

DOI: doi.org/10.21009/03.1301.FA11

IDENTIFIKASI JENIS GEMPA VULKANIK BERDASARKAN DATA REKAMAN SEISMIC PADA GUNUNG SINABUNG MENGGUNAKAN *PRINCIPAL COMPONENT ANALYSIS*

Bagas Anwar Arif Nur^{1, a)}, Bambang Heru Iswanto¹, Mohammad Hasib²,
Ahmad Basuki³

¹*Program Studi Fisika, Universitas Negeri Jakarta, Jakarta, Indonesia.*

²*Pusat Riset Kebencanaan Geologi, Badan Riset dan Inovasi Nasional, Bandung, Indonesia*

³*Pusat Vulkanologi dan Mitigasi Bencana Geologi, Bandung, Indonesia*

Email: ^{a)}bagasanwararifnur@gmail.com

Abstrak

Erupsi vulkanik merupakan peristiwa alam yang memiliki potensi merusak yang signifikan terhadap manusia dan lingkungan. Identifikasi jenis gempa vulkanik menjadi kunci dalam upaya mitigasi risiko bencana dalam memberikan informasi proses dan lokasi kegiatan magma pada terjadi. Dalam penelitian ini, kami mengusulkan sebuah pendekatan menggunakan Principal Component Analysis (PCA) untuk mengidentifikasi jenis gempa vulkanik berdasarkan data rekaman seismik. Identifikasi diawali dengan melakukan reduksi dimensi fitur menggunakan Principal Component Analysis (PCA). Hasil PCA kemudian dilakukan clustering yang setelahnya dievaluasi dengan Silhouette Score, Adjusted Random Indeks (ARI), David-Bouldin Indeks (DB-Indeks), dan Calinski-Harabasz index (CH-Indeks). Ekperimen dilakukan dengan menggunakan data rekaman berjumlah 329 sampel. Untuk tiap rekaman dilakukan ekstraksi fitur berupa fitur statistik, fitur entropi, dan fitur bentuk dengan total berjumlah 16 fitur dalam domain waktu dan frekuensi Hasil PCA pada dua komponen utama PC1 menjelaskan 49.2741% dan PC2 24.5507% dari varians data dan Hasil evaluasi menggunakan *Silhouette Score* sebesar 0.53, ARI 0.8, CH-Indeks 529.34, dan DB-Indeks 0.6

Kata-kata kunci: tuliskan kata-kata kunci tidak lebih dari satu baris.

Abstract

Volcanic eruptions are natural events that have the potential for significant damage to humans and the environment. Identifying the type of volcano earthquake is key in disaster risk mitigation by providing information on the process and the location of magma activity beneath the volcano. In this research, we propose an approach using Principal Component Analysis (PCA) to identify types of volcanic earthquakes based on seismic recording data. Identification begins by reducing feature dimensions using Principal Component Analysis (PCA). The PCA results were then clustered and then evaluated Silhouette Score, ARI, CH-Indeks, DB-Indeks. Experiments were carried out using recorded data totaling 329 samples. For each recording, feature extraction was carried out in the form of statistical features, entropy features and shape features with a total of 16 features in the time and frequency domains. PCA results on the two main components PC1 explained 49.2741% and PC2 24.5507% of the data variance and evaluation results using *Silhouette Score* were equal to 0.53, ARI 0.8, CH-Index 529.34, and DB-Index 0.6

Keywords: tuliskan kata-kata kunci tidak lebih dari satu baris.

PENDAHULUAN

Indonesia merupakan negara yang terletak di wilayah ring of fire, menyebabkan Indonesia menjadi negara yang memiliki jumlah api terbanyak di dunia sebanyak 629 dengan 129 [1] atau sekitar 16 % dari jumlah seluruh gunung api aktif di dunia [2]. Akibat hal ini Indonesia menjadi salah satu negara yang rawan dengan bencana alam disebabkan oleh gunung api.

Pada tahun 1883 dimana terjadi erupsi Krakatau yang dampaknya dapat dirasakan seluruh dunia [3]. Pada tahun 2010 terjadi letusan pada gunung Merapi dimana sebanyak 150 juta m³ lahar dingin dimuntahkan. Letusan ini menyebabkan korban jiwa sebanyak 346 jiwa [4]. Bencana letusan gunung api juga terjadi pada gunung sinabung. Letusan pertama gunung sinabung pada abad 21 terjadi pertama kali pada 27 Agustus 2010. Letusan ini menyebabkan gunung sinabung menjadi gunung api Tipe-A setelah sebelumnya merupakan gunung api Tipe-B [5]

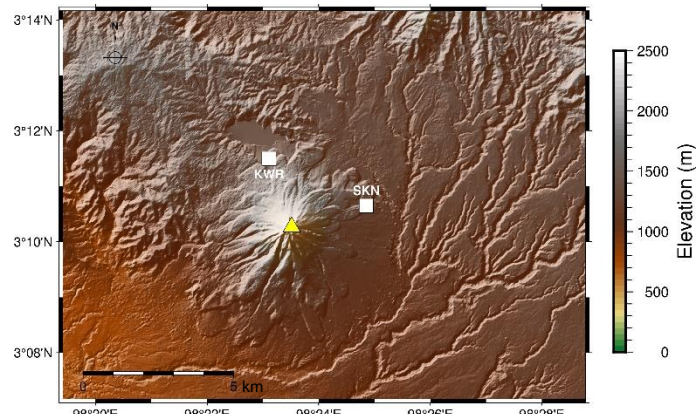
Dampak dari bencana alam yang disebabkan menjadi hal yang penting untuk diperhatikan dan karena itu mitigasi bencana alam akibat gunung api memegang peran penting dalam mengurangi resiko yang ditimbulkan. Mitigasi bencana pada gunung api umumnya dilakukan dengan melakukan pemantauan aktivitas seismiknya. Aktivitas seismik pada gunung dapat menggambarkan pergerakan magma di bawah permukaan bumi, pemahaman mengenai rekaman seismik dapat dilakukan dengan memperhatikan jumlah kegempaan yang terjadi. Informasi kegempaan pada gunung api dapat digunakan sebagai pengetahuan untuk menentukan status bahaya gunung api dan informasi resiko erupsi yang terjadi [6], [7]

Namun permasalahan dalam menentukan kegempaan pada gunung api secara manual terjadi karena sifatnya yang subjektif dan memakan banyak waktu. Pada rekaman seismik, untuk tiap jenis kegempaan akan memiliki bentuk gelombang dan konten frekuensi yang berbeda [8]. sehingga memungkinkan untuk mengidentifikasinya dengan cara mengekstrak informasi dari gelombang aslinya. Penelitian kali ini bertujuan untuk mengidentifikasi jenis gempa vulkanik dengan mengekstraksi fitur dan kemudian dilakukan reduksi dimensi menggunakan Principal Component Analysis (PCA). Setelahnya dilakukan clustering untuk melihat seberapa baik identifikasi yang dilakukan.

METODOLOGI

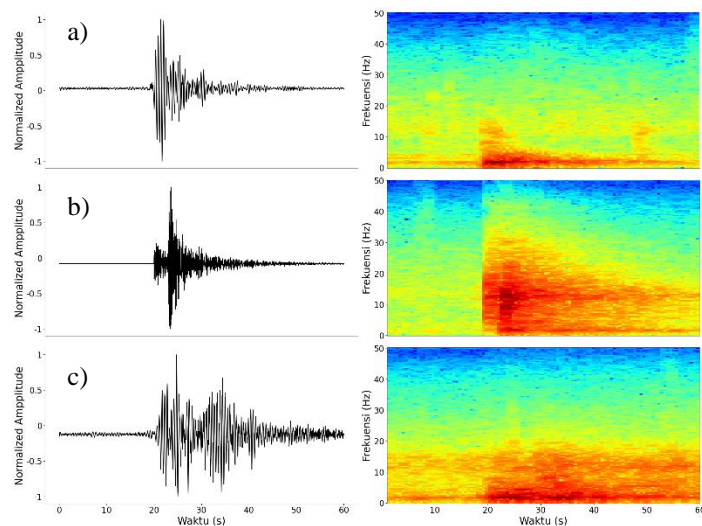
Gunung Sinabung merupakan gunung tipe-B yang kembali aktif pada tahun 2010. Sejak tahun 2010 gunung Sinabung mengalami 4 periode erupsi dalam rentang tahun 2010-2021 [9]. Januari – April 2016 termasuk periode dimana gunung Sinabung memiliki aktivitas seismik yang tinggi sehingga analisis dilakukan pada periode tersebut karena adanya ketersediaan data. Data rekaman yang didapatkan sifatnya sekunder yang artinya data tidak didapatkan secara langsung. Pada penelitian ini data rekaman berasal dari Pusat Vulkanologi dan Mitigasi Bencana Geologi (PVMBG).

Pada penelitian ini stasiun yang digunakan untuk analisis adalah stasiun Sukanali (SKN) yang terletak pada 03°10'35.84" N, 98°24'51.52" E dan Laukawar (KWR) yang terletak pada 03°11'29.70" N, 98°23'06.30" E dengan kordinat kawah aktif berada di 03°10'12" N, 98°23'33" E (kordinat ditulis dalam format Degree, Minute, Second). Perekaman pada kedua stasiun dilakukan menggunakan short-period seismometer dengan satu komponen vertikal (L4C) dengan resolusi A/D 24-bit [10]. Kedua stasiun tersebut dipilih karena merupakan stasiun yang memiliki signal-noise-ratio tinggi. Data diterima dalam format WIN dengan rekaman selama satu menit untuk setiap file. Perekaman dilakukan dengan 100 sampling rate yang artinya untuk setiap satu detik terdapat 100 titik data.



Gambar 1. Detail lokasi gunung Sinabung dan lokasi stasiun Laukawar dan Sukanalu

Sebelum dilakukan pengolahan data, dilakukan preparasi terlebih dahulu, preparasi ini meliputi penggabungan data menjadi data berdurasi satu hari. Koreksi instrument untuk mengubah nilai pada rekaman menjadi nilai yang sebenarnya. Setelahnya dilakukan pengambilan dan pelabelan sample untuk jenis gempa yang diteliti dimana tiap sample yang akan diteliti akan memiliki durasi selama 60 detik. Pelabelan pada penelitian kali ini didasari pada literatur [10] dan katalog letusan yang diberikan oleh PVMBG. Jenis gempa yang diteliti diantaranya adalah gempa Volcano-Tectonic (VT), Low-Frequency (LF), dan Letusan dengan sample pada tiap kelas berjumlah 129, 100, dan 100.



Gambar 2. Contoh *waveform* dan spektrogram dari tiap jenis gempa a). *Low-Frequency*, b). *Volcano-Tectonic*, c). Letusan

Penelitian dilanjutkan dengan melakukan ekstraksi fitur pada tiap sample. Fitur yang diekstrak berasal dari domain waktu dan frekuensi, serta spektrogramnya. Untuk domain waktu fitur ini terdiri dari *Zero-Crossing Rate*, *Energy*, dan *Entropy*. Pada domain frekuensi terdapat frekuensi dominan, dan fitur statistik berupa *Kurtosis*, *Mean*, dan *Skewness*. Sementara pada spektrogram fitur yang diekstraksi merupakan fitur statistik. Total dari fitur yang digunakan pada penelitian kali ini adalah 11 fitur

Fitur yang telah diekstraksi kemudian akan dilakuakn reduksi dimensi menggunakan PCA. PCA sendiri merupakan sebuah teknik statistik untuk mereduksi dimensi data dengan mempertahankan informasi paling relevannya. PCA dapat memberikan informasi varian dari sebuah data [11]

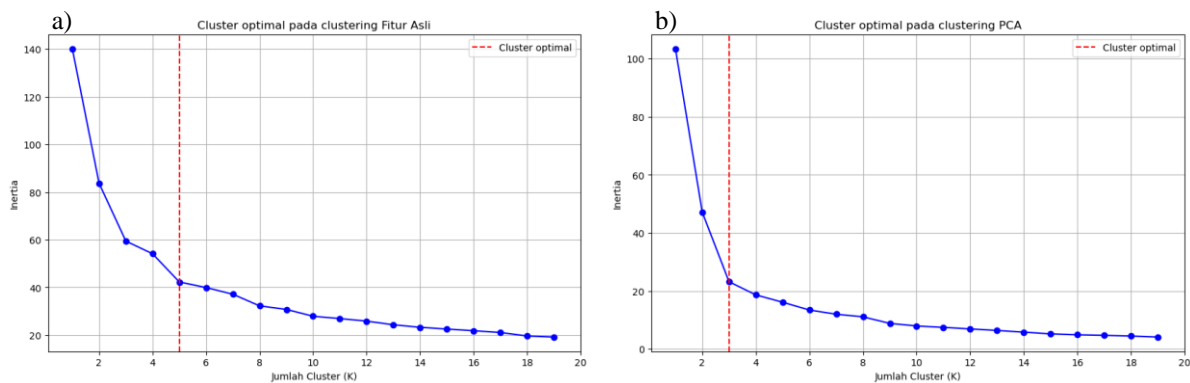
Identifikasi dilakukan dengan teknik clustering seperti *K-Means* yang berbasis pada jarak euclidian, *X-Means* yang merupakan pengembangan *K-means* dengan keunggulan untuk memperkirakan kemungkinan jumlah cluster yang ada, *Gaussian Mixture Model (GMM)* berdasarkan persebaran gaussian pada data, dan *Agglomerative Clustering (AC)* yang dibangun berdasarkan hirarki [12], [13], [14], [15]. Clustering akan dilakukan menggunakan dua macam masukkan, yaitu masukkan Ketika fitur asli digunakan tanpa direduksi dan fitur Ketika sudah

direduksi menggunakan PCA. Untuk tiap hasil clustering kemudian akan dievaluasi menggunakan empat macam metrik diantaranya adalah *Sillhouette Score*, *David-Bouldin Indeks* (DB-Indeks), *Calinski-Harabasz index* (CH-Index), dan *Adjusted Random Index* (ARI)

HASIL DAN PEMBAHASAN

Pada penelitian ini digunakan 4 model clustering diantaranya adalah X-means, K-means, GMM, dan AC. Pada tahap awal dilakukan reduksi dimensi menggunakan PCA dimana PCA menunjukkan bahwa penggunaan dua komponen utama PC1 dan PC2 secara berurutan dapat menjelaskan 49.2741% dan 24.5507% dari total varians data. Setelahnya dilakukan clustering menggunakan X-means untuk melihat jumlah optimal kelompok dari suatu data. X-means digunakan pada dua keadaan yaitu pada fitur asli tanpa reduksi dan pada data yang sudah direduksi menggunakan PCA, X-means sendiri merupakan model clustering yang dalam proses pelatihannya tidak memerlukan inialisasi awalan sehingga dapat memberikan informasi mengenai jumlah cluster yang memungkinkan pada suatu data [15]. Clustering menggunakan X-means dapat dilihat pada Tabel 1. Pada Tabel 1. ditunjukkan bahwa X-means memprediksi kemungkinan adanya 6 cluster yang terdapat pada masukkan fitur asli. Sementara itu pada masukkan hasil PCA X-means memprediksi kemungkinan adanya 3 cluster yang berarti sesuai dengan jumlah kelompok pada penelitian kali ini. Hal ini mengindikasikan bahwa PCA dapat mengambil informasi penting pada data.

Pada gambar 3. Ditunjukkan juga evaluasi elbow untuk mengetahui jumlah optimal cluster. Hasil pada penggunaan fitur asli menunjukkan jumlah cluster yang optimal adalah 5 sementara pada penggunaan PCA jumlah optimal cluster adalah 3. Baik X-means dan Evaluasi elbow, cluster yang dibentuk menggunakan masukkan PCA lebih sesuai dengan jumlah kelompok pada penelitian ini

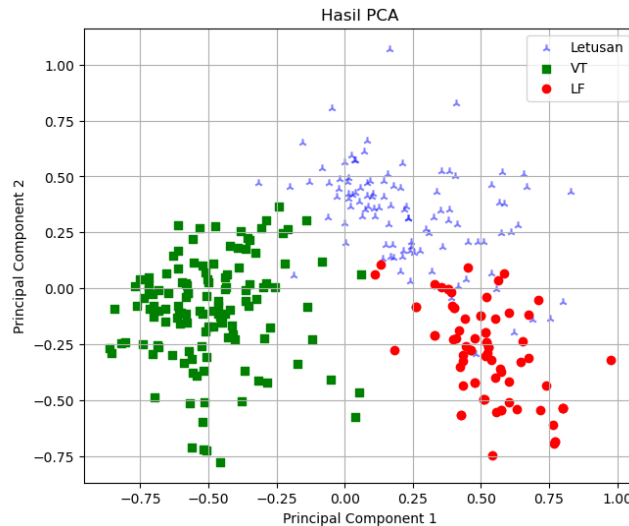


Gambar 3. Evaluasi Elbow a). Penggunaan fitur asli b). Penggunaan PCA

TABEL 1. Jumlah Cluster yang dihasilkan X-means

Fitur Masukkan	Jumlah Cluster
Fitur Asli	6
Hasil PCA	3

Pada Gambar 4. ditunjukkan plot PCA, secara visual ketiga kelompok sudah cukup terpisah dan menempati ruangnya sendiri. Hal ini mengindikasikan bahwa ketiga jenis gempa sudah dapat tergambar dengan baik dan memiliki karakteristiknya masing-masing



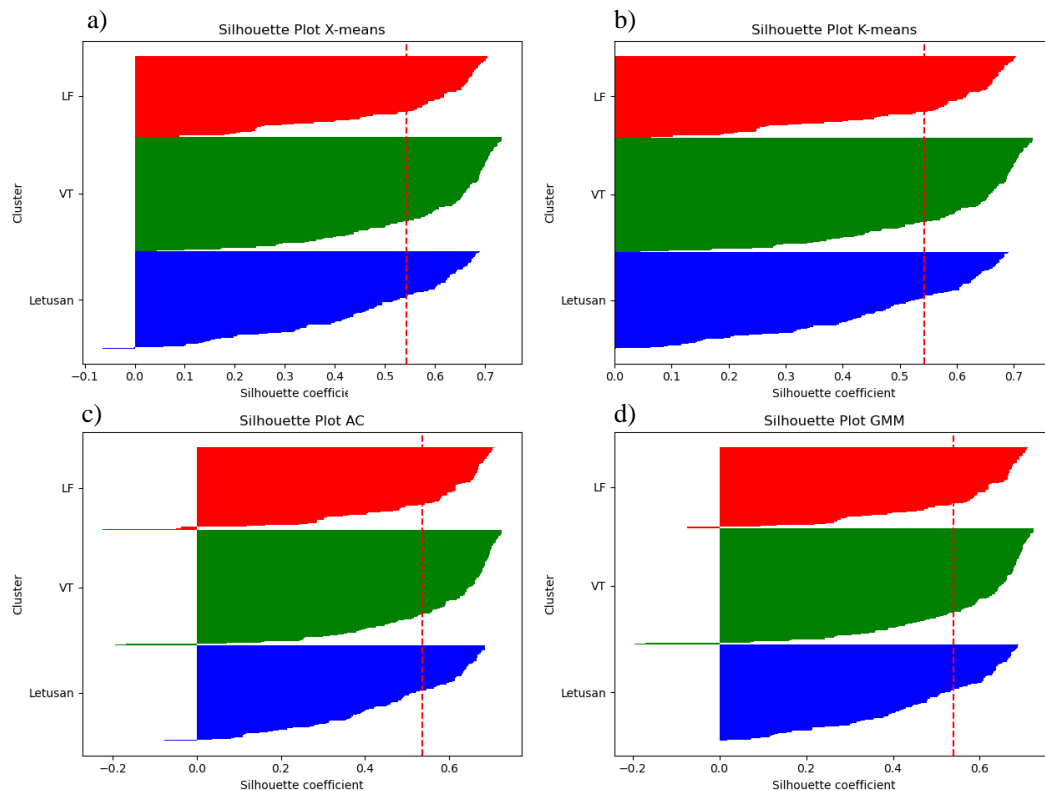
Gambar 4. Hasil PCA

Selain dilakukan analisis visual dilakukan juga analisis secara kuantitatif. Analisis kuantitatif dilakukan menggunakan empat metrik diantaranya adalah *Silhouette Score* yang memiliki rentang -1 hingga 1 dimana nilai 1 menunjukkan pemisahan yang sempurna, ARI dengan nilai terbaiknya adalah 1, CH-Index yang menunjukkan semakin besar nilainya maka semakin baik hasil *clustering*, dan DB-Index yang menunjukkan *clustering* semakin baik apabila nilainya mendekati 0 [16]. Pada hasil tabel 2. Ditunjukkan hasil perhitungan metrik untuk *clustering* keempat model yang dilakukan menggunakan masukkan fitur asli dan masukkan PCA. Secara keseluruhan penggunaan PCA sebagai masukkan menunjukkan kenaikan performa untuk setiap model *clustering*, mengindikasikan bahwa PCA dapat mengoptimalkan model dalam memahami pola pada suatu data.

TABEL 2. Perhitungan empat metrik

Clustering	Fitur Masukkan	Metrik			
		<i>Silhouette Score</i>	ARI	CH-Index	DB-Index
X-means	Fitur Asli	0.2791	0.4668	164.7903	1.2666
	Hasil PCA	0.5435	0.8084	529.2857	0.6280
K-means	Fitur Asli	0.3836	0.8060	207.5458	1.0566
	Hasil PCA	0.5435	0.8060	529.5437	0.6286
GMM	Fitur Asli	0.3471	0.7450	179.7175	1.1696
	Hasil PCA	0.5395	0.8182	517.4301	0.6274
AC	Fitur Asli	0.3768	0.8722	201.0545	1.0700
	Hasil PCA	0.5360	0.7951	511.3170	0.6322

Gambar 4. Menunjukkan plotting silhouette score untuk keempat model. Silhouette plot menunjukkan ukuran jarak suatu anggota cluster dengan pusat clusternya dan pusta cluster lainnya. Jika diperhatikan pada beberapa model *clustering* terdapat warna yang menuju kearah minus, hal ini menunjukkan bahwa beberapa anggota dari cluster yang bersangkutan lebih dekat dengan pusat cluster lain. Hal ini sesuai dengan plot PCA pada gambar 3. Dimana beberapa anggota dari salah satu kelompok masih berada di wilayah kelompok yang lain. Pada penelitian kali ini gempa letusan merupakan kelompok yang anggotanya paling banyak berada di wilayah kelompok gempa LF. Namun secara dominan tiap anggota kelompok sudah berada pada wilayahnya masing-masing.



Gambar 5. *Silhouette Plot* a). X-means, b). K-means, c). AC, d). GMM

KESIMPULAN

Pada penelitian kali ini dilakukan identifikasi jenis gempa vulkanik berdasarkan data rekaman seismiknya pada gunung Sinabung dengan jenis gempa yang diteliti adalah *Volcano-Tectonic*, *Low-Frequency*, dan Letusan. Penelitian dilakukan dengan empat model clustering diantaranya adalah K-means, X-means, Gaussian Mixture Models, dan Agglomerative Clustering. Untuk setiap model dilakukan pengujian dengan dua masukan yaitu masukan menggunakan fitur asli tanpa reduksi dan hasil reduksi menggunakan PCA. Hasil penelitian menunjukkan bahwa PCA dapat meningkatkan performa model dalam memahami pola pada data.

UCAPAN TERIMA KASIH

Penulis ingin mengungkapkan terima kasih kepada pihak Badan Riset dan Inovasi Nasional dan Pusat Vulkanologi dan Mitigasi Bencana Geologi yang telah bersedia mengizinkan penggunaan data kepada penulis dalam penelitian ini

REFERENSI

- [1] L. C. Gosal, R. C. Tarore, and H. H. Karongkong, "Analisis Spasial Tingkat Kerentanan Bencana Gunung Api Lokon Di Kota Tomohon," *Spasial*, vol. 5, no. 2, pp. 229–237, 2018.
- [2] A. Nugroho, "Pengembangan model pembelajaran mitigasi bencana gunung meletus di sekolah dasar lereng gunung slamet," *J. Pengabd. Masy. Multidisiplin*, vol. 1, no. 2, pp. 131–137, 2018.
- [3] A. L. Madden-Nadeau *et al.*, "The magmatic and eruptive evolution of the 1883 caldera-forming eruption of Krakatau: Integrating field-to crystal-scale observations," *J. Volcanol. Geotherm. Res.*, vol. 411, p. 107176, 2021.
- [4] T. Trirahayu, "Manajemen bencana erupsi gunung merapi oleh badan penanggulangan bencana daerah kabupaten sleman," *J. Public Policy Adm. Res.*, vol. 1, no. 6, pp. 1138–1151,

- 2016.
- [5] S. B. Kusumayudha, P. Lestari, and E. T. Paripurno, "Eruption characteristic of the sleeping volcano, Sinabung, North Sumatera, Indonesia, and SMS gateway for disaster early warning system," *Indones. J. Geogr.*, vol. 50, no. 1, pp. 70–77, 2018.
 - [6] R. S. J. Sparks, J. Biggs, and J. W. Neuberg, "Monitoring volcanoes," *Science (80-.)*, vol. 335, no. 6074, pp. 1310–1311, 2012.
 - [7] M. Hasib *et al.*, "Event classification of volcanic earthquakes based on K-Means clustering: Application on Anak Krakatau Volcano, Sunda Strait," in *IOP Conference Series: Earth and Environmental Science*, 2024, p. 12045.
 - [8] S. R. McNutt and D. C. Roman, "Volcanic seismicity," in *The encyclopedia of volcanoes*, Elsevier, 2015, pp. 1011–1034.
 - [9] E. Program., Global Volcanism. compiled by Venzke, "Sinabung (261080)," [*Database*] *Volcanoes World (v. 5.1.7; 26 Apr 2024)*.
 - [10] H. Gunawan *et al.*, "Overview of the eruptions of Sinabung Volcano, 2010 and 2013--present and details of the 2013 phreatomagmatic phase," *J. Volcanol. Geotherm. Res.*, vol. 382, pp. 103–119, 2019.
 - [11] M. A. Abdillah, B. H. Iswanto, and H. Suhendar, "EKSTRAKSI FITUR BUNYI KETUKAN BUAH KELAPA BERBASIS POWER-NORMALIZED CEPSTRAL COEFFICIENTS (PNCC)," in *PROSIDING SEMINAR NASIONAL FISIKA (E-JOURNAL)*, 2024.
 - [12] K. P. Sinaga and M.-S. Yang, "Unsupervised K-means clustering algorithm," *IEEE access*, vol. 8, pp. 80716–80727, 2020.
 - [13] D. Xu and Y. Tian, "A comprehensive survey of clustering algorithms," *Ann. data Sci.*, vol. 2, pp. 165–193, 2015.
 - [14] R. Xu and D. Wunsch, "Survey of clustering algorithms," *IEEE Trans. neural networks*, vol. 16, no. 3, pp. 645–678, 2005.
 - [15] A. G. Jondya and B. H. Iswanto, "Indonesian's traditional music clustering based on audio features," *Procedia Comput. Sci.*, vol. 116, pp. 174–181, 2017.
 - [16] F. Pedregosa *et al.*, "Scikit-learn: Machine Learning in {P}ython," *J. Mach. Learn. Res.*, vol. 12, pp. 2825–2830, 2011