

PENERAPAN K-MEANS CLUSTERING UNTUK PEMETAAN WILAYAH RAWAN BENCANA ALAM KOTA MALANG

I Kadek Riski Dwi Putra, Ahmad Faisol, Mira Orisa
Teknik Informatika, Institut Teknologi Nasional Malang
Jalan Raya Karanglo km 2 Malang, Indonesia
2018069@scholar.itn.ac.id

ABSTRAK

Bencana alam merupakan salah satu fenomena alam yang membahayakan dan dapat mengancam kelangsungan hidup manusia. Pemetaan wilayah rawan bencana yang dilakukan oleh BPBD Kota Malang dapat menjadi jawaban dari permasalahan tersebut. Namun pada halaman resmi website BPBD Kota Malang belum memadainya pengelompokan data bencana. Dalam ilmu *science* dan teknologi telah banyak permasalahan yang dapat diselesaikan oleh disiplin ilmu data mining untuk mengklasterkan data tertentu. Penggunaan Algoritma *K-Means Clustering* salah satu metode analisis data yang bisa digunakan dalam pengelompokan wilayah berdasarkan karakteristik yang mirip. Tujuan pengelompokan tersebut diharapkan dapat membantu tugas BPBD Kota Malang untuk mengidentifikasi wilayah dengan tingkat resiko yang sama. Menurut hasil pengujian menggunakan metode K-Means Clustering dalam pemetaan wilayah bencana di Kota Malang pada sistem dan dibandingkan dengan perhitungan manual memiliki tingkat akurasi sebesar 77,19%. Pengembangan pada aplikasi untuk perhitungan K-Means Clustering data ditambahkan untuk perhitungan untuk pemilihan optimasi pada pusat centroid dengan tujuan untuk menetapkan titik pusat agar jarak yang dihitung tidak berubah - ubah.

Kata kunci : *Pengelompokan, Pemetaan, K-Means Clustering, BPBD, Bencana Alam, Cluster*

1. PENDAHULUAN

Bencana alam merupakan salah satu fenomena alam yang membahayakan dan dapat mengancam kelangsungan hidup manusia. Kondisi tersebut membuat Indonesia dilanda berbagai jenis bencana alam seperti bencana banjir, tanah longsor, banjir, cuaca ekstrem, kekeringan, gempa bumi, tsunami dan gunung meletus. Selain itu tak jarang juga Indonesia kerap kali dilanda berbagai jenis bencana yang terjadi akibat manusia seperti pencemaran udara dan kebakaran hutan. Tidak hanya itu, situasi topografi di Indonesia dengan banyaknya daerah pegunungan menjadi faktor terjadinya bencana alam. Oleh karena itu penting adanya perhatian khusus yang harus segera dirancang dan diaplikasikan sebagai bentuk tindak pencegahan kerugian lingkungan sekitar. Pemetaan wilayah rawan bencana yang dilakukan oleh BPBD Kota Malang dapat menjadi salah satu jawaban dari permasalahan tersebut. Namun pada halaman resmi website BPBD Kota Malang belum memadainya pengelompokan data bencana yang dapat membantu masyarakat untuk mempersiapkan hal dalam menghadapi suatu bencana kedepannya.

Dalam ilmu *science* dan teknologi telah banyak permasalahan yang dapat diselesaikan oleh disiplin ilmu data mining untuk mengklasterkan data tertentu. Seperti penelitian yang dilakukan oleh Rif'atul Maliah yang berjudul "Pengelompokan Data Bencana Alam Berdasarkan Wilayah Menggunakan Algoritma K-Means" pada penelitian tersebut peneliti menggunakan Algoritma K-Means Clustering yang merupakan salah satu metode analisis data yang bisa digunakan dalam pengelompokan wilayah berdasarkan karakteristik yang mirip. Dengan penerapan metode ini, diharapkan wilayah yang memiliki karakteristik

serupa dalam hal risiko bencana alam dapat dikelompokkan dalam satu kelompok. Tujuan pengelompokan tersebut diharapkan dapat membantu tugas BPBD Kota Malang untuk mengidentifikasi wilayah dengan tingkat resiko yang sama untuk segera memberikan bantuan ataupun penanggulangan masalah yang terjadi. Pembagian karakteristik kelompok bencana alam berdasarkan tingkat kawasannya pun dapat dibantu melalui metode ini.

Penelitian ini menerapkan K-Means Clustering untuk pemetaan wilayah rawan bencana di Kota Malang dapat membantu dalam pengembangan sistem informasi yang lebih baik khususnya bagi website BPBD Kota Malang. Adanya sistem informasi pemetaan wilayah rawan bencana ini diharapkan dapat membantu untuk membaca pola – pola tertentu untuk pemetaan kelompok daerah rawan bencana alam sehingga dapat dikelompokkan dan dipetakan menjadi lebih mudah sehingga dapat dilakukannya peringatan dini maupun evakuasi lebih cepat dan akurat. Pemetaan dan pengelompokan ini sangat penting untuk dilakukan guna mengurangi efek dari kejadian bencana alam kepada penduduk serta dapat menjadi acuan dalam penentuan untuk mengambil keputusan di kemudian harinya. Penggunaan K-Means Clustering untuk pemetaan wilayah rawan bencana di Kota Malang ditujukan dapat menambah kesiapan komunitas dalam menghadapi bencana alam.

2. TINJAUAN PUSTAKA

2.1. Penelitian Terdahulu

Pada penelitian yang berjudul "Pengelompokan Data Bencana Alam berdasarkan Wilayah Menggunakan Algoritma K-Means", yang dilakukan oleh Rif'atul Amaliah dan rekannya dengan tujuan yaitu untuk

memberikan hasil informasi dalam bentuk pola pengelompokan daerah rawan bencana yang dimulai dari tahun 2021-2022 dimana akan dijadikan sebagai acuan dalam pengambilan keputusan pihak-pihak terkait untuk mengidentifikasi berbagai jenis wilayah bencana alam paling sering terjadi.[1]

Menurut Tutut Suryani, Ahmad Faisol, dan Nurlaily Vendyansyah dalam penelitiannya yang berjudul “Sistem Informasi Geografis Pemetaan Kerusakan Jalan di Kota Malang Menggunakan Metode K-Means” yang bertujuan membuat SIG atau sistem informasi geografis untuk pemetaan kerusakan pada jalanan yang ditujukan untuk pemerintah agar dapat dipetakan dan dikelompokkan dengan adanya tingkat kerusakan jalan yang terjadi di Kabupaten Malang. [2]

Penelitian yang dilakukan oleh Edi Wahyudin, Rizki Amir Rudin, dan rekan – rekannya yang berjudul “Penerapan Data Mining Pengelompokan Produktivitas Padi Menggunakan Algoritma K-Means Pada Provinsi Jawa Barat”, bertujuan untuk memberikan wawasan terkait daerah tertentu. Informasi ini nantinya tidak hanya bermanfaat bagi pemerintah, namun disini lain dapat juga sebagai acuan untuk penelitian masa depan dalam memperhatikan potensi produktivitas padi [3].

2.2. Data Mining

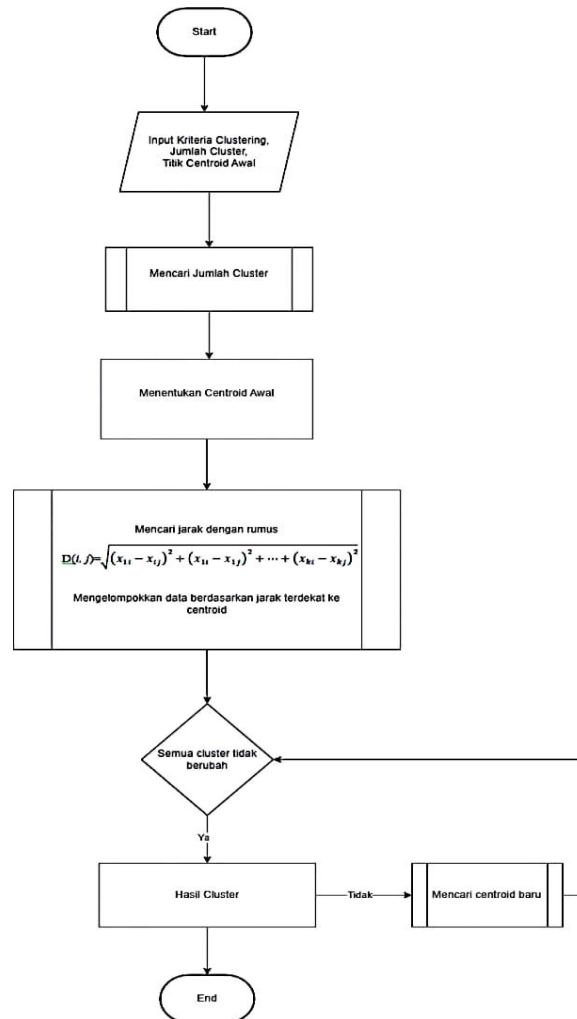
Data mining adalah salah satu bidang ilmu computer yang telah banyak digunakan dalam menyelesaikan berbagai masalah dalam penelitian. Ada beberapa macam algoritma dalam data mining diantaranya K-Means, KMedoid, C4.5, Naïve Bayes, Regresi Linier dan Lainnya. Terdapat beberapa istilah dalam data mining seperti pattern recognition yang artinya data mining dapat melakukan pengelompokan pola – pola pada data yang bersifat tersembunyi pada suatu data besar atau kompleks.[4]

Secara sederhana data mining adalah proses mengolah informasi agar berguna dari data yang besar dan kompleks. Tujuan dari data mining adalah menemukan informasi yang dapat digunakan oleh pengguna. Penggunaan data mining ini banyak digunakan di berbagai bidang lain seperti teknologi informasi, bisnis, ilmu pengetahuan, dan sebagainya untuk mendapatkan informasi pada data yang didapatkan untuk menentukan keputusan yang baik dan tepat.[5]

2.3. Algoritma K-Means

K-means Clustering adalah strategi mengolah data dengan data ke dalam kelompok-kelompok tertentu. Prinsip dasarnya adalah mengelompokkan data sedemikian rupa dengan data yang mempunyai karakteristik yang serupa dan data dengan karakteristik tidak sama dapat dikelompokkan dalam kelompok lain.[6]

Diagram *flowchart* metode pada aplikasi pemetaan daerah yang rentan terhadap bencana alam ditampilkan dalam gambar 1.



Gambar 1. Flowchart Perhitungan K-Means Clustering

Pada diagram 1, terdapat diagram *flowchart* perhitungan K-Means Clustering yang menjelaskan proses dari algoritma *K-Means Clustering*. Langkah awal melibatkan masukan kriteria *clustering*, jumlah *cluster*, dan penentuan titik awal *centroid* untuk digunakan dalam perhitungan. Setelah itu, proses melibatkan penentuan jumlah *cluster* dan penentuan titik awal *centroid*. Selanjutnya, data dikelompokkan berdasarkan jarak terdekat ke *centroid* dengan menggunakan perhitungan sesuai rumus yang telah ditetapkan. Jika tidak ada perubahan pada seluruh *cluster*, hasil *cluster* akan ditentukan. Namun, jika ada data yang berubah, proses akan kembali mencari *centroid* baru.

Tahap – tahap untuk melakukan proses pengelompokan data menggunakan *K-Means Clustering* sebagai berikut:

- Menentukan jumlah *cluster* (k) atau kelompok yang akan dibuat.
- Melakukan permissalan *cluster* (k) sebagai pusat *cluster*. Permissalan ini nantinya dapat dihitung beberapa macam metode, pada perhitungan k-means clustering pemilihan pusat *cluster* dengan pemilihan secara random atau acak. Pada pusat

cluster akan diberikan nilai awal Dimana akan dipilih secara random atau acak.

- c. Mendistribusikan seluruh objek dan data ke cluster atau kelompok dengan jarak terdekat. Jarak kedekatan antara dua objek atau data ditentukan dengan menghitung jarak antara kedua data atau objek. Cara menghitung jarak dari pusat cluster dengan data atau objek, dapat menggunakan teori menghitung jarak *Euclidean*, dirumuskan sebagai berikut:

$$D(x_i, x_j) = \sqrt{(x_{i1} - x_{j1})^2 + (x_{i2} - x_{j2})^2 + \dots + (x_{ki} - x_{kj})^2} \dots\dots (1)$$

Penjelasan :

$D(x_i, x_j)$ = Jarak data ke I ke usat cluster j

x_{ki} = Data ke i pada atribut ke k

x_{kj} = Data ke j pada atribut ke k

- d. Memperbaharui pusat *cluster* dengan menganalisis banyak anggota *cluster* saat ini. Pusat *cluster* yang baru dihitung sebagai rata – rata pada data atau objek yang mencakup dalam *cluster* tertentu. Jika diinginkan, median dari *cluster* tersebut juga dapat digunakan sebagai alternatif ukuran. Oleh sebab itu, penggunaan perhitungan rata-rata (*mean*) tidaklah satu - satunya metode yang dapat digunakan.
- e. Memperbaharui nilai dari *centroid* dapat diperoleh dengan menghitung rata-rata dari jumlah anggota *cluster* yang bersangkutan, menggunakan rumus :

$$C_{m(q)} = \frac{1}{n_m} \sum_{i=1}^{n_m} x_{i(q)} \dots\dots\dots (2)$$

Penjelasan:

$C_{m(q)}$ = Centroid dari klaster m pada perhitungan ke-q

N_m = Jumlah data yang termasuk dalam klaster m

$X_{i(q)}$ = Data ke-i pada iterasi ke-q

- f. Dilakukan peroses perulangan pada seluruh data dengan menggunakan pusat cluster yang sudah diperbaharui. Apabila pusat cluster tidak mengalami perubahan, maka proses pengelompokan selesai. Jika terjadi perbuahan, lakukan kembali proses langkah nomor tiga dan lakukan terus proses tersebut sampai pusat cluster tidak terdapat perbuahan.

2.4. Website

Website adalah kumpulan dari semua halaman web yang berfungsi untuk menampilkan berbagai informasi dalam bentuk tulisan, gambar, dan suara dari domain yang terbentuk dalam suatu rangkain yang saling terkait. Website juga merupakan kumpulan halaman media informasi dalam suatu domain yang diakses oleh pengguna menggunakan jaringan internet. Suatu halaman web yang sudah terhubung dengan halaman web lainnya disebut dengan hyperlink, sedangkan teks yang terhubung dengan teks lain disebut hypertext[7].

2.5. Database

Database dipergunakan untuk menyimpan data yang terkait dengan suatu sistem atau organisasi, dan memfasilitasi pengguna dalam menarik data sesuai dengan keperluan mereka. Jenis-jenis database juga dapat dikelompokkan, seperti *database* relasional, database berorientasi objek, dan database terdistribusi, bergantung pada cara penyimpanan dan organisasi data dalam sistem[8] .

2.6. PHP

PHP adalah bahasa pemrograman sebagai penerjemahan baris kode sumber menjadi kode mesin yang dimengerti oleh sistem secara langsung ketika kode dijalankan. PHP juga merupakan bahasa pemrograman *open source* yaitu dengan hak cipta terbuka sehingga pengguna dapat mengembangkan kode PHP sesuai dengan kebutuhan[9].

2.7. Laravel

Laravel adalah kerangka kerja yang mendukung pengembang *web* dalam mengoptimalkan penggunaan *PHP* pada tahapan pembangunan situs web.

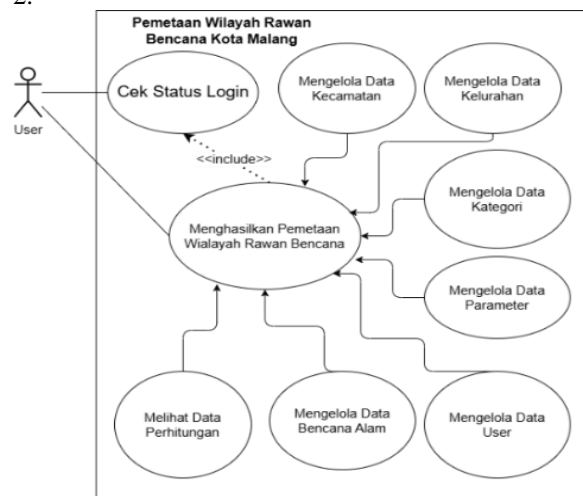
Laravel memungkinkan membuat seluruh proses pengembangan *web* lebih cepat dan lebih mudah bagi pengembang dengan menghilangkan semua masalah yang terkait dengan penanganan kode *PHP* yang rumit. [10]

3. METODE PENELITIAN

Dalam penelitian ini memerlukan perhatian khusus dalam opsi metode penelitian agar mendapatkan hasil yang baik. Pada penelitian ini memanfaatkan metode *K-Means Clustering* Algoritma dalam melakukan pemetaan dan pengelompokan wilayah yang rawan bencana alam di Kota Malang.

3.1. Use Case Diagram

Use Case diagram untuk *website* pemetaan daerah rawan bencana alam dapat dilihat pada gambar 2.

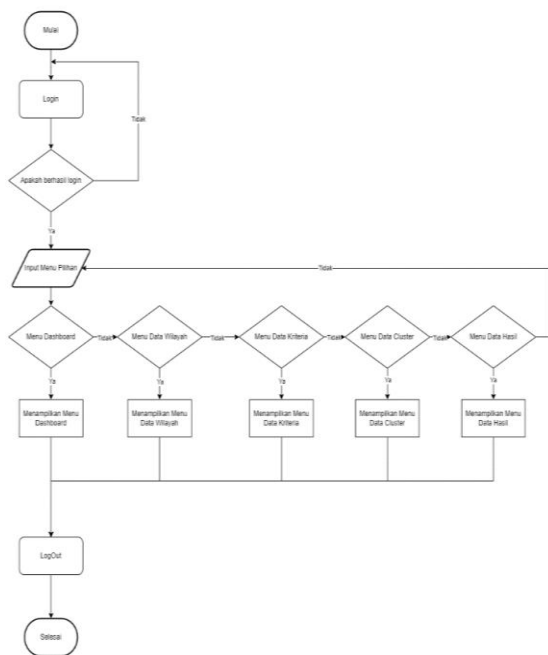


Gambar 2. Use Case Diagram

Pada ilustrasi diatas yaitu dalam gambar 2 terdapat hanya satu pengguna yaitu user atau admin. Admin atau user ini bertanggung jawab untuk mengelola data yang akan digunakan untuk melakukan pemetaan dan pengelompokan dengan metode algoritma K-Means Clustering. User atau admin dapat menambah, mengubah, dan menghapus data yang digunakan dalam sistem seperti data kecamatan, data kelurahan, data kategori, data parameter, dan data bencana.

3.2. Flowchart Sistem

Flowchart sistem yang terdapat pada website pemetaan wilayah rawan bencana alam terdapat pada gambar 3.

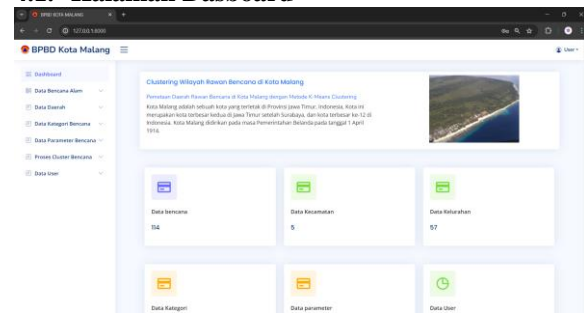


Gambar 3. Flowchart sistem

Pada Gambar 3 Flowchart sistem, terdapat diagram alur yang menggambarkan tahap - tahap menggunakan sistem. Pertama user yaitu admin akan melakukan login di halaman login, apabila user berhasil melakukan login, dilanjutkan ke halaman dashboard. Apabila login tidak berhasil dilakukan, user akan diarahkan kembali pada halaman login. Pada website terdapat beberapa menu, menu untuk melakukan operasi *CRUD* untuk data nama kelurahan yang terkena dampak, nama kecamatan, data kategori bencana, data parameter, data bencana, proses clustering dan pemetaan wilayah rawan bencana. Pada menu clustering ditugaskan untuk melakukan perhitungan pengelompokan daerah rawan bencana alam menggunakan metode K-Means Clustering.

4. HASIL DAN PEMBAHASAN

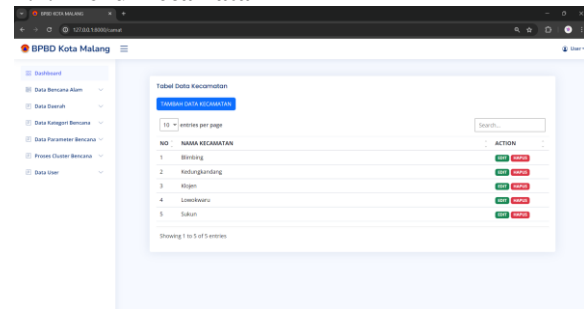
4.1. Halaman Dashboard



Gambar 4. Halaman dashboard

Pada gambar 4 yaitu Halaman dashboard merupakan tampilan halaman dashboard yang sudah dibuat. Pada halaman dashboard terdapat profile Kota Malang, menu data bencana, data kecamatan, data kelurahan, data parameter, data kategori dan data user.

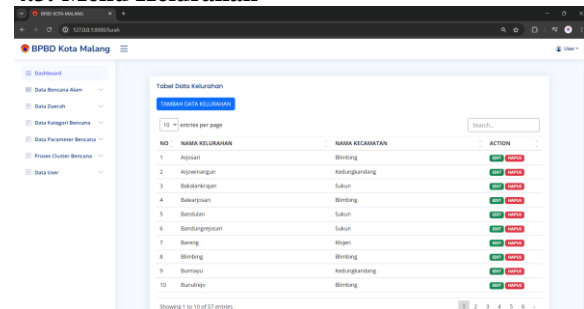
4.2. Menu Kecamatan



Gambar 5. Tampilan Halaman Menu Kecamatan

Pada gambar 5 adalah tampilan halaman data kecamatan yang menunjukkan informasi mengenai data kecamatan. Pada data kecamatan tersebut terdapat data nama kecamatan.

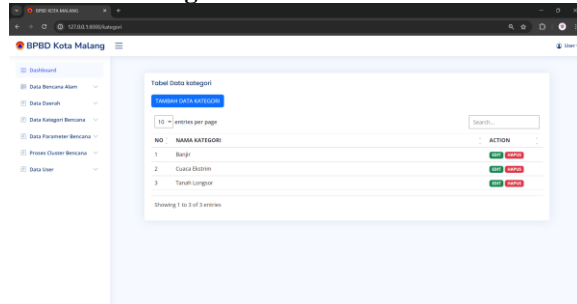
4.3. Menu Kelurahan



Gambar 6. Tampilan Halaman Menu Kelurahan

Pada gambar 6 yaitu tampilan data kelurahan adalah tampilan halaman yang menunjukkan informasi mengenai data kelurahan yang dibuat. Pada tabel data kecamatan tersebut terdapat data seperti nama kelurahan dan nama kecamatan.

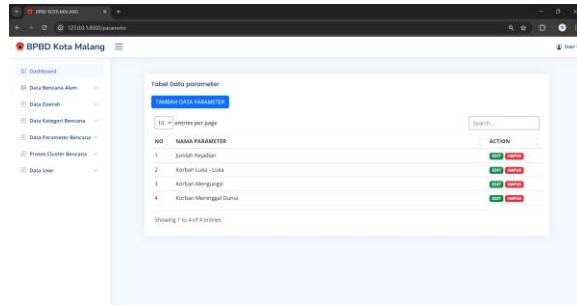
4.4. Menu Kategori Bencana



Gambar 7. Tampilan Halaman Menu Kategori

Pada gambar 7 yaitu tampilan data kategori adalah tampilan data antarmuka yang menunjukkan informasi terhadap data kategori bencana yang dibuat. Pada tabel data kategori tersebut terdapat data layaknya nama bencana yang akan dikelompokkan menggunakan metode K-Means Clustering.

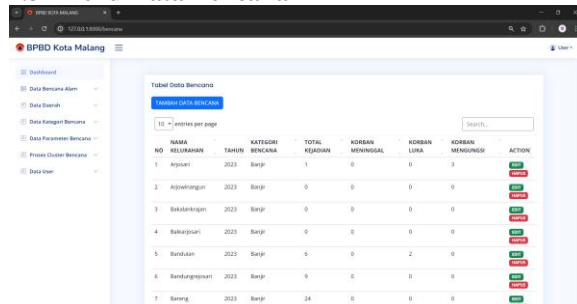
4.5. Menu Parameter Bencana



Gambar 8. Halaman Parameter

Pada gambar 8 yaitu tampilan data parameter adalah tampilan data antarmuka dimana menunjukkan informasi mengenai data parameter bencana yang diproses. Pada tabel data kategori itu terdapat data seperti nama parameter yang akan menjadi parameter dalam pengelompokan.

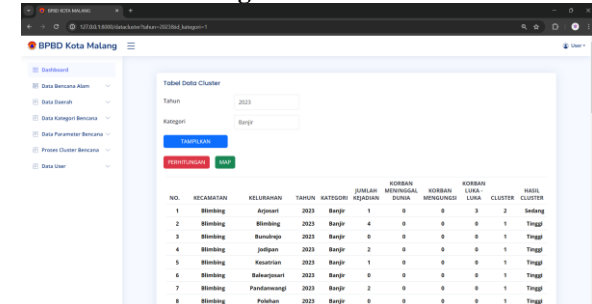
4.3 Menu Data Bencana



Gambar 9. Halaman Data Bencana

Pada gambar 9 yaitu tampilan data bencana adalah tampilan data antarmuka yang menunjukkan informasi perihail data bencana yang dibuat. Pada tabel data bencana tersebut mencakup data seperti nama kelurahan yang terjadi bencana, tahun kejadian, kategori bencana, dan parameter bencana.

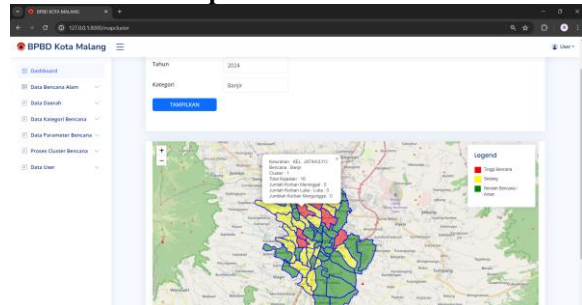
4.4 Menu Clustering



Gambar 10. Halaman Menu Clustering

Pada gambar 10 yaitu tampilan menu clustering adalah tampilan data antarmuka yang menunjukkan proses clustering atau pengelompokan yang dilakukan sistem dengan metode K-Means Clustering. Pada halaman ini terdapat filter yang dapat memilih tahun kejadian bencana dan kategori bencana.

4.5 Halaman Map



Gambar 11. Halaman Map

Pada gambar 11 yaitu tampilan halaman map berisikan peta yang berikan kelompok atau cluster yang sudah dibagi menjadi beberapa cluster atau kelompok dan diberi warna merah sebagai cluster 1, warna kuning sebagai cluster 2, dan warna hijau sebagai cluster 3.

4.6. Implementasi Metode K-Means Clustering

Tabel 1. Input data wilayah

No	Kelurahan	Kecamatan
1	Arjosari	Blimbing
2	Arjowinangun	Kedungkandang
...
57	Wonokoyo	Kedungkandang

Perhitungan nilai proses iterasi pada tahap pertama akan menggunakan perhitungan nilai K-n yaitu dengan melibatkan rumus tertentu. Rumus itu mengarah pada nilai kejadian bencana di suatu daerah jumlah korban yang meninggal, jumlah korban yang terluka dan jumlah korban yang mengungsi yang telah diiniliasasikan dengan kode A, B, C, dan D. Nilai – nilai tersebut digunakan sebagai parameter bencana yang diinputkan ke dalam sistem:

Tabel 2. Parameter bencana

No	(A)	(B)	(C)	(D)
1	0	0	0	0
2	1	0	0	0
....
57	0	0	0	0

Penjelasan parameter diatas:

- Jumlah Kejadian (A): Jumlah kejadian bencana mengacu pada jumlah kejadian bencana yang terjadi pada setiap kelurahan di Kota Malang
- Korban Meninggal (B): Korban Meninggal merujuk pada jumlah korban yang meninggal diakibatkan yang ada pada setiap kelurahan.
- Korban Luka (C): Korban Luka merujuk pada jumlah korban yang terluka diakibatkan bencana alam pada setiap kelurahan.
- Korban Mengungsi (D): Korban Mengungsi merujuk pada jumlah korban yang mengungsi diakibatkan bencana alam pada setiap kelurahan

Berikut merupakan implementasi perhitungan dengan metode *K-Means* Clustering pada aplikasi:

- Program untuk menentukan jumlah *cluster* dan menentukan pusat *cluster* / *centroid* secara acak

```

1 $k = 3;
2 $lengthData = count($this->dataBencana_old);
3 $centroid_data = [];
4 $c1 = Session::get('c1');
5 $c2 = Session::get('c2');
6 $c3 = Session::get('c3');

```

Gambar 11. Source code menentukan jumlah *cluster*

Pada gambar 11 adalah source code yang digunakan menentukan titik *centroid* awal.

- Menentukan Jarak

```

1 $hitungC[$j] = sqrt(pow(($this->dataBencana_old[$i]['data']->nilai_1
2 - $dataCentroid[$j]['nilai_1']), 2) +
3 pow(($this->dataBencana_old[$i]['data']->nilai_2
4 - $dataCentroid[$j]['nilai_2']), 2) +
5 pow(($this->dataBencana_old[$i]['data']->nilai_3
6 - $dataCentroid[$j]['nilai_3']), 2) +
7 pow(($this->dataBencana_old[$i]['data']->nilai_4
8 - $dataCentroid[$j]['nilai_4']), 2));
9

```

Gambar 12. Source code perhitungan Euclidean Distance

Pada gambar 12 adalah source code untuk perhitungan Euclidean Distance yaitu menghitung jarak data dengan pusat cluster.

- Memperbaharui *Cluster*

Pada gambar 13 merupakan tampilan source code untuk memperbaharui data cluster dengan melakukan perulangan sampai dengan anggota setiap cluster tidak terjadinya perubahan cluster.

```

1 $centroid_data = [];
2 for ($i = 0; $i < count($dataCentroid); $i++) {
3     $nilai_1 = 0;
4     $nilai_2 = 0;
5     $nilai_3 = 0;
6     $nilai_4 = 0;
7     $jmlData = 0;
8     for ($j = 0; $j < count($dataBaru); $j++) {
9         // dd($dataBaru[$j]['cluster']);
10        if ($dataBaru[$j]['cluster'] == ($i + 1)) {
11            // $centroid_data[$i][$j] = $dataBaru[$j];
12            $nilai_1 += $dataBaru[$j]['data']->nilai_1;
13            $nilai_2 += $dataBaru[$j]['data']->nilai_2;
14            $nilai_3 += $dataBaru[$j]['data']->nilai_3;
15            $nilai_4 += $dataBaru[$j]['data']->nilai_4;
16            $jmlData += 1;
17        }
18    }
19 }

```

Gambar 13. Source code memperbaharui data *cluster*

Perhitungan Manual K-Means

- Menentukan jumlah *cluster*

Tabel 3. Tabel *cluster*

Cluster	Nama Cluster
C1	Tinggi
C2	Sedang
C3	Rendah

- Menetapkan pusat *cluster* secara acak

Perhitungan yang dilakukan menggunakan excel, nilai k atau jumlah cluster terdapat 3. Setelah menentukan jumlah cluster atau k tahap selanjutnya adalah memilih pusat cluster secara acak atau random dalam perhitungan manual ini menggunakan data ke 5, 8, dan 2 sebagai pusat cluster pertama.

Tabel 4. Tabel Pusat Cluster (*Centroid*)

Data ke	Pusat cluster	A	B	C	D
5	C1	5	0	4	0
8	C2	2	0	0	0
2	C3	1	0	0	0

- Menghitung jarak data dengan pusat cluster

Setelah menentukan pusat cluster awal, selanjutnya dilakukan perhitungan jarak menggunakan rumus Euclidean dan menglompokan data menurut data terdekat. Selanjutnya nilai centroid baru yang berhasil dihasilkan akan digunakan untuk menjadi acuan pada perhitungan selanjutnya, dilakukan perulangan pada perhitungan ini sampai nilai centroid sama pada perhitungan sebelum dan sesudahnya.

- Hasil perhitungan jarak pada titik pusat 1

(5, 0,4,0) :

$D(1,1)$

$$= \sqrt{(0-5)^2 + (0-0)^2 + (0-4)^2 + (0-0)^2} = 6,403124237$$

$D(57,1)$

$$= \sqrt{(0-5)^2 + (0-0)^2 + (0-4)^2 + (0-0)^2} = 6.403124237$$

- Hasil perhitungan jarak pada titik pusat 2

(2, 0,0,0) :

$$D(1,2) = \sqrt{(0-2)^2 + (1-0)^2 + (0-0)^2 + (1-0)^2} = 2$$

$$D(57,2) = \sqrt{(0-2)^2 + (1-0)^2 + (0-0)^2 + (1-0)^2} = 2$$

- 3) Hasil perhitungan jarak pada titik pusat 3.

$$(1, 0, 0, 0) :$$

$$D(1,3) = \sqrt{(0-1)^2 + (0-0)^2 + (0-0)^2 + (1-0)^2} = 1$$

$$D(57,3) = \sqrt{(0-1)^2 + (0-0)^2 + (1-0)^2 + (0-0)^2} = 0$$

Setelah melakukan perhitungan untuk seluruh data dengan menggunakan rumus Euclidean langkah selanjutnya dilakukan pengelompokan data ke dalam cluster berdasarkan jarak terdengannya dengan nilai terkecil ke tiap – tiap cluster. Tahapan berikutnya dilakukan menghitung nilai centroid baru dengan menggunakan persamaan 1.

- d. Menentukan Jarak *Centroid*

Mengumpulkan data menurut jarak terdekat atau nilai terkecil (minimum) ke pusat cluster (*centroid*). Dimana, C1 dikelompokkan dengan cluster 1 (Tinggi), C2 dikelompokkan dengan cluster 2 (Sedang), dan C3 dikelompokkan dengan cluster 3 (Rendah/Tidak Beresiko)

Tabel 5. Tabel Jarak Data

No	Kelurahan	Centroid			Min	C
		C1	C2	C3		
1	Arjosari	6.403124237	2	1	1	3
2	Arjowinangun	5.656854249	1	0	0	3
...		
57	Wonokoyo	6.403124237	2	1	1	3

- e. Pengelompokan seluruh data berdasarkan jarak terdekat dengan pusat cluster.

Tabel 6. Klasifikasi setiap data dengan *centroid*

No	Nama Kelurahan	C1	C2	C3
1	Arjosari			1
2	Arjowinangun			1
...
57	Wonokoyo			1

- f. Memperbaharui Pusat *Cluster*

Menguji apakah terdapat data terjadi perubahan atau tidak. Jika terjadi perubahan dilakukan perulangan dari tahap ketiga hingga tahap kelima. Jika tidak terjadi perubahan nilai pada iterasi terakhir maka tidak dilakukan perulangan. Setelah itu, dilakukan pembaruan nilai *centroid* dengan menghitung rata-rata cluster menggunakan rumus:

$$C_{m(q)} = \frac{1}{n_m} \sum_{i=1}^{n_m} x_{i(q)} \dots\dots\dots (4)$$

Tabel 7. Klasifikasi Berdasarkan Kedekatan Dengan *Centroid*

No	Nama Kelurahan	C1	C2	C3
1	Arjosari			1
2	Arjowinangun			1
...
57	Wonokoyo			1

Setelah mendapatkan nilai *centroid* baru, dilanjutkan pada menghitung nilai jarak *Euclidean* untuk setiap data terhadap *centroid* baru pada iterasi kedua. Selanjutnya, lakukan perhitungan *centroid* baru secara berulang hingga nilai *centroid* baru tidak mengalami perubahan dari nilai *centroid* lama atau iterasi sebelumnya. Berikut ini adalah hasil perhitungan data yang telah dilakukan perhitungan ditampilkan pada tabel 8.

Tabel 8. Hasil *Clustering*

No	C1	C2	C3	Minimum	Cluster
1.	6.9442 222	2.6199 056	0.4601 942	0.460194 202	Clust er3
2.	5.9907 336	1.6226 941	0.6567 274	0.656727 416	Clust er3
...
57.	6.9442 222	2.6199 056	0.4601 942	0.460194 202	Clust er3

4.7. Pengujian Metode *K-Means Clustering*

Pengujian yang dilakukan dalam menilai ukuran Tingkat keakuratan perhitungan metode dengan perhitungan manual dengan cara membandingkan data perhitungan pada program dengan perhitungan manual. Pada pengujian ini peneliti menggunakan 57 data keluarahan pada bencana Cuaca Ekstrem di tahun 2022

Berikut adalah gambaran dari hasil pengelompokan atau clustering pada program dan sistem.

Tabel 9. Data Asli

No	Nama Kelurahan	A	B	C	D	Kelas
5	Kesatrian	2	0	0	0	Cluster 2
14	Buring	5	0	4	0	Cluster 3
16	Kedungkandang	3	0	1	0	Cluster 2
21	Sawojajar	2	0	1	0	Cluster 2
25	Gadingasri	6	0	1	0	Cluster 1
26	Kasin	2	0	0	0	Cluster 2
27	Kauman	3	0	0	0	Cluster 2
29	Klojen	1	0	0	8	Cluster 3
30	Oro – oro dowo	9	0	0	3	Cluster 1
40	Mojolangu	3	0	0	0	Cluster 2
41	Sumbersari	2	0	0	0	Cluster 2
44	Tulusrejo	2	0	0	0	Cluster 2
46	Tunjung sekar	2	0	0	0	Cluster 2

Pada tabel 12 terdapat perbedaan data antara, data perhitungan sistem dengan data perhitungan dengan

excel. Perbedaan ini terletak pada kelurahan buring, gading asri, kasin kauman, kedungkandang, kesian, mojolangu, oro-oro dowo, sawojajar, sumpangsari, tulusrejo, tunjung sekar, dan klojen.

5	Blimbing	Kesatrian	2023	Cuaca Ekstrem	2	0	0	0	3	Rendah
14	Kedungkandang	Buring	2023	Cuaca Ekstrem	5	0	4	0	2	Sedang
16	Kedungkandang	Kedungkandang	2023	Cuaca Ekstrem	3	0	1	0	3	Rendah
21	Kedungkandang	Sawojajar	2023	Cuaca Ekstrem	2	0	1	0	3	Rendah
25	Klojen	Gadingari	2023	Cuaca Ekstrem	6	0	1	0	2	Sedang

Gambar 14. Hasil Perhitungan Sistem

26	Klojen	Kasin	2023	Cuaca Ekstrem	2	0	0	0	3	Rendah
27	Klojen	Kauman	2023	Cuaca Ekstrem	3	0	0	0	3	Rendah
29	Klojen	Klojen	2023	Cuaca Ekstrem	1	0	0	0	1	Tinggi
30	Klojen	Oro-Oro Dowo	2023	Cuaca Ekstrem	9	0	0	3	2	Sedang
40	Lowokwaru	Mojoagung	2023	Cuaca Ekstrem	3	0	0	0	3	Rendah
41	Lowokwaru	Sumpangsari	2023	Cuaca Ekstrem	2	0	0	0	3	Rendah
44	Lowokwaru	Tulusrejo	2023	Cuaca Ekstrem	2	0	0	0	3	Rendah
46	Lowokwaru	Tunjungsekar	2023	Cuaca Ekstrem	2	0	0	0	3	Rendah

Gambar 15. Hasil Perhitungan Sistem

Berdasarkan hasil perbandingan antara kedua data yaitu perhitungan data sistem dengan perhitungan data manual dengan excel dilakukan perhitungan persentase Tingkat akurasi menggunakan persamaan 5.

$$\text{Akurasi} = \frac{\text{Data Yang Sesuai}}{\text{Total Jumlah Data}} \times 100\% \dots (5)$$

Dari 57 data keluarahan yang diuji dengan bencana cuaca ekstrem pada tahun 2022 ditemukan terdapat 13 perbedaan data. Data yang sesuai dengan sistem terdapat 44 data. Berdasarkan data dihitung nilai tingkat akurasi yang terdapat pada persamaan 6.

$$\text{Akurasi} = \frac{44}{57} \times 100\% = 77,19\% \dots (6)$$

5. KESIMPULAN DAN SARAN

Menurut hasil pengujian menggunakan metode K-Means Clustering dalam pemetaan wilayah bencana di Kota Malang pada sistem dan dibandingkan dengan perhitungan manual memiliki tingkat akurasi sebesar 77,19%. Pengembangan pada aplikasi untuk perhitungan K-Means Clustering data ditambahkan untuk perhitungan untuk pemilihan optimasi pada pusat centroid dengan tujuan untuk menetapkan titik pusat agar jarak yang dihitung tidak berubah - ubah.

DAFTAR PUSTAKA

- [1] R. Amaliah, E. Tohidi, E. Wahyudin, A. Rizki Rinaldi, and I. Iin, "Pengelompokan Data Bencana Alam Berdasarkan Wilayah Menggunakan Algoritma K-Means," *JATI (Jurnal Mhs. Tek. Inform., vol. 7, no. 6, pp. 3572–3579, 2024, doi: 10.36040/jati.v7i6.8253.*
- [2] T. Suryani, A. Faisol, and N. Vendyansyah, "Sistem Informasi Geografis Pemetaan Kerusakan Jalan Di Kabupaten Malang Menggunakan Metode K-Means," *JATI (Jurnal Mhs. Tek. Inform., vol. 5, no. 1, pp. 380–388, 2021, doi: 10.36040/jati.v5i1.3259.*
- [3] E. Wahyudin, R. Amir Rudin, and S. Eka Permana, "Penerapan Data Mining Pengelompokan Produktivitas Padi Menggunakan Algoritma K-Means Pada Provinsi Jawa Barat," *J. Mhs. Tek. Inform., vol. 8, no. 1, 2024, [Online]. Available: <https://opendata.jabarprov.go.id/id/dataset/produktivitas>*
- [4] M. Rafi Nahjan, Nono Heryana, and Apriade Voutama, "Implementasi Rapidminer Dengan Metode Clustering K-Means Untuk Analisa Penjualan Pada Toko Oj Cell," *JATI (Jurnal Mhs. Tek. Inform., vol. 7, no. 1, pp. 101–104, 2023, doi: 10.36040/jati.v7i1.6094.*
- [5] A. P. Riani, A. Voutama, and T. Ridwan, "Penerapan K-Means Clustering Dalam Pengelompokan Hasil Belajar Peserta Didik Dengan Metode Elbow," *J-SISKO TECH (Jurnal Teknol. Sist. Inf. dan Sist. Komput. TGD), vol. 6, no. 1, p. 164, 2023, doi: 10.53513/jsk.v6i1.7351.*
- [6] Nurjoko, D. Dwirohayati, and N. H. Sudibyo, "Sistem Informasi Pemetaan Wilayah Rawan Kriminalitas Polresta Bandar Lampung Menggunakan K-Means Clustering," *Teknika, vol. 14, no. 2, pp. 127–135, 2020.*
- [7] T. A. Kinaswara, N. R. Hidayati, and F. Nugrahanti, "Rancang Bangun Aplikasi Inventaris Berbasis Website Pada Kelurahan Bantengan | Kinaswara | Prosiding Seminar Nasional Teknologi Informasi dan Komunikasi (SENATIK)," *Pros. Semin. Nas. Teknol. Inf. dan Komun., vol. 2, no. 1, pp. 71–75, 2019, [Online]. Available: <http://prosiding.unipma.ac.id/index.php/SENATIK/article/view/1073>*
- [8] N. Novita, "Manajemen Proyek Sistem Informasi Pengolahan Data Apotek Berbasis Database," *Methosisfo J. Ilm. Sist. Inf., vol. 2, no. 1, pp. 9–17, 2022.*
- [9] P. A. W. Purnama and T. A. Putra, "Perancangan Sistem Penjualan Berbasis Web (E-Commerce) Pada Toko DMX Factory Outlet Dengan Menggunakan Bahasa Pemrograman PHP-MYSQL Dan Java Script," *Ris. dan E-Jurnal Manaj. Inform. Komput., vol. 5, no. 1, pp. 129–133, 2020.*
- [10] S. Arfan and S. I. G. Novian, "Implementasi Aplikasi Framework Laravel PT XYZ," *J. Tek. Inform. Stmik Antar Bangsa, vol. V, no. 1, pp. 18–24, 2019.*