

PEMETAAN KEJADIAN BENCANA ALAM DI JAWA BARAT MENGUNAKAN ALGORITMA K-MEANS

[Natural Disaster Event Mapping In West Java Using K-Means Algorithm]

Sagaino, sagaino18@gmail.com¹⁾, Teady Matius Surya Mulyana,
10999@lecturer.ubm.ac.id²⁾, I Gusti Ngurah Suryantara, 10181@lecturer.ubm.ac.id³⁾, Jusia
Amanda Ginting, 11717@lecturer.ubm.ac.id⁴⁾, dan Fransiskus Adikara,
11915@lecturer.ubm.ac.id⁵⁾

¹⁾²⁾³⁾⁴⁾⁵⁾Teknik Informatika, Fakultas Teknik dan Desain, Universitas Bunda Mulia

ABSTRACT

Natural disaster is an event that cannot be avoided, therefore a mapping of occurrence of the natural disaster is needed. In addition to mapping, clustering of natural disaster events is also needed to determine which areas have low to high intensity events. In performing the clustering, a method or algorithm can be used, namely the k-means algorithm.

In the research conductor, the scope of natural disasters is West Java Province with the attributes used are floods, landslides and tornadoes. And also from this research, it was conducted to find out how many optimal number of clusters that can be clustered.

The method that used in this study is the K-Means algorithm which is used to perform clustering. The Elbow method is used to determine the optimal K value from the dataset by calculate the SSE (Sum Square Error) of each predetermined cluster.

From the result of its application, the K-Means algorithm can cluster datasets of Natural Disaster in West Java with predetermined attributes. Based on the calculation results from the elbow method, the value of K from the dataset is 4. And from the research conductor, the accuracy rate of each cluster is 0,04% to 0,56%.

Keywords: K-Means, Natural Disaster, Elbow Method, Cluster, Machine Learning

ABSTRAK

Bencana alam adalah peristiwa yang tidak bisa dihindarkan, oleh karena itu dibutuhkan sebuah pemetaan dari kejadian bencana alam tersebut. Selain untuk melakukan pemetaan, pengklasteran kejadian bencana alam dibutuhkan juga untuk menentukan daerah mana yang memiliki intensitas kejadian rendah sampai dengan tinggi. Dalam melakukan pengelompokkan tersebut dapat di gunakan sebuah metode atau algoritma yaitu algoritma K-means.

Pada penelitian yang dilakukan, ruang lingkup kejadian bencana alam yaitu provinsi Jawa Barat dengan atribut yang digunakan adalah kejadian bencana banjir, kejadian bencana tanah longsor dan kejadian bencana puting beliung. Dan juga dari penelitian ini dilakukan untuk mengetahui jumlah *cluster* optimal dari dataset yang dimiliki yang dapat mengklasterkan dataset.

Metode peneltian yang dipergunakan pada penelitian ini merupakan algoritma K-Means yang digunakan untuk melakukan pengklasteran. Metode *Elbow* digunakan untuk menentukan nilai K optimal dari dataset dengan menghitung SSE (*Sum Square Error*) dari tiap kluster yang sudah ditentukan.

Dari hasil penerapannya, metode K-Means dapat mengklasterkan dataset kejadian bencana lama di Jawa Barat dengan atribut yang sudah ditentukan. Berdasarkan hasil perhitungan dari metode *Elbow* didapatkan nilai K optimal dari dataset yaitu berjumlah 4. Dari penelitian yang dilakukan tingkat akurasi dari setiap klaster yaitu sebesar 0,04% sampai 0,56%.

Kata Kunci: K-Means, Bencana Alam, Metode Elbow, *Cluster*, *Machine Learning*.

PENDAHULUAN

Dalam bagian pendahuluan ini akan akan dibagi kedalam beberapa bagian yaitu

a. Latar Belakang

Bencana alam atau dalam Bahasa Inggris yaitu *Natural Disaster* merupakan sebuah peristiwa yang tidak dapat dielakkan lagi, kita tidak tahu kapan dan dimana bencana lama akan terjadi. Adapun terjadinya bencana alam di akibatkan oleh alam itu sendiri seperti angin topan atau puting beliung dan juga di akibatkan oleh manusia yang merusak alam yang memicu terjadinya bencana alam. Seperti membuang sampah sembarangan dan penebangan pohon sembarangan yang dapat menyebabkan banjir maupun tanah longsor[1]. Dan juga menurut Badan Nasional Penanggulangan Bencana (BNPB), bencana alam yang berlangsung pada tahun 2020 ini diungguli dengan bencana yang memiliki sifat hidrometeorologi, contohnya yaitu banjir, banjir bandang, tanah longsor dan angin puting beliung[2].

Oleh sebab itu dibutuhkan informasi serta data bencana alam yang sudah pernah terjadi di berbagai tempat, sehingga dapat memperkirakan, mempersiapkan dan menghindari terjadinya bencana tersebut oleh masyarakat yang berada di wilayah yang rawan akan musibah alam[3]. Dalam mengantisipasi bencana alam yang terutama di provinsi Jawa Barat berdasarkan kejadian yang sudah terjadi selama selang waktu dari tahun 2012 sampai dengan tahun 2020, dibutuhkan pengelompokan berdasarkan kejadian tersebut agar dapat menentukan daerah yang memiliki intensitas kejadian bencana alamnya rendah sampai dengan sangat tinggi. Dalam

melakukan pengklasteran tersebut dapat digunakannya metode algoritma *K-Means*.

Algoritma *K-Means* yang merupakan bagian dari *machine learning*[4], yang dapat mengklasterkan kejadian bencana alam yang terjadi di kabupaten/kota yang ada di provinsi Jawa Barat. Dimana pengklasteran dalam algoritma K-Means ini menggolongkan setiap data ke dalam *cluster* yang berdasarkan titik pusat dari setiap *cluster* yang di inisialisasi secara acak. Setelah mendapatkan hasil atau data dari penghitungan menggunakan K-Means, hasil tersebut dapat digunakan untuk mengetahui daerah atau wilayah kluster mana yang memiliki intensitas kejadian bencana alam. Dengan demikian, berdasarkan permasalahan di atas dibuat penelitian dengan judul “Pemetaan Bencana Alam di Jawa Barat menggunakan Algoritma K-Means”.

b. Identifikasi Masalah

Berdasarkan permasalahan yang sudah di paparkan pada bagian latar belakang berikut merupakan rumusan masalahnya:

1. Bagaimana hasil pengklasteran dengan memanfaatkan *Elbow method* untuk memastikan nilai K yang optimal terhadap dataset yang sudah dimiliki?
2. Bagaimana tingkat akurasi dari setiap cluster yang sudah ditentukan?

c. Tujuan dan Manfaat Penelitian

Berikut merupakan tujuan serta manfaat dari penelitian yang dibuat:

1. Untuk menerapkan metode K-Means pada aplikasi yang dibuat.
2. Untuk mengetahui berapa jumlah cluster terbaik yang didapatkan dari data yang sudah didapatkan.
3. Untuk mengetahui berapa tingkat akurasi dari setiap cluster yang dimiliki.

4. Untuk memberikan tambahan pengetahuan dan wawasan tentang daerah di Jawa Barat yang rentan akan terjadinya bencana alam dan juga dapat dilakukannya pemetaan dari hasil pengklasteran yang sudah dilakukan.

METODOLOGI PENELITIAN

Dataset yang digunakan dalam penelitian ini didapatkan dari website pemerintahan daerah Jawa Barat, yaitu opendata.jabarprov.go.id. Dengan jenis bencana alam yang digunakan yaitu banjir, puting beliung dan tanah longsor, dengan jangka waktu dari tahun 2012 sampai tahun 2020 dan berdasarkan kabupaten/kota.

Metode Studi Pustaka

Studi Pustaka adalah sebuah metode *collecting data* yang digunakan dalam penelitian ini. Metode ini merupakan metode pengumpulan data yang didapatkan dengan cara mendalami, mengamati, mengkaji dan membaca *e-book*, *paper*, skripsi baik bersifat softcopy maupun hardcopy.

Algoritma K-Means

Algoritma K-Means yang merupakan termasuk kedalam Unsupervised Learning yaitu merupakan bagian dari *Machine Learning*[5]. Algoritma K-Means berfungsi untuk melakukan pengklasteran data, dimana tanpa membutuhkan label dari data tersebut. Algoritma ini juga ialah salah satu metode data *clustering non-hierarchical* atau *Partitional Clustering*[6]. Proses yang terjadi dalam algoritma K-Means dilakukan secara berulang – ulang hingga mendapatkan atau menghasilkan data klaster. Untuk mendapatkan hasil output *cluster*, maka dibutuhkannya inputan berupa jumlah *cluster* atau K yang diinginkan. Untuk melakukan penghitungan terlebih dahulu harus menentukan jumlah K, dan juga menentukan titik pusat *Centroids* yang dapat dilakukan secara

random[7]. Jumlah perulangan atau biasa disebut dengan iterasi tergantung dari titik pusat yang ditentukan di awal secara *random*. Setelah mendapatkan nilai titik pusat awal, maka nilai tersebut akan di hitung jaraknya dengan data yang dimiliki dengan menggunakan rumus *Ecludian Distance*[8]. Rumus ini berguna untuk menghitung jarak antar titik pusat terhadap data, dan data yang memiliki jarak terpendek terhadap titik pusat akan di alokasikan ke titikpusat tersebut yang nantinya akan membentuk sebuah *cluster*[9].

Adapaun proses yang dilakukan dalam algoritma K-Means:

1. Tentukan jumlah K atau *cluster* yang dibutuhkan.
2. Inisialisasi *centroids* awal, dimana inisialisasi dapat dilakukan secara *random*.
3. Hitung jarak antara titik pusat terhadap masing – masing data yang dimiliki menggunakan rumus *Ecludian Distance*.

$$ED = \sqrt{(X_i - W_i)^2 + (Y_i - Z_i)^2} \dots\dots\dots(1)$$

Rumus *Ecludian Distance* yang dipergunakan untuk menghitung jarak di peragakan pada rumus 1. Dimana variabel i merupakan banyaknya objek data, variabel X dan Y merupakan koordinat dari objek data dan variabel W dan Z merupakan koordinat dari *centroid*.

4. Setelah mendapatkan hasil perhitungan menggunakan rumus *Ecludian Distance*, maka akan dilakukan pengalokasian data dimana data akan di alokasikan ke titik pusat yang memiliki jarak terdekat.
5. Setelah melakukan pengalokasian, maka akan menentukan titik pusat baru dari setiap *cluster*. Penentuan titik pusat

baru dilakukan dengan cara menentukan rata – rata (*Mean*) dari setiap *cluster*.

6. Lakukan lagi perhitungan antara titik pusat terhadap semua data menggunakan rumus 1. Apabila titik pusat *cluster* tidak mengalami perubahan maka proses *clustering* dianggap selesai. Dan apabila masih terjadi perubahan maka proses akan di lanjutkan kembali ke proses ke 3 sampai mendapatkan titik pusat *cluster* tidak berubah.

Metode Elbow

Elbow method adalah salah satu *optional method* dari algoritma K-Means, dimana berguna untuk menentukan jumlah K yang paling optimal dari data yang dimiliki[10]. Cara kerja dari metode *Elbow* adalah dengan membandingkan nilai persentase dari setiap cluster. Untuk mendapatkan perbandingan tersebut digunakan rumus SSE (*Sum Square Error*) yang di hitung tiap masing – masing *cluster*[11].

$$SSE = \sum_{k=1}^K \sum_{x_i \in S_k} \|x_i - c_k\|_2^2 \dots \dots \dots (2)$$

Rumus SSE (*Sum Square Error*) yang di pergunakan untuk menghitung perbandingan tiap *cluster*. Dimana variable X_i merupakan nilai dari atribut ke-i dan variable C_k merupakan nilai dari atribut ke-k.

Hasil perhitungan dengan menggunakan rumus SSE dapat di tunjukkan dalam bentuk grafik. Apabila nilai dari *cluster* pertama dan membuat sebuah sudut dalam grafik atau memiliki perbandingan yang besar, maka cluster tersebut merupakan cluster terbaik[12].

Seperti terlihat di gambar 2 di tentukan terlebih dahulu jumlah *cluster*. Setelah itu dilakukan perhitungan menggunakan rumu SSE terhadap masing – masing *cluster* dan mendapatkan hasil seperti gambar 2. Terjadi penurunan yang besar mulai dari K=1 sampai K = 4, dan dari k=4 sampai

K=5 terjadi penurunan secara perlahan – lahan. Maka dapat di ketahui bahwa nilai K ideal dari gambar 2 tersebut adalah K=4.

HASIL DAN PEMBAHASAN

Berdasarkan hasil pengujian terhadap keseluruhan data yang dimiliki dengan total data sebanyak 241. Dan fitur yang digunakan yaitu kejadian bencana alam Banjir, kejadian bencana alam Tanah Longsor, kejadian bencana alam Puting Beliung dan Kabupaten/Kota yang berada dalam Provinsi Jawa Barat. Dengan cluster berjumlah 4 yang didapatkan melalui perhitungan metode *Elbow* dan titik pusat awal di tentukann secara *random*. Hasil pengujian data yang berjumlah 241 dengan menggunakan metode K-Means bisa dilihat dalam tabel 1.

Tabel 1. Hasil Pengklasteran dengan K-Means

No	Cluster	Kabupaten/Kota	Tahun
1	0	Kab. Kuningan	2013
2	0	Kab. Sukabumi	2017
3	0	Kota Bogor	2017
4	0	Kota Bogor	2018
5	0	Kab. Bogor	2019
..
40	1	Kab. Garut	2020
41	1	Kab. Kuningan	2020
42	1	Kab. Majalengka	2020
43	1	Kab. Sumedang	2020
44	1	Kota Bogor	2020
...
100	2	Kab. Subang	2020
101	2	Kab. Purwakarta	2020
102	2	Kab. Bandung Barat	2020
103	2	Kab. Pangandaran	2020
104	2	Kota Banjar	2020
...
237	3	Kota Cirebon	2020
238	3	Kota Bekasi	2020
239	3	Kota Depok	2020
240	3	Kota Cimahi	2020
241	3	Kota Tasikmalaya	2020

Berdasarkan Tabel 1 yang merupakan hasil pengklasteran menggunakan algoritma K-Means terhadap 241 data dan jumlah K yaitu 4. Pada *cluster* 0 memiliki total anggota sebesar 10 dengan titik pusat yaitu [16.1], [64], [107.4] dan *range* kejadian bencana banjir mulai dari 10 sampai 43 kejadian, *range* kejadian puting beliung mulai dari 11 sampai 183 kejadian, *range* kejadian bencana tanah longsor mulai dari 68 sampai 133 kejadian. Pada *cluster* 1 memiliki total anggota sebesar 34 dengan titik pusat yaitu [10.7], [17.8], [50.5] dan *range* kejadian bencana banjir mulai dari 2 sampai 29 kejadian, *range* kejadian puting beliung mulai dari 5 sampai 39 kejadian, *range* kejadian bencana tanah longsor mulai dari 32 sampai 82 kejadian. Pada *cluster* 2 memiliki total anggota sebesar 60 dengan titik pusat yaitu [11.8], [10.7], [18.7] dan *range* kejadian bencana banjir mulai dari 1 sampai 71 kejadian, *range* kejadian puting beliung mulai dari 1 sampai 35 kejadian, *range* kejadian bencana tanah longsor mulai dari 0 sampai 34 kejadian. Pada *cluster* 3 memiliki total anggota sebesar 137 dengan titik pusat yaitu [2.8], [3], [2.6] dan *range* kejadian bencana banjir mulai dari 0 sampai 16 kejadian, *range* kejadian puting beliung mulai dari 0 sampai 19 kejadian, *range* kejadian bencana tanah longsor mulai dari 0 sampai 15 kejadian.

Berdasarkan dari hasil pengklasteran dapat di ketahui bahwasannya *cluster* 0 merupakan *cluster* dengan tingkat kejadian bencana alam sangat tinggi di karenakan memiliki total jumlah titik pusat paling besar yaitu 187.5, *cluster* 1 merupakan *cluster* dengan tingkat kejadian bencana alam tinggi dengan total jumlah titik pusat sebesar 79, *cluster* 2 merupakan *cluster* dengan tingkat kejadian bencana alam sedang dengan total jumlah titik pusat sebesar 41.2 dan *cluster* 3 merupakan *cluster* dengan tingkat kejadian bencana alam rendah dengan total jumlah titik pusat sebesar 8.4.

Setelah mendapatkan hasil pengklasteran, maka akan dilakukan perhitungan akurasi, dengan rumus:

$$AR = \frac{\text{jumlah data cluster}}{\text{Total jumlah data}} \times 100\% \dots \dots \dots (3)$$

Rumus akurasi yang digunakan untuk menentukan tingkat akurasi dari setiap cluster dengan jumlah data tiap *cluster* dibagi dengan total data dan di kali dengan 100%.

Berikut merupakan hasil perhitungan tingkat akurasi dari tiap *cluster* berdasarkan hasil perhitungan algoritma K-Means:

1. Nilai akurasi dari *cluster* 0 yang memiliki 10 data yaitu:

$$AR = \frac{10}{241} \times 100\% = 0,04\%$$
2. Nilai akurasi dari *cluster* 1 yang memiliki 34 data yaitu:

$$AR = \frac{34}{241} \times 100\% = 0,14\%$$
3. Nilai akurasi dari *cluster* 2 yang memiliki 60 data yaitu:

$$AR = \frac{34}{241} \times 100\% = 0,14\%$$
4. Nilai akurasi dari *cluster* 3 yang memiliki 137 data yaitu:

$$AR = \frac{137}{241} \times 100\% = 0,56\%$$

Berdasarkan hasil perhitungan, tingkat akurasi yang didapatkan dari *cluster* 0 sebesar 0,04%, *cluster* 1 bernilai 0,14%, *cluster* 2 bernilai 0,24% dan *cluster* 4 bernilai 0,56%. Berdasarkan hasil penghitungan yang dilakukan bisa dilihat juga di tabel 2.

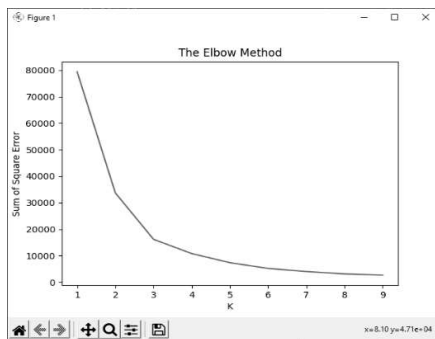
Implementasi Metode Elbow

Dari pengimplementasian metode *Elbow* terhadap total data sebanyak 241, dengan jumlah *cluster* sebanyak 10. Untuk pengujian data dibagi secara acak menjadi 3 bagian dengan jumlah data yang berbeda.

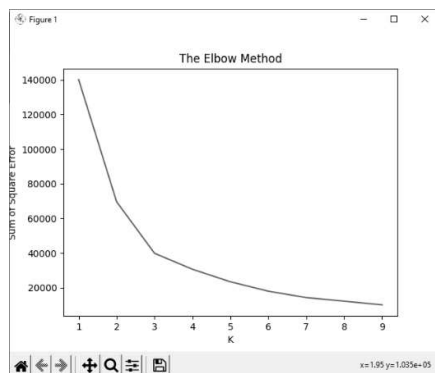
- a. Uji coba 1 terhadap 60 data
 Pada uji coba pertama menggunakan 60 data, hasil dari perhitungan SSE dari K=2 sampai K=3 mendapati degradasi yang besar dengan selisih nilai 17512,81 dan K=4 sampai K=9 mendapati degradasi secara berangsur - angsur. Maka didapatkan hasil bahwa K=4 merupakan K optimal pada uji coba menggunakan 60 data. Pengujian bisa dilihat berdasarkan tabel 3 dan juga gambar 4.

Tabel 2. Hasil SSE terhadap 60 data

Cluster	Hasil SSE	Selisih
K1	79402,31	79402,31
K2	33698,74	45703,57
K3	16185,93	17512,81
K4	10812,01	5373,92
K5	7375,33	3436,68
K6	5204,34	2170,99
K7	4031,91	1172,43
K8	3142,91	889
K9	2695,55	447,36



Gambar 1. Grafik Metode Elbow Terhadap 60 data



Gambar2. Grafik Metode Elbow terhadap 120 data

- b. Uji coba 2 terhadap 120 data
 Pada uji coba pertama menggunakan 120 data, hasil dari perhitungan SSE

dari K=2 sampai K=3 mendapati degradasi yang besar dengan selisih nilai 29769,56 dan K=4 sampai K=9 mendapati degradasi secara berangsur - angsur. Maka didapatkan hasil bahwa K=4 merupakan K optimal pada uji coba menggunakan 120 data. Pengujian bisa dilihat berdasarkan tabel 4 dan juga gambar 5.

Tabel 3. Hasil SSE terhadap 120 data

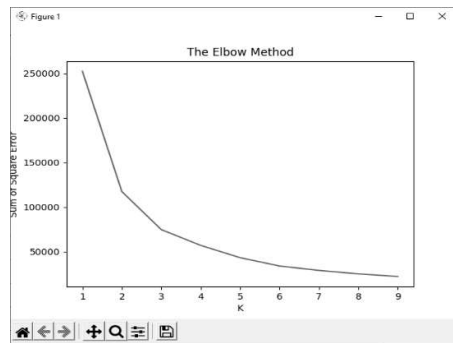
Cluster	Hasil SSE	Selisih
K1	140100,75	140100,75
K2	69642,60	70458,15
K3	39873,04	29769,56
K4	30649,53	9223,51
K5	23425,82	7223,71
K6	17339,18	6086,64
K7	14300,76	3038,42
K8	12240,00	2060,76
K9	10097,67	2142,33

- c. Uji coba 3 terhadap 241 data
 Pada uji coba pertama menggunakan 241 data, hasil dari perhitungan SSE dari K=2 sampai K=3 mendapati degradasi yang besar dengan selisih nilai 42595,59 dan K=4 sampai K=9 mendapati degradasi secara berangsur - angsur. Maka didapatkan hasil bahwa K=4 merupakan K optimal pada uji coba menggunakan 241 data. Pengujian bisa dilihat berdasarkan tabel 5 dan juga gambar 6.

Tabel 4. Hasil SSE terhadap 241 data

Cluster	Hasil SSE	Selisih
K1	252156,46	252156,46
K2	117406,75	134781,91
K3	7477,96	42595,59
K4	57084,50	17842,12
K5	4342,36	13656,48

K6	33958,26	9322,1
K7	29013,34	4944,92
K8	25141,00	3833,83
K9	22205,30	2797,47



Gambar 3. Grafik Metode Elbow terhadap 241 data

SIMPULAN

Berdasarkan hasil pengimplementasian dari algoritma dan metode pada aplikasi yang dibuat, penulis bisa membuat hasil kesimpulan dari penelitian:

1. Algoritma K-Means bisa dipakai dalam melakukan pengklasteran terhadap kejadian bencana alam yang ada di Jawa Barat dengan atribut yang digunakan yaitu bencana alam banjir, tanah longsor, puting beliung dan kabupaten/kota.
2. Aplikasi dapat mengvisualisasikan pemetaan kejadian bencana alam berdasarkan jumlah cluster yang telah di dapatkan dari perhitungan algoritma k-means.
3. Metode *elbow* berhasil dalam menentukan jumlah K optimal dari data yang dimiliki dengan K optimal adalah 4.
4. Hasil K ini juga digunakan dalam menentukan *cluster* menggunakan algoritma K-Means.
4. Klaster 0 memiliki jumlah anggota 10 yang merupakan klaster dengan tingkat kejadian bencana alam sangat tinggi dengan total jumlah titik pusat sebesar 187.5, klaster 1 memiliki anggota 34 yang merupakan klaster dengan tingkat

kejadian bencana alam sangat tinggi dengan total jumlah titik pusat sebesar 79, klaster 2 memiliki jumlah anggota 60 yang merupakan klaster dengan tingkat kejadian bencana alam sangat tinggi dengan total jumlah titik pusat sebesar 41.2, klaster 3 memiliki jumlah anggota 137 yang merupakan klaster dengan tingkat kejadian bencana alam sangat tinggi dengan total jumlah titik pusat sebesar 8.4. Dengan tingkat akurasi dari C0, C1, C2, C3 yaitu 0,04%, 0,14%, 0,24% dan 0,56%.

DAFTAR PUSTAKA

- [1] Martini, "Identifikasi Sumber Bencana Alam Dan Upaya Penanggulangannya Di Sulawesi Tengah," *J. Infrastruktur, Tek. Sipil Univ. Tadulako, Palu.*, vol. 1, pp. 96–102, 2011, [Online]. Available: <http://jurnal.untad.ac.id/jurnal/index.php/JTSl/article/view/689/593>.
- [2] D. Arifin, "Sebanyak 2.925 Bencana Alam Terjadi Pada 2020 di Tanah Air, Bencana Hidrometeorologi Mendominasi," *BNPB*, 2020. <https://www.bnpb.go.id/berita/sebanyak-2-925-bencana-alam-terjadi-pada-2020-di-tanah-air-bencana-hidrometeorologi-mendominasi> (accessed Oct. 07, 2021).
- [3] A. Priambodo, N. Widyaningrum, and H. K. Rahmat, "Strategi Komando Resor Militer 043/ Garuda Hitam dalam Penanggulangan Bencana Alam di Provinsi Lampung," *Perspektif*, vol. 9, no. 2, pp. 307–313, 2020, doi: 10.31289/perspektif.v9i2.3588.
- [4] M. R. Muttaqin and M. Defriani, "Algoritma K-Means untuk Pengelompokan Topik Skripsi Mahasiswa," *Ilk. J. Ilm.*, vol. 12, no. 2, pp. 121–129, 2020, doi: 10.33096/ilkom.v12i2.542.121-129.
- [5] P. Bradley, K. Bennett, and A. Demiriz, "Constrained k-means

- clustering,” *Tech. Rep.*, p. 9, 2000, [Online]. Available: <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.33.3257>.
- [6] I. Bin Mohamad and D. Usman, “Standardization and its effects on K-means clustering algorithm,” *Res. J. Appl. Sci. Eng. Technol.*, vol. 6, no. 17, pp. 3299–3303, 2013, doi: 10.19026/rjaset.6.3638.
- [7] J. A. Ginting and A. M. Jadera, “Analisa Indeks Vegetasi Menggunakan Citra Satelit Lansat 7 dan Lansat 8 Menggunakan Metode K-Means di Kawasan Gunung Sinabung,” *Indones. J. Comput. Model.*, vol. 1, no. 1, pp. 42–48, 2018, doi: 10.24246/j.icm.2018.v1.i1.p42-48.
- [8] G. Hamerly and C. Elkan, “Learning the in -means,” *Adv. Neural Inf. Process. Syst.*, 2004.
- [9] W. Sutikno, J. F. Palandi, and C. A. Oktavia, “Klasifikasi E-book Berbahasa Inggris Dengan Menggunakan Metode K-Means Clustering Studi Kasus Perpustakaan STIKI Malang,” *J-Intech*, vol. 7, no. 01, pp. 80–85, 2019, doi: 10.32664/j-intech.v7i01.409.
- [10] N. Putu, E. Merliana, and A. J. Santoso, “Analisa Penentuan Jumlah Cluster Terbaik pada Metode K-Means,” pp. 978–979.
- [11] A. B. H. Kiat, Y. Azhar, and V. Rahmayanti, “PENERAPAN METODE K-MEANS DENGAN METODE ELBOW UNTUK SEGMENTASI PELANGGAN MENGGUNAKAN MODEL RFM(Recency, Frequency, & Monetary),” *J. Repos.*, vol. 2, no. 7, p. 945, 2020, doi: 10.22219/repositor.v2i7.973.
- [12] N. T. Hartanti, “Metode Elbow dan K-Means Guna Mengukur Kesiapan Siswa SMK Dalam Ujian Nasional,” *J. Nas. Teknol. dan Sist. Inf.*, vol. 6, no. 2, pp. 82–89, 2020, doi: 10.25077/teknosi.v6i2.2020.82-89.