



Analisis Pola Aktivitas Gempa Bumi di Pulau Sumatera dengan Metode *K-Means Clustering* dan Rantai Markov

Putri Anissa Rahmi¹, Triyana Muliawati², Danni Gathot Harbowo³

^{1,2}Program Studi Matematika, Institut Teknologi Sumatera

³Program Studi Teknik Geologi, Institut Teknologi Sumatera

¹putri.120160024@student.itera.ac.id

³danni.gathot@gl.itera.ac.id

Corresponding author email: triyana.muliawati@ma.itera.ac.id

Abstract: This study aims to identify and understand the pattern of earthquake activity on the island of Sumatra. Using data on magnitude (*mag*), depth, and time of earthquake occurrence (*t*-day of the year), this study provides in-depth insight into the distribution and characteristics of earthquakes in the region. Data containing outliers is detected using boxplot visualization, then the data is further processed with the *K-Means Clustering* method to group the data into different clusters. Each cluster formed is represented as a state space in a Markov chain model, enabling the prediction of future earthquake patterns based on the transition between states. From the results of the study, the probability of earthquake activity patterns with the highest state transition for magnitude (*mag*) is 0.5339 (from medium to medium magnitude), for depth is 0.7165 (from medium to medium depth), and for earthquake occurrence time is 0.9942 (from the beginning of the year to the beginning of the year). This approach is expected to contribute significantly to disaster risk mitigation and earthquake management planning on the island of Sumatra.

Keywords: *K-Means Clustering*, Markov chain, probability, earthquake.

Abstrak: Penelitian ini bertujuan untuk mengidentifikasi dan memahami pola aktivitas gempa bumi di Pulau Sumatera. Dengan menggunakan data magnitudo (*mag*), kedalaman (*depth*), dan waktu kejadian gempa (hari ke-*t* dalam satu tahun), penelitian ini memberikan wawasan mendalam tentang distribusi dan karakteristik gempa di wilayah tersebut. Data yang mengandung pencilan (*outlier*) dideteksi menggunakan visualisasi *boxplot*, kemudian data diolah lebih lanjut dengan metode *K-Means Clustering* untuk mengelompokkan data ke dalam *cluster-cluster* yang berbeda. Setiap *cluster* yang terbentuk direpresentasikan sebagai ruang keadaan (*state*) dalam model *Markov Chain*, memungkinkan prediksi pola gempa di masa depan berdasarkan transisi antar *state* tersebut. Dari hasil penelitian, diperoleh peluang pola aktivitas gempa bumi dengan peralihan keadaan tertinggi untuk magnitudo (*mag*) yaitu sebesar 0.5339 (dari magnitudo sedang ke sedang), untuk kedalaman (*depth*) yaitu sebesar 0.7165 (dari kedalaman menengah ke menengah), dan untuk waktu terjadinya gempa yaitu sebesar 0.9942 (dari awal tahun ke awal tahun). Pendekatan ini diharapkan memberikan kontribusi signifikan dalam mitigasi risiko bencana dan perencanaan penanggulangan gempa bumi di Pulau Sumatera.

Kata kunci: *K-Means Clustering*, Markov Chain, peluang, gempa bumi.

I. PENDAHULUAN

Indonesia dikelilingi oleh beberapa lempeng tektonik aktif yang saling berbenturan, menyebabkan tingginya risiko gempa bumi [1]. Secara geografis, Indonesia berada di wilayah “*Ring of Fire*” tempat pertemuan lempeng Eurasia, Indo-Australia, dan Samudra Pasifik [2, 3]. Pulau Sumatera, bagian dari Indonesia yang berada di Cincin Api Pasifik, memiliki wilayah tektonik aktif sehingga rentan terhadap gempa bumi [4]. Kerentanan ini disebabkan oleh dua zona utama: zona subduksi Sunda dan Sesar Sumatera (*Sumateran Fault Zone/SFZ*), yang dapat memicu gempa dan tsunami [5].

Menurut Kementerian ESDM, dari 25 wilayah rawan gempa di Indonesia, 6 berada di Pulau Sumatera: Sumatera Utara, Aceh, Jambi, Bengkulu, Sumatera Barat, dan Lampung. Data BMKG mencatat 5937 gempa di Pulau Sumatera selama 2009-2018, dengan 36% di antaranya adalah gempa besar [6]. Gempa besar tidak hanya dapat menyebabkan tanah longsor, likuifaksi, tsunami, longsoran bawah laut [7, 8], dan gempa susulan lainnya, tetapi juga dapat menyebabkan kerugian material dan korban jiwa. Dampak-dampak ini membuat masyarakat lebih menyadari pentingnya perencanaan darurat dan infrastruktur yang tahan gempa di Pulau Sumatera.

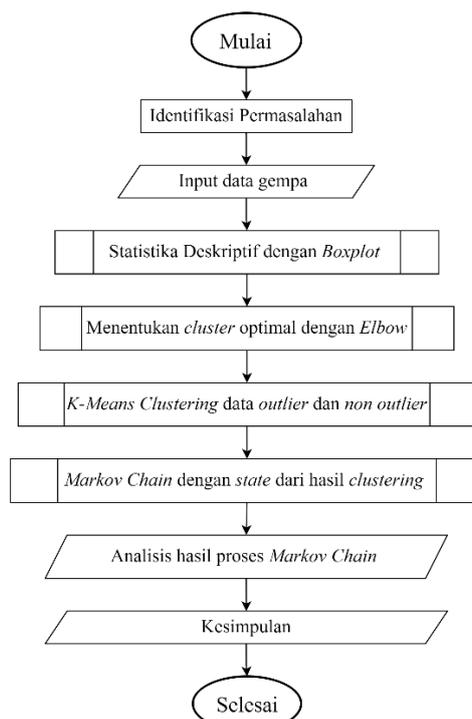


Berdasarkan hal tersebut, gempa bumi termasuk dalam bencana alam yang sulit untuk dihindari dan terjadi secara tiba-tiba, sehingga perlu diadakannya penelitian tentang peluang pola aktivitas gempa bumi di Pulau Sumatera untuk memahami bagaimana probabilitas dari pola aktivitas gempa bumi pada Pulau Sumatera selanjutnya sehingga dapat merencanakan mitigasi bencana, meningkatkan kewaspadaan, serta mengembangkan strategi adaptasi untuk mengurangi risiko dan dampak negatif gempa bumi. Aktivitas gempa bumi dapat dilihat berdasarkan magnitudo, kedalaman (*depth*) dan waktu (hari ke-*t* dalam satu tahun).

Penelitian sebelumnya menggunakan *Markov Chain* untuk memprediksi probabilitas transisi amplitudo maksimum dan durasi aktivitas seismik dalam data gempa Krakatau. Ruang keadaan (*state*) diperoleh dengan mengelompokkan variabel-variabel menggunakan *K-Means Clustering* [9]. Penelitian ini mengadaptasi pendekatan serupa dengan mengelompokkan data gempa Pulau Sumatera berdasarkan kedalaman, magnitudo, dan waktu kejadian menggunakan *K-Means Clustering*. Setelah menentukan jumlah *cluster* optimal dengan metode *Elbow*, data gempa dianalisis untuk memahami dan memprediksi pola aktivitas gempa di masa depan menggunakan model *Markov Chain*. Tujuannya adalah memberikan wawasan untuk mitigasi risiko dan perencanaan penanggulangan bencana gempa bumi di wilayah tersebut.

II. METODE PENELITIAN

Data sekunder yang digunakan dalam penelitian ini diperoleh dari *website* resmi *United States Geological Survey* (USGS). Data dikumpulkan dari gempa bumi Pulau Sumatera, yang memiliki luas wilayah -8.051, 6.891 dan 91.23, 108.193. Data tersebut berupa data magnitudo (X_1), kedalaman (X_2) dan hari ke-*t* dalam satu tahun (X_3) dalam periode waktu 01 Januari 1990 sampai 31 Juli 2023 dengan total masing-masing data sebanyak 14260. Metode yang digunakan yaitu statistika deskriptif dengan *boxplot*, *K-Means Clustering* dan *Markov Chain* dengan bantuan platform *google colab*.



Gambar 1. Diagram Alir Penelitian



2.1 Statistika Deskriptif

Statistik deskriptif adalah teknik untuk menguraikan dan memberikan informasi tentang data, keadaan, atau fenomena tertentu. Ini melibatkan penyajian data dalam bentuk diagram, tabel, atau grafik, serta perhitungan statistik seperti median, modus, rata-rata (*mean*), nilai minimum, dan nilai maksimum [10]. Dalam penelitian ini, *boxplot* digunakan untuk menampilkan distribusi data secara visual. Kuartil pertama (Q1) dan kuartil ketiga (Q3) digunakan untuk mengidentifikasi titik tengah dari setengah bawah dan setengah atas data setelah diurutkan. Rentang antarkuartil (IQR) dihitung sebagai selisih antara Q3 dan Q1, yang memberikan gambaran tentang sebaran data dalam *boxplot*. Selanjutnya, batas bawah dan atas dari *boxplot* digunakan untuk mengidentifikasi potensial *outlier* yang perlu diperiksa lebih lanjut [11].

2.2 Metode Elbow

Metode *Elbow* adalah salah satu teknik yang digunakan dalam analisis *clustering* untuk menentukan jumlah *cluster* optimal dalam sebuah dataset. Pendekatan ini melibatkan penghitungan SSE (*Sum of Square Error*), yang merupakan metrik evaluasi untuk mengukur seberapa baik setiap titik data dalam sebuah *cluster* cocok dengan pusat *cluster* nya, dengan menjumlahkan kuadrat jarak antara setiap titik data dan pusat *cluster* nya. Adapun persamaan SSE untuk *K-Means Clustering* dapat dilihat pada persamaan (1) dengan k adalah jumlah *cluster*, x_i adalah nilai atribut data ke- i dalam *cluster* ke- j , C_j adalah *centroid cluster* ke- j , dan S_k adalah himpunan data dalam *cluster* ke- j [12]:

$$SSE = \sum_{k=1}^n \sum_{x_i \in S_k} \|x_i - C_j\|^2 \quad (1)$$

Nilai SSE ini kemudian dipetakan terhadap jumlah *cluster* yang dieksplorasi. Kurva yang dihasilkan menunjukkan bagaimana SSE akan menurun seiring dengan penambahan *cluster* (k). Pada analisis, titik “*Elbow*” adalah titik di kurva di mana penurunan SSE menunjukkan perubahan yang paling signifikan, sering kali terbentuk sudut yang tajam (*Elbow*). Titik “*Elbow*” ini menandai jumlah *cluster* optimal di mana penambahan *cluster* tambahan tidak lagi memberikan penurunan SSE yang signifikan, sehingga memberikan rekomendasi jumlah *cluster* terbaik untuk dataset yang dianalisis. Metode *Elbow* sangat penting karena membantu menghindari *overfitting* atau *underfitting* model *clustering*, dan memastikan pengelompokan data yang lebih efektif.

2.3 K-Means Clustering

K-Means Clustering adalah metode yang menggabungkan data ke dalam *cluster* yang berbeda sehingga data yang dikumpulkan dalam satu *cluster* memiliki tingkat kesamaan yang lebih tinggi daripada data yang dikumpulkan dalam *cluster* lain [13]. Adapun tahapan-tahapan pada algoritma *K-Means Clustering* sebagai berikut [14]:

1. Menentukan jumlah *cluster*;
2. Menentukan centroid awal secara acak;
3. Menghitung jarak ke setiap *centroid* menggunakan *Euclidean Distance*;

$$D(x, C) = \sqrt{\left(\sum_{i=1}^n x_i \right) - C_j} \text{ dengan } j = 0,1,2,3, \dots, k \quad (2)$$

dengan $D(x, C)$ adalah *Euclidean Distance* dari data ke- i (x_i) ke *centroid cluster* ke- j (C_j).

4. Mengelompokkan data berdasarkan jarak terdekat ke *centroid*;



5. Memperbarui nilai *centroid* dengan rata-rata data dalam *cluster*;
6. Melakukan iterasi hingga tidak ada perpindahan data antar *cluster*.

2.4 Analisis Rantai Markov (*Markov Chain*)

Dalam proses Markov, keadaan masa lalu tidak mempengaruhi keadaan masa depan. Dengan kata lain, jika keadaan saat sekarang diketahui, peluang keadaan proses di satu langkah ke depan hanya dipengaruhi oleh keadaan saat sekarang [15]. Rantai Markov adalah proses stokastik yang memiliki sifat markovian, jika peluang perpindahan dari *state i* ke *state j* (P_{ij}) ditulis sebagai berikut:

$$\{X_{t+1} = j | X_0 = i_0, X_1 = i_1, \dots, X_t = i\} = P\{X_{t+1} = j | X_t = i\} = P_{ij} \quad (3)$$

Sifat Markovian menyatakan bahwa peluang bersyarat dari kejadian masa depan (X_{t+1}) adalah independen dari kejadian di masa lalu dan hanya bergantung pada keadaan saat ini (X_t). Ruang keadaan (*state*), diagram transisi, dan matriks peluang transisi adalah konsep penting dalam penyelesaian rantai Markov [16]. Matriks peluang transisi mengatur pergerakan sistem antar *state*. Sebagai contoh, peluang perpindahan dari *state i* ke *j* ditunjukkan oleh (P_{ij}). Matriks peluang transisi P dengan dimensi $(i + 1) \times (j + 1)$ dijelaskan sebagai berikut.

$$P = P_{ij} = \begin{matrix} \text{State} & & 0 & 1 & 2 & \dots \\ & 0 & P_{00} & P_{01} & \dots & P_{0j} \\ & 1 & P_{10} & P_{11} & \dots & P_{1j} \\ & 2 & \dots & \dots & \dots & \dots \\ & \vdots & P_{i0} & P_{i1} & \dots & P_{ij} \end{matrix} \quad (4)$$

Semua elemen dalam matriks P harus positif ($P_{ij} \geq 0$), dan total probabilitas dalam setiap baris harus sama dengan 1 ($\sum_{j=1}^m P_{ij} = 1$). Artinya, probabilitas semua perpindahan dari satu keadaan ke keadaan lain dalam satu langkah harus menyusun total 1 [17]. Dengan demikian, matriks transisi adalah representasi matematis dari peluang transisi antar keadaan dalam sistem, dengan elemen-elemen matriks yang menggambarkan probabilitas perpindahan dari satu keadaan ke keadaan lain.

III. HASIL DAN PEMBAHASAN

3.1 *Boxplot* Data Gempa Bumi

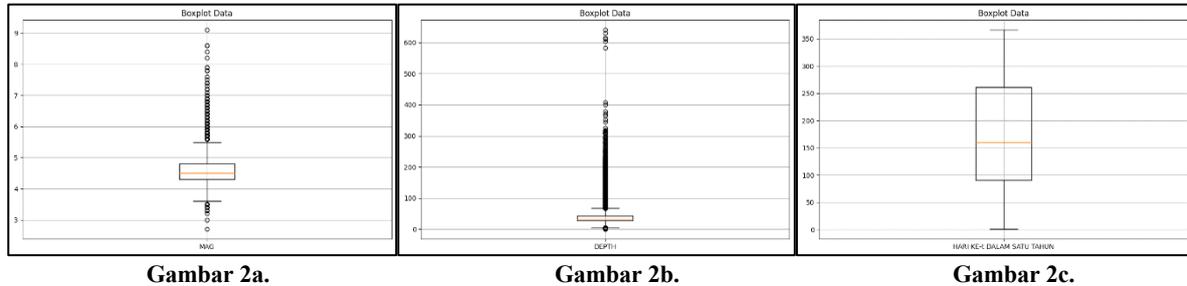
Untuk menganalisis pola peluang aktivitas gempa bumi di Pulau Sumatera, 3 variabel data akan diuji menggunakan *boxplot*: magnitudo, kedalaman (*depth*), dan waktu terjadinya gempa (hari ke- t dalam satu tahun). Hasil *boxplot* akan digunakan untuk analisis pengelompokan (*clustering*), dengan memisahkan data *outlier*. Hasil *clustering* dari data *outlier* akan digabungkan dengan hasil *clustering* data *non-outlier*. Hasil *clustering* ini akan digunakan dalam pemodelan *Markov Chain*. Berikut adalah tabel statistik deskriptif dari ketiga variabel yang digunakan:

Tabel 1. Statistika deskriptif data gempa bumi Pulau Sumatera

Variabel	Satuan	Min	Q1	Median	Q3	Max	Outlier
Magnitudo (mag)	mb	3.6	4.3	4.5	4.8	5.5	494
Kedalaman (<i>depth</i>)	km	5.55	28.5	33	44	67.22	1928
Hari ke- t dalam satu tahun	hari ke-	1	91	160	261	366	0

Berdasarkan Tabel 1 diatas, maka variabel gempa bumi yang mempunyai data pencilan (*outlier*) adalah magnitudo dengan 494 data *outlier*, yang berarti terdapat 13.766 data *non-outlier*. Kemudian, variabel kedalaman dengan 1928 data *outlier* dan 12.332 data *non-outlier*. Sedangkan variabel hari ke-

t dalam satu tahun tidak memiliki data pencilan (*outlier*). Berikut merupakan *boxplot* dari masing-masing variabel data gempa bumi.



Gambar 2. *Boxplot* data gempa: (2a) *boxplot* magnitudo, (2b) *boxplot* kedalaman, (2c) *boxplot* hari ke- t dalam satu tahun

3.2 Menentukan *Cluster* Optimal dengan Metode *Elbow*

Dalam penelitian ini, penentuan *cluster* optimal (k) dilakukan menggunakan metode *Elbow* dengan bantuan *Python* di platform *Google Colab*. Metode *Elbow* menampilkan grafik perbandingan nilai *SSE* untuk setiap *cluster*, yang membantu menentukan jumlah *cluster* optimal. *Cluster* optimal yang ditentukan adalah untuk masing-masing variabel gempa bumi yang terdiri dari data *outlier* dan data *non-outlier*. Berikut adalah tabel hasil *output* metode *Elbow* untuk masing-masing variabel:

Tabel 2. Hasil *Cluster* optimal dengan metode *Elbow* pada data gempa bumi Pulau Sumatera

Data	Variabel	Nilai SSE Sig	Sudut (Min)	Total Cluster (k)
<i>Outlier</i>	Magnitudo (mag)	128.6973990	147.09°	3
<i>Outlier</i>	Kedalaman (depth)	343.7244800	156.63°	3
<i>Non-Outlier</i>	Magnitudo (mag)	2410.001330	153.37°	3
<i>Non-Outlier</i>	Kedalaman (depth)	4115.539335	143.83°	3
<i>Non-Outlier</i>	Hari ke- t dalam satu tahun	1761.189435	153.02°	3

Dari Tabel 2 tersebut, hasil *cluster* optimal untuk setiap variabel gempa bumi yang memiliki *outlier* dan tanpa *outlier* mempunyai jumlah *cluster* (k) yang sama yaitu sebanyak 3 *cluster* optimal. Jumlah *cluster* tersebut akan digunakan sebagai acuan dalam membagi banyaknya *cluster* dalam proses berikutnya yaitu *K-Means Clustering*.

3.3 Pengelompokan Data Gempa dengan *K-Means Clustering*

Berdasarkan hasil pengujian menggunakan metode *Elbow* pada proses sebelumnya, secara keseluruhan telah diperoleh banyaknya *cluster* (k) optimal pada masing-masing variabel data gempa bumi. Selanjutnya dilakukan proses pengelompokan dengan *K-Means Clustering* yang dilakukan secara terpisah untuk data *outlier* dan *non-outlier*. Hasil pengelompokan tersebut akan digabung menjadi satu sehingga terbentuk hasil akhir *clustering* dari masing-masing data gempa bumi. Berikut adalah ringkasan jumlah hasil *cluster* pada masing-masing variabel.

Tabel 3. Hasil akhir *Clustering* pada masing-masing data gempa bumi Pulau Sumatera

Variabel	Jenis Rentang	Rentang Cluster	Cluster
Magnitudo (mag)	<i>Outlier</i>	$2.7 \leq X_1 \leq 3.5$	0
	<i>Non-Outlier</i>	$3.5 < X_1 \leq 4.2$	1
	<i>Non-Outlier</i>	$4.2 < X_1 \leq 4.7$	2
	<i>Non-Outlier</i>	$4.7 < X_1 \leq 5.5$	3
	<i>Outlier</i>	$5.5 < X_1 \leq 6.4$	4
	<i>Outlier</i>	$6.4 < X_1 \leq 9.1$	5
Kedalaman (depth)	<i>Outlier</i>	$0.3 \leq X_2 \leq 5.13$	0
	<i>Non-Outlier</i>	$5.13 < X_2 \leq 21.98$	1
	<i>Non-Outlier</i>	$21.98 < X_2 \leq 42.56$	2



Variabel	Jenis Rentang	Rentang Cluster	Cluster
Hari ke- <i>t</i> dalam satu tahun	Non-Outlier	$42.56 < X_2 \leq 67.22$	3
	Outlier	$67.22 < X_2 \leq 300$	4
	Outlier	$300 < X_2 \leq 640.21$	5
	Non-Outlier	$1 \leq X_3 \leq 125$	0
	Non-Outlier	$125 < X_3 \leq 242$	1
	Non-Outlier	$242 < X_3 \leq 366$	2

Berdasarkan Tabel 3 diatas, telah diperoleh hasil pengelompokan untuk masing-masing variabel data gempa bumi. Untuk variabel magnitudo dan kedalaman masing-masing memiliki total 6 *cluster*, sedangkan variabel hari ke-*t* dalam satu tahun memiliki total 3 *cluster*. Dari hasil pengelompokan tersebut, jumlah *cluster* yang terbentuk akan digunakan sebagai ruang keadaan (*state*) pada proses analisis *Markov Chain*. Analisis ini akan membantu dalam memahami dan memprediksi pola peluang aktivitas gempa bumi di Pulau Sumatera.

3.4 Analisis Peluang Pola Aktivitas Gempa dengan Metode *Markov Chain*

Penentuan *state* atau ruang keadaan dalam proses *Markov Chain* sangat penting, memungkinkan identifikasi langkah selanjutnya dan perhitungan nilai probabilitas. Dalam penelitian ini, *state* diperoleh dengan menggabungkan hasil *cluster K-Means Clustering* untuk data *non-outlier* dan *outlier*. Informasi mengenai *state* dapat ditemukan dalam Tabel 3 di mana kolom “*cluster*” menunjukkan *state* yang bersesuaian. Misalnya, data dalam *cluster* 3 dianggap sebagai *state* 3 dalam proses *Markov Chain*. Penentuan keterangan untuk variabel magnitudo dan kedalaman pada setiap *state* didasarkan pada klasifikasi yang telah ditetapkan berdasarkan literatur terkait [18]. Sedangkan untuk variabel hari ke-*t* dalam satu tahun di klasifikasikan berdasarkan hasil analisis dari penulis sendiri. Berikut merupakan tabel pengklasifikasian gempa bumi berdasarkan magnitudo dan kedalamannya.

Tabel 4. Klasifikasi gempa bumi berdasarkan kedalaman gempa

Kategori Kedalaman	Rentang Kedalaman (km)
Dangkal	< 70
Menengah	70 – 300
Dalam	> 300

Tabel 5. Klasifikasi gempa bumi berdasarkan kekuatan magnitudo gempa

Kategori Magnitudo	Rentang Kekuatan (mb)
Sangat Lemah	< 1.0 – 2.9
Lemah	3.0 – 3.9
Sedang	4.0 – 4.9
Kuat	5.0 – 5.9
Sangat Kuat	6.0 – 6.9
Besar	7.0 – 7.9
Sangat Besar	≥ 8.0

Berdasarkan pengklasifikasian data gempa bumi menurut referensi terkait yang ditampilkan pada Tabel 4 dan Tabel 5 sebelumnya, berikut ini disajikan Tabel 6 yang merupakan hasil klasifikasi data gempa bumi untuk setiap *state* pada masing-masing variabel penelitian.

Tabel 6. Keterangan masing-masing *state* pada data gempa bumi Pulau Sumatera

Variabel	Kategori	Rentang Cluster	State ke-
Magnitudo (mag)	LEMAH	$2.7 \leq X_1 \leq 3.5$	0
	CUKUP LEMAH	$3.5 < X_1 \leq 4.2$	1
	SEDANG	$4.2 < X_1 \leq 4.7$	2
	CUKUP KUAT	$4.7 < X_1 \leq 5.5$	3
	KUAT	$5.5 < X_1 \leq 6.4$	4



Variabel	Kategori	Rentang Cluster	State ke-
Kedalaman (<i>depth</i>)	SANGAT KUAT	$6.4 < X_1 \leq 9.1$	5
	DANGKAL	$0.3 \leq X_2 \leq 5.13$	0
	CUKUP DANGKAL	$5.13 < X_2 \leq 21.98$	1
	MENENGAH	$21.98 < X_2 \leq 42.56$	2
	CUKUP DALAM	$42.56 < X_2 \leq 67.22$	3
	DALAM	$67.22 < X_2 \leq 300$	4
Hari ke- <i>t</i> dalam satu tahun	SANGAT DALAM	$300 < X_2 \leq 640.21$	5
	AWAL TAHUN	$1 \leq X_3 \leq 125$	0
	PERTENGAHAN TAHUN	$125 < X_3 \leq 242$	1
	AKHIR TAHUN	$242 < X_3 \leq 366$	2

Kemudian, setelah menentukan *state* untuk masing-masing variabel maka langkah berikutnya adalah menentukan matriks peluang transisi. Untuk menentukan matriks peluang transisi, akan melibatkan identifikasi peralihan *state* pada setiap variabel data penelitian. Langkah ini memungkinkan pembentukan matriks frekuensi, yang kemudian digunakan untuk menghitung matriks peluang transisi dengan menghitung probabilitas dari setiap elemen dalam matriks frekuensi. Dengan demikian, dapat diperoleh analisis pola perubahan *state* dan memprediksi dinamika gempa bumi di Pulau Sumatera. Berikut merupakan matriks peluang transisi untuk setiap variabel penelitian.

Tabel 7. Matriks peluang transisi variabel magnitudo pada data gempa bumi Pulau Sumatera

Matriks Peluang Transisi Magnitudo (%)						
State	State 0	State 1	State 2	State 3	State 4	State 5
State 0	0	35.71	39.28	25.00	0	0
State 1	0.23	26.09	49.88	21.41	2.03	0.33
State 2	0.16	20.21	53.39	23.76	2.18	0.27
State 3	0.22	17.89	47.77	29.08	4.12	0.90
State 4	0.25	12.37	46.71	32.07	7.32	1.26
State 5	0	25.71	40.00	27.14	2.85	4.28

Nilai peluang data magnitudo tertinggi terjadi pada saat *state* 2 beralih ke *state* 2 atau pada kondisi magnitudo dengan kekuatan gempa “SEDANG” ke kekuatan gempa “SEDANG” yaitu sebesar 53.39%, seperti yang ditunjukkan dalam Tabel 7 matriks peluang transisi di atas. Dimana diperoleh hasil analisis untuk peluang pola aktivitas dari data magnitudo di Pulau Sumatera sebagai berikut.

- Jika magnitudo hari ini Lemah, besok berpeluang 35.71% “Cukup Lemah”, 39.28% “Sedang”, dan 25% “Cukup Kuat”.
- Jika magnitudo hari ini “Cukup Lemah”, besok berpeluang 0.23% “Lemah”, 26.09% “Cukup Lemah”, 49.88% “Sedang”, 21.41% “Cukup Kuat”, 2.03% “Kuat”, dan 0.33% “Sangat Kuat”.
- Jika magnitudo hari ini “Sedang”, besok berpeluang 0.16% “Lemah”, 20.21% “Cukup Lemah”, 53.39% “Sedang”, 23.76% “Cukup Kuat”, 2.18% “Kuat”, dan 0.27% “Sangat Kuat”.
- Jika magnitudo hari ini “Cukup Kuat”, besok berpeluang 0.22% “Lemah”, 17.89% “Cukup Lemah”, 47.77% “Sedang”, 29.08% “Cukup Kuat”, 4.12% “Kuat”, dan 0.90% “Sangat Kuat”.
- Jika magnitudo hari ini “Kuat”, besok berpeluang 0.25% “Lemah”, 12.37% “Cukup Lemah”, 46.71% “Sedang”, 32.07% “Cukup Kuat”, 7.32% “Kuat”, dan 1.26% “Sangat Kuat”.
- Jika magnitudo hari ini “Sangat Kuat”, besok berpeluang 25.71% “Cukup Lemah”, 40% “Sedang”, 27.14% “Cukup Kuat”, 2.85% “Kuat”, dan 4.28% “Sangat Kuat”.

Kemudian untuk variabel data kedalaman, nilai peluang data kedalaman tertinggi terjadi pada saat *state* 2 beralih ke *state* 2 atau pada kondisi kedalaman dengan kedalaman gempa “MENENGAH” ke kedalaman gempa “MENENGAH” yaitu sebesar 71.65%, seperti yang ditunjukkan dalam Tabel 8 matriks peluang transisi dibawah ini.



Tabel 8. Matriks peluang transisi variabel kedalaman pada data gempa bumi Pulau Sumatera

Matriks Peluang Transisi Kedalaman (%)						
<i>State</i>	<i>State 0</i>	<i>State 1</i>	<i>State 2</i>	<i>State 3</i>	<i>State 4</i>	<i>State 5</i>
<i>State 0</i>	0	33.33	44.44	11.11	11.11	0
<i>State 1</i>	0.43	45.85	29.42	12.24	11.63	0.39
<i>State 2</i>	0.21	8.13	71.65	9.99	9.88	0.10
<i>State 3</i>	0.22	16.23	45.03	19.81	18.52	0.16
<i>State 4</i>	0.21	14.17	44.49	17.33	23.45	0.32
<i>State 5</i>	0	13.79	37.93	17.24	24.13	6.89

Dimana diperoleh hasil analisis untuk peluang pola aktivitas dari data kedalaman pada gempa bumi di Pulau Sumatera sebagai berikut.

- Jika kedalaman gempa hari ini Dangkal, besok berpeluang 33.33% “Cukup Dangkal”, 44.44% “Menengah”, 11.11% “Cukup Dalam” dan 11.11% “Dalam”.
- Jika kedalaman gempa hari ini “Cukup Dangkal”, besok berpeluang 0.43% “Dangkal”, 45.85% “Cukup Dangkal”, 29.42% “Menengah”, 12.24% “Cukup Dalam”, 11.63% “Dalam” dan 0.39% “Sangat Dalam”.
- Jika kedalaman gempa hari ini “Menengah”, besok berpeluang 0.21% “Dangkal”, 8.13% “Cukup Dangkal”, 71.65% “Menengah”, 9.99% “Cukup Dalam”, 9.88% “Dalam” dan 0.10% “Sangat Dalam”.
- Jika kedalaman gempa hari ini “Cukup Dalam”, besok berpeluang 0.22% “Dangkal”, 16.23% “Cukup Dangkal”, 45.03% “Menengah”, 19.81% “Cukup Dalam”, 18.52% “Dalam” dan 0.16% “Sangat Dalam”.
- Jika kedalaman gempa hari ini “Dalam”, besok berpeluang 0.21% “Dangkal”, 14.17% “Cukup Dangkal”, 44.49% “Menengah”, 17.33% “Cukup Dalam”, 23.45% “Dalam” dan 0.32% “Sangat Dalam”.
- Jika kedalaman gempa hari ini “Sangat Dalam”, besok berpeluang 13.79% “Cukup Dangkal”, 37.93% “Menengah”, 17.24% “Cukup Dalam”, 24.13% “Dalam” dan 6.89% “Sangat Dalam”.

Selanjutnya untuk variabel data waktu terjadi nya gempa (hari ke- t dalam satu tahun), nilai peluang waktu terjadi nya gempa tertinggi terjadi pada saat *state 0* beralih ke *state 0* atau pada kondisi waktu terjadi nya gempa di “AWAL TAHUN” ke kedalaman gempa “AWAL TAHUN” yaitu sebesar 99.42%, seperti yang ditunjukkan dalam Tabel 9 matriks peluang transisi dibawah ini.

Tabel 9. Matriks peluang transisi variabel hari ke- t dalam satu tahun pada data gempa bumi Pulau Sumatera

Matriks Peluang Transisi Hari ke-t dalam satu tahun (%)			
<i>State</i>	<i>State 0</i>	<i>State 1</i>	<i>State 2</i>
<i>State 0</i>	99.42	0	0.57
<i>State 1</i>	0.80	99.19	0
<i>State 2</i>	0	0.76	99.23

Dimana diperoleh hasil analisis untuk peluang pola aktivitas dari data hari ke- t dalam satu tahun pada gempa bumi di Pulau Sumatera sebagai berikut.

- Jika hari ini waktu terjadi nya gempa di Awal Tahun, besok berpeluang 99.42% di Awal Tahun dan 0.57% di Akhir Tahun.
- Jika hari ini waktu terjadi nya gempa di Pertengahan Tahun, besok berpeluang 0.80% di Awal tahun dan 99.19% di Pertengahan Tahun.
- Jika hari ini waktu terjadi nya gempa di Akhir Tahun, besok berpeluang 0.76% di Pertengahan Tahun dan 99.23% di Akhir Tahun.



IV. KESIMPULAN

Berdasarkan hasil analisis yang telah dilakukan dapat diambil beberapa kesimpulan sebagai berikut.

1. Hasil pengelompokan *K-Means Clustering* untuk variabel data Magnitudo (*mag*) terdapat 6 *cluster* (3 *non-outlier*, 3 *outlier*), variabel data Kedalaman (*depth*) terdapat 6 *cluster* (3 *non-outlier*, 3 *outlier*) dan variabel data waktu terjadinya gempa bumi (hari ke-*t* dalam satu tahun) terdapat 3 *cluster*.
2. Peluang pola aktivitas gempa bumi dari hasil *Markov Chain* untuk data Magnitudo (*mag*) dengan peralihan kekuatan “SEDANG” ke “SEDANG” memiliki peluang transisi tertinggi yaitu sebesar 53.39%, untuk data Kedalaman (*depth*) dengan peralihan kedalaman “MENENGAH” ke “MENENGAH” memiliki peluang transisi tertinggi yaitu sebesar 71.65% dan untuk data hari ke-*t* dalam satu tahun dengan peralihan dari “AWAL TAHUN” ke “AWAL TAHUN” memiliki peluang transisi tertinggi yaitu sebesar 99.42%.

UCAPAN TERIMAKASIH

Dalam kesempatan ini, kami ingin menyampaikan ucapan terima kasih kepada panitia acara Seminar Nasional Sains Data 2024 (SENADA 2024) di UPN Veteran Jawa Timur. Terima kasih atas kesempatan yang diberikan untuk berbagi pengetahuan dan pengalaman. Kami juga mengucapkan terima kasih kepada semua pembicara, peserta, dan pihak terkait atas kontribusi dalam kesuksesan acara ini. Semoga hasil penelitian yang kami sajikan dalam jurnal ini dapat memberikan kontribusi positif dalam berbagai bidang ilmu pengetahuan dan teknologi, serta menginspirasi penelitian mendalam di masa depan. Terima kasih.

REFERENSI

1. Cummins, Geohazards in Indonesia: Earth science for disaster risk reduction introduction, Geological Society Special Publication, 2017.
2. Metrikasari and Choiruddin, "Pemodelan Risiko Gempa Bumi di Pulau Sumatera Menggunakan Model Inhomogeneous Neyman-Scott Cox Process.," *Jurnal Sains dan Seni ITS*, 2020.
3. T. Muliawati and D. G. Harbowo, "A Statistical review of the dates and patterns of volcanic activity of Lewotolo Volcano, East Nusa Tenggara, Indonesia," *IOP Conference Series: Earth and Environmental Science*, vol. 1245, no. 1, p. 012006, 2023.
4. Ricky, Ridho, Marizka, Jesi and Novisa, "Implementasi Business Intelligence Menentukan Daerah Rawan Gempa Bumi di Indonesia dengan Fitur Geolokasi," *JEPIN (Jurnal Edukasi dan Penelitian Informatika)*, 2018.
5. Asnita, Sugiyanto and Rusydy, "Kajian Statistik Seismistas Kawasan Sumatera," *Jurnal Natural*, vol. 16, no. 2, pp. 11-12, 2016.
6. Hafizhah and Pujiastuti, "Perbandingan Anomali Frekuensi Kritis Lapisan F2 (Fof2) Ionosfer Pada Gempa Bumi Laut Dan Gempa Bumi Darat Pulau Sumatera," *Jurnal Fisika Unand*, vol. 10, no. 1, pp. 41-47, 2021.
7. L. K. Agustina, R. Setiawan, D. G. Harbowo, B. Alfarishi and D. Raditio, "Determine of Potential Evacuation Locations as Disaster Mitigation Against Landslides," *IOP Conference Series: Earth and Environmental Science*, vol. 1209, no. 1, p. 012006, 2023.
8. D. G. Harbowo, "An assessment of the scientific value of Krakatoa, Indonesia from a geoheritage perspective," *Journal of Applied Geoscience and Engineering*, vol. 2, no. 1, pp. 11-25, 2023.
9. D. G. Harbowo and T. Muliawati, "State transition matrix and Markov-chain diagram for frequent volcanic eruptions: Krakatoa, Indonesia," *E3S Web of Conferences*, vol. 479, p. 02005, 2024.
10. Arisena, Buku Ajar Pengantar Statistika, 2018, pp. 1-46.
11. Abdurrahman and Rachman, Inequality in Human Development Index and Its Components on Regencies/Municipalities In Kalimantan Selatan Province 2010-2020, BESTARI: Buletin Statistikan dan Aplikasi Terkini, 2021.



SENADA
Seminar Nasional Sains Data

Seminar Nasional Sains Data 2024 (SENADA 2024)
UPN "Veteran" Jawa Timur

E-ISSN 2808-5841

P-ISSN 2808-7283

12. P. Irwanto and Soelaiman, "Optimasi Kinerja Algoritma Klasterisasi K-Means Untuk Kuantitasi Warna Citra," *Jurnal Teknik ITS*, pp. 197-202, 2012.
13. Agusta, "K-Means-Penerapan, Permasalahan dan Metode Terkait.," *Jurnal Sistem dan Informatika*, pp. 47-60, Februari 2007.
14. Yana, Setiawan, Ulfa and Rusyana, "Penerapan Metode K-Means dalam Pengelompokan Wilayah Menurut Intensitas Kejadian Bencana Alam di Indonesia Tahun 2013-2018," *Journal of Data Analysis*, vol. 1, no. 2, pp. 93-102, 2018.
15. A. Papoulis, Probabilitas, Variabel Random, dan Proses Stokastik, Yogyakarta: Gadjah Mada University Press, 1992.
16. Ross, Introduction to Probability Models 9th Edition, 2007.
17. Winston, Operations Research Application and Algorith (Third Edition), 2004.
18. Sunarjo, T. Gunawan and S. Pribadi, Gempa Bumi Edisi Populer 2nd Edition, Jakarta: Badan Meteorologi Klimatologi dan Geofisika, 2012.