

Perbandingan *Algoritma Density-Based Spatial Clustering Algorithm with Noise (DBSCAN)* dan *Self-Organizing Map (SOM)* untuk *Clustering Data Gempa Bumi*

ROSITA K. WATI¹, HASIH PRATIWI², WINITA SULANDARI³

^{1, 2, 3}Program Studi Statistika Fakultas MIPA Universitas Sebelas Maret, Indonesia
e-mail: hasihpratiwi@staff.uns.ac.id

ABSTRAK

Gempa bumi merupakan bencana alam yang kerap melanda Indonesia karena letak geografisnya berada pada batas pertemuan tiga lempeng aktif dunia. Dampak kerusakan yang timbul akibat gempa bumi bergantung pada magnitudo dan kedalamannya. Oleh karena itu, perlu upaya mitigasi bencana dan manajemen risiko bencana melalui pengolahan data untuk mengetahui karakteristik dari data gempa tersebut. Penelitian ini bertujuan untuk *clustering* data gempa bumi di Indonesia berdasarkan magnitudo dan kedalaman dengan menerapkan algoritma *Density-Based Spatial Clustering Algorithm With Noise (DBSCAN)* dan *Self-Organizing Map (SOM)* dengan validasi kebaikan *cluster* menggunakan koefisien silhouette. Penerapan algoritma DBSCAN dengan nilai *Eps* dan *MinPts* optimal sebesar 1,6 dan 12 membentuk dua *cluster* dan 23 data diidentifikasi sebagai *noise*, sedangkan menggunakan algoritma SOM dengan *learning rate* 0,05 dan maksimal *epoch* 1.000 membentuk dua *cluster*. Pada analisis ini SOM mampu melakukan *clustering* yang lebih baik jika dibandingkan dengan DBSCAN karena memberikan nilai koefisien *silhouette* yang lebih besar, yaitu sebesar 0,717 sedangkan DBSCAN sebesar 0,677. Hasil *clustering* terbaik memiliki karakteristik yaitu *cluster* 1 dikategorikan sebagai gempa sedang berkekuatan sedang dan *cluster* 2 dikategorikan sebagai gempa dangkal berkekuatan sedang.

Kata Kunci: *Clustering, DBSCAN, Gempa Bumi, SOM.*

ABSTRACT

Earthquakes are natural disasters that occur frequently in Indonesia because of the geographical location at the convergence of three active tectonic plates. The severity of an earthquake's impact is influenced by magnitude and depth. Therefore, disaster mitigation efforts and disaster risk management through data mining are needed to understand the characteristics of earthquakes. This research aims to cluster earthquake data in Indonesia based on magnitude and depth by applying a Density-Based Spatial Clustering Algorithm with Noise (DBSCAN) and Self-Organizing Map (SOM) algorithms and cluster results are evaluated using the silhouette coefficient. Using the DBSCAN algorithm with optimal Eps and MinPts values of 1.6 and 12 formed two clusters and 23 data were identified as noise while using the SOM algorithm with a learning rate of 0.05 and a maximum epoch of 1000 formed two clusters. SOM can perform clustering better than DBSCAN because it provides a larger silhouette coefficient value, which is 0.717 while DBSCAN is 0.677. The clustering results obtained show that cluster 1 is categorized as moderate earthquakes of moderate intensity and cluster 2 is categorized as shallow earthquakes of moderate intensity.

Keywords: *Clustering, DBSCAN, Earthquakes, SOM.*

1. PENDAHULUAN

Indonesia berisiko tinggi mengalami gempa bumi karena posisinya yang berada di pertemuan tiga lempeng tektonik dunia yaitu Hindia-Australia, Pasifik, dan Eurasia yang terus bergerak (Damayanti dkk., 2020; Utomo dkk., 2019). Pergerakan antara ketiga lempeng ini dan kondisi wilayah Indonesia yang berada dalam sistem seismotoniik yang kompleks, berpengaruh pada

tingginya frekuensi kejadian gempa bumi. Berdasarkan informasi dari Data dan Informasi Bencana Indonesia (DIBI) tahun 2022, jumlah gempa bumi yang dirasakan mencapai 776 kejadian dengan kerugian yang ditimbulkan dari bencana ini yaitu 16.500 korban meninggal dan 1.529.254 kerusakan rumah dan fasilitas umum. Berdasarkan kerugian yang ditimbulkan, perlu dilakukan mitigasi bencana untuk memperkecil kerusakan dari bencana tersebut.

Untuk mencapai tujuan mitigasi yang efektif, penelitian ini dilakukan untuk mendapatkan informasi yang lebih mendalam tentang pengelompokan karakteristik gempa bumi di Indonesia. Salah satu langkah yang dapat dilakukan yaitu analisis *clustering* menggunakan teknik *data mining*. *Data mining* merupakan suatu proses guna menemukan dan mengidentifikasi informasi dan pengetahuan yang bermanfaat dari berbagai basis data yang besar dengan menggunakan teknik matematika, statistik, dan *machine learning* (Sharda dkk., 2018). Dalam *data mining* ada beberapa proses, diantaranya: prediksi, klasifikasi, *clustering*, asosiasi dan estimasi. *Clustering* adalah metode untuk mengelompokkan data berdasar atas kesamaan karakteristik yang dimiliki dengan kriteria dalam satu *cluster* yang terbentuk mempunyai kemiripan yang tinggi (Han dkk., 2022; Halim dan Widodo, 2017). Dalam penelitian ini akan dibahas dua algoritma dalam *clustering* yaitu *Density-Based Spatial Clustering Algorithm With Noise* (DBSCAN) dan *Self-Organizing Map* (SOM).

DBSCAN adalah metode *clustering* berbasis densitas yang membentuk *cluster* berdasarkan tingkat kepadatan, dimana dalam satu *cluster* memiliki kepadatan yang tinggi sedangkan yang tidak masuk dalam *cluster* menjadi *noise* (Risman dkk., 2019). Kelebihan algoritma DBSCAN yaitu mampu mengenali adanya *noise*, mampu menangani *cluster* dengan bentuk dan ukuran yang tidak teratur, tahan terhadap *outlier*, dan mampu mendeteksi anomali. Sedangkan SOM adalah metode *clustering* yang berbasis *Artificial Neural Network* (ANN) yang digunakan untuk memetakan data berdasarkan karakteristiknya melalui mekanisme pengaturan diri (Nisrina dkk., 2022). SOM memiliki keunggulan jika dibandingkan dengan algoritma lain yaitu mampu menjalani proses *clustering* dengan tingkat kestabilan yang baik, menghasilkan sedikit percabangan, dan mampu mencapai nilai *centroid* yang konsisten pada setiap pengujian.

Penelitian sebelumnya dilakukan oleh Rahman dan Wijayanto (2021) yang menerapkan algoritma DBSCAN pada data gempa bumi Indonesia dari 5 November hingga 5 Desember 2020, dan berhasil mengidentifikasi tiga *cluster* wilayah yang berisiko mengalami gempa. Dalam studi tersebut, kombinasi nilai *Eps* 0,28 dan *MinPts* 3 yang menghasilkan nilai koefisien *silhouette* sebesar 0,81. Selanjutnya penelitian oleh Hartatik dan Cahya (2020) menggunakan algoritma SOM pada data gempa di Pulau Jawa dari Januari hingga Juni 2019 menghasilkan tiga *cluster* berdasarkan tingkat kerusakan dengan akurasi tertinggi mencapai 0,96.

Berdasarkan pemaparan di atas, peneliti melakukan *clustering* data gempa bumi di Indonesia tahun 2013 hingga 2022 menggunakan algoritma DBSCAN dan SOM. Hasil *clustering* terbaik dari dua algoritma akan digunakan untuk informasi mitigasi bencana di masa mendatang.

2. METODE PENELITIAN

2.1 Sumber Data

Data dalam penelitian ini adalah data sekunder yang didapatkan dari laman United States Geological Survey (USGS, 2023) di <https://earthquake.usgs.gov/earthquakes/>. Penelitian ini menggunakan data gempa bumi Indonesia dari tanggal 1 Januari 2013 sampai dengan 31 Desember 2022. Data yang diambil berjumlah 15.116 dengan 4 variabel, yaitu kedalaman dan magnitudo (digunakan untuk analisis), lintang dan bujur (digunakan untuk pemetaan).

2.2 Metode yang Digunakan

Tujuan dari penelitian ini adalah *clustering* dengan menerapkan algoritma DBSCAN dan SOM kemudian melakukan validasi kebaikan *cluster* menggunakan koefisien *silhouette*. Koefisien *silhouette* yaitu sebuah metode yang diperkenalkan oleh Rousseeuw pada tahun 1987 guna menunjukkan seberapa baik objek terletak dalam suatu *cluster* dibandingkan pada *cluster* lain (Pratiwi dkk., 2021). *Clustering* optimal terjadi jika nilai koefisien *silhouette* mencapai maksimum yaitu mendekati satu, yang berarti meminimalkan jarak dalam satu *cluster* sekaligus memaksimalkan jarak antar-*cluster*.

2.2.1 DBSCAN

DBSCAN adalah algoritma berbasis densitas yang membentuk *cluster* dengan memperhatikan tingkat kepadatan antar objek dimana objek yang cukup tinggi masuk ke dalam satu *cluster*, sementara objek yang tidak masuk ke *cluster* manapun disebut sebagai *noise* (Risman dkk., 2019). Dalam penerapannya, ada dua parameter untuk membentuk *cluster* dengan algoritma ini, yaitu *Epsilon (Eps)* yang merupakan jarak maksimal antara dua data dalam satu *cluster* yang diizinkan dan *minimum points (MinPts)* merupakan jumlah minimum data yang diperlukan agar suatu *cluster* dapat terbentuk. Penentuan jarak maksimal menggunakan konsep jarak Euclidean, yaitu akar kuadrat dari jumlah selisih kuadrat dalam data untuk setiap variabel (Joseph dkk., 2019), dengan rumus pada persamaan (1) berikut:

$$d_{(ij)} = \sqrt{\sum_{i=1}^n (x_i - x_j)^2 + (y_i - y_j)^2} \quad \dots(1)$$

dengan

- $d_{(ij)}$: jarak antara objek i dan objek j ,
- x_i : objek pada pengamatan ke- i dari x ,
- x_j : objek pada pengamatan ke- j dari x ,
- y_i : objek pada pengamatan ke- i dari y ,
- y_j : objek pada pengamatan ke- j dari y ,
- n : banyak pengamatan.

Terdapat tiga jenis tipe titik pada algoritma DBSCAN yaitu : (1) *Core point*, yaitu satu titik data yang terdapat dalam suatu *cluster* serta memiliki jumlah tetangga minimal sebanyak *MinPts*, (2) *border point*, yaitu titik data yang berada dalam suatu *cluster* tetapi tidak memenuhi jumlah tetangga minimal sebanyak *MinPts*, (3) *noise point*, yaitu titik data yang tidak masuk dalam *cluster* manapun.

2.2.2 SOM

SOM adalah teknik pelatihan *Artificial Neural Network* (ANN) yang mengadopsi konsep *winner takes all*, di mana hanya *centroid* data (*neuron*) yang memiliki jarak Euclidean paling dekat dengan bobotnya yang akan diperbarui (Baca dkk., 2005). Walaupun berasal dari keluarga ANN, SOM tidak memerlukan nilai target kelas yang biasanya digunakan dalam tugas pengelompokan. Dalam *clustering* menggunakan algoritma SOM, ada beberapa parameter yang harus diperhatikan yaitu bobot awal, *learning rate*, dan banyaknya *epoch*.

Terdapat tiga komponen utama dalam SOM (Haykin, 1999), yaitu:

- a. Kompetisi merupakan tahap menentukan nilai minimum dari perhitungan jarak vektor $d_{(w_{ij}, x_{in})}$ sebagai *winning neuron*. Perhitungan jarak vektor $d_{(w_{ij}, x_{in})}$ menggunakan rumus pada persamaan (2) berikut:

$$d_{(w_{ij}, x_{in})} = \sqrt{\sum_{i=1}^n (w_{ij} - x_{in})^2} \quad \dots(2) \quad (1)$$

dengan

- $d_{(w_{ij}, x_{in})}$: jarak antara data ke- i pada bobot ke- j dan *input* ke- n ,
 - w_{ij} : nilai bobot data ke- i pada *output* ke- j ,
 - x_{in} : nilai data ke- i pada *input* ke- n .
- b. Kerja sama merupakan tahap *winning neuron* menentukan lokasi spasial lingkungan *topologi excited neuron* untuk membentuk dasar kerjasama dalam suatu lingkungan *neuron*,
 - c. Adaptasi sinaptik merupakan tahap *neuron* yang diaktivasi menurunkan nilai fungsi diskriminan yang terkait dengan pola input melalui penyesuaian bobot terkait, sehingga respons dari *winning neuron* terhadap aplikasi berikutnya dengan pola input yang sama akan meningkat.

Pada penelitian ini, seluruh komputasi dilakukan menggunakan *software Python*. Tahapan analisis yang dilakukan pada penelitian ini adalah sebagai berikut:

- 1) Melakukan *Preprocessing Data*
 - a. *Cleaning data* dilakukan dengan mengecek *missing value* kemudian imputasi *missing value* dengan nilai median.
 - b. Konversi data pada variabel magnitudo menjadi satuan Mw sesuai dengan Tabel 1 menggunakan persamaan (3) sebagai berikut :

$$y = bx + a \quad \dots(3)$$

dengan

y : variabel terikat (Mw)

x : variabel bebas (variabel selain Mw)

b : koefisien variabel x

a : konstanta

Tabel 1. Hasil Penghitungan Hubungan Linear Magnitudo Katalog BMKG

Tipe Magnitudo	Persamaan Linear
MLv	Mw = 0,8736 * MLv + 0,3078
Mb	Mw = 1,198 * mb - 1,2053
mB	Mw = 1,1426 * mB - 1,3386

- c. Transformasi data, dilakukan dengan mengubah skala data ke bentuk lain sesuai dengan distribusi yang diinginkan. Besar kekuatan gempa berbanding lurus dengan efek dari gempa, artinya semakin besar kekuatan gempa semakin besar pula efek yang ditimbulkan (Utomo dkk., 2019). Untuk itu, perlu transformasi data untuk menyelaraskan dampak dari gempa bumi. Untuk langkah transformasi digunakan rumus pada persamaan (4) berikut:

$$h'_i = h_{maks} - h_i \quad \dots(4)$$

dengan

h_i : kedalaman objek ke $-i$,

h_{maks} : nilai maksimum kedalaman gempa, dan

h'_i : hasil transformasi kedalaman objek ke $-i$.

- 2) *Clustering* dengan algoritma DBSCAN
- 3) *Clustering* dengan algoritma SOM
- 4) Membandingkan nilai koefisien *silhouette* algoritma DBSCAN dan SOM.
- 5) Menyusun peta persebaran gempa bumi Indonesia menggunakan hasil *clustering* yang lebih baik untuk mengetahui karakteristik dari data gempa bumi di Indonesia.

Karakteristik data gempa bumi dapat diketahui berdasarkan kedalaman dan kekuatannya (BMKG, 2021). Berdasarkan kedalaman, gempa bumi dangkal (< 70 km) menghasilkan dampak paling besar karena berada dekat dengan permukaan. Gempa sedang (70-300 km) memiliki efek moderat, sementara gempa dalam (>300 km) menyebabkan kerusakan yang lebih ringan karena getaran melemah sebelum mencapai permukaan. Berdasarkan kekuatannya, gempa ringan (<4,9 Mw) umumnya tidak merusak, gempa sedang (4,9 – 6,0 Mw) menyebabkan kerusakan yang tidak parah, sedangkan gempa besar (>6,0 Mw) menyebabkan kerusakan parah.

3. HASIL DAN PEMBAHASAN

3.1 Deskripsi Data

Data pada penelitian ini bersumber dari data sekunder yang diperoleh dari situs USGS tahun 2023, dengan fokus pada data gempa bumi yang terjadi di Indonesia selama periode 1 Januari 2013 hingga 31 Desember 2022. Data terdiri dari 15.116 data dengan empat variabel, meliputi magnitudo dan kedalaman untuk analisis data, serta lintang dan bujur untuk tujuan pemetaan hasil analisis.

Tabel 2. Sampel Data Gempa Bumi Indonesia

Lokasi	Lintang	Bujur	Kedalaman	Magnitudo
185 km WNW of Pariaman, Indonesia	-0,2300	98,499	35	4,4
Pulau Pulau Tanimbar, Indonesia	-7,1100	130,368	17,8	4,4
Pulau-Pulau Sula, Maluku Utara, Indonesia	-1,0850	126,956	20,6	4,9
96 km WSW of Kotabumi, Indonesia	-5,3020	104,149	80,7	4,5
25 km SW of Sinabang, Indonesia	2,3330	96,207	41	4,7

3.2 Preprocessing Data

a. Data Cleaning

Tahap awal dalam *preprocessing* data melibatkan *data cleaning* untuk memeriksa apakah terdapat nilai yang hilang (*missing value*) dalam dataset. Hasil analisis menunjukkan bahwa terdapat 85 nilai yang hilang pada variabel kedalaman dan 61 nilai yang hilang pada variabel magnitudo. Untuk mengatasi nilai yang hilang diterapkan imputasi dengan menggunakan nilai median dari data sebelum dan sesudahnya. Pendekatan ini dipilih karena nilai yang hilang memiliki karakteristik yang mirip dengan data sebelum dan sesudahnya.

b. Konversi Nilai Magnitudo dengan Tipe Mw

Variabel magnitudo memiliki beberapa jenis yang perlu dikonversi menjadi tipe Mw. Terdapat tiga jenis magnitudo dalam data: mb (14.302 data), Mw (804 data), dan MLv (2 data). Konversi dilakukan sesuai dengan Tabel 1.

c. Transformasi Data

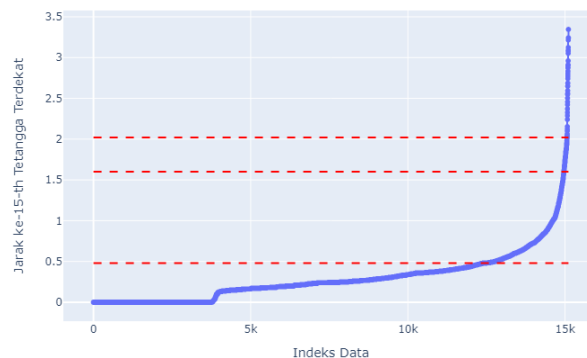
Data telah dibersihkan dan disesuaikan dengan satuan yang digunakan. Transformasi data hanya dilakukan pada variabel kedalaman, karena variabel ini memiliki hubungan berbanding terbalik dengan dampak yang dihasilkan dari gempa. Transformasi dilakukan untuk menjadikan interpretasi dampak yang dihasilkan menjadi berbanding lurus. Tabel 3 merupakan hasil transformasi data.

Tabel 3. Data Hasil Transformasi

Lokasi	Lintang	Bujur	Kedalaman (km)	Magnitudo (Mw)
185 km WNW of Pariaman	-0,2300	98,499	264,580	4,0659
Pulau Pulau Tanimbar, Indonesia	-7,1100	130,368	281,780	4,0659
Pulau-Pulau Sula, Maluku Utara, Indonesia	-1,0850	126,956	278,980	4,6649
96 km WSW of Kotabumi	-5,3020	104,149	218,880	4,1857

3.3 Clustering dengan Algoritma DBSCAN

Nilai *Eps* dan *MinPts* dapat ditentukan melalui penerapan K-NN (*K-Nearest Neighbors*). Penelitian ini dilakukan melalui plot K-NN dengan $l = 15$ untuk data gempa bumi di Indonesia tahun 2013-2022 yang divisualisasikan dalam Gambar 1.



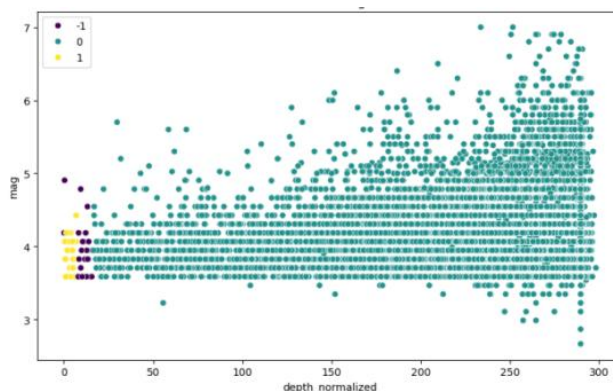
Gambar 1. Plot K-NN untuk $l = 15$

Gambar 1 mengindikasikan adanya sebuah patahan pada titik 0,48; 1,6; dan 2,02. Selanjutnya, dilakukan pengujian untuk menentukan nilai *Eps* optimal dengan memperhatikan nilai koefisien *silhouette*-nya dengan melibatkan variasi *MinPts* dari 11 hingga 15 serta *Eps* sebesar 0,48; 1,6; dan 2,02.

Tabel 4. Perbandingan Nilai *MinPts* dan *Eps*

<i>Eps</i>	<i>MinPts</i>	Jumlah Cluster	Jumlah Noise	Koefisien <i>Silhouette</i>
0,48	11	27	1010	0,454
	12	23	1159	0,441
	13	20	1272	0,493
	14	17	1417	0,472
	15	15	1568	0,447
1,6	11	2	21	0,676
	12	2	23	0,677
	13	5	29	0,646
	14	7	30	0,646
	15	5	50	0,645
2,02	11	2	1	0,670
	12	2	9	0,673
	13	2	16	0,675
	14	2	18	0,675
	15	3	22	0,676

Tabel 4 menunjukkan bahwa kombinasi nilai *Eps* 1,6 (jarak maksimal antara dua data dalam satu *cluster* adalah 1,6) dengan *MinPts* 12 (jumlah minimum data dalam satu *cluster* yaitu 12) memberikan nilai koefisien *silhouette* terbesar yaitu sebesar 0,677 dengan jumlah *noise* sebanyak 23 data. Setelah memperoleh nilai *Eps* dan *MinPts* optimal, dilakukan visualisasi untuk tiap titik *cluster* pada Gambar 2.

Gambar 2. Hasil *Clustering* Algoritma DBSCAN

Gambar 2 menunjukkan pembentukan dua *cluster* dan adanya data *noise*. *Cluster* 1 ditandai dengan warna kuning dan terdiri dari 28 data dengan karakteristik gempa dangkal berkekuatan sedang, sementara *cluster* 2 ditandai dengan warna hijau dengan total 15.057 data dengan karakteristik gempa yang bervariasi dari sisi kedalaman dan kekuatan. Data *noise* ditampilkan dalam warna ungu dan terdiri dari 23 data. Berdasarkan Gambar 2, dapat diketahui bahwa *cluster* terbagi berdasarkan kedalaman.

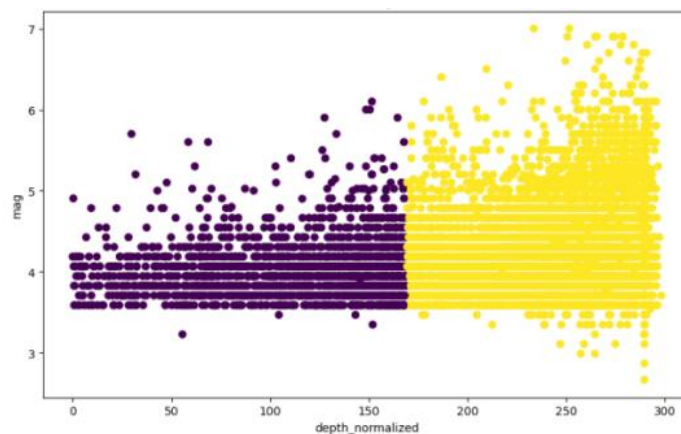
3.4 Clustering dengan Algoritma SOM

Dalam analisis *Clustering* menggunakan algoritma SOM, nilai optimal jumlah *cluster* (k) ditentukan menggunakan algoritma *K-Means* dengan koefisien *silhouette* terbaik sebesar 0,717 pada $k = 2$. Oleh karena itu, untuk tahap *Clustering* dengan algoritma SOM selanjutnya, digunakan nilai $k = 2$. Pada penelitian ini akan dianalisis dengan banyaknya *epoch* sebanyak 1.000 dengan beberapa *learning rate* yang akan diujikan pada data yang selanjutnya akan diperoleh nilai koefisien *silhouette*.

Tabel 5. Kombinasi *Learning rate* beserta Koefisien *Silhouette*

<i>Learning rate</i>	Koefisien <i>Silhouette</i>
0,05	0,717
0,1	0,716
0,7	0,714
1,0	0.714

Tabel 5 menunjukkan bahwa *learning rate* 0,05 memberikan nilai koefisien *silhouette* tertinggi yaitu sebesar 0,717. Berikut hasil visualisasi untuk *clustering* dengan nilai *learning rate* optimal pada *epoch* terakhir.

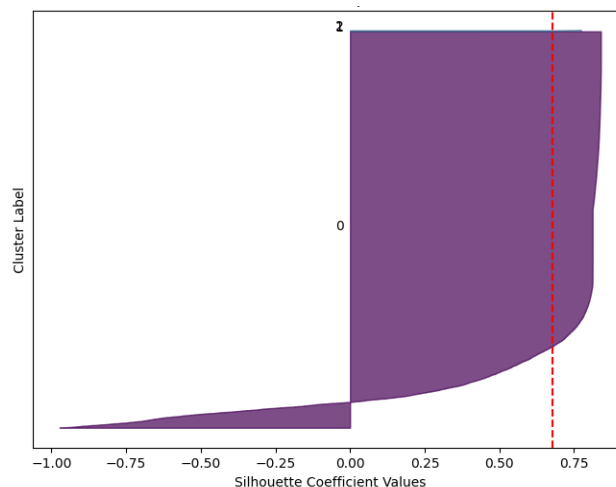


Gambar 3. Hasil *Clustering* Algoritma SOM

Gambar 3 merupakan visualisasi hasil *Clustering* algoritma SOM, terdapat dua *cluster* yang terbentuk. *Cluster* 1 berwarna ungu dengan jumlah 2.079 data, sementara *cluster* 2 berwarna kuning dengan jumlah 13.029 data. Gambar 3 menunjukkan bahwa nilai magnitudo untuk kedua *cluster* ini mirip dan terdapat pemisahan yang jelas dalam hal kedalaman.

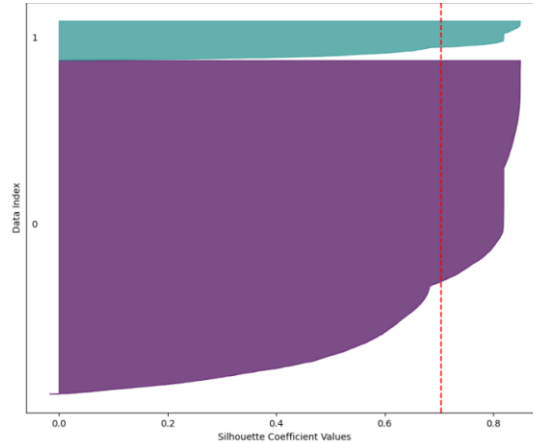
3.5 Perbandingan Nilai Koefisien *Silhouette* Algoritma DBSCAN dan SOM

Setelah melakukan analisis dengan algoritma DBSCAN dan SOM pada data gempa bumi di wilayah Indonesia tahun 2013-2022, dilakukan perbandingan nilai koefisien *silhouette* antara dua algoritma tersebut untuk menentukan algoritma yang lebih baik dalam melakukan *clustering*. Gambar 4 dan Gambar 5 merupakan visualisasi rata-rata nilai koefisien *silhouette* untuk setiap *cluster* yang terbentuk dari kedua algoritma.



Gambar 4. Plot Koefisien *Silhouette* Algoritma DBSCAN

Gambar 4 menunjukkan plot koefisien *silhouette* dari algoritma DBSCAN. Rata-rata koefisien *silhouette* yang dihasilkan dari plot tersebut adalah 0,677 yang menunjukkan adanya ikatan yang cukup kuat antar-objek dalam satu *cluster*. Namun, terdapat beberapa koefisien *silhouette* negatif di dalam *cluster*, menandakan adanya data yang mungkin masuk ke dalam *cluster* yang tidak sesuai.



Gambar 5. Plot Koefisien *Silhouette* Algoritma SOM

Gambar 5 menunjukkan plot dari rata-rata nilai koefisien *silhouette* dari algoritma SOM. Rata-rata koefisien *silhouette* yang dihasilkan dari plot tersebut adalah 0,717 yang artinya terdapat ikatan yang sangat kuat antar-objek dalam satu *cluster*.

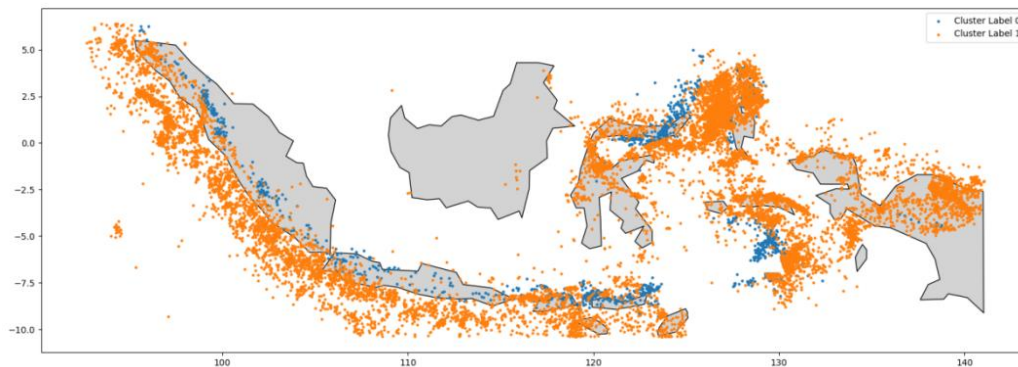
3.6 Clustering Gempa Bumi di Indonesia

Berdasar hasil perbandingan nilai koefisien *silhouette* dari algoritma DBSCAN dan SOM terlihat bahwa algoritma SOM menghasilkan nilai koefisien *silhouette* yang lebih tinggi yaitu sebesar 0,717. Oleh karena itu, dilakukan deskripsi data gempa bumi dari hasil *clustering* pada Tabel 6.

Tabel 6. Statistik Deskriptif dari Hasil *Clustering*

Cluster	Variabel	Statistik Deskriptif			Anggota Cluster
		Minimal	Rata-Rata	Maksimal	
1	Kedalaman	130,83	184,004	299,24	2.079
	Magnitudo	3,227	4,025	6,1	
2	Kedalaman	1,410	43,863	130,65	13.029
	Magnitudo	2,667	4,192	7,0	

Cluster 1 menunjukkan bahwa gempa bumi dalam kisaran kedalaman 130,83 hingga 299,24 km, dengan rata-rata kedalaman 184,004 km. Rentang magnitudo gempa dalam *cluster* 1 berkisar dari 3,227 hingga 6,1 Mw, dengan rata-rata magnitudo sebesar 4,025 Mw. Secara umum, *cluster* 1 dapat dikategorikan sebagai gempa sedang berkekuatan sedang. Sementara itu, *cluster* 2 mencakup gempa bumi dalam rentang kedalaman 1,41 hingga 130,65 km, dengan rata-rata kedalaman sekitar 43,863 km. Rentang magnitudo gempa di *cluster* 2 berkisar dari 2,667 hingga 7,0 Mw, dengan rata-rata magnitudo sebesar 4,192 Mw. *Cluster* 2 dapat diinterpretasikan sebagai gempa dangkal berkekuatan sedang. Secara umum, *cluster* 1 dapat dikategorikan sebagai gempa kedalaman sedang berkekuatan sedang dan *cluster* 2 dapat dikategorikan sebagai gempa berkedalaman dangkal berkekuatan sedang.



Gambar 6. Hasil Pemetaan *Clustering* dengan Algoritma SOM

Berdasarkan Gambar 6, data gempa bumi di Indonesia berhasil dikelompokkan menjadi 2 *cluster*. *Cluster 1* menunjukkan potensi gempa di daratan Pulau Sumatra, Aceh, Gunung Masurai, Lampung, Pulau Jawa, NTB, NTT, Bali, selatan Laut Fores, Laut Tengah, timur Banda Laut, barat Maluku Utara, dan Sulawesi Utara. Gempa dalam *cluster* ini cenderung tidak terasa karena kedalaman yang dalam. *Cluster 2* menunjukkan potensi gempa di sepanjang pesisir barat Pulau Sumatra, pesisir selatan Pulau Jawa, pesisir utara Nusa Tenggara dan Bali, Laut Sewu, Pulau Timor, selatan Laut Fores, timur Banda Laut, daratan Papua, Maluku Utara, utara Laut Molucca, Sulawesi Barat, dan Kalimantan Selatan dan Utara. Gempa pada *cluster 2*, terutama di daerah laut dan pesisir, memiliki potensi tsunami yang berdampak merugikan. Oleh karena itu, edukasi masyarakat sekitar perlu dilakukan sebagai upaya mitigasi bencana gempa dan tsunami.

4. SIMPULAN DAN SARAN

Algoritma DBSCAN menggunakan nilai *Eps* dan *MinPts* optimal 1,6 dan 12 membentuk 2 *cluster* dengan 23 data sebagai *noise* dan menghasilkan koefisien *silhouette* sebesar 0,677. Sementara itu, algoritma SOM dengan *output layer 2*, *learning rate* 0,05, dan maksimal *epoch* 1.000 membentuk 2 *cluster* dengan koefisien *silhouette* 0,717. Algoritma SOM lebih baik dibandingkan DBSCAN karena menghasilkan koefisien *silhouette* yang lebih besar. *Cluster 1* dari SOM dikategorikan sebagai gempa sedang berkekuatan sedang, sementara *cluster 2* sebagai gempa dangkal berkekuatan sedang. *Cluster 2* memiliki potensi gempa dengan dampak kerugian lebih besar dibandingkan *cluster 1* sehingga mitigasi risiko dapat difokuskan di area yang masuk dalam *cluster 2*. Terdapat saran untuk penelitian selanjutnya yaitu dapat memperdalam eksplorasi terhadap parameter yang digunakan, mencoba algoritma *clustering* lain, atau menambahkan metode validasi *cluster* yang berbeda untuk meningkatkan akurasi hasil *clustering*.

DAFTAR PUSTAKA

- Bação, F., Lobo, V., & Painho, M. (2005). *Self-Organizing Maps as Substitutes for K-Means Clustering*. Computational Science-ICCS 2005: 5th International Conference, Atlanta, GA, USA, May 22-25, 2005, Proceedings, Part III 5, 476-483.
- Badan Meteorologi, Klimatologi, dan Geofisika. (2021). Katalog Gempabumi Indonesia: Relokasi Hiposenter dan Implikasi Tektonik. Badan Meteorologi Klimatologi dan Geofisika: Jakarta.
- Damayanti, C., Yamko, A. K., Souisa, C. J., Barends, W., & Naroly, I. L. P. T. (2020). Pemodelan Segmentasi Mentawai-Pagai: Studi Kasus Gempa Megathrust di Indonesia. *Jurnal Geosains dan Remote Sensing*, 1(2), 105-110.
- Halim, N. N., & Widodo, E. (2017, July). Clustering Dampak Gempa Bumi di Indonesia Menggunakan Kohonen Self Organizing Maps (SOM). In *Prosiding SI MaNIs (Seminar Nasional Integrasi Matematika dan Nilai-Nilai Islami)* (Vol. 1, No. 1, pp. 188-194).
- Han, J., Pei, J., & Tong, H. (2022). *Data Mining: Concepts and Techniques*. Morgan Kaufmann.
- Hartatik, H., dan Cahya, A. S. D. (2020). Clusterisasi Kerusakan Gempa Bumi di Pulau Jawa Menggunakan Self Organizing Maps (SOM). *Jurnal Ilmiah Intech: Information Technology Journal of UMUS*, 2(02).
- Haykin, S. (1999). *Neural Networks: A Comprehensive Foundation*. Singapore: Pearson Education.

- Hair, J.F., Babin, B.J., Black, W.C., & Anderson, R.E. (2019). *Multivariate Data Analysis*. United Kingdom : Cengage.
- Nisrina, S., Nurmawati, W. P., & Gazali, M. (2022). Penerapan Metode Clustering SOM dan DBSCAN dalam Mengelompokkan Unmet Need Keluarga Berencana di Nusa Tenggara Barat. *J Statistika: Jurnal Ilmiah Teori dan Aplikasi Statistika*, 15(2), 237-244.
- Pratiwi, H., Handajani, S. S., & Respatiwan, (2021). Clustering dan Penerapannya dalam Bidang Seismologi. Sleman: Deepublish Publisher.
- Putriana, S., Ernawati, E., & Andreswari, D. (2021). Clustering Data Titik Gempa dengan Metode Fuzzy Possibilistic C-Means (Studi Kasus: Titik Gempa Pulau Sumatera Tahun 2013-2018). *Rekursif: Jurnal Informatika*, 9(1).
- Rahman, R. R. A. dan Wijayanto, A. W. (2021). Pengelompokan Data Gempa Bumi Menggunakan Algoritma Density-Based Spatial Clustering Algorithm with Noise (DBSCAN). *Jurnal Meteorologi dan Geofisika*, 22(1), 31-38.
- Risman, R., Syaripuddin, S., & Suyitno, S. (2019, May). Implementasi Metode DBSCAN pada Pengelompokan Kabupaten/Kota di Pulau Kalimantan Berdasarkan Indikator Kesejahteraan Rakyat. In *Prosiding Seminar Nasional Matematika dan Statistika* (Vol. 1, pp. 22-28).
- Sharda, R., Delen, D., Turban, E., & King, D. (2018). *Business Intelligence, Analytics, and Data Science: A Managerial Perspective (4th Edition)*. Pearson.
- Utomo, D. P., & Purba, B. (2019, September). Penerapan Data Mining pada Data Gempa Bumi terhadap Potensi Tsunami di Indonesia. In *Prosiding Seminar Nasional Riset Information Science (SENARIS)* (Vol. 1, pp. 846-853).