

Analisis Sebaran Titik Rawan Bencana dengan *K-Means Clustering* dalam Penanganan Bencana

Teguh Iman Hermanto¹, Yusuf Muhyidin²

Sekolah Tinggi Teknologi Wastukencana

Jl. Cikopak No. 53, Sadang, Purwakarta, Jawa Barat

teguh@stt-wastukencana.ac.id

Abstract

Puwakata Regency has fertile land, agricultural products and abundant natural resources. However, the area is also vulnerable to disaster risk. Based on the data collected, the disasters that occurred in Puwakata Regency included several categories, namely landslides, droughts, hurricanes and floods. The trend of increasing numbers of disasters requires further investigation to prevent an increase in the number of victims. Given the large amount of data available, this information can be obtained through data mining analysis methods. For natural disaster data, the clustering method in data mining is very useful for grouping disaster data based on the same characteristics, so that it can be used as a basis for classifying future disaster events. The *k-means* algorithm is a model used to form clusters by measuring how close it is to the data set. Therefore, in terms of the location of the disaster, the type of disaster and its impact on the disaster, it is hoped that this research can use the clustering technique with the *k-means* algorithm to classify disaster-prone points. The results obtained 3 clusters, namely, the type of drought disaster is cluster 0, the type of landslide is cluster 1, and the type of landslide is cluster 2. After forming three clusters, disaster management strategies are drawn up at each disaster-prone point in the Purwakarta area.

Keywords: Clustering, *k-means*. Datamining, disasters

Abstrak

Kabupaten Puwakata memiliki tanah yang subur, hasil pertanian dan sumber daya alam yang melimpah. Namun kawasan tersebut juga rentan terhadap risiko bencana. Berdasarkan data yang terkumpul, bencana yang terjadi di Kabupaten Puwakata meliputi beberapa kategori yaitu longsor, kekeringan, angin topan dan banjir. Tren peningkatan angka bencana memerlukan penyelidikan lebih lanjut untuk mencegah peningkatan jumlah korban. Mengingat sejumlah besar data yang tersedia, informasi ini dapat diperoleh melalui metode analisis data mining. Untuk data bencana alam, metode clustering pada data mining sangat berguna untuk pengelompokan data bencana berdasarkan karakteristik yang sama, sehingga dapat digunakan sebagai dasar untuk pengklasifikasian kejadian bencana yang akan datang. Algoritma *k-means* adalah model yang digunakan untuk membentuk cluster dengan mengukur seberapa dekat jarak yang ada dengan kumpulan data. Oleh karena itu, ditinjau dari lokasi bencana, jenis bencana dan dampaknya terhadap bencana, diharapkan penelitian ini dapat menggunakan teknik clustering dengan algoritma *k-means* untuk mengelompokkan titik-titik rawan bencana. Hasil penelitian didapatkan 3 cluster yaitu, tipe bencana kekeringan adalah cluster 0, jenis bencana longsor adalah cluster 1, dan jenis bencana longsor adalah cluster 2. Setelah terbentuk tiga cluster, disusun strategi penanggulangan bencana di setiap titik rawan bencana di wilayah Purwakarta.

Kata kunci: Datamining, *k-means*, bencana, clustering

1. PENDAHULUAN

Dinas Kebakaran dan Penanggulangan Bencana Kabupaten Purwakarta merupakan dinas yang terlatih dan bertanggung jawab dalam penanganan



kebakaran, selain dilatih untuk menyelamatkan korban kebakaran, mereka juga dilatih untuk menyelamatkan korban bencana alam dan bencana lainnya. Sejak tahun 2017, Biro Penanggulangan Kebakaran dan Bencana telah menjadi badan layanan bersama, yang merupakan tugas eksekutif pemerintah, bertanggung jawab untuk menangani kebakaran dan bencana lainnya termasuk dalam layanan darurat atau penyelamatan / (penyelamatan).

Biro Kebakaran dan Penanggulangan Bencana juga merupakan salah satu instansi yang memberikan informasi tentang kejadian bencana kepada masyarakat. Namun informasi tentang bencana masih bersifat acak, tidak dapat dipahami, dan belum terbukti efektifitasnya, sehingga banyak kendala untuk membuatnya menjadi masalah. Purwakarta memiliki sejarah bencana alam seperti longsor, banjir, kekeringan dan angin topan yang terus meningkat dari tahun ke tahun, sehingga diperlukan informasi data, kemudian informasi data tersebut diolah untuk menentukan strategi yang harus dilaksanakan dan diolah.

Analisis *Clustering* merupakan sebuah teknik pengolahan data untuk mengelompokkan objek dengan mengukur jarak terdekat[1]. Metode *Clustering* yang digunakan untuk menganalisis data titik rawan bencana menggunakan algoritma *K-Means*. Analisis data menggunakan algoritma *K-Means* sudah diterapkan untuk menganalisis daerah pelaku kejahatan pencurian[2], mengelompokkan jumlah penumpang pada alat transportasi umum[3], dan pengelompokan pada transaksi pembelian[4].

2. METODOLOGI PENELITIAN

Penelitian ini melakukan pengelompokan sebaran titik rawan bencana di Kabupaten Purwakarta menggunakan algoritma *K-Means Clustering* dengan tahapan-tahapan sebagai berikut :

a) Pengumpulan Data

Pada tahapan ini dilakukan pengumpulan data-data yang dibutuhkan sebagai pendukung utama. Data yang diambil disesuaikan dengan topik dan batasan permasalahan pada penelitian. Dalam hal ini, data didapatkan dari dinas pemadam kebakaran kabupaten purwakarta:

- 1) Laporan data bencana alam dalam kurun waktu tahun 2017-2019
- 2) Laporan data bencana lainnya dalam kurun waktu 2017-2019

Selain pengumpulan data, dilakukan juga wawancara kepada narasumber terkait kondisi terkini keadaan di wilayah kabupaten purwakarta yang berhubungan dengan clustering sebagai informasi pendukung dalam proses penelitian[5].

b) *Preprocessing Data*

Pada tahap ini akan dilakukan pengolahan data mentah untuk memenuhi kebutuhan analisis. *Preprocessing* data akan melakukan pemilihan atribut, membersihkan baris data dengan nilai kosong, dan menggabungkan data asli yang diperoleh. Output dari proses ini adalah data siap untuk diproses ke tahap *cleaning*.

- c) *Data Cleaning*
- d) Pembersihan data adalah proses pada tahap preprocessing untuk mengisi data blank, mengulang data, mengoreksi data yang tidak sesuai dengan ketentuan atau salah ketik (seperti huruf yang tidak mencukupi dan huruf yang berlebihan), mengubah dan memodifikasi persyaratan konfirmasi data. Data yang diolah merupakan data yang tidak sesuai dengan ketentuan atau kesalahan pencetakan. Jaga konsistensi data, sortir data dengan kapitalisasi tidak beraturan, dan ubah format angka dan huruf sesuai kebutuhan.[6].
- e) *Data Transformation*
 Metode *K-means clustering* merupakan metode yang umum digunakan apabila data yang digunakan dalam bentuk digital maka diperlukan proses konversi ini. Proses konversi merupakan tahap mengubah data atribut selain angka menjadi nilai numerik agar datanya dapat diolah dengan menggunakan algoritma *K-Means clustering*.
- f) Implementasi *K-Means Clustering* dengan *Rapidminer*
 Data yang sudah ditransformasi akan diolah menggunakan aplikasi *Rapidminer* dengan model algoritma *K-Means* agar data dapat divisualisasikan hasil clusternya.
- g) Evaluasi
 Berdasarkan hasil pengolahan data menggunakan algoritma *K-means* dengan aplikasi *Rapidminer*, maka hasil cluster yang terbentuk akan diberikan strategi-strategi penanggulangan bencana sesuai cluster yang terbentuk[7].

3. HASIL DAN PEMBAHASAN

3.1. Pengumpulan Data

Pengumpulan data dilakukan dengan mengambil data, data didapatkan dari dinas pemadam kebakaran kabupaten purwakarta adalah Laporan data bencana alam dan non alam dalam kurun waktu tahun 2017-2019 sebanyak 571 seperti yang terlihat pada tabel 1.

Tabel 1. Hasil pengumpulan data

id	Tanggal bencana	Lokasi	Jenis	Korban	Dampak	Kerusakan
1	16/02/17	Kp. Karang anyar, jatiluhur	Tanah Longsor		Retakan	Ringan
2	17/02/17	Kp. Sidang Sari, Jatiluhur	Tanah Longsor		Retakan	Ringan
3	20/02/17	Kp. Pangadungan sukasari	Tanah Longsor		Retakan	Sedang
4	01/03/17	Kp. Cigonewah, jatiluhur	Tanah Longsor		Retakan	Ringan
5	05/03/17	Gang kamboja, purwakarta	Tanah Longsor		Ambruk	Berat
6	17/03/17	Kp. Cihuni purwakarta	Tanah Longsor		Retakan	Berat

id	Tanggal bencana	Lokasi	Jenis	Korban	Dampak	Kerusakan
7	02/04/17	Kp. Berbedahan tegalwaru	Tanah Longsor		Retakan	Ringan
8	09/05/17	Desa cimahi, campaka	Tanah Longsor		Amblas	Sedang

3.2. Preprocessing Data

Pada tahap ini data yang sudah diseleksi akan dilakukan *preprocessing* dengan cara mengurangi atribut yang tidak digunakan dalam proses data mining atau hanya mengambil atribut yang sesuai untuk dilanjutkan ke tahap berikutnya[8]. Pada proses ini diambil 5 atribut yaitu id, tanggal bencana, lokasi bencana, jenis bencana dan dampak bencana. Hasil dari *Preprocessing Data* dapat dilihat pada Tabel 2.

Tabel 2. Hasil *Preprocessing Data*

id	Tanggal bencana	Lokasi	Jenis	Dampak
1	16/02/17	Kp. Karang anyar, jatiluhur	Tanah Longsor	Retakan
2	17/02/17	Kp. Sidang Sari, Jatiluhur	Tanah Longsor	Retakan
3	20/02/17	Kp. Pangadungan sukasari	Tanah Longsor	Retakan
4	01/03/17	Kp. Cigonewah, jatiluhur	Tanah Longsor	Retakan
5	05/03/17	Gang kamboja, purwakarta	Tanah Longsor	Ambruk
6	17/03/17	Kp. Cihuni purwakarta	Tanah Longsor	Retakan
7	02/04/17	Kp. Berbedahan tegalwaru	Tanah Longsor	Retakan
8	09/05/17	Desa cimahi, campaka	Tanah Longsor	Amblas

3.3. Data Cleaning

Pada dataset yang telah melewati *Preprocessing* data selanjutnya adalah tahapan data cleaning yang dimana penulis menemukan beberapa data yang inkonsisten dan data yang redundan/duplikasi[9]. Maka penulis melakukan perubahan data yang inkonsisten sehingga sesuai dan menghapus data yang redundan/duplikasi seperti yang terlihat pada gambar dibawah ini

194	26 September 2019	kp. Krajan desa marancang kecamatan babakancikao
195	27 September 2019	kp. Gunung hejo desa gunung hejo kecamatan darangdan
196	27 September 2019	Kp. Pangkalan Desa Palinggihan Kecamatan Plered Kabupaten Purwakarta
197	28 September 2019	esa marancang kecamatn babakancikao
198	28 September 2019	eesa margasari kecamatan pasawahan
199	28 September 2019	eesa darangdan kecamatan darangdan
200	28 September 2019	kp. Cilimus desa darangdan kecamatan darangdan
201	28 September 2019	kp. Bungur jaya dusun desa pasawahan kecamatan pasawahan
202	30 September 2019	kp. Cicariu desa nagrak kecamatan darangdan
203	30 September 2019	kp. Cikokosan desa nagrak kecamtan darangdan
204	30 September 2019	kp. Krajan desa parakan lima kecamatan jatiluhur

Gambar 1. Data inkonsisten

3.4. Data Transformation

Penulis memperoleh atribut sesuai kebutuhan untuk memenuhi persyaratan memasuki proses algoritma *K-Means*. Pada algoritma *clustering K-Means*, data yang dapat diolah hanya data dalam bentuk digital, sedangkan

data dengan nilai dalam atribut masih berupa teks / kategori, sehingga penulis harus melakukan konversi data tersebut ke bentuk digital.[10].

Langkah pertama yang penulis lakukan adalah merubah alamat atau lokasi bencana menjadi titik koordinat latitude longitude dengan bantuan google maps dan menginisilaisasi jenis bencana dan dampak bencana kedalam bentuk angka[11].

No	Lokasi_Bencana	Longitude (Garis Bujur)	Latitude (Garis Lintang)
1	Kp. Karang Anyar Rt.18 Rw.05 Desa Jati mekar Kecamatan Jatiluhur Kabupaten Purwakarta	107.39736	-6.52687
2	Kp. Sindang Sari Rt. 20 Rw.07 Desa Parakanlima Kecamatan Jatiluhur Kabupaten Purwakarta	107.44794	-6.60859
3	Kp. Pagadungan Rt. 02 Rw.02 Desa Kertamanah Kecamatan Sukasari	107.45242	-6.20448
4	Gang Kamboja I Rt. 36 Rw. 07 Kelurahan Nagri Tengah Kecamatan Purwakarta	107.44205	-6.5496
5	Kp. Cigonewah Rt.11 Rw. 02 Desa Jatiluhur Kabupaten Purwakarta	107.56112	-6.93662
6	Kp. Cihuni Rt.10 Rw.05 Desa Cihuni Kecamatan Pasawahan Kabupaten Purwakarta	107.46758	-6.57882
7	Kp. Ciminyak Rt.13 Rw.04 Desa Cirama hilir Kecamatan Maniis	107.26384	-6.68155
8	Desa Cimahi Kecamatan Campaka Rt.01 Rw.01 Kabupaten Purwakarta	107.48315	-6.50514
9	Kp. Berbedahan Desa Tegalwaru Kecamatan Tegalwaru Kabupaten Purwakarta	107.34477	-6.64495
10	Kp. Bedahan Desa Tegalwaru Kecamatan Tegalwaru Kabupaten Purwakarta	107.34499	-6.63901

Gambar 2. Hasil perubahan alamat ke titik kordinat

Untuk tranformasi jenis kategori menjadi numeric atrribut jenis bencana dan dampak bencana bisa dilihat pada table dibawah ini:

Tabel 3. Inisiasi dampak bencana

No	Dampak Bencana	Inisialisasi
1	Bangunan Rusak	1
2	Amblas	2
3	Krisis Air Bersih	3
4	Pohon tumbang	4
5	Tergenangnya rumah warga	5
6	Kerusakan fasilitas Umum	6
7	Gagal Panen	7

Tabel 4. Inisiasi jenis bencana

No	Jenis Bencana	Inisialisasi
1	Tanah Longsor	1
2	Kekeringan	2
3	Angin Putting Beliung	3
4	Banjir	4

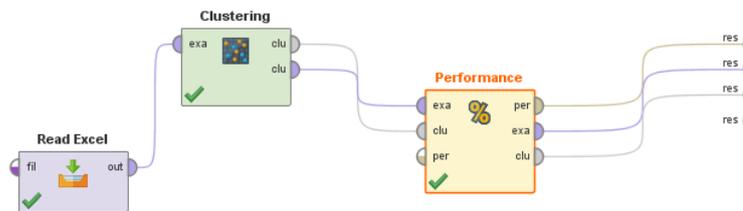
Data lengkap yang akan diproses untuk masuk pada tahapan algoritma K-Means adalah atribut ID, Longitude dan Latitude, jenis bencana, dampak bencana seperti yang terlihat pada tabel di bawah ini[12]

Tabel 5. Dataset akhir untuk proses algoritma *K-Means*

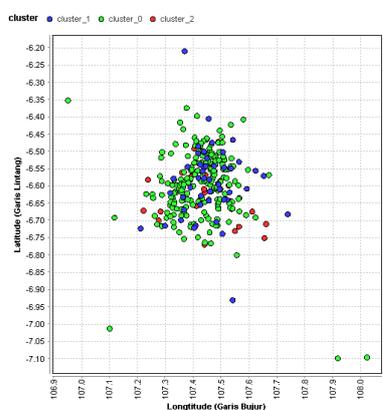
No	Longitude (Garis Bujur)	Latitude (Garis Lintang)	Jenis Bencana	Dampak Bencana
1	107.39736	-6.526873	1	1
2	107.447942	-6.608592	1	1
3	107.452419	-6.204482	1	1
4	107.442049	-6.549602	1	1
5	107.561121	-6.936617	1	1
6	107.467581	-6.578822	1	1
7	107.263839	-6.681549	1	1
8	107.483147	-6.505144	1	2
9	107.344773	-6.644947	2	3
10	107.344985	-6.639006	2	3

3.5. Implementasi *K-means Clustering* dengan *Rapidminer*

Pada tahap ini data yang sudah ditransformasi mulai dibuat model pengolahan data menggunakan aplikasi *rapidminer* dengan modul algoritma *K-Means Clustering*[13]. Dari model yang sudah dirancang dengan tahapan hasil transformasi data (*read excel*) dan operator *Clustering* maka didapatkan hasil pengolahan data titik rawan bencana menggunakan algoritma *K-Means* dapat dilihat pada gambar berikut.

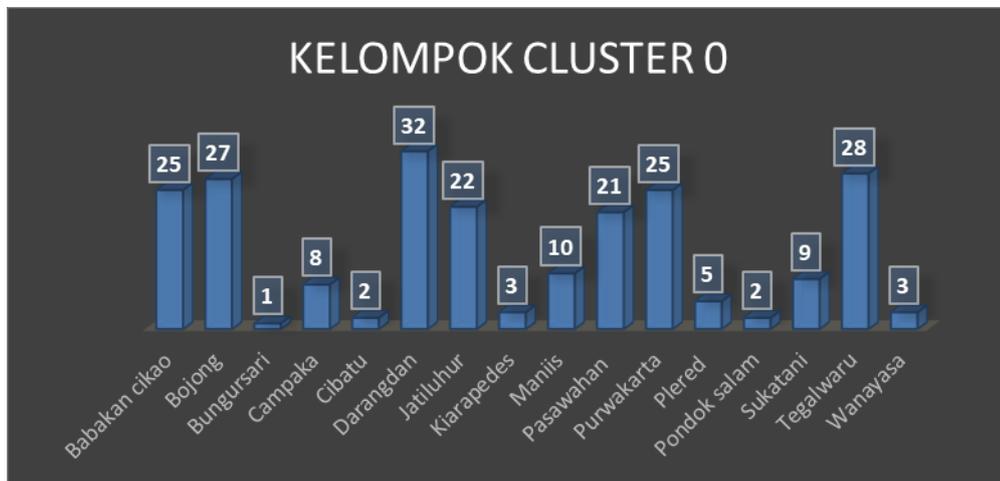


Gambar 3 Model *K-Means clustering*

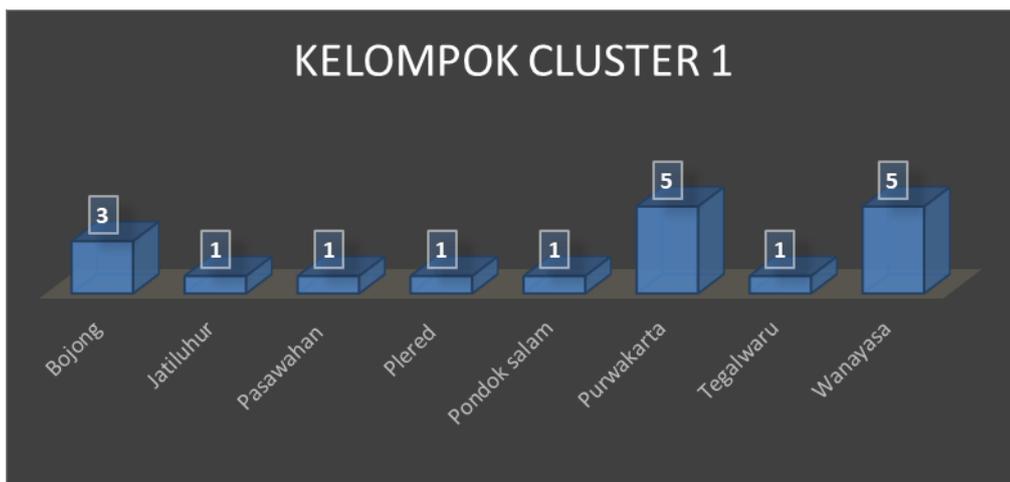


Gambar 4. Sebaran cluster berdasarkan garis lintang dan bujur

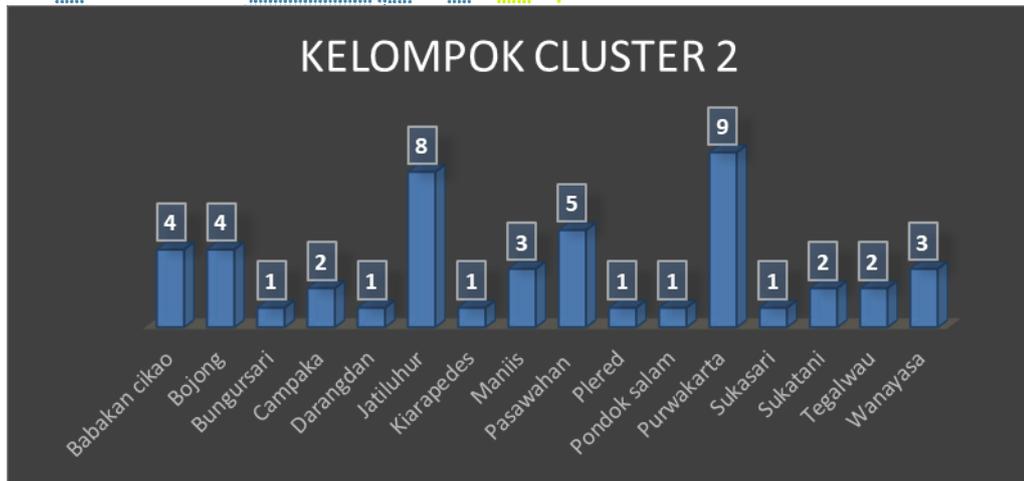
Informasi Yang didapat dari hasil pengolahan dataset dengan menggunakan aplikasi rapid miner adalah bahwa daerah yang termasuk kedalam cluster_0 adalah daerah intensitas bencana rendah, data daerah yang masuk kedalam cluster_1 adalah daerah dengan intensitas bencana sedang dan cluster_2 intensitas bencana tinggi. Informasi Yang didapat dari hasil pengolahan dataset dengan menggunakan aplikasi rapid miner adalah bahwa daerah yang termasuk kedalam cluster_0 adalah daerah intensitas bencana rendah, data daerah yang masuk kedalam cluster_1 adalah daerah dengan intensitas bencana sedang dan cluster_2 intensitas bencana tinggi. Berikut adalah Hasil akhir proses kalkulasi, penulis menjumlahkan data daerah percluster[14].



Gambar 5. Grafik hasil kalkulasi cluster 0



Gambar 6. Grafik hasil kalkulasi cluster 1



Gambar 7. Grafik hasil kalkulasi cluster 2

3.6. Evaluasi

Hasil dari pengolahan data bencana dengan menggunakan RapidMiner dan Excel, terdapat sebanyak 3 cluster yaitu Cluster_0, Cluster_1 dan Cluster_2. Dari hasil tersebut maka di cluster_0 jenis bencana yang paling sering terjadi yaitu kekeringan dengan 179 kali, Angin puting Beliung 20 kali dan Tanah longsor 1 kali, cluster_1 Tanah longsor 28 kali, Angin Puting beliung 14 kali, Banjir 6 kali, Cluster_2 Tanah longsor 15 kali, Angin puting beliung 3kali, Banjir 8kali adapun dampak bencana yang sering terjadi pada cluster_0 adalah Krisis Air Bersih serta Pohon Tumbang[15]. Oleh karena itu, strategi yang harus dilaksanakan oleh aparat DPKPB Purwakarta adalah melaksanakan rencana peningkatan kapasitas peralatan dan rencana pencegahan dini, dan mengubahnya menjadi kegiatan seperti pelatihan dan pendidikan sumber daya peralatan yang ada, sosialisasi tentang potensi bencana alam, dan pemberian konsultasi. Dan upaya untuk mengurangi risiko bencana pada cluster_0 adalah :

- a) mencari tempat yang aman dan jauh dari pepohonan besar
- b) berusaha untuk berlindung di bawah jembatan atau jalan layang
- c) Tidak mengendarai kendaraan saat akan menghindari dari angin puting beliung
- d) Waspada dari benda-benda yang terbang akibat angin puting beliung

Upaya ini dilakukan untuk mengurangi resiko saat angin puting beliung datang, selain itu upaya mengatasi kekeringan diantaranya :

- a) Menyediakan atau ada pembangunan Embung untuk menampung air dimusim hujan.
- b) Membangun sumur dangkal pada lahan yang mengalami kekeringan.
- c) Pompanisasi dan Pipanisasi, berguna untuk menarik air dari sumber-sumber yang ada baik dari sungai maupun mata air.

Jika dilihat dari data di atas cluster_1 paling sering terjadi bencana tanah longsor maka upaya untuk cluster_1 adalah :

- a) Tidak menebang pohon di lereng perbukitan atau pegunungan
- b) Tidak memotong tebing secara tidak lurus
- c) Membuat Terasering
- d) Lakukan upaya Preventif

Adapun data dari cluster_2 bencana Banjir lebih sering terjadi di banding cluster lainnya, maka upaya yang harus dilakukan adalah :

- a) Membuat saluran air dan terowongan air bawah tanah.
- b) Membuang sampah pad tempatnya.
- c) Menanam banyak pohon di lahan kosong dan sekitar rumah
- d) Tidak mendirikan bangunan dekat sungai.
- e) Penciptaan Green Open Space adalah kawasan yang khusus ditujukan untuk penanaman pohon di daerah perkotaan.
- f) Mendirikan bangunan atau konstruksi pencegahan banjir.

4. SIMPULAN

Berdasarkan hasil penelitian penulis terhadap kumpulan data bencana alam Kabupaten Purwakarta dengan menggunakan metode K-Means Clustering, jumlah atribut yang dipakai 4 yaitu Longitude, Latitude, Jenis Bencana dan Dampak Bencana, Penulis menghasilkan jumlah cluster sebanyak 3 dengan jumlah cluster 0 sebanyak 200 titik daerah, cluster 1 sebanyak 48 titik daerah dan cluster 2 sebanyak 26 titik daerah.

Hasil dari pengolahan data dengan K-Means Clustering menunjukkan bahwa kekeringan dan krisis air bersih masuk kedalam cluster 0, sedangkan untuk cluster 1 menunjukkan bahwa daerah tersebut sering terjadi bencana tanah longsor dan cluster 2 hasil dari clustering menunjukkan bahwa daerah tersebut sering terjadi bencana angin puting beliung.

DAFTAR PUSTAKA

- [1] M. Tanzil Furqon and L. Muflikhah, "Clustering the Potential Risk of Tsunami Using Density-Based Spatial Clustering of Application With Noise (DbSCAN)," *J. Environmental Eng. Sustain. Technol.*, vol. 3, no. 1, pp. 1-8, 2016, doi: 10.21776/ub.jeest.2016.003.01.1.
- [2] R. R. E. G. Sihombing, L. S. Dewi, and E. Arisawati, "Analisis Algoritma Datamining pada Kasus Daerah Pelaku Kejahatan Pencurian Berdasarkan Provinsi," *J-SAKTI (Jurnal Sains Komput. dan Inform.)*, vol. 4, no. 1, p. 77, 2020, doi: 10.30645/j-sakti.v4i1.189.
- [3] A. Supriyatna, I. Carolina, S. Janti, and A. Haidir, "Cluster ing Koridor Transjakarta Berdasarkan Jumlah Penumpang Dengan Algoritma K-Means," *J-SAKTI (Jurnal Sains Komput. dan Inform.)*, vol. 4, no. September, pp. 682-693, 2020.
- [4] S. Susliansyah, H. Sumarno, H. Priyono, and N. Hikmah, "Pengelompokkan Data Pembelian Tinta Dengan Menggunakan Metode



- K-Means,” *J-SAKTI (Jurnal Sains Komput. dan Inform.,* vol. 3, no. 2, p. 381, 2019, doi: 10.30645/j-sakti.v3i2.156.
- [5] P. S. Bhattacharjee, A. K. Md Fujail, and S. A. Begum, “A Comparison of Intrusion Detection by K-Means and Fuzzy C-Means Clustering Algorithm over the NSL-KDD Dataset,” *2017 IEEE Int. Conf. Comput. Intell. Comput. Res. ICCIC 2017*, pp. 1–6, 2018, doi: 10.1109/ICCIC.2017.8524401.
- [6] P. Prihandoko and B. Bertalya, “a Data Analysis of the Impact of Natural Disaster Using K-Means Clustering Algorithm,” *Kursor*, vol. 8, no. 4, p. 169, 2017, doi: 10.28961/kursor.v8i4.109.
- [7] M. S. Basarlan and I. D. Argun, “Classification of a bank data set on various data mining platforms | Bir Banka Müşteri Verilerinin Farklı Veri Madenciliği Platformlarında Siniflandırılması,” *2018 Electr. Electron. Comput. Sci. Biomed. Eng. Meet. EBBT 2018*, pp. 1–4, 2018, doi: 10.1109/EBBT.2018.8391441.
- [8] M. S. Geetha Devasena, R. Kingsy Grace, and G. Gopu, “PDD: Predictive diabetes diagnosis using datamining algorithms,” *2020 Int. Conf. Comput. Commun. Informatics, ICCCI 2020*, pp. 22–25, 2020, doi: 10.1109/ICCCI48352.2020.9104108.
- [9] Prihandoko, Bertalya, and M. I. Ramadhan, “An analysis of natural disaster data by using K-means and K-medoids algorithm of data mining techniques,” *QiR 2017 - 2017 15th Int. Conf. Qual. Res. Int. Symp. Electr. Comput. Eng.*, vol. 2017-Decem, pp. 221–225, 2017, doi: 10.1109/QIR.2017.8168485.
- [10] N. Puspitasari, J. A. Widians, and N. B. Setiawan, “Customer segmentation using bisecting k-means algorithm based on recency, frequency, and monetary (RFM) model,” *J. Teknol. dan Sist. Komput.*, vol. 8, no. 2, pp. 78–83, 2020, doi: 10.14710/jtsiskom.8.2.2020.78-83.
- [11] B. E. V. Comendador, L. W. Rabago, and B. T. Tanguilig, “An educational model based on Knowledge Discovery in Databases (KDD) to predict learner’s behavior using classification techniques,” *ICSPCC 2016 - IEEE Int. Conf. Signal Process. Commun. Comput. Conf. Proc.*, pp. 1–6, 2016, doi: 10.1109/ICSPCC.2016.7753623.
- [12] A. S. Devi, I. K. G. D. Putra, and I. M. Sukarsa, “Implementasi Metode Clustering DBSCAN pada Proses Pengambilan Keputusan,” *Lontar Komput. J. Ilm. Teknol. Inf.*, vol. 6, no. 3, p. 185, 2015, doi: 10.24843/lkjiti.2015.v06.i03.p05.
- [13] S. Sugriyono and M. U. Siregar, “Preprocessing kNN algorithm classification using K-means and distance matrix with students’ academic performance dataset,” *J. Teknol. dan Sist. Komput.*, vol. 8, no. 4, pp. 311–316, 2020, doi: 10.14710/jtsiskom.2020.13874.
- [14] A. Masrur, G. Thakur, K. Sparks, R. Palumbo, and D. J. Peuquet, “Co-location Pattern Mining of Geosocial Data to Characterize Urban Functional Spaces,” *Proc. - 2019 IEEE Int. Conf. Big Data, Big Data 2019*, pp. 4099–4102, 2019, doi: 10.1109/BigData47090.2019.9006263.

- [15] M. Murdiaty, A. Angela, and C. Sylvia, "Pengelompokan Data Bencana Alam Berdasarkan Wilayah, Waktu, Jumlah Korban dan Kerusakan Fasilitas Dengan Algoritma K-Means," *J. Media Inform. Budidarma*, vol. 4, no. 3, p. 744, 2020, doi: 10.30865/mib.v4i3.2213.