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

Sistem Informasi Geografis Pemetaan Berbasis Web Untuk Prediksi Longsor Di Kecamatan Brastagi Dengan Metode K-Means Clustering

Mages Pren[1)*], Megaria Purba[2)]
[1]Sekolah Tinggi Manajemen Informatika Dan Komputer Pelita Nusantara
[2]Sekolah Tinggi Manajemen Informatika Dan Komputer Pelita Nusantara
[1]magespren239@gmail.com, [2]megariapurba18@gmail.com

Received: 23 August 2025 Accepted: 25 Juni 2026 Published: 10 July 2026
Pp. 278-289

***Penulis Koresponden**
Kata Kunci:
Longsor,
Sistem Informasi Geografis,
K-Means Clustering,
Web,
GIS,
Pemetaan Kerawanan

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

Kecamatan Brastagi merupakan wilayah yang memiliki potensi tinggi terhadap bencana tanah longsor akibat kondisi geografisnya yang berbukit serta intensitas curah hujan yang tinggi. Dalam upaya mitigasi dan manajemen risiko bencana, pemetaan tingkat kerawanan longsor menjadi hal yang krusial. Penelitian ini mengembangkan sistem informasi geografis (SIG) berbasis web untuk mengidentifikasi zona kerawanan longsor dengan memanfaatkan algoritma K-Means Clustering sebagai metode klasifikasi spasial. Proses klasterisasi dilakukan dengan mengelompokkan data berdasarkan tiga parameter utama, yaitu kemiringan lereng, curah hujan, dan jenis tanah. Melalui proses iteratif yang mencakup penentuan centroid awal, perhitungan jarak Euclidean, dan pembaruan centroid hingga tercapai konvergensi, sistem mampu mengelompokkan wilayah ke dalam tiga zona kerawanan: tinggi (zona merah), sedang (zona kuning), dan rendah (zona hijau). Implementasi sistem dilakukan menggunakan bahasa pemrograman PHP dan database MySQL, serta ditampilkan dalam bentuk peta interaktif untuk memberikan representasi visual kondisi spasial wilayah. Hasil pengujian menunjukkan bahwa sistem SIG yang dibangun mampu menyajikan informasi kerawanan longsor secara cepat, akurat, dan responsif. Selain itu, sistem juga menyediakan fitur-fitur pendukung seperti input data lokasi, visualisasi hasil clustering, dan cetak laporan. Sistem ini diharapkan dapat digunakan oleh instansi pemerintah maupun masyarakat sebagai alat bantu dalam perencanaan tata ruang, evakuasi bencana, dan peningkatan kesadaran terhadap risiko longsor.

PENDAHULUAN

Di era digital saat ini, teknologi informasi berperan penting dalam berbagai aspek kehidupan, termasuk dalam penanggulangan bencana alam. Salah satu teknologi yang semakin berkembang adalah Sistem Informasi Geografis (SIG), yang mampu memvisualisasikan data spasial untuk mendukung pengambilan keputusan berbasis lokasi. SIG merupakan alat yang sangat berguna dalam berbagai aplikasi, termasuk dalam analisis bencana alam, seperti longsor. Dalam hal ini, SIG digunakan untuk memantau kondisi lingkungan, mengidentifikasi wilayah yang rawan longsor, serta memberikan visualisasi yang membantu perencanaan evakuasi bencana. Seiring dengan perkembangan teknologi, SIG terus mengalami kemajuan baik dalam hal fungsionalitas maupun kemudahan penggunaan. SIG juga berkembang untuk meningkatkan akurasi dalam memprediksi potensi longsor berdasarkan pola spasial dan temporal. Metode K-Means Clustering memungkinkan pengelompokan wilayah berdasarkan karakteristik spasial secara otomatis dan objektif, sehingga dapat meningkatkan akurasi dalam prediksi daerah rawan longsor. Penerapan metode ini di Kecamatan Brastagi diharapkan dapat memperkuat efektivitas perencanaan mitigasi bencana dan strategi evakuasi yang lebih tepat sasaran.

Kecamatan Brastagi, yang terletak di Kabupaten Karo, Sumatera Utara, merupakan salah satu wilayah dengan tingkat kerawanan tinggi terhadap bencana tanah longsor. Yang menjadi masalah utama yang menyebabkan tingginya longsor di wilayah ini adalah topografi yang berbukit, curah hujan yang tinggi, serta jenis tanah vulkanik yang gembur. Selain itu, peningkatan aktivitas pembangunan, seperti perubahan tata guna lahan yang tidak terkontrol dan penggundulan hutan, turut memperburuk kondisi stabilitas tanah di daerah tersebut. Ketidakseimbangan antara eksploitasi sumber daya alam dan upaya konservasi tanah telah menyebabkan meningkatnya kejadian longsor, yang tidak hanya berdampak pada infrastruktur tetapi juga terhadap kehidupan masyarakat setempat.

Kecamatan Brastagi, Kabupaten Karo, Sumatera Utara, merupakan wilayah dengan tingkat kerentanan tinggi terhadap longsor akibat curah hujan yang tinggi, kemiringan lereng yang berbukit, serta kondisi tanah yang labil. Selain faktor alam, perubahan penggunaan lahan yang tidak terkendali semakin meningkatkan risiko bencana ini, mengakibatkan kerusakan properti, infrastruktur, serta gangguan terhadap aktivitas ekonomi dan transportasi. Hingga saat ini, belum

tersedia sistem yang secara khusus memantau potensi longsor di Kecamatan Brastagi. Informasi yang disediakan oleh BMKG masih bersifat umum, sehingga upaya mitigasi belum dapat dilakukan secara optimal. Oleh karena itu, diperlukan penerapan Sistem Informasi Geografis (SIG) berbasis web dengan metode K-Means Clustering untuk menganalisis dan memetakan potensi longsor secara lebih akurat dan terklasifikasi berdasarkan pola karakteristik wilayah.

Dengan adanya sistem ini, diharapkan kesiapsiagaan masyarakat dan pemerintah dalam menghadapi bencana longsor dapat meningkat. Beberapa penelitian terdahulu telah mengembangkan berbagai pendekatan dalam analisis longsor menggunakan teknologi SIG dan metode analisis spasial Misalnya, menerapkan Metode K-Means Clustering, K-means clustering adalah metode pengelompokan data yang membagi data menjadi beberapa kelompok berdasarkan kedekatannya dengan titik pusat atau centroid (Ali & Arnomo, 2024). Keunggulan metode ini adalah cepat dan efisien, sehingga cocok untuk data set besar, seperti data kecamatan di Kabupaten Grobogan. K-means juga memungkinkan kita untuk menentukan jumlah cluster terlebih dahulu, yang membantu mengelompokkan kecamatan berdasarkan tingkat kerentanannya terhadap bencana alam, seperti banjir atau gempa (Rahmi, Muliawati, & Harbowo, 2024).

Salah satu metode yang dapat digunakan untuk prediksi potensi longsor adalah metode K-Means Clustering. Metode ini merupakan metode dasar dalam analisis pengelompokan yang digunakan untuk memetakan wilayah berdasarkan kemiripan karakteristik data spasial. K-Means Clustering telah banyak digunakan dalam pemetaan zona kerentanan dan terbukti mampu menghasilkan klasifikasi wilayah yang akurat, terutama pada daerah dengan kondisi heterogen seperti Kecamatan Brastagi. Hal ini menjadi alasan peneliti dalam pemilihan metode pada penelitian prediksi longsor di kecamatan brastagi dengan metode K-Means Clustering.

Dari latar belakang yang telah dijelaskan sebelumnya, penelitian ini bertujuan untuk membantu pemerintah daerah Kecamatan Brastagi dalam memetakan wilayah rawan longsor menggunakan algoritma clustering K-Means serta memvisualisasikan hasilnya dalam bentuk peta melalui Sistem Informasi Geografis berbasis web. Dengan data spasial yang lebih akurat mengenai potensi bencana longsor, pihak terkait dapat lebih efektif dalam menetapkan skala prioritas untuk upaya mitigasi dan penanggulangan bencana di wilayah yang berisiko . Namun, Penelitian yang menggunakan Sistem Informasi Geografis (SIG) berbasis web dengan metode K-Means Clustering untuk prediksi longsor di Kecamatan Brastagi masih sangat terbatas sehingga penelitian ini menjadi relevan untuk dikembangkan lebih lanjut guna menyediakan sistem pemetaan yang lebih akurat dan mudah diakses. Sistem Informasi Geografis (SIG) merupakan teknologi berbasis komputer yang dirancang untuk mengintegrasikan proses pengumpulan, penyimpanan, pengolahan, analisis, hingga penyajian data serta informasi yang berkaitan dengan lokasi atau posisi geografis suatu objek atau fenomena di permukaan bumi (Nasution & Harahap, 2024). Dalam studi ini, SIG dibuat sebagai aplikasi di web yang menggunakan metode K-Means Clustering karena efektif dalam mengelompokkan data geospasial secara otomatis berdasarkan kesamaan nilai, seperti sudut lereng, jenis tanah, dan jumlah curah hujan.

Pendekatan ini memungkinkan pemetaan potensi longsor yang lebih objektif dengan membagiwilayah ke dalam beberapa kelompok kerawanan. Selain itu, penerapan SIG berbasis web mendukung visualisasi data secara interaktif, aksesibilitas yang lebih luas, serta pembaruan informasi secara real-time. Hal ini mempermudah pemangku kepentingan, seperti pemerintah daerah, akademisi, hingga masyarakat umum, dalam mengambil keputusan yang tepat untuk mitigasi dan penanggulangan bencana longsor.

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.

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).

Langkah- langkah melakukan 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 sebuah teknologi yang mengintegrasikan elemen geografis dengan metode analisis data, sehingga dapat memberikan pemahaman yang lebih mendalam mengenai berbagai fenomena. SIG tidak hanya berfokus pada pembuatan peta, tetapi juga memberikan kesempatan bagi pengguna untuk menganalisis data ruang dan mengidentifikasi informasi penting yang terkandung dalam data tersebut (Wiguna, 2024).

Sistem Informasi Geografis (SIG) ialah sistem yang berbasis komputer dan digunakan untuk menyimpan serta memproses data-data geografis. SIG dirancang untuk mengumpulkan, menyimpan, serta menganalisis objek dan fenomena. Salah satu aspek utama yang akan dianalisis ialah lokasi geografis (Sutedi, Heryanto, & Hidayat, 2023). Sistem Informasi Geografis (SIG) memberikan kesempatan kepada pengguna untuk memilih berbagai pilihan saat membahas distribusi geografis yang berkaitan dengan isu-isu kesehatan. Basis data Sistem Informasi Geografis (SIG) terhubung dengan teknik analisis spasial serta pengelolaan dan pengolahan data yang dapat mengidentifikasi hubungan antara persebaran penyakit di suatu tempat dengan keadaan lingkungan di area tersebut.



Sumber:(Utomo, 2010)

Gambar 1 Siklus Pengembangan SIG Berbasis Web

Secara umum, sistem informasi geografis terdiri dari empat langkah, uraian dibawah ini:

1. Tahap Akuisisi Data
   Tahapan awal ini mencakup proses perolehan dan pencatatan data spasial beserta atributnya, yang menggambarkan posisi geografis, bentuk, ukuran, serta karakteristik suatu objek di permukaan bumi. Proses ini melibatkan aktivitas perencanaan, penentuan tujuan, pengumpulan informasi lapangan, digitalisasi peta, hingga memasukkan data ke dalam sistem komputer.
2. Tahap Pengolahan Awal Data
   Pada tahap ini, data yang telah dikumpulkan akan melalui proses pemrosesan dasar seperti klasifikasi, stratifikasi, penggabungan data, serta analisis geospasial awal seperti pemotongan (clipping), penggabungan (merge), atau penyederhanaan bentuk (dissolve). Kegiatan ini bertujuan untuk menyiapkan data agar siap dianalisis lebih lanjut.
3. Tahap Analisis dan Transformasi Data
   Tahap ini ialah inti dari pemanfaatan SIG, dimana data spasial dianalisis untuk menghasilkan informasi yang bermanfaat. Beberapa metode yang dapat digunakan antara lain:
   a. Pemberian Bobot (Scoring): Proses penilaian terhadap parameter tertentu berdasarkan kriteria yang telah ditentukan.
   b. *Clustering* K-means: Teknik pengelompokan data spasial ke dalam beberapa *cluster* berdasarkan kemiripan nilai, guna mengidentifikasi pola sebaran dan menghasilkan klasifikasi wilayah baru.

c. Buffering: Pembuatan zona pengaruh (buffer) di sekitar objek titik, garis, atau area berdasarkan jarak tertentu, digunakan untuk analisis kedekatan spasial.

4. Tahap Penyajian Data (Output)
Ialah tahap akhir dalam alur SIG, dimana hasil analisis disajikan dalam bentuk visual seperti peta tematik, grafik, atau laporan digital. Informasi spasial yang dihasilkan dapat digunakan sebagai dasar pengambilan keputusan dalam berbagai bidang, termasuk mitigasi bencana, perencanaan wilayah, dan pengelolaan sumber daya.

Hasil Penelitian

Pada tahap ini, dilakukan proses analisis terhadap jenis dan sumber data yang diperlukan dalam pengembangan Sistem Informasi Geografis (SIG) berbasis web untuk memprediksi daerah rawan longsor di wilayah Kecamatan Brastagi. Data yang diperoleh akan digunakan sebagai masukan utama dalam proses pengelompokan menggunakan algoritma K-Means Clustering, sekaligus untuk keperluan visualisasi dalam bentuk peta digital interaktif. Adapun jenis data yang dibutuhkan dalam penelitian ini meliputi:

1. Data Lokasi
Data ini mencakup nama lokasi atau desa di Kecamatan Brastagi yang berpotensi mengalami longsor, serta titik koordinat dalam bentuk latitude dan longitude. Data ini digunakan untuk menampilkan lokasi secara spasial di peta digital.
2. Data Kemiringan Lereng
Kemiringan lereng diukur dalam satuan derajat (°) dan menjadi salah satu parameter utama dalam menentukan tingkat kerawanan longsor. Semakin besar nilai kemiringan, maka potensi longsor akan semakin tinggi.
3. Data Curah Hujan
Curah hujan dinyatakan dalam satuan milimeter (mm) dan diperoleh berdasarkan data klimatologi dari wilayah setempat. Curah hujan tinggi menjadi faktor pemicu utama terjadinya longsor, khususnya di daerah lereng dan perbukitan.
4. Data Jenis Tanah
Jenis tanah diklasifikasikan berdasarkan karakteristik geologis wilayah Brastagi, seperti Andosol, Regosol, Grumosol, Litosol, dan sebagainya. Data ini kemudian dikonversi menjadi numerik agar dapat digunakan dalam proses clustering.
5. Data Zona Rawan Longsor (hasil)
Merupakan hasil akhir dari proses pengelompokan (clustering) yang mengelompokkan setiap lokasi ke dalam zona rawan longsor, yaitu zona merah (sangat rawan), zona kuning (sedang), dan zona hijau (rendah). Data ini akan ditampilkan dalam peta untuk memberikan informasi visual kepada pengguna.

Tabel 1 Data Lokasi

No	Nama Lokasi	Latitude	Longitude	Kemiringan	Curah Hujan	Jenis Tanah
1	Desa Doulu	3.222.000	98.531.000	32	250	Andosol
2	Sempajaya	3.200.500	98.515.500	25	220	Regosol
3	Sikulikap	3.198.200	98.512.093	37	280	Litosol
4	Bukit Kubu	3.195.000	98.510.000	28	200	Podsolik Merah Kuning
5	Gundaling I	3.194.500	98.505.000	30	210	Andosol
6	Gundaling II	3.189.000	98.500.000	34	260	Regosol
7	Rumah Kaban	3.208.500	98.511.500	26	230	Aluvial
8	Rih Tengah	3.203.500	98.513.500	29	240	Grumosol
9	Peceren	3.197.500	98.515.000	33	270	Andosol
10	Jaranguda	3.191.500	98.495.000	27	210	Litosol
11	Kandibata	3.209.000	98.504.000	36	290	Podsolik Merah Kuning
12	Tambak Lau Mulgap I	3.218.000	98.522.000	31	260	Aluvial
13	Tambak Lau Mulgap II	3.220.000	98.528.000	38	300	Andosol
14	Desa Raya	3.215.000	98.520.000	30	225	Grumosol
15	Sembahe Km 36.2	3.194.500	98.507.500	35	295	Litosol

Tabel 2 Kategori Jenis Tanah

No	Jenis Tanah	Nilai
1	Litosol	0
2	Regosol	1
3	Grumosol	2
4	Andosol	3

5	Aluvial	4
6	Podsolik Merah Kuning	5

Tabel ini menyajikan jenis-jenis tanah yang diklasifikasikan ke dalam kategori numerik. Tujuan pengkategorian ini adalah untuk mempermudah proses analisis kuantitatif, khususnya dalam sistem berbasis komputer atau perhitungan statistik. Berikut uraian singkat tiap jenis tanah:

1. Litosol (Nilai: 0)
   Tanah berbatu dengan lapisan tanah sangat tipis, sering ditemukan di daerah perbukitan atau pegunungan. Tidak cocok untuk pertanian intensif.
2. Regosol (Nilai: 1)
   Tanah muda yang terbentuk dari material vulkanik atau pasir, memiliki tingkat kesuburan sedang dan mudah diolah.
3. Grumosol (Nilai: 2)
   Tanah berat yang memiliki kandungan lempung tinggi, subur, tetapi bisa keras saat kering dan lengket saat basah. Cocok untuk tanaman tertentu.
4. Andosol (Nilai: 3)
   Tanah yang berasal dari abu vulkanik, sangat subur, gembur, dan baik untuk pertanian. Banyak ditemukan di daerah vulkanik.
5. Aluvial (Nilai: 4)
   Tanah hasil endapan sungai yang subur dan cocok untuk berbagai jenis pertanian, terutama padi.
6. Podsolik Merah Kuning (Nilai: 5)
   Tanah yang umum di wilayah tropis basah, memiliki kandungan besi dan aluminium tinggi. Kesuburan sedang hingga rendah, tergantung pengelolaan.

Tabel 3 Data Zona Longsor

No	Zona
1	Zona Merah
2	Zona Kuning
3	Zona Hijau

Penentuan titik pusat awal (*centroid*) untuk setiap cluster dilakukan dengan cara memilih sejumlah data secara acak dari himpunan data yang tersedia. Pemilihan secara acak ini bertujuan untuk memberikan variasi awal dalam proses pengelompokan, di mana setiap centroid mewakili posisi awal dari masing-masing cluster. Nilai-nilai centroid tersebut kemudian digunakan sebagai acuan dalam tahap iterasi selanjutnya pada algoritma K-Means Clustering.

Perhitungan rata-rata:

Kemiringan : (32 + 25 + 37 + 28 + 30 + 34 + 26 + 29 + 33 + 27 + 36 + 3 1 + 38 + 30 + 35) / 15 = 30.93

Curah Hujan: Total = 3710 / 15 = 247.33

Jenis Tanah: Total = 38 / 15 = 2.53

Tabel 6 Cluster Zona

Centroid	Zona	Kemiringan	Curah Hujan	Jenis Tanah	Keterangan
K-1	Zona Merah	38	300	3	Max
K-2	Zona Kuning	30,93	247,33	2,53	Average
K-3	Zona Hijau	25	220	1	Min

Setelah nilai awal centroid ditentukan, langkah selanjutnya adalah mengelompokkan setiap data ke dalam cluster berdasarkan tingkat kedekatannya dengan centroid tersebut. Pengukuran kedekatan dilakukan menggunakan rumus Euclidean Distance, yang menghitung jarak lurus antara titik data dengan masing-masing centroid. Setiap data akan dimasukkan ke dalam cluster yang memiliki jarak terpendek dari titik pusat (centroid) tersebut. Proses ini bertujuan untuk membentuk pengelompokan awal berdasarkan kemiripan nilai.

1. Jarak antara Lokasi nomor pertama dengan titik K-1
   $D_{ij} = \sqrt{\sum_{i=1}^{n}(x_1 - y_1)^2 + (x_2 - y_2)^2 + \ldots + (x_i - y_i)^2}$
   $K1 = \sqrt{(32-38)^2 + (250-300)^2 + (3-3)^2}$
   K1= 50,359
2. Jarak antara Lokasi nomor pertama dengan titik K-2
   $D_{ij} = \sqrt{\sum_{i=1}^{n}(x_1 - y_1)^2 + (x_2 - y_2)^2 + \ldots + (x_i - y_i)^2}$
   $K2 = \sqrt{(32-30{,}93)^2 + (250-247{,}33)^2 + (3-2{,}53)^2}$
   $K2$= 2,915

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

$D_{ij} = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2}$

$K3 = \sqrt{(32-25)^2 + (250-220)^2 + (3-1)^2}$

$K3$= 30,871

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

Tabel 7 Hasil Perhitungan Iterasi Ke 1

No	K-1	K-2	K-3	Jarak Terdekat	Nilai Terdekat
1	50,359	2,915	30,871	K-2	2,915
2	81,074	28,008	0,000	K-3	0,000
3	20,248	33,325	61,196	K-1	20,248
4	100,519	47,485	20,616	K-3	20,616
5	90,355	37,345	11,358	K-3	11,358
6	40,249	13,126	41,000	K-2	13,126
7	71,028	18,077	10,488	K-3	10,488
8	60,679	7,598	20,421	K-2	7,598
9	30,414	22,769	50,675	K-2	22,769
10	90,719	37,621	10,247	K-3	10,247
11	10,392	43,041	70,972	K-1	10,392
12	40,620	12,755	40,559	K-2	12,755
13	0,000	53,144	81,074	K-1	0,000
14	75,432	22,356	7,141	K-3	7,141
15	6,557	47,910	75,670	K-1	6,557
Total					**156,211**

Within-Cluster Variability (WCV) adalah ukuran yang menunjukkan seberapa rapat data dalam suatu cluster terhadap pusatnya (centroid). Nilai WCV dihitung dari total jarak kuadrat antara tiap data dan centroid dalam cluster. Semakin kecil WCV, semakin baik hasil pengelompokan karena data dalam cluster lebih homogen.

WCV = (2,915 + 0 + 20,248 + 20,616 + 11,358 + 13,126 + 10,488 + 7,598 + 22,769 + 10,247 + 10,392 + 12,755 + 0 + 7,141 + 6,557)
= 156,211

Dari proses *cluster* Tabel 8 di dapat pemetaan longsor berastagi sebagai berikut :

Tabel 8 Hasil Cluster 1, 2, dan 3 Iterasi-1

No	Cluster	Jumlah Cluster
1	K-1	4
2	K-2	5
3	K-3	6

Dari Tabel 8 adapun daftar hasil *cluster* anggota K1-K3 sebagai berikut

1. K-1 ={ 3, 11, 13, 15}
2. K-2 ={ 1, 6, 8, 9, 12}
3. K-3 ={ 2, 4, 5, 7, 10, 14}

Between-Cluster Variability (BCV) adalah ukuran yang menunjukkan seberapa jauh posisi antar centroid satu dengan lainnya, atau dengan kata lain, mengukur variasi antar cluster. Nilai BCV dihitung dari jarak kuadrat antara masing-masing centroid terhadap centroid keseluruhan (global mean). Semakin besar nilai BCV, maka semakin baik pemisahan antar cluster, karena menunjukkan bahwa masing-masing cluster cukup berbeda satu sama lain.

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

$D_{ij} = \sqrt{\sum_{i=1}^{n}(x_1 - y_1)^2 + (x_2 - y_2)^2 + \ldots + (x_i - y_i)^2}$

$K1 = \sqrt{(38-30{,}93)^2 + (300-247{,}33)^2 + (3-2{,}53)^2}$

$K1$= 53,144

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

$D_{ij} = \sqrt{\sum_{i=1}^{n}(x_1 - y_1)^2 + (x_2 - y_2)^2 + \ldots + (x_i - y_i)^2}$

$K2 = \sqrt{(38-25)^2 + (300-220)^2 + (3-1)^2}$

$K2$= 28,008

3. Jarak antara Lokasi K-2 dengan titik K-3

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

$$K3 = \sqrt{(30{,}93 - 25)^2 + (247{,}33 - 220)^2 + (2{,}53 - 1)^2}$$

$K3$= 81,074

BCV= 53,144 + 28,008 + 81,074
= 162,226

Menghitung kembali nilai rasio dengan membandingkan nilai BCV dan WCV.

BCV/WCV = 162,226 / 156,211
= 1,039

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

Diketahui :

K-1 = Zona Merah
K-2 = Zona Kuning
K-3 = Zona Hijau

1. Pada *cluster* 1 terdapat 4 data maka nilai rata-rata pada *cluster*-1 adalah
   K-1,1 = (37 + 36 + 38 + 35) / 4= 36,50
   K-1,2 = (280 + 290 + 300 + 295) / 4 = 291,25
   K-1,3 = (0 + 5 + 3 + 0) / 4 = 2,00
2. Pada *cluster* 2 terdapat 5 data maka nilai rata-rata pada *cluster*-2 adalah
   K-2,1 = (32 + 34 + 29 + 33 + 31) / 5 = 31,80
   K-2,2 = (250 + 260 + 240 + 270 + 260) / 5 = 256,00
   K-2,3 = (3 + 1 + 2 + 3 + 4) / 5 = 2,60
3. Pada *cluster* 3 terdapat 6 data maka nilai rata-rata pada *cluster*-3 adalah
   K-3,1 = (25 + 28 + 30 + 26 + 27 + 30) / 6 = 27,67
   K-3,2 = (220 + 200 + 210 + 230 + 210 + 225) / 6 = 215,83
   K-3,3 = (1 + 5 + 3 + 4 + 0 + 2) / 6 = 2,50

Tabel 9 Titik Pusat *Cluster* pada Iterasi-2

Titik Pusat Awal	Kemiringan	Curah Hujan	Jenis Tanah
K-1	36,50	291,25	2,00
K-2	31,80	256,00	2,60
K-3	27,67	215,83	2,50

Jika proses pengelompokan belum stabil—yaitu anggota tiap cluster masih mengalami perubahan—maka algoritma K-Means harus mengulangi langkah penghitungan dari tahap ke-2 hingga ke-5. Perulangan ini terus dilakukan hingga komposisi anggota cluster tetap dan nilai centroid tidak mengalami perubahan dari iterasi sebelumnya. Setelah jumlah cluster (k) dan titik pusat awal (centroid) ditentukan, langkah selanjutnya adalah menghitung jarak antar centroid menggunakan Euclidean Distance, yang kemudian menghasilkan matriks jarak antar cluster, seperti K-1, K-2, dan K-3.

Tabel 10 Hasil Perhitungan Iterasi Ke 2

No	K-1	K-2	K-3	Jarak Terdekat	Jarak Centroid Terkecil
1	41,507	6,017	34,444	K-2	6,017
2	72,179	36,672	5,169	K-3	5,169
3	11,437	24,694	64,890	K-1	11,437
4	91,694	56,180	16,033	K-3	16,033
5	81,516	46,037	6,303	K-3	6,303
6	31,366	4,837	44,644	K-2	4,837
7	62,176	26,676	14,343	K-3	14,343
8	51,796	16,254	24,209	K-2	16,254
9	21,560	14,057	54,431	K-2	14,057
10	81,828	46,323	6,381	K-3	6,381
11	3,288	34,342	74,675	K-1	3,288
12	31,793	4,313	44,318	K-2	4,313
13	8,934	44,436	84,800	K-1	8,934
14	66,568	31,058	9,472	K-3	9,472
15	4,507	39,217	79,545	K-1	4,507
Total					**131,346**

Dari Tabel 10 adapun proses pencarian nilai WCV (*Within-Cluster Variability*) berikut :

WCV = (5,244 + 5,169 + ... + 15,880)
= 131,346

Dari proses *cluster* Tabel 10 di dapat Pemetaan zona longsor sebagai berikut :

Tabel 11 Hasil Cluster 1, 2, dan 3 Iterasi-2

No	Cluster	Jumlah Cluster
1	K-1	4
2	K-2	5
3	K-3	6

Dari Tabel 11 adapun daftar hasil *cluster* anggota K1-K3 sebagai berikut

1. K-1 ={ 3, 11, 13, 15}
2. K-2 ={ 1, 6, 8, 9, 12}
3. K-3 ={ 2, 4, 5, 7, 10, 14}

Dari Tabel 11 adapun proses pencarian nilai BCV(*Between-Cluster Variability*) berikut :

1. Jarak antara Lokasi K-1 dengan titik K-2
$= \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2}$
$= \sqrt{(36{,}50 - 31{,}80)^2 + (291{,}25 - 256{,}00)^2 + (2{,}00 - 2{,}60)^2}$
= 35,567
2. Jarak antara Lokasi K-1 dengan titik K-3
$= \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2}$
$= \sqrt{(36{,}50 - 27{,}67)^2 + (291{,}25 - 215{,}83)^2 + (2{,}00 - 2{,}50)^2}$
= 75,934
3. Jarak antara Lokasi K-2 dengan titik K-3
$= \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2}$
$= \sqrt{(31{,}80 - 27{,}67)^2 + (256{,}00 - 215{,}83)^2 + (2{,}60 - 2{,}50)^2}$
= 40,379

BCV = 35,567 + 75,934 + 40,379
= 151,880

Langkah akhir dalam evaluasi hasil clustering adalah menghitung rasio antara Between-Cluster Variability (BCV) dan Within-Cluster Variability (WCV). Rasio ini digunakan untuk mengukur kualitas pengelompokan yang dihasilkan. Rumus yang digunakan adalah:

BCV/WCV = 151,880 / 131,346
= 1,156

Tabel 12 Titik Pusat *Cluster* pada Iterasi-3

Titik Pusat Awal	Kemiringan	Curah Hujan	Jenis Tanah
Cluster 1	36,50	291,25	2,00
Cluster 2	31,80	256,00	2,60
Cluster 3	27,67	215,83	2,50

Tabel 13 Hasil Perhitungan Iterasi Ke 3

No	K-1	K-2	K-3	Jarak Terdekat	Jarak Centroid Terkecil
1	41,507	6,017	34,444	K-2	6,017
2	72,179	36,672	5,169	K-3	5,169
3	11,437	24,694	64,890	K-1	11,437
4	91,694	56,180	16,033	K-3	16,033
5	81,516	46,037	6,303	K-3	6,303
6	31,366	4,837	44,644	K-2	4,837
7	62,176	26,676	14,343	K-3	14,343
8	51,796	16,254	24,209	K-2	16,254
9	21,560	14,057	54,431	K-2	14,057
10	81,828	46,323	6,381	K-3	6,381
11	3,288	34,342	74,675	K-1	3,288
12	31,793	4,313	44,318	K-2	4,313

13	8,934	44,436	84,800	K-1	8,934
14	66,568	31,058	9,472	K-3	9,472
15	4,507	39,217	79,545	K-1	4,507
Total					**131,346**

Dari proses *cluster* Tabel 13 di dapat Pemetaan zona longsor sebagai berikut :

Tabel 14 Hasil Cluster 1, 2, dan 3 Iterasi-2

No	Cluster	Jumlah Cluster
1	K-1	5
2	K-2	4
3	K-3	6

Dari Tabel 14 adapun daftar hasil *cluster* anggota K1-K3 sebagai berikut

1. K-1 ={ 3, 11, 13, 15}
2. K-2 ={ 1, 6, 8, 9, 12}
3. K-3 ={ 2, 4, 5, 7, 10, 14}

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 zona pemetaan longsor daerah berastagi diambil dari iterasi 2 dikarenakan rasio iterasi 3 sampa ke 4 hasilnya sama, maka adapun hasil pengelompokkan hasil cluster adalah sebagai berikut:

Tabel 4.15 Pengelompokan Hasil Cluster dari iterasi 3

Cluster	Nomor	Nama Lokasi
Zona Merah	3, 11, 13, 15	Sikulikap, Kandibata, Tambak Lau Mulgap II, Sembahe Km 36.2
Zona Kuning	1, 6, 8, 9, 12	Desa Doulu, Gundaling II, Rih Tengah, Peceren, Tambak Lau Mulgap I
Zona Hijau	2, 4, 5, 7, 10, 14	Sempajaya, Bukit Kubu, Gundaling I, Rumah Kaban, Jaranguda, Desa Raya

Hasil clustering iterasi ketiga menunjukkan pembagian wilayah rawan longsor di Kecamatan Brastagi menjadi tiga zona: Merah (tinggi), Kuning (sedang), dan Hijau (rendah), berdasarkan kemiringan lereng, curah hujan, dan kestabilan tanah. Temuan ini menegaskan efektivitas metode K-Means sebagai dasar mitigasi bencana. Sistem SIG yang dikembangkan menampilkan antarmuka intuitif untuk input data dan pemrosesan clustering, dengan menu utama sebagai pusat navigasi menuju fitur-fitur aplikasi.

1. Menu Login
   Menu login merupakan lapisan keamanan yang membatasi akses sistem hanya kepada pengguna yang sah. Pengguna harus mengisi nama pengguna dan kata sandi sebelum dapat melanjutkan ke dalam sistem.



Gambar 2 Menu Login

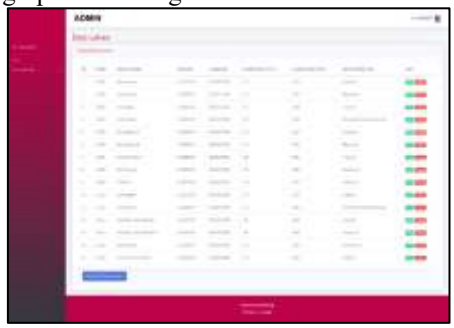
2. Menu Utama
   Setelah proses autentikasi berhasil, sistem akan menampilkan halaman utama yang menyediakan akses menuju halaman data lokasi, pemrosesan K-Means, dan laporan.



Gambar 3 Menu Utama

Halaman administrator menampilkan antarmuka untuk mengelola dan memproses data yang disimpan ke dalam basis data, termasuk input data lokasi dan eksekusi algoritma K-Means.

1. Menu Data Lokasi Longsor
   Halaman ini digunakan untuk menginput dan mengelola data lokasi rawan longsor di wilayah Kecamatan Brastagi.



Gambar 4 Menu Data Lokasi Longsor

2. Menu Data Centroid
   Menu ini digunakan untuk mengelola data pusat cluster (centroid) awal yang digunakan dalam proses clustering dengan K-Means.



Gambar 5 Menu Data Centroid

3. Menu Proses K-Means
   Halaman ini berfungsi untuk menjalankan proses klasterisasi menggunakan algoritma K-Means berdasarkan parameter data yang telah dimasukkan.



Gambar 6 Menu Proses K-Means

Tahap pengujian dilakukan untuk menguji keakuratan dan validitas sistem yang telah dirancang. Pengujian dilakukan dengan menggunakan data uji yang belum digunakan sebelumnya untuk memastikan sistem dapat memproses dan mengelompokkan data dengan benar. Hasil dari proses klasterisasi divisualisasikan dalam bentuk peta dan laporan.



Gambar 7 Hasil Mengelompokkan K-Means



Gambar 8 Hasil Pemetan GIS Wilayah Longsor Berastagi



Gambar 9 Cetak Laporan

KESIMPULAN

Berdasarkan hasil implementasi, sistem informasi geografis berbasis web yang dikembangkan dengan metode K-Means Clustering terbukti mampu memetakan tingkat kerawanan longsor di Kecamatan Brastagi secara lebih terstruktur. Melalui analisis parameter kemiringan lereng, curah hujan, dan jenis tanah, wilayah penelitian berhasil diklasifikasikan ke dalam tiga kategori utama, yaitu zona merah (tinggi), zona kuning (sedang), dan zona hijau (rendah). Hasil ini memberikan gambaran spasial yang komprehensif mengenai distribusi potensi bencana longsor di wilayah tersebut.

Proses pengelompokan dilakukan secara iteratif, dimulai dari penentuan centroid awal, perhitungan jarak menggunakan metode Euclidean, hingga terbentuk klaster yang stabil tanpa perubahan anggota. Prosedur ini memungkinkan sistem menghasilkan analisis yang representatif terhadap kondisi geologis setempat. Dengan integrasi basis data MySQL dan pemrograman PHP, sistem mampu menyajikan data dalam bentuk peta interaktif, sehingga memudahkan pengguna dalam memahami hasil klasifikasi kerawanan longsor pada 15 titik lokasi penelitian.

Selain itu, sistem yang dikembangkan memiliki antarmuka yang ramah pengguna dengan fitur input data, pengolahan menggunakan algoritma K-Means, visualisasi hasil, hingga pencetakan laporan. Hal ini menjadikan sistem tidak hanya sebagai sarana analisis spasial, tetapi juga sebagai alat pendukung pengambilan keputusan yang dapat dimanfaatkan oleh pemerintah daerah maupun lembaga terkait. Dengan demikian, sistem ini diharapkan mampu mendukung strategi mitigasi bencana longsor di Kecamatan Brastagi secara lebih efektif dan berbasis data ilmiah.

REFERENCES

Londa, G. O., Witi, F. L., & Bhae, B. Y. (2022). Sistem Informasi Pendataan Penduduk Desa Detusoko Barat Kecamatan Detusoko Kabupaten Ende Berbasis Web. JURNAL JITEK, II(2).

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.

Aprianti, W., & Maliha, U. (2016). Sistem Informasi Kepadatan Penduduk Kelurahan Atau Desa Studi Kasus Pada Kecamatan Bati-Bati Kabupaten Tanah Laut. Jurnal Sains dan Informatika, II(1).

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

Hendini, A. (2016). PEMODELAN UML SISTEM INFORMASI MONITORING PENJUALAN DAN STOK BARANG (STUDI KASUS: DISTRO ZHEZHA PONTIANAK). JURNAL KHATULISTIWA INFORMATIKA,, IV(2).

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

Kinanti, A., Awaluddin, M., & Yusuf, M. A. (2022). Analisis Pemetaan Risiko Bencana Tanah Longsor Berbasis Sistem Informasi Geografis (Studi Kasus: Kecamatan Candisari, Kota Semarang). Jurnal Geodesi Undip, 1-10.

Marzuki, I., Ariyanti, D., & Firmansyah, D. (2020). Aplikasi Sebaran Pengangguran Di Kota Probolinggo Berbasis Web. Jurnal Energy, X(2), 83-91.

Maulani, M. R. (2022). Rancang Bangun Sistem Informasi Inventori Furniture Menggunakan Metode Mark Up Pricing Pada Toko XYZ. Jurnal Teknik Informatik, XIV(1).

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.

Novria, R., Kurniawan, B., & Suryanto. (2022). Aplikasi Pemesanan Makanan Di Bebek dan Ayam Tekaeng Menggunakan Php dan Mysql. Jurnal Informatika dan Komputer (JIK), XIII(1).

Nurprihatin, F., Tannady, H., Lusiani, M., Karo-Karo, G., & Renatha . (2024). Algoritma K-Means Clustering Untuk Menentukan Lokasi Pusat Distribusi Ritel Dengan Mempertimbangkan Jumlah Permintaan. Jurnal PASTI Volume, 32 - 44.

Prihandoyo, M. T. (2018). Unified Modeling Language (UML) Model Untuk Pengembangan Sistem Informasi Akademik Berbasis Web. Jurnal Informatika: Jurnal Pengembangan IT (JPIT), III(2477-5126), 126-129.

Rivaldi, A., Irawan, J. D., & Faisol, A. (2023). Sistem Informasi Geografis Pemetaan Sebaran Keluarga Miskin Menggunakan Metode K-Means. JATI (Jurnal Mahasiswa Teknik Informatika), 2436-2442.

Sinaga, K. P., & Yang, M. S. (2025). A Globally Collaborative Multi-View k-Means Clustering. Electronics, 1-31.

Suendri. (2018). Implementasi Diagram UML (Unified Modelling Language) Pada Perancangan Sistem Informasi Remunerasi Dosen Dengan Database Oracle (Studi Kasus: UIN Sumatera Utara Medan). Jurnal Ilmu Komputer dan Informatika, III(2598-6341), 1-9.

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

Taruna, A., Soemantri, L., & Setiawan, I. (2021). Pemetaan Kawasan Potensi Rawan Longsor Berbasis Data Analisis Sistem Informasi Geografis Di Kota Cimahi. Jurnal Kajian Ilmu dan Pendidikan Geografi, 7-13.

Ufriani, S., Jasmir, D., & Arviita , Y. (2023). Penerapan Algoritma Clustering K-Means Untuk Menentukan Prioritas Penerima Bantuan Dana Sosial PKH Di Kelurahan Kampung Singkep. Jurnal Informatika Dan Rekayasa Komputer (JAKAKOM), 342-350.