

**IMPLEMENTASI DATA MINING UNTUK CLUSTERING
DAERAH PENYEBARAN PENYAKIT DEMAM BERDARAH
DI KOTA TANGERANG SELATAN MENGGUNAKAN
ALGORITMA K-MEANS
(Studi Kasus : Dinas Kesehatan Tangerang Selatan)
Nurul Chafid, Ismail Ari Wibowo**

ABSTRAK

Demam Berdarah merupakan penyakit yang menular yang terdapat di wilayah tropis maupun subtropis. Semakin meningkatnya kepadatan penduduk di Kota Tangerang Selatan pada Tahun 2016 jumlah 1.593.812 jiwa yang tersebar di 7 Kecamatan. Untuk melihat daerah penyebarannya perlu dibuat sebuah pengelompokan penyebaran agar dapat memperoleh pusat titik penyebaran. Algoritma K-Means cocok untuk pengelompokan daerah endemis, potensial dan sporadis. Pengelompokan C1 yaitu Kelurahan Rawa Buntu, C2 yaitu Kelurahan Pd. Ranji dan C3 yaitu Kelurahan Serua Indah sebagai pusat cluster. Hasil pengelompokan C1 ada 20 Kelurahan, C2 ada 13 Kelurahan dan C3 ada 19 dari total Kelurahan 52 akan menjadi bahan untuk melakukan penyehatan lingkungan sesuai dengan kelompok yang akan dikerjakan oleh Dinas Kesehatan Tangerang Selatan.

Kata Kunci : Demam Berdarah, Algoritma K-Means

ABSTRACT

Dengue fever is a contagious disease found in tropical and subtropical regions. The increasing population density in South Tangerang City in 2016 amounts to 1,593,812 people spread in 7 subdistricts. To see the spreading area there needs to be a dispersion grouping in order to obtain a central point of deployment. The K-Means algorithm is suitable for grouping of endemic, potential and sporadic areas. Grouping C1 is Rawa Buntu Village, C2 that is Pd. Ranji and C3 is Serua Indah Urban Village as the center of cluster. The result of C1 grouping is 20 Kelurahan, C2 there are 13 Kelurahan and C3 there are 19 out of total Village 52 will be material to do environmental sanitation in accordance with group to be done by Health Service of South Tangerang.

Keywords : *Dengue Fever, K-Means Algorithm*

PENDAHULUAN

Demam berdarah merupakan salah satu jenis penyakit menular yang terdapat di wilayah tropis maupun subtropis. Demam berdarah dapat menular karena gigitan Nyamuk Aedes Aegypti. Dampak dari Demam Berdarah dapat membuat suhu tubuh penderita menjadi sangat tinggi dan pada umumnya disertai demam, mual/muntah, sakit kepala, nyeri perut, dan lekopenia. Penyakit Demam Berdarah masih merupakan salah satu permasalahan kesehatan masyarakat. Dinas Kesehatan mempunyai tugas pokok untuk membantu penyelenggaraan kegiatan penyehatan lingkungan. ▯

Semakin meningkatnya kepadatan penduduk di Kota Tangerang Selatan pada tahun 2016 jumlah 1.593.812 jiwa tersebar di 7 Kecamatan dan meningkatnya penderita penyakit

Demam Berdarah karena kurangnya kesadaran masyarakat terhadap kesehatan lingkungan. Timbulah masalah yang menyulitkan Dinas Kesehatan untuk mengelompokkan daerah yang akan menghasilkan titik-titik pusat penyebaran penderita Demam Berdarah. Agar program kegiatan penyehatan lingkungan berjalan dengan baik dan mudah dilakukan. Maka dilakukan pengelompokan berdasar jumlah penduduk, jumlah penderita. Supaya program kerja Dinas Kesehatan Kota Tangerang Selatan bisa melakukan budi daya tamanan anti nyamuK, program 3M (menguras, menutup, mengubur) yang bertujuan untuk menurunkan angka penderita penyakit Demam Berdarah dan penyuluhan untuk memberi informasi tentang bahaya penyakit Demam Berdarah.

Pada penelitian ini bertujuan mengimplementasikan Algoritma K-Means untuk pengelompokan daerah penyebaran Demam Berdarah dengan memaksimalkan kemiripan dalam satu klaster maka akan menghasilkan titik-titik pusat penyebaran penderita penyakit Demam Berdarah. Dari hasil penelitian di atas penulis mencoba menuangkan dalam bentuk jurnal penelitian ini dengan sebuah gagasan walaupun sebelumnya sudah pernah ada namun tidak menutup kemungkinan untuk menganalisa dengan metode yang disebutkan diatas guna mengelompokan penyebaran penyakit DBD di daerah tersebut.

TINJAUAN PUSTAKA

Dari penelitian sebelumnya, terdapat banyak penelitian yang dilakukan mengenai pengelompokan daerah rentan penyakit, sebagai berikut, Penelitian ini dilakukan oleh Fina Nasari dan Charles Jhony Manto Sianturi, Program Studi Sistem Informasi, Universitas Potensi Utama berjudul “Penerapan Algoritma K-Means Clustering Untuk Pengelompokan Penyebaran Diare Di Kabupaten Langkat” (Fina Nasari, Charles Jhony Manto Sianturi, Desember, 2016). Kemudian dalam penelitian ini dilakukan oleh Windha Mega Pradnya Duhita, STMIK AMIKOM Yogyakarta berjudul “Clustering Menggunakan Metode K-Means Untuk Menentukan Status Gizi Balita” (Duhita, Desember, 2015). Dan yang terakhir didapatkan dalam penelitian ini dilakukan oleh Erga Aprina Sari, Universitas Dian Nuswantoro berjudul “Penerapan Algoritma K-Means Untuk Menentukan Tingkat Kesehatan Bayi Dan Balita Pada Kabupaten dan Kota Di Jawa Tengah” (Sari).

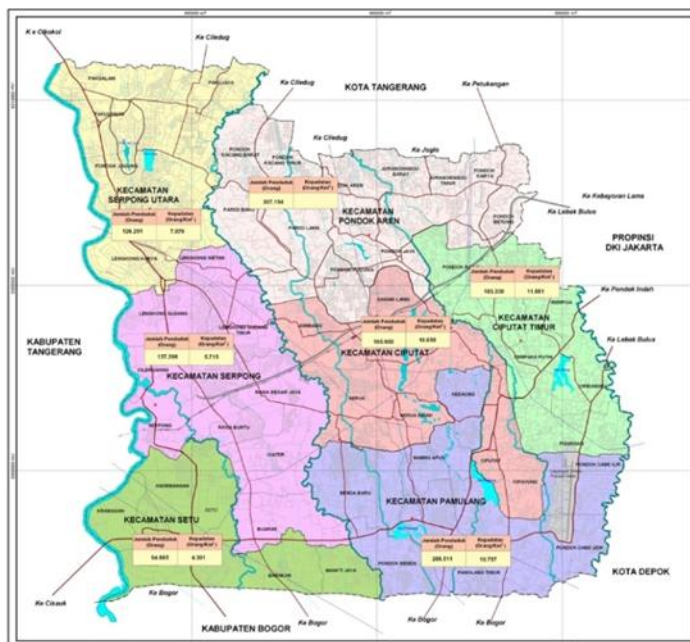
Pada penelitian diatas menyimpulkan pencarian hasil pengelompokan berdasarkan Kecamatan dengan jumlah atribut yang dihitung hanya 2. Penelitian tersebut mensarankan pengelompokan berdasarkan desa atau kelurahan untuk lebih detail dan penambahan atribut.

LANDASAN TEORI

Peta Tangerang Selatan

Kota Tangerang Selatan adalah sebuah kota yang terletak di Provinsi Banten, Indonesia. Kota ini terletak 30 km sebelah barat Jakarta dan 90 km sebelah tenggara Serang, ibu kota Provinsi Banten. Kota Tangerang Selatan berbatasan dengan Kota Tangerang di sebelah utara, Kabupaten Bogor (Provinsi Jawa Barat) di sebelah selatan, Kabupaten Tangerang di sebelah barat, serta Daerah Khusus Ibukota Jakarta di sebelah timur. Dari segi jumlah penduduk, Tangerang Selatan merupakan kota terbesar kedua di Provinsi Banten setelah Kota Tangerang serta terbesar kelima di kawasan Jabodetabek setelah Jakarta, Bekasi, Tangerang, dan Depok. Wilayah Kota Tangerang Selatan merupakan hasil pemekaran dari Kabupaten Tangerang. Letak geografis Tangerang Selatan terletak di bagian timur Provinsi Banten yaitu pada titik koordinat 106°38' – 106°47' Bujur Timur dan 06°13'30' – 06°22'30' Lintang Selatan.

Wilayah Kota Tangerang Selatan diantaranya dilintasi oleh Kali Angke, Kali Pesanggrahan dan Sungai Cisadane sebagai batas administrasi kota di sebelah barat. Letak geografis Tangerang Selatan yang berbatasan dengan Provinsi DKI Jakarta pada sebelah utara dan timur memberikan peluang pada Kota Tangerang Selatan sebagai salah satu daerah penyangga provinsi DKI Jakarta, selain itu juga sebagai daerah yang menghubungkan Provinsi Banten dengan DKI Jakarta. Selain itu, Tangerang Selatan juga menjadi salah satu daerah yang menghubungkan Provinsi Banten dengan Provinsi Jawa Barat (<https://kabartangsel.com/direktori-tangsel/peta-kota-tangerang-selatan/>). Gambar dapat kita lihat berikut ini:



Gambar 1 Peta Tangerang Selatan

Pengertian Demam Berdarah

Penyakit Demam Berdarah Dengue (DBD) merupakan salah satu masalah kesehatan masyarakat di Indonesia yang jumlah penderitanya cenderung meningkat *browser* dan melayani dengan menyediakan respon HTTP berupa konten data, biasanya berupa halaman web yang terdiri dari dokumen HTML dan objek terkait seperti gambar dan lain-lain. dan penyebarannya semakin luas dan penyakit ini merupakan penyakit menular yang terutama menyerang anak-anak dan remaja (Widoyono, 2008).

Demam Berdarah Dengue (DBD) atau Dengue Haemorrhagic Fever (DHF) adalah suatu penyakit yang disebabkan oleh virus Dengue Family Flaviviridae, dengan genusnya adalah Flavivirus. Virus mempunyai empat serotipe yang dikena dengan DEN-1, DEN-2, DEN-3, dan DEN-4. Selama ini secara klinik mempunyai tingkatan manifestasi yang berbeda tergantung dari serotipe virus dengue. MORbiditas penyakit DBD menyebar di negara-negara tropis dan sub tropis. Disetiap Negara penyakit DBD mempunyai manifestasi klinik yang beda.

Pengertian Data Mining

Data mining adalah proses untuk mendapatkan informasi yang berguna dari gudang basis data yang besar. Data mining juga dapat diartikan sebagai pengekstrakan informasi baru yang diambil dari bongkahan data bebas yang membantu dalam pengambilan keputusan. Istilah data mining kadang disebut juga knowledge discovery (Tan P, Steinbach M, Kumar V, 2006).

Pada prosesnya data mining akan mengekstrak informasi yang berharga dengan cara menganalisis adanya pola-pola ataupun hubungan keterkaitan tertentu dari data-data yang berukuran besar. Data mining berkaitan dengan bidang ilmu-ilmu lain, seperti Database System, Data Warehousing, Statistic, Machine Learning, Information Retrieval, dan Komputasi Tingkat Tinggi. Selain itu data mining didukung oleh ilmu lain seperti Neural Network, Pengenalan Pola, Spatial Data Analysis, Image Database, Signal Processing.

Pekerjaan yang berkaitan dengan data mining dapat dibagi menjadi empat kelompok, yaitu model prediksi (prediction modelling), analisis kelompok (cluster analysis), analisis asosiasi (association analysis) dan deteksi anomaly (anomaly detection) (Prasetyo, 2012)

Data mining dibagi menjadi beberapa kelompok berdasarkan tugas yang dapat dilakukan yaitu :

a) Model prediksi

Model prediksi berkaitan dengan pembuatan sebuah model yang dapat melakukan pemetaan dari setiap himpunan variabel ke setiap targetnya, kemudian menggunakan model tersebut untuk memberikan nilai target pada himpunan baru yang didapat. Ada dua jenis model prediksi, yaitu klasifikasi dan regresi. Klasifikasi digunakan untuk variabel target diskret, sedangkan regresi untuk variabel target kontinu. Misalnya pekerjaan untuk melakukan deteksi jenis penyakit pasien berdasarkan sejumlah nilai parameter penyakit yang diderita masuk dalam jenis klasifikasi karena di sini target yang diharapkan adalah diskret, hanya beberapa jenis kemungkinan nilai target yang didapatkan, tidak ada nilai deret waktu (time series) yang harus didapatkan untuk mendapat target nilai akhir. Sementara, pekerjaan prediksi jumlah penjualan yang didapatkan pada tiga bulan ke depan termasuk regresi karena untuk mendapatkan nilai penjualan bulan ketiga, nilai penjualan bulan kedua harus didapatkan dan untuk mendapatkan nilai penjualan bulan kedua, nilai penjualan bulan pertama harus didapatkan. Di sini ada nilai deret waktu yang harus dihitung untuk sampai pada target akhir yang diinginkan, ada nilai kontinu yang harus dihitung untuk mendapatkan nilai target akhir yang diinginkan.

b) Analisis kelompok

Contoh pekerjaan yang berhubungan dengan analisis kelompok (cluster analysis) adalah bagaimana cara mengetahui pola pembelian barang oleh para

konsumen pada waktu-waktu tertentu. Dengan mengetahui pola kelompok pembelian tersebut, perusahaan/ pengecer dapat menentukan jadwal promosi yang dapat diberikan sehingga omzet penjualan bisa ditingkatkan. Analisis kelompok melakukan pengelompokan data-data ke dalam sejumlah kelompok (cluster) berdasarkan kesamaan karakteristik masing-masing data pada kelompok-kelompok yang ada. Data-data yang masuk dalam batas kesamaan dengan kelompoknya akan bergabung dalam kelompok tersebut, dan akan terpisah dalam kelompok yang berbeda jika keluar dari batas kesamaan dengan kelompok tersebut.

c) Analisis asosiasi

Analisis asosiasi (association analysis) digunakan untuk menemukan pola yang menggambarkan kekuatan hubungan fitur dalam data. Pola yang ditemukan biasanya merepresentasikan bentuk aturan implikasi atau subset fitur. Tujuannya adalah untuk menemukan pola yang menarik dengan cara yang efisien. Penerapan yang paling dekat dengan kehidupan sehari-hari adalah analisis data keranjang belanja. Sebagai contoh, pembeli adalah ibu rumah tangga yang akan membeli barang kebutuhan rumah tangga di sebuah supermarket. Jika ibu tersebut membeli beras, sangat besar kemungkinannya bahwa ibu itu juga akan membeli barang lain, misalnya minyak, telur, dan tidak mungkin atau jarang membeli barang lain seperti topi atau buku. Dengan mengetahui hubungan yang lebih kuat antara beras dengan telur daripada beras dengan topi, pengecer dapat menentukan barang-barang yang sebaiknya disediakan dalam jumlah yang cukup banyak.

d) Deteksi anomali

Pekerjaan deteksi anomali (anomaly detection) berkaitan dengan pengamatan sebuah data dari sejumlah data yang secara signifikan mempunyai karakteristik yang berbeda dari sisa data yang lain. Data-data yang karakteristiknya menyimpang (berbeda) dari data yang lain disebut outlier. Algoritma deteksi anomali yang baik harus mempunyai laju deteksi yang tinggi dan laju eror yang rendah. Deteksi anomali dapat diterapkan pada sistem jaringan untuk mengetahui pola data yang memasuki jaringan sehingga penyusupan bisa ditemukan jika pola kerja data yang datang berbeda. Perilaku kondisi cuaca mengalami anomali juga dapat dideteksi dengan algoritma ini (Prasetyo, 2012).

Algoritma K-Means

K-Means merupakan salah satu metode clustering yang non hirarki yang berusaha mempartisi data yang ada kedalam bentuk satu atau lebih cluster. Metode ini mempartisi data kedalam cluster sehingga data yang memiliki karakteristik yang sama dikelompokkan kedalam satu cluster yang sama dan data yang mempunyai karakteristik yang berbeda dikelompokkan kedalam cluster yang lain (Prasetyo, 2012).

Metodologi Analisa K-Means

Metodologi dasar analisis Algoritma K-Means Clustering adalah sebagai berikut (Prasetyo, 2012) : Tentukan jumlah kluster (k), tetapkan pusat kluster secara random. Hitung jarak setiap data ke pusat kluster. Kelompokkan data ke dalam kluster yang dengan jarak yang paling pendek. Hitung pusat kluster yang baru. Ulangi langkah 2 sampai dengan 4 hingga sudah tidak ada lagi data yang berpindah ke kluster yang lain. Proses klustering dimulai dengan mengidentifikasi data yang akan dikluster, X_{ij} ($i=1, \dots, n$; $j=1, \dots, m$) dengan n adalah jumlah data yang akan dikluster dan m adalah jumlah variabel. Pada awal iterasi, pusat setiap kluster ditetapkan secara random, C_{kj} ($k=1, \dots, k$; $j=1, \dots, m$). Kemudian dihitung jarak antara setiap data dengan setiap pusat kluster. Untuk melakukan penghitungan jarak data ke- i (x_i) pada pusat kluster ke- k (c_k), diberi nama (d_{ik}), dapat digunakan formula Euclidean seperti pada persamaan berikut yaitu:

$$d_{ij} = \sqrt{(\sum_{k=1}^m X_{ij} - C_{jk})^2}$$

Dimana :

d_{ij} : Jarak objek antar nilai data dan nilai pusat cluster

m : Jumlah dimensi data

X_{ij} : Nilai data dari dimensi ke- k

X_{jk} : Nilai pusat cluster dari dimensi ke- k

Sedangkan formula membangkitkan cendroid baru, yaitu :

$$C = (\sum m) / n$$

Dimana :

C : centroid data

m : anggota data yang termasuk kedalam centroid tertentu

n : jumlah data yang menjadi anggota centroid tertentu

Penambahan langkah perhitungan yaitu Ratio antara (Between Cluster Variation) dengan WCV (Within Cluster Variation) (Jasmir, April 2017) sebagai berikut : Menghitung BCV:

$M1 = (3; 2,59; 1)$ $M2 = (1; 3,46; 2)$ $M3 = (1; 3,02; 3)$

$$d(M1, M2) = \sqrt{((3-1))^2 + ((2,59-3,46))^2 + ((1-2))^2} = 2.40$$

$$d(M1, M3) = \sqrt{((3-1))^2 + ((2,59-3,02))^2 + ((1-3))^2} = 2.86$$

$$d(M2, M3) = \sqrt{((1-1))^2 + ((3,46-3,02))^2 + ((2-3))^2} = 1.09$$

$$BCV = d(m1, m2) + d(m1, m3) + d(m2, m3) = 2.40 + 2.86 + 1.09 = 6,35$$

Maka Nilai Between Cluster Variation nya adalah 6,35.

Menghitung WCV :

$$WCV = \sum (\text{jarak pusat tiap cluster paling minimum})^2$$

Menghitung Rasio :

$$\text{Rasio} = (BCV / WCV)$$

METODE PENELITIAN

Analisa Kebutuhan

Pada penelitian ini, penulis menggunakan alat penelitian berupa perangkat keras dan perangkat lunak.

Perangkat Keras	Perangkat Lunak
1. Laptop Lenovo AMD A8, RAM 4 Gb, 2,5 Ghz	1. Sistem Operasi Windows 8.1 2. Browser Mozilla Firefox dan Google Chrome 3. Xampp 4. Notepad ++

Tabel 1. Analisa Kebutuhan Perangkat Lunak

Teknik Analisa

Adapun teknik analisa yang penulis lakukan dengan melakukan beberapa cara dan beberapa metode berikut analisa yang penulis lakukan demi memudahkan untuk melakukan proses penulisan data yang akan diperoleh guna mengetahui data penyakit tersebut yang nantinya akan dijadikan sebagai sampling dalam metode K-Means sebagai berikut:

a. Data Primer

Data Primer (primary data): data primer adalah data yang dikumpulkan sendiri oleh perorangan/suatu organisasi secara langsung dari objek yang diteliti dan untuk kepentingan studi yang bersangkutan yang dapat berupa interview, observasi

b. Metode Pengumpulan Data

Adapun teknik pengumpulan data yang penulis gunakan sebagai berikut:

- 1) Wawancara, dengan cara mengajukan pertanyaan langsung kepada pengurus Dinas Kesehatan Tangerang Selatan untuk mendapatkan informasi yang lebih akurat.
- 2) Studi Pustaka, dengan cara mempelajari dan membaca literature - literatur, catatan - catatan, dan laporan-laporan yang ada hubungannya dengan masalah yang berhubungan dengan permasalahan yang menjadi obyek penelitian.
- 3) Observasi, pada tahap ini penulis memperoleh berbagai data secara pengamatan dan peninjauan langsung terhadap objek penelitian untuk mengetahui gambaran yang terjadi pada Dinas Kesehatan Tangerang Selatan.

Variabel dan Skala Pengukuran

Variabel yang akan digunakan sebagai data proses sampling dalam K-Means dalam penelitian ini penulis mengambil sebagai proses awal untuk menghasilkan nilai persentase untuk mengetahui seberapa besar jumlah penduduk yang terserang penyakit dan berapa besar persentase penyakit yang sering timbul dimasyarakat. Dengan demikian variabel yang diambil salah satunya adalah penduduk, penderita dengan skala pengukuran C1, C2, dan C3.

HASIL DAN PEMBAHASAN

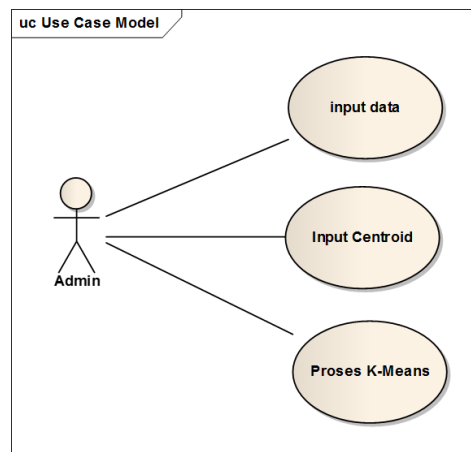
Adapun desain yang akan dirancang untuk menunjang dalam penelitian ini penulis melakukan desain rancangan dibantu dengan *tools* UML sebagai sarana untuk membangun sebuah aplikasi pengukuran dengan metode K-Means, dengan demikian guna memenuhi peneleitian tersebut maka desain yang ingin dirancang sebagai berikut:

Desain Rancangan Sistem

Desain perancangan sistem yang akan digunakan penulis adalah permodelan berorientasi objek, dimana dalam prosesnya akan di gambarkan melalui UML (Unified Modeling Language). Adapun desain dari permodelan sistem pengelompokan daerah endemic adalah sebagai berikut :

Use Case Diagram Pengelompokan K-Means

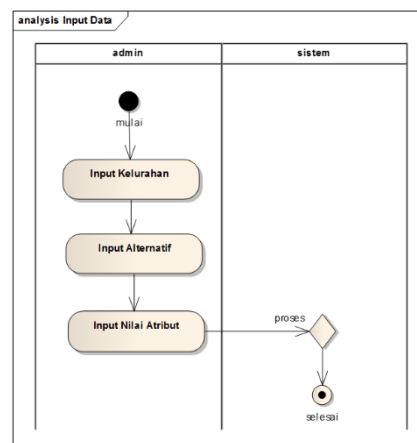
Use case diagram yang dirancang dibawah ini adalah menggambarkan suatu aktivitas antara aplikasi dan *user*, sebagai berikut:



Gambar 2 Use Case Diagram

Activity Diagram Pengelompokan

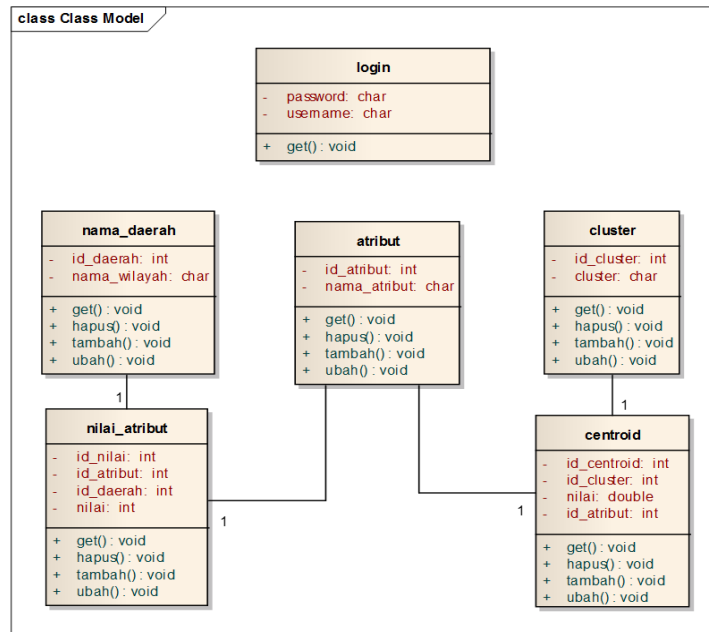
Berikut adalah bentuk rancangan *tools* dari activity diagram yang menggambar secara detail dari apa yang terjadi pada *use case*.



Gambar 3 Activity Diagram Pengelompokan

Class Diagram

Berikut merupakan salah satu proses *tools* desain rancangan aplikasi lanjutan dari setelah proses *Activity* berikut gambarnya:



Gambar 4 Class Diagram Aplikasi K-Means

Data Perhitungan

K-Means merupakan salah satu metode clustering yang non hirarki yang berusaha mempartisi data yang ada kedalam bentuk satu atau lebih cluster. Metode ini mempartisi data kedalam cluster sehingga data yang memiliki karakteristik yang sama dikelompokkan kedalam satu cluster yang sama dan data yang mempunyai karakteristik yang berbeda dikelompokkan kedalam cluster yang lain. Untuk menentukan cluster dalam penelitian ini pusat awal random dari data yang ada didapat titik pusat dari setiap pengclusteran dapat dilihat dalam tabel berikut:

Titik Pusat Awal	Kelurahan	DBD	Penduduk
C1	Rawa Buntu	25	33768
C2	Pd. Ranji	18	39956
C3	Serua Indah	5	19796

Tabel 1 Titik Pusat Centroid Awal

Berikut contoh dari sampel yang dihitung berdasarkan cluster dengan rumus Euclidean, sebagai berikut:

$$d_{ij} = \sqrt{\sum_{k=1}^m X_{ij} - C_{jk}^2}$$

$$d(C 1, data 1) = \sqrt{(data 1 dimensi 1 - C1 dimensi 1)^2 + (data 1 dimensi 2 - C1 dimensi 2)^2}$$

$$d(c1, data 1) = \sqrt{(12 - 25)^2 + (38318 - 33768)^2}$$

$$d(c1, data 1) = \sqrt{165 + 20702500}$$

$$d(c1, data 1) = \sqrt{20702665}$$

$$d(c1, data 1) = 4450.0185713907$$

$$d(C 2, data 1) = \sqrt{(data 1 dimensi 1 - C1 dimensi 1)^2 + (data 1 dimensi 2 - C1 dimensi 2)^2}$$

$$d(c2, data 1) = \sqrt{(12 - 18)^2 + (38318 - 39956)^2}$$

$$d(c2, data 1) = \sqrt{36 + 2683044}$$

$$d(c2, data 1) = \sqrt{2683080}$$

$$d(c2, data 1) = 1638.0109889741$$

$$d(C 3, data 1) = \sqrt{(data 1 dimensi 1 - C1 dimensi 1)^2 + (data 1 dimensi 2 - C1 dimensi 2)^2}$$

$$d(c3, data 1) = \sqrt{(12 - 5)^2 + (38318 - 19796)^2}$$

$$d(c3, data 1) = \sqrt{49 + 343064484}$$

$$d(c3, data 1) = \sqrt{343064530}$$

$$d(c3, data 1) = 18522.001322751$$

Hasil Perbandingan

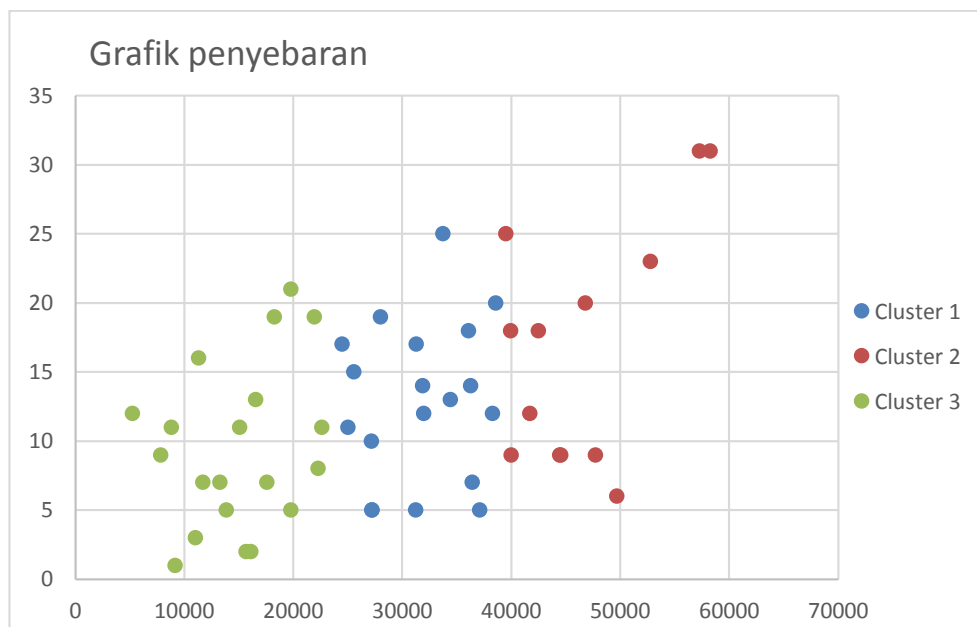
Hasil yang didapat berikut ini merupakan proses dari beberapa sampling dengan melakukan perbandingan dari sekian banyak angka yang didapat berdasarkan kriteria penyebaran penyakit dengan mengcluster, sebagai berikut :

Proses	WCV	BCW	Rasio
Iterasi 1	2315459644.611	40320.022577865	1.7413398964523

Iterasi 2	1440784367.056	57468.077375346	3.9886660828207
Iterasi 3	1385403142.1625	62176.230557493	4.4879521826723
Iterasi 4	1378406009.0571	63313.878717333	4.5932677528475
Iterasi 5	1378406009.0571	63313.878717333	4.5932677528475

Tabel 2 Tabel Perbandingan Cluster Centroid

Setelah melakukan perbandingan maka akan dapat dilihat dalam bentuk penyebaran melalui grafik, sebagai berikut:



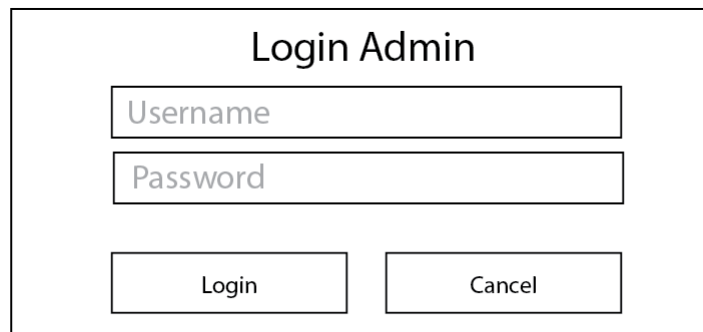
Gambar 5 Grafik Hasil Penyebaran

- C1 = Paku Alam, Pd. Jaya, Parigi Lama, Setu, Ciputat, Serua Indah, LK. Karya, LK. Wetan, Bakti Jaya, Kranggan, Pakulonan, Parigi Baru, Muncul, Buaran, Rawa Mekar Jaya, LK. Gudang Timur, LK. Gudang Barat, Serpong, Cilenggang, Babakan.
- C2 = Pam. Barat, Pam. Timur, Pd. Betung, Jombang, Pd. Ranji, Pd. Kacang Timur, JR. Mangu Barat, JR. Mangu Timur, Pd. Karya, Pd. Kacang Barat, Pd. Benda, Benda Baru, Kedaung.
- C3 = Pisangan, Cirendeu, Paku Jaya, Serua, Cabe Ilir, Cabe Udik, Pd. Aren, Pd. Jagung, Sawah Baru, Rawa Buntu, Ciater, Bambu Apus, Cipayung, Jelumpang, Pd. Jagung Timur, Pd. Pucung, Kademangan, Rengas, Rempoa.

Rancangan Antar Muka Program

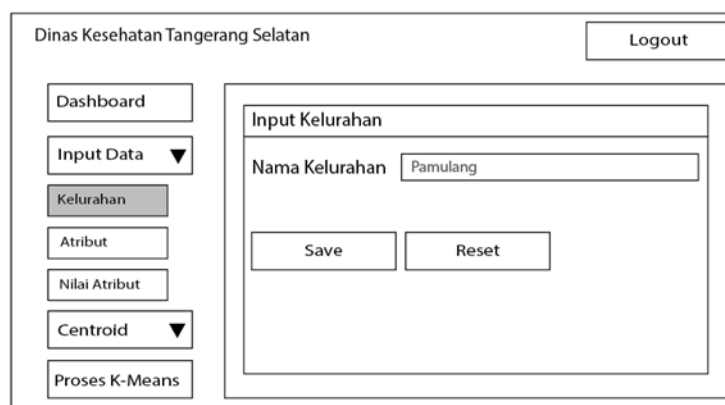
Dari proses yang telah dibuat dari mulai pembuatan *tools* dengan UML ke perhitungan dengan demikian dibuatlah Perancangan antar muka digunakan sebagai gambaran untuk membuat antar muka sistem yang akan dibuat. Berikut ini adalah rancangan yang dibuat :

Form Rancangan Menu Login Admin



The image shows a simple login form titled "Login Admin". It features two input fields: "Username" and "Password". Below the input fields are two buttons: "Login" and "Cancel".

Form Rancangan Menu Input



The image shows a more complex form titled "Dinas Kesehatan Tangerang Selatan" with a "Logout" button in the top right corner. On the left side, there is a vertical menu with several options: "Dashboard", "Input Data" (with a dropdown arrow), "Kelurahan" (highlighted), "Atribut", "Nilai Atribut", "Centroid" (with a dropdown arrow), and "Proses K-Means". The main content area is titled "Input Kelurahan" and contains a text input field labeled "Nama Kelurahan" with the value "Pamulang" entered. Below this input field are two buttons: "Save" and "Reset".

KESIMPULAN DAN SARAN

Kesimpulan

Kesimpulan yang dapat diperoleh pada pengembangan sistem dengan menggunakan metode klasterisasi k-means dapat digunakan untuk mengelompokkan 3 daerah yang dimana C1 di Kelurahan Rawa Buntu, C2 di Kelurahan Pondok Ranji dan C3 di Kelurahan Serua Indah yang ditentukan secara random dari 52 Kelurahan. Prototipe aplikasi clustering untuk menganalisis daerah endemik penyakit Demam Berdarah dengan menggunakan metode klasterisasi k-means telah berhasil dikembangkan dan diharapkan dapat membantu Dinas Kesehatan dalam program penyehatan lingkungan. Dari hasil klasterisasi C1 ada 20 Kelurahan, C2 ada 13 Kelurahan, dan C3 ada 19 Kelurahan dari 52 Kelurahan.

Saran

Saran yang diberikan untuk pengembangan aplikasi clustering daerah penyebaran penyakit Demam Berdarah ini dapat dtindak lebih lanjut guna mencari nilai cluster pertama dibantu dengan algoritma tertentu, supaya hasil cluster lebih optimal dan dapat digunakan ke arah yang lebih baik.

DAFTAR PUSTAKA

- Dhuhita, W. M. (Desember, 2015). Clustering Menggunakan Metode K-Means untuk Menentukan Gizi Balita. *Jurnal Informatika*, Vol 15, No 2, 160.
- Fina Nasari, Charles Jhony Manto Sianturi. (Desember, 2016). Penerapan Algoritma K-Means Clustering Untuk Pengelompokan Penyebaran Diare Di Kabupaten. *Cogito Smart Journal*, Vol 2, No 2, 108.
- Jasmir. (April 2017). Analisis Profil Alumni Dengan Klasterisasi K-Means. *MEDIA SISFO*, Vol 1, No 1, 722.
- Kristanto, A. (2007). *Perancangan Sistem Informasi dan Aplikasi* (Vol. 12). Klaten: Gava Media.
- Munawar. (2005). *Pemodelan Visual dengan UML*. Yogyakarta: Graha Ilmu.
- Mustakini, J. H. (2010 : 59). *Analisis dan Desain*.
- Nugroho, B. (2008). *Aplikasi Pemrograman Web Dinamis dengan PHP dan MySQL*. Yogyakarta: Gava Media.
- Prasetyo, E. (2012). *Data Mining - Konsep dan Aplikasi Menggunakan MATLAB*. Gresik: Andi.
- Pressman. (2012). *Rekayasa Perangkat Lunak*. Yogyakarta: Andi.
- Sari, E. A. (n.d.). *Penerapan Algoritma K-Means Untuk Menentukan Tingkat Kesehatan Bayi Dan Balita Pada Kabupaten dan Kota Di Jawa Tengah*.
- Tan P, Steinbach M, Kumar V. (2006). *Introduction to Data Mining*. In E. Prasetyo, *Data Mining - Konsep dan Aplikasi Menggunakan MATLAB* (p. 2). Gresik: ANDI Yogyakarta.
- Widoyono. (2008). *Penyakit Tropis Epidemiologi, Penularan, Pencegahan, dan Pemberantasan*. Jakarta: Erlangga.
- <https://kabartangsel.com/direktori-tangsel/peta-kota-tangerang-selatan/>
- <https://id.wikipedia.org/wiki/HTML5>
- <https://tangselkota.bps.go.id/linkTableDinamis/view/id/49>