JDSP
Jurnal Data Science Penusa
E-ISSN: 1234-5678

Sistem Informasi Geografis Pemetaan Berbasis Web Untuk Prediksi Gempa Bumi Di Kabupaten Karo Dengan Metode K-Means Clustering

Milkha Jabenayegi Pillai[1)*], Megaria Purba[2)]
[1, 2]Sekolah Tinggi Manajemen Informatika Dan Komputer Pelita Nusantara
[1]milkhajabenayegipillai@gmail.com, [2]megariapurba18@gmail.com

Received: 12 September 2025 Accepted: 25 Juni 2026 Published: 13 July 2026
Pp. 302-313

***Penulis Koresponden**

Kata Kunci:
SIG
K-means Clustering
Gempa Bumi
Kabupaten Karo

JDSP: Jurnal Data Science Penusa is licensed under a Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0).

Penelitian ini mengembangkan Sistem Informasi Geografis (SIG) berbasis web untuk memetakan tingkat kerawanan gempa bumi di Kabupaten Karo dengan menerapkan algoritma K-means Clustering. Tujuan utama penelitian adalah mengelompokkan wilayah berdasarkan parameter spasial seperti kepadatan penduduk, jarak ke patahan aktif, dan frekuensi kejadian gempa. Hasil clustering membagi wilayah ke dalam tiga kategori risiko: tinggi, sedang, dan rendah, yang divisualisasikan dalam peta digital interaktif. Sistem dibangun menggunakan teknologi PHP dan MySQL, serta dilengkapi fitur input data lokasi, pengolahan data otomatis, pemetaan tingkat risiko, dan pencetakan laporan. Proses normalisasi data dan penentuan centroid awal dilakukan untuk meningkatkan akurasi klasifikasi. Hasil pengujian menunjukkan bahwa beberapa kecamatan berada pada zona risiko tinggi, sementara lainnya tergolong sedang hingga rendah. Penerapan metode K-means dalam SIG terbukti efektif untuk menyajikan informasi spasial yang dapat dijadikan dasar dalam mitigasi bencana. Sistem yang dirancang mampu mendukung pengambilan keputusan oleh pemerintah daerah dan instansi terkait secara cepat dan tepat. Penelitian ini menunjukkan bahwa kombinasi SIG dan data mining merupakan pendekatan yang efisien dalam analisis wilayah rawan bencana.

PENDAHULUAN

Di era digital ini, perkembangan teknologi informasi telah membawa perubahan signifikan dalam berbagai bidang, termasuk dalam bidang prediksi bencana. Salah satu teknologi yang berperan penting ialah Sistem Informasi Geografis (SIG), yang memungkinkan pemetaan dan analisis spasial untuk mendukung pengambilan keputusan berbasis data geospasial. SIG ialah alat yang digunakan dalam berbagai aplikasi, mulai dari perencanaan tata ruang hingga memprediksi terjadinya suatu bencana alam, seperti gempa bumi. Dalam konteks bencana, SIG digunakan untuk memantau kondisi lingkungan, mengidentifikasi wilayah rawan gempa bumi, serta memberikan visualisasi yang membantu dalam perencanaan evakuasi bencana. Teknologi ini semakin berkembang dengan integrasi algoritma kecerdasan buatan, seperti clustering, untuk meningkatkan akurasi prediksi terhadap bencana alam yang berbasis pola spasial dan temporal. Metode K-means dapat memungkinkan analisis bencana yang lebih akurat dalam pengelompokkan wilayah yang rawan gempa bumi, sehingga dapat memperbaiki efektivitas perencanaan penanganan dan evakuasi, khususnya di Kabupaten Karo.

Kabupaten Karo terletak di Sumatera Utara yang ialah di dataran tinggi sekitar 66 Km dari Medan dengan jalur menuju Kawasan wisata dan pegunungan seperti Gunung Sibayak dan Gunung Sinabung. Berastagi terkenal sebagai daerah wisata dengan udara yang sejuk, perkebunan hortikultura, serta budaya Karo yang kuat. Kabupaten Karo ialah salah satu wilayah yang tergolong memiliki persentase yang tinggi terhadap bencana gempa bumi. Faktor utama yang menyebabkan tingginya persentase gempa bumi yaitu, aktivitas gunung yang berapi aktif, yang dimana aktivitas magma dalam perut bumi dapat menyebabkan gempa bumi, sehingga daerah ini memiliki kerentanan yang sangat tinggi terhadap gempa bumi. Gempa bumi ialah getaran atau guncangan yang terjadi di permukaan bumi akibat pelepasan energi secara tiba-tiba dari dalam kerak bumi. Pelepasan energi ini biasanya disebabkan oleh pergerakan lempeng tektonik, aktivitas gunung berapi, atau runtuhan batuan di dalam bumi. Bencana alam yang terjadi ini, sering merugikan berbagai aspek, seperti manusia yang mencakup korban jiwa, korban luka, orang hilang bahkan trauma psikologis. Selain itu kerusakan infrastruktur, kerugian ekonomi, dampak lingkungan dan gangguan sosial juga menjadi aspek yang dirugikan oleh bencana alam ini. Gempa bumi ialah getaran di permukaan bumi akibat pelepasan energi mendadak dari dalam kerak bumi, biasanya karena pergeseran lempeng, aktivitas gunung berapi, atau runtuhan batuan. Di Kabupaten Karo, gempa sering terjadi karena adanya gunung berapi aktif seperti Gunung Sinabung. Aktivitas magma di gunung ini

sering memicu gempa, terutama saat erupsi. Selain itu, kegiatan manusia seperti peledakan tambang juga bisa menimbulkan getaran kecil yang memengaruhi patahan aktif. Dengan kondisi geologinya, Karo termasuk wilayah dengan risiko gempa tinggi, namun belum ada sistem pemantauan gempa yang khusus di daerah ini.

Informasi gempa bumi yang disediakan oleh BMKG selama ini masih bersifat umum karena disajikan dalam skala nasional atau regional luas tanpa pemetaan detail untuk wilayah kecil seperti Kabupaten Karo. Data yang diberikan biasanya hanya mencakup waktu, lokasi, magnitudo, dan kedalaman gempa setelah kejadian, tanpa analisis spasial mendalam mengenai zona rawan gempa di tingkat lokal. Selain itu, BMKG belum menyediakan klasifikasi tingkat risiko gempa berdasarkan data historis yang diolah secara khusus untuk setiap daerah, sehingga upaya mitigasi oleh pemerintah daerah maupun masyarakat menjadi kurang optimal. Oleh karena itu, diperlukan penerapan Sistem Informasi Geografis (SIG) berbasis web dengan metode K-means untuk menganalisis dan memetakan potensi gempa bumi secara lebih akurat dan spesifik (Rahmi, Muliawati, & Harbowo, 2024). Metode K-means dapat mengelompokkan wilayah berdasarkan tingkat kerawanan gempa, seperti zona risiko tinggi, sedang, atau rendah, berdasarkan data frekuensi, magnitudo, dan kedalaman gempa di masa lalu (Wiguna, 2024). Dengan adanya sistem ini, pemerintah daerah dapat merancang kebijakan mitigasi yang lebih tepat sasaran, dan masyarakat pun dapat meningkatkan kesiapsiagaan terhadap bencana gempa di wilayahnya. Oleh karena itu, diperlukan penerapan Sistem Informasi Geografis berbasis web dengan metode K-means untuk menganalisis dan memetakan potensi gempa bumi secara lebih, akurat. Dengan adanya sistem ini, diharapkan antisipasi masyarakat serta pemerintah dapat meningkat dalam menaggapi (Nasution & Harahap, 2024).

Penelitian terdahulu telah mengembangkan berbagai pendekatan dalam analisis gempa bumi menggunakan metode analisis spasial, misalnya menerapkan metode Analytical Hierarchy Process (AHP) untuk menentukan Resiko rawan gempa bumi berdasarkan faktor geologi dan hidrologi di Kabupaten Purworejo, Selain itu peneliti sebelumnya memanfaatkan Sistem Informasi Geografis (SIG) untuk memetakan kerawanan gempa bumi dengan metode K-means. K-means ialah kemampuan untuk menempatkan grafis satu peta diatas grafis peta yang lain dan menampilkan hasil nya dilayar komputer atau layar plot (Pitaloka et al., 2018). Salah satu metode yang dapat digunakan untuk analisis potensi gempa bumi ialah metode K-means. Yang dimana metode ini ialah metode dasar yang digunakan untuk melakukan pemetaan menggunakan K-means dari beberapa layer yang telah diberi label sesuai dengan intensitasnya. Metode ini banyak digunakan dalam pemetaan dan telah menunjukkan hasil yang signifikan (Ali & Arnomo, 2024). Hal ini menjadi alasan peneliti dalam pemilihan metode dalam penelitian analisis gempa bumi di Kabupaten Karo dengan metode K-means. Namun, Penelitian yang menggunakan Sistem Informasi Geografis (SIG) berbasis web dengan metode K-means untuk analisis gempa bumi di Kabupaten karo masih sangat terbatas sehingga penelitian ini menjadi relevan untuk dikembangkan lebih lanjut guna menyediakan sistem pemetaan yang lebih akurat dan mudah diakses.

Penelitian ini menggunakan metode K-means dalam Sistem Informasi Geografis (SIG) berbasis web karena kemampuannya dalam menampilkan serta menganalisis data geospasial secara efektif. Algoritma K-means termasuk dalam kategori algoritma partisional karena bekerja dengan cara membagi data ke dalam sejumlah kelompok (cluster) yang telah ditentukan sebelumnya. Proses awal algoritma ini dimulai dengan menentukan jumlah cluster (k) dan menetapkan pusat cluster awal (initial centroids) secara acak atau berdasarkan kriteria tertentu. Selanjutnya, algoritma K-means melakukan proses iteratif untuk mengelompokkan data ke dalam cluster berdasarkan kedekatan terhadap centroid, kemudian memperbarui posisi centroid hingga konvergen. Proses ini berlanjut hingga tidak terjadi perubahan signifikan pada pembentukan cluster, dan hasil akhirnya ialah posisi centroid akhir serta pembagian data ke dalam cluster yang stabil. Pendekatan ini memungkinkan pemetaan tingkat kerawanan gempa bumi dengan mempertimbangkan berbagai faktor lingkungan yang berpengaruh terhadap pergerakan lempeng bumi. Selain itu, SIG berbasis web memberikan akses data yang lebih luas, interaktif, dan real-time, sehingga mendukung pengambilan keputusan yang lebih cepat dan akurat dalam upaya mitigasi bencana. Teknologi ini juga memungkinkan pembaruan data secara lebih efisien serta dapat diakses oleh berbagai pihak, termasuk pemerintah daerah, akademisi, dan masyarakat umum (Barreto, Kelen, & Lestari, 2025).

TINJAUAN LITERATUR

K-Means

Algoritma K-means termasuk dalam kelompok algoritma partisional karena memerlukan penentuan jumlah klaster (k) di awal, serta pemilihan nilai centroid awal. Proses klasterisasi dilakukan secara iteratif, dimana data dibagi ke dalam kelompok berdasarkan kedekatannya dengan centroid, kemudian centroid diperbarui hingga mencapai konvergensi. Output dari algoritma ini ialah posisi akhir centroid dan struktur klaster yang terbentuk (Fatlun, Hilabi, Priyatna, & Hananto, 2025).

Pemilihan awal centroid dilakukan secara acak, dan hal ini memengaruhi jumlah iterasi yang dibutuhkan untuk

mencapai hasil akhir. Oleh karena itu, dalam pengembangan K-means, strategi penentuan centroid berdasarkan kepadatan awal data dapat diterapkan untuk meningkatkan efisiensi algoritma. Hung et al. dalam penelitiannya memodifikasi pendekatan pemilihan centroid guna menyederhanakan proses dan mempercepat kinerja klasterisasi. (Indraputra & Fitriana, 2020).

Adapun cara penerapan algoritma clustering dengan metode K-means adalah sebagai berikut (Dinata, Safwandi, Hasdyna, & Azizah, 2020):

1. Menentukan Jumlah Klaster (k)
   Langkah awal ialah menetapkan nilai k sebagai jumlah klaster yang ingin dibentuk dari data.
2. Inisialisasi Centroid Awal Secara Acak
   Sebanyak k centroid awal dipilih secara acak. Selanjutnya, dilakukan perhitungan jarak antara setiap data terhadap masing-masing centroid menggunakan rumus Euclidean Distance:

$$D_{ij} = \sqrt{\sum_{i=1}^{n} (x_{1-}y_1)^2 + (x_{2-}y_2)^2 + \ldots + (x_{i-}y_i)^2}$$

   Dimana:

   D_{ij} = Jarak Euclidean antara **data ke-i** dengan **centroid *cluster* ke-j**
   x_i = Nilai atribut ke-**i** dari suatu data (misalnya: kemiringan, curah hujan, jenis tanah, dll).
   n = Jumlah atribut atau dimensi yang digunakan (misal: 3 jika ada 3 atribut → kemiringan, curah hujan, jenis tanah).
   $\sum_{i=1}^{n}$ = Menunjukkan bahwa kita menjumlahkan selisih kuadrat antara masing-masing atribut dari data dan centroid-nya.
3. Menentukan Keanggotaan Klaster
   Setiap data akan dikelompokkan ke dalam klaster dengan centroid terdekat berdasarkan jarak terkecil.
4. Memperbarui Centroid
   Centroid setiap klaster diperbarui dengan menghitung rata-rata dari seluruh data yang termasuk dalam klaster tersebut.
5. Iterasi Hingga Konvergen
   Langkah 2 hingga 4 diulang hingga tidak terjadi lagi perubahan keanggotaan data dalam klaster, atau hingga nilai centroid stabil.

Sistem Informasi Geografis

Sistem Informasi Geografis (SIG) ialah suatu sistem berbasis komputer yang dirancang untuk mengelola, memproses, dan menganalisis data spasial serta data atribut yang terkait guna menghasilkan informasi berbasis lokasi. SIG berfungsi sebagai platform integratif yang menghubungkan informasi geografis dengan berbagai jenis data lainnya untuk mendukung proses pengumpulan, penyimpanan, pengelolaan, analisis, hingga visualisasi data yang memiliki referensi keruangan. Melalui pemanfaatan teknologi komputer dan perangkat lunak khusus, SIG mampu menggabungkan berbagai sumber data seperti peta digital, citra satelit, informasi demografis, dan atribut tambahan lainnya dalam satu sistem terpadu yang dapat digunakan untuk kegiatan pemetaan, analisis spasial, serta mendukung pengambilan keputusan secara berbasis lokasi (Barreto, Kelen, & Lestari, 2025).

Secara sederhana, SIG dapat dipahami sebagai suatu teknologi berbasis komputer yang dikembangkan untuk mengelola, memproses, dan memanfaatkan data yang berkaitan dengan lokasi di permukaan bumi. Melalui SIG, pengguna dapat melakukan visualisasi, analisis, serta interpretasi informasi spasial secara lebih efektif dan efisien. Dalam konteks ilmiah, SIG berfungsi sebagai instrumen yang mampu mengintegrasikan data keruangan dengan data atribut, sehingga mempermudah dalam mengidentifikasi pola, hubungan, serta tren yang berkaitan dengan fenomena geografi. Sejalan dengan pendapat Bernhardsen, SIG pada dasarnya merupakan seperangkat sistem yang dirancang untuk mengumpulkan, menyimpan, mengolah, serta menganalisis objek atau fenomena yang memiliki dimensi keruangan, di mana aspek lokasi menjadi komponen utama dalam proses analisis tersebut.

SIG memiliki peran strategis dalam berbagai sektor, seperti mitigasi bencana, perencanaan tata ruang, serta pengelolaan sumber daya alam. Dalam bidang penanggulangan bencana, SIG berfungsi sebagai alat untuk mengidentifikasi dan memetakan wilayah dengan potensi risiko tinggi, seperti daerah rawan banjir, longsor, maupun gempa bumi. Melalui pemanfaatan data spasial yang akurat, SIG memungkinkan penyediaan informasi yang lebih tepat guna, sehingga dapat mendukung pengambilan keputusan yang cepat dan efektif oleh pemerintah maupun masyarakat dalam upaya peningkatan kesiapsiagaan dan respons terhadap bencana (Sutedi, Heryanto, & Hidayat, 2023).

Sig juga memungkinkan analisis spasial dengan menggabungkan berbagai parameter, seperti jarak sungai, curah

hujan, kemiringan lereng, serta penggunaan lahan untuk mengidentifikasi wilayah yang memiliki risiko tinggi terhadap bencana, khususnya banjir. Hasil analisis ini kemudian dapat divisualisasikan dalam bentuk peta tematik yang dapat diakses oleh berbagai pihak terkait, termasuk melalui platform berbasis web untuk mempermudah penyebaran informasi.

Salah satu komponen krusial dalam penerapan SIG ialah adanya alur kerja yang sistematis, di mana setiap tahapan analisis menghasilkan keluaran yang menjadi dasar bagi proses selanjutnya. Meskipun demikian, pelaksanaan secara manual kerap kali memerlukan waktu yang cukup lama dan sumber daya yang besar, terutama ketika terjadi perubahan data input atau diperlukan penyesuaian terhadap parameter analisis. Untuk menjawab tantangan tersebut, diterapkan metode K-Means Clustering yang memungkinkan integrasi berbagai lapisan data spasial, sehingga proses identifikasi wilayah rawan banjir dapat dilakukan secara lebih efisien, otomatis, dan terstruktur (Wahyuni, 2025).

Hasil Penelitian

Pada tahap ini dilakukan analisis terhadap jenis dan sumber data yang diperlukan dalam pembangunan Sistem Informasi Geografis (SIG) berbasis *web* untuk memetakan dan memprediksi potensi gempa bumi di Kabupaten Karo. Data yang dikumpulkan akan digunakan sebagai input dalam proses pengelompokan menggunakan metode K-means *Clustering*, serta divisualisasikan dalam bentuk peta digital interaktif. Adapun data yang dibutuhkan dalam penelitian ini terdiri atas:

1. Data Lokasi
   Data ini mencakup koordinat geografis (lintang dan bujur) dari titiCtitik pengamatan atau wilayah-wilayah di Kabupaten Karo. Lokasi menjadi dasar untuk pemetaan dan visualisasi spasial pada SIG, serta digunakan sebagai variabel utama dalam analisis spasial.
2. Data Kepadatan
   Data ini menunjukkan jumlah penduduk per satuan luas (misalnya jiwa/km²) di setiap wilayah. Kepadatan penduduk menjadi indikator penting dalam menentukan tingkat kerentanan suatu wilayah terhadap dampak gempa, terutama dalam konteks perencanaan mitigasi bencana.
3. Data Jarak Ke Patahan
   Ialah data jarak setiap wilayah atau titik pengamatan terhadap patahan aktif terdekat. Jarak ini berpengaruh signifikan terhadap potensi kerusakan akibat gempa. Semakin dekat suatu wilayah ke patahan aktif, semakin tinggi potensi risikonya.
4. Data Frekuensi
   Data ini mencakup jumlah kejadian gempa dalam rentang waktu tertentu di setiap wilayah. Informasi ini penting untuk mengetahui pola historis kejadian gempa dan digunakan sebagai variabel dalam prediksi dan klasifikasi tingkat kerawanan.
5. Data Resiko Rawan Gempa Bumi
   Data ini ialah klasifikasi atau indeks yang menggambarkan tingkat risiko bencana gempa bumi di suatu wilayah berdasarkan parameter-parameter sebelumnya. Risiko ini biasanya dihitung berdasarkan kombinasi faktor bahaya (hazard), kerentanan (vulnerability), dan kapasitas (capacity). Analisis risiko bencana gempa bumi memerlukan identifikasi sejumlah variabel penting yang dapat memengaruhi tingkat bahaya, kerentanan, dan kapasitas suatu wilayah. Variabel-variabel tersebut digunakan untuk memetakan dan mengukur sejauh mana suatu daerah berisiko terhadap kejadian gempa bumi, baik dari segi intensitas kejadian maupun dampak potensial terhadap populasi. Penjelasan masing-masing variabel dalam Tabel 1 disajikan uraian dibawah ini:
   a. Kepadatan
      Menyatakan jumlah penduduk atau pemukiman per satuan luas di suatu wilayah, biasanya dinyatakan dalam jiwa/km². Kepadatan penduduk berkorelasi langsung dengan tingkat kerentanan (vulnerability) — semakin tinggi kepadatan, maka semakin tinggi potensi korban dan kerusakan akibat gempa.
   b. Jarak ke Patahan (Km)
      Ialah jarak suatu lokasi terhadap sesar (fault line) aktif terdekat, dinyatakan dalam kilometer. Sesar aktif ialah sumber utama gempa bumi. Lokasi yang lebih dekat ke patahan cenderung memiliki risiko lebih tinggi terhadap aktivitas seismik.
   c. Frekuensi
      Menggambarkan jumlah atau intensitas kejadian gempa bumi yang terjadi di wilayah tersebut dalam kurun waktu tertentu (misalnya 10 tahun terakhir). Frekuensi digunakan sebagai indikator tingkat bahaya (hazard) — semakin sering gempa terjadi, semakin tinggi nilai frekuensinya.

Tabel 1 Data Lokasi

No	Nama Lokasi	Latitude	Longitude	Kepadatan	Jarak Ke Patahan	Frekuensi

						/Tahun
1	Kabanjahe	3.100000°	98.490000°	1250 Jiwa	4 Km	4
2	Berastagi	3.190000°	98.500000°	1280 Jiwa	3.8 Km	4
3	Simpang Empat	3.120000°	98.472000°	920 Jiwa	6.5 Km	2
4	Tigapanah	3.210000°	98.454200°	910 Jiwa	6.7 Km	2
5	Merek	3.170000°	98.493200°	900 Jiwa	6.4 Km	2
6	Tiga Binanga	3.070000°	98.371500°	660 Jiwa	9.6 Km	1
7	Barusjahe	3.203000°	98.502100°	1220 Jiwa	4.2 Km	4
8	Lau Baleng	3.267800°	98.335900°	640 Jiwa	9.7 Km	1
9	Mardingding	3.234000°	98.272400°	650 Jiwa	9.4 Km	1
10	Payung	3.156800°	98.535100°	670 Jiwa	9.5 Km	1
11	Juhar	3.243500°	98.419100°	915 Jiwa	6.6 Km	2
12	Dolat Rayat	3.149000°	98.498700°	905 Jiwa	6.3 Km	2
13	Kutabuluh	3.193400°	98.344000°	630 Jiwa	9.6 Km	1

Tabel 2 Data Resiko Gempa Bumi

Resiko	Kepadatan	Jarak Ke Patahan	Frekuensi
Resiko Tinggi	1500	2	7
Resiko Sedang	910	6,5	2
Resiko Rendah	650	9,5	1

Data pada Tabel 1 perlu melalui proses normalisasi sebelum dapat diolah menggunakan algoritma K-means. Tahapan ini mencakup eliminasi data yang tidak relevan serta standarisasi data pengiriman dan alamat. Selanjutnya, penerapan algoritma K-means dilakukan dengan menentukan jumlah *cluster*, misalnya $k = 3$ sebagai langkah awal dalam proses klasterisasi.

Tabel 3 Data Konversi Penilaian

No	Nama Lokasi	Kepadatan	Jarak Ke Patahan	Frekuensi
1	Kabanjahe	1250	4	4
2	Berastagi	1280	3,8	4
3	Simpang Empat	920	6,5	2
4	Tigapanah	910	6,7	2
5	Merek	900	6,4	2
6	Tiga Binanga	660	9,6	1
7	Barusjahe	1220	4,2	4
8	Lau Baleng	640	9,7	1
9	Mardingding	650	9,4	1
10	Payung	670	9,5	1
11	Juhar	915	6,6	2
12	Dolat Rayat	905	6,3	2
13	Kutabuluh	630	9,6	1

Tabel 4 *Cluster*

No	*Cluster*	Keterangan
1	K-1	**Resiko Tinggi**
2	K-2	**Resiko Sedang**
3	K-3	**Resiko Rendah**

Penentuan pusat *cluster* (centroid) dilakukan dengan memilih sejumlah data secara acak dari dataset sebagai inisialisasi awal. Pemilihan acak ini bertujuan untuk memulai proses klasterisasi algoritma K-means dengan mendistribusikan centroid secara representatif terhadap data sumber.

Tabel 5 *Cluster* Resiko Gempa

Centroid	Resiko	Kepadatan	Jarak Ke Patahan	Jarak Lokasi	Keteranga
K-1	**Resiko Tinggi**	1500	2	7	Max
K-2	**Resiko Sedang**	910	6,5	2	Averang
K-3	**Resiko Rendah**	650	9,5	1	Min

Setiap data diklasifikasikan ke dalam *cluster* berdasarkan jarak terdekatnya terhadap centroid. Pengukuran jarak

dilakukan menggunakan rumus Euclidean, sehingga tiap data akan tergabung pada *cluster* dengan pusat yang memiliki jarak minimum.

1. Jarak antara Lokasi Gempa Bumi nomor pertama dengan titik K-1

$$D_{ij} = \sqrt{\sum_{i=1}^{n}(x_{1-}y_1)^2 + (x_{2-}y_2)^2 + \ldots + (x_{i-}y_i)^2}$$

$$K1 = \sqrt{(1250 - 1500)^2 + (4 - 2)^2 + (4 - 7)^2}$$

$K1$= 250,026

2. Jarak antara Lokasi Gempa Bumi nomor pertama dengan titik K-2

$$D_{ij} = \sqrt{\sum_{i=1}^{n}(x_{1-}y_1)^2 + (x_{2-}y_2)^2 + \ldots + (x_{i-}y_i)^2}$$

$$K2 = \sqrt{(1250 - 910)^2 + (4 - 6{,}5)^2 + (4 - 2)^2}$$

$K2$= 340,015

3. Jarak antara Lokasi Gempa Bumi nomor pertama dengan titik K-3

$$D_{ij} = \sqrt{\sum_{i=1}^{n}(x_{1-}y_1)^2 + (x_{2-}y_2)^2 + \ldots + (x_{i-}y_i)^2}$$

$$K3 = \sqrt{(1250 - 650)^2 + (4 - 9{,}5)^2 + (4 - 1)^2}$$

$K3$= 600,033

Untuk lebih lengkapnya jarak pada setiap baris data, hasilnya seperti pada tabel berikut:

Tabel 6 Hasil Perhitungan Iterasi Ke 1

No	K-1	K-2	K-3	Jarak Terdekat	Nilai Terdekat
1	250,026	340,015	600,033	K-1	250,026
2	220,028	370,015	630,033	K-1	220,028
3	580,039	10,000	270,019	K-2	10,000
4	590,040	0,200	260,017	K-2	0,200
5	600,037	10,000	250,021	K-2	10,000
6	840,056	250,021	10,000	K-3	10,000
7	280,025	310,015	570,033	K-1	280,025
8	860,055	270,021	10,002	K-3	10,002
9	850,053	260,018	0,100	K-3	0,100
10	830,056	240,021	20,000	K-3	20,000
11	585,039	5,001	265,018	K-2	5,001
12	595,037	5,004	255,022	K-2	5,004
13	870,054	280,019	20,000	K-3	20,000
Total					840,387

Dari Tabel 6 WCV (Within-*Cluster* Variability) ialah ukuran jumlah variasi atau penyebaran data di dalam setiap klaster, dihitung sebagai total jarak atau deviasi antara setiap data dan pusat klasternya (centroid); nilai WCV yang lebih kecil menunjukkan klaster yang lebih kompak dan homogen. adapun proses pencarian nilai WCV (*Within-Cluster Variability*) berikut :

WCV = (250,026 + 220,028 +... + 20,000)
= 840,387

Dari proses *cluster* Tabel 7 di dapat pemetaan Gempa Bumi berastagi uraian dibawah ini :

Tabel 7 Hasil *Cluster* 1, 2, dan 3 Iterasi-1

No	*Cluster*	Jumlah *Cluster*
1	K-1	3
2	K-2	5
3	K-3	5

Dari Tabel 8 adapun daftar hasil *cluster* anggota K-1-K-3 uraian dibawah ini:

1. K-1 ={ 1, 2, 7}
2. K-2 ={ 3, 4, 5, 11, 12}
3. K-3 ={ 6, 8, 9, 10, 13}

Dari Tabel 7 BCV (Between-*Cluster* Variability) ialah ukuran seberapa jauh pusat (centroid) tiap klaster dari pusat data global, mencerminkan pemisahan antar klaster; semakin besar nilai BCV, semakin baik pemisahan antar klaster. Adapun proses pencarian nilai BCV (*Between-Cluster Variability*) berikut

1. Jarak antara Lokasi Gempa Bumi K-1 dengan titik K-2

$$D_{ij} = \sqrt{\sum_{i=1}^{n}(x_{1-}y_1)^2 + (x_{2-}y_2)^2 + \ldots + (x_{i-}y_i)^2}$$

$$K1, K2 = \sqrt{(1500 - 910)^2 + (2 - 6{,}5)^2 + (7 - 2)^2}$$
$K1, K2 = 590{,}038$

2. Jarak antara Lokasi Gempa Bumi K-1 dengan titik K-3

$$D_{ij} = \sqrt{\sum_{i=1}^{n}(x_1 - y_1)^2 + (x_2 - y_2)^2 + \ldots + (x_i - y_i)^2}$$
$$K1, K3 = \sqrt{(1500 - 650)^2 + (2 - 9{,}5)^2 + (7 - 1)^2}$$
$K1, K3 = 260{,}019$

3. Jarak antara Lokasi Gempa Bumi K-1 dengan titik K-3

$$D_{ij} = \sqrt{\sum_{i=1}^{n}(x_1 - y_1)^2 + (x_2 - y_2)^2 + \ldots + (x_i - y_i)^2}$$
$$K2, K3 = \sqrt{(910 - 650)^2 + (6{,}5 - 9{,}5)^2 + (2 - 1)^2}$$
$K2, K3 = 850{,}054$

BCV = 590,038 + 260,019 + 850,054
= 1700,112

Menghitung kembali nilai rasio dengan membandingkan nilai BCV dan WCV.

BCV/WCV = 1700,112 / 840,387
= 2,023

Kemudian langkah berikutnya pada iterasi ke-2 ialah Menghitung kembali pusat-pusat baru dari *cluster* yang baru terbentuk dengan menghitung nilai rata-rata pada *cluster*.

Diketahui :

K-1 = Resiko Tinggi
K-2 = Resiko Sedang
K-3 = Resiko Rendah

1. Pada *cluster* 1 terdapat 3 data maka nilai rata-rata pada *K*-1 ialah

K-1 = (1250 + 1280 + 1220) / 3 = 1250
K-2 = (4 + 3,8 + 4,2) / 3 = 4
K-3 = (4 + 4 + 4) / 3 = 4

2. Pada *cluster* 2 terdapat 5 data maka nilai rata-rata pada *K*-2 ialah

K-1 = (920 + 910 + 900 + 915 + 905) / 5 = 910
K-2 = (6,5 + 6,7 + 6,4 + 6,6 + 6,3) / 5 = 6,50
K-3 = (2 + 2 + 2 + 2 +2) / 5 = 2

3. Pada *cluster* 3 terdapat 5 data maka nilai rata-rata pada *K*-3 ialah

K-1 = (660 + 640 + 650 + 670 + 630) / 5 = 650
K-2 = (9,6 + 9,7 + 9,4 + 9,5 + 9,6) / 5 = 9,56
K-3 = (1 + 1 + 1 + 1 + 1) / 5 = 1

Tabel 8 Titik Pusat *Cluster* pada Iterasi-2

Titik Pusat Awal	Kemiringan	Jarak Ke Patahan	Jarak Lokasi
Cluster 1	1250	4	4
Cluster 2	910	6,5	2
Cluster 3	650	9,56	1

Apabila data setiap *cluster* belum berhenti, lakukan perulangan dari langkah 2 hingga 5, sampai anggota tiap *cluster* tidak ada yang berubah. Nilai *Centroid* berubah dari nilai *Centroid* sebelumnya, maka algoritma dilanjutkan ke langkah berikutnya. Setelah diketahui nilai k dan pusat *cluster* awal selanjutnya mengukur jarak antara pusat *cluster* menggunakan *Euclidian distance,* kemudian akan didapatkan matriks jarak yaitu K-1, K-2 dan K-3 uraian dibawah ini:.

Tabel 9 Hasil Perhitungan Iterasi Ke 2

No	K-1	K-2	K-3	Jarak Terdekat	Jarak Centroid Terkecil
1	0,000	340,015	600,033	K-1	0,000
2	30,001	370,015	630,033	K-1	30,001
3	330,016	10,000	270,019	K-2	10,000
4	340,017	0,200	260,018	K-2	0,200
5	350,014	10,000	250,022	K-2	10,000
6	590,034	250,021	10,000	K-3	10,000
7	30,001	310,015	570,033	K-1	30,001
8	610,034	270,021	10,001	K-3	10,001
9	600,032	260,018	0,160	K-3	0,160

10	580,034	240,021	20,000	K-3	20,000
11	335,016	5,001	265,018	K-2	5,001
12	345,013	5,004	255,023	K-2	5,004
13	620,033	280,019	20,000	K-3	20,000
Total					150,368

Dari Tabel 9 adapun proses pencarian nilai WCV (*Within-Cluster Variability*) berikut :

WCV = (0,000+ 30,001 + ... + 20,000)
= 150,368

Dari proses *cluster* Tabel 10 di dapat Pemetaan lokasi Gempa Bumi uraian dibawah ini :

Tabel 10 Hasil *Cluster* 1, 2, dan 3 Iterasi-2

No	*Cluster*	Jumlah *Cluster*
1	K-1	3
2	K-2	5
3	K-3	5

Dari Tabel 11 adapun daftar hasil *cluster* anggota K-1-K-3 uraian dibawah ini

1. K-1 = {1, 2, 7}
2. K-2 = {3, 4, 5, 11, 12}
3. K-3 = {6, 8, 9, 10, 13}

Setelah dilakukan perbandingan antara hasil iterasi ke-2 dan ke-3, diketahui bahwa hasil pembagian anggota *cluster* tidak mengalami perubahan. Dengan demikian, dapat disimpulkan bahwa proses iterasi telah mencapai konvergensi, sehingga hasil *clustering* dinyatakan stabil dan tidak memerlukan iterasi lanjutan

PEMBAHASAN

Hasil pengelompokkan *cluster* dari Resiko pemetaan Gempa Bumi daerah berastagi diambil dari iterasi 2 dikarenakan rasio iterasi 3 sampa ke 4 hasilnya sama, maka adapun hasil pengelompokkan hasil *cluster* ialah uraian dibawah ini:

Tabel 11 Pengelompokan Hasil *Cluster* dari iterasi 3

Cluster	Nomor	Nama Lokasi
Risiko Tinggi	1, 2, 7	Kabanjahe, Berastagi, Barusjahe
Risiko Sedang	3, 4, 5, 11, 12	Simpang Empat, Tigapanah, Merek, Juhar, Dolat Rayat
Risiko Rendah	6, 8, 9, 10, 13	Tiga Binanga, Lau Baleng, Mardingding, Payung, Kutabuluh

Berdasarkan hasil pengelompokan menggunakan metode K-means *Clustering*, diketahui bahwa wilayah dengan risiko tinggi gempa bumi meliputi Kabanjahe, Berastagi, dan Barusjahe yang memiliki kepadatan penduduk tinggi, frekuensi gempa yang sering, serta jarak yang relatif dekat dengan patahan aktif. Wilayah dengan risiko sedang seperti Simpang Empat, Tigapanah, Merek, Juhar, dan Dolat Rayat menunjukkan karakteristik kepadatan dan jarak ke patahan yang sedang, serta frekuensi gempa yang tidak terlalu tinggi. Sementara itu, wilayah dengan risiko rendah seperti Tiga Binanga, Lau Baleng, Mardingding, Payung, dan Kutabuluh memiliki kepadatan yang rendah, frekuensi gempa yang jarang, serta jarak yang cukup jauh dari patahan aktif. Pengelompokan ini penting dalam mendukung upaya mitigasi bencana dan penentuan prioritas penanganan risiko gempa bumi di Kabupaten Karo.

Tahapan ini menunjukkan implementasi antarmuka aplikasi setelah melalui proses analisis dan perancangan. Antarmuka berfungsi sebagai media input-output yang membantu pengguna dalam mengoperasikan sistem klasifikasi lokasi gempa berbasis metode K-Means. Aplikasi terdiri atas beberapa komponen utama: Login, Menu Utama, dan Halaman Administrator.

4.1.1 Menu Utama

Dalam Menu Utama untuk menampilkan pada tampilan *menu* pada awal sistem yaitu *menu login* dan *menu* utama. Adapun *menu* Menu Utama uraian dibawah ini.

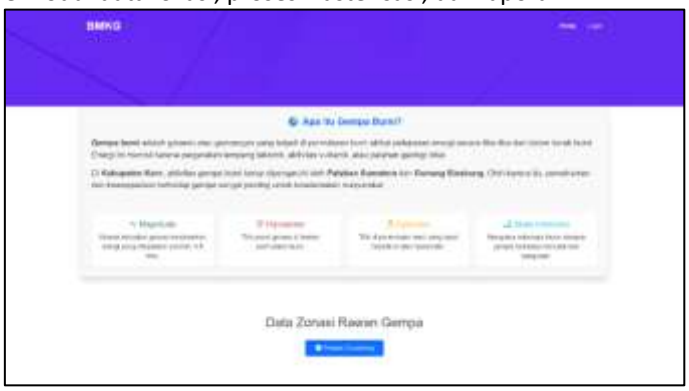
1. *Menu Login*

   Berfungsi sebagai sistem otentikasi pengguna untuk mencegah akses ilegal terhadap sistem.



Gambar 1 *Menu Login*

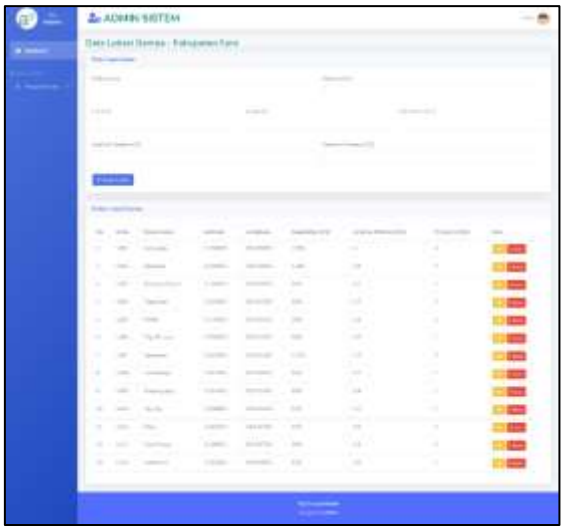
2. *Menu* Utama
   Menyediakan navigasi ke modul data lokasi, proses klasterisasi, dan laporan.



Gambar 2 *Menu* Utama

Dalam adminstrator untuk menampilkan *menu* pengolahan data pada penyimpanan data kedalam *database* yaitu *menu* Lokasi Gempa Bumi dan proses K-means Adapun *menu* halaman adminstrator utama uraian dibawah ini.

1. *Menu* Data Lokasi Gempa Bumi
   Menyediakan fasilitas manajemen data wilayah terdampak gempa di Kabupaten Karo.



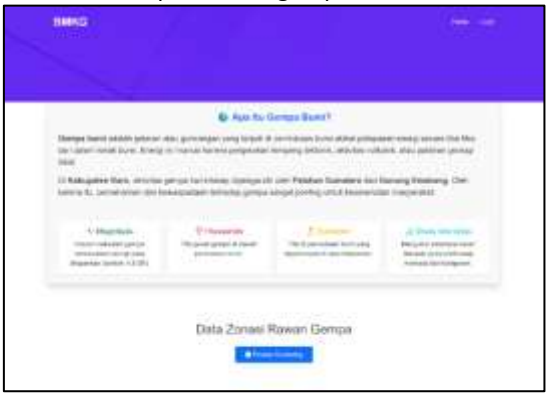
Gambar 3 *Menu* Data Lokasi Gempa Bumi

2. *Menu* Data Centroid
   Menyajikan titik pusat klaster untuk proses K-Means. Adapun *menu* data centroid ialah uraian dibawah ini.



Gambar 4 *Menu* Data Centroid

3. *Menu* Proses K-means
   Melakukan klasterisasi wilayah berdasarkan parameter gempa.

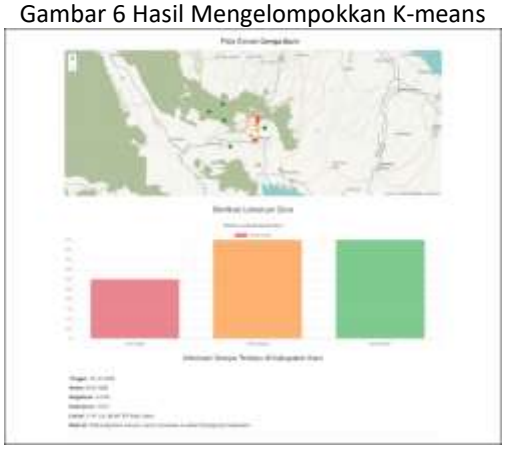


Gambar 5 *Menu* Proses K-means

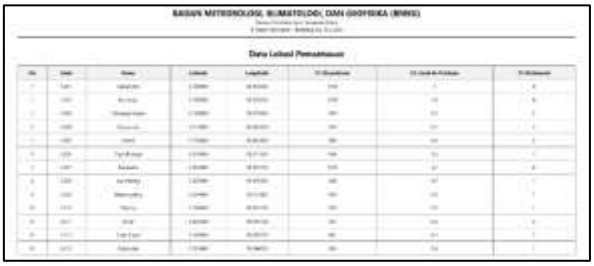
Pengujian dilakukan menggunakan data baru untuk menguji akurasi sistem klasifikasi. Hasil visualisasi pemetaan menunjukkan klaster wilayah rawan gempa melalui metode K-Means serta representasi geografis dalam bentuk peta digital berbasis GIS. Fitur cetak laporan juga tersedia untuk dokumentasi hasil.



Gambar 6 Hasil Mengelompokkan K-means



Gambar 7 Hasil Pemetan GIS Wilayah Gempa Bumi Kabupaten karo



Gambar 8 Cetak Laporan

KESIMPULAN

Berdasarkan hasil analisis dan implementasi Sistem Informasi Geografis (SIG) berbasis web yang dirancang untuk memetakan tingkat kerentanan gempa bumi di Kabupaten Karo menggunakan algoritma K-Means Clustering, dapat disimpulkan bahwa metode ini efektif dalam mengelompokkan wilayah ke dalam kategori risiko tinggi, sedang, dan rendah. Dengan mempertimbangkan parameter kepadatan penduduk, kedekatan terhadap patahan aktif, serta frekuensi kejadian seismik, sistem mampu memberikan klasifikasi spasial yang akurat sehingga dapat digunakan sebagai dasar pertimbangan dalam penyusunan strategi mitigasi bencana.

Selain itu, SIG berbasis web yang dikembangkan dengan dukungan teknologi PHP dan MySQL terbukti mampu menyajikan informasi spasial secara interaktif melalui peta digital. Sistem ini tidak hanya menampilkan hasil clustering, tetapi juga menyediakan fitur untuk input data lokasi, pemrosesan otomatis, visualisasi tingkat kerentanan, hingga pencetakan laporan. Kehadiran fitur-fitur tersebut menjadikan sistem ini tidak hanya bersifat analitis, tetapi juga aplikatif dalam mendukung proses pengambilan keputusan.

Secara keseluruhan, penerapan SIG berbasis web dengan algoritma K-Means Clustering dapat menjadi solusi inovatif dalam upaya mitigasi dan perencanaan kebencanaan di Kabupaten Karo. Antarmuka yang sederhana dan mudah digunakan memungkinkan sistem ini diakses oleh berbagai pihak, mulai dari pemerintah daerah hingga masyarakat umum. Dengan demikian, sistem ini diharapkan dapat menjadi instrumen pendukung dalam meningkatkan kesadaran, kesiapsiagaan, serta efektivitas penanganan risiko bencana gempa bumi.

REFERENSI

Rahmi, P. A., Muliawati, T., & Harbowo, D. G. (2024). Analisis Pola Aktivitas Gempa Bumi di Pulau Sumatera dengan Metode K-Means Clustering dan Rantai Markov. *Seminar Nasional Sains Data 2024 (SENADA 2024)* , 592-601.

Wiguna, W. (2024). Pengembangan Sistem Informasi Geografis Pariwisata Bandung Barat Menggunakan Metode Agile Berbasis Website. *Jurnal Nasional Komputasi dan Teknologi Informasi (JNKTI)*, 2337-2347.

Ali, P. A., & Arnomo, S. A. (2024). Clustering Dengan Algoritma K-Means Untuk Meningkatkan Penjualan Pada Pt Pulau Bintang Emas Batam. *JURNAL COMASIE*, 98-104.

Barreto, J., Kelen, Y. P., & Lestari, A. K. (2025). Implementasi Sistem Informasi Geografis Pemetaan Lahan Pertanian Dan Komoditas Pertanian Di Kabupaten Belu Berbasis Dashboard. *Jurnal Sistem Informasi*, 293-303.

Dinata, R. K., Safwandi, Hasdyna, N., & Azizah, N. (2020). Analisis K-Means Clustering pada Data Sepeda Motor. *Informatics Journal*, *V*(1).

Fatlun, A., Hilabi, S. S., Priyatna, B., & Hananto, A. L. (2025). Penggunaan Algoritma K-Means Clustering CRISP-DM Dalam Mengelompokkan Drama Korea Sebagai Rekomendasi Film. *Jurnal Teknik Informatika dan Sistem Informasi*, 130-140.

Indraputra, R. A., & Fitriana, R. (2020). K-Means Clustering Data COVID-19. *Jurnal Teknik Industri*, *X*(2622-5131), 275-282.

Nasution, A. B., & Harahap, N. M. (2024). Sistem Informasi Geografis pada Persebaran Pembangunan Madrasah di Provinsi Sumatera Utara berbasis Web. *Remik: Riset dan E-Jurnal Manajemen Informatika Komputer*, 153-164.

Sutedi, A., Heryanto, D., & Hidayat, M. (2023). Rancang Bangun Sistem Informasi Geografis Pariwisata Berbasis Web Studi Kasus Kabupaten Bandung. *Jurnal Algoritma*, 328-339.

Wahyuni, R. (2025). Implementasi Sistem Informasi Geografis (SIG) untuk Pemetaan dan Pengelolaan Sumber Daya Alam di Kawasan Pertanian. *Journal of Human And Education*, 545-550.