencana alam tinggi. Dalam tahap ini akan dianalisis hasil dari pengolahan data.

III. PEMBAHASAN DAN HASIL

3.1 Data Bencana Alam

Data yang digunakan dalam penelitian ini adalah data jumlah bencana alam di 27 Kabupaten atau kota di Jawa Barat dari tahun 2012 sampai 2021 dengan 5 variabel bencana alam yaitu Tanah Longsor, Gempa Bumi, Banjir, Angin Putting Beliung, dan Kebakaran Hutan dan Lahan.

Tabel 1. Data Gabungan Jumlah Bencana Alam di Jawa Barat

No	Kabupaten/Kota	Tanah Longsor	Gempa Bumi	Banjir	Angin Putting Beliung	Kebakaran
1	Bogor	801	16	202	695	123
2	Sukabumi	890	95	139	289	420
3	Cianjur	162	18	58	56	34
4	Bandung	219	27	272	108	218
5	Garut	388	99	100	110	234
6	Tasikmalaya	273	89	47	82	53
7	Ciamis	365	8	87	238	228



DOI: 10.52362/jisamar.v6i4.939

Ciptaan disebarluaskan di bawah [Lisensi Creative Commons Atribusi 4.0 Internasional](https://creativecommons.org/licenses/by/4.0/).

8	Kuningan	506	3	59	90	144
9	Cirebon	26	0	88	75	85
10	Majalengka	306	2	54	71	127
11	Sumedang	352	2	82	65	242
12	Indramayu	5	2	43	30	11
13	Subang	90	4	78	76	25
14	Purwakarta	59	2	10	38	120
15	Karawang	15	0	122	94	59
16	Bekasi	8	3	96	29	37
17	Bandung Barat	245	3	30	77	77
18	Pangandaran	47	33	27	107	141
19	Kota Bogor	607	2	73	266	151
20	Kota Sukabumi	54	71	22	50	41
21	Kota Bandung	40	0	54	45	362
22	Kota Cirebon	7	1	14	35	48
23	Kota Bekasi	7	0	80	14	23
24	Kota Depok	11	0	14	12	15
25	Kota Cimahi	47	3	61	58	135
26	Kota Tasikmalaya	85	4	35	112	36
27	Kota Banjar	47	6	7	57	36

Sumber : Open Data Jabar (<https://opendata.jabarprov.go.id>)

3.2 Penerapan Algoritma K-Means

1. Transformasi Data

Pada tahap transformasi data, data diakumulasikan kemudian diambil nilai rata-rata dari setiap 860variable. Rata-rata = jumlah bencana ÷ jumlah tahun

Sampel :

▪ Bogor

- Tanah Longsor : $801 \div 10 = 80,1$
- Gempa Bumi : $16 \div 10 = 1,6$
- Banjir : $202 \div 10 = 20,2$
- Angin Puting Beliung : $695 \div 10 = 69,5$
- Kebakaran : $123 \div 10 = 12,3$

Tabel 2 Data Hasil Transformasi

No	Kabupaten/Kota	Tanah Longsor	Gempa Bumi	Banjir	Angin Puting Beliung	Kebakaran
1	Bogor	80,1	1,6	20,2	69,5	12,3
2	Sukabumi	89	9,5	13,9	28,9	42
3	Cianjur	16,2	1,8	5,8	5,6	3,4
4	Bandung	21,9	2,7	27,2	10,8	21,8
5	Garut	38,8	9,9	10	11	23,4
6	Tasikmalaya	27,3	8,9	4,7	8,2	5,3
7	Ciamis	36,5	0,8	8,7	23,8	22,8
8	Kuningan	50,6	0,3	5,9	9	14,4
9	Cirebon	2,6	0	8,8	7,5	8,5
10	Majalengka	30,6	0,2	5,4	7,1	12,7
11	Sumedang	35,2	0,2	8,2	6,5	24,2



DOI: 10.52362/jisamar.v6i4.939

Ciptaan disebarluaskan di bawah [Lisensi Creative Commons Atribusi 4.0 Internasional](https://creativecommons.org/licenses/by/4.0/).

12	Indramayu	0,5	0,2	4,3	3	1,1
13	Subang	9	0,4	7,8	7,6	2,5
14	Purwakarta	5,9	0,2	1	3,8	12
15	Karawang	1,5	0	12,2	9,4	5,9
16	Bekasi	0,8	0,3	9,6	2,9	3,7
17	Bandung Barat	24,5	0,3	3	7,7	7,7
18	Pangandaran	4,7	3,3	2,7	10,7	14,1
19	Kota Bogor	60,7	0,2	7,3	26,6	15,1
20	Kota Sukabumi	5,4	7,1	2,2	5	4,1
21	Kota Bandung	4	0	5,4	4,5	36,2
22	Kota Cirebon	0,7	0,1	1,4	3,5	4,8
23	Kota Bekasi	0,7	0	8	1,4	2,3
24	Kota Depok	1,1	0	1,4	1,2	1,5
25	Kota Cimahi	4,7	0,3	6,1	5,8	13,5
26	Kota Tasikmalaya	8,5	0,4	3,5	11,2	3,6
27	Kota Banjar	4,7	0,6	0,7	5,7	3,6

2. Inisialisasi

Dalam tahap ini dilakukan penentuan titik tengah atau centroid awal dengan mengambil nilai terkecil (min) untuk *cluster* tingkat rendah (C1), nilai rata-rata (*average*) untuk *cluster* tingkat sedang (C2) dan nilai terbesar (*maks*) untuk *cluster* tingkat tinggi (C3).

Tabel 3 *Centroid* Awal

	Cluster	A	B	C	D	E
Cluster Rendah	C1	0,5	0	0,7	1,2	1,1
Cluster Sedang	C2	20,97	1,83	7,24	11,03	11,94
Cluster Tinggi	C3	89	9,9	27,2	69,5	42

3. Iterasi 1 (satu)

Hasil perhitungan jarak setiap data ke *centroid* terdekat untuk menentukan kelompok *K-Means* baru berdasarkan jarak terdekat dari centroid.

Rumus *Euclidean Distance* :

$$d_{Euclidean}(x,y) = \sqrt{\sum_i (x_i - y_i)^2}$$

Keterangan :

d : Jarak *Euclidean*

i : Banyak data

c : *Centroid*

x : Data

Perhitungan jarak *centroid* dilakukan menggunakan *Microsoft Excel* dengan memasukan rumus diatas, berikut tabel hasil perhitungan pada iterasi 1 :

Tabel .4 Jarak *Centroid* Iterasi 1

Kabupaten/Kota	C1	C2	C3
Bogor	107,281406	85,00397579	32,85102738



DOI: 10.52362/jisamar.v6i4.939

Ciptaan disebarluaskan di bawah [Lisensi Creative Commons Atribusi 4.0 Internasional](https://creativecommons.org/licenses/by/4.0/).

Sukabumi	102,649111	72,03530384	42,72481714
Cianjur	17,3317627	14,04147428	106,7547657
Bandung	41,0871026	23,29345402	91,69503803
Garut	47,3784761	23,0683116	81,14240568
Tasikmalaya	29,6663109	15,70665464	97,05009016
Ciamis	48,3972107	23,43243479	75,08954654
Kuningan	52,6770348	32,04654428	80,26468713
Cirebon	12,8245858	22,34652098	113,4362376
Majalengka	33,1286885	16,00659239	93,44592019
Sumedang	42,685829	19,19122195	87,38060426
Indramayu	4,02988834	25,27043727	120,6060115
Subang	12,8740048	17,34032872	110,7179751
Purwakarta	12,4442758	21,55089325	113,5897443
Karawang	14,9509197	23,5249017	113,5538639
Bekasi	9,43610089	24,99221879	118,6743865
Bandung Barat	25,8300213	13,64628154	99,16541736
Pangandaran	17,081569	20,7637906	109,4812769
Kota Bogor	67,1475986	44,31409482	62,08872684
Kota Sukabumi	10,0054985	21,80502236	114,9715617
Kota Bandung	35,7384947	21,93032148	109,8038706
Kota Cirebon	4,41814441	25,46474033	119,5759591
Kota Bekasi	7,40337761	25,52091299	120,3214029
Kota Depok	1,00498756	25,93563186	121,6355211
Kota Cimahi	14,8932871	20,95822034	111,8650973
Kota Tasikmalaya	13,3510299	17,74471752	109,5702514
Kota Banjar	6,67083203	21,93050387	115,9320059

Tahap selanjutnya adalah menentukan jarak terpendek dari tiap data. Jarak terpendek ialah nilai terkecil di antara C1, C2, dan C3 pada tiap-tiap data. Berikut table *centroid* jarak terpendek pada tiap *cluster* :

Tabel 5. Jarak *Centroid*, Jarak Terpendek dan *Cluster* pada Iterasi 1

Kabupaten/Kota	C1	C2	C3	Jarak Terpendek	Cluster
Bogor	107,281406	85,00397579	32,85102738	32,85102738	3
Sukabumi	102,649111	72,03530384	42,72481714	42,72481714	3
Cianjur	17,3317627	14,04147428	106,7547657	14,04147428	2
Bandung	41,0871026	23,29345402	91,69503803	23,29345402	2
Garut	47,3784761	23,0683116	81,14240568	23,0683116	2
Tasikmalaya	29,6663109	15,70665464	97,05009016	15,70665464	2
Ciamis	48,3972107	23,43243479	75,08954654	23,43243479	2
Kuningan	52,6770348	32,04654428	80,26468713	32,04654428	2
Cirebon	12,8245858	22,34652098	113,4362376	12,82458576	1
Majalengka	33,1286885	16,00659239	93,44592019	16,00659239	2
Sumedang	42,685829	19,19122195	87,38060426	19,19122195	2
Indramayu	4,02988834	25,27043727	120,6060115	4,029888336	1
Subang	12,8740048	17,34032872	110,7179751	12,87400482	1
Purwakarta	12,4442758	21,55089325	113,5897443	12,44427579	1
Karawang	14,9509197	23,5249017	113,5538639	14,9509197	1
Bekasi	9,43610089	24,99221879	118,6743865	9,43610089	1
Bandung Barat	25,8300213	13,64628154	99,16541736	13,64628154	2
Pangandaran	17,081569	20,7637906	109,4812769	17,08156901	1
Kota Bogor	67,1475986	44,31409482	62,08872684	44,31409482	2



DOI: 10.52362/jisamar.v6i4.939

Ciptaan disebarluaskan di bawah [Lisensi Creative Commons Atribusi 4.0 Internasional](https://creativecommons.org/licenses/by/4.0/).

Kota Sukabumi	10,0054985	21,80502236	114,9715617	10,00549849	1
Kota Bandung	35,7384947	21,93032148	109,8038706	21,93032148	2
Kota Cirebon	4,41814441	25,46474033	119,5759591	4,418144407	1
Kota Bekasi	7,40337761	25,52091299	120,3214029	7,403377608	1
Kota Depok	1,00498756	25,93563186	121,6355211	1,004987562	1
Kota Cimahi	14,8932871	20,95822034	111,8650973	14,89328708	1
Kota Tasikmalaya	13,3510299	17,74471752	109,5702514	13,35102992	1
Kota Banjar	6,67083203	21,93050387	115,9320059	6,670832032	1

Setelah hasil centroid dari tiap data didapatkan, kemudian langkah selanjutnya menghitung masing-masing *K-Means* baru berdasarkan kelompok data dalam setiap *cluster*-nya. Perhitungan algoritma *K-Means* akan terus berlanjut sampai setiap *cluster* dalam jarak *centroid* sudah tidak ada yang berubah.

Tabel IV.6. *Centroid* Baru Iterasi 1

Centroid Iterasi 1	A	B	C	D	E
C1	3,63	0,98	4,98	5,62	5,72
C2	31,48	2,3	8,33	10,98	17
C3	84,55	5,55	17,05	49,2	27,15

4. Iterasi 6 (enam)

Berdasarkan perhitungan iterasi 2 dan iterasi 6 disetiap cluster dalam jarak *centroid* memiliki kesamaan atau tidak ada perubahan, maka perhitungan algoritma *K-Means Clustering* dihentikan di iterasi ke-6. Dengan hasil *centroid* akhir iterasi 6 pada *cluster* 1 terdapat 17 data, *cluster* 2 terdapat 9 data, dan *cluster* 3 terdapat 1 data. Hasil centroid akhir iterasi 6 pada tabel 7 dan hasil akhir pada Tabel 8.

Tabel 7. Jarak *centroid*, Jarak Terpendek, dan *Cluster* Iterasi 6

Kabupaten/Kota	C1	C2	C3	Jarak Terpendek	Cluster
Bogor	81,87740714	70,99198617	52,07456193	52,07456193	3
Sukabumi	8,734632219	16,82878784	91,9403611	8,734632219	1
Cianjur	24,1439392	19,73737825	83,50682607	19,73737825	2
Bandung	27,02849977	16,34325855	73,64835368	16,34325855	2
Garut	14,78647355	12,29854056	82,9943974	12,29854056	2
Tasikmalaya	27,57204744	18,22953922	65,05835842	18,22953922	2
Ciamis	34,00673169	25,06208491	68,85557349	25,06208491	2
Kuningan	14,74536537	25,04080869	99,98604903	14,74536537	1
Cirebon	14,02104846	8,461566049	81,02573665	8,461566049	2
Majalengka	22,55530536	14,29972377	79,19962121	14,29972377	2
Sumedang	20,15603632	30,96843716	105,5396608	20,15603632	1
Indramayu	12,07335082	21,88899495	95,5932006	12,07335082	1
Subang	12,86576076	23,59919702	100,9594968	12,86576076	1
Purwakarta	17,38366475	26,80556099	99,48612969	17,38366475	1
Karawang	18,92574437	29,24339413	104,460806	18,92574437	1
Bekasi	9,297408241	11,00972752	85,02523155	9,297408241	2
Bandung Barat	13,19271769	22,44067958	97,23672146	13,19271769	1
Pangandaran	47,74040008	37,99942236	48,91809481	37,99942236	2
Kota Bogor	15,83476555	25,48140695	100,8059026	15,83476555	1
Kota Sukabumi	28,00038928	31,85997646	103,9664369	28,00038928	1
Kota Bandung	18,70983164	29,81452834	105,224997	18,70983164	1
Kota Cirebon	19,90416539	30,41542536	105,798724	19,90416539	1



Kota Bekasi	20,52169096	31,4537454	106,6701927	20,52169096	1
Kota Depok	12,45936595	22,70458324	99,72356793	12,45936595	1
Kota Cimahi	12,10214031	22,0585607	94,24155135	12,10214031	1
Kota Tasikmalaya	15,66951818	26,56113891	101,0571126	15,66951818	1
Kota Banjar	22,87500383	33,73573921	108,6653119	22,87500383	1

5. Hasil Clustering Data

Tabel 8. Hasil Akhir

No	Kabupaten/Kota	Cluster 1	Cluster 2	Cluster 3
1	Bogor			1
2	Sukabumi	1		
3	Cianjur		1	
4	Bandung		1	
5	Garut		1	
6	Tasikmalaya		1	
7	Ciamis		1	
8	Kuningan	1		
9	Cirebon		1	
10	Majalengka		1	
11	Sumedang	1		
12	Indramayu	1		
13	Subang	1		
14	Purwakarta	1		
15	Karawang	1		
16	Bekasi		1	
17	Bandung Barat	1		
18	Pangandaran		1	
19	Kota Bogor	1		
20	Kota Sukabumi	1		
21	Kota Bandung	1		
22	Kota Cirebon	1		
23	Kota Bekasi	1		
24	Kota Depok	1		
25	Kota Cimahi	1		
26	Kota Tasikmalaya	1		
27	Kota Banjar	1		
Jumlah		17	9	1
Total		27		

Perhitungan *K-Means Clustering* menghasilkan *cluster* tingkat rendah (C1) adalah daerah Sukabumi, Kuningan, Sumedang, Indramayu, Subang, Purwakarta, Karawang, Bandung Barat, Kota Bogor, Kota Sukabumi, Kota Bandung, Kota Cirebon, Kota Bekasi, Kota Depok, Kota Cimahi, Kota Tasikmalaya, Kota Banjar. Kemudian *cluster* tingkat sedang (C2) diantaranya Cianjur, Bandung, Garut, Tasikmalaya, Ciamis, Majalengka, Sumedang, Bekasi, Pangandaran. Dan untuk *cluster* tingkat tinggi/rawan (C3) ialah Kabupaten Bogor.

3.2 Penerapan Algoritma *K-Means* pada Aplikasi *RapidMiner*



DOI: 10.52362/jisamar.v6i4.939

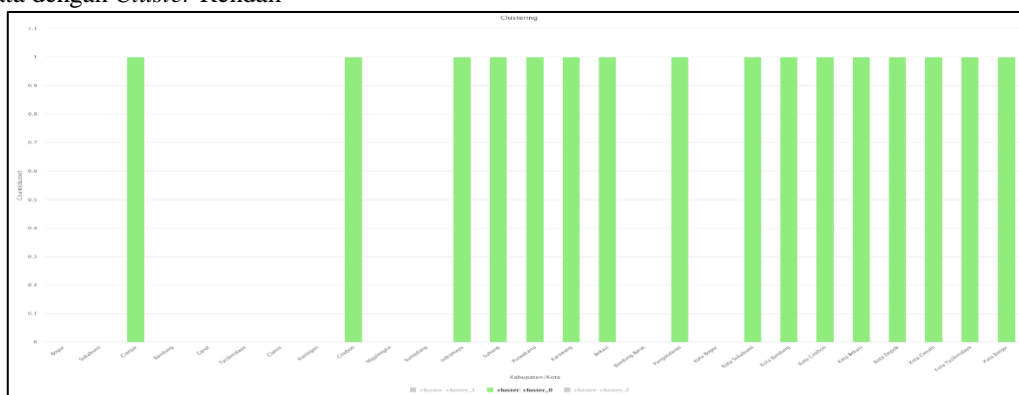
Ciptaan disebarluaskan di bawah [Lisensi Creative Commons Atribusi 4.0 Internasional](https://creativecommons.org/licenses/by/4.0/).

Penerapan pengolahan data bencana alam dengan K-Means pada penelitian ini dilakukan dengan menggunakan aplikasi RapidMiner



Gambar 2. Tampilan Visualisasi data pada aplikasi Rapid Miner

1. Data dengan *Cluster Rendah*



Gambar 3. *Cluster Rendah*

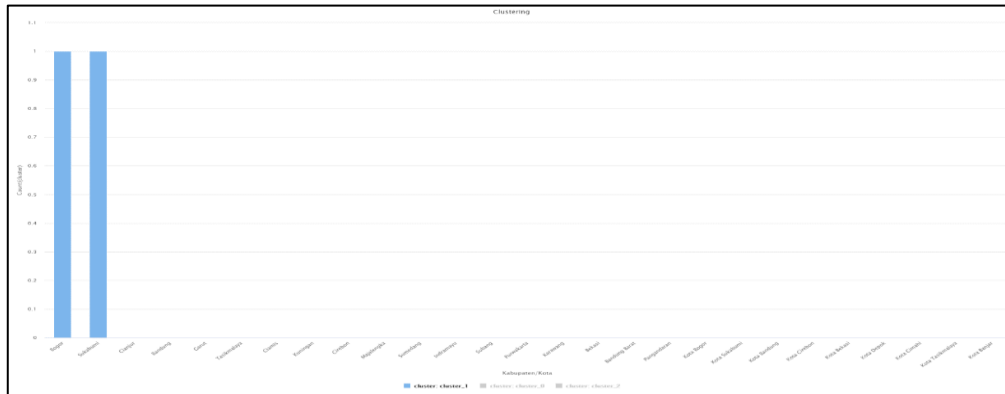
Gambar 3 menampilkan hasil perhitungan algoritma k-means dengan kelompok *cluster-0* atau *cluster rendah* diantaranya Cianjur, Cirebon, Indramayu, Subang, Purwakarta, Karawang, Bekasi, Pangandaran, Kota Sukabumi, Kota Bandung, Kota Cirebon, Kota Bekasi, Kota Depok, Kota Cimahi, Kota Tasikmalaya, dan Kota Banjar.

2. Data dengan *Cluster Sedang*



DOI: 10.52362/jisamar.v6i4.939

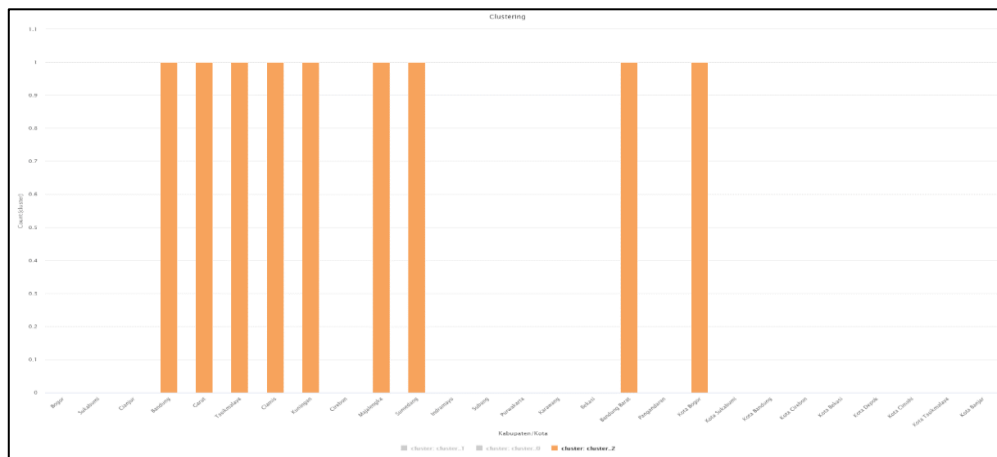
Ciptaan disebarluaskan di bawah [Lisensi Creative Commons Atribusi 4.0 Internasional](https://creativecommons.org/licenses/by/4.0/).



Gambar 4. Cluster Sedang

Gambar 4. menampilkan hasil perhitungan algoritma k-means dengan kelompok *cluter-1* atau *cluster* sedang diantaranya Bogor, dan Sukabumi.

3. Data dengan Cluster tinggi



Gambar 5. Cluster Tinggi

Gambar 5 menampilkan hasil perhitungan algoritma k-means dengan kelompok *cluster-2* atau *cluster* tinggi (rawan) diantaranya Bandung, Garut, Tasikmalaya, Ciamis, Kuningan, Majalengka, Sumedang, Bandung Barat, Kota Bogor.

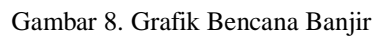
3.3 Grafik Bencana Alam Tiap Kabupaten/Kota

1. Grafik Bencana Tanah Longsor

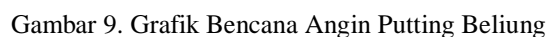


DOI: 10.52362/jisamar.v6i4.939

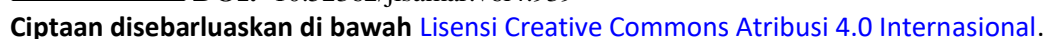
Ciptaan disebarluaskan di bawah [Lisensi Creative Commons Atribusi 4.0 Internasional](https://creativecommons.org/licenses/by/4.0/).

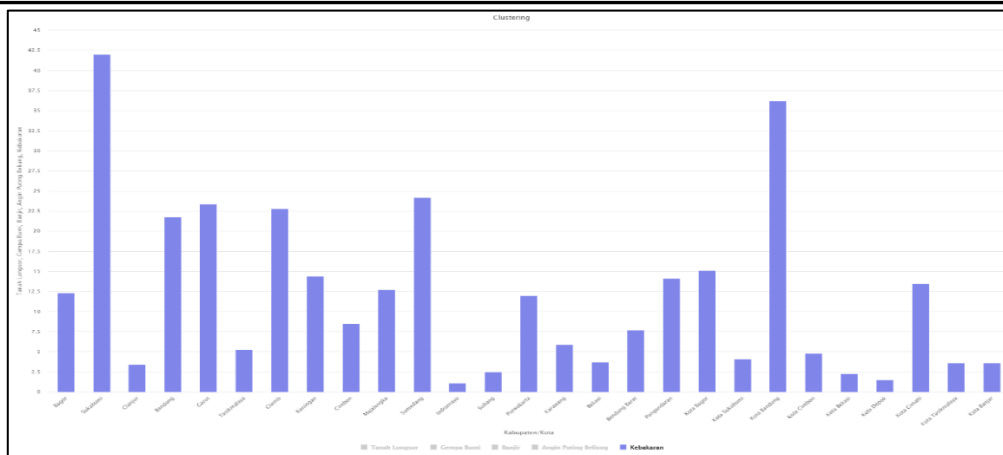


4. Grafik Bencana Angin Puting Beliung



5. Grafik Bencana Kebakaran Lahan dan Bangunan





Gambar 10. Grafik Bencana Kebakaran Lahan dan Bangunan

Gambar 10. menampilkan grafik bencana alam Kebakaran Lahan dan Bangunan dengan tingkat kerawanan diatas 50% diantaranya terdapat pada daerah Sukabumi, Garut, Ciamis, Sumedang, dan Kota Bandung.

IV. KESIMPULAN

Berdasarkan hasil penelitian penulis terhadap kumpulan data bencana alam pada 27 Kabupaten/Kota di Provinsi Jawa Barat tahun 2012 sampai 2021 menggunakan Algoritma *K-Means Clustering* untuk prediksi daerah rawan bencana alam dengan 5 variable bencana alam yaitu Tanah Longsor, Gempa Bumi, Banjir, Angin Putting Beliung dan Kebakaran. Pengolahan data menggunakan *Microsoft Excel* dan aplikasi *RapidMiner* untuk memproses dan menentukan nilai *centroid* dalam 3 (tiga) *cluster*, diantaranya C1 (*cluster* tingkat rendah), C2 (*cluster* tingkat sedang), C3 (*cluster* tingkat tinggi/rawan).

Dari hasil perhitungan *K-Means Clustering* menggunakan aplikasi *RapidMiner* menghasilkan *cluster* tingkat rendah (C0) diantaranya terdapat pada daerah Cianjur, Cirebon, Indramayu, Subang, Purwakarta, Karawang, Bekasi, Pangandaran, Kota Sukabumi, Kota Bandung, Kota Cirebon, Kota Bekasi, Kota Depok, Kota Cimahi, Kota Tasikmalaya, dan Kota Banjar. Sedangkan untuk *cluster* sedang (C1) berada di Bogor, dan Sukabumi. Selanjutnya untuk *cluster* tinggi/rawan (C2) berada pada daerah Bandung, Garut, Tasikmalaya, Ciamis, Kuningan, Majalengka, Sumedang, Bandung Barat, dan Kota Bogor.

Dari hasil perhitungan pada aplikasi *RapidMiner* terdapat perhitungan tingkat kerawanan bencana alam dengan presentase diatas 50% pada tiap daerah diantaranya bencana alam Tanah Longsor terdapat pada daerah Bogor, Sukabumi, Kuningan, dan Kota Bogor. Bencana alam Gempa Bumi terdapat pada daerah Sukabumi, Garut, Tasikmalaya, dan Kota Sukabumi. Bencana alam Banjir terdapat pada daerah Kabupaten Bogor, dan Bandung. Bencana alam Angin Putting Beliung terdapat di daerah Kabupaten Bogor. Dan bencana alam Kebakaran terdapat di daerah Sukabumi, Garut, Ciamis, Sumedang, dan Kota Bandung.

REFERENSI

- [1] D. Oleh, B. Nasional, dan P. Bencana, "IRBI INDEKS RISIKO BENCANA INDONESIA," 2018.
- [2] M. Iqbal Ramadhan, "PENERAPAN DATA MINING UNTUK ANALISIS DATA BENCANA MILIK BNPB MENGGUNAKAN ALGORITMA K-MEANS DAN LINEAR REGRESSION," 2017.



DOI: 10.52362/jisamar.v6i4.939

Ciptaan disebarluaskan di bawah [Lisensi Creative Commons Atribusi 4.0 Internasional](https://creativecommons.org/licenses/by/4.0/).

-
- [3] M. B. Santoso, A. Buchari, dan I. Darmawan, “MEKANISME MASYARAKAT LOKAL DALAM MENGENALI BENCANA DI KABUPATEN GARUT,” *Share : Social Work Journal*, vol. 8, no. 2, hlm. 142, Jan 2019, doi: 10.24198/share.v8i2.18885.
 - [4] J. Oktaviani, S. Sari, D. Taufan, dan H. Akbar, “MEKANISME PENANGGULANGAN BENCANA ALAM OLEH PALANG MERAH INDONESIA (PMI) KABUPATEN BANDUNG,” 2019.
 - [5] A. Nur Khomarudin, “Teknik Data Mining : Algoritma K-Means Clustering,” 2018. [Daring]. Available: <https://agusnkhom.wordpress.com>
 - [6] Anjar Wanto dan dkk, *Data Mining : Algoritma dan Implementasi*, 1 ed. Medan: Yayasan Kita Menulis, 2020.
 - [7] M. Murdiaty, A. Angela, dan C. Sylvia, “Pengelompokan Data Bencana Alam Berdasarkan Wilayah, Waktu, Jumlah Korban dan Kerusakan Fasilitas Dengan Algoritma K-Means,” *JURNAL MEDIA INFORMATIKA BUDIDARMA*, vol. 4, no. 3, hlm. 744, Jul 2020, doi: 10.30865/mib.v4i3.2213.
 - [8] F. N. Dhewayani dkk., “Implementasi K-Means Clustering untuk Pengelompokan Daerah Rawan Bencana Kebakaran Menggunakan Model CRISP-DM,” *Jurnal Teknologi dan Informasi*, vol. 12, no. 1, Mar 2022, doi: 10.34010/jati.v12i1.
 - [9] Halilintar R dan Farina Nur I, “Penerapan Algoritma K-Means Clustering Untuk Prediksi Prestasi Nilai Akademik Mahasiswa,” *Jurnal Sains dan Informatika*, vol. 4, no. 2, 2018.
 - [10] K. Fatmawati dan A. P. Windarto, “DATA MINING: PENERAPAN RAPIDMINER DENGAN K-MEANS CLUSTER PADA DAERAH TERJANGKIT DEMAM BERDARAH DENGUE (DBD) BERDASARKAN PROVINSI,” 2018. [Daring]. Available: <https://www.depkes.go.id/>.
 - [11] Alfasaleh, “K Means Clustering : Contoh Sederhana Penerapan Algoritma K-Means Clustering,” 2019. <https://www.alfasoleh.com/2019/11/k-means-clustering-contoh-sederhana.html> (diakses Jul 04, 2022).
 - [12] S. Ag. , M. Pd. I. Iwan Hermawan, *METODOLOGI PENELITIAN PENDIDIKAN KUANTITATIF, KUALITATIF DAN MIXED METHODE*, 1 ed. Kuningan: Hidayatul Quran Kuningan, 2019.

