

Algoritma K-Means untuk Mengelompokkan Hotel di Sekitar Wilayah Indonesia yang Rentan Gempa Bumi

Nurfidah Dwitianti*, Erlin Windia Ambarsari, Noni Selvia

Fakultas Teknik dan Ilmu Komputer, Program Studi Teknik Informatika, Universitas Indraprasta PGRI, Jakarta, Indonesia

Email: ^{1,*}nurfidah.pulungan@gmail.com, ²erlinunindra@gmail.com, ³noni.selvia@gmail.com

Email Penulis Korespondensi: nurfidah.pulungan@gmail.com

Abstrak—Karena lokasinya di Cincin Api Pasifik, Indonesia sering mengalami gempa bumi. Fenomena ini membahayakan masyarakat serta infrastruktur, termasuk industri perhotelan. Hotel dikelompokkan berdasarkan lokasi di sekitar wilayah Indonesia yang rentan terhadap gempa bumi, sehingga pemilik hotel dapat menggunakan desain dan konstruksi hotel yang tahan gempa saat membangun. Ini karena daerah dengan gempa tinggi meningkatkan risiko kerusakan bangunan hotel. Penelitian ini menggunakan data lokasi geografis hotel dan data seismik Indonesia untuk mengklasifikasikan hotel di wilayah Indonesia yang rawan gempa dengan algoritma K-Means. Data diambil dari USGS dari Januari 2022 hingga November 2023. Dengan menggunakan algoritma K-Means, hotel dikelompokkan menjadi kluster berdasarkan kedekatan mereka dengan episenter gempa dan magnitudo gempa yang tercatat. Hasil penelitian menunjukkan bahwa ada dua kluster yang ideal berdasarkan nilai shilhoutte tertinggi, yaitu 0,954, yang berarti dapat dikelompokkan menjadi dua kluster yang berbeda berdasarkan latitude, longitude, dan alamat hotel, dengan empat hotel dan tiga hotel sebagai anggota masing-masing kluster. Hotel Mutiara, Saparua, dan Tiara memiliki tingkat kejadian rentan yang lebih tinggi dibandingkan dengan hotel lainnya karena berada di dekat lokasi gempa. Penelitian ini dapat membantu pemilik hotel dan pembuat kebijakan membuat strategi untuk mengurangi risiko, meningkatkan keselamatan, dan mengembangkan berkelanjutan di wilayah yang rawan gempa.

Kata Kunci: Gempa Bumi; Clustering; Algoritma; K-Means; Nilai Shilhoutte

Abstract—Due to its location on the Pacific Ring of Fire, Indonesia frequently experiences earthquakes. This phenomenon endangers society and infrastructure, including the hotel industry. Hotels are placed based on locations around areas of Indonesia vulnerable to earthquakes so that hotel owners can use hotel designs and construction that are earthquake-resistant when built. This is because areas with high earthquakes increase the risk of damage to hotel buildings. This research uses hotel geographic location data and Indonesian seismic data to classify hotels in earthquake-prone regions of Indonesia using the K-Means algorithm. Data was taken from USGS from January 2022 to November 2023. Using the K-Means algorithm, hotels were aggregated into clusters based on their proximity to the earthquake epicenter and recorded earthquake magnitude. The research results show that there are two ideal clusters based on the highest Shilhoutte value, namely 0.954, which means it can be sent into two different clusters based on latitude, longitude, and hotel address, with four hotels and three hotels as members of each cluster. Hotels Mutiara, Saparua, and Tiara have a higher level of vulnerability than other hotels because they are located near the earthquake. This research can help hotel owners and policymakers create strategies to reduce risk, increase safety, and develop sustainability in earthquake-prone regions.

Keywords: Earthquake; Clustering; Algorithm; K-Means; Shilhoutte Value

1. PENDAHULUAN

Seiring dengan pertumbuhan industri perhotelan yang pesat di seluruh dunia, penting bagi pengembang, manajemen, dan pemilik hotel untuk memahami dan mengelola ancaman gempa bumi. Gempa bumi adalah bencana alam yang tidak dapat diprediksi dan memiliki potensi kerusakan yang sangat besar. Terletak di Cincin Api Pasifik, Indonesia sering mengalami gempa bumi [1]. Fenomena ini menimbulkan bahaya bagi penduduk dan infrastruktur, termasuk industri perhotelan. Berdasarkan data dari BMKG, sebanyak 57 ribu rumah dan fasum mengalami kerusakan akibat gempa dengan magnitudo di atas 5 SR di sepanjang tahun 2022 [2]. Sangat penting untuk memahami tingkat rentang gempa bumi di suatu wilayah karena gempa bumi pasti menyebabkan kerugian besar. Di daerah dengan rentang gempa yang tinggi, desain dan konstruksi bangunan harus lebih ketat. Hotel yang terletak di daerah ini harus dirancang dengan desain struktural yang tahan gempa. Sehingga dibutuhkan penelitian mengenai pengelompokkan hotel yang rentan terdampak oleh kejadian gempa bumi sebagai pedoman bagi manajemen hotel untuk memiliki strategi mitigasi risiko sebagai langkah mengurangi risiko bagi tamu dan kerugian finansial.

Mitigasi gempa bumi dapat dilakukan dengan mengolah data dengan clustering dan mengidentifikasi karakteristik gempa [3], [4]. Dalam teknik clustering, suatu objek ditempatkan dalam kelompok yang sama dengan objek lain yang terkait, tetapi dalam kelompok yang berbeda jika melibatkan objek yang tidak terkait [5]. Karena karakteristiknya, metode ini berkonsentrasi pada pengelompokan objek berdasarkan kesamaan tertentu. Metode ini dapat dibagi menjadi dua kategori, yaitu pengelompokan hierarkis dan pengelompokan terpartisi [6]. Metode pengelompokan objek yang dikenal sebagai "clustering" melibatkan pembagian objek menjadi kelompok, atau kluster, sehingga semua objek dalam suatu kluster "mirip" satu sama lain dan "tidak mirip" satu sama lain [7]. Tujuan dari clustering adalah meminimalkan jarak di dalam cluster dan memaksimalkan jarak antar cluster [8]. Ada sejumlah algoritma, salah satunya adalah K-means, yang dapat digunakan untuk menyelesaikan pendekatan clustering [9]. Algoritma k-means dapat menangani jumlah data yang besar dengan cepat dan efisien [3], [10]–[12]. Algoritma K-Means juga memiliki potensi sebagai alat yang dapat digunakan untuk analisis kluster gempa.

Penelitian tentang algoritma K-Means dan pengelompokan data gempa bumi di Indonesia telah banyak dilakukan. Salah satu contohnya adalah penelitian yang dilakukan oleh Hadi, Pratiwi dan Slamet, menggunakan data gempa bumi Indonesia tahun 2012-2021, algoritma K-Means dan DBSCAN berhasil dikelompokkan, tetapi penelitian menunjukkan

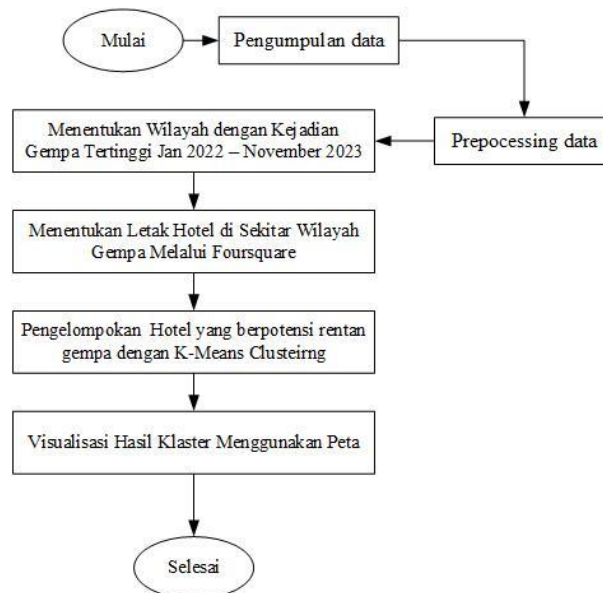
bahwa K-Means memiliki tingkat kesalahan kelompok yang lebih rendah daripada DBSCAN [3]. Studi terkait lainnya, dengan menggunakan algoritma K-Means untuk mengelompokkan kejadian gempa bumi. Hasil penelitian menunjukkan bahwa hasil klaster yang diperoleh adalah $k = 7$, dengan satu klaster mengalami jumlah kejadian gempa maksimum 571 dan satu klaster lainnya mengalami jumlah kejadian gempa minimum 180 [13]. Penelitian tambahan tentang penggunaan algoritma K-Means tentang pengelompokan data gempa bumi di Provinsi Nusa Tenggara Barat. Hasil penelitian menunjukkan bahwa jumlah klaster optimal yang digunakan dalam pengelompokan ini adalah 2 berdasarkan skor silhouette tertinggi sebesar 0,749 dengan episenter gempa cukup tersebar, dan jumlah gempa di wilayah timur Nusa Tenggara Barat lebih banyak dibandingkan wilayah barat sebelum tahun 2018 [14].

Tujuan dari penelitian ini adalah untuk mengelompokkan hotel di Indonesia berdasarkan tingkat kerentanan hotel terhadap gempa bumi. Hal ini dilakukan dengan menggunakan algoritma K-Means untuk menganalisis data lokasi geografis hotel dan data seismik di Indonesia. Metode klasterisasi memungkinkan penilaian kerentanan gempa yang lebih tepat. Hal ini juga memudahkan pembuatan strategi yang ditargetkan untuk mengurangi risiko. Selain itu, proses klasterisasi tetap skalabel dan dapat diterapkan pada sejumlah besar hotel di Indonesia berkat kemampuan algoritma K-Means untuk menangani dataset besar. Pendekatan ini memastikan bahwa setiap hotel, terlepas dari ukuran atau lokasinya, dapat memperoleh manfaat dari pengelompokan kerentanan gempa yang tepat. Penelitian ini menggunakan algoritma K-Means menawarkan perspektif baru tentang cara hotel dapat dikategorikan berdasarkan risiko gempa dan menawarkan kontribusi praktis bagi pemangku kepentingan industri perhotelan untuk meningkatkan kesiapsiagaan dan mitigasi bencana diantaranya dapat memberikan panduan berharga bagi pemilik hotel dalam memilih lokasi dan desain konstruksi yang tahan gempa, terutama di daerah dengan risiko gempa tinggi dan bagi pembuat kebijakan dapat menjadi acuan dalam merumuskan kebijakan pembangunan berkelanjutan dan tata ruang yang mempertimbangkan risiko gempa.

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Gambar 1 menunjukkan prosedur pengambilan data, pre-processing data, dan clustering K-Means yang digunakan dalam penelitian ini.



Gambar 1. Tahapan penelitian

Tahapan penelitian dimulai dari melakukan pengumpulan data tentang wilayah di Indonesia yang memiliki kejadian gempa bumi. Data yang digunakan dalam penelitian ini diperoleh dari United States Geological Survey (USGS), sebuah agensi ilmiah pemerintah Amerika Serikat. Data ini diambil dari 1661 gempa bumi yang terjadi di Indonesia dari Januari 2022 hingga November 2023 dengan magnitude $\geq 4,5$ dan kedalaman maksimal sampai 310 km. Tahapan selanjutnya, proses preprocessing data dilakukan meliputi tiga hal utama, diantaranya seleksi data, pembersihan data (preprocessing data) dan integrasi data dalam analisis ini, yang berarti memilih data yang akan digunakan dalam proses data mining, menghapus data yang tidak penting, melengkapi data dan menggabungkan dua atau lebih kumpulan data menjadi satu kumpulan data [15]. Proses selanjutnya adalah menentukan daerah mana yang paling sering mengalami gempa. Berdasarkan lokasi tersebut, serta garis lintang dan bujur akan dihubungkan melalui Foursquare dan menggunakan API-nya untuk mencari hotel di dekat lokasi gempa. Tujuannya adalah untuk menemukan hotel yang perlu meningkatkan dalam kesiapsiagaan dan mitigasi bencana. Kemudian, tahapan berikutnya adalah untuk mengelompokkan hotel yang diperoleh dari Foursquare API, algoritma K-Means Clustering menggunakan unsupervised learning. Label dari kumpulan data dihilangkan, dan kemudian, sebagai parameter, kumpulan data yang tidak berlabel diteruskan ke

algoritma K-Means Clustering. Hasil dari kluster K-Means membagi hotel ke dalam k kelompok berbeda dan memplotnya ke dalam peta. Tahapan penelitian ini diadaptasi berdasarkan implementasi penelitian yang telah dilakukan oleh Nath dengan menggunakan bahasa pemrograman Python [16].

2.2 Algoritma K-Means Clustering

Tujuan awal algoritma ini adalah menentukan jumlah kluster yang ingin dibentuk, kemudian memilih variabel pertama yang akan digunakan sebagai titik tengah cluster, dan kemudian melakukan proses pengulangan. Metode ini dianggap sebagai algoritma pembelajaran yang sederhana. Proses ini akan dilakukan hingga data menjadi stabil sehingga objek data tidak akan bergerak [17]. Pertama, data dibagi. Setelah menghitung mean untuk setiap kelompok, data dibagi kembali dengan mengatur setiap data ke posisi cluster mean terdekatnya. Algoritma ini, dalam versi paling sederhananya, terdiri dari langkah-langkah berikut [7]:

1. Menghitung jumlah k-kelompok secara acak;
2. Menghasilkan nilai acak untuk setiap kelompok centroid sebanyak k-kelompok;
3. Menghitung jarak setiap data masukan ke setiap centroid menggunakan jarak geometris untuk mengetahui jarak terdekat setiap data dengan centroid. Euclidean adalah salah satu dari banyak metode untuk mengukur jarak data ke pusat kelompok[18]. Persamaan 2 dapat digunakan untuk menghitung jarak Euclidean[19].

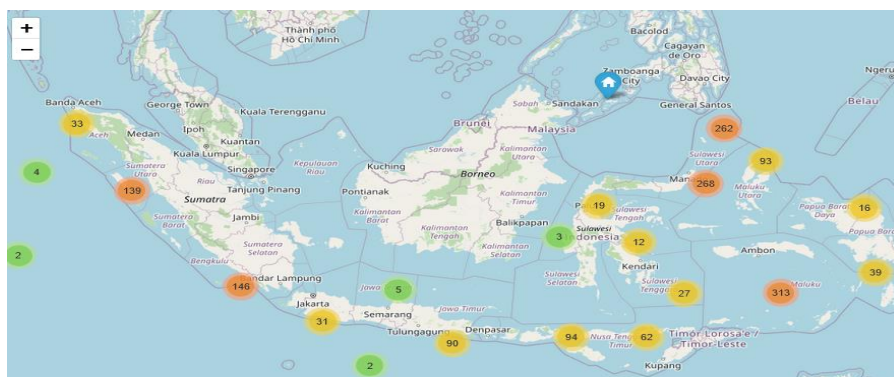
$$d = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \quad (1)$$

4. Membagi setiap data berdasarkan jarak terkecil ke centroid.
5. Perbarui nilai centroid massa. Nilai centroid baru diperoleh dari nilai rata-rata cluster.
6. Ulangi langkah ini hingga anggota setiap cluster tidak berubah.
7. Jika langkah-langkah di atas telah dilakukan, nilai rata-rata pusat cluster pada iterasi terakhir akan digunakan sebagai parameter untuk menentukan pengelompokan data.

3. HASIL DAN PEMBAHASAN

3.1 Pengolahan Data

Penelitian ini akan menyelidiki algoritma K-Means clustering untuk mengelompokkan hotel berdasarkan data seismik dan lokasi geografis yang berada di wilayah rentan gempa di Indonesia. Sumber data dalam penelitian ini diperoleh dari United States Geological Survey (USGS). Data yang diambil berupa data gempa bumi di Indonesia mulai dari bulan Januari 2022 hingga November 2023 dengan magnitude $\geq 4,5$ dan kedalaman maksimal sampai 310 km. Pada tahap preprocessing data, dilakukan cleaning terhadap data yang tidak diperlukan dan mengisi data yang kosong serta mengubah kolom dengan tipe data yang tidak sesuai ke kolom yang sesuai untuk memudahkan pengelolaan data. Selanjutnya, data yang telah dibersihkan diplot seperti yang ditampilkan pada gambar 2 di bawah ini.



Gambar 2. Plot data gempa di wilayah Indonesia, Januari 2022-November 2023

Gambar 2 menampilkan plot data gempa di wilayah Indonesia, dimana dari peta tersebut didapatkan bahwa distribusi kejadian gempa hampir merata terjadi di setiap wilayah kepulauan Indonesia di sepanjang tahun 2022 hingga 2023 dengan magnitude $\geq 4,5$ kecuali di pulau kalimantan yang jarang terjadi. Ini berarti bahwa wilayah di Indonesia sangat rawan akan bencana gempa bumi. Berdasarkan distribusi gempa bumi, dapat ditentukan urutan wilayah di Indonesia yang sering mengalami kejadian gempa di sepanjang tahun 2022 hingga 2023 yang ditunjukkan pada tabel 1.

Tabel 1. Kejadian gempa di Indonesia, Jan 2022 – November 2023

Wilayah	Banyak Kejadian Gempa
Banda Sea	70
Molucca Sea	32

Wilayah	Banyak Kejadian Gempa
Kepulauan Babar, Indonesia	15
Southwest of Sumatra, Indonesia	11
South of Java, Indonesia	10
Halmahera, Indonesia	8
Java, Indonesia	8
Pulau Tanimbar, Indonesia	8
Southern Sumatra, Indonesia	7
Kepulauan Barat Daya, Indonesia	7
Mindanao, Philippines	7
Pulau-Pulau Talaud, Indonesia	7
Seram, Indonesia	6
Minahasa, Sulawesi, Indonesia	6
Northern Sumatra, Indonesia	5
Flores Sea	5
Flores Region, Indonesia	5
67 km NNE of Ruteng, Indonesia	4

Pada tabel 1, ditemukan bahwa ada 18 tempat di mana jumlah gempanya telah meningkat atau setara dengan 5 dalam hampir 2 tahun terakhir ini. Ditemukan bahwa tempat-tempat ini benar-benar rentan terhadap gempa bumi, terutama Banda Sea, Molucca Sea, dan Kepulauan Babar. Selanjutnya, dari wilayah yang sering terjadi gempa, dihubungkan melalui API Foursquare untuk mencari hotel di sekitar lokasi gempa berdasarkan pada garis lintang dan bujur. Gambar 3 merupakan hasil nama-nama hotel yang diperoleh dari API Foursquare yang nantinya akan diubah kedalam bentuk dataframe pandas untuk memudahkan pengolahan data.

```

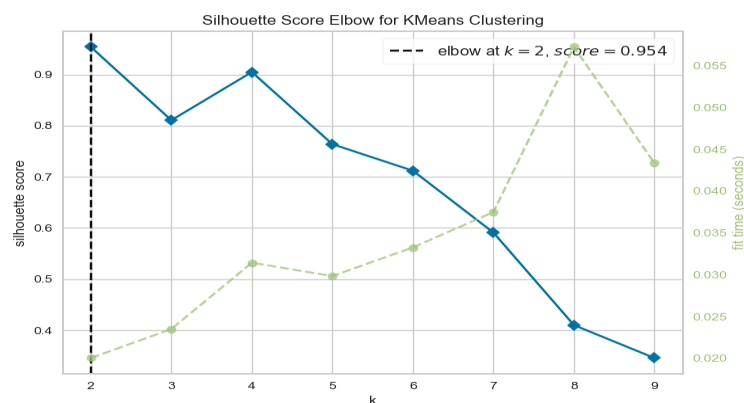
0          name  latitude  longitude \
1      Room 14 Hotel Perdana Saparua -3.577253  128.652934
2      Room 8 Perdana Hotel, Saparua Island -3.571335  128.658465
3      Hotel Arena Bula -3.101972  130.482060
4      Hotel Mutiara Bula Kab. Seram Bag Timur -3.095324  130.476283
5      Hotel Mutiara -3.095059  130.476253
6      Mutiara Hotel -3.106419  130.488174
7      Hotel Tiara 4.268199  96.364630
8      Hotel Oasis 4.230702  96.307819
9      Hotel Jw Marriot 4.225820  96.308410
10     hotel tiara 4.198821  96.085353

0          address
1      [Saparua, Saparua, Molucas 97525, Indonesia]
2      [Maluku (Pl. Benteng Duustede), Saparua, Maluku]
3      [Bula, Maluku, Indonesia]
4      [Bula, Maluku, Indonesia]
5      [Indonesia]
6      [Bula, Seram Bagian Timur, Maluku, Indonesia]
7      [Meulaboh (Meulaboh), Meulaboh, Aceh 52473, In...]
8      [Indonesia]
9      [Indonesia]
10     [Meulaboh, Kabupaten Aceh Barat, Aceh, Indonesia]

```

Gambar 3. Daftar nama hotel sekitar wilayah gempa di Indonesia

Tahapan berikutnya adalah dilakukan pengelompokkan hotel yang memiliki potensi berisiko gempa. Algoritma clustering yang digunakan adalah K-Means. Untuk memulai algoritma K-Means, langkah pertama adalah menentukan jumlah k (cluster) yang akan dibentuk pada data. Untuk menemukan nilai k optimal, metode silhouette adalah salah satu metode yang dapat digunakan bersama dengan algoritma ini. Informasi nilai k optimal yang dihasilkan oleh metode silhouette adalah nilai k optimal yang dipilih untuk setiap silhouette dari nilai rata-ratanya; nilai rata-rata yang lebih tinggi merupakan nilai yang lebih optimal [3], [20]. Pembangkitan nilai k optimal dengan metode silhouette digambarkan dalam Gambar 4.



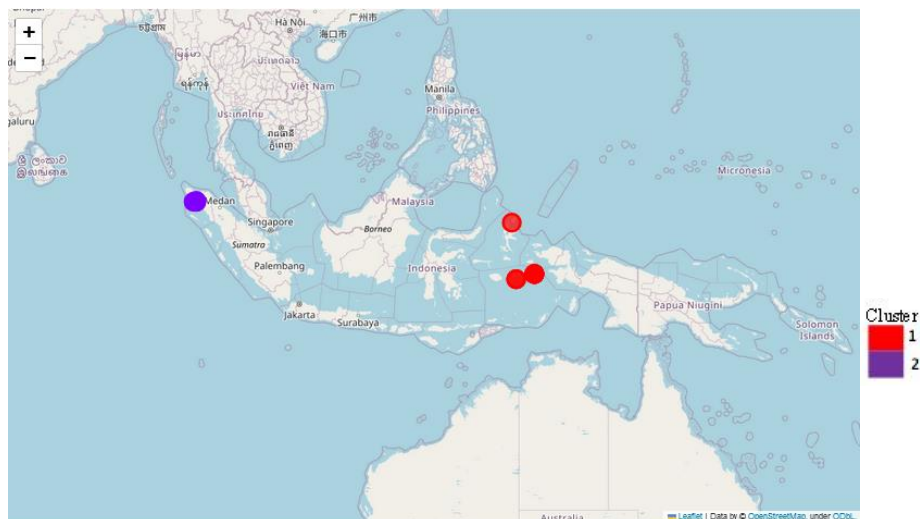
Gambar 4. Metode Silhouette untuk menemukan K Optimal

Jumlah kluster k yang ideal adalah 2, dan nilai silhouette adalah 0.954, seperti yang ditunjukkan pada Gambar 4. Selanjutnya, hasil ditampilkan dalam bentuk tabel. Tabel 2 menunjukkan hasil klusterisasi untuk nama-nama hotel di Indonesia yang memiliki risiko gempa.

Tabel 2. Kejadian gempa di Indonesia, Jan 2022 – November 2023

Nama	Latitude	Longitude	Alamat	Cluster Labels	Banyak Kejadian
Hotel Perdana Morotai	2.100.150	128.201.622	['Indonesia']	1	1
Hotel Perdana Saparua	- 3.577.253	128.652.934	['Saparua', 'Saparua', 'Molucas 97525', 'Indo...']	1	2
Hotel Mutiara	- 3.095.324	130.476.283	['Maluku', 'Indonesia']	1	3
Hotel Arema Bula	- 3.101.972	130.482.060	['Bula', 'Maluku', 'Indonesia']	1	1
Hotel Tiara	4.268.199	96.364.630	['MeuLaboh (MeuLaboh)', 'Meulaboh', 'Aceh 5247...']	2	2
Hotel Jw Mariot	4.225.820	96.308.410	['Indonesia']	2	1
Hotel Oasis	4.230.702	96.307.819	['Indonesia']	2	1

Selanjutnya, dilakukan pemetaan dari hasil algoritma K-Means yang ditampilkan pada gambar 5.

**Gambar 5.** Hasil pemetaan kluster menggunakan algoritma K-Means

Dari gambar 5 terlihat letak kedua kluster yang sudah terbentuk pada peta, dimana kluster yang pertama ditunjukkan pada lingkaran berwarna merah berada di wilayah kepulauan Maluku dan kluster yang kedua yang ditunjukkan oleh lingkaran berwarna ungu berada di Banda Aceh. Hal ini menjelaskan bahwa letak lokasi hotel yang rentan dengan kejadian gempa bumi berada di kedua kluster.

3.2 Pembahasan

Kebanyakan gempa bumi sering terjadi di sekitar bagian timur Indonesia, seperti yang ditunjukkan pada gambar 2. Hal ini karena sesar-sesar di sana lebih aktif dibandingkan di Indonesia bagian barat. Misalnya, di daerah Maluku, yang terletak di tempat beberapa lempeng bumi berkumpul [16]. Hasil penelitian ini juga serupa dengan hasil penelitian sebelumnya [11]. Dengan menggunakan algoritma DBSCAN dan data gempa dari Indonesia mulai dari tahun 2018 hingga 2020, penelitian tersebut menemukan bahwa kegempaan tersebar hampir di seluruh wilayah Indonesia, tetapi wilayah barat memiliki titik atau frekuensi kegempaan yang agak rendah dibandingkan dengan wilayah tengah dan timur. Analisis menunjukkan 1281 lokasi gempa unik yang tercatat dalam dua tahun terakhir. Di antaranya, 28 lokasi mengalami peningkatan frekuensi gempa dalam dua tahun terakhir, dengan jumlah tertinggi 70 gempa di Laut Banda selama dua tahun, 32 gempa di Laut Molluca selama dua tahun, dan 15 gempa di Kepulauan Bar Bar selama dua tahun. Ketiga lokasi ini merupakan lokasi gempa bumi yang paling tidak stabil dan konsisten. Hasil analisis saat ini menunjukkan bahwa tujuh hotel berbeda memiliki risiko gempa bumi karena lokasinya. Hotel Mutiara, Saparua, dan Tiara adalah yang memiliki risiko gempa bumi yang lebih tinggi, karena ketiga hotel tersebut berada di dekat lokasi gempa. Gambar 5 menunjukkan kesamaan antara hotel-hotel dalam dua kluster yang berbeda berdasarkan latitude, longitude, dan alamatnya.

4. KESIMPULAN

Hasil penelitian ini menunjukkan bahwa terdapat tujuh hotel yang rentan terhadap gempa bumi karena lokasinya; dari tujuh hotel tersebut, Hotel Mutiara, Hotel Saparua, dan Hotel Tiara memiliki tingkat kejadian rentan yang lebih tinggi. Berdasarkan data gempa bumi di Indonesia bersumber pada USGS, mulai dari Januari 2022 hingga November 2023, ditemukan bahwa ketiga hotel tersebut berada di dekat lokasi gempa. Berdasarkan latitude, longitude, dan Alamat, hotel

dapat dimasukkan ke dalam dua klaster, dengan empat hotel dan tiga hotel sebagai anggota masing-masing klaster. Klasterisasi ini dapat membantu mengidentifikasi hotel-hotel yang memiliki karakteristik serupa untuk memahami potensi faktor risiko gempa. Hasil penelitian ini dapat membantu para pihak yang bertanggung jawab untuk mengawasi hotel-hotel untuk meningkatkan keselamatan terjadinya gempa di masa mendatang. Dalam penelitian mendatang, diharapkan dapat digunakan algoritma clustering lainnya pada pengelompokan hotel-hotel yang berada di sekitar daerah rawan gempa bumi di Indonesia, sehingga hasil pengelompokan dengan algoritma K-Means akan dibandingkan.

REFERENCES

- [1] I. Setiawan, D. Krismawati, S. Pramana, and E. Tanur, "Klasterisasi Wilayah Rentan Bencana Alam Berupa Gerakan Tanah dan Gempa Bumi di Indonesia," in *Seminar Nasional Official Statistics*, 2022, pp. 669–676.
- [2] Daryono, "Aktivitas Gempa Bumi di Indonesia tahun 2022," Jakarta, 2022.
- [3] A. P. W. Hadi, H. Pratiwi, and I. Slamet, "Pengelompokan Data Gempa Bumi di Indonesia dengan Algoritma K-Means dan DBSCAN," in *Seminar Nasional Pendidikan Matematika Ahmad Dahlan 2023*, 2023, pp. 52–60.
- [4] M. Muhajir and N. N. Sari, "K-Affinity Propagation (K-AP) and K-Means Clustering for Classification of Earthquakes in Indonesia," 2018 Int. Symp. Adv. Intell. Informatics, pp. 6–10, 2018.
- [5] A. Ahmad and N. Irsalinda, "Cluster Analysis of Earthquake ' s Data Clustering in Indonesia using Fuzzy K-means Clustering," vol. 3, no. April, pp. 3–7, 2020.
- [6] I. H. Rifa, H. Pratiwi, N. Sciences, and U. S. Maret, "Clustering of earthquake risk in Indonesia using k-medoids and k-means algorithms," *Media Stat.*, vol. 13, no. 2, pp. 194–205, 2020, doi: 10.14710/medstat.13.1.194-205.
- [7] J. Han, M. Kamber, and J. Pei, *Data Mining: Concepts and Techniques Edition: 3rd*. Amsterdam: Morgan Kaufmann-Elsevier, 2022.
- [8] I. H. Rifa, H. Pratiwi, and Respatiwan, "Implementasi Algoritma Clara untuk Data Gempa Bumi di Indonesia," in *Seminar Nasional Penelitian Pendidikan Matematika (SNP2M) 2019*, 2019, pp. 161–166.
- [9] F. Reviantika et al., "Analisis Gempa Bumi Pada Pulau Jawa Menggunakan Clustering Algoritma K-Means," *J. Din. Inform.*, vol. 9, no. 1, pp. 51–60, 2020.
- [10] A. Amalia, U. Harmoko, and G. Yuliyanto, "Clustering of Seismicity in the Indonesian Region for the 2018-2020 Period using the DBSCAN Algorithm," *J. Phys. Its Appl.*, vol. 4, no. 1, pp. 1–6, 2021.
- [11] F. R. Senduk and F. Nhita, "Clustering of Earthquake Prone Areas in Indonesia Using K-Medoids Algorithm," *Ind. J. Comput.*, vol. 4, pp. 65–76, 2019, doi: 10.21108/indoic.2019.4.3.359.
- [12] A. S. Ahmar, D. Napitupulu, and R. Rahim, "Using K-Means Clustering to Cluster Provinces in Indonesia Using K-Means Clustering to Cluster Provinces in Indonesia," in *2nd International Conference on Statistics, Mathematics, Teaching, and Research*, 2018, pp. 0–6.
- [13] I. M. B. Sukmadana and C. Ramadhani, "Pengklasteran Kejadian Gempa Wilayah Indonesia Menggunakan Algoritma k-Means," vol. 8, no. 2, pp. 62–67, 2022.
- [14] I. Rahadi, B. A. Novianti, K. Syahidi, M. Putra, M. Gazali, and R. H. Hirzi, "Applying K-Means Algorithm for Clustering Analysis Earthquakes Data in West Nusa Tenggara Province," *Indones. Phys. Rev.*, vol. 5, no. 3, pp. 197–207, 2022, doi: <https://doi.org/10.29303/ipr.v5i3.148>.
- [15] A. P. Joshi and B. V. Patel, "Data Preprocessing : the Techniques for Preparing Clean and Quality Data for Data Analytics Process," *Orient. J. Comput. Sci. Technol.*, vol. 13, no. 2–3, pp. 78–81, 2020.
- [16] S. Nath, "Analyzing earthquakes in USA to determine the possibly risky to quakes hotels.," *Towards Data Science*, 2019. <https://towardsdatascience.com/analyzing-the-earthquakes-in-usa-to-determine-the-possibly-risky-to-quakes-hotels-2a9ff162e747> (accessed Nov. 17, 2023).
- [17] S. S. Nagari, L. Inayati, U. Airlangga, E. Java, M. V. Midwife, and E. Java, "Implementation of Clustering Using K-Means Method To Determine Nutritional Status," *J. Biometrika dan Kependud.*, vol. 9, no. July, pp. 62–68, 2020, doi: 10.20473/jbk.v9i1.2020.62.
- [18] G. Shmueli, P. C. Bruce, I. Yahav, N. R. Patel, and K. C. Lichtendahl, *Data Mining for Business Analytics : Concepts, Techniques, and Application in R*, First. John Wiley & Sons, Inc, 2018.
- [19] A. Wanto, M. N. H. Siregar, A. P. Windarto, and D. Hartama, *Data Mining : Algoritma & Implementasi*, 1st ed. Medan: Yayasan Kita Menulis, 2020.
- [20] M. S. Yana, L. Setiawan, E. M. Ulfa, and A. Rusyana, "Penerapan Metode K-Means dalam Pengelompokan Wilayah Menurut Intensitas Kejadian Bencana Alam di Indonesia Tahun 2013-2018," *J. Data Anal.*, vol. 1, no. 2, pp. 93–102, 2018.