

# Clustering Data Gempa Bumi Berdasarkan Lokasi Menggunakan K-Medoids

Vika Febri Muliati<sup>#1</sup>

<sup>#</sup>Program Studi Sistem Informasi, Universitas Siber Asia

Jl. Harsono RM No.1, Ragunan Pasar Minggu, Jakarta Selatan, (021) 2780-6189

<sup>1</sup>vikamuliati@lecturer.unsia.ac.id

**Abstrak**— Salah satu bencana alam yang sering terjadi di Indonesia adalah Gempa Bumi. Berdasarkan BMKG, di Indonesia sering terjadi gempa bumi dengan magnitudo 1 hingga 3 skala richter, peristiwa ini disebut dengan gempa micro. Dari tahun 2017 hingga 2018 telah terjadi 1624 data kejadian gempa bumi yang berkekuatan mulai dari 3 hingga 7 skala richter. Letak pusat gempa bumi yang jauh dari pemukiman juga menjadi faktor besarnya getaran gempa yang dirasakan. Identifikasi penyebaran gempa yang terjadi dapat dilakukan dengan metode Clustering dalam Data Mining. Pada penelitian ini metode *clustering* yang dilakukan adalah dengan menggunakan metode K-Medoids. Penelitian ini menggunakan 8 atribut pada data, data gempa yang ada di cluster menjadi 3 bagian, yang terdiri daerah rawan, sedang dan rendah. Hasil dari penelitian ini adalah terdapat 669 daerah yang berada pada tingkat rawan, 512 di tingkat sedang, dan 443 di tingkat jarang. Nilai rata-rata silhouette yang diperoleh adalah 0.32.

**Kata kunci**— BMKG, Clustering, Data Mining, Gempa Bumi, K-Medoids

## I. PENDAHULUAN

### A. Pendahuluan

Indonesia adalah salah satu negara dengan tingginya tingkat rawan terjadinya bencana alam. Letak dan kondisi Indonesia secara geografis, biologi dan demografis mengakibatkan Indonesia menjadi negara yang muda terkena bencana alam. Bencana tersebut dapat berasal dari alam maupun non-alam. Secara geografis, Indonesia merupakan tempat bertemunya 4 lempengan tektonik yaitu lempengan dari Samudra Hindia dan Samudra Pasifik. Bagian selatan hingga Timur Indonesia terdapat sabuk vulkanik atau dikenal dengan volcanic Arc. Sabuk vulkanik ini memanjang dimulai dari pulau Sumatra, Jawa, Nusa Tenggara, dan Sulawesi. Daerah tersebut merupakan daerah dengan sisinya berupa pegunungan vulkanik serta dataran rendah dengan didominasi oleh rawa-rawa [1]. Iklim yang tropis juga menyebabkan pergerakan tanah yang tidak stabil. Infrastruktur dari daerah juga dapat menentukan alasan terjadinya rawan bencana, misalnya tata letak kota tidak sesuai dengan kondisi alam pada daerah tersebut. Bencana alam yang sering terjadi antara lain tsunami, gempa bumi, tanah longsor dan lain sebagainya.

Pergeseran 4 lempengan bumi selain membentuk lipatan dan gunung, juga menyebabkan terjadi getaran pada kulit

bumi dan mempengaruhi fenomena yang ada di permukaan bumi [2]. Gempa bumi merupakan salah satu bencana alam yang rawan dialami di Indonesia. Gempa bumi terdiri dari gempa bumi vulkanik yang disebabkan aktifitas gunung berapi maupun tektonik yang diakibatkan oleh pergerakan lempengan bumi. Berdasarkan BMKG, di Indonesia sering terjadi gempa bumi dengan magnitudo 1 hingga 3 skala richter, peristiwa ini disebut dengan gempa micro [3]. Dari tahun 2017 hingga 2018 telah terjadi 1624 data kejadian gempa bumi yang berkekuatan mulai dari 3 hingga 7 skala richter [4]. Salah satu faktor yang mengakibatkan getaran gempa bumi terasa lebih kuat adalah letak dari pusat gempa bumi yang jauh dari pemukiman. Oleh karena itu, dibutuhkan analisa untuk mengelompokkan daerah dengan tingkat intensitas kejadian rawan, sedang, dan rendah.

Dengan berkembangnya teknologi, data kejadian gempa bumi yang terjadi dari tahun 2017-2018 dapat dikelompokkan dengan menggunakan metode clustering dalam data mining. Data mining merupakan sebuah metode cerdas yang dapat mengekstrak pola dari sebuah data [5]. Sehingga didapat cluster daerah rawan gempa dan daerah rawan tsunami yang disebabkan oleh gempa bumi. Clustering merupakan salah satu metode dalam data mining dengan unsupervised learning, dimana data dikelompokkan berdasarkan kemiripan [6]. Pada penelitian ini menggunakan metode clustering dengan algoritma K-medoids. K-medoids merupakan suatu algoritma yang digunakan untuk mengatasi kelemahan k-means yang sensitif dengan outlier [6]. K-medoids adalah algoritma untuk menemukan k cluster dalam n objek untuk menemukan wakil dari objek (medoids) dalam setiap cluster [7]. Oleh karena itu, penelitian ini mengambil topik tentang clustering daerah di Indonesia rawan gempa dengan menggunakan K-Medoids dan diharapkan penelitian ini dapat memberikan referensi kepada BNPB dalam melakukan penanggulangan bencana, dan dapat berguna untuk masyarakat dalam mengetahui daerah dengan tingkat rawan gempa bumi.

### B. Dasar Teori Mutakhir

#### Gempa Bumi

Pelepasan energi oleh gelombang seismic secara tiba-tiba merupakan peristiwa yang disebut gempa bumi. Peristiwa ini terjadi karena deformasi lempeng tektonik pada kerak bumi.

Gempa bumi di Indonesia terjadi mencapai 3 hingga 5 kali selama setahun, dan hal ini mengalami banyak kerusakan [8].

Berdasarkan proses terjadinya, gempa dibedakan menjadi 3 macam yaitu gempa vulkani, gempa tektonik dan gempa runtuh atau disebut dengan gempa buatan. Di Indonesia, gempa yang sering terjadi adalah gempa vulkanik dan juga gempa tektonik. Gempa vulkanik merupakan gempa yang diakibatkan oleh aktivitas vulkanik. Gempa vulkani dapat dirakan saat gunung berapi akan meletus, sedang meletus atau setelah meletus. Sedangkan gempa tektonik merupakan gempa yang terjadi karena adanya pertemuan antara lempengan di lapisan litosfer pada kulit bumi yang dikarenakan tenaga tektonik [8].

Berdasarkan dampaknya, gempa terbagi menjadi dua yaitu dampak dari gempa vulkanik dan dampak dari gempa tektonik [8]. Dampak dari gempa tersebut adalah:

- Dampak dari gempa vulkanik. Gempa yang terjadi karena gunung berapi ini memberikan dampak dengan adanya lava yang keluar dari gunung berapi, lahar, dan juga lontaran material serta adanya abu letusan yang berasal dari gunung berapi.
- Dampak dari gempa tektonik. Dampak dari gempa ini dibagi menjadi dua, yaitu dampak primer dan dampak sekunder. Dampak primer adalah dampak atau efek yang diakibatkan oleh gempa tersebut, contohnya seperti korban jiwa dan kerusakan infrastruktur bangunan serta kerugian lainnya. Sedangkan yang termasuk dalam dampak sekunder seperti terjadinya tsunami, tanah longsor dan kebakaran serta penyakit yang menyebar.

### C. Metode Penelitian

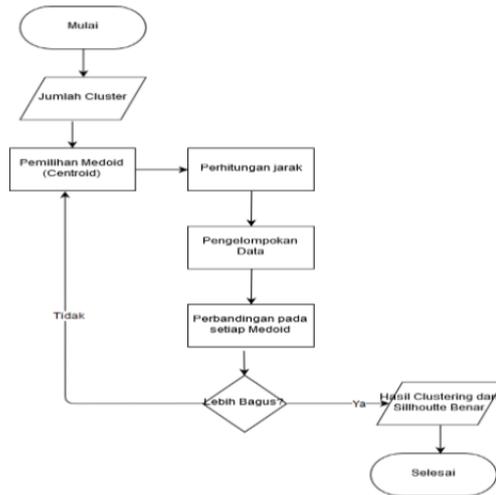
#### K-Medoids

K-Medoids adalah algoritma yang memiliki hubungan dengan K-Means. K-means dan K-medoids merupakan algoritma yang memiliki sifat dengan memecahkan data ke beberapa cluster atau sering disebut dengan partisional. Nilai  $n$  pada kedua algoritma bertujuan untuk meminimalkan jarak pada setiap cluster [6]. Variabel yang digunakan pada data ini adalah depth, mag, gap, dmin, depthError, magError, magNst, place. Analisis yang digunakan menggunakan metode K-Medoids. K-medoids digunakan dalam peng-clusteran dengan objek terdekat. Langkah-langkah penggunaan metode K-Medoids adalah sebagai berikut [6]:

- Menentukan jumlah dari cluster yang digunakan.
- Memilih acak dari data yang akan digunakan dalam medoid awal sebanyak jumlah cluster yang digunakan.
- Menghitung jarak terhadap medoid awal
- Mengelompokkan data dengan medoid awal yang berjarak paling dekat dengan data tersebut dan menghitung cost
- Mengganti salah satu medoid dengan data dari kelompoknya dan melakukan proses perhitungan. Jika cost yang dihasilkan lebih kecil maka medoid akan diganti dengan medoid baru dan terjadi perubahan data yang baru.

Metode Penelitian merupakan rancangan pelaksanaan penelitian dengan mengarah pada hasil yang dituliskan secara jelas, memuat teknik pengambilan data dan analisis yang digunakan dalam memperoleh hasil, selain itu dapat memuat jumlah responden yang digunakan dalam penelitian. Penulisan

rumus dapat menggunakan *equation* dengan mencantumkan sumber utama yang digunakan sebagai acuan.



Gambar. 1 Metode K-Medoids

#### Sillhouette Coefficient

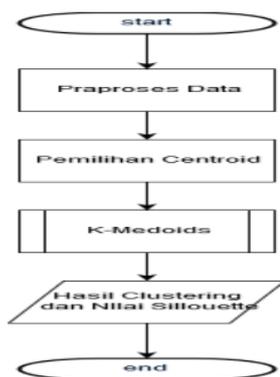
Sillhouette coefficient adalah metode yang digunakan untuk melihat kualitas penempatan objek pada cluster tersebut. [6]. Tahapan dalam perhitungan sillhoutte adalah:

- Melakukan perhitung rata-rata objek ke-i terhadap semua objek yang terdapat di kelompok rata-rata jarak tersebut.
- melakukan perhitung rata-rata objek pada cluster lain dengan simbol  $b(i)$ , dan ambil nilai terkecil.
- Nilai sillhoutte coefficient didapat dengan menggunakan rumus.

Nilai dari sillhoutte berada berkisar antara -1 sampai 1. Jika nilai mendekati nilai 1 maka nilai tersebut berada pada cluster yang benar. jika nilai mendekati nilai 0, maka nilai tersebut berada pada antara 2 cluster. jika nilai sillhoutte bernilai negatif, maka nilai tersebut tidak berada pada cluster yang benar.

#### Metode Implementasi

Tahap pengerjaan dalam penelitian ini dimulai dari praproses data, kemudian melakukan pemilihan nilai centroid, dan melakukan implementasi dengan menggunakan R Studio, dan menganalisa hasil cluster dan nilai sillhoutte yang ada. Diagram dari tahapan implementasi dapat dilihat pada Gambar 2.



Gambar. 2 Diagram Implementasi

**D. Hasil dan Analisa**

Data penelitian ini didapat dari website United States Geological Survey (USGS) yang diunduh pada website. Data yang digunakan adalah data dari 1 Januari 2017 hingga 18 April 2018, mempunyai 1624 data dengan 8 atribut. Atribut yang digunakan adalah depth, mag, gap, dmin, depthError, magError, magNst, dan place. Keterangan dari atribut tersebut dapat dilihat pada tabel 1.

TABEL I  
ATRIBUT

Atribut	Keterangan
Depth (Kedalaman)	Kedalaman gempa
Mag (Magnitude)	Kekuatan gempa yang dirasakan
Gap	Perhitungan kepastian lokasi gempa
Dmin	Jarak lokasi gempa ke stasiun terdekat
DepthError	Kesalah kedalaman
MagError	Kesalahan kekuatan gempa
MagNst	Jumlah stasiun untuk menghitung kekuatan gempa
Location	Lokasi Gempa

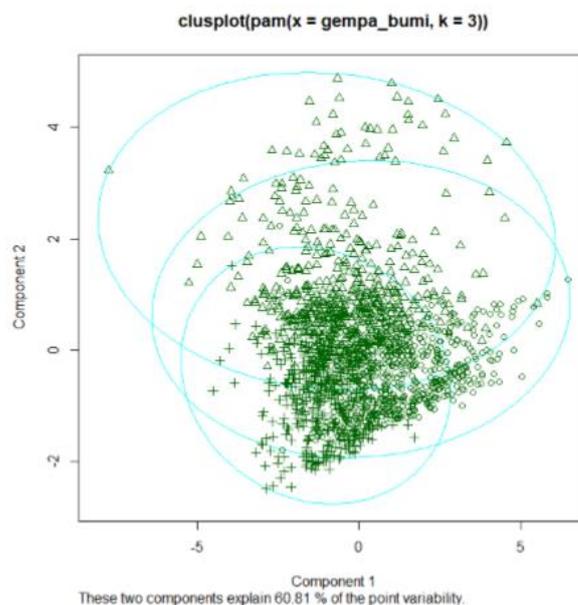
Data gempa yang terdiri dari 1624 data, dilakukan preprocessing dengan cara mengisi data N/A dengan nilai 0 di excel. Selanjutnya dengan menggunakan tools R Studio, dilakukan peng-clusteran dengan menggunakan k-medoids. Dengan menggunakan R maka didapat hasil cluster dengan 3 cluster dengan kategori rawan, sedang dan rendah. Hasil cluster dilakukan setiap kabupaten pada 31 provinsi di Indonesia. Hasil dari pembagian 3 cluster dapat dilihat pada Gambar 3.

```

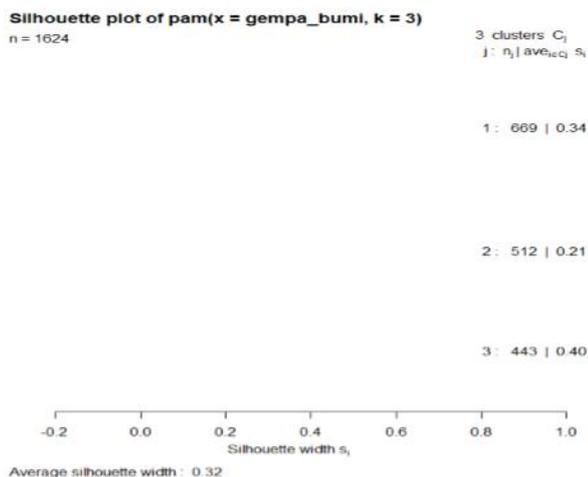
Medoids:
  ID depth mag gap dmin depthError magError magNst
1041 1041 38.94 4.6 73 0.761 6.7 0.097 32
236 236 173.68 4.4 85 3.502 9.2 0.109 24
800 800 42.98 4.2 143 2.366 9.5 0.132 18
  
```

Gambar. 3 Hasil Clustering

Cluster di atas menunjukkan bahwa cluster pertama merupakan cluster rawan, dengan tingkat kekuatan gempa tinggi dan dengan tingkat Error yang kecil dan jumlah perhitungan stasiun yang lebih banyak. Cluster kedua mempunyai nilai error menengah dalam melakukan perhitungan, dan cluster ketiga merupakan cluster dengan tingkat kedalaman dan kekuatan yang rendah dan tingkat kesalahan yang tinggi dalam melakukan perhitungan. Dari riwayat gempa bumi yang didapat, maka di dapat 3 cluster dengan total cluster 1 dengan jumlah 669 kota, cluster 2 dengan jumlah data 512 kabupaten, dan 443 kabupaten berada pada cluster 3. Hasil cluster dapat dilihat pada Gambar 4.



Gambar. 4 Hasil Plot dengan Jumlah Cluster 3



Gambar. 5 Hasil Perhitungan Sillhouette

Dari hasil cluster didapat nilai rata-rata sillhoutte 0.32. cluster pertama memiliki nilai sillhoutte 0.34, cluster kedua 0.21 dan cluster ketiga 0.40. Data sillhoutte yang didapat dapat dilihat pada Figure 5. Data sillhoutte yang didapat dari 3 cluster menandakan bahwa data yang ada mendekati nilai 1 dan menandakan bahwa data telah berada pada cluster yang benar. semakin besar cluster yang dilakukan maka nilai sillhoutte akan semakin besar dan jarak yang ada pada setiap nilai cluster semakin kecil. Hal ini menghasilkan nilai sillhoutte yang baik. Setelah di dapat hasil cluster, maka didapat daerah yang berada pada cluster 1 hingga 3 yang dapat dilihat pada Gambar 6.

depth	mag	gap	dmin	depthE	magErr	magNsi	place	Hasil Cluster
64.48	5	91	0.545	6.7	0.075		57 64km SSE of Bengkulu, Indonesia	1
34.51	4.2	79	3.03	4.2	0.152		12 7km N of Radamata, Indonesia	1
10	4.6	45	1.728	1.9	0.083		44 189km NNE of Tual, Maluku, Indonesia	1
28.2	5.1	49	0.719	4.3	0.054		111 25km WNW of Ogotua, Sulawesi Teng	1
12.59	5.1	30	3.514	1.5	0.073		18 37km WNW of Ogotua, Sulawesi Teng	1
10	4.7	101	2.388	1.9	0.075		54 207km S of Kangin, Indonesia	1
10	4.7	91	2.962	1.7	0.078		50 South of Java, Indonesia	1
88.25	5.4	17	6.4	6.5	0.073		18 250km E of Enarotali, Indonesia	1
35.04	4.3	71	0.617	10.3	0.119		20 68km W of Kota Ternate, Indonesia	1
57.54	4.7	45	0.489	8.1	0.076		52 53km NW of Kota Ternate, Indonesia	1
35	5	76	2.859	1.9	0.059		93 238km NNW of Tobelo, Indonesia	1
44.61	4.6	47	1.263	8	0.072		58 114km ESE of Bitung, Indonesia	1
10	5.3	64	1.275	1.8	0.068		21 113km ESE of Bitung, Indonesia	1
60.42	5.3	53	2.738	5.8	0.071		19 79km SSW of Padangcermin, Indonesia	1
30.85	4.6	81	1.206	5.9	0.067		67 51km SE of Sinabang, Indonesia	1
56.29	4.2	77	1.013	9.2	0.159		11 113km SW of Kota Ternate, Indonesia	1
24.45	4.9	60	3.213	5.7	0.091		38 51km W of Dobo, Indonesia	1
26.69	4.8	64	6.923	5	0.056		98 178km WNW of Abepura, Indonesia	1
39.74	4.9	66	5.887	6.9	0.052		116 Near the north coast of Papua, Indonv	1
10	4.3	96	1.302	2	0.188		8 90km NE of Amuntai, Indonesia	1
29.65	5.2	32	1.349	3.4	0.073		18 149km NW of Kota Ternate, Indonesia	1
44.5	4.9	68	1.633	6.2	0.04		192 145km WNW of Tobelo, Indonesia	1

Gambar. 5 Hasil Perhitungan Sillhouette

E. Kesimpulan

Dari penelitian di atas maka dapat disimpulkan bahwa cluster yang dapat dibentuk pada data gempa bumi dari tahun 2017 sampai 2018 yaitu 3 cluster. Cluster terbentuk menjadi rawan, sedang, dan jarang. cluster rawan terdiri dari 669 dengan nilai sillhoutte 0.34. Cluster kedua yaitu cluster sedang dengan 512 kota dengan nilai sillhoutte 0.21, dan pada cluster

3 yaitu daerah jarang dengan jumlah 443 kota dan dengan nilai 0.40. Rata-rata nilai sillhoutte adalah 0.32, yang berarti nilai yang di cluster telah berada pada cluster yang benar.

REFERENSI

- [1] BNPB, "Potensi dan ancaman bencana," 2018.
- [2] R. Pebria, B. Heru, I. Sugihartono et al., "Identifikasi penyebaran gempa di indonesia dengan metode clustering," in PROSIDING SEMINAR NASIONAL FISIKA (E-JOURNAL), vol. 3, 2014, pp. 366–370.
- [3] A. Febriani, B. Siska, and R. Hakim, "Analisis clustering gempa bumi selama satu bulan terakhir dengan menggunakan algoritma self-organizing maps (soms) kohonen," 2015.
- [4] USGS, "Earthquake hazard programs," 2018.
- [5] J. Han, M. Kamber, and J. Pei, "Data mining: concepts and techniques (the morgan kaufmann series in data management systems)," Morgan Kaufmann, 2000.
- [6] D. A. S. Simamora, "Clustering data kejadian tsunami yang disebabkan oleh gempa bumi dengan menggunakan algoritma k-medoids," Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer, 2017.
- [7] W. A. Triyanto, "Algoritma k-medoids untuk penentuan strategi pemasaran produk," Simetris: Jurnal Teknik Mesin, Elektro dan Ilmu Komputer, vol. 6, no. 1, pp. 183–188, 2015.
- [8] Hartuti. 2009. Buku Pintar Gempa. Yogyakarta: Diva Press.