

**PENGELOMPOKAN DAMPAK GEMPA BUMI DARI SEGI
KERUSAKAN FASILITAS PADA PROVINSI YANG BERPOTENSI
GEMPA DI INDONESIA MENGGUNAKAN K-MEANS-CLUSTERING**

Hepita Artatia, RB Fajriya Hakim

Program Studi Statistika Fakultas MIPA Universitas Islam Indonesia

hepitaartatia36@gmail.com, hakimf@fmipa.uii.ac.id

ABSTRAK. Gempa bumi adalah salah satu bencana alam besar yang sering terjadi di Indonesia, dengan tingkat skala richter yang bervariasi. Gempa bumi disebabkan karena pergeseran lempeng (kerak bumi). Jika daerah potensi rawan gempa berada didaerah puncak maka akan terjadi letusan gunung api yang diawali dengan adanya gempa vulkanik sedangkan jika daerah potensi gempa berada dibawah dan terjadi patahan pada lempengan maka akan terjadi gempa tektonik. Banyaknya bencana gempa bumi yang terjadi, menyebabkan banyak dampak kerusakan baik dari segi ekonomi, sosial, bisnis hingga fasilitas umum. Fasilitas umum seperti jalan raya, sekolah, rumah sakit dan lain-lain merupakan aspek yang penting pada suatu wilayah. Data kerusakan fasilitas yang didapat akibat gempa bumi yang terjadi di Indonesia akan dilakukan pengelompokan menggunakan Teknik Data Mining dengan Metode Clustering menggunakan Algoritma K-Means, sehingga akan didapatkan cluster dampak Gempa Bumi dari segi kerusakan fasilitas.

Kata Kunci: *Gempa Bumi; Kerusakan Fasilitas; K-Means Clustering.*

1. PENDAHULUAN

Indonesia merupakan daerah pertemuan 3 lempeng tektonik besar, yaitu lempeng Indo-Australia, Eurasia dan lempeng Pasifik. Lempeng Indo-Australia bertabrakan dengan lempeng Eurasia di lepas pantai Sumatra, Jawa dan Nusa Tenggara, sedangkan dengan Pasifik di utara Irian dan Maluku utara. Pertemuan tiga lempeng tersebut menyebabkan kontur tanah dan relief di Indonesia bervariasi, dari dataran rendah hingga dataran tinggi, dari suhu yang rendah hingga suhu tinggi dan dari tanah gambut hingga tanah bertekstur kapur. Sehingga tidak heran jika Indonesia menjadi negara yang rawan bencana. Pertemuan tiga lempeng tersebut membuat Indonesia dikelilingi banyak gunung api dan lempengan tersebut kerap kali bergeser dan melepaskan energi yang besar. Bencana alam merupakan bencana yang tidak dapat dihindari dan tidak dapat diramalkan kapan terjadi dan akibat yang ditimbulkan.

Gempa bumi adalah salah satu bencana alam besar yang sering terjadi di Indonesia, dengan tingkat skala richter yang bervariasi. Kerusakan yang ditimbulkan pun tidak main-main. Karena gempa bumi tidak hanya terjadi di darat, namun terjadi di dasar laut yang menyebabkan gelombang tsunami. Jika daerah potensi rawan gempa berada didaerah puncak maka akan terjadi letusan gunung api yang diawali dengan adanya gempa vulkanik

sedangkan jika daerah potensi gempa berada dibawah dan terjadi patahan pada lempengan maka akan terjadi gempa tektonik.

Indonesia kerap kali menjadi sorotan dunia akibat bencana alam yang terjadi di Negara ini sehingga Negara ini dijuluki dengan museum bencana. Banyaknya bencana gempa bumi yang terjadi menyebabkan banyak dampak kerusakan baik dari segi ekonomi, sosial, bisnis hingga fasilitas umum. Fasilitas umum seperti jalan raya, sekolah, rumah sakit dan lain-lain merupakan aspek yang penting pada suatu wilayah. Dari hal tersebut penelitian ini akan membahas tentang klasifikasi data kerusakan fasilitas di sejumlah provinsi. Sehingga dengan clustering ini kita dapat mengetahui tingkat kerusakan yang ditimbulkan sehingga masyarakat atau pemerintah dapat meningkatkan kemampuan antisipasi bencana untuk mengurangi resiko yang ditimbulkan bencana.

Dari uraian latar belakang diatas, maka permasalahannya adalah ingin mengetahui bagaimana dampak kerusakan fasilitas akibat gempa dengan menggunakan metode clustering dengan algoritma K Means yang bertujuan untuk mengetahui klastering kerusakan yang diakibatkan.

2. METODE PENELITIAN

Normalisasi data menggunakan metode Min-max merupakan metode yang menghasilkan *transformasi linier* pada data asal dimana Min-Max Normalization memetakan sebuah nilai v dari A menjadi v' dalam *range* nilai minimal dan maksimal yang baru.

Rumus Min-max normalization adalah :

$$v' = \frac{v - \min A}{\max A - \min A} \times (\text{new}_{\max A} - \text{new}_{\min A}) + \text{new}_{\min A}$$

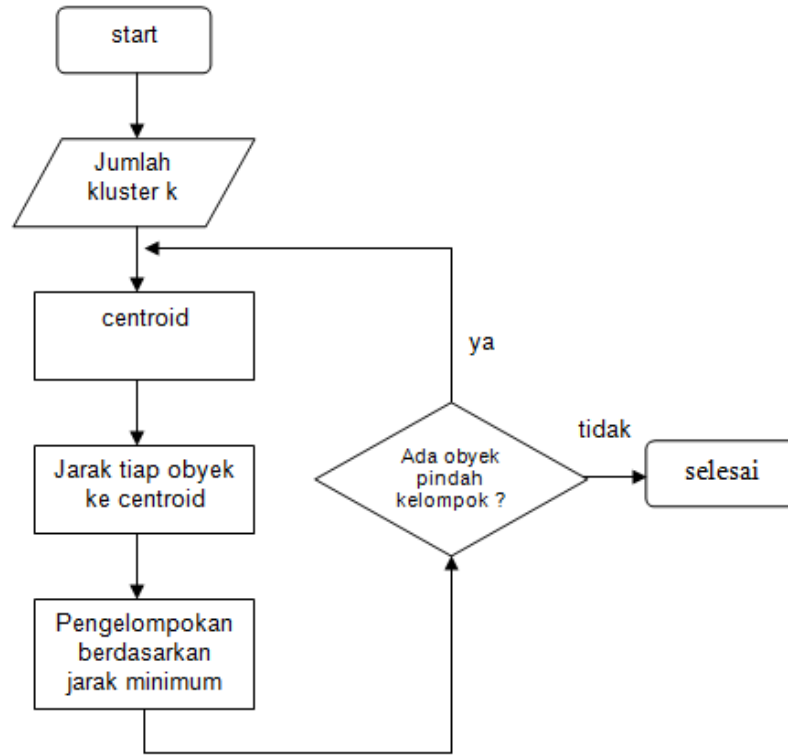
Dimana $\min A$, $\max A$, $\text{new}_{\max A}$ dan $\text{new}_{\min A}$ masing-masing adalah merupakan nilai minimum dan nilai maksimum dari atribut A serta nilai maksimum dan nilai minimum pada skala baru dari atribut A .

Clustering merupakan proses membagi data dalam suatu himpunan ke dalam beberapa kelompok yang kesamaan datanya dalam suatu kelompok lebih besar dari pada kesamaan tersebut dengan data kelompok lainnya (Santosa, [2]).

Clustering mengelompokkan data tanpa berdasarkan acuan kelas data tertentu. Bahkan clustering dapat dipakai untuk menentukan kelas data bagi data dengan kelas yang belum diketahui, teknik clustering ini sering digolongkan sebagai metode *unsupervised learning*. Prinsip kerja dari teknik clustering adalah memaksimalkan kesamaan antar anggota dalam satu kelas dan meminimumkan kesamaan antar kelas. Banyak algoritma clustering menggunakan fungsi jarak untuk mengukur kemiripan antar data, diperlukan juga metode untuk normalisasi bermacam atribut data.

K-Means merupakan salah satu metode pengelompokan data nonhierarki (sekatan) yang berusaha mempartisi data yang ada ke dalam bentuk dua atau lebih kelompok. Metode ini mempartisi data kedalam kelompok sehingga data yang berkarakteristik berbeda dikelompokkan ke dalam kelompok yang lain. Adapun tujuan pengelompokan data ini adalah untuk meminimalkan variasi di dalam suatu kelompok dan memaksimalkan variasi antar kelompok (Prasetyo, [4]).

Diagram alir algoritma k-means :



Gambar 1 diagram alir algoritma k-means clustering

Pengelompokan data dengan metode K- means ini secara umum dilakukan dengan algoritma sebagai berikut :

1. Mulai dengan menetapkan nilai k = jumlah kelompok
2. Ambil partisi awal yang mengelompokkan data ke dalam k kelompok. Bisa digunakan sampel pelatihan secara random, atau secara sistematis seperti berikut,
 - a. Ambil sampel pelatihan awal sebanyak k sebagai anggota kelompok tunggal
 - b. Tentukan untuk setiap data lainnya ($N-k$) menjadi anggota kelompok kepada centroid yang terdekat. Setelah penentuan dari setiap sisa data tadi, hitung kembali centroid dari kelompok yang telah terbentuk
3. Ambil setiap sampel pada barisan dan hitung jaraknya dari centroid yang baru pada setiap kelompok. Jika sebuah sampel tidak berada pada kelompok awal, maka sampel tersebut pasti menjadi anggota kelompok yang lain.
4. Ulangi langkah 3 sampai kekonvergenan tercapai, yakni sampai pada tidak ada perubahan sebuah data menjadi anggota kelompok yang lain.

Jika jumlah data lebih kecil daripada jumlah kelompok maka setiap data menjadi centroid kelompok. Setiap centroid akan mempunyai sejumlah kelompok. Jika jumlah data lebih besar daripada jumlah kelompok, untuk setiap data, dihitung jarak dari setiap seluruh centroid dan

cari jarak minimum. Data dikatakan masuk dalam sebuah kelompok jika memiliki jarak minimum ke dalam kelompok tersebut.

Jumlah minimum dari atribut data adalah satu, jika jumlah atribut data adalah satu maka setiap titik sampel mewakili sebuah titik pada suatu distribusi. Algoritma k-means menjadi cara untuk menghitung nilai rata-rata dari k distribusi.

Berikut ini adalah penelitian terdahulu tentang algoritma *K-means-clustering* yang pernah dilakukan untuk Pengenalan Signal EKG Menggunakan Dekomposisi Paket Wavelet dan K-means-clustering (Rizal & Suryani, [6]). Pernah juga dilakukan untuk Verifikasi Citra Sidik Jari Poin Minutiae dalam Visum Et Repertum (VER) Menggunakan K-means Cluster (Pratama, [5]). Pernah juga dilakukan untuk Pembentukan Cluster dalam Knowledge Discovery in Database dengan Algoritma K-means (Andayani, [1]). Juga pernah dilakukan pada penelitian Data Mining dengan Teknik Clustering dalam Pengklasifikasian Data Mahasiswa Studi Kasus Prediksi Lama Studi Mahasiswa Universitas Bina Nusantara (Lindawati, [3]).

Metode yang digunakan dalam analisis ini adalah metode *clustering* dengan *Algoritma K-means* untuk data dampak bencana gempa bumi di provinsi Indonesia. Tahapan yang dilakukan pada penelitian ini dibagi menjadi 4 tahapan, yaitu pengumpulan data, analisa, standarisasi, penentuan jumlah cluster dan clustering data dengan algoritma k-means. Pada tahapan ini juga ditentukan atribut-atribut yang digunakan yaitu rumah rusak berat, rumah rusak ringan, fasilitas kesehatan, fasilitas pendidikan, kerusakan jalan dan kerusakan lahan.

Data bersumber dari website Data dan Informasi Bencana Indonesia (Badan Nasional Penanggulangan Bencana) yaitu <http://dibi.bnpb.go.id>.

Berikut adalah jumlah data kerusakan fasilitas di provinsi Indonesia yang telah diakumulasi per provinsi :

Tabel 1 Data Jumlah Kerusakan

PROVINSI	RUMAH RUSAK BERAT	RUMAH RUSAK RINGAN	FASILITAS KESEHATAN	FASILITAS PENDIDIKAN	KERUSAKAN JALAN KM	KERUSAKAN LAHAN
Bali	4178	0	16	321	0	0
banten	41	179	0	1	0	0
bengkulu	24175	36221	327	904	0	0
DI Yogyakarta	93318	101236	186	1950	0	0
DKI Jakarta	0	0	0	0	0	0
Gorontalo	233	347	2	20	0	0
Jambi	643	937	3	23	0.03	0
Jawa Barat	75983	186696	0	5656	0	93
Jawa Tengah	38165	69439	121	579	0	500
Jawa Timur	1878	5638	137	144	0	0
Kalimantan Timur	7	141	4	6	0	0
Lampung	1120	8266	0	0	0	0
Maluku	1925	172	0	5	0	0

Maluku Utara	873	1609	3	23	0	0
Nusa Tenggara Barat	6601	20284	22	125	1	0
Nusa Tenggara Timur	27169	24632	9	556	0	25
Papua	6056	7761	109	478	53	0
Papua Barat	2842	1251	34	112	14	0
Pemerintah Aceh	12069	15442	332	1472	22	0
Sulawesi Barat	1287	38	2	34	50	0
Sulawesi Selatan	0	0	0	0	50	0
Sulawesi Tengah	6158	6819	6	25	0	0
Sulawesi Tenggara	101	0	0	1	0	0
Sulawesi Utara	647	3735	0	37	0	0
Sumatera Barat	146877	210432	415	7224	296.6	0
Sumatera Selatan	194	682	1	3	0	0
Sumatera Utara	47683	1869	135	908	1650.33	1968

3. HASIL PENELITIAN DAN PEMBAHASAN

Dalam penyelesaian kasus ini, pertama dilakukan standarisasi menggunakan metode transformasi min-max dengan tujuan merubah skala data yang berbeda pada masing-masing variabel menjadi skala baru yang sama. Tahapan selanjutnya dilakukan analisis K-means clustering untuk mengetahui ada tidaknya perbedaan yang nyata (signifikan) antar kelompok yang terbentuk, melihat profil serta kecenderungan-kecenderungan dari masing-masing cluster dan melihat posisi masing-masing obyek terhadap obyek lainnya dari cluster yang terbentuk.

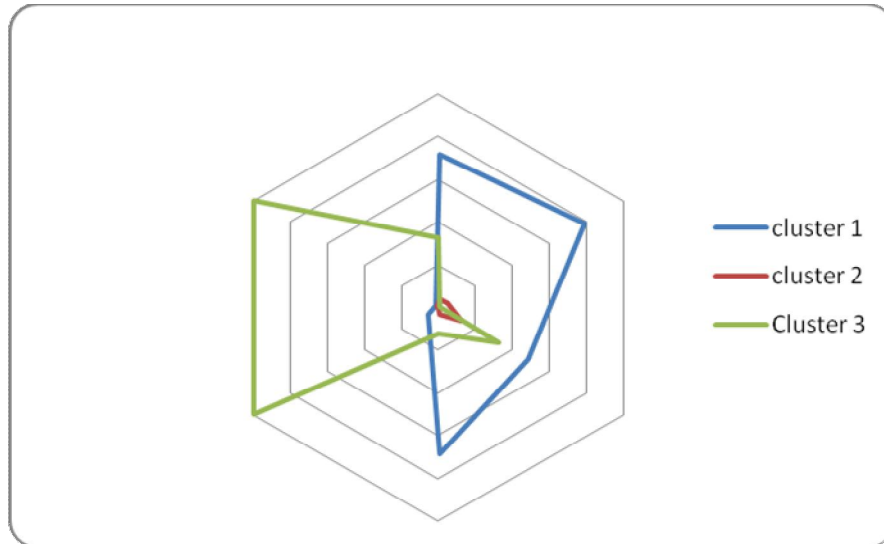
Metode clustering yang digunakan yaitu algoritma K-means. Hasil dari rata-rata clustering data disajikan dalam tabel berikut :

Tabel 2 Hasil Cluster

	Rumah Rusak Ringan	Rumah Rusak Berat	Fasilitas Kesehatan	Fasilitas Pendidikan	Kerusakan Jalan	Kerusakan Lahan
cluster 1	7.175573212	7.89429998	4.827309237	6.842930971	0.599072105	0.157520325
cluster	0.421403933	0.438299654	1.234392114	0.306176382	0.052339394	0.121258315

2						
Cluster						
3	3.246457921	0.08881729	3.253012048	1.256921373	10	10

Hasil diatas, diolah dengan menggunakan statistik deskriptif grafik sarang laba-laba sehingga diperoleh hasil sebagai berikut :



Grafik 1 Hasil Analisis

Grafik diatas dapat menjelaskan bahwa, hasil pada cluster 1 gempa hanya berdampak pada rumah baik rusak berat maupun rusak ringan, fasilitas kesehatan dan fasilitas pendidikan dengan kata lain cluster 1 lebih berdampak pada kerusakan bangunan. Sedangkan cluster 2 tidak memperlihatkan dampak gempa yang terlalu signifikan karena dampak yang diakibatkan tidak begitu berbeda pada setiap atribut kerusakan fasilitas yang semua nilainya sama-sama rendah. Berbeda dengan cluster ke 3, dampak gempa terlihat begitu kontras antar atribut yakni cluster 3 lebih condong berdampak pada kerusakan lahan dan kerusakan jalan yang begitu tinggi dibandingkan dengan kerusakan fasilitas lainnya.

3. SIMPULAN

Dari pembahasan diatas, dapat disimpulkan bahwa dampak yang paling tinggi akibat gempa di Indonesia adalah kerusakan lahan, jalan dan rumah rusak ringan dan metode clustering dengan algoritma K-means dapat digunakan untuk mengclusterkan kerusakan fasilitas dari dampak gempa bumi di Indonesia. Sehingga didapatkan hasil cluster yang nantinya dapat digunakan untuk mengetahui pengelompokan kerusakan agar nantinya kualitas fasilitas-fasilitas yang ada dapat di tingkatkan seperti bangunan tahan gempa sekaligus mengurangi resiko bencana.

DAFTAR PUSTAKA

- [1] Andayani, S. (2007). Pembentukan Cluster dalam Knowledge Discovery in Database dengan Algoritma K-means. *SEMNAS Matematika dan Pend Matematika*, (pp. 1-10). Yogyakarta.
- [2] B, Santosa. 2007. *Data Mining Teknik Pemanfaatan data untuk Keperluan Bisnis, First Edition*. Yogyakarta: Graha Ilmu.
- [3] Lindawati. (2008). Data Mining dengan Teknik Clustering dalam Pengklasifikasian Data Mahasiswa Studi Kasus Prediksi Lama Studi Mahasiswa Universitas Bina Nusantara. *Seminar Nasional Informatika UPN Veteran* (pp. 174-180). Yogyakarta: Informatika.
- [4] Prasetyo, E. (2012). *Data Mining Konsep dan Aplikasi Menggunakan Matlab*. Yogyakarta: Andi.
- [5] Pratama, A. B. 2008. *Verifikasi Citra Sidik Jari Poin Minutiae dalam Visum Et Repertum (VER) Menggunakan K-means Cluster*. Ilmu Komputer UB , 10.
- [6] Rizal, A., & Suryani, V. (2008). Pengenalan signal EKG meggunakan dekomposisi paket Wavelet dan K-Means clusering. *Seminar Nasional Aplikasi Teknologi Informasi 2008* (pp. J-51). Yogyakarta: SNATI.