

PERBANDINGAN ANTARA ALGORITMA K-MEANS DAN ALGORITMA BISECTING K-MEANS DALAM MENGANALISIS GEMPA BUMI DI INDONESIA

Adela Tania^{*1}, Teny Handhayani², Janson Hendryli³
Fakultas Teknologi Informasi Universitas Tarumanagara
e-mail : tenyh@fti.untar.ac.id

Abstrak Gempa bumi, khusus nya gempa tektonik adalah gempa yang paling sering terjadi di Indonesia. Hal itu dikarenakan kondisi geografis Indonesia yang terletak pada daerah pertemuan 3 batas lempeng tektonik dunia. Kondisi tersebut mendorong berbagai pihak dalam berupaya untuk siaga saat muncul potensi yang dapat ditimbulkan. Salah satu upaya yang dilakukan adalah dengan mengelompokkan wilayah kejadian gempa bumi di Indonesia yang memiliki potensi akan rawan terjadinya gempa bumi berdasarkan kedalaman dan kekuatan gempa bumi dengan menggunakan metode clustering. Metode clustering yang digunakan adalah Algoritma K-Means. Tujuan dari penelitian ini adalah untuk menganalisa pola spasial dari persebaran gempa di Indonesia. Data yang digunakan adalah data titik gempa di seluruh daerah di Indonesia dari November 2008 hingga Juni 2022 yang dicatat oleh Badan Meteorologi Klimatologi dan Geofisika (BMKG). Hasil clustering dengan menggunakan Algoritma K-Means menghasilkan 3 cluster dengan nilai rata-rata Silhouette Coefficient yaitu 0.7390 dan Davies Bouldin Index yaitu 0.4475. Selain itu dari penelitian ini juga didapatkan bahwa Algoritma K-Means memiliki nilai rata-rata Silhouette Coefficient dan Davies Bouldin Index lebih baik dibandingkan dengan Algoritma Bisecting K-Means.

Kata Kunci—Algoritma Bisecting K-Means, Algoritma K-Means, Clustering, Gempa Bumi, Indonesia

I. PENDAHULUAN

Gempa bumi adalah peristiwa terjadinya getaran di permukaan bumi akibat adanya pergerakan pada lapisan luar bumi yang menimbulkan gelombang seismik[1]. Berdasarkan penyebabnya, gempa bumi dibedakan menjadi tektonik, vulkanik, puing, dan buatan[2]. Gempa tektonik merupakan gempa yang paling sering terjadi di Indonesia. Hal ini karena Indonesia merupakan negara maritim atau dikenal juga sebagai negara kepulauan dengan jumlah pulau sebanyak 16.771 buah. Secara geografis, Indonesia berada pada pertemuan 3 batas lempeng tektonik dunia (Eurasia, Indo-Australia, dan Pasifik) yang membuat Indonesia rawan gempa[3].

Menurut data Badan Meteorologi, Klimatologi, dan Geofisika (BMKG), sejak tahun 2008 hingga 2021 frekuensi gempa bumi di Indonesia terus meningkat setiap tahunnya dimana telah terjadi 185 kali gempa yang merusak dimana gempa tersebut mengakibatkan kerugian materil dan immateril seperti korban jiwa, kerusakan bangunan, dan lain-lain, serta gempa bumi dengan kekuatan kurang dari 5 Skala Kecil (SR) juga sering terjadi di Indonesia setiap hari; yang biasa disebut sebagai gempa mikro. Hal ini mendorong berbagai pihak untuk waspada ketika potensi yang dapat ditimbulkan dari kondisi geografis Indonesia. Salah satu upaya yang dapat dilakukan sebagai langkah awal pencegahan adalah dengan melakukan analisis yang dapat mengelompokkan wilayah kejadian gempa bumi di Indonesia yang berpotensi rawan gempa berdasarkan kedalaman dan kekuatan gempa dengan menggunakan metode clustering.

Metode *clustering* adalah metode yang melakukan pembentukan kelompok dimana objek akan ditempatkan ke dalam kelompok yang sama dengan objek lain yang terkait tetapi akan menempati kelompok yang berbeda ketika objek tidak memiliki keterkaitan, yang dibagi menjadi dua metode yaitu *clustering* hirarkis dan partisi[4]. Metode hirarki mengelompokkan data dengan cara membuat hirarki berupa dendrogram dengan data yang sejenis ditempatkan pada hirarki yang berdekatan sedangkan yang berbeda pada skala yang jauh. Sementara itu, metode partisi membagi data menjadi beberapa kelompok dengan masing-masing kelompok mewakili sebuah *cluster*. Metode *clustering* memiliki beberapa algoritma yang dapat digunakan untuk implementasi antara lain Algoritma K-Means dan Bisecting K-Means. Algoritma K-Means dapat mengklasifikasikan data dalam jumlah besar dengan waktu komputasi yang cepat dan efisien. Algoritma ini adalah algoritma pengelompokan yang menggunakan metode partisi berbasis *centroid*. Sementara itu, Algoritma Bisecting K-Means merupakan variasi dari algoritma K-Means dimana satu *cluster* dibagi menjadi dua sub-cluster untuk membentuk sebuah hirarki.

Penelitian terkait yang pernah dilakukan sebelumnya oleh Pepi Novianto, et al. (2017) yaitu melakukan analisis kejadian gempa bumi di Provinsi Bengkulu menggunakan

Algoritma K-Means dimana menghasilkan jumlah *cluster* yang optimal sebanyak 7 *cluster* berdasarkan pengelompokan episentrum gempa. Hasil analisis yang diperoleh menyatakan bahwa terdapat 5 *cluster* dengan kejadian gempa yang terletak di laut dan daratan, 1 *cluster* terjadi di sekitar kepulauan Mentawai Provinsi Sumatera Barat dan 1 *cluster* lainnya terjadi di laut sekitar Provinsi Lampung. Lalu penelitian terkait lainnya yang dilakukan oleh Ferin Reviantika, et al. (2020) adalah melakukan perbandingan antara Algoritma K-Means dan Algoritma DBSCAN pada data kejadian gempa bumi di pulau Jawa dengan rentang magnitudo > 3 pada periode 2018-2020, dimana hasil analisis yang diperoleh bahwa pembagian *cluster* pada Algoritma DBSCAN lebih banyak yaitu 3 *cluster* sedangkan pada Algoritma K-Means sebanyak 2 *cluster*. Apabila dilihat dari nilai indeks metode Silhouette antara kedua metode tersebut, Algoritma K-Means lebih unggul 0,3733 dari pada Algoritma DBSCAN.

Oleh karena itu, penelitian ini bertujuan untuk mengelompokkan data gempa bumi di Indonesia menggunakan perbandingan hasil *cluster* dari algoritma K-Means dan Bisecting K-Means untuk menentukan metode pengelompokan yang tepat.

II. LANDASAN TEORI

A. Gempa Bumi

Gempa bumi merupakan salah satu dari bencana alam yang ada di Indonesia dimana terjadi akibat adanya getaran di permukaan bumi yang dihasilkan dari pergerakan lempeng-lempeng tektonik sehingga menciptakan gelombang seismik. Pergerakan satu lempeng tektonik relatif terhadap lempeng lainnya menghasilkan 3 kemungkinan, yaitu kedua lempeng dapat bergerak saling menjauhi (*spreading*), saling mendekati (*collision*), atau saling bergeser (*transform*) dimana umumnya gerakan ini berlangsung lambat dan tidak dapat dirasakan oleh manusia namun terukur terdapat pergerakan sebesar 0-15 cm pertahunnya. Gelombang seismik sendiri merupakan gelombang yang menjalar ke segala arah di seluruh bagian bumi yang disebabkan oleh deformasi struktur akibat terdapat tekanan maupun tarikan karena keelastisitasan bumi[5]. Karakteristik dari gempa bumi lainnya antara lain yaitu terjadi di lokasi tertentu, berpotensi akan terulang kembali, belum dapat diprediksi, dan lain-lain.

Apabila terjadi gempa bumi dan dalam penanganannya tidak dilakukan dengan baik maka akan menimbulkan ancaman dari segi materil dan immateril. Berikut ini faktor-faktor yang dapat mengakibatkan kerusakan akibat dari gempa bumi, yaitu kekuatan yang dihasilkan dari gempa bumi, kedalaman gempa bumi, jarak hiposentrum gempa bumi, lama getaran gempa bumi, kondisi tanah setempat, dan kondisi bangunan.

B. Data Mining

Data mining terdiri dari kata data dan *mining*, dimana kata *mining* sendiri berartikan usaha dalam mendapatkan hal berharga dari sejumlah besar material dasar. Maka

dapat didefinisikan data mining atau sering disebut juga *Knowledge Discovery in Database* (KDD) merupakan proses pengumpulan informasi penting dari data dalam jumlah yang banyak dengan penggunaan teknik matematika, statistik, dan kecerdasan buatan[6]. Fungsi dari data mining terbagi menjadi 2 yaitu, deskriptif dan prediktif. Namun, apabila terdapat penambahan data maka memiliki fungsi lain seperti asosiasi, klasifikasi, *clustering*, *forecasting*, dan *sequencing*.

Data mining didefinisikan sebagai proses untuk menemukan pola dalam data secara otomatis dan semiotomatis. Pola-pola yang ditemukan harus bermakna dan memberikan keuntungan. Pengenalan pola adalah bidang yang meneliti bagaimana dalam mengklasifikasikan objek ke dalam beberapa kelas atau kategori dan mengidentifikasi tren dalam data. Proses data mining terdiri dari pembersihan data, integrasi data, seleksi data, transformasi data, proses *mining*, evaluasi pola, dan presentasi pengetahuan[7].

C. Algoritma K-Means

Algoritma K-Means merupakan salah satu algoritma metode *clustering* yang menerapkan metode *non-hierarchical* atau *partitional* dimana mempartisi objek ke dalam satu atau lebih *cluster* dengan memiliki kesamaan karakteristik antar objek. Konsep yang diterapkan dari algoritma ini adalah dengan melakukan pencarian pusat *cluster* secara iteratif[8]. Berikut ini langkah-langkah dari Algoritma K-Means.

- 1) Tentukan banyak nilai *cluster* *K*
- 2) Menentukan titik pusat (*centroid*) secara acak.
- 3) Menghitung jarak antara objek dan titik pusat menggunakan metode Euclidian Distance (1).

$$d_{(i,j)} = \sqrt{\sum_{k=1}^n (x_{ik} - x_{jk})^2} \tag{1}$$

$d_{(i,j)}$ = jarak antar objek *i* dengan *j*

x_{ik} = nilai objek *i* pada variabel ke-*k*

x_{jk} = nilai objek *j* pada variabel ke-*k*

n = banyaknya variabel yang diamati

- 4) Mengelompokkan objek sesuai dengan nilai minimum dari jarak objek ke titik pusat.

- 5) Menghitung nilai rata-rata dari objek yang terdapat dalam satu *cluster* (2).

$$\mu_k = \frac{1}{N_k} \sum_{q=1}^{N_k} x_q \tag{2}$$

μ_k = titik pusat dari *cluster* ke-*k*

N_k = banyaknya data pada *cluster* ke-*k*

x_q = data ke-*q* pada *cluster* ke-*k*

6) Apabila titik pusat berubah, maka dapat mengulangi kembali langkah 3-5 hingga tidak ada lagi objek yang berpindah *cluster*.

D. *Algoritma Bisecting K-Means*

Algoritma Bisecting K-Means merupakan algoritma variasi dari Algoritma K-Means dengan menerapkan metode clustering hirarki yang dimana pada setiap langkah pembagian cluster dilakukan pembagian menjadi 2 sub *cluster* menggunakan Algoritma K-Means. Selain itu, Bisecting K-Means efektif dalam mengatasi situasi apabila memasuki kondisi optimal lokal hingga batas tertentu. Bisecting K-Means juga termasuk ke dalam *hard clustering* dimana setiap objek hanya memiliki satu keanggotaan *cluster*.

E. *Silhouette Coefficient*

Metode Silhouette Coefficient merupakan metode evaluasi cluster untuk melihat kualitas seberapa baik suatu objek ditempatkan dalam suatu *cluster* dengan menggabungkan metode *cohesion* dan *separation* [9]. Tahap perhitungan untuk mendapatkan nilai silhouette sebagai berikut.

1) Hitung rata-rata jarak dari suatu objek ke-*i* dengan semua objek lain yang berada dalam satu *cluster* (diasumsikan apabila data ke-*i* berada pada *cluster* A) (3).

$$a(i) := \frac{1}{|A| - 1} \sum_{j \in A, j \neq i} d(i, j) \tag{3}$$

A = banyaknya data di *cluster* A

j = objek lain dalam satu *cluster* A

d(i, j) = jarak antara objek ke-*i* dengan *j*

2) Hitung rata-rata jarak dari objek ke-*i* dengan semua objek yang berada pada *cluster* yang berbeda dan mengambil hasil dengan nilai terkecil (diasumsikan bahwa cluster yang berbeda selain *cluster* A yaitu *cluster* C) (4), (5).

$$d(i, C) := \frac{1}{|C|} \sum_{j \in C} d(i, j) \tag{4}$$

Dengan

$$b(i) := \min_{C \neq A} d(i, C) \tag{5}$$

d(i, C) = jarak rata-rata objek ke-*i* dengan semua objek pada *cluster* lain

C = banyak data di *cluster* C

3) Setelah nilai *a(i)* dan *b(i)*, menghitung nilai silhouette coefficient (6).

$$s(i) = \frac{b(i) - a(i)}{\max [a(i), b(i)]} \tag{6}$$

Nilai *s(i)* berada di antara -1 dan 1, dimana setiap nilai dapat diinterpretasi sebagai berikut¹.

s(i) ≈ 1 => data ke-*i* digolongkan dengan baik.

s(i) ≈ 0 => data ke-*i* berada di tengah antara dua *cluster*.

s(i) ≈ -1 => data ke-*i* digolongkan dengan lemah.

F. *Davies Bouldin Index*

Davies Bouldin Index (DBI) pertama diperkenalkan oleh David L. Davies dan Donald W. Bouldin pada tahun 1979 dengan tujuan untuk mengevaluasi hasil dan untuk menentukan jumlah cluster yang paling optimal dalam proses *clustering* [10]. Tahap perhitungan Davies Bouldin Index (DBI) adalah sebagai berikut.

1) Hitung terlebih dahulu varians dari masing-masing *cluster* (7).

$$var(x) = \frac{1}{N - 1} \sum_{i=1}^N (x_i - \bar{x})^2 \tag{7}$$

N = jumlah anggota *cluster*

x_i = *cluster* ke-*i*

x̄ = rata-rata dari *cluster* x

2) Menghitung nilai DBI (8), (9), (10).

$$DBI = \frac{1}{k} \cdot \sum_{i=1}^k R_i \tag{8}$$

Dengan

$$R_i = \max_{j = 1, \dots, k, i \neq j} R_{ij} \tag{9}$$

Dan

$$R_{ij} = \frac{var(C_i) + var(C_j)}{\|c_i - c_j\|} \tag{10}$$

C_i = *cluster* ke-*i*

C_j = *cluster* ke-*j* (*cluster* lain selain *cluster* i)

c_i = *centroid* dari *cluster* ke-*i*

c_j = *centroid* dari *cluster* ke-*j*

III. METODE PENELITIAN

Penelitian dilakukan dengan menggunakan tahapan Algoritma K-Means yang dimana hasil *clustering*nya akan dibandingkan dengan algoritma lainnya. Objek penelitian adalah kejadian gempa bumi yang terjadi di Indonesia pada periode waktu mulai dari 1 November 2008 hingga 17 Juni 2022 dengan variabel yang digunakan antara lain yaitu garis lintang (*latitude*), garis bujur (*longitude*), mangitudo, dan kedalaman. Dataset diperoleh melalui *website*

repositori resmi milik BMKG dengan data yang didapatkan sebanyak 86.001 data.

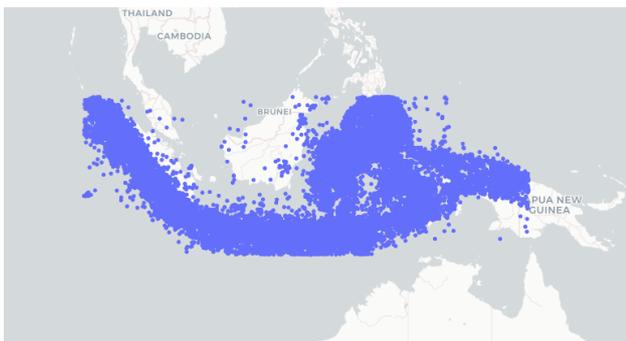
Data yang diperoleh merupakan data bersih yang tidak mengalami *missing value* sehingga hanya dilakukan pemilihan pada kolom data yang akan digunakan untuk penelitian. Eksplorasi data gempa juga dilakukan sebelum proses clustering dengan tujuan untuk mendeskripsikan kejadian gempa bumi di Indonesia berdasarkan variabel-variabel yang digunakan dalam penelitian.

Setelah itu akan dilanjutkan ke tahap pengujian kualitas yaitu dengan menerapkan 2 metode yaitu metode Silhouette Coefficient dan Davies Bouldin Index dalam menentukan jumlah *cluster* yang optimal yang akan diterapkan dalam proses pengelompokan data.

IV. HASIL DAN PEMBAHASAN

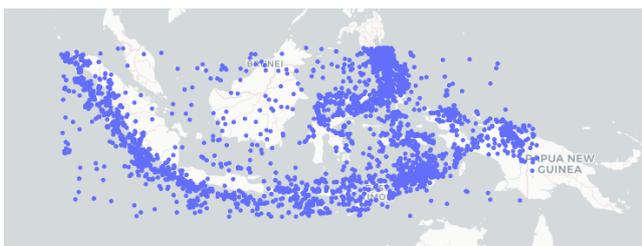
A. Deskripsi Data

Pada penelitian ini, berdasarkan data yang diperoleh menunjukkan bahwa koordinat kejadian gempa bumi terletak pada 94 BT (Bujur Timur) – 142 BT dan -11 LS (Lintang Selatan) dan 6 LU (Lintang Utara). Penyebaran kejadian gempa bumi berdasarkan magnitudo dapat dilihat melalui Gambar 1 hingga Gambar 3.



Gambar 1. Peta penyebaran gempa bumi berdasarkan magnitudo < 5 SR

Pada Gambar 1 merupakan visualisasi dalam bentuk peta dari data aktual kejadian gempa bumi di Indonesia yang memiliki magnitudo kurang dari 5 SR yang dimana memiliki kejadian yang sangat banyak dikarenakan Indonesia setiap harinya dapat terjadi gempa bumi mikro kecuali pada Pulau Kalimantan yang memiliki kejadian gempa bumi yang sedikit dikarenakan Pulau Kalimantan terletak di daerah yang jauh dari zona tumbukan lempeng dan tidak memiliki gunung berapi yang aktif.



Gambar 2. Peta penyebaran gempa bumi berdasarkan magnitudo 5 - 7 SR

Pada Gambar 2 merupakan visualisasi dalam bentuk peta dari data aktual kejadian gempa bumi di Indonesia yang memiliki magnitudo mulai dari 5 SR hingga 7 SR,

dimana dapat dilihat bahwa seluruh wilayah di Indonesia pernah mengalami kejadian gempa bumi dengan kekuatan magnitudo 5 SR hingga 7 SR.



Gambar 3. Peta penyebaran gempa bumi berdasarkan magnitudo > 7 SR

Pada Gambar 3 merupakan visualisasi dalam bentuk peta dari data aktual kejadian gempa bumi di Indonesia yang memiliki magnitudo lebih dari 7 SR. Pada gambar tersebut, didapatkan bahwa kejadian terjadi beberapa kali pada Pulau Sumatra bagian barat, Pulau Jawa bagian timur, Pulau Papua, dan Pulau Sulawesi bagian utara.

Statistik deskriptif dari kejadian gempa bumi di Indonesia dapat dilihat pada Tabel 1 dan Tabel 2.

Tabel 1.

| Statistik deksriptif dari kedalaman (Km) | |
|--|-------|
| Keterangan | Nilai |
| Rata-rata | 49,29 |
| Standart Devisiasi | 77,44 |
| Minimal | 2 |
| Maksimal | 750 |
| Median | 16 |
| Modus | 10 |

Pada tabel 1, menunjukkan bahwa kedalaman rata-rata pada kejadian gempa bumi di Indonesia adalah 49,29 km dan memiliki kedalaman minimum 2 km serta maksimum 750km. berdasarkan dari tabel diatas juga terlihat bahwa gempa bumi di Indonesia banyak terjadi di kedalaman 10 km yaitu sebanyak 32.631 kali.

Tabel 2.

| Statistik deksriptif dari magnitudo (SR) | |
|--|-------|
| Keterangan | Nilai |
| Rata-rata | 3,6 |
| Standart Devisiasi | 0,82 |
| Minimal | 1 |
| Maksimal | 7,9 |
| Median | 3,5 |
| Modus | 3,3 |

Berdasarkan pada Tabel 2, rata-rata magnitudo gempa bumi di Indonesia yaitu 3,6 SR dengan memiliki nilai minimum 1 SR dan maksimum 7,9 SR. Gempa bumi dengan magnitudo sebesar 7,9 SR terjadi pada tahun 2009 di Irian Jaya, Papua.

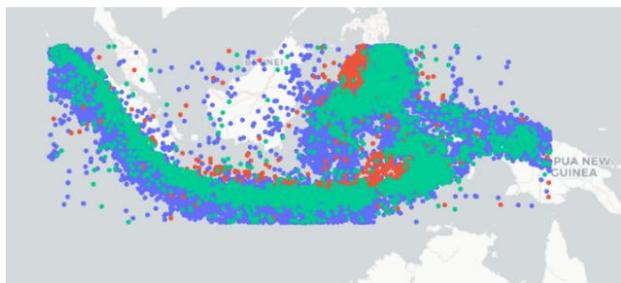
B. Analisis Clustering Dengan Algoritma K-Means

Clustering menggunakan algoritma K-Means dilakukan terhadap 86.001 data kejadian gempa bumi di Indonesia dengan melakukan pengujian dengan metode Silhouette dan Davies Bouldin dalam penentuan jumlah cluster yang optimal mulai dari k=2 hingga k=12. Berikut ini dapat dilihat pada Tabel 3 yang merupakan hasil dari nilai rata-rata Silhouette dan Davies Bouldin untuk algoritma K Means.

Tabel 3.
Statistik deskriptif dari kedalaman (Km)

| Jumlah Cluster | Silhouette Coefficient | Davies Bouldin |
|----------------|------------------------|----------------|
| K=2 | 0,7586 | 0,5647 |
| K=3 | 0,7390 | 0,4475 |
| K=4 | 0,6621 | 0,5458 |
| K=5 | 0,6233 | 0,5512 |
| K=6 | 0,5771 | 0,5970 |
| K=7 | 0,4437 | 0,6858 |
| K=8 | 0,4466 | 0,6669 |
| K=9 | 0,4418 | 0,6839 |
| K=10 | 0,4406 | 0,6985 |
| K=11 | 0,4403 | 0,6950 |
| K=12 | 0,3687 | 0,7464 |

Tabel 3 menunjukkan bahwa jumlah *cluster* yang optimal adalah 3 *cluster* dengan memiliki nilai rata-rata Silhouette dan Davies Bouldin yaitu 0,7390 dan 0,4475. Walaupun nilai rata-rata yang dihasilkan metode Silhouette pada k=2 lebih tinggi dibandingkan dengan k=3, namun nilai Silhouette untuk k=3 masih termasuk nilai lebih dari 0,71 yang dimana bahwa struktur yang dihasilkan kuat. Sedangkan, pada nilai rata-rata yang dihasilkan metode Davies Bouldin untuk k=3 merupakan nilai paling kecil dibandingkan dengan nilai dari jumlah cluster lainnya, dikarenakan pada metode Davies Bouldin semakin kecil nilai yang dihasilkan menunjukkan hasil yang baik.



Gambar 4. Visualisasi dalam bentuk peta hasil clustering dengan k=3

Pada gambar 4, menunjukkan informasi ditunjukkan dengan *cluster* 1 menggunakan warna ungu yang terdiri dari 70.575 anggota yang berisikan gempa dangkal dengan kedalaman 2 km hingga 84 km dan magnitudo 1 SR hingga 7,9 SR. Pada *cluster* 2 menggunakan warna kuning terdiri dari 1.479 anggota yang berisikan gempa dalam dengan kedalaman 306 km hingga 750 km dan magnitudo 2,1 SR hingga 7,2 SR. Sedangkan *cluster* 3 menggunakan warna hijau memiliki anggota sebanyak 13.947 yang berisikan data gempa sedang dengan kedalaman 84 km – 305 km dan magnitudo 1,1 SR hingga 7,3 SR.

Tabel 4.

Frekuensi kejadian gempa bumi pada masing-masing cluster berdasarkan magnitudo

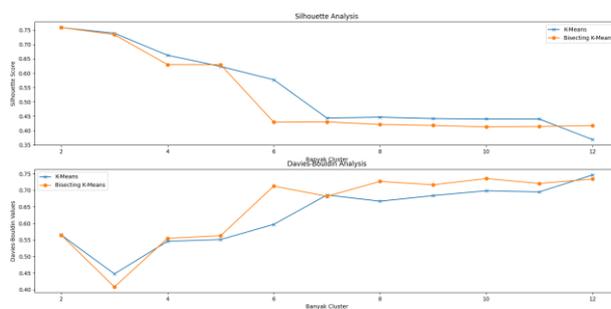
| Magnitudo | Cluster | | |
|-----------|---------|-------|--------|
| | 1 | 2 | 3 |
| < 5 | 66.988 | 1.287 | 12.932 |
| 5 – 7 | 3.562 | 190 | 1.007 |
| > 7 | 25 | 2 | 8 |

Berdasarkan tabel 4, semua besaran tersebar di semua *cluster*. Namun jika diperhatikan lebih lanjut, pada *cluster* 1 merupakan cluster yang memiliki jumlah kejadian gempa

terbanyak dan memiliki kejadian gempa yang paling signifikan hasilnya untuk semua besaran magnitudo. Gempa dengan magnitudo lebih kecil dari 5 SR memiliki persentase sebesar 82,49%, magnitudo antara 5 hingga 7 SR memiliki persentase sebesar 74,84%, dan magnitudo lebih besar dari 7 SR memiliki persentase sebesar 71,42%.

C. Analisis Algoritma K-Means vs Algoritma Bisecting K-Means

Penelitian ini juga menggunakan Algoritma Bisecting K-Means sebagai pembandingan dengan metode yang digunakan yaitu Algoritma K-Means. Tahapan *preprocessing* data yang dilakukan untuk pengolahan data dengan Algoritma Bisecting K-Means sama dengan yang dilakukan pada Algoritma K-Means. Pengujian dilakukan dengan metode Silhouette dan Davies Bouldin, didapatkan bahwa Algoritma Bisecting K-Means juga menghasilkan k=3 untuk jumlah cluster yang optimal dengan hasil nilai rata-rata yaitu 0,7340 dan 0,4081.



Gambar 5. Visualisasi algoritma K-Means vs algoritma Bisecting K-Means dengan metode pengujian Silhouette dan Davies Bouldin

Pada gambar 5 juga dapat dilihat hasil perbandingan dari Algoritma K-Means dan Bisecting K-Means yang menunjukkan bahwa Algoritma K-Means memiliki hasil yang lebih baik dan unggul jika dibandingkan dengan Bisecting K-Means. Walaupun perbedaan nilai yang diperoleh antara kedua algoritma kecil.

V. KESIMPULAN DAN SARAN

Dari penelitian ini dapat disimpulkan bahwa:

- 1) Pengelompokkan menggunakan Algoritma K-Means berhasil dilakukan terhadap 86.001 data dengan menghasilkan jumlah cluster yang optimal yaitu sebanyak 3 *cluster*. Hasil nilai rata-rata pengujian yang diperoleh dengan metode Silhouette dan Davies Bouldin yaitu 0.7390 dan 0.4475.
- 2) Hasil *clustering* pada *cluster* 1 merupakan daerah yang perlu mendapatkan perhatian lebih dalam penerapan mitigasi akan terjadinya gempa bumi dikarenakan memiliki kedalaman yang dangkal di bawah 100 km dan memiliki persentase magnitudo di semua besaran di atas 70%. Kejadian gempa bumi yang terjadi pada kedalaman dangkal akan memberikan efek getaran yang lebih kuat dan akan mempengaruhi efek kerusakan yang diberikan.
- 3) Dari penelitian ini juga ditemukan bahwa Algoritma K-Means memiliki kinerja yang lebih baik daripada

Algoritma Bisecting K-Means sebagai metode analisis gempa bumi.

Pada penelitian selanjutnya disarankan untuk dapat menggunakan metode lainnya seperti Intelligent K-Means agar dapat dilakukan perbandingan hasil untuk mendapatkan hasil cluster yang lebih baik dan menambahkan analisis bahaya atau kerusakan yang diakibatkan oleh gempa bumi agar lebih dapat memperkirakan upaya mitigas yang lebih tepat.

UCAPAN TERIMA KASIH

Penulis A.T. mengucapkan terima kasih kepada Direktorat Pendidikan Tinggi, Departemen Pendidikan dan Kebudayaan Republik Indonesia, Lembaga Penelitian dan Pengabdian Kepada Masyarakat (LPPM) Universitas Tarumanagara, dan Fakultas Teknologi Infomasi atas dukungan dana dan bimbingannya sehingga penelitian ini dapat terlaksana dengan baik.

DAFTAR PUSTAKA

- [1] D.S, Rini, I. Sriliana, P. Novianti, S. Nugroho, and P. Jana, "Spherical K-Means Method To Determine Earthquake Clusters", *Journal of Physics: Conference Series* (2021), 1.
- [2] S. Urip, "Katalog Gempa Bumi Signifikan dan Merusak 1821 – 2018", Badan Meteorologi Klimatologi dan Geofisika (BMKG): Jakarta (2019), v
- [3] H.I. Rifa, H. Pratiwi, and Respatiwulan, "Clustering of Earthquake Risk in Indonesia Using K-Medoids and K-Means Algorithms", in *Media Statistika* Vol. 13, No. 2 (2020), 195.
- [4] P. Novianti, D. Setyorini, and U. Rafflesia, "K-Means Cluster Analysis in Earthquake Epicenter Clustering" in *Internasional Journal of Advances in Intelligent Informatics* Vol.3, No. 2, (2017), 82.
- [5] Gemeliarini, I.A. Kusdiah, Helmi, and Muhammad, "Strategi Mitigasi Berdasarkan Model Geospasial Risiko Bencana Gempa Bumi di Kabupaten Lombok Utara Nusa Tenggara Barat", Universitas Diponegoro: Semarang (2022).
- [6] D. Jollyta, W. Ramadhana, and M. Zarlis, "Konsep Data Mining dan Penerapan", Deepublish: Yogyakarta (2020), 45-46.
- [7] J. Han and M. Kamber, "Data Mining: Concepts and Techniques, 2nd Edition", Morgan Kaufmann: San Francisco (2006), 47-94.
- [8] M. Wahyudi, "Data Mining: Penerapan Algoritma K-Means Clustering dan K-Medoids Clustering", Yayasan Kita Menulis:, Medan (2020), 5.
- [9] S. Paeembonan and H. Abduh, "Penerapan Metode Silhouette Coefficient Untuk Evaluasi Clustering Obat", in *Jurnal Ilmiah Ilmu (Ilmu Teknik)*, Vol.6, No.2, (2021, September), 50-51.
- [10] P.N. Tan, M. Steinbach, and V. Kumar, "Introduction to Data Mining", Addison-Wesley: United States (2006).