

**PROPOSAL SKIRIPSI**

**IMPLEMENTASI ALGORITMA *K-MEANS* DALAM PENGELOMPOKAN DATA KECELAKAAN DI**

**KABUPATEN PADANG LAWAS UTARA**

**Disusun Sebagai Syarat Memperoleh Gelar Sarjana Komputer**

**Prodi Teknik Informatika Fakultas Teknik**

**Universitas Malikussaleh**

**DISUSUN OLEH:**

**NAMA : KHOPIPAH PARAWANSAH SIREGAR**

**NIM : 200170081**

**PRODI : TEKNIK INFORMATIKA**

**JURUSAN INFORMATIKA**

**FAKULTAS TEKNIK**

**UNIVERSITAS MALIKUSSALEH**

**LHOKSEUMAWE**

**2024**

# 

# **DAFTAR ISI**

[**DAFTAR ISI** **i**](#_Toc161637548)

[**DAFTAR TABEL** **iii**](#_Toc161637549)

[**DAFTAR GAMBAR** **iv**](#_Toc161637550)

[**BAB I PENDAHULUAN** **1**](#_Toc161637551)

[1.1 Latar Belakang 1](#_Toc161637553)

[1.2 Rumusan Masalah 3](#_Toc161637554)

[1.3 Batasan Masalah 4](#_Toc161637555)

[1.4 Tujuan Penelitian 4](#_Toc161637556)

[1.5 Manfaat Penelitian 4](#_Toc161637557)

[**BAB 11 TINJAUAN PUSTAKA** **6**](#_Toc161637558)

[2.1 Kondisi Umum Kabupaten Padang Lawas Utara 6](#_Toc161637560)

[2.2 Kecelakaan 7](#_Toc161637561)

[2.3 Zonasi 7](#_Toc161637562)

[2.4 Data Mining 8](#_Toc161637563)

[2.4.1 Pengertian Data Mining 8](#_Toc161637564)

[2.4.2 Metode Data Mining 10](#_Toc161637565)

[2.4.3 Clustering 11](#_Toc161637566)

[2.5 Algoritma *K-Means* 11](#_Toc161637567)

[2.6 Sistem Informasi Geografis 12](#_Toc161637568)

[2.6.1 Data Spasial dan Non Spasial 13](#_Toc161637569)

[2.6.2 Quantum Gis 13](#_Toc161637570)

[2.6.3 OpenStreetMap 13](#_Toc161637571)

[2.7 PHP 14](#_Toc161637572)

[2.8 Database MYSQL 14](#_Toc161637573)

[2.9 XAMPP 15](#_Toc161637574)

[2.10 Unified Modeling Language (UML) 16](#_Toc161637575)

[2.10.1 Use Case Diagram 16](#_Toc161637576)

[2.10.2 Actifity Diagram 17](#_Toc161637577)

[2.10.3 Squense Diagram 17](#_Toc161637578)

[2.10.4 Class Diagram 19](#_Toc161637579)

[**BAB III METODELOGI PENELITIAN** **21**](#_Toc161637580)

[3.1 Waktu dan Tempat Pelaksanaan Penelitian 21](#_Toc161637582)

[3.2 Metode Penelitian 21](#_Toc161637583)

[3.3 Analisa Kebutuhan Sistem 22](#_Toc161637584)

[3.3.1 Perangkat Keras (*hardware*) 23](#_Toc161637585)

[3.3.2 Perangkat Lunak (*software*) 23](#_Toc161637586)

[3.3.3 Analisis Kebutuhan Sistem 23](#_Toc161637587)

[3.3.4 Analisis Kebutuhan Output Sistem 24](#_Toc161637588)

[3.4 Skema 25](#_Toc161637589)

[**DAFTAR PUSTAKA** **28**](#_Toc161637590)

# **DAFTAR TABEL**

[Tabel 2. 1 Atribut Use Case Diagram 16](#_Toc161344039)

[Tabel 2. 2 Atribut Actifity Diagram 17](#_Toc161344040)

[Tabel 2. 3 Atribut Squense Diagram 18](#_Toc161344041)

[Tabel 2. 4 Atribut Class Diagram 19](#_Toc161344042)

# **DAFTAR GAMBAR**

[Gambar 2. 1 Peta Daerah Kabupaten Padang Lawas 6](#_Toc161642886)

[Gambar 2. 2 Proses Knowledge Discovery in Database (KDD) 9](#_Toc161642887)

[Gambar 3. 1 Langkah Penelitian 21](#_Toc161642888)

[Gambar 3. 2 Skema Sistem 26](#_Toc161642889)

# **BAB I**

# **PENDAHULUAN**

## **1.1 Latar Belakang**

Di era modern saat ini, keselamatan lalu lintas merupakan topik yang selalu diperbincangkan karena semakin seringnya insiden kecelakaan di jalan raya. Kecelakaan lalu lintas membutuhkan penanganan serius supaya tingkat kerugiannya tidak semakin besar. Berbagai upaya yang dilakukan untuk menanggulangi kecelakaan lalu lintas telah di terapkan oleh pemerintah antara lain yaitu pemberian rambu-rambu lalu lintas namun masalah tersebut tidak dapat terselesaikan dengan mudah begitu saja. Tingginya jumlah kejadian kecelakaan lalu lintas merupakan salah satu permasalahan yang cukup besar dibidang kesehatan masyarakat. Selain menjadi penyebab kematian tertinggi di dunia, kecelakaan lalu lintas juga merupakan salah satu penyebab yang seringkali terjadi diantara tenggelam, terbakar, cedera akibat terjatuh dan keracun. Oleh karena itu perlu dilakukan upaya penanggulangan untuk memanimalisir kecelakaan.

Kabupaten Padang Lawas Utara merupakan salah satu kabupaten di Provinsi Sumatera Utara yang berada di Indonesia dengan disertai pertumbuhan dan perkembangan penduduk yang tinggi. Sejalan dengan hal itu kegiatan transportasi di Kabupaten Padang Lawas Utara semakin meningkat, karena era sekarang kendaraan bukan hanya sebagai kebutuhan skunder melainkan menjadi kebutuhan primer, sehingga pertumbuhan jumlah kendaraan semakin pesat baik itu kendaraan roda 2 maupun kendaraan roda 4. Dan hal tersebut akan menuntut peningkatan kualitas sarana dan prasarana di Kabupaten Padang Lawas Utara. Dengan segala perkembangan dan pertumbuhan yang terjadi, tentu akan menimbulkan beberapa masalah transportasi salah satunya kecelakaan lalu lintas.

Kasus kecelakaan di Kabupaten Padang Lawas Utara dari tahun ke tahun dimungkinkan semakin bertambah banyak, dengan data yang banyak tersebut sangat sulit untuk menentukan daerah yang rawan kecelakaan. sehingga, penggunaan teknologi data mining diperlukan untuk mananggulangi masalah tersebut. Alasan digunakannya teknologi data mining karena teknologi tersebut

termasuk teknologi sederhana dan cepat dalam melakukan proses clustering daerah mana saja yang rawan terjadi kecelakaan lalu lintas.

Data mining merupakan bagian dari *Big Data,* sedangkan *Big Data* adalah salah satu bahasa yang paling banyak digunakan di kalangan industri dimana *Big Data* merupakan suatu konsep abstark. *Big Data* umumnya digambarkan sebagai perilaku integrasi data yang sangat banyak dan kompleks yang susah dikerjakan dengan alat manajemen *database* yang ada. Dan *Big Data* sering dicirikan sebagai lima yaitu kuantitas, kecepatan, keragaman, nilai, dan akurasi (Yang et al. 2020).

Jika *Big Data* adalah bagian dari data mining maka data mining juga tidak akan terlepas dari *Big Data.* Menerapkan teknologi data mining dapat dengan mudah menganalisa data seperti, data daerah rawan kecelakaan. Dan penerapan Teknik data mining dapat kita pergunakan untuk mempermudah dalam menganalisa data dengan *clustering* yaitu Algoritma *K-Means*. Algoritma *K-Means* termasuk dalam metode sederhana dan cepat dalam melakukan proses *clustering.* Dengan adanya Algoritma *K-Means* dapat mempermudah untuk membuat penentuan daerah mana saja yang rawan kecelakaan lalu lintas. Penggunaan Algoritma *K-Means* dapat lebih mudah dan optimal pada pekerjaan yang dilakukan.

Kurangnya Pemanfaatan teknologi saat ini dalam sistem pelayanan pengaduan korban kecelakaan pada Badan Pusat Statistik (BPS) kabupaten Padang Lawas Utara dikarenakan sistem pengaduan masih menggunakan konvensional dimana masyarakat melakukan pelaporan korban kecelakaan melalui telepon kepada pihak Badan Pusat Statistik (BPS) Kabupaten Padang Lawas Utara. Dikarenakan penggunaan sistem pelaporan masih konvensional pihak Badan Pusat Statistik (BPS) Kabupaten Padang Lawas Utara mengalami kesulitan dalam mendapatkan atau mengumpulkan informasi tentang pemetaan daerah yang sering terjadi kecelakaan atau daerah rawan kecelakaan. Oleh karena itu petugas kesulitan dalam merangkum informasi tentang daerah yang rawan terjadi kecelakaan untuk analisis dan pemetaan kecelakaan yang akan datang.

Dari permasalahan di atas peneliti membuat model *Clustering,* Perancangan dan juga implementasi perangkat lunak untuk pengembangan zonasi rawan kecelakaan yang berfungsi sebagai pemetaan daerah rawan kecelakaan untuk acuan mengantisipsi para pengendara untuk berhati-hati di daerah rawan kecelakaan. Serta mendukung kecepatan layanan informasi kecelakaan di Kabupaten Padang Lawas Utara. Dengan penggunaan Algoritma *K-Means* peneliti akan Mengklasterkan menjadi 3 kriteria yaitu sangat rawan, rawan, dan tidak rawan.

Sebelumnya telah dilakukan penelitian terdahulu terkait tentang clustering daerah rawan kecelakaan, yaitu: Penelitian oleh Risawandi dan Yesy Afrillia “Sistem Informasi Geografis Pemetaan Daerah Rawan Kriminalitas di Lhokseumawe Menggunakan Metode *K-Means*” yang bertujuan untuk menginformasikan kepada masyarakat dimana wilayah kecamatan yang tinggi rawan kriminalitas (Risawandi and Afrillia 2022). Penelitian oleh Ilsya Hidayat Nasution, Eva Darnila, dan Yesy Afrillia “*Clustering* Zonasi Bencana Alam Di Kabupaten Mandailing Natal Menggunakan Algoritma *K-Means*” yang bertujuan untuk menginformasikan kepada masyarakat dimana wilayah kecamatan yang rawan zonasi bencana banjir di mandailing natal (Hidayat, Darnila, and Afrillia 2023). Penelitian Oleh Adimas Ketut Nalendra, M.Mujiono, Rafika Akhsani dan Adiguna Sasama Wahyu U “Implementasi Algoritma *K-Means* Dalam Pengelompokan Data Kecelakaan (Studi Kasus Kabupaten Kediri” yang bertujuan untuk memberikan informasi kepada Masyarakat wilayah daerah yang rawan kecelakaan di kabupaten Kediri (Nalendra et al. 2020).

Atas dasar uraian diatas, maka peneliti tertarik untuk melakukan penelitian dengan judul “**Implementasi Algoritma *K-Means* Dalam Pengelompokan Data Kecelakaan di Kabupaten Padang Lawas Utara**”.

## **1.2 Rumusan Masalah**

Kecelakaan yang sering terjadi marupakan permasalahan yang serius dan butuh penanganan secara tepat. Maka berdasarkan latar belakang diatas diambil suatu rumusan masalah adalah:

1. Bagaimana merancang *Clustering* pemetaan daerah rawan terjadi kecelakaan di Kecamatan Kabupaten Padang Lawas Utara?

2. Bagaimana Penerapan Algoritma *K-Means* untuk mengetahui daerah rawan terjadi kecelakaan di Kecamatan Kabupaten Padang Lawas Utara?

## **1.3 Batasan Masalah**

Berikut ini merupakan batasan masalah supaya penyajiannya lebih terarah serta saling terkait. Adapun Batasan masalahnya sebagai berikut:

1. Daerah penelitian hanya berada di wilayah Kabupaten Padang Lawas Utara.

2. Data di ambil dari wilayah Badan Pusat Statistik (BPS) Kabupaten Padang Lawas Utara pada tahun 2020-2023.

3. Algoritma yang digunkana pada penelitian ini adalah Algoritma *K-Means*.

4. Kelompok *Clustering* yang akan dibentuk adalah sangat rawan, rawan, dan tidak rawan.

5. Kasus yang akan diteliti adalah Kecelakaan.

6. Kriteria dari kecelakaan yang akan digunakan yaitu jumlah kejadian kecelakaan, jumlah meninggal, jumlah luka berat, dan jumlah luka ringan.

7. Sistem yang dibuat di kembangkan dalam bentuk website yang menampilkan hasil berbentuk peta lokasi sangat rawan, rawan, dan tidak rawan yang dapat di akses oleh admin dan user.

## **1.4 Tujuan Penelitian**

Adapun tujuan dari penelitian ini yaitu:

1. Untuk merancang *Clustering* pemetaan daerah rawan kecelakaan di Kabupaten Padang Lawas Utara menggunakan Algoritma *K-Means.*

2. Untuk mengetahui daerah rawan kecelakaan di Kabupaten Padang Lawas Utara menggunakan Algoritma *K-means.*

## **1.5 Manfaat Penelitian**

Adapun manfaat dari penelitian ini adalah:

1. Memberikan informasi tentang wilayah yang rawan terjadi kecelakaan di Kabupaten Padang Lawas Utara kepada masyarakat yang berada di daerah Kabupaten Padang Lawas Utara maupun yang di luar Kabupaten.

2. Penelitian ini dapat berguna untuk menambah informasi kepada masyarakat untuk menjadi acuan daerah yang sering terjadi kecelakaan di Kabupaten Padang Lawas Utara. Dan memberikan informasi tentang kecelakaan dan upaya-upaya yang dapat dilakukan untuk pencegahan terjadinya kecelakaan.

3. Untuk peneliti, menambah pengetahuan tentang pentingnya peta daerah rawan terjadi kecelakaan serta menambah pengetahuan atau wawasan maupun pengalaman dalam pengelolahan data dengan menggunakan Algoritma *K-means Clustering.*

# **BAB 11**

# **TINJAUAN PUSTAKA**

## **2.1 Kondisi Umum Kabupaten Padang Lawas Utara**

Padang Lawas Utara atau sering dikenal PALUTA adalah salah satu Kabupaten yang berada di Provinsi Sumatera Utara, Indonesia. Padang Lawas Utara merupakan hasil pemekaran dari Kabupaten Tapanuli Selatan pada tahun 2007. Sesuai dengan Undang-Undang Republik Indonesia Nomor 37 tahun 2007, tentang pembentukan Kabupaten Padang Lawas Utara. Ibu kota Kabupaten ini berada di Gunung Tua. Diperkirakan jumlah penduduk yang berada di Kabupaten Padang Lawas Utara yaitu sebanyak 269.845 jiwa dengan kepadatan 69 jiwa/. Dibawah ini, adalah gambar yang memperlihatkan daerah Padang Lawas Utara.



Gambar 2. 1 Peta Daerah Kabupaten Padang Lawas

(https://id.wikipedia.org/wiki/Daftar\_kecamatan\_dan\_kelurahan\_di\_Kabupaten\_Padang\_Lawas\_Utara#/media/Berkas:(Peta\_Wilayah)\_Kabupaten\_Padang\_Lawas\_Utara.svg)

Kabupaten Padang Lawas Utara memiliki jumlah Kecamatan sebanyak 12 Kecamatan yaitu batang onang, dolok, dolok sigompulon, halongonan, halongonan

timur, hulu sihapas, padang bolak, padang bolak julu, padang bolak tenggara, portibi, simangambat, dan ujung batu.

## **2.2 Kecelakaan**

merupakan suatu peristiwa yang tidak diinginkan yang dapat menimbulkan kerugian baik secara fisik maupun material. Kecelakaan dapat terjadi di berbagai tempat dan situasi, mulai dari kecelakaan lalu lintas, kecelakaan kerja, kecelakaan di rumah, hingga kecelakaan di tempat umum. Kecelakaan juga dapat disebabkan oleh berbagai faktor, seperti kelalaian, ketidaksengajaan, atau faktor alam. Kecelakaan lalu lintas dapat didefinisikan sebagai suatu peristiwa yang tidak diinginkan yang terjadi di jalan raya yang melibatkan kendaraan bermotor dan/atau pejalan kaki. Ada banyak faktor yang dapat menyebabkan terjadinya kecelakaan lalu lintas. Salah satu faktor utama adalah human error, yaitu kesalahan yang dilakukan oleh pengemudi kendaraan atau pejalan kaki. Contoh dari human error adalah melanggar rambu-rambu lalu lintas, mengemudi dalam keadaan mabuk, atau tidak memperhatikan kondisi jalan. Selain human error, kondisi jalan yang buruk juga dapat menjadi faktor penyebab kecelakaan lalu lintas. Jalan yang berlubang, licin akibat hujan, atau tidak terdapatnya lampu penerangan malam hari dapat meningkatkan risiko terjadinya kecelakaan. Selain itu, faktor cuaca yang buruk juga dapat menyebabkan kecelakaan lalu lintas, seperti hujan deras, kabut tebal, atau angin kencang. Kendaraan yang tidak dalam kondisi baik juga dapat menjadi faktor penyebab kecelakaan lalu lintas, seperti rem yang blong, ban kempes, atau lampu yang mati.

## **2.3 Zonasi**

Zonasi merupakan wilayah yang mempunyai fungsi serta karakteristik kawasan yang spesifik. Adapun peraturan zonasi berfungsi untuk mengelompokkan suatu area supaya Sebagian area dengan peraturan hukum, bertujuan untuk pemisahan pembangunan pabrik supaya memiliki keuntungan dari area perumahan. Pada waktu akhir abad ke-19, Jerman mengembangkan rancangan zonasi, dan juga setelah itu rancangan itu tersiar ke negara lain yaitu Amerika Serikat dan juga Kanada pada awal mula abad ke-20. Rancangan ini terlihat sebagai jawaban kepada industrialisasi serta melonjaknya keluh kesah publik berkenaan dengan gangguan semacam musibah alam. Penguasa semestinya cepat melakukan tindakan mencari teknik solusi buat menangani gangguan-gangguan dengan efek tidak baik dari urbanisasi serta perkembangan populasi masyarakat (Heryana 2020).

## **2.4 Data Mining**

### **2.4.1 Pengertian Data Mining**

Data mining merupakan suatu metode yang digunakan untuk mengolah data untuk menemukan pola yang tersembunyi dari data yang diolah. Data yang diolah dengan data mining ini kemudian menghasilkan suatu pengetahuan baru yang bersumber dari data lama, hasil dari pengolahan data tersebut dapat digunakan dalam menentukan Keputusan dimasa depan. Data mining merupakan hasil dari ekstraksi atau pemahaman pattern yang menarik pada data, atau dapat juga diartikan sebagai serangkaian proses untuk menggali nilai tambah dari satu atau beberapa kumpulan data pengetahuan yang selama ini tidak diketahui secara manual. Tujuan dilakukan data mining adalah menemukan hubungan atau pola yang memberikan informasi yang bermanfaat bagi pembacanya.

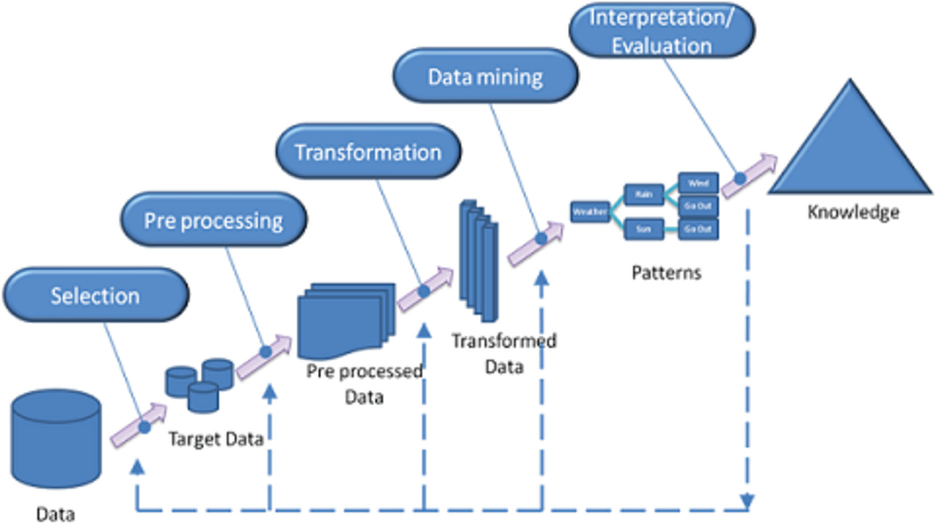
Dalam hal ini data mining dapat disimpulkan yaitu cara yang digunakan dengan mengungkap ilmu yang belum pernah terlihat, Dimana cara yang dilakukan yaitu *machine leraning,* matematika, kecerdasan buatan, dan statistik untuk mengidentifikasi data yang telah disimpan kedalam penyimpanan *database.* Adapun macam-macam kepentingan yang digunakan dalam data mining yaitu:

1. Analisa data dengan menerapkan data mining ialah data yang diolah dalam skala besar.

2. Melibatkan data mining dengan cara otomatisasi yang dilakukan dengan data tersedia.

3. Penelusuran informasi melalui proses data mining yang diterapkan untuk mengolah data dalam basis data yang cukup besar. Metode *K-Means* *Clustering* dipandang sebagai pilihan yang tepat untuk mengelompokkan data yang mempuanyai dimensi sangat lebar.

Selain itu proses Knowledge Discovery in Database atau bisa di singkat dengan *(KDD*) termasuk komponen terhadap proses Data mining dengan menukar data awal menjadi sangat bermanfaat dan bernilai (Damanik et al. 2021).



Gambar 2. 2 Proses Knowledge Discovery in Database (KDD)

Tahapan dalam proses KDD sebagai berikut :

a. Data Selection

Pada tahapan ini, data yang relevan dengan masalah yang sedang diteliti dipilih dan diambil dari berbagai sumber data yang tersedia.

b. *Pre-Processing/Cleaning*

Tahapan ini melibatkan pembersihan dan persiapan data agar dapat digunakan dalam proses analisis lebih lanjut.Termasuk di dalamnya adalah penghilangan data yang tidak lengkap atau tidak relevan, penanganan nilai yang hilang dan normaliasi data.

c. *Transformation*

Perubahan atau transformasi menyertakan integrasi informasi yang sudah dipilih maka cocok buat cara data mining.

d. Data Mining

Penentuan kewajiban data mining yang jadi tujuan dari prosedur KDD, semacam klasifikasi, karakterisasi, pengelompokan, regresi, asosiasi, dan sejenisnya.

e. *Interpretation/Evaluation*

Langkah ini menyertakan catatan serta pengecekan pola ataupun informasi yang telah didapati bertentangan terhadap bukti ataupun asumsi yang bertentangan terhadap bukti ataupun asumsi yang terdapat sebelumnya terjadi.

### **2.4.2 Metode Data Mining**

Data mining menerapkan pendekatan dengan menggunakan 2 metode Langkah, yaitu *Unsupervised learning* dan *Supervised learning*. *Unsupervised learning* merupakan Tahap yang menerapkan tanpa pengggunaan bimbingan (*training*) serta tanpa memiliki guru (*teacher*). Nama dari data bertindak sebagai guru. Sementara itu *Supervised learning* merupakan cara pembelajaran dengan melaksanakan bimbingan serta instruktur. Dalam suatu pendekatan, sebagian jenis informasi dengan output ataupun identitas yang di menggunakan sebagai manfaat pemisah, manfaat keputusan maupun manfaat regresi sepanjang prosedur latihan. Metode yang dimanfaatkan oleh data mining untuk memastikan kewajiban yang bisa dilakukan adalah sebagai berikut:

1. **Deskripsi (Gambaran)**, Para ilmuan selalu mencari Teknik untuk mengartikan pola serta gaya yang tertutup ke dalam informasi.

2. **Estimasi (Perkiraan)**, sesuai pengelompokkan menyertakan faktor sasaran yang berkarakter numerik daripada kategori.

3. **Prediksi Serta Ramalan**, mendekati dengan perkiraan serta pengelompokan, akan tetapi membuktikan suatu yang belum pernah terjadi, (bisa jadi terjadi dimasa depan).

4. **Klasifikasi Ataupun Pengelompokan**, mengaitkan pengelompokan faktor objek yang berkarakter kategorikal. Sebagai ilustrasi kami, mampu mengkategorikan kecelakaan jadi tiga kelompok yaitu sangat rawan, rawan, dan tidak rawan.

5. ***Clustering* Ataupun Klasterisasi**, cara yang lebih menitikberatkan pada pengelompokan rekaman, observasi, ataupun masalah kedalam group bersumber pada kecocokan mereka.

6. **Asosiasi**, menyertakan pengidentifikasian ikatan antara berbagai kejadian yang terjadi dengan cara berbarengan.

### **2.4.3 Clustering**

*Clustering* adalah sebuah metode dalam analisis data yang digunakan untuk mengelompokkan objek-objek data ke dalam kelompok-kelompok yang serupa berdasarkan pada karakteristik atau atribut yang dimiliki. Tujuan utama dari clustering adalah untuk membagi data menjadi beberapa kelompok sedemikian rupa sehingga objek-objek dalam satu kelompok memiliki kesamaan yang tinggi di antara satu sama lain, sedangkan objek-objek di kelompok yang berbeda memiliki perbedaan yang signifikan. Metode *clustering* umumnya digunakan dalam berbagai bidang, termasuk ilmu komputer, statistik, ilmu sosial, ilmu hayati, dan lain sebagainya. Contoh penerapan clustering meliputi segmentasi pasar, analisis pola genetik, pengelompokan dokumen teks, dan pengelompokan pengguna dalam media sosial. *Clustering* dapat diartikan sebagai "pengelompokan" atau "penggolongan".

## **2.5 Algoritma *K-Means***

Algoritma *K-means* adalah salah satu metode yang paling umum digunakan dalam clustering. Tujuannya adalah untuk membagi himpunan data menjadi beberapa kelompok yang disebut cluster, sedemikian rupa sehingga objek dalam satu kelompok memiliki kesamaan yang tinggi dengan objek-objek dalam kelompok yang sama dan perbedaan yang tinggi dengan objek-objek dalam kelompok lainnya.

Metode *clustering K-Means* melibatkan cara iteratif untuk prosesnya. Dalam konteks *K-Means*, huruf K mewakili banyak cluster yang akan dibentuk. Seterusnya, nilai K ditentukan secara acak/random sebagai inisialisasi. Di sisi lain “Means” merupakan nilai tengah yang berfungsi sebagai pusat dari setiap klister, juga dikenal sebagai centroid (Harahap et al. 2022), yang merepresentasikan karakteristik kelompok tersebut. Secara keseluruhan, *K-Means* merupakan Cara data mining yang ampuh supaya mudah dalam melaksanakan pengelompokkan (*Clustering*) untuk analisis data. Prosedur metode *K-Means* dapat dijabarkan dibawah ini:

1. Tentukan brapa banyak Cluster (k) terhadapa data set.
2. Menginisialisasi titik pusat (Centroid) secara acak.
3. Menggunakan metode perhitungan jarak yang paling dekat dengan centroid, dapat menggunakan rumus (2.1) di bawah ini:

........................................................ (2.1)

Keterangan :

*d* = Euclidean Distance.

*i* = banyak objek.

*x,y* = Titik koordinat objek.

*s,t* = Titik koordinat centroid

1. Lakukan cara pengelompokan objek berlandaskan jarak yang paling dekat dengan centroid.
2. Hitung rata-rata untuk setiap kelompok menggunakan rumus (2.2) di bawah:

 *............................................................................* (2.2)

Keterangan :

*V\_ij* = Centroid rata-rata cluster ke-i untuk variable ke-j.

*N\_i* = Jumlah anggota cluster ke-i.

*i,k* = Indeks dari cluster.

*j* = Indeks dari variable.

*X\_kj*= Nilai data ke-k variable ke-j untuk cluster tersebut.

1. Lakukan perulanagan dari poin 3 serta poin 4 dan iterasi hingga mencapai centroid dengan nilai optimal (Risawandi and Afrillia 2022).

## **2.6 Sistem Informasi Geografis**

Sistem Informasi Geografis adalah sebuah sistem yang menyampaikan informasi berbasis computer dengan menggunakan data dengan komponen informasi spasial. Sistem ini melakukan pengambilan, analisis, pemeriksaan, integrasi, manajemen, pengolahan, dan visualisasi data secara spasial yang berkaitan dengan bentuk atau kondisi di bumi. Dengan memperhatikan unsur-unsur utamanya, dapat di rangkum bahwa Gis ialah sebuah sistem informasi yang mempunyai kefokusan terhadapa bagian "Informasi Geografis". SIG ialah satu buah alat lunak yang mengizinkan konsume untuk memasukkan, menyimpan, memanipulasi, menampilkan serta membuahkan informasi geografis bersama atributnya (Jannah et al. 2022). Implementasi teknologi Sistem Informasi Geografis (SIG) menyatukan bermacam cara pengerjaan informasi yang biasa dipakai masa ini, termasuk pengumpulan informasi yang relevan serta analisa statistik. Kelebihan teknologi ini terdapat pada penggambaran yang khas serta memberikan banyak manfaat dengan analisa geografis dengan dukungan ilustrasi yang ada.

### **2.6.1 Data Spasial dan Non Spasial**

Data spasial dan non-spasial adalah dua jenis data yang sering digunakan dalam berbagai bidang, terutama dalam ilmu geografi, ilmu lingkungan, dan sistem informasi geografis (SIG). Informasi atribut meliputi informasi memiliki kaitan dengan lokasi geografis, seperti nama, karakteristik dan jenis lainnya. Data spasial dapat disajikan dalam format vektor (polygon, line, point) atau raster. Sementara itu, data atribut berisi informasi dalam bentuk tabel yang menjelaskan karakteristik objek dalam data spasial (Kurniawati et al. 2020).

### **2.6.2 Quantum Gis**

*Quantum GIS* (*QGIS*) adalah sebuah perangkat lunak sistem informasi geografis (*SIG*) sumber terbuka yang digunakan untuk melihat, mengedit, dan menganalisis data spasial. *QGIS* menyediakan berbagai fitur untuk mengolah data geografis dan membuat peta, serta mendukung berbagai format data spasial, termasuk vektor dan raster, serta *database*.

### **2.6.3 OpenStreetMap**

*OpenStreetMa*p (*OSM)*, merupakan salah satu inisiatif yang timbul pada tahun 2004 di Inggris oleh lembaga nirlaba, makin menunjukkan keberadaannya dalam memberikan informasi geospasial berplatform komunitas dengan cara gratis pada rakyat ,*OSM* mengadakan informasi geospasial dengan beragam tema, termasuk prasarana transportasi seperti jalan, jalur kereta api aliran sungai dan informasi lainnya sesuai kawasan yang sangat penting seperti gedung, fitur alam, pemakaian tanah, garis pesisir serta batasan administratif. Selain itu, *OpenStreetMap* (OSM) telah terbukti memiliki data geografis yang lebih komprehensif di beberapa wilayah dan memiliki tingkat akurasi yang lebih tinggi secara lokal dan semantik jika dibandingkan dengan dataset lainnya. Hal ini menunjukkan betapa pentingnya penggunaan API OSM dalam penelitian ini untuk mencapai tujuannya. Proses pemetaan pada OSM diatur sedemikian rupa sehingga pengguna dapat mendaftar dan login, membuat, mengedit, dan menghapus peta, mengakses seluruh kumpulan data, serta mengakses riwayat lengkap data secara gratis, sehingga pengguna dapat melacak setiap tindakan yang dilakukan (Samah et al. 2020).

## **2.7 PHP**

*PHP* adalah singkatan dari "*Hypertext Preprocessor*". Ini adalah bahasa pemrograman server-side yang digunakan secara luas untuk pengembangan web. PHP pertama kali dibuat oleh Rasmus Lerdorf pada tahun 1994 dan sejak itu telah berkembang menjadi salah satu bahasa pemrograman web paling populer di dunia. Mulanya *PHP* dipakai sebagai bahasa pemrograman buat server-side HTML-embedded yang di promosikan sebagai Individu Home Pages. Rasmus Lerdorf merupakan pembuat sebuah Bahasa pemograman yaitu *PHP* waktu masa 1995. Pada masa itu, *PHP* di juluki FI (*Form Interpreted*), yang yang tergolong ke dalam sekelompok script yang di gunakan buat mengoperasikan keterangan form pada web.

## **2.8 Database MYSQL**

Karena karakteristik arsitektur dari *MySQL* yang flexible, *MySQL* dapat berjalan dengan baik di desktop application maupun web application. *MySQL* dapat digunakan untuk embedded application, data warehouses, content indexing and delivery software, highly reduntant systems, online transactio roccesing(OLTP), dan masih banyak lagi. Dengan konfigurasi yang benar, *MySQL* dapat digunakan pada berbagai macam perangkat keras, dan juga *MySQL* mendukung banyak tipe data. Salah satu fitur penting dari *MySQL* adalah storage-engine architecture yang designnya memisahkan query processing dan pekerjaan - pekerjaan server lainnya dari penyimpanan dan pengambilan data. Dengan adanya separasi ini dapat membuat pengguna memilih bagaimana data disimpan, performa apa yang ingin dicapai, fitur - fitur apa aja yang diinginkan, dan karakteristik lainnya. *MySQL* juga memiliki fitur Partitioned Tables sebagai logical table yang merupakan komposisi dari beberapa sub-tables. Tujuan dari partitioning ini adalah untuk membentuk coarse form of indexing dan data clustering dari table. Dengan begini sebagian besar dari table tidak perlu diakses dan untuk menyimpan rows yang berelasi dengan erat. Keuntungan *MySQL* adalah:

1. *MySQL* dapat dijalankan dibanyak platform, Saat ini dapat dijalankan di Linux, Windows, Mac, solaris, OS2.

2. Banyak API bahasa pemrograman yang terintegrasi dengan *MySQL*. *MySQL* merupakan database sistem yang memiliki performa yang baik, sehingga banyak bahasa pemrograman yang memiliki API untuk mengintegrasikan *MySQL* untuk dipakai sebagai database. Berikut adalah contoh - contohnya: C, C++, PHP, Java, Perl, Phyton, Tcl, Ruby, dan lain-lain.

3. Performa dapat dioptimasikan dengan arsitektur yang dimiliki *MySQL*, pengguna dapat melakukan konfigurasi *MySQL* untuk mendapatkan performa yang optimal.

4. *MySQL* merupakan DBMS yang open-sources& freeware, Perbaikan error & pengembangan lebih mudah karena adanya forum yang aktif.

5. *MySQL* irit resources. *MySQL* menggunakan engine default bernama MyISAM yang menggunakan sedikit disk space, memory, dan CPU usage.

Kekurangan *MySQL* adalah:

1. *MySQL* tidak memiliki fitur - fitur seperti Analysis Services, Integration Services dan service - service lainnya yang disediakan dari Database System yang berbayar.

2. *MySQL* tidak memiliki algoritma join selengkap Oracle, SQL Server, dan PostgreSQL.

## **2.9 XAMPP**

*XAMPP* adalah software open source berbasis web server yang berisi berbagai program. Aplikasi ini mendukung berbagai sistem operasi seperti Linux, Windows, MacOS, dan Solaris. Fungsi *XAMPP* adalah sebagai server lokal/localhost, di dalamnya sudah mencakup program Apache, *MySQL* dan *PHP*. Kemunculan *XAMPP* diawali dengan adanya kesulitan dalam menginstall Apache dan jika akan menambahkan dukungan *PHP* dan *MySQL*. Hal ini kemudian menjadikan munculnya *XAMPP*, sebagai aplikasi untuk mempermudah developer yang membutuhkan web server di localhost hanya dengan satu aplikasi.*XAMPP* sudah berdiri selama lebih dari 10 tahun, sehingga komunitas pengembangnya sudah banyak. Jika mengalami kendala terkait *XAMPP*, kamu bisa gabung di komunitas *XAMPP* untuk mencari solusinya. Komunitas tersebut bernama [Apache Friends Forums](https://community.apachefriends.org/f).

## **2.10 Unified Modeling Language (UML)**

Bahasa pemodelan visual yang dipakai buat mengespesifikasikan, memvisualkan menciptakan serta mengabadikan desain sistem sebuah perangkat lunak lunak ialah *Unified Modeling Language* (*UML*). Pemodelan ini memberikan keterangan yang nyata mengenai bentuk serta manfaat sistem yang akan diciptakan.*UML* bisa diaplikasikan pada segala langkah pengembangan, daur sistem, serta bermacam ranah aplikasi. Di dalam *UML* ada rancangan semantik, catatan serta tiap tipe bagan, *UML* pula melingkupi aspek-aspek statis, energik ruang lingkup, serta organisasional. Tujuan penting *UML* ialah menyatukan cara pemodelan mengarah sasaran jadi satu standar. Sebagian bagan *UML* yang biasa dipakai mencakup Bagan *Use Case*, Bagan Keaktifan Bagan Antrean serta Bagian Kategori.

### **2.10.1 Use Case Diagram**

*Use Case Diagram* adalah jenis diagram *UML* yang menggambarkan fungsi, ruang lingkup, dan interaksi pengguna dengan system tersebut. Dibawah ini merupakan penjelasan atribut yang digunakan pada *use case diagram*.

Tabel 2. 1 Atribut Use Case Diagram

|  |  |
| --- | --- |
| **Simbol** | **Keterangan** |
|  | *Use case* merupakan teknik untuk melukiskan fungsionalitas sistem sebagai unit-unit yang berhubungan dengan aktor, yang diterangkan melalui penerapan kata-kata. |
|  | Actor, ataupun Aktor, adalah representasi abstrak dari individu maupun sistem lain yang menyebabkan manfaat dalam sistem yang dituju. Buat mengenali aktor, harus didefinisikan alokasi tugas serta kewajiban yang terkait  dengan kondisi sistem yang dituju. |
|  | Relasi asosiasi adalah relasai yang di gunakan buat membuktikan ikatan antara actor serta use case*.* |
|  | Asosiasi antara aktor serta use case dalam skema *use* *case* dicirikan dengan panah terbuka, menandakan kalau aktor berkorelasi dengan cara pasif dengan sistem. |
| <<include>> | include dalam use case yaitu selagi satu *use case* (required) termasuk maupun memanggil use case lain, mendekati dengan pemanggilan dengan manfaat dalam  satu buah program. |
| <<extend>> | Extend, ialah perluasan dari *use case* lain bila situasi ataupun kondisi terlaksana. |

### **2.10.2 Actifity Diagram**

*Actifity* *diagram* adalah diagram yang dapat memodelkan proses-proses yang terjadi pada sebuah sistem. Runtutan proses dari suatu sistem digambarkan secara vertikal. *Actifity* merupakan pengembangan dari *Use Case* yang memiliki alur aktifitas. Dibawah ini merupakan penjelasan atribut yang digunakan pada *Actifity* *diagram*.

Tabel 2. 2 Atribut Actifity Diagram

|  |  |
| --- | --- |
| **Gambar** | **Keterangan** |
|  | Start Point ditempatkan di pojok kiri atas dengan menandai tahapan pertama dari kegiatan yang  terjadi. |
|  | End Point menandai titik akhir dari keaktifan yang  lagi terjadi. |
|  | Activities, menjelaskan sebuah proses/kegiatan bidang usaha. |
|  | Fork (percabangan), dipakai buat membuktikan aktivitas yang dijalani dengan cara paralel ataupun buat menyatukan dua aktivitas paralel jadi satu. |
|  | Join (pencampuran) ataupun rake, dipakai buat membuktikan terdapatnya penguraian. |
|  | *Decision Points*, mengilustrasikan alternatif buat pengumpulan ketentuan true ataupun false. |
|  | Swimlane, penghitungan activity skema buat memperlihatkan siapa melaksanakan apa. |

### **2.10.3 Squense Diagram**

*Squense diagram* adalah sebuah diagram yang digunakan untuk menjelaskan dan menampilkan interaksi antar objek-objek yang ada dalam sebuah Sistem terperinci. Selain itu squense diagram juga menampilkan pesan atau perintah yang dikirim. Dibawah ini merupakan penjelasan dari atribut yang digunakan pada *squense diagram*.

Tabel 2. 3 Atribut Squense Diagram

|  |  |  |  |
| --- | --- | --- | --- |
| **Gambar** | | | **Keterangan** |
|  | | | Entity Class ialah bagian dalam sistem yang terdiri dari kelompok kategori yang mewakili entitas-entitas yang menciptakan kerangka awal sistem serta sebagai  dasar untuk mendesain basis data. |
|  | | | Boundary Class yaitu bagian sistem yang terdiri dari kelompok kategori yang berperan sebagai antarmuka ataupun penghubung antara satu ataupun lebih aktor dengan sistem, sesuai bentuk form input serta form  cetak. |
|  | | | Dalam kondisi sistem, Control Class yaitu suatu  target yang mempunyai logika aplikasi serta tidak mempunyai peranan kepada entitas-entitas lainnya. |
|  | | | Message, lambang yang digunakan untuk mengirim  perintah antar class. |
|  |  |  | Recursive mendeskripsikan metode pengiriman pesan yang dikirim balik terhadap dirinya sendiri. |
|  |  |
|  | | |
|  | | | Activation dipakai buat menggantikan sistem eksekusi aktivitas dari target serta panjang kotak tersebut berkorelasi dengan waktu aktivasi aktivitas  itu. |
|  | | | Lifeline ialah garis yang terdiri dari titik-titik yang tersambung dengan target serta di sepanjang lifeline itu ada activation yang melukiskan kegiatan target  itu. |

### **2.10.4 Class Diagram**

*Class* *diagram* adalah salah satu jenis diagram berbentuk struktur pada model UML. Diagram ini menggambarkan struktur, atribut, kelas, hubungan dan metode dengan sangat jelas dari setiap objeknya. Dibawah ini merupakan penjelasan atribut yang digunakan pada *class* *diagram.*

Tabel 2. 4 Atribut Class Diagram

|  |  |
| --- | --- |
| **Gambar** | **Keterangan** |
| Description: fs | *Nary Association*, usaha buat menjauhi federasi  lebih dari 2 target. |
| Description: C:\Users\ACER\AppData\Local\Microsoft\Windows\INetCache\Content.Word\garis.png | Generalization, ikatan dimana sasaran anak beragam prilaku serta bentuk informasi dari target  yang terlihat di atasnya. |
| Description: af | Class, gabungan dari objek-objek yang memberi atribut serta aktivitas yang sesuai. |
| Description: C:\Users\ACER\AppData\Local\Microsoft\Windows\INetCache\Content.Word\Sc.png | Collaboration, gambaran dari deretan aksi yang diperlihatkan sistem yang membuahkan sesuatu  hasil yang terukur buat sesuatu aktor. |
| Description: ad | Realization, aktivitas yang benar-benar dijalani oleh sesuatu target. |
| Description: fa | Dependency, ikatan dimana perubahan yang terjadi pada sebuah komponen mandiri akan mempengaruhi komponen yang bergantung  padanya komponen yang tidak mandiri. |
| Description: aaa | Association, merupakan apa yang menyambung  antara target satu dengan target lainnya. |

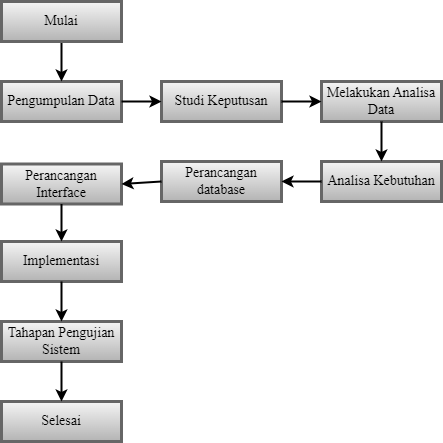
# **BAB III**

# **METODELOGI PENELITIAN**

## **3.1 Waktu dan Tempat Pelaksanaan Penelitian**

Peneliti melakukan penelitian ini dimulai pada bulan Februari. Lokasi penelitian yang digunakan untuk melakukan pengumpulan data berada di Kabupaten Padang Lawas Utara pada Kantor Badan Pusat Statistik.

## **3.2 Metode Penelitian**

****

Gambar 3. 1 Langkah Penelitian

Dalam Penelitian ini sudah disusun berbagai langkah-langkah penelitian yang nantinya akan dilakukan secara sistematis. Berikut adalah penjelasan dari langkah penelitian yang dilakukan:

1. Pengumpulan Data

Pada tahapan ini peneliti melakukan pengumpulan data kecelakaan Kabupaten Padang Lawas Utara, data yang dikumpulkan yaitu data kriteria kecelakaan seperti data jumlah kecelakaan, jumlah meninggal, jumlah luka berat, dan jumlah luka ringan.

2. Studi Keputusan

Tahapan ini peneliti melakukan pengumpulan-pengumpulan informasi yang dilakukan oleh peneliti terdalu yang bertujuan untuk menemukan suatu masalah untuk diteliti.

3. Analisis Data

Dari data yang telah dikumpulkan akan dilakukan pengolahan data dengan menggunakan Algoritma *K-Means.*

4. Analisa Kebutuhan

Pada tahapan ini peneliti melakukan tahapan atau proses dalam pencarian kebutuhan-kebutuhan yang diperlukan dalam merancang sebuah sistem.

5. Perancangan Database

Melakukan perancangan database pada data kecelakaan di daerah kabupaten Padang Lawas Utara menggunakan *MySQL.*

6. Perancangan Interface

Peneliti melakukan perancangan *Web* atau Aplikasi menggunakan Bahasa pemrograman *PHP*.

7. Implementasi

Hasil pengelompokan daerah rawan kecelakaan di Kabupaten Padang Lawas Utara dengan menggunakan Algoritma *K-Means* di implementasikan ke dalam pemrograman.

8. Pengujian Sistem

Pada tahapan ini peneliti melakukan pengujian terhadap sistem yang telah dibuat, yang bertujuan untuk mengetahui apakah sistem sudah berjalan atau sesuai dengan yang diinginkan.

## **3.3 Analisa Kebutuhan Sistem**

Analisa Kebutuhan Sistem merupakan tahap menganalisa sistem yang akan dilakukan atau dibangun. Tahapan kebutuhan sistem juga juga merupakan tahap penting dalam memahami dan menentukan persyaratan untuk pengembangan suatu sistem. Setelah Analisa didapatkan, maka langkah selanjutnya adalah membuat membuat sebuah hasil analisis. Hasil analisis tersebut akan menjadi acuan dari perancangan sistem yang dibangun. Dalam penelitian ini, perlu dipertimbangkan alat dan bahan penelitian. Alat penelitian tersebut ialah perangkat keras (*hardware*) dan perangkat lunak (*software*).

### **3.3.1 Perangkat Keras (*hardware*)**

Perangkat keras (*hardware*) adalah segala komponen yang ada pada ystemn yang sifatnya dapat dilihat dengan secara kasat mata dan diraba secara langsung atau berbentuk nyata. Spesifikasi perangkat keras (*hardware*) dilakukan bertujuan agar sebuah sistem berjalan baik. Perangkat keras yang digunakan pada pembuatan perancangan sistem ini adalah sebagai berikut:

1. Laptop LENOVO ThinkPad T450.

2. IntelI CoreI i5-5300U CPU @ 2.30GHZ, 2295 Mhz.

3. RAM 8,00 GB.

### **3.3.2 Perangkat Lunak (*software*)**

Perangkat Lunak (*softwar*e) adalah sekumpulan data elektronik yang disimpan dan dikendalikan oleh perangkat sistem. Perangkat lunak juga merupakan langkah awal dalam merancang dan menghasilkan suatu perangkat. Berikut adalah spesifikasi umum perangkat lunak yang digunakan dalam perancangan sistem ini:

1. Sistem Operasi : Microsoft Windows 10 pro.

2. Web Browser : Google Chrome.

3. Local Server : XAMPP Control Panel v3.3.0

4. Text Editor : Visual Studio Code

5. DBMS : MySQL.

### **3.3.3 Analisis Kebutuhan Sistem**

Kebutuhan Input merupakan salah satu data yang akan dimasukkan ke dalam sistem oleh admin. Selanjutnya data ini akan di olah supaya menjadi sebuah informasi dengan menggunakan sistem. Dalam sistem informasi pemetaan Kecelakaan terdapat beberapa kebutuhan input yaitu:

a. Peta Kabupaten Padang Lawas Utara

Peta Kabupaten Padang Lawas Utara merupakan sebuah data yang berisi tentang informasi wilayah Kabupaten Padang Lawas Utara.

b. Data Kecelakaan

Data Kecelakaan merupakan data yang dimasukan berupa suatu rawan kecelakaan yang didalamnya berisikan tahun dan lokasi kejadian beserta kriteria jenis kecelakaan.

c. Data Kriteria Kecelakaan

Data Kriteria Kecelakaan yaitu data yang dimasukan berupa kriteria dari kecelakaan yaitu jumlah kecelakaan, jumlah meninggal, jumlah luka berat, dan jumlah luka ringan.

d. Data Tahun

Data Tahun merupakan data yang dimasukan berupa angka tahun.

e. Data Desa

Data Desa merupakan data yang dimasukan berupa nama dari desa.

### **3.3.4 Analisis Kebutuhan Output Sistem**

Adapun gambaran output yang akan dilakukan pada sistem adalah sebagai berikut:

a. Informasi Peta Kabupaten Padang Lawas Utara

Sistem ini akan memberikan informasi tentang gambaran peta Kabupaten Padang Lawas Utara yang ditampilkan di dalam sistem dalam batas wilayah desa.

b. Informasi Kecelakaan

Sistem ini akan memberikan informasi lokasi rawan terjadi Kecelakaan yang terjadi berdasarkan data yang telah dimasukkan kedalam sistem pada setiap desa.

c. Informasi Jumlah Kecelakaan

Sistem ini akan memberikan informasi tentang jumlah korban kecelakaan di Kabupaten Padang Lawas Utara dalam 1 tahun.

d. Jumlah Meninggal

Sistem ini akan memberikan informasi tentang jumlah korban meninggal kecelakaan di Kabupaten Padang Lawas Utara dalam 1 tahun.

e. Jumlah Luka Berat

Sistem ini akan memberikan informasi tentang jumlah korban luka berat kecelakaan di Kabupaten Padang Lawas Utara dalam 1 tahun.

f. Jumlah Luka Ringan

Sistem ini akan memberikan informasi tentang jumlah korban luka ringan kecelakaan di Kabupaten Padang Lawas Utara dalam 1 tahun.

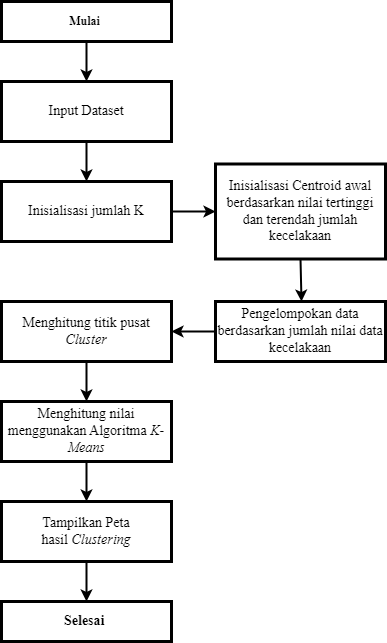
g. Informasi *Clustering* Data

Sistem ini akan memberikan informasi tentang hasil klasterisasi data. Informasi yang akan ditampilkan yaitu rawan, sangat rawan, dan tidak rawan yang telah dihitung menggunakan Algoritma *k-Means.*

## **3.4 Skema**

Skema sistem merupakan sistem dari beberapa elemen yang saling berhubungan satu dengan yang lainnya , sehingga dapat dikatakan sebagai sistem. Sistem terdiri dari kesatuan yang utuh dari beberapa bagian yang saling berhubungan dan saling berinteraksi sehingga mencapai tujuan tertentu. Dari alur skema sistem ini dapat dijelaskan bahwa *Clustering* menggunakan Algoritma *K-Means* yang bertujuan untuk mengelompokan data yang berkarakteristik sama dalam satu cluster dan data yang berkarakteristik berbeda ke dalam cluster yang lain.

Adapun skema sistem daerah rawan kecelakaan menggunakan Algoritma *K-Means Clustering* adalah sebagai berikut:



Gambar 3. 2 Skema Sistem

Keterangan:

1. Input Dataset, menginput Dataset sebagai bahan yang akan digunakan dalam pengolahan data.

2. Inisialisasi Jumlah K, menentukan berapa cluster (K) data tersebut akan dikelompokan.

3. Inisialisasi Centroid awal berdasarkan nilai tertinggi dan terendah jumlah kecelakaan.

4. Pengelompokan data berdasarkan jumlah nilai data kecelakaan, pengelompokan terdiri dari 4 kriteria yaitu jumlah kecelakaan, jumlah meninggal, jumlah luka berat, dan jumlah luka ringan.

5. Menghitung titik pusat cluster, Dimana berdasarkan jumlah yang sering terjadi kecelakaan.

6. Menghitung nilai menggunakan Algoritma K-*Means.*

7. Tampilkan Peta hasil *Clustering*, pengelompokan clustering menjadi 3 kriteria yaitu sangat rawan, rawan, dan tidak rawan.

# **DAFTAR PUSTAKA**

Damanik, Yeusi Fani Selfrida Yani, Sumarno Sumarno, Indra Gunawan, Dedy Hartama, and Ika Okta Kirana. 2021. “Penerapan Data Mining Untuk Pengelompokan Penyebaran Covid-19 Di Sumatera Utara Menggunakan Algoritma K-Means.” *Jurnal Ilmu Komputer Dan Informatika* 1 (2): 109–32. https://doi.org/10.54082/jiki.13.

Harahap, Lina Mardiana, Wahyu Fuadi, Lidya Rosnita, Eva Darnila, and Rini Meiyanti. 2022. “Klastering Sayuran Unggulan Menggunakan Algoritma K-Means.” *Jurnal Teknik Informatika Dan Sistem Informasi* 8 (3): 567–79. https://doi.org/10.28932/jutisi.v8i3.5277.

Heryana, Ade. 2020. “Pengertian Dan Jenis Bencana.” *Researchgate.Net*, no. January: 1–4. https://www.researchgate.net/publication/338537206\_Pengertian\_dan\_Jenis\_Bencana.

Hidayat, Ilsa, Eva Darnila, and Yesy Afrillia. 2023. “Clustering Zonasi Daerah Rawan Bencana Alam Di Kabupaten Mandailing Natal Menggunakan Algoritma K-Means.” *G-Tech: Jurnal Teknologi Terapan* 7 (3): 1218–26. https://doi.org/10.33379/gtech.v7i3.2880.

Jannah, Misbahul, Muthmainnah Muthmainnah, Safwandi Safwandi, Mochamad Ari Saptari, Muhammad Muhammad, Rahmad Wahyudi, and Mirza Farhan. 2022. “Implementation of Geographic Information System for Tourist Locations and Lodging Services in Lhokseumawe City Based on Android.” *International Journal of Engineering, Science and Information Technology* 2 (4): 39–47. https://doi.org/10.52088/ijesty.v2i4.320.

Kurniawati, Ummi Fadlilah, KDM Erli Handayeni, Siti Nurlaela, Hertiari Idajati, Fendy Firmansyah, Nursakti Adhi Pratomoadmojo, and Riswan Sianturi Septriadi. 2020. “Pengolahan Data Berbasis Sistem Informasi Geografis (SIG) Di Kecamatan Sukolilo.” *Sewagati* 4 (3): 190. https://doi.org/10.12962/j26139960.v4i3.8048.

Nalendra, Adimas Ketut, M. Mujiono, Rafika Akhsani, and Adiguna Sasama Wahyu Utama. 2020. “Implementasi Algoritma K-Mean Dalam Pengelompokan Data Kecelakaan Di Kabupaten Kediri.” *VOCATECH: Vocational Education and Technology Journal* 1 (2): 53–60. https://doi.org/10.38038/vocatech.v1i2.28.

Risawandi, Risawandi, and Yesy Afrillia. 2022. “Geographic Information System Mapping Of Criminality Villed Areas In Lhokseumawe Using K-Means Method.” *Journal of Informatics and Telecommunication Engineering* 5 (2): 442–51. https://doi.org/10.31289/jite.v5i2.6265.

Samah, K. A.F.A., S. Ibrahim, N. Ghazali, M. Suffian, M. Mansor, and W. A. Latif. 2020. “Mapping a Hospital Using Openstreetmap and Graphhopper: A Navigation System.” *Bulletin of Electrical Engineering and Informatics* 9 (2): 661–68. https://doi.org/10.11591/eei.v9i2.2082.

Yang, Jin, Yuanjie Li, Qingqing Liu, Li Li, Aozi Feng, Tianyi Wang, Shuai Zheng, Anding Xu, and Jun Lyu. 2020. “Brief Introduction of Medical Database and Data Mining Technology in Big Data Era.” *Journal of Evidence-Based Medicine* 13 (1): 57–69. https://doi.org/10.1111/jebm.12373.