

# ANALISIS *K-MEDOIDS CLUSTERING* PADA EPISENTRUM GEMPA BUMI DI PROVINSI SUMATERA BARAT DAN SEKITARNYA

Jeri Jefrianto<sup>1</sup>, Devni Prima Sari<sup>2</sup>

<sup>1,2</sup>Prodi Matematika, Fakultas Matematika Ilmu Pengetahuan dan Alam Universitas Negeri Padang (UNP)

<sup>3</sup> Prodi Matematika, Universitas Negeri Padang

## Article Info

### Article history:

Received February 04, 2022

Revised July 15, 2022

Accepted July 16, 2022

### Keywords:

K-Medoids Clustering Analisis  
Sillhouette Coefficient  
Euclidean Distance  
Earthquake

### Kata Kunci:

Analisis *K-Medoids Clustering*  
*Sillhouette Coefficient*  
*Euclidean Distance*  
Gempa Bumi

## ABSTRACT

Sumatera is located along the confluence line of the Eurasian and the Indo-Australia plate, which has high seismic activity. The Sumatera island is prone to earthquakes which pose a considerable danger, and one of the provinces on this island that is prone to earthquakes is the Province of West Sumatra, so it is necessary to earthquake characteristic in the Province of West Sumatra and its surroundings using *K-Medoids Clustering* analysis. The earthquake data used in this research is tectonic earthquake data from December 1970 to September 2021 with attributes: latitude, longitude, and magnitude. In this study, several clusters were tried, namely  $K=2$ ,  $K=3$ ,  $K=4$ , and  $K=5$  using the Euclidean Distance. The most optimum cluster is  $K=2$  0.448072 using the Silhouette coefficient validation test, the epicenter was at  $99.29^{\circ}\text{E}$  and  $-1.05^{\circ}\text{S}$  with a magnitude of 5.38 on the SR for the first group and the second group produced an epicenter at  $101.15^{\circ}\text{E}$  and  $-3.44^{\circ}\text{S}$  with a magnitude of 5.36 SR.

## ABSTRAK

Sumatera terletak sepanjang garis pertemuan pertemuan lempeng Eurasia dan lempeng Indo-Australia, yang memiliki aktivitas seismik yang tinggi. Pulau Sumatera merupakan wilayah rawan gempa yang menimbulkan bahaya yang cukup besar, dan salah satu Provinsi di pulau ini yang rawan gempa adalah Provinsi Sumatera Barat, sehingga perlu dilakukan kajian setiap karakteristik gempa di Provinsi Sumatera Barat dan sekitarnya dengan menggunakan analisis *K-Medoids Clustering*. Data gempa yang digunakan adalah data gempa tektonik bulan Desember 1970 sampai September 2021 dengan atribut: *latitude*, *longitude*, dan *magnitude*. Pada penelitian ini dicoba beberapa kelompok yaitu  $K=2$ ,  $K=3$ ,  $K=4$ , dan  $K=5$  menggunakan jarak *Euclidean Distance*. Kelompok yang paling optimum adalah  $K=2$  sebesar 0,448072 menggunakan uji validitas *sillhouette coefficient*, menghasilkan titik pusat gempa di  $99,29^{\circ}\text{E}$  dan  $-1,05^{\circ}\text{S}$  dengan *magnitude* 5,38 SR untuk kelompok pertama dan kelompok kedua menghasilkan titik pusat gempa di  $101,15^{\circ}\text{E}$  dan  $-3,44^{\circ}\text{S}$  dengan *magnitude* sebesar 5,36 SR.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



## Penulis pertama

(Jeri Jefrianto)

Prodi Matematika, Jurusan Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam,  
Universitas Negeri Padang, Jl. Prof. Dr. Hamka, Air Tawar barat, Padang Utara, Padang, 25171 Padang, Sumatera Barat  
Email: [Jeri.Jefrianto98@gmail.com](mailto:Jeri.Jefrianto98@gmail.com)



## 1. PENDAHULUAN

Sumatera merupakan pulau yang terletak pada sepanjang garis pertemuan lempeng Eurasia dan lempeng Indo-Australia. Garis ini bertemu di dasar laut pada batas lempeng yang disebut dengan palung subduksi Sumatera, dan pada lempeng inilah samudera Indo-Australia perlahan turun menunjam di bawah lempeng benua Eurasia pada kecepatan sebesar 50-70 mm/tahun. Selain itu, pada sepanjang 1900 km Sumatera melewati zona patahan, di dekat atau di dalam busur vulkanik aktif. Sumatera terbelah dari teluk Semongko hingga ke banda Aceh yang disebabkan oleh sesar aktif. Dengan kondisi yang demikian menyebabkan pulau Sumatera rawan terhadap gempa bumi dan mengakibatkan bahaya yang sangat besar pada daerah-daerah yang berpenduduk tinggi disekitar jejak sesar aktif.

Pada 30 September 2009 Provinsi Sumatera Barat mengalami gempa bumi sebesar 7,9 SR dengan pusat gempa (episentrum) 57 km di barat daya Kota Pariaman (0,84 LS dan 99,65 BT) pada kedalaman (hiposentrum) 71 km dengan korban luka berat sebanyak 383.431 orang, luka ringan sebanyak 771 orang dan hilang sebanyak 2 orang. Selain itu, gempa bumi dapat mengakibatkan kerusakan pada bangunan seperti, rumah ibadah, sarana kesehatan, sekolah, dan lainnya. Getaran gempa bumi dirasakan di seluruh wilayah Sumatera Barat serta provinsi-provinsi yang berada di sekitar Sumatera Barat.

Kerusakan yang sangat parah dan berbahaya akibat gempa bumi terjadi pada daerah rawan gempa. Hal ini harus diminimalisir oleh pemerintah dan masyarakat agar menjadi perhatian penting akan dampak tersebut. Mitigasi dan kesiapsiagaan dalam mengantisipasi gempa masih sangat rendah dan belum memiliki *roadmap* secara terencana dan sistematis, sehingga perlu dilakukan kajian pola gempa secara spesial untuk mengantisipasi gempa bumi dan akibatnya. Gempa bumi menyangkut sejumlah besar data, terutama data deret waktu. *Data mining* merupakan pengolahan data dalam jumlah besar yang dimana data mining ini semakin berkembang seiring dengan kemajuan teknologi komputer [1]. Dari itu diperlukan cara yang tepat untuk menghasilkan kesimpulan yang lebih baik. Dalam penelitian ini, data dianalisis berdasarkan klasifikasi seismik berdasarkan magnitudo dan lokasi gempa secara statistik.

Pada era teknologi saat ini membantu dalam mengumpulkan data-data statistik dari kajian gempa yang terjadi pada masa lalu, dan mempelajari karakteristik dari setiap kajian tersebut dengan melakukan pengelompokan data. Teknologi ini dapat membantu secara cepat, efektif dan efisien dalam pengelompokan data atau disebut juga dengan *clustering*. Pada penelitian ini dengan memanfaatkan algoritma *K-Medoids clustering*, yang dapat dikembangkan pada pengelompokan data kejadian gempa bumi, untuk mengetahui karakteristik setiap kelompok dan bisa berguna sebagai peringatan dan penanggulangan bencana gempa bumi. *K-Medoids* atau disebut juga dengan PAM (*Partitioning Around Medoids*) adalah sebuah algoritma *data mining unsupervised* yang menggunakan *medoids* sebagai titik pusat dalam setiap *cluster* [2].

Menurut [3] proses algoritma *K-Medoids* :

1. Menentukan  $k$  (jumlah kelompok) yang diinginkan.
2. Memilih secara random *medoids* awal sebanyak  $k$  dari  $n$  data.
3. Menghitung setiap data (objek) ke *medoids* awal menggunakan rumus *Euclidean Distance*.

Rumus *Euclidean Distance* adalah

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} ; 1, 2, 3, \dots, n \quad (1)$$

dengan :

$d(x, y)$  = jarak data ke  $x$  ke pusat kelompok  $y$

$n$  = jumlah atribut dalam suatu data

$i$  = indeks data

$x_i$  = nilai atribut ke- $i$  dari  $x$

$y_i$  = nilai atribut ke- $i$  dari  $y$

4. Memilih objek secara acak dari setiap kelompok sebagai kandidat *medoids* baru.
5. Menghitung jarak antara setiap objek dari setiap kelompok dengan kandidat *medoids* baru.
6. Menghitung simpangan total ( $S$ )

$$S = \text{Total cost baru} - \text{Total cost lama} \quad (2)$$





di Samudera Hindia dan hanya sebagian kecil yang terjadi di daratan dengan kekuatan *magnitude* yang paling sering 5,0 *SR* sampai dengan 6,0 *SR*.

### 3.1 Analisis pengelompokkan dengan Algoritma *K-Medoids*

Data kejadian gempa bumi dengan atribut *latitude*, *longitude*, dan *magnitude* di Provinsi Sumatera Barat dan sekitarnya menggunakan *K-Medoids Clustering* dicobakan beberapa kelompok yaitu  $K=2$ ,  $K=3$ ,  $K=4$ , dan  $K=5$  dengan bantuan *RStudio* untuk mendapatkan hasil setiap kelompok lebih cepat dan baik. Kelompok  $K=2$  memiliki rata-rata *magnitude* paling besar adalah kelompok pertama sebesar 5,38 *SR* yang memiliki titik pusat gempa di  $99,29^{\circ}\text{E}$  dan  $-1,05^{\circ}\text{S}$  dengan jumlah aktivitas seismik sebanyak 372, kelompok  $K=3$  memiliki rata-rata *magnitude* paling besar adalah kelompok kedua sebesar 5,38 *SR* yang memiliki titik pusat gempa di  $101,28^{\circ}\text{E}$  dan  $-3,51^{\circ}\text{S}$  dengan jumlah aktivitas seismik sebanyak 305, kelompok  $K=4$  memiliki rata-rata *magnitude* paling besar adalah kelompok ketiga sebesar 6,34 *SR* yang memiliki titik pusat gempa di  $100,05^{\circ}\text{E}$  dan -dengan jumlah aktivitas seismik sebanyak 69, kelompok  $K=5$  memiliki rata-rata *magnitude* paling besar adalah kelompok keempat sebesar 6,36 *SR* yang memiliki titik pusat gempa di  $100,03^{\circ}\text{E}$  dan  $-2,15^{\circ}\text{S}$  dengan jumlah aktivitas seismik sebanyak 66. Hasil pengelompokkan dapat dilihat pada Gambar 2 yang telah divisualisasi ke peta.

### 3.2 Evaluasi Hasil Setiap Pengelompokkan dengan Metode *Silhouette Coefficient*

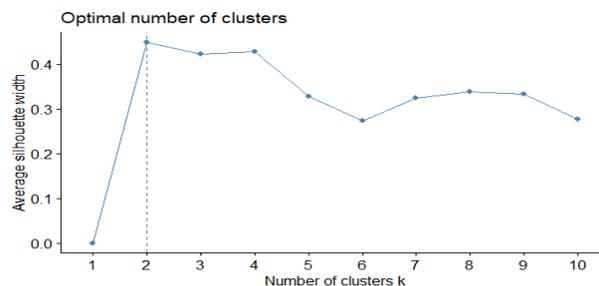
Evaluasi hasil pengelompokkan dengan metode *silhouette coefficient* dilakukan dengan tujuan untuk mengetahui validitas seberapa kuat struktur data pada kelompok.

**Tabel 1.** Hasil Uji Validitas kelompok dengan *Silhouette Coefficient*

Kelompok	Hasil <i>Silhouette Coefficient</i>
2	0,448072
3	0,422957
4	0,4265823
5	0,3281799

Pada Tabel 1 menunjukkan bahwa evaluasi hasil pengelompokkan dengan metode *silhouette coefficient* yang paling optimum terdapat pada kelompok  $K=2$  yaitu sebesar 0,448072.

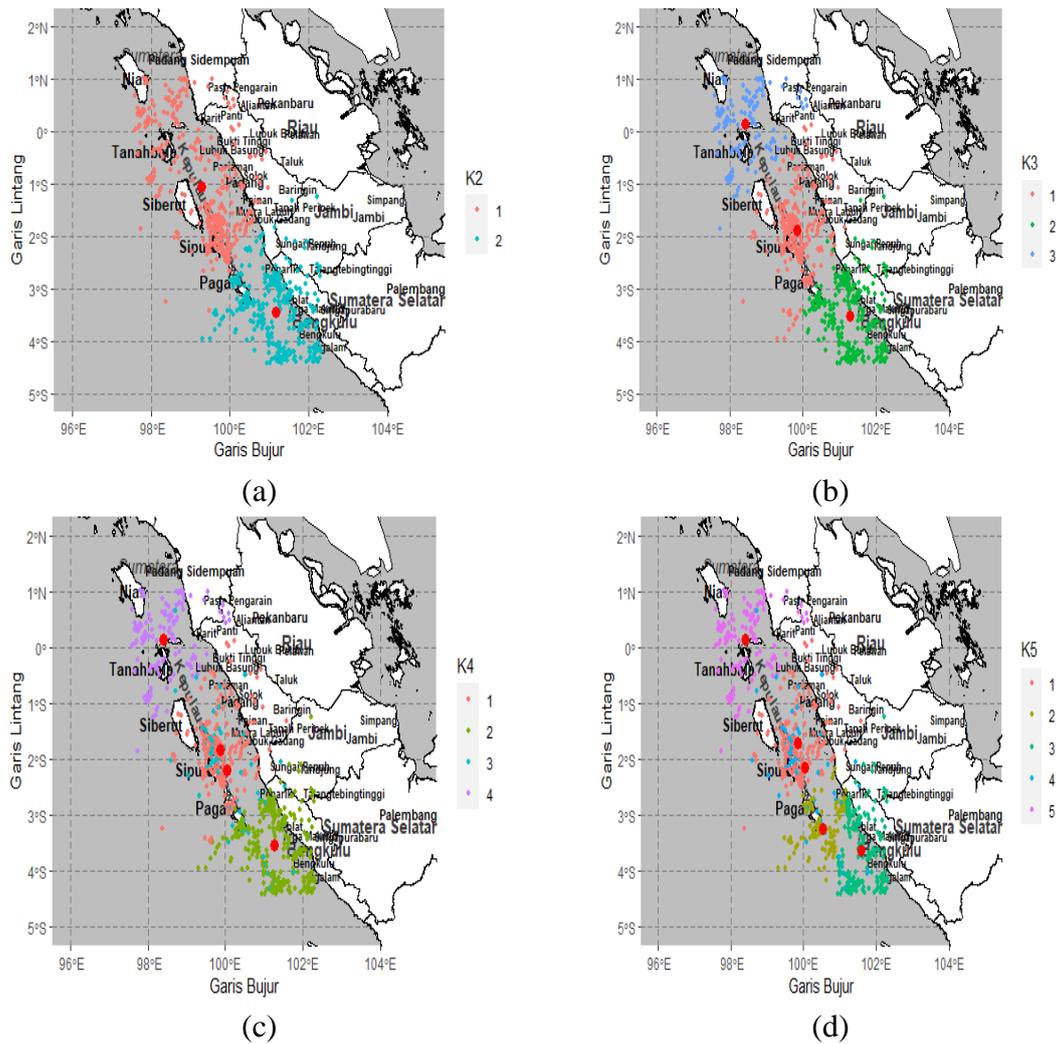
**Gambar 2.** Plots Hasil Uji Validasi *Silhouette Coefficient*



### 3.3 Visualisasi Setiap Kelompok *K-Medoids* ke Peta

Pada tahap ini pengelompokkan data gempa bumi ke peta sesuai dari hasil setiap kelompok *K-Medoids* sebelumnya yaitu  $K=2$ ,  $K=3$ ,  $K=4$ , dan  $K=5$ . Hasil dari setiap kelompok diinput ke peta dengan bantuan *RStudio* dengan cara menggabungkan data asli dari *latitude* dan *longitude* dengan tujuan untuk mengetahui kelompok-kelompok sebaran gempa di peta. Untuk hasil visualisasi setiap kelompok *K-Medoids* ke peta terlebih dahulu menentukan titik pusat kelompok pada peta. Titik pusat didapatkan dengan cara menentukan rata-rata dari setiap kelompok yang ada pada setiap kelompok.

**Gambar 3.** Hasil Visualisasi dengan  $K=2$ ,  $K=3$ ,  $K=4$ , dan  $K=5$



Gambar 3 merupakan hasil keluaran *software RStudio* di mana sumbu- $x$  sebagai posisi bujur dan sumbu- $y$  sebagai posisi lintang. kelompok yang dihasilkan cenderung berkumpul disekitar pusat gempa. Hasil dari setiap kelompok diberikan dengan warna yang berbeda. Pada Gambar 3(a) menunjukkan hasil kelompok gempa tektonik yang terjadi di Provinsi Sumatera Barat dan sekitarnya terhadap kelompok  $K=2$  menghasilkan titik pusat gempa di  $99,29^{\circ}\text{E}$  dan  $-1,05^{\circ}\text{S}$  dengan *magnitude*  $5,38$  *SR* Untuk kelompok pertama. Pada kelompok kedua menghasilkan titik pusat gempa di  $101,15^{\circ}\text{E}$  dan  $-3,44^{\circ}\text{S}$  dengan *magnitude* sebesar  $5,36$  *SR*. Jumlah aktivitas seismik pada kelompok pertama sebanyak 372 objek dan pada kelompok kedua sebanyak 344 objek.

Gambar 3(b) menunjukkan hasil kelompok  $K=3$ . Jika dibandingkan dengan hasil kelompok  $K=2$ , kelompok pertama dan kelompok ketiga pada kelompok  $K=3$  merupakan hasil pembagian dari kelompok pertama pada kelompok  $K=2$ , sedangkan untuk kelompok kedua pada kelompok  $K=3$  hampir sama dengan hasil kelompok kedua pada kelompok  $K=2$ . Gambar 3(c) dan Gambar 3(d) merupakan hasil kelompok dengan  $K=4$  dan  $K=5$ . Untuk  $K=4$  terdapat tumpang tindih antar kelompok yaitu kelompok ketiga yang bersinggungan dengan semua kelompok. Sedangkan Gambar 3(d) untuk  $K=5$  juga menunjukkan hasil kelompok yang tumpang tindih yaitu kelompok keempat yang menyinggung kelompok lain. Berdasarkan nilai *sillhouette coefficient* menggunakan metode *K-Medoids clustering*, kelompok terbaik berada pada  $K=2$ .



#### 4. Kesimpulan

Berdasarkan evaluasi hasil pengelompokkan dengan uji validitas *sillhoutte coefficient*, kelompok yang paling optimum yaitu  $K=2$  dengan nilai sebesar 0,448072 yang menghasilkan titik pusat gempa di  $99,29^{\circ}\text{E}$  dan  $-1,05^{\circ}\text{S}$  dengan *magnitude* 5,38 *SR* yang sebagian besar aktivitas gempanya terjadi di Sumatera Barat termasuk Kepulauan Mentawai, dan kejadian gempa juga terlihat di wilayah Mandiling Natal serta Kepulauan Nias pada kelompok pertama. Magnitudo gempa tertinggi yang pernah terjadi pada kelompok pertama adalah gempa tahun 30 September 2009 di laut lepas Pariaman pada titik pusat gempa  $99,87^{\circ}\text{E}$  dan  $-0,72^{\circ}\text{S}$  dengan magnitudo 7,6 *SR*. Pada kelompok kedua menghasilkan titik pusat gempa di  $101,15^{\circ}\text{E}$  dan  $-3,44^{\circ}\text{S}$  dengan *magnitude* sebesar 5,36 *SR* yang aktivitas gempanya sebagian besar terjadi di Bengkulu Utara, Muko Muko, Pulau Pagai Selatan, dan kejadian gempa juga terlihat di wilayah Jambi, dan Pesisir Selatan. Magnitudo gempa tertinggi yang pernah terjadi pada kelompok kedua adalah gempa tahun 12 September 2007 di Bengkulu pada titik pusat gempa  $100,84^{\circ}\text{E}$  dan  $-2,63^{\circ}\text{S}$  dengan magnitudo 7,9 *SR*. Jumlah aktivitas seismik pada kelompok pertama sebanyak 372 aktivitas dan pada kelompok kedua sebanyak 344 aktivitas.

#### REFERENSI

- [1] Novianti, P., Setyorini, D., & Rafflesia, U. (2017). *K-Means Cluster Analysis in Earthquake Epicenter Clustering*. Vol. 3, No. 2, July 2017, 81-89.
- [2] Han, J., & Kamber, M. (2006). *Data Mining : Concepts and Techniques*. San Francisco, CA: Morgan Kaufmann Publishers.
- [3] ICM2E. (2020). *The Improved Discretization of Peak Ground Acceleration Variable Using K-Medoids*. Padang.
- [4] Kodinariya, T., & Makwana, P. (2013). *Review on Determining Number of Cluster in K-Means Clustering*, "Int. J. Adv. Res. Comput. Sci. Manag. Stud.", vol. 1, no. 6.
- [5] Struyf, A., Hubert, M., & Rousseeuw, P. J. (1997). *Integrating Robust Clustering Techniques in S-PLUS*. 26(1), 17–37.