

Clustering Daerah Rawan Bencana Alam Di Indonesia Berdasarkan Provinsi Dengan Metode *K-Means*

Azizah Baldah¹⁾, After Valent Duarisah²⁾, Rizky Akbar Maulana³⁾

¹⁾²⁾³⁾ Program Studi S1 Sistem Informasi, Institut Teknologi Telkom Purwokerto
Jl. DI Panjaitan No.128, Karangreja, Purwokerto Kidul, Kec. Purwokerto Sel., Kabupaten Banyumas, Jawa Tengah
Email : azizahbaldah@gmail.com¹⁾, valentaja1502@gmail.com²⁾, maullana018@gmail.com³⁾

ABSTRACT

Indonesia is a country located in a strategic geological position and is at the confluence of three world plates. This position causes Indonesia to become a country frequently hit by natural disasters such as earthquakes, landslides and other natural disasters. Based on data obtained from the Regional Disaster Management Agency (BPBD), from year to year the number of natural disasters always increases. The increase in the occurrence of natural disasters has led to the need for further research to find out how vulnerable areas are in Indonesia and can later reduce the risks posed when natural disasters occur. Researchers will use the *K-Means Clustering* method to classify natural disaster areas based on their level of vulnerability to natural disasters. Clustering is a process for grouping data into several clusters or groups. *K-Means* is a method for grouping non-hierarchical data (partitions) which can divide data into two or more groups. The source of this dataset comes from the Central Bureau of Statistics (BPS) and consists of 12 columns and the data used are landslides and no natural disasters. Then a search for the optimal *K* value was carried out using the *elbow* method and *Silhouette Analysis* which produced 2 clusters namely cluster 0 and cluster 1. The results of the clustering indicated that cluster 1 was classified as a disaster-prone area in Indonesia including Aceh, North Sumatra, West Java, Central Java, East Java.

Keywords : Natural disasters, Cluster, *K-Means*, Landslide

ABSTRAK

Indonesia merupakan negara yang terletak pada posisi geologis yang strategis serta berada di pertemuan tiga lempeng dunia. Posisi itu menyebabkan Indonesia menjadi negara yang sering dilanda bencana alam seperti gempa bumi, tanah longsor, dan bencana alam lainnya. Berdasarkan data yang didapat dari Badan Penanggulangan Bencana Daerah (BPBD), dari tahun ke tahun jumlah kejadian bencana alam selalu meningkat. Adanya peningkatan kejadian bencana alam tersebut menyebabkan perlunya penelitian lebih lanjut untuk mengetahui seberapa rawan daerah yang ada di Indonesia serta nantinya dapat mengurangi resiko-resiko yang ditimbulkan ketika terjadi bencana alam. Peneliti akan menggunakan metode *K-Means Clustering* untuk mengelompokkan daerah bencana alam berdasarkan tingkat kerawanan terhadap bencana alam. *Clustering* adalah sebuah proses untuk mengelompokkan data ke dalam beberapa *cluster* atau kelompok. *K-Means* merupakan salah satu metode untuk mengelompokkan data *non hierarki* (partisi) yang dapat membagi data menjadi dua kelompok atau lebih. Sumber dataset ini berasal dari Badan Pusat Statistik (BPS) dan terdiri dari 12 kolom dan data yang digunakan yaitu Tanah longsor dan tidak ada bencana alam. Selanjutnya dilakukan pencarian nilai *K* yang optimal menggunakan metode *elbow* dan *Silhouette Analysis* dimana menghasilkan 2 klaster yakni klaster 0 dan klaster 1. Hasil dari klasterisasi mengindikasikan bahwa klaster 1 dikelompokkan sebagai daerah rawan bencana di indonesia diantaranya Aceh, Sumatera Utara, Jawa Barat, Jawa Tengah, Jawa Timur.

Kata Kunci : Bencana alam, Cluster, *K-Means*, Tanah Longsor



Article History

Received : 21/07/2023
Revised : 28/07/2023
Accepted : 01/08/2023
Online : 02/08/2023



This is an open access article under the
CC BY-SA 4.0 License

1. Pendahuluan

Bencana alam merupakan suatu peristiwa yang berdampak bagi populasi manusia. Indonesia terletak pada cincin Asia Pasifik, dimana adanya aktivitas tektonik di wilayah tersebut. Hal ini menyebabkan Indonesia harus menghadapi resiko dari aktivitas tersebut yang tidak lain adalah bencana alam (Wicaksono and Susetyo 2023). Indonesia termasuk kedalam negara yg memiliki 2 musim yaitu penghujan dan kemarau. Hal tersebut menjadi faktor yg dapat memicu terjadinya bencana di Indonesia. Saat kemarau tiba terjadi peningkatan ancaman seperti kebakaran hutan serta lahan, kekeringan sumber air, namun ketika musim penghujan datang ancaman bencana banjir, tanah longsor serta puting beliung juga kian meningkat (Ramadhani et al. 2022).

Menurut data yang dihimpun dari seluruh Badan Penanggulangan Bencana Daerah (BPBD) selama tahun 2021 telah terjadi 5.402 kejadian bencana, dan 99,5% dari kejadian sepanjang tahun 2021 yang merupakan bencana hidrometeorologi. Jumlah kejadian tersebut didominasi antara lain bencana banjir yang terjadi 1.794 kejadian, 1.577 cuaca ekstrem, 1.321 tanah longsor, 579 kebakaran hutan dan lahan, 91 gelombang pasang dan abrasi, 24 gempa bumi, 15 kekeringan dan 1 erupsi gunung api. Dampak dari kejadian tersebut adalah 728 orang meninggal dunia, 87 orang hilang, 14.915 luka-luka, 7.630.692 menderita dan mengungsi, 158.658 rumah rusak, 4.445 fasilitas umum rusak, 664 kantor rusak dan 505 jembatan rusak (Muhari Abdul 2022).

Dari data yang dikumpulkan, bencana alam yang terjadi di Indonesia terdiri dari beberapa kategori yaitu banjir, tanah longsor, gempa bumi, letusan gunung berapi, puting beliung, dan tsunami yang terjadi di berbagai provinsi di Indonesia. Banyaknya peristiwa bencana alam di Indonesia ini menimbulkan korban, baik korban jiwa maupun korban luka-luka, menghancurkan wilayah sekitarnya serta menghancurkan infrastruktur dan menimbulkan kerugian harta benda (Murdiaty, et al., 2020). Tidak dapat dipungkiri dengan terjadinya bencana alam menimbulkan banyak sekali dampak negatif bagi kehidupan masyarakat, di antaranya timbulnya korban jiwa yang tidak sedikit, hilangnya harta benda, kerusakan lingkungan, dan terganggunya fungsi psikologis para korban bencana alam (Khairul Rahmat et al. 2021).

Berdasarkan permasalahan diatas, Peneliti akan menggunakan metode *K-Means Clustering* untuk mengelompokkan daerah bencana alam berdasarkan tingkat kerawanan terhadap bencana alam. Metode *K-Means Clustering* digunakan untuk mengelompokkan daerah bencana alam sehingga memudahkan dalam menentukan kebijakan dalam merencanakan tindakan pencegahan dan mitigasi bencana. Tujuan dari penelitian ini untuk mengidentifikasi daerah-daerah rawan bencana alam di Indonesia pada tahun 2021 dan memberikan rekomendasi kepada pemerintah untuk fokus pada

daerah-daerah dengan potensi bencana yang tinggi. Hasil penelitian ini dapat menjadi panduan dalam perencanaan kebijakan untuk mengurangi dampak bencana di masa depan.

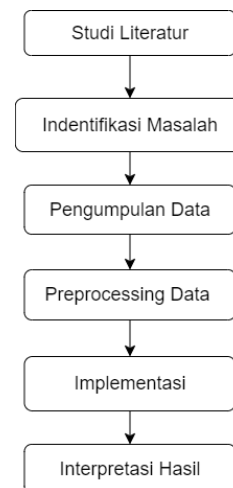
1.1 Pengertian *Clustering*

Clustering adalah proses membagi informasi dalam sesuatu himpunan ke dalam sebagian kelompok yang kesamaan informasinya dalam sesuatu kelompok lebih besar dari pada kesamaan informasi tersebut dengan informasi dalam kelompok lain (Nuryani and Darwis 2021). *Clustering* digunakan untuk pengelompokan maupun mengidentifikasi data atau informasi yang mempunyai karakteristik tertentu (Firman, Halik, and Septiana 2022)

1.2 Pengertian *K-Means*

K-Means merupakan salah satu metode untuk mengelompokkan data *non hierarki* (partisi) yang dapat membagi data menjadi dua kelompok atau lebih. Metode ini membagi data menjadi satu kelompok, dimana data dengan karakteristik yang sama akan dimasukkan ke dalam kelompok yang sama, dan data dengan karakteristik yang berbeda akan dikelompokkan ke dalam kelompok lain. Pengelompokan memiliki tujuan untuk meminimalkan fungsi objektif yang ditetapkan selama proses pengelompokan, biasanya akan mencoba meminimalkan perbedaan dalam kelompok dan memaksimalkan perbedaan antar kelompok. Penelitian terdahulu (Gustientiedina et al., 2019).

Metode pada penelitian ini dapat dilihat pada penjelasan dibawah ini:



Gambar 1. Alur Penelitian

Adapun keterangan pada alur penelitian yaitu:

a. Studi Literatur: Melakukan mempelajari referensi dari berbagai sumber penelitian sebelumnya yang berkaitan dengan metode *K-Means*. Beberapa penelitian sebelumnya yang diteliti oleh penelitian dengan judul “*Clustering* Kejadian Bencana Alam di Jawa Barat Berdasarkan Jenis Bencana Menggunakan *K-Means*” oleh Indah Rosaliyah dkk di tahun 2023. Artikel ini membahas mengelompokkan peristiwa bencana alam di Jawa Barat berdasarkan jenis bencana dengan

menggunakan teknik pengelompokan data mining. Algoritma *K-Means* digunakan untuk pengelompokan data, dan pendekatan Knowledge Discovery in Database (KDD) digunakan untuk tahap desain. Data diproses menggunakan alat RapidMiner, dan proses pengelompokan menghasilkan 6 kelompok. Jenis bencana yang paling sering terjadi adalah tanah longsor, diikuti oleh tornado dan banjir (Rosaliyah, Nurhakim, and Manajemen 2023).

Penelitian dengan judul “Pengelompokan Daerah Bencana Alam Menggunakan Algoritma *K-Means Clustering*” oleh Isni Rinjani dkk pada tahun 2023. Jurnal ini menjelaskan tujuan penelitian yang dilakukan untuk memprediksi daerah bencana alam di Jawa Barat dengan menganalisis data dari situs West Java Open Data. Metode penelitian menggunakan algoritma *K-Means clustering* dalam mengelompokkan daerah bencana alam. Hasil penelitian menunjukkan adanya tiga kelompok daerah bencana alam, yaitu tinggi, sedang, dan rendah. Hasil ini memberikan informasi yang berguna untuk manajemen bencana (Rinjani Isni, Anwar Saeful, and Herdiana Ruli 2023).

Penelitian lainnya yang diteliti oleh Muhamad Firman Al Halik dkk dengan judul “Analisa Data Untuk Prediksi Daerah Rawan Bencana Alam Di Jawa Barat Menggunakan Algoritma *K-Means Clustering*”. Jurnal ini membahas penggunaan teknik data mining, khususnya algoritma *K-Means Clustering*, untuk memprediksi daerah-daerah yang rentan terhadap bencana alam di Jawa Barat, Indonesia. Studi ini bertujuan untuk mengelompokkan data bencana alam ke dalam kluster berdasarkan tingkat risikonya. Hasil studi menunjukkan adanya tiga kluster: risiko rendah, risiko sedang, dan risiko tinggi. Jurnal ini menekankan pentingnya informasi tentang bencana alam untuk manajemen bencana dan peran data mining dalam memproses dan menganalisis data tersebut (Firman, Halik, and Septiana 2022).

b. Identifikasi masalah: mengidentifikasi masalah dengan melakukan pemahaman dan penentuan solusi serta metode penanggulangan. Berdasarkan konteks yang dibahas dalam latar belakang, peneliti memilih untuk menyoroti permasalahan terkait daerah rawan bencana alam di Indonesia.

c. Pengumpulan data: Pengumpulan data berdasarkan data dan keterangan yang terangkum dalam Badan Pusat Statistika (<https://www.bps.go.id/>) sebagai portal data terbuka yang menyediakan data akurat.

d. Preprocessing Data: Proses seleksi data yang sesuai dan memiliki keterkaitan tentang bencana alam dilakukan pada tahap ini. Penulis melakukan pembersihan data dengan penentuan kolom sebagai acuan dalam klasterisasi dan melakukan pembersihan data pada nilai data yang terlalu besar. Dalam tahap ini, peneliti menggunakan kolom “Tanah Longsor” dan “Tidak ada bencana alam” sebagai penentuan kluster daerah rawan bencana di Indonesia.

e. Implementasi: Selanjutnya data akan dicluster ke dalam 2 cluster yakni cluster tidak rawan bencana dan cluster rawan bencana alam. Dalam tahap ini akan dianalisis hasil dari pengolahan data dengan menggunakan metode *K-Means* sesuai proses yang sudah diawal.

f. Interpretasi Hasil: Dalam penelitian ini, pengujian hipotesis melibatkan langkah pengklasteran data menggunakan dua kelompok yang akan diberi label "0" dan "1". Label tersebut akan menunjukkan rentang nilai dari terendah hingga tertinggi. Setelah dilakukan klasterisasi data, hasilnya akan digabungkan dengan variabel-variabel bencana, dan klasifikasi kluster akan ditentukan berdasarkan jenis label yang diberikan. Untuk mengevaluasi keabsahan hasil klasterisasi, digunakan metode *silhouette coefficient* yang akan mengukur kualitas kluster.

2. Pembahasan

2.1 Dataset

Dataset Desa/Kelurahan Menurut Jenis Bencana Alam tahun 2021 yang digunakan untuk pengklusteran terdiri dari 29 kolom data dalam bentuk format CSV. Sumber dataset ini berasal dari Badan Pusat Statistik (BPS) dan terdiri dari 12 kolom. Kolom-kolom tersebut mencakup ID, Provinsi, Tanah Longsor, Banjir, Banjir Bandang, Gempa Bumi, Tsunami, Gelombang Pasang Laut, Angin Puyuh / Angin Puting Beliung / Topan, Gunung Meletus, Kebakaran Hutan, Kekeringan, dan Tidak Ada Bencana Alam.

	Provinsi	Tanah Longsor	Banjir	Banjir Bandang	Gempa Bumi	Tsunami	Gelombang Pasang Laut	Angin Puyuh / Angin Puting Beliung / Topan	Gunung Meletus	Kebakaran Hutan	Kekeringan	Tidak Ada Bencana Alam
0	ACEH	186	1436	81	493	2	106	108	1	43	173	4406
1	SUMATERA UTARA	483	732	82	964	4	78	483	82	89	127	3827
2	SUMATERA BARAT	222	342	65	364	0	57	248	0	18	43	513
3	RIAU	21	485	1	0	0	16	83	0	194	81	1224
4	JAWA BARAT	87	478	17	36	0	2	44	0	16	16	1001
5	SUMATERA SELATAN	103	380	36	49	0	3	90	0	64	96	2844
6	BENGKULU	81	171	15	66	0	14	17	0	4	19	1195
7	LAMPUNG	70	324	23	47	0	35	166	0	11	50	2093
8	KEP. BANGKA BELITUNG	1	89	0	0	0	17	77	0	14	0	268
9	KEP. RIAU	26	61	1	0	0	69	40	0	87	27	241
10	DKI JAKARTA	7	109	0	1	0	5	0	0	1	0	162
11	JAWA BARAT	1238	1193	100	601	0	91	412	0	76	442	3081
12	JAWA TENGAH	1190	1249	89	114	0	123	619	26	49	278	5937
13	DIYODIYAKARTA	89	80	1	82	0	12	24	3	4	31	233
14	JAWA TIMUR	811	1176	96	2449	4	27	488	74	80	166	4656
15	BANTEN	96	419	41	162	1	39	107	0	8	77	914
16	BAU	105	39	5	79	0	34	46	0	6	2	482
17	NUBA TENDOSARA BARAT	44	187	84	100	0	44	66	0	19	88	748
18	NUBA TENDOSARA TIMUR	610	870	161	246	0	218	1219	84	144	464	1427
19	KALIMANTAN BARAT	66	664	22	0	0	42	83	0	126	76	1160
20	KALIMANTAN TENGAH	24	736	9	0	0	12	16	0	66	37	789
21	KALIMANTAN SELATAN	74	929	31	0	0	30	70	0	60	36	973
22	KALIMANTAN TIMUR	74	343	5	6	0	9	26	0	82	32	628
23	KALIMANTAN UTARA	42	200	0	16	0	3	3	0	13	6	246
24	SULAWESI UTARA	218	288	37	349	0	96	107	0	26	24	1078
25	SULAWESI TENGAH	133	987	87	866	0	117	85	0	29	42	964
26	SULAWESI SELATAN	291	620	52	272	0	39	393	0	22	60	1863
27	SULAWESI TENGGARA	46	240	10	43	0	58	104	0	12	16	1682
28	GORONTALO	66	289	24	96	0	37	34	0	9	87	382
29	SULAWESI BARAT	144	125	7	486	0	36	26	0	8	22	106

Gambar 2. Dataset awal

2.2 Preprocessing Data

Data yang akan digunakan sebagai referensi dalam proses pengklusteran adalah informasi tentang "Tanah Longsor" dan "Tidak ada bencana". Untuk melaksanakan proses ini, data tersebut akan dipecah menjadi bagian-bagian yang terpisah dan diubah menjadi format array. Selanjutnya, data ini akan disimpan dalam variabel yang disebut sebagai dataset.

```
In [3]: X = dataset.iloc[:, [1,11]].values
X
Out[3]: array([[ 198, 4406],
 [ 483, 3827],
 [ 222, 513],
 [ 21, 1224],
 [ 57, 1001],
 [ 103, 2644],
 [ 81, 1195],
 [ 70, 2093],
 [ 1, 265],
 [ 25, 241],
 [ 7, 152],
 [1288, 3081],
 [1190, 5637],
 [ 59, 233],
 [ 511, 4655],
 [ 96, 914],
 [ 105, 482],
 [ 44, 748],
 [ 610, 1487],
 [ 66, 1158],
 [ 24, 789],
 [ 74, 973],
 [ 74, 628],
 [ 42, 248],
 [ 218, 1075],
 [ 133, 984],
 [ 251, 1863],
 [ 46, 1852],
 [ 65, 352],
 [ 144, 105],
 [ 58, 798],
 [ 80, 423],
 [ 72, 1482],
 [ 146, 5142]], dtype=int64)
```

Gambar 3. Pembentukan Array Data

Langkah berikutnya adalah melakukan normalisasi data sebelum melakukan perhitungan menggunakan metode klusterisasi. Normalisasi data dilakukan menggunakan fungsi "MinMaxScaler" yang tersedia dalam library "sklearn.preprocessing". Setelah proses normalisasi selesai, hasil dari normalisasi data dapat dilihat dalam Gambar 3.

```
In [4]: from sklearn.preprocessing import MinMaxScaler
from sklearn import preprocessing
min_max_scaler = preprocessing.MinMaxScaler ( feature_range = (0,1) )
col = dataset.columns
result = min_max_scaler.fit_transform(dataset)
min_max_scaler_df = pd.DataFrame(result, columns=col)
print(min_max_scaler_df)
```

	Tanah Longsor	Tidak Ada Bencana Alam
0	0.153069	0.777477
1	0.374514	0.672813
2	0.171717	0.073753
3	0.015540	0.202278
4	0.043512	0.161967
5	0.079254	0.458966
6	0.062160	0.197035
7	0.053613	0.359364
8	0.000000	0.028923
9	0.018648	0.024584
10	0.004662	0.008496
11	1.000000	0.537961
12	0.923854	1.000000
13	0.045066	0.023138
14	0.396270	0.822487
15	0.073815	0.146240
16	0.080808	0.068149
17	0.033411	0.116233
18	0.473193	0.249819
19	0.050505	0.190347
20	0.017871	0.123644
21	0.056721	0.156905
22	0.056721	0.094541
23	0.031857	0.025850
24	0.168609	0.175343
25	0.102564	0.158894
26	0.194250	0.317787
27	0.034965	0.315799
28	0.049728	0.044649
29	0.111111	0.000000
30	0.044289	0.125271
31	0.061383	0.057484
32	0.055167	0.248915
33	0.112665	0.910521

Gambar 4. Normalisasi Data

2.3 Proses Clustering

Di dalam *Jupyter Notebook*, proses pengklusteran menggunakan bahasa pemrograman Python dimulai dengan mencari nilai K yang optimal menggunakan metode *Elbow*. Dalam proses ini, digunakan random state dengan nilai 42.

```
In [8]: from sklearn.cluster import KMeans
wcss = []
for i in range (1,11):
    kmeans = KMeans(n_clusters=i, init='k-means++', random_state=42)
    kmeans.fit(X)
    wcss.append(kmeans.inertia_)
plt.plot(range(1,11), wcss)
plt.title('Elbow Method')
plt.xlabel('Cluster Number')
plt.ylabel('WCSS')
plt.show()
```

Gambar 5 Proses mencari nilai K

Nilai K yang optimal ditentukan ketika grafik eksperimen pada metode *Elbow* menunjukkan adanya penurunan yang signifikan dan hampir membentuk sudut tajam. Gambaran visual dari grafik eksperimen ini dapat ditemukan dalam Gambar 4.

- Indonesia tahun 2021 berdasarkan variabel ‘longsor’ dan ‘tidak ada bencana’.
2. Dataset Desa/Kelurahan Menurut Jenis Bencana Alam tahun 2021 yang digunakan untuk pengklusteran terdiri dari 29 kolom data dalam bentuk format CSV. Sumber dataset ini berasal dari Badan Pusat Statistik (BPS) dan terdiri dari 12 kolom. Menghasilkan 2 kluster yaitu kluster 0 dan kluster 1.
 3. Dengan menggunakan metode *K-Means* bisa digunakan untuk mengelompokkan daerah rawan bencana alam di Indonesia
 4. Hasil dari klusterisasi mengindikasikan bahwa kluster 1 dikelompokkan sebagai daerah rawan bencana diantaranya Aceh, Sumatera Utara, Jawa Barat, Jawa Tengah, Jawa Timur.
 5. Hasil klusterisasi dan identifikasi daerah-daerah dengan potensi rawan bencana yang tinggi dapat digunakan sebagai masukan bagi pemerintah untuk menyampaikan fokus yang lebih pada daerah-daerah tersebut. Dengan demikian, pemerintah dapat mengambil langkah-langkah pencegahan, mitigasi, dan persiapan yang lebih intensif khususnya daerah yang termasuk ke dalam *cluster* 1. Hal ini bertujuan untuk meningkatkan kesiapsiagaan dan respons terhadap kemungkinan bencana di masa depan.

Daftar Pustaka

- Firman, Muhamad, Al Halik, and Laila Septiana. 2022. “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* 6(4): 856–70. <http://journal.stmikjayakarta.ac.id/index.php/jisamar>.
- Gustientiedina, Gustientiedina, M. Hasmil Adiya, and Yenny Desnelita. 2019. “Penerapan Algoritma *K-Means* Untuk *Clustering* Data Obat-Obatan.” *Jurnal Nasional Teknologi dan Sistem Informasi* 5(1): 17–24.
- Khairul Rahmat, Hayatul et al. 2021. 1 Counselling Research and Applications *THE URGENCY OF ALTRUISM AND RESILIENCE IN NATURAL DISASTER MANAGEMENT VOLUNTEERS: A LITERATURE STUDY*.
- Muhari Abdul. 2022. “BNPB Verifikasi 5.402 Kejadian Bencana Sepanjang Tahun 2021.” *Badan Nasional Penanggulangan Bencana*.
- Nuryani, Indah, and Dedi Darwis. 2021. 1 Prosiding Seminar Nasional Ilmu Komputer *ANALISIS CLUSTERING PADA PENGGUNA BRAND HP MENGGUNAKAN METODE K-MEANS*.
- Ramadhani, Devy Isya, Oki Damayanti, Osa Thaushiyah, and Abdul Rahman Kadafi. 2022. “Penerapan Metode *K-Means* Untuk *Clustering* Desa Rawan Bencana Berdasarkan Data Kejadian Terjadinya Bencana Alam.” *JURIKOM (Jurnal Riset Komputer)* 9(3): 749.
- Rinjani Isni, Anwar Saeful, and Herdiana Ruli. 2023. “PENGELOMPOKAN DAERAH BENCANA ALAM MENGGUNAKAN ALGORITMA *K-MEANS CLUSTERING*.” *Jurnal Ilmiah Sistem Informasi Dan Ilmu Komputer* 3.
- Rosaliyah, Indah, Bani Nurhakim, and Jurusan Manajemen. 2023. “*CLUSTERING* KEJADIAN BENCANA ALAM Di JAWA BARAT BERDASARKAN JENIS BENCANA MENGGUNAKAN *K-MEANS*.” 18: 10–16. <https://opendata.jabarprov.go.id/>.
- Wicaksono, David Aryo, and Yeremia Alfa Susetyo. 2023. “*CLUSTERING* ZONASI DAERAH RAWAN BENCANA ALAM DI PROVINSI SUMATERA BARAT MENGGUNAKAN ALGORITMA *K-MEANS* DAN LIBRARY GEOPANDAS.” *Jurnal Indonesia: Manajemen Informatika dan Komunikasi* 4(2): 426–38.
- Murdiaty, Angela & Sylvia, C., 2020. Pengelompokan Data Bencana Alam Berdasarkan Wilayah, Waktu, Jumlah Korban dan Kerusakan Fasilitas Dengan Algoritma *K-Means*. *JURNAL MEDIA INFORMATIKA BUDIDARMA*, Volume 4, pp. 744 - 752.