

CLUSTERING DAERAH RAWAN BANJIR DI KABUPATEN TUBAN DENGAN K-MEANS DISERTAI VISUALISASI SISTEM INFORMASI GEOGRAFIS

Ratih Aprillina¹, Alfian Nurlifa², Andy Haryoko³, Rizki Eka Putri⁴, Alfa Nurfahma Rosalita⁵

* Departement of Informatics, PGRI Rongolawe University
Correspondence Author: lifa.nurlifa13@gmail.com

| Info Artikel : | ABSTRACT (in English) |
|---|--|
| Sejarah Artikel : Menerima : 23 Juli 2022 Revisi : - Diterima : 31 Juli 2022 Online : 31 Juli 2022 Keyword : Data Mining, Clustering, Flood Prone Areas, K- Means, GIS Visualization. | <p>In this research, there are grouping of flood prone areas in Tuban Regency using data mining techniques. This study is based on the lack of information to know and anticipate which areas are potentially flooded. So that the handling due to flood in areas flooded still less than the maximum. Therefore, in optimizing the knowledge of information and anticipation by mapping or classifying potentially flood-prone areas or not in the area of Tuban Regency by clustering method, using data obtained from the BPS-Statistics "Kabupaten Tuban Dalam Angka (Tuban Regency In Figure)" from year 2014 to 2017. K-means algorithm is a method that will be implemented to reach information analysis in grouping flood prone areas with data attributes used are rainfall, soil type and height of soil in every sub-district in Tuban Regency. The result obtained a system to classify flood-prone areas by producing 3 groups of categories, namely prone, less prone and not vulnerable. Result of map visualization with result of year 2013 as many as three colors orange, yellow and green, year 2014 as two color that is yellow and green, year 2015 three color that is orange, yellow and green and year 2016 three color orange, yellow and green. From the research that has been done on 20 equation data is able to produce 45% accuracy testing for 2013, 40% for 2014, 40% for 2015 and 60% for the year 2016. This device can be developed to assist other regions in classifying vulnerable areas flood. This research method can be developed for different research objects.</p> |
| | INTISARI (in Indonesia) |
| Kata Kunci : Data Mining, Clustering, | <p>Pada penelitian ini telah dilakukan pengelompokan daerah rawan banjir di Kabupaten Tuban dengan menggunakan teknik <i>data mining</i>. Penelitian ini didasarkan adanya kurangnya informasi untuk mengetahui dan mengantisipasi daerah mana yang berpotensi banjir. Sehingga penanganan</p> |

| | |
|---|---|
| <p>Daerah Rawan Banjir, K-Means, Visualisasi SIG.</p> | <p>akibat banjir di daerah-daerah yang tergenang banjir masih kurang maksimal. Oleh karena itu dalam mengoptimalkan pengetahuan informasi dan antisipasi dengan memetakan atau mengelompokkan daerah yang berpotensi rawan banjir atau tidak di daerah Kabupaten Tuban dengan metode <i>clustering</i>, dengan memanfaatkan data yang diperoleh dari katalog BPS “Kabupaten Tuban Dalam Angka (<i>Tuban Regency In Figure</i>)” dari tahun 2014 hingga 2017. Algoritma k-means adalah metode yang akan diimplementasikan untuk menggali informasi analisa dalam mengelompokkan daerah rawan banjir dengan atribut data yang digunakan adalah curah hujan, jenis tanah dan ketinggian tanah di setiap kecamatan di Kabupaten Tuban. Hasilnya diperoleh suatu sistem untuk mengelompokkan daerah rawan banjir dengan menghasilkan 3 kelompok kategori, yakni rawan, kurang rawan dan tidak rawan. Hasil visualisasi peta dengan hasil tahun 2013 sebanyak tiga warna yaitu oranye, kuning dan hijau, tahun 2014 sebanyak dua warna yaitu kuning dan hijau, tahun 2015 tiga warna yaitu oranye, kuning dan hijau serta tahun 2016 tiga warna yaitu oranye, kuning dan hijau. Dari penelitian yang telah dilakukan pada 20 data kecamatan mampu menghasilkan akurasi pengujian sebesar 45% untuk tahun 2013, 40% untuk tahun 2014, 40% untuk tahun 2015 dan 60% untuk tahun 2016. Perangkat ini dapat dikembangkan untuk membantu daerah lain dalam mengelompokkan daerah rawan banjir. Metode penelitian ini dapat dikembangkan untuk objek penelitian yang berbeda.</p> |
|---|---|

1. PENDAHULUAN

Fenomena banjir yang terjadi di beberapa daerah di Indonesia masih didominasi oleh adanya curah hujan yang tinggi dan luapan air sungai. Seperti halnya banjir yang terjadi dikarenakan oleh luapan sungai di daerah aliran sungai (DAS) maupun dataran rendah pada wilayah Kabupaten Tuban, ketika curah hujan tinggi dan sungai tidak dapat menampung air yang berasal dari air hujan, maka terjadi luapan dan mengakibatkan banjir. Hal ini telah menjadi fenomena rutin ketika musim penghujan datang. Hal ini telah menjadi fenomena rutin ketika musim penghujan datang. Namun penanganan akibat banjir di daerah-daerah yang tergenang banjir masih kurang maksimal dikarenakan kurangnya informasi untuk mengetahui dan mengantisipasi daerah mana yang berpotensi banjir. Untuk itu dalam mengoptimalkan pengetahuan informasi dan antisipasi dengan memetakan atau mengelompokkan daerah yang berpotensi rawan banjir atau tidak di daerah Kabupaten Tuban dengan metode *clustering*.

Data mining merupakan proses ekstraksi informasi data berukuran besar, sedangkan *clustering* berguna untuk menemukan pola distribusi di dalam sebuah dataset yang berguna untuk proses analisa data. Kesamaan objek biasanya diperoleh dari kedekatan nilai-nilai atribut yang menjelaskan objek-objek data, sedangkan objek-objek data biasanya direpresentasikan sebagai sebuah titik dalam ruang multidimensi. *Clustering* membantu proses pengelompokkan dan penelusuran pada data historis untuk mengidentifikasi pola dalam jumlah data yang besar. Selanjutnya clustering akan divisualisasikan melalui peta dengan berbasis Sistem Informasi Geografis (SIG) atau yang sering disebut GIS (*Geographic Information System*). Pengelompokkan

tersebut dapat menggunakan metode pengelompokan dengan algoritma *k-means*. Pemilihan algoritma *k-means* umum karena mudah diimplementasikan dan lebih sensitif, dengan data. *Geographic Information System* (GIS) juga mampu mengintegrasikan berbagai macam sistem, data dan informasi. Identifikasi kerentanan banjir menggunakan GIS dapat dilakukan dengan beberapa atribut penyebab banjir.

Tujuan dari penelitian ini adalah untuk mengembangkan GIS dengan metode *clustering* dalam mengelompokkan data dan menampilkan pemetaan daerah yang rawan terhadap banjir. Data yang diambil sebagai sampel berdasarkan Katalog BPS “Kabupaten Tuban Dalam Angka (Tuban Regency In Figure)” dari tahun 2014 hingga 2017, yang diimplementasikan untuk mengetahui daerah yang rawan terhadap banjir di Wilayah Kabupaten Tuban. Sehingga diharapkan dapat memberikan wawasan teoritis dan aplikatif untuk identifikasi dan pemetaan kawasan mengenai peranan SIG dalam antisipasi bencana banjir. Dalam melakukan penelitian ini, Data yang digunakan adalah data dari Katalog BPS “Kabupaten Tuban Dalam Angka (Tuban Regency In Figure)” tahun 2014 hingga tahun 2017. Data tersebut kemudian digunakan untuk mengelompokkan daerah dengan nilai tingkat kerawanannya terhadap banjir di Kabupaten Tuban menggunakan *clustering* dengan menggunakan algoritma *k-means*.

Manfaat dari dilakukannya penelitian ini bagi penulis adalah *clustering* dengan menggunakan algoritma *k-means*. Adapun pula manfaat bagi masyarakat sekitar adalah sebagai bahan studi masyarakat daerah aliran sungai khususnya dalam mengantisipasi bencana banjir tiap tahunnya, sebagai bahan studi masyarakat daerah aliran sungai khususnya dalam mengantisipasi bencana banjir tiap tahunnya, memberikan pola sebaran kawasan rawan banjir pada daerah yang rentan terhadap bencana banjir sehingga dapat dijadikan pertimbangan dalam perencanaan dan pengembangan wilayah secara optimal dan berkelanjutan. Dilingkungan akademik pula, bermanfaat Sebagai tolak ukur sejauh mana pemahaman dan penguasaan materi terhadap teori yang diajukan dan sebagai bahan referensi bagi mereka yang mengadakan penelitian untuk dikembangkan lebih lanjut dengan permasalahan yang berbeda.

Metode Penelitian yang digunakan dalam penelitian ini adalah Observasi yaitu melakukan pengamatan ke lembaga yang bersangkutan yaitu Badan Pusat Statistik Kabupaten Tuban untuk proses analisa data penyebab banjir serta kunjungan pada lokasi yang rentan terhadap banjir, mengadakan wawancara dengan pihak-pihak yang berkaitan langsung dengan permasalahan untuk memperoleh gambaran dan penjelasan secara mendasar, dan Studi pustaka dilakukan dengan mempelajari buku-buku, literatur-literatur, artikel-artikel dan jurnal-jurnal yang berhubungan dengan konsep *data mining*, penerapan *data mining*, serta teknik-teknik data mining khususnya teknik *clustering* dengan algoritma *k-means* serta penerapannya pada GIS.

2. TINJAUAN PUSTAKA

Dalam tinjauan pustaka, peneliti mengawali dengan menelaah penelitian terdahulu yang berkaitan serta relevansi dengan penelitian yang akan dilakukan peneliti. Dengan demikian, peneliti mendapatkan rujukan pendukung, pelengkap serta pembandingan dalam menyusun skripsi ini sehingga lebih memadai. Selain itu, telaah pada penelitian terdahulu berguna untuk memberikan gambaran awal mengenai kajian terkait dengan masalah dalam penelitian ini. Dalam penelitian ini penulis memaparkan beberapa penelitian terdahulu.

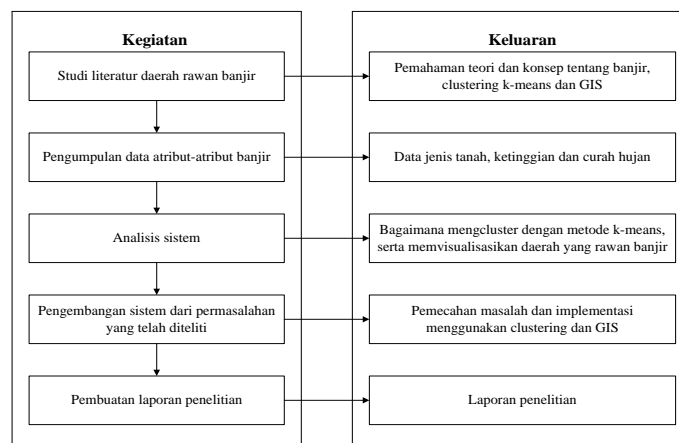
Dalam bagian ini peneliti akan menampilkan beberapa jurnal dari penelitian terdahulu yang berkaitan dengan data mining metode *cluster* algoritma *k-means* yang dilengkapi dengan sebuah

visualisasi peta. Contoh yang pertama adalah jurnal milik Slamet Handoko, Eko Sediono dan Suhartono (2011) dengan judul “Sistem Informasi Geografis Berbasis Web untuk Pemetaan Sebaran Alumni Menggunakan Metode K-Means”, dengan permasalahan untuk memetakan sebaran alumni sesuai dengan jurusan mereka menggunakan metode *k-means* dan menghasilkan sebuah sistem yang mampu menganalisa dan mengkorelasikan jurusan para alumni dengan pekerjaan yang mereka dapatkan.

Selanjutnya jurnal penelitian milik M. Rosyid Ridlo, Sofi Defiyanti dan Aji Primajaya (2017) yang berjudul “Implementasi Algoritma K-Means Untuk Pemetaan Produktivitas Panen Padi Di Kabupaten Karawang”, dengan permasalahan yaitu meningkatkan target produktivitas panen untuk mengimbangi jumlah penduduk. Sistem yang di jurnalkan ini menghasilkan sebuah pemetaan dengan pengelompokan data mining menjadi 3 kelompok dengan persentase kenaikan target sebesar 3%.

3. METODE PENELITIAN

Seperti yang telah disebutkan sebelumnya metode penelitian dilakukan dengan cara observasi, wawancara dan studi Pustaka. Dari metode tersebut didapatkan bahan prnrllitian. Dalam melakukan penelitian bahan yang digunakan mencakup panduan dan sampel data yang bertujuan sebagai tolak ukur untuk mendapatkan data, supaya keakuratan dalam mendapatkan data tetap terjaga kebenarannya. Bahan-bahan penelitian tersebut adalah Data primer dan data sekunder. Data primer adalah data yang diperoleh secara langsung dari sumber dilapangan Adapun data diperoleh dari hasil penelitian dari Katalog Kabupaten Tuban Dalam Angka (*Tuban Regency In Figures*) Nomor 1102001.3523, Tahun 2014 hingga 2017 yang diterbitkan oleh Badan Pusat Statistik Kabupaten Tuban. Selain memperoleh data secara langsung, penulis juga memperoleh data secara tidak langsung yaitu dengan membaca, mencari, dan mengkaji materi yang berkaitan dengan penelitian. Dalam melakukan penelitian terdapat pula prosedur penelitian adalah seperti pada alur berikut :

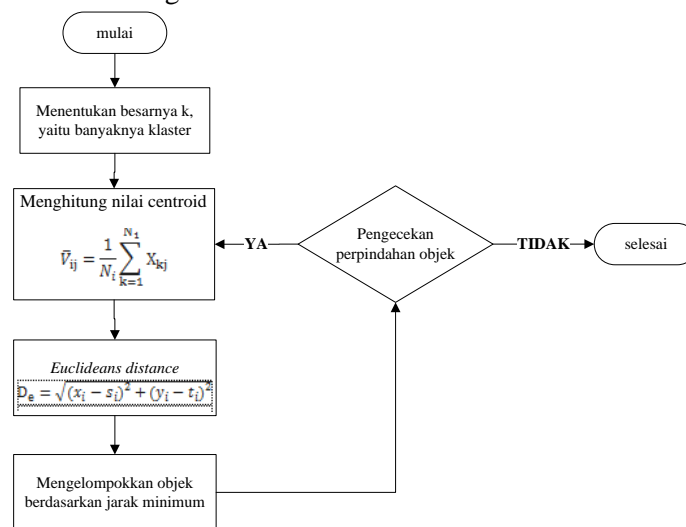


Gambar 1 Kerangka Penelitian

Dari Langkah penelitian yang dilakukan, sehingga didapatkan permasalahan yang telah disebutkan sebelumnya yaitu bagaimana mengolah data atribut banjir yang sudah ada menjadi ilmu pengetahuan baru berupa informasi yang akan datang, yang dapat berguna untuk menciptakan suatu strategi baru untuk memperoleh solusi dalam penyelesaian masalah yang sedang terjadi, pada kasus ini adalah menganalisis pola pengelompokan sebaran daerah rawan banjir di Kabupaten Tuban. Dari sini akan didapatkan sebuah data. Setelah data dianalisis, selanjutnya penentuan system

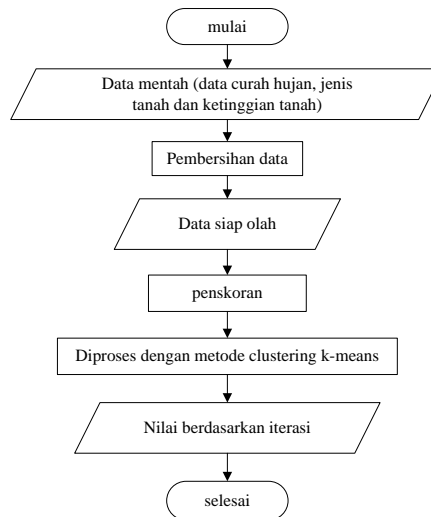
yang akan digunakan yaitu dengan sebuah sistem dengan fungsi utama untuk melakukan pengelompokan daerah rawan banjir.

Selanjutnya adalah analisis penyelesaian masalah. Untuk penerapan *clustering* menggunakan algoritma *K-Means*, disini akan menentukan pola daerah yang rawan, cukup rawan dan tidak rawan dalam pengelompokan daerah rawan banjir dengan melihat dari atribut yang telah ditentukan sebelumnya. Adapun yang menjadi konsep penyelesaian dalam masalah ini yakni dengan pola *waterfall* dengan mengubah data yang ada menjadi data penskoran, serta dari data yang telah di *cluster* akan diubah kedalam bentuk visualisasi peta digital. Data penelitian ini berdasarkan data dari Katalog Kabupaten Tuban Dalam Angka (*Tuban Regency In Figures*) Nomor 1102001.3523, diterbitkan oleh Badan Pusat Statistik Kabupaten Tuban pada tahun 2014 hingga 2017 sebanyak 20 data untuk tiap tahunnya. Pengelompokan proses yang terjadi pada pengelompokan daerah rawan banjir menggunakan algoritma *k-means* ini, dapat digambarkan dengan menggunakan *flowchart* sebagai berikut:



Gambar 2. Flowchart metode *clustering k-means*

Dari gambar 3 yang digunakan untuk penyelesaian masalah yaitu pengelompokan daerah rawan banjir dengan menggunakan algoritma *k-means*, yang merupakan tipe pengelompokan pada metode *clustering* dimana dalam pemilihan atribut menggunakan suatu kesamaan antar objek yang biasanya diperoleh dari kedekatan nilai-nilai atribut. Dalam menganalisa sistem tentunya juga harus memiliki sebuah desain alur yang dilakukan agar mendapat hasil yang Analisa yang dilakukan memiliki desain sistem yang dijalankan. Memiliki tujuan untuk memahami pemecahan masalah yang didapat pada tahap analisis melalui suatu pemodelan dengan gambaran flowchart sebagai berikut:



Gambar 2. Flowchart pengolahan data

Analisis kerawanan banjir merupakan penentuan nilai kerawanan suatu daerah terhadap bencana banjir. Analisis ini dimulai dengan penskoran dan pembobotan kemudian menghitung nilai kerawanan. Nilai kerawanan ini bisa diketahui dengan menggunakan persamaan sebagai berikut:

$$X = \sum_{i=1}^n (W_i \times X_i) \dots\dots\dots (1)$$

Dimana:

X = Nilai kerawanan

Wi = Bobot untuk variabel ke-i

Xi = Skor pada variabel ke-i

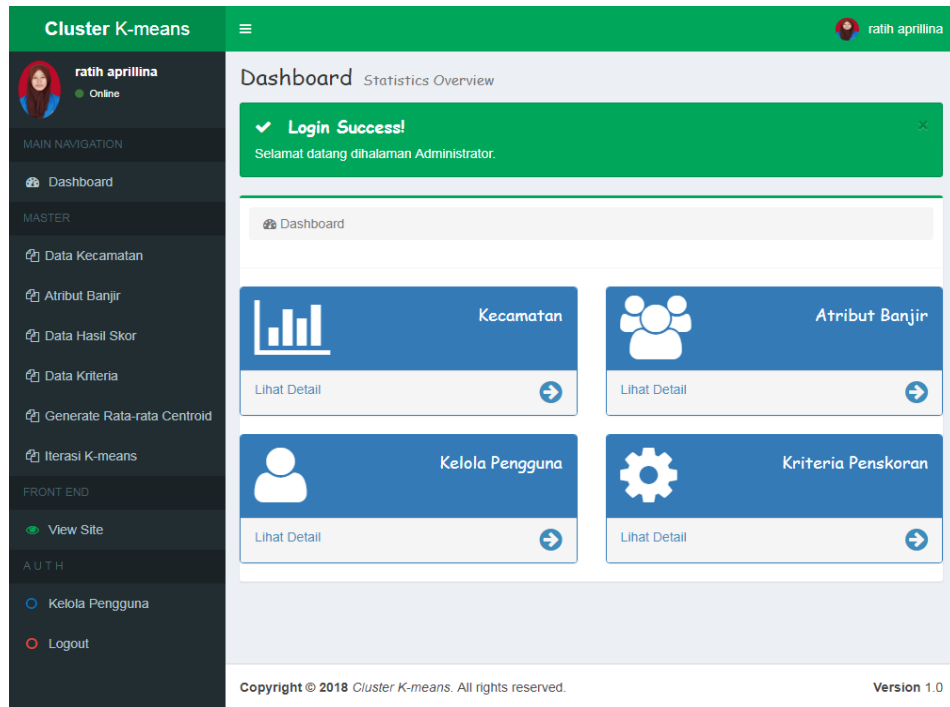
Daerah yang rawan terhadap banjir akan mempunyai total nilai yang tinggi, dan sebaliknya daerah yang tidak rawan maka akan mempunyai nilai yang rendah. Tabel 3.5 dibawah ini menunjukkan nilai tingkat kerawanan banjir menurut Asep Purnama (2008).

Tabel 1. Nilai Tingkat Kerawanan Banjir (Purnama, 2008)

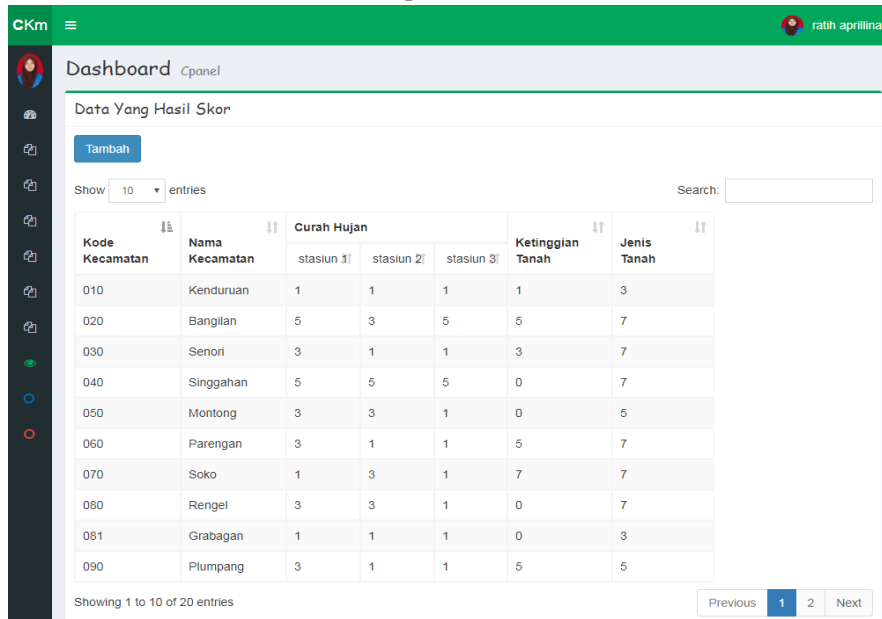
| Tingkat Kerawanan Banjir | Jumlah Nilai Semua Parameter |
|--------------------------|------------------------------|
| Sangat rawan banjir | 6,75 – 9 |
| Rawan banjir | 4,5 – 6,75 |
| Kurang rawan banjir | 2,25 – 4,5 |
| Tidak rawan banjir | < 2,25 |

4. HASIL DAN ANALISA

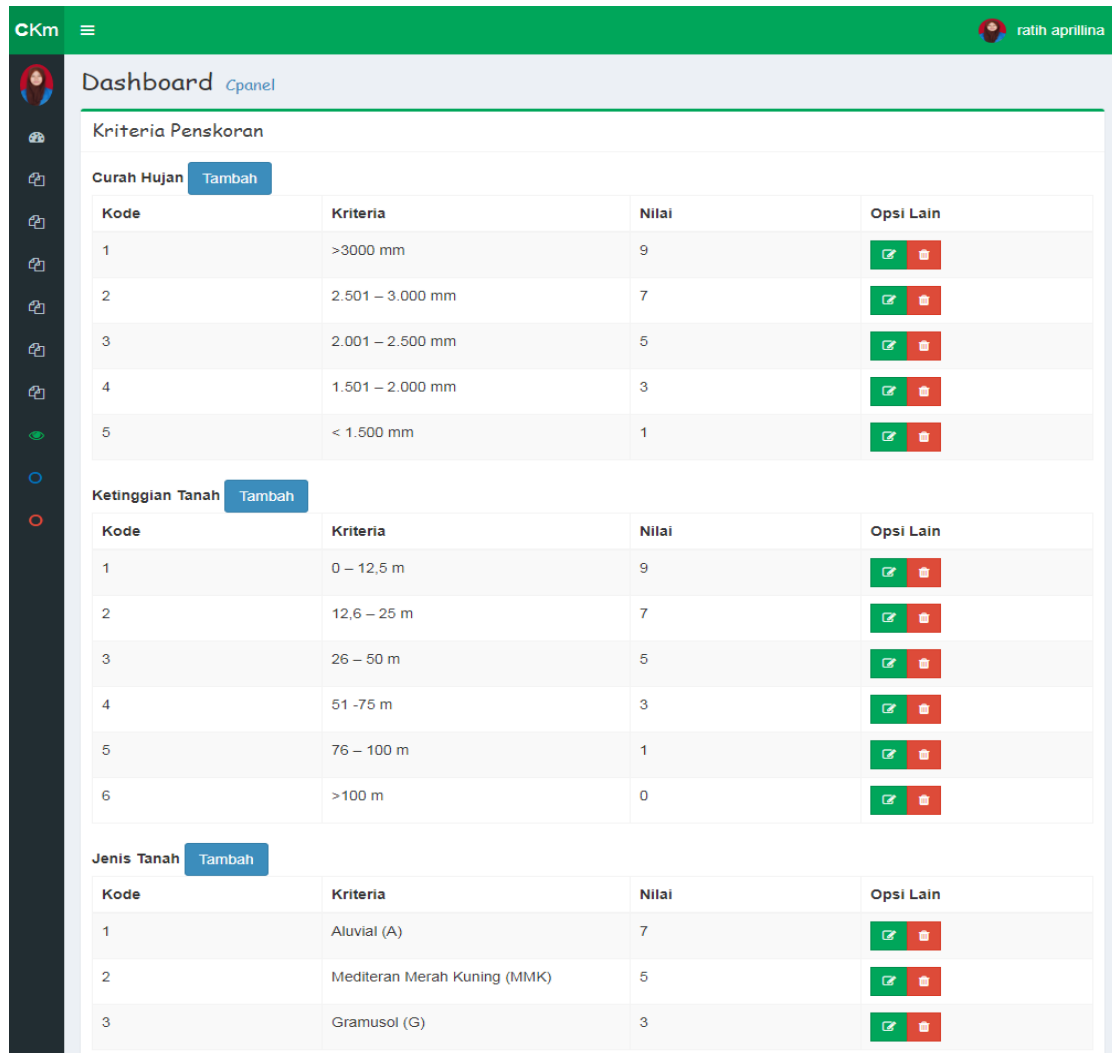
Implementasi sistem merupakan tahapan pembuatan sistem yang dilakukan berdasarkan hasil analisa dan perancangan yang telah dilakukan sebelumnya. Pada tahapan implementasi sistem ini diharapkan sistem yang telah dirancang siap untuk dioperasikan pada keadaan yang sebenarnya, sehingga akan diketahui apakah sistem yang dibuat benar-benar dapat menghasilkan tujuan yang diinginkan. Dari hasil penelitian ini akan menghasilkan sebuah program aplikasi dengan desain sebagai berikut :



Gambar 4. Tampilan *dashboard* administrator



Gambar 5. Tampilan halaman data hasil skor



Gambar 6. Tampilan halaman data penskoran

Untuk mengetahui hasil dari *Clustering* Daerah Rawan Banjir Di Kabupaten Tuban Dengan K-Means Disertai Visualisasi Sistem Informasi Geografis ini, maka proses perhitungan dengan algoritma *Clustering* metode k-means dilakukan setelah data atribut banjir pada tiap kecamatan telah dilakukan penskoran sebelumnya untuk menghasilkan suatu *cluster*. kecamatan telah dilakukan penskoran sebelumnya untuk menghasilkan suatu *cluster*. Tiap hasil *cluster* tersebut digunakan untuk memprediksi kecamatan sesuai tingkat kerawannya, yang kemudian akan divisualisasikan. Seperti yang sudah dijelaskan sebelumnya bahwa atribut yang digunakan adalah curah hujan, ketinggian tanah dan jenis tanah. Dari hasil observasi dan data yang ada maka dilakukan penskoran sehingga dihasilkan table sebagai berikut :

Tabel 2. Hasil pengolahan dengan pemberian skor

| No | Kode | Nama Kecamatan | jenis tanah | Ketinggian | curah hujan | | | tahun |
|-----|------|----------------|-------------|------------|-------------|-----------|-----------|-------|
| | | | | | Stasiun 1 | Stasiun 2 | Stasiun 3 | |
| (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| 1 | 10 | Kenduruan | 3 | 1 | 1 | 1 | 1 | 2013 |
| 2 | 20 | Bangilan | 7 | 5 | 5 | 3 | 5 | |
| 3 | 30 | Senori | 7 | 3 | 3 | 1 | 1 | |

| | | | | | | | | |
|-----|-----|------------|-----|-----|-----|-----|-----|--|
| 4 | 40 | Singgahan | 7 | 1 | 5 | 5 | 5 | |
| 5 | 50 | Montong | 5 | 1 | 3 | 3 | 1 | |
| 6 | 60 | Parengan | 7 | 5 | 3 | 1 | 1 | |
| 7 | 70 | Soko | 7 | 7 | 1 | 3 | 1 | |
| 8 | 80 | Rengel | 7 | 1 | 3 | 3 | 1 | |
| 9 | 81 | Grabagan | 3 | 1 | 1 | 1 | 1 | |
| 10 | 90 | Plumpang | 5 | 5 | 3 | 1 | 1 | |
| 11 | 100 | Widang | 5 | 9 | 3 | 1 | 1 | |
| 12 | 110 | Palang | 5 | 9 | 3 | 3 | 1 | |
| 13 | 120 | Semanding | 5 | 3 | 3 | 1 | 1 | |
| 14 | 130 | Tuban | 7 | 9 | 5 | 1 | 1 | |
| 15 | 140 | Jenu | 5 | 9 | 1 | 1 | 1 | |
| ... | ... | ... | ... | ... | ... | ... | ... | |
| ... | ... | ... | ... | ... | ... | ... | ... | |
| 63 | 30 | Senori | 7 | 3 | 5 | 1 | 5 | |
| 64 | 40 | Singgahan | 7 | 5 | 5 | 5 | 5 | |
| 65 | 50 | Montong | 5 | 1 | 7 | 7 | 1 | |
| 66 | 60 | Parengan | 7 | 5 | 5 | 1 | 1 | |
| 67 | 70 | Soko | 7 | 7 | 3 | 5 | 1 | |
| 68 | 80 | Rengel | 7 | 9 | 5 | 3 | 1 | |
| 69 | 81 | Grabagan | 3 | 1 | 1 | 1 | 1 | |
| 70 | 90 | Plumpang | 5 | 5 | 5 | 1 | 1 | |
| 71 | 100 | Widang | 5 | 9 | 5 | 1 | 1 | |
| 72 | 110 | Palang | 5 | 9 | 3 | 3 | 1 | |
| 73 | 120 | Semanding | 5 | 1 | 3 | 1 | 1 | |
| 74 | 130 | Tuban | 7 | 9 | 7 | 1 | 1 | |
| 75 | 140 | Jenu | 5 | 7 | 1 | 1 | 1 | |
| 76 | 150 | Merakurak | 5 | 9 | 3 | 3 | 3 | |
| 77 | 160 | Kerek | 5 | 3 | 3 | 1 | 1 | |
| 78 | 170 | Tambakboyo | 5 | 9 | 1 | 1 | 1 | |
| 79 | 180 | Jatirogo | 3 | 3 | 1 | 1 | 1 | |
| 80 | 190 | Bancar | 7 | 9 | 1 | 1 | 1 | |

Setelah semua data curah hujan, jenis tanah dan ketinggian tanah pada tahun 2013 sampai 2016 dirubah ke dalam bentuk angka dengan cara penskoran, maka data-data tersebut telah dapat dikelompokkan dengan menggunakan algoritma *K-Means Clustering*. Untuk dapat melakukan pengelompokan data-data tersebut menjadi beberapa *cluster* perlu dilakukan beberapa langkah, yaitu:

1. Tentukan jumlah *cluster* yang diinginkan. Pada penelitian ini, terdapat tiga cluster yaitu C1, C2 dan C3.
2. Tentukan titik pusat awal dari setiap *centroid*. Dalam penelitian ini titik pusat awal ditentukan secara random yang diambil dari table 3.1. dan didapat titik pusat dari setiap *centroid* dapat dilihat pada tabel 4.2.

Tabel 3. Titik pusat awal *centroid*

| 3. Centroid | Curah hujan | Ketinggian tanah | Jenis tanah |
|-------------|-------------|------------------|-------------|
| C1 | 6,75 | 2,25 | 2,25 |
| C2 | 2,25 | 6,75 | 2,25 |
| C3 | 2,25 | 2,25 | 6,75 |

3. Menghitung jarak antara titik centroid dengan titik tiap objek menggunakan *Euclidean Distance* pada iterasi pertama. Sebagai contoh yakni kecamatan Kenduruan pada stasiun 1 tahun 2013 dari tabel 4.2, yang kemudian akan didapatkan matriks jarak yaitu D1, D2 dan D3 sebagai berikut:

Perhitungan jarak Kenduruan dengan pusat *cluster* pertama adalah:

$$D_1 = \sqrt{(3 - 6,75)^2 + (1 - 2,25)^2 + (1 - 2,25)^2} = 5,931905259$$

Perhitungan jarak Kenduruan dengan pusat *cluster* kedua adalah:

$$D_2 = \sqrt{(3 - 2,25)^2 + (1 - 6,75)^2 + (1 - 2,25)^2} = 5,931905259$$

Perhitungan jarak Kenduruan dengan pusat *cluster* ketiga adalah:

$$D_3 = \sqrt{(3 - 2,25)^2 + (1 - 2,25)^2 + (1 - 6,75)^2} = 4,145780988$$

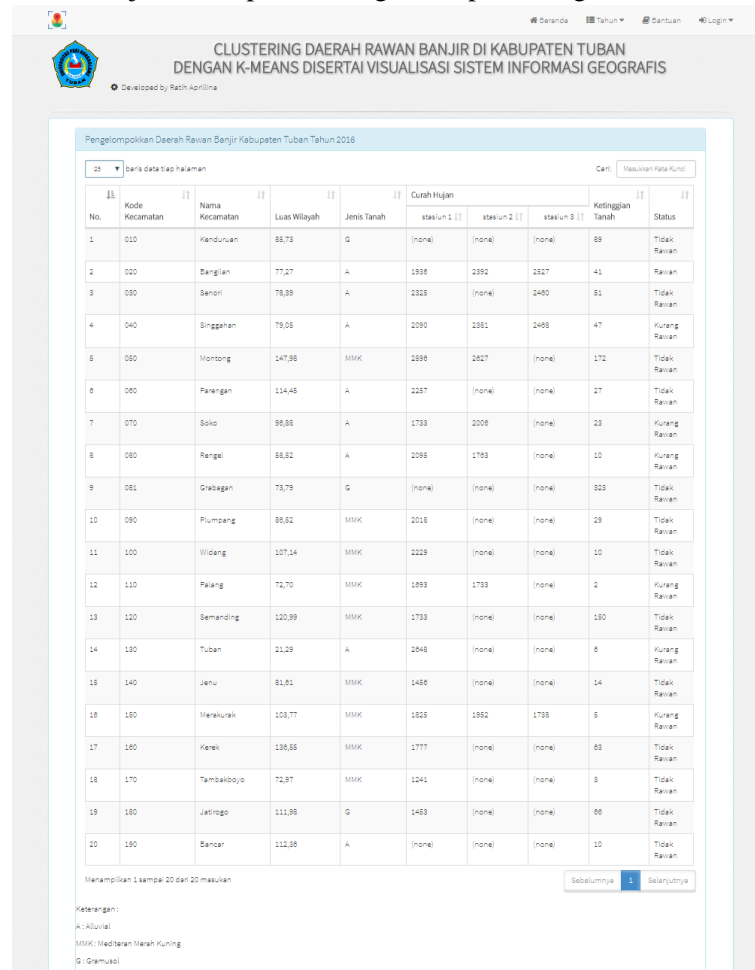
4. Langkah *clustering* yang terakhir yaitu dengan mengelompokkan obyek berdasarkan jarak minimum antara data dengan titik pusatnya. Sehingga dalam proses perhitungan didapat beberapa iterasi hingga mencapai hasil hingga perpindahan jarak antar *cluster* tidak terjadi. Sehingga didapatkan hasil yang berhenti pada iterasi ke-3 dari hasil *clustering*, dan telah mencapai stabil dan konvergen. Sehingga didapatkan hasil akhir sebahai berikut :

Tabel 4 Hasil akhir iterasi

| kode kec | Kecamatan | Curah hujan | ketinggian | Jenis tanah | Jarak Terpendek |
|----------|------------|-------------|------------|-------------|-----------------|
| 10 | Kenduruan | 661,042 | 5,095,524 | 7,398,643 | 5,095,524 |
| 20 | Bangilan | 2,274,938 | 1,987,738 | 2,787,934 | 1,987,738 |
| 30 | Senori | 3,680,557 | 2,319,606 | 4,910,478 | 2,319,606 |
| 40 | Singgahan | 6,628,846 | 4,432,823 | 7,007,062 | 4,432,823 |
| 50 | Montong | 6,457,264 | 4,286,843 | 7,074,747 | 4,286,843 |
| 60 | Parengan | 2,009,975 | 2,258,549 | 3,685,412 | 2,009,975 |
| 70 | Soko | 306,564 | 4,687,373 | 4,986,122 | 306,564 |
| 80 | Rengel | 6,542,649 | 4,560,686 | 7,410,033 | 4,560,686 |
| 81 | Grabagan | 7,446,536 | 5,753,618 | 8,185,669 | 5,753,618 |
| 90 | Plumpang | 1,711,722 | 1,636,522 | 2,953,932 | 1,636,522 |
| 100 | Widang | 2,813,714 | 511,069 | 3,415,355 | 2,813,714 |
| 110 | Palang | 2,813,714 | 511,069 | 3,415,355 | 2,813,714 |
| 120 | Semanding | 3,526,542 | 1,719,806 | 4,388,194 | 1,719,806 |
| 130 | Tuban | 3,187,842 | 5,233,741 | 3,272,845 | 3,187,842 |
| 140 | Jenu | 384,469 | 5,938,779 | 504,721 | 384,469 |
| 150 | Merakurak | 2,813,714 | 511,069 | 3,415,355 | 2,813,714 |
| 160 | Kerek | 3,526,542 | 1,719,806 | 4,388,194 | 1,719,806 |
| 170 | Tambakboyo | 2,813,714 | 511,069 | 3,415,355 | 2,813,714 |

| | | | | | |
|-----|----------|-----------|-----------|-----------|-----------|
| 180 | Jatirogo | 4,396,191 | 2,921,454 | 4,732,832 | 2,921,454 |
| 190 | Bancar | 306,564 | 4,687,373 | 4,986,122 | 306,564 |

Dengan hasil diatas akan tersaji dalam aplikasi dengan tampilan sebagai berikut :



Gambar 7. Tampilan hasil perhitungan

Suatu pengujian akurasi sangat dibutuhkan untuk mengetahui apakah sistem yang sudah dibangun tersebut sesuai dengan apa yang dibutuhkan, maka penulis menguji hasil pengolahan data dengan *clustering* algoritma *k-means* ini sudah sesuai dengan data kejadian sebenarnya.

Frekuensi genangan menurut Wika Ristya (2012), Semakin sering terjadinya genangan disuatu wilayah maka kerugian dan kerusakan akan semakin besar. Klasifikasi untuk frekuensi genangan akan dibagi menjadi empat kelas yaitu:

- 1) 0 kali kejadian (tidak berpotensi)
- 2) 1-2 kali kejadian (rendah)
- 3) 3-5 kali kejadian (sedang)
- 4) 6-20 kali kejadian (tinggi)

Dengan contoh sebagai berikut :

Perbandingan hasil *cluster* dengan data kejadian sebenarnya

Tabel 5. Perbandingan hasil *cluster* dengan data kejadian sebenarnya

| Kecamatan | Kejadian sebenarnya | Hasil cluster | Perbandingan |
|------------|---------------------|---------------|--------------|
| Tambakboyo | Tidak rawan | Tidak rawan | Sesuai |

| | | | |
|----------|-------------|-------------|--------|
| Jatirogo | Tidak rawan | Tidak rawan | Sesuai |
| Bancar | Tidak rawan | Tidak rawan | Sesuai |

Dengan rumus persentase perbandingan yaitu:

$$\text{persentase kesesuaian} = \frac{12}{20} \times 100\% = 60\%$$

Maka 12 atau 60% data sesuai dan 8 atau 40% data tidak sesuai.

5. KESIMPULAN

Berdasarkan hasil validasi keakuratan perhitungan, algoritma k-means menunjukkan nilai yang kurang memuaskan, dikarenakan hasil persentase hampir seluruhnya menunjukkan nilai kesesuaian kurang dari 50%. Menurut Dedy Atmajaya (2016), "Algoritma k-means diinisialisasikan secara random sehingga data yang dihasilkan kurang baik, jika nilai random kurang baik maka hasilnya pun kurang maksimal atau biasa disebut dengan *curse of dimentionality*, selain dari inialisasi, banyaknya data juga mempengaruhi hasil". Jadi dapat disimpulkan jika algoritma k-means kurang sesuai digunakan dengan data atribut banjir pada penelitian ini

DAFTAR PUSTAKA

- Atmajaya, Dedy., dan Azis, W.S. 2016, Pengelompokan Minat Baca Mahasiswa Menggunakan Metode K-Means, Fakultas Ilmu Komputer, UMI, 8 Agustus.
- De Bay, 1999. *Prinsiples of Geographic Information Systems*. ITC Core Modules Textbook, Netherlands: ITC.
- Fahmi Muhammad, Farid., Suprpto, Yoyon K. 2013, Penentuan Prioritas Rehabilitasi DAS Menggunakan Algoritma K-Means Clustering JAVA Journal of Electrical and Electronics Engineering Volume 11, Nomor 2, Oktober 2013.
- Han, J. and Kamber, M. "*Data mining: Concepts and Techniques*", Second edition. The Morgan Kaufmann series in Data Management System, Jim Grey, series Editor. 2006.
- Handoko, Slamet., Sediono, Eko., dan Suhartono, 2011, Sistem Informasi Geografis Berbasis Web untuk Pemetaan Sebaran Alumni Menggunakan Metode K-Means, Jurnal Sistem Informasi Bisnis 02 (SINBIS), Semarang, agustus 2011, ISSN- 2088 3587.
- Haviluddin, 2009, Memahami Penggunaan Diagram Arus Data, Samarinda, Jurnal Informatika Mulawarman, Vol 4, No. 3, September.
- Ika Wardati Dian, Prabawani. 2010. Aplikasi Penginderaan Jauh dan Sistem Infomasi Geografi Untuk Pemodelan Genangan Banjir di Kecamatan Jebres Kota Surakarta. Yogyakarta: Fakultas Geografi UGM.
- Kang-Tsung, Chang. 2002. *Introduction to Geographic Information System*, Mc.Graw-Hill.
- Katalog BPS. 1102001.3523, 2014, Kabupaten Tuban Dalam Angka (Tuban Regency In Figures) 2014 bagian Geografi (*Geographycal*) 3-18.
- Katalog BPS. 1102001.3523, 2015, Kabupaten Tuban Dalam Angka (Tuban Regency In Figures) 2015 bagian Geografi (*Geographycal*) 3-18.
- Katalog BPS. 1102001.3523, 2016, Kabupaten Tuban Dalam Angka (Tuban Regency In Figures) 2016 bagian Geografi (*Geographycal*) 3-18.
- Katalog BPS. 1102001.3523, 2017, Kabupaten Tuban Dalam Angka (Tuban Regency In Figures) 2017 bagian Geografi (*Geographycal*) 3-18.
- Kodoatie Robert J. dan Sugiyanto. 2002. *Banjir: Beberapa Penyebab dan Metode Pengdaliaannya Dalam Perspektif Lingkungan*. Yogyakarta: Gadjah Mada University Press.
- Larose D, T. 2005, *Discovering knowledge in data: an introduction to data mining*, Jhon Wiley & Sons Inc.