

BAB I

PENDAHULUAN

1.1 Latar Belakang

Gempa bumi menjadi salah satu bencana alam yang kerap melanda di Indonesia. Fenomena ini melibatkan adanya getaran dan guncangan yang muncul di permukaan bumi. Kebanyakan gempa bumi diakibatkan adanya pelepasan kumpulan energi akibat tekanan lempeng yang bergerak. Tekanan tersebut mengalami peningkatan hingga akhirnya mencapai titik di mana tepi lempeng tidak mampu lagi menahan tekanan sehingga terjadi gempa bumi [1]. Besaran gempa berkisar dari yang cukup kecil hingga sulit dirasakan sampai dengan yang besar atau kuat. Kekuatan gempa bumi dikenal dengan magnitudo. Salah satu dampak gempa bumi adalah bangunan stabil dapat runtuh akibat percepatan tanah [2].

Indonesia merupakan negara yang sering kali terjadi gempa bumi. Berdasarkan aktivitas tektoniknya, terdapat tujuh wilayah di Indonesia yang berpotensi besar munculnya gempa bumi yaitu Sumatra, Jawa, Sulawesi, Nusa Tenggara, Papua, Laut Maluku, dan Laut Banda [3]. Besarnya resiko dari terjadinya gempa di Indonesia disebabkan kondisi dari geografis wilayahnya yang bertempat di bertemunya 3 lempeng-lempeng tektonik yaitu Lempeng Eurasia, Lempeng Indo-Australia, dan Lempeng Pasifik [4]. Berdasarkan astronomisnya, wilayah Indonesia bertempat di antara 6°LU - 11°LS serta 95°BT - 141°BT [5]. Adanya pertemuan lempeng tersebut disebut subduksi yang menyebabkan seringkali adanya bencana alam gempa bumi [6].

Kejadian gempa bumi yang sering kali di wilayah Indonesia membuat jumlah data menjadi cukup besar. Setiap kejadian dari gempa bumi tersebut sangat beragam baik dari besarnya magnitudo maupun lokasi pusat gempa bumi. Menurut data dari Badan Geologi Amerika Serikat yaitu *United States Geological Survey* (USGS) tercatat sebanyak 7.746 kejadian gempa bumi di Indonesia sepanjang tahun 2012 hingga tahun 2023 dengan besar magnitudo dan kedalaman yang berbeda-beda. Banyaknya data gempa bumi dapat dikelompokkan untuk mengetahui karakteristik

tertentu seperti lokasi, kedalaman, magnitudo, dan waktu di suatu wilayah. Faktor-faktor penyebab kerusakan akibat gempa bumi antara lain skala gempa atau magnitudo, kedalaman gempa, jarak ke pusat gempa, lama guncangan gempa, kondisi tanah setempat, dan kondisi bangunan [7].

Jumlah data yang cukup besar rumit dan kompleks akan tidak mudah untuk dapat dianalisis dengan metode konvensional. Sehingga perlu memanfaatkan teknologi untuk mendapatkan informasi dari data salah satunya yaitu pendekatan *data mining* [8]. *Data mining* adalah suatu proses mengambil dan identifikasi informasi dan pemahaman yang berguna dari dataset besar menggunakan metode terkini [9]. Metode *Clustering* menjadi salah satu dari beberapa metode *data mining* yang ada [10]. *Clustering* adalah metode untuk mengelompokkan dan membagi pola data menjadi beberapa kumpulan, memisahkannya dengan membentuk pola yang serupa dalam satu kelompok dan pola yang berbeda dalam kelompok yang lain. *Clustering* memungkinkan identifikasi pola dalam data gempa bumi yang bervariasi [11].

Terdapat beberapa algoritma yang termasuk dalam clustering. Algoritma yang umum digunakan yaitu algoritma K-Means yang merupakan metode untuk mengelompokkan data ke dalam beberapa kelompok sesuai karakteristik data yang sama. Algoritma K-Means memiliki kelebihan yakni dapat mengumpulkan data dengan kinerja yang cukup cepat serta efisien dalam jumlah yang besar [11]. Metode K-Means adalah metode *clustering* non-hirarki yang melakukan pengelompokan kumpulan data ke sejumlah kelompok yang telah ditetapkan sebelumnya. [12].

Penelitian terkait pengelompokkan gempa bumi menggunakan algoritma K-Means telah banyak dilakukan. Penelitian [2] membahas pengelompokkan gempa bumi dari rentang tanggal 09/12/2018 hingga tanggal 05/03/2020 di Pulau Jawa dengan algoritma K-Means kemudian dibandingkan dengan algoritma DBSCAN. Hasil penelitian menyatakan bahwa ketika algoritma DBSCAN dibandingkan dengan algoritma K-Means dalam melakukan clustering gempa bumi diperoleh bahwa algoritma K-Means lebih unggul. Penelitian [13] menjelaskan pengelompokkan kumpulan data lokasi gempa mulai rentang tahun 2013 sampai

2018 di Pulau Sumatra menggunakan algoritma *Fuzzy Possibilistic C-Means*. Hasil penelitian menunjukkan data kluster berdasarkan karakter kedalaman saja, sedangkan magnitudo gempa tersebar di setiap kluster. Evaluasi terhadap hasil pengelompokan menjadi langkah penting, seperti yang dilakukan dalam penelitian [14], menggunakan metode *Sum of Square Error* (SSE) untuk mengevaluasi hasil penentuan jumlah k terbaik. Evaluasi tersebut dilakukan dengan menghitung nilai SSE dari jumlah kelompok 1 sampai kelompok 9. Hasil perhitungan diperoleh jumlah k terbaik adalah 2 kelompok dengan nilai selisih SSE 4611.379920.

Berdasarkan hal tersebut, maka akan dilakukan penelitian pengelompokan gempa bumi di wilayah yang berbeda yaitu Indonesia menggunakan algoritma K-Means. Data gempa bumi yang digunakan juga berbeda yaitu dari Januari 2012 sampai dengan Juni 2024. Hasil pengelompokan penelitian ini akan dievaluasi dengan *Sum of Square Error* (SSE), *Davies-Bouldin Index* (DBI), dan *Silhouette Score*. Dengan dilakukan pengelompokan gempa bumi berdasarkan latitude, longitude, depth, magnitude, dan place sebagai informasi pendukung untuk mengetahui lokasi gempa bumi yang diharapkan dapat bermanfaat dalam membantu upaya tanggap bencana di Indonesia. Hal ini penting untuk memahami karakteristik atau profil gempa bumi di Indonesia, yang kemudian dapat digunakan dalam perencanaan mitigasi risiko.

1.2 Rumusan Masalah

Rumusan masalah pada penelitian ini adalah gempa bumi seringkali terjadi di Indonesia setiap tahunnya. Setiap titik lokasi gempa bumi memiliki keragaman kedalaman dan magnitudo. Untuk mengetahui karakteristik yang sama antar satu titik yang lain perlu dilakukan pengelompokan data menggunakan salah satu metode data mining yaitu K-Means.

1.3 Pertanyaan Penelitian

1. Bagaimana karakteristik gempa berdasarkan hasil pengelompokan menggunakan K-Means?
2. Berapa jumlah nilai k yang paling optimal?

1.4 Tujuan Penelitian

1. Mengelompokkan gempa bumi di Indonesia menggunakan algoritma K-Means
2. Mengetahui jumlah nilai k yang optimal berdasarkan nilai *Sum of Square Error* (SSE), *Davies-Bouldin Index* (DBI), dan *Silhouette Coefficient*

1.5 Batasan Masalah

Batasan masalah dari penelitian ini adalah sebagai berikut:

1. Data yang digunakan adalah data sekunder gempa bumi di Indonesia dari Januari 2012 sampai Juni 2024 yang diambil dari website *United States Geological Survey* (USGS)
2. Jumlah dataset adalah 8.070 baris dan 5 kolom
3. Pengelompokan dilakukan menggunakan algoritma K-Means
4. Pengelompokan gempa bumi dilakukan berdasarkan latitude, longitude, depth, magnitudo dan untuk mengetahui posisi pusat gempa place
5. Evaluasi hasil jumlah kelompok dianalisis menggunakan *Sum of Square Error* (SSE), *Davies-Bouldin Index* (DBI), dan *Silhouette Coefficient*

1.6 Manfaat Penelitian

Manfaat yang dapat diperoleh dari penelitian ini adalah sebagai berikut:

1. Mengetahui pengelompokan gempa bumi di Indonesia berdasarkan latitude, longitude, depth, dan magnitudo
2. Menambah referensi bagi para ilmuwan dan ahli gempa bumi untuk memahami karakter istik gempa bumi di wilayah Indonesia
3. Pengelompokan gempa bumi menggunakan K-Means dapat memberikan gambaran yang lebih jelas tentang pola dan karakteristik gempa bumi di berbagai wilayah, yang berguna untuk mendukung perencanaan kebijakan pembangunan yang lebih aman dan berkelanjutan di daerah rawan gempa, ini erat kaitannya dengan tujuan Sustainable Development Goals (SDGs) nomor 11 yaitu sustainable cities and communities.