

## Aplikasi Data Mining untuk Clustering Daerah Penyebaran Penyakit Diare di DKI Jakarta Menggunakan Algoritma *K-MEANS*

Teguh Budi Santoso<sup>1</sup>, Ahmad Sofian Hadi<sup>2</sup>

<sup>1,2</sup>Fakultas Teknik, Universitas Satya Negara Indonesia  
Jl. Arteri Pondok Indah No. 11, Jakarta Selatan  
e-mail : <sup>1</sup>teguh.santos12@gmail.com, <sup>2</sup>ahmadsof48@gmail.com

### Abstract

Diarrhea is a disease that makes sufferers defecate more than 4 times, diarrhea usually strikes all people, including toddlers, teens and adults. The increasing population density in DKI Jakarta Province in 2017 around 10,177,924 inhabitants spread over 6 cities consisting of 44 districts and 267 villages. To be able to see the area of diarrhea distribution, it is necessary to make a grouping based on the attributes used consisting of JSP, JSSP, SHARING and OD in order to obtain the center of the diarrhea spread point. K-Means algorithm is very suitable in clustering the spread of diarrhea Kelurahan. To determine the centroid starting point based on diarrhea sufferers as a parameter in determining C1, C2 and C3 which get C1 is Duren Tiga Village, C2 is Kebon Baru Village and C3 is Bendungan Hilir Village. The results of grouping based on the attributes used get C1 totaling 11 Kelurahans, C2 totaling 34 Kelurahans and C3 totaling 43 Kelurahans from 88 Kelurahan data used in which C1 is categorized as a diarrhea prone point, C2 is categorized as a possible spread of diarrhea and C3 is categorized as a Kelurahan that is safe from spread of diarrhea.

Keywords: Diarrheal Disease, K-Means Application

### Abstrak

Diare merupakan penyakit yang membuat penderitanya buang air besar lebih dari 4 kali, biasanya diare menyerang ke semua kalangan baik usia balita, remaja dan dewasa. Semakin meningkatnya kepadatan jumlah penduduk di Provinsi DKI Jakarta pada tahun 2017 sekitar 10.177.924 jiwa yang tersebar 6 Kota terdiri dari 44 Kecamatan serta 267 Kelurahan. Untuk dapat melihat daerah penyebaran diare maka perlu dibuat sebuah pengelompokan berdasarkan atribut yang dipakai terdiri dari JSP, JSSP, SHARING dan OD agar dapat memperoleh pusat titik penyebaran diare. Algoritma K-Means sangat cocok dalam pengelompokan Kelurahan penyebaran diare. Untuk menentukan titik awal centroid berdasarkan penderita diare sebagai Parameter dalam menentukan C1, C2 dan C3 yang dimana mendapatkan C1 adalah Kelurahan Duren Tiga, C2 adalah Kelurahan Kebon Baru dan C3 adalah Kelurahan Bendungan Hilir. Hasil dari pengelompokan berdasarkan atribut yang dipakai mendapatkan C1 berjumlah 11 Kelurahan, C2 berjumlah 34 Kelurahan dan C3 berjumlah 43 Kelurahan dari 88 data Kelurahan yang digunakan dimana C1 dikategorikan sebagai titik rawan diare, C2 dikategorikan sebagai kemungkinan penyebaran diare dan C3 dikategorikan sebagai Kelurahan yang aman dari penyebaran diare.

Kata Kunci : Penyakit Diare, K-Means Aplikasi

## I. PENDAHULUAN

Diare adalah penyakit yang membuat penderitanya sering buang air besar lebih dari 4 kali biasanya menyerang ke semua kalangan usia baik usia remaja, dewasa dan balita dengan kondisi buang-buang air besar menjadi encer dapat berwarna hijau atau dapat pula bercampur lender darah atau lender saja. Sementara itu penyebab dari penyakit Diare [12] itu sendiri bisa jadi antara lain virus yaitu *Rotavirus* (40-60%), *Bakteri Escherichia Coli* (20-30%), *Shigella Sp.* (1-2%) dan *Parasit Entamoeba Hystolitica* (<1%), diare juga sering kali terjadi karena beberapa factor, seperti keracunan makanan, infeksi bakteri, pemanis buatan, Fruktosa dan-lain lain. Diare dapat terjadi karena *Higiene* dan *Sanitasi* yang buruk, *Malnutrisi*, lingkungan padat dan sumber daya medis yang buruk [13]. Pada tahun 2016 jumlah penderita Diare semua umur (SU) yang dilayani di sarana kesehatan sebanyak 274.803 penderita [4]. Semakin meningkatnya kepadatan penduduk di Provinsi DKI Jakarta pada tahun 2017 sekitar 10.177.924 jiwa yang tersebar 6 Kota terdiri dari 44 Kecamatan dan 267 Kelurahan [2] dan meningkatnya penderita penyakit Diare karena kurangnya kesadaran masyarakat terhadap pentingnya kesehatan. Pada tahun 2017 di Provinsi DKI Jakarta penderita Diare di fasilitas kesehatan sejumlah 280.104 jiwa dari jumlah penderita yang ditangani sejumlah 250.234 jiwa dalam hitungan persen sebanyak 89,3% [5]. Timbulah masalah yang menyulitkan Dinas Kesehatan dalam mengelompokkan daerah yang akan menghasilkan titik-titik pusat penyebaran penderita penyakit Diare. Perlunya mengklompokkan data kesehatan penduduk untuk menentukan kesehatan masyarakat [1]. Agar program Dinas Kesehatan dalam memberikan edukasi pada titik pusat penyebaran penyakit Diare dapat berjalan dengan baik. Maka dilakukan pengelompokkan berdasarkan Jamban Sehat Permanen (JSP), Akses Jamban Sehat Semi Permanen (JSSP), Masih Numpang ke Jamban Sehat (SHARING), Open Defecation (OD), Penderita Diare.

Adapun pada penelitian terdahulu [3] Algoritma K-Means merupakan metode yang digunakan untuk mengklustering atau mengelompokkan suatu data [6] penyebaran penyakit TBC merupakan penyakit menular yang masih menjadi perhatian dunia hingga saat ini, pada penelitian ini bertujuan mengklompokkan daerah-daerah penyebaran penyebarannya. Adapun pada penelitian [10] melakukan proses klasifikasi kelayakan pemberian kredit pada koperasi, proses klasifikasi data menentukan berdasarkan himpunan data dan atribut menjadikan keputusan pada pola klasifikasi. Sedangkan pada penelitian [9] membahas mengenai masalah penyebaran penyakit diare yang dikelompokkan menggunakan algoritma K-Means dengan menentukan 2 titik centroid yaitu  $c_1$  dan  $c_2$ , hasil pada penelitian ini diperoleh  $c_1$  berada pada kecamatan Batang serangan, brandan barat dan pematang jaya dan  $c_2$  berada pada kecamatan Hinai, untuk kelompok penyebaran diare,  $c_1$  jumlah penderita tingkat menengah,  $c_2$  daerah penyebaran diare. Pada penelitian ini bertujuan akan dilakukan transformasi data untuk menghasilkan 6 kelompok berdasarkan himpunan kecamatan, pengolahan data selanjutnya dilakukan untuk menentukan nika  $k$  dari jumlah cluster yang dibentuk menjadi 3 Cluster yang bertujuan untuk mengklompokkan berdasarkan data kelurahan yang diambil berdasarkan data penderita terbanyak setiap tahunnya yang akan dijadikan data sampling proses penerapan algoritma k-means, yang ada di daerah jakbar. Setelah itu tentukan titik pusat (*centroid*) awal dari setiap cluster. Hasil nya didapat 3 titik centroid yaitu  $c_1$ ,  $c_2$ , dan  $c_3$ . Pada penelitian ini menerapkan Algoritma Kluster K-Means dengan himpunan JSP (Jamban Sehat Permanen), JSSP (Akses Jamban Sehat Semi Pemanen), SHARING (Masih Menumpang ke Jamban ke Jamban Sehat), OD (Open Defecation), Penderita Diare, dengan membuat skala pengukuran data dengan  $C_1, C_2, C_3$ . Bertujuan untuk mengkluster daerah daerah penyebaran penyakit diare di daerah DKI Jakarta.

## II. METODE PENELITIAN

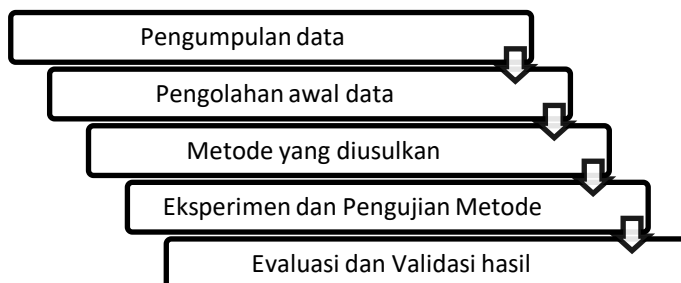
### 2.1 Analisa Kebutuhan

Dalam penelitian ini digunakan metode eksperimen dengan kebutuhan seperti: Pengumpulan data, Pengolahan awal data, Metode yang diusulkan, Eksperimen dan pengujian metode. Dalam penelitian ini akan dibuat suatu aplikasi berbasis *Web*. System yang diterapkan kedalam aplikasi Web nanti akan digunakan dengan skala pengukuran  $C_1, C_2, C_3$  untuk titik cluster dengan himpunan JSP (Jamban Sehat Permanen), JSSP (Akses Jamban Sehat Semi Pemanen), SHARING (Masih Menumpang ke Jamban ke Jamban Sehat), OD (Open Defecation), Penderita Diare. Guna menunjang pada penelitian ini dibutuhkan satu unit laptop. Dan aplikasi Weka untuk menguji data training set.

### 2.2 Perancangan Penelitian

Pada penelitian ini Algoritma K-Means digunakan untuk mengklustering data-data penyebaran penyakit

diare di DKI Jakarta. Yang akan dibuat kedalam aplikasi berbasis web dan dilakukan proses perhitungan secara manual dan menggunakan aplikasi matlab guna untuk memvalidasikan hasil perhitungan[11]. Tahapan pada penelitian ini adalah Pengumpulan data, Pengolahan awal data, Metode yang dikembangkan, Eksperimen dan Pengujian metode serta Evaluasi. Seperti terlihat pada gambar 1 Tahapan Penelitian.



Gambar 1 Tahapan Penelitian

a. Pengumpulan Data

Pengumpulan data pada penelitian ini dilakukan yang bersumber dari jurnal, paper, buku serta beberapa informasi yang berhubungan dengan penelitian dan kutipan data dari berbagai instansi seperti kementerian kesehatan, Kelurahan, Kecamatan, karena data yang digunakan dalam penelitian ini adalah sekunder.

b. Pengolahan Data

Yang dimaksud dengan pengolahan data pada penelitian ini adalah proses pengelompokan data data yang sudah didapatkan sebelumnya untuk menentukan titik centroid berdasarkan dengan 6 himpunan yang akan digunakan.

c. Metode yang diusulkan

Pada penelitian ini metode atau Algoritma Klustering yang digunakan adalah K-Means

d. Eksperimen dan Pengujian Metode

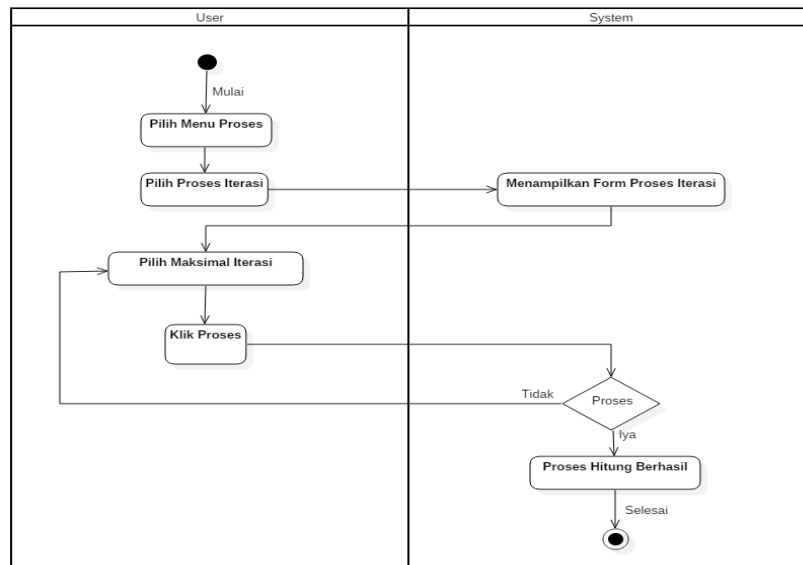
Eksperimen dan pengujian metode dalam penelitian ini adalah proses penghitungan dan simulasi dengan memasukan nilai dari masing-masing variabel input yang digunakan sebagai parameter ukur untuk mencari metode yang terbaik

### 2.3 Perancangan Perangkat Lunak

Pada tahap ini dijelaskan rancangan Aplikasi untuk mengkluster penyebaran penyakit diare menggunakan algoritma K-Means dengan menggunakan design UML(*Unified Modeling Language*)[7].

1. Activity Diagram

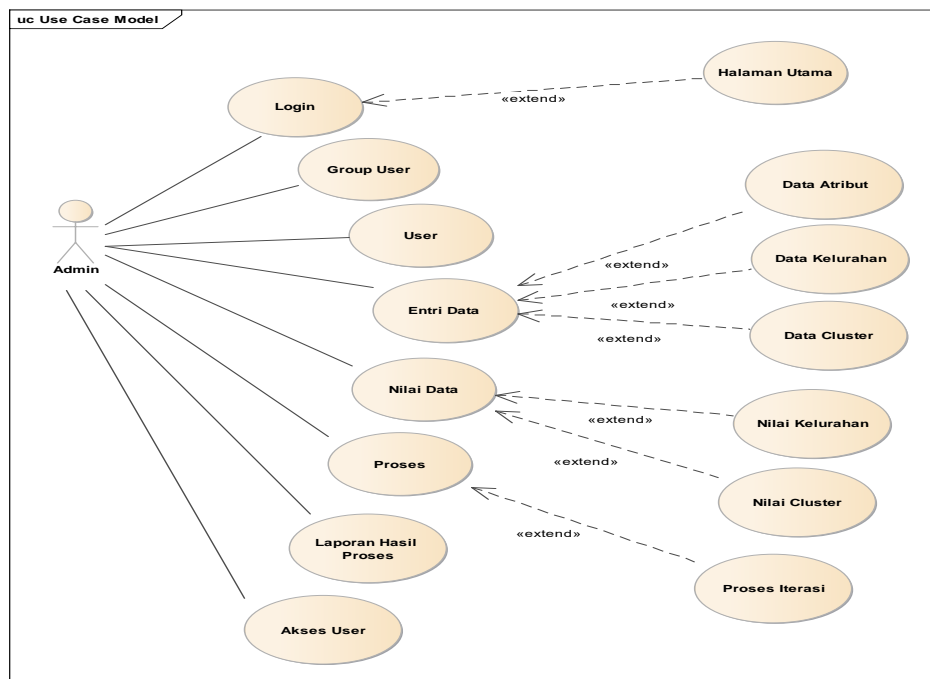
Activity yang menggambarkan proses bisnis urutan aktifitas dalam sebuah proses system. Dapat dilihat pada gambar 2 aktifitas proses Iterasi



Gambar 2. Activity Diagram Proses Iterasi

2. Use Case Diagram

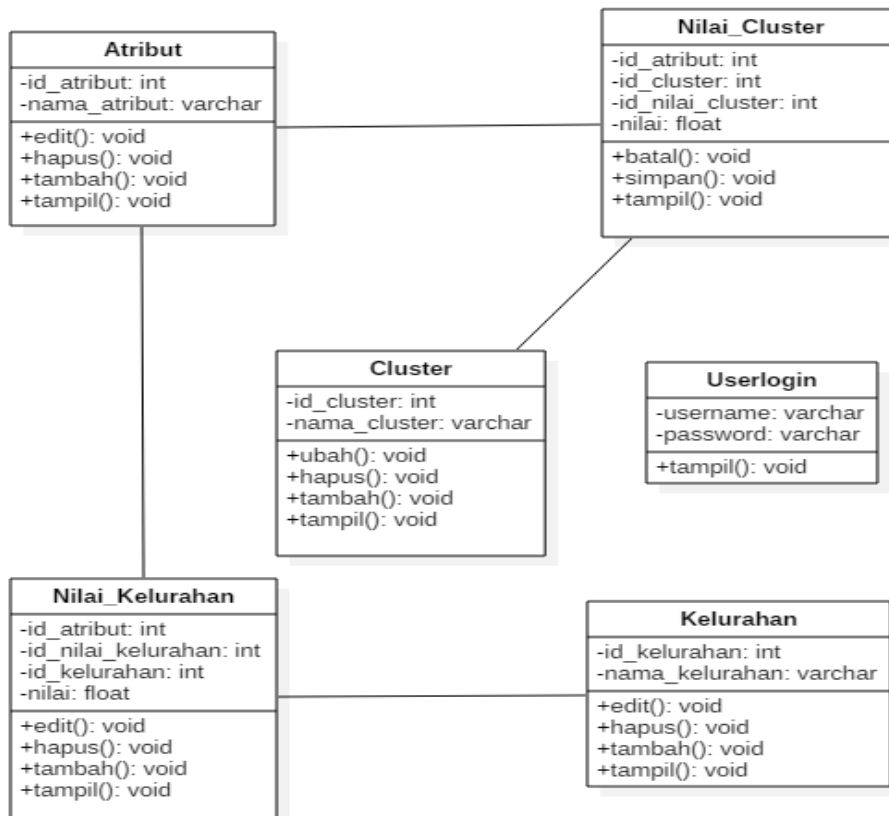
Use case diagram menjelaskan fungsi dari admin, user dan umum yang diharapkan dari sebuah system. Dapat dilihat pada gambar 3 Use Case Diagram system.



Gambar 3 Use Case Diagram system

3. Class Diagram

Class diagram pada aplikasi pengelompokan daerah, sebagai berikut:



Gambar 3. Class Diagram

#### IV. HASIL DAN PEMBAHASAN

Dari hasil simulasi dilakukan pada penelitian ini dengan menerapkan metode clustering K-Means berdasarkan himpunan cluster dimana titik pusat sudah ditentukan dari beberapa atribut diantaranya: JSP, JSSP, SHARING, OD, Penderita Diare untuk dikelompokkan kedalam suatu cluster yang diperoleh titik terdekat data. Dimana dapat dilihat pada tabel 1 data laporan penyakit diare 2018.

Tabel 1. Data Laporan Penyakit Diare 2018

No	Kelurahan	JSP	JSSP	SHARING	OD	Penderita Diare
1	Pulau Harapan	391	0	150	15	120
2	Pulau Kelapa	1077	596	64	113	377
3	Pulau Pari	368	349	164	17	335
4	Pulau Tidung	1090	66	7	0	225
5	Cengkareng Barat	20252	1166	115	116	6.445
6	Cengkareng Timur	23776	538	1971	131	549
7	Grogol	4829	977	2712	78	913
	Dst....					

#### Menentukan Titik Jumlah Cluster dan Tentukan Titik Pusat Awal

Dalam penelitian ini titik pusat awal dipilih secara random dan mendapatkan titik pusat dari setiap cluster dapat dilihat pada tabel 2.

Tabel 2. Titik Pusat Centroid Awal

Titik Pusat Awal	Kelurahan	JSP	JSSP	SHARING	OD
CI	Duren Tiga	8092	0	0	650
C2	Kebon Baru	4628	0	0	1196
C3	Bendungan Hilir	1037	0	0	0

**Hitung dengan Rumus**

Dalam perhitungan mencari titik terdekat yang sebelumnya titik terdekat dengan sebelumnya titik centroid telah ditentukan dengan rumus

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

$$d(\text{data 1}, C1) = \sqrt{(\text{data 1 dimensi 1} - C1 \text{ dimensi 1})^2 + (\text{data 1 dimensi 2} - C1 \text{ dimensi 2})^2 + (\text{data 1 dimensi 3} - C1 \text{ dimensi 3})^2 + (\text{data 1 dimensi 4} - C1 \text{ dimensi 4})^2}$$

dibawah ini perhitungan data 1 dengan titik centroid C1:

$$\begin{aligned} d(\text{data 1}, C1) &= \sqrt{(391 - 8092)^2 + (0 - 0)^2 + (150 - 0)^2 + (15 - 650)^2} \\ d(\text{data 1}, C1) &= \sqrt{59.305.402 + 0 + 22.500 + 403.225} \\ d(\text{data 1}, C1) &= \sqrt{59.731.126} \\ d(\text{data 1}, C1) &= 7728,59 \end{aligned}$$

Dengan titik centroid c2:

$$\begin{aligned} d(\text{data 1}, C2) &= \sqrt{(391 - 4628)^2 + (0 - 0)^2 + (150 - 0)^2 + (15 - 1196)^2} \\ d(\text{data 1}, C2) &= \sqrt{17.952.169 + 0 + 22.500 + 1.394.761} \\ d(\text{data 1}, C2) &= \sqrt{19.369.430} \\ d(\text{data 1}, C2) &= 4401,07 \end{aligned}$$

Dengan titik centrod C3

$$\begin{aligned} d(\text{data 1}, C3) &= \sqrt{(391 - 1037)^2 + (0 - 0)^2 + (150 - 0)^2 + (15 - 0)^2} \\ d(\text{data 1}, C3) &= \sqrt{417.316 + 0 + 22.500 + 225} \\ d(\text{data 1}, C3) &= \sqrt{440.041} \\ d(\text{data 1}, C3) &= 663,356 \\ \text{Dst.....} \end{aligned}$$

**Kelompok Cluster yang terpendek**

Dalam penelitian ini menentukan titik terdekat dapat dilihat berdasarkan cluster 1,2 dan cluster 3, dimana nilai terkecil akan dijadikan titik terdekat cluster dan kelompok cluster terpendek dapat dilihat pada tabel 3 Iterasi 1.

Tabel 3 Iterasi 1

No	Nama Kelurahan	Cluster1	Cluster2	Cluster3	Terdekat
1	Pulau Harapan	7728.59	4401.07	663.356	663.356
2	Pulau Kelapa	7061.01	3760.56	611.294	611.294
3	Pulau Pari	7759.48	4436.93	772.365	772.365
4	Pulau Tidung	7032.42	3735.27	849.353	849.353
88	Dst..				

Dilanjutkan dengan cluster 1 seperti ditunjukkan pada Tabel 4

Tabel 4 Cluster 1

Nama Kelurahan	C1			
	JSP	JSSP	SHARING	OD
Ancol	7302	0	1483	516
Angke	2676	2341	4064	1848
Balekambang	0	0	0	0
Bali Mester	1229	2	520	342
Bambu Apus	5427	2560	194	10
Bangka	6506	1557	162	148
Dst.....				

Dan seterusnya sampai dengan Iterasi ke-3 dan Cluster ke-3

Menghitung Centroid Baru

Berikut ini proses membangkitkan centroid baru dengan rumus sebagai berikut:

$$C = \frac{\sum m}{n}$$

Menghitung centroid C1 data 1 (D1)

C1, D1

$$= \frac{20252 + 23776 + 7481 + 7304 + 15917 + 12392 + 19844 + 8969 + 10164 + 7441 + 6833 + 9271 + 11502 + 6965 + 7892 + 18672 + 12139 + 11254 + 12611 + 6502 + 6506 + 8092 + 7893 + 12426 + 17956 + 17170 + 11946 + 10515 + 11225 + 8058 + 8190 + 9813 + 6417 + 18362 + 7862 + 18538 + 13895 + 13057 + 8011 + 18780 + 7302 + 26999 + 14411 + 6888}{44}$$

$$C1, D1 = \frac{527493}{44}$$

$$C1, D1 = \frac{527493}{44}$$

$$C1, D1 = 11988.5$$

Menghitung centroid C1 data 2 (D2)”

$C1, D2$

$$= \frac{1166 + 538 + 180 + 576 + 2471 + 2192 + 19844 + 732 + 0 + 0 + 868 + 4753 + 2095 + 60 + 20 + 0 + 0 + 0 + 0 + 5458 + 1557 + 0 + 0 + 5000 + 0 + 180 + 1164 + 0 + 0 + 0 + 0 + 0 + 433 + 27 + 0 + 0 + 167 + 0 + 0 + 0 + 0 + 0 + 0 + 0 + 668}{44}$$

$$C1, D2 = \frac{50149}{44}$$
$$C1, D2 = 1139.75$$

**Dst....**

Ulangi Langkah Ke-2

Pada proses ini mengulangi untuk menentukan cluster pada setiap iterasi sampai iterasi ke-n dengan posisi data cluster tertentu tidak berubah setelah dihitung sampai iterasi ke-11 tidak ada data yang berpindah cluster, maka iterasi berhenti pada iterasi ke -11

Hasil Centroid

Hasil akhir perhitungan berhenti iterasi ke 11, dapat dilihat di tabel 5. Dan seterusnya, sebagai berikut :

Tabel 5 Hasil Iterasi Terakhir C1

C1
Kelurahan
Cengkareng Barat
Cengkareng Timur
Duri Kepa
Kalideres
Ciganjur
Pesangrahan
Dst

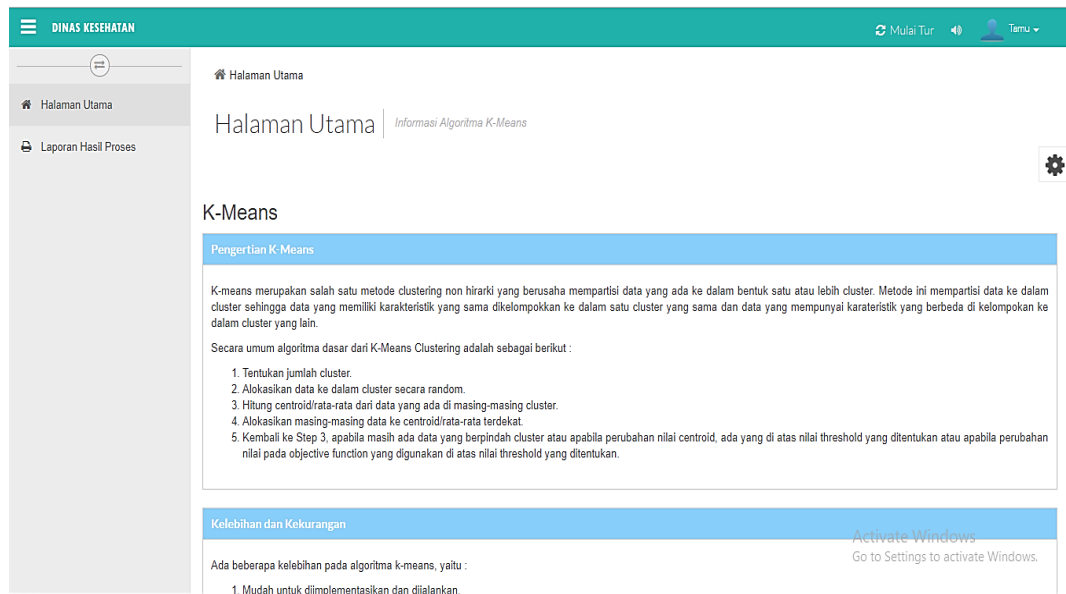
#### 4.1. Tampilan Interface Aplikasi

Berikut ini gambaran tampilan interface aplikasi untuk mempermudah pengguna dalam menggunakan system. Berikut ini tampilan interface berbasis web:

##### a. Tampilan Halaman Utama

Pada gambar dibawah ini menjelaskan tampilan utama untuk digunakan tanpa harus login untuk mengakses aplikasi.

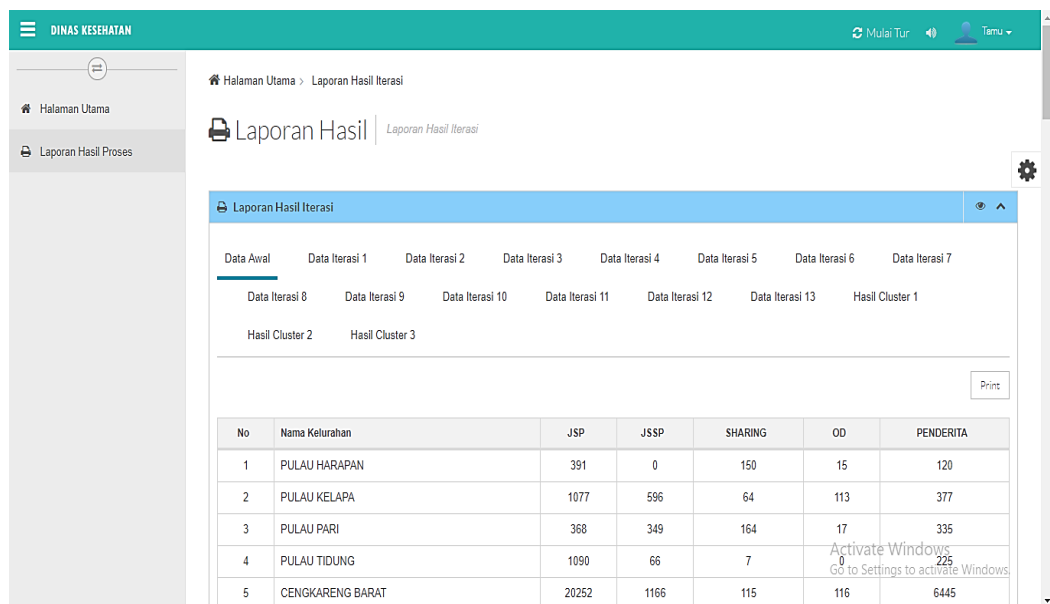




Gambar 4. Tampilan Utama Umum

**b. Tampilan Laporan Hasil Iterasi Umum**

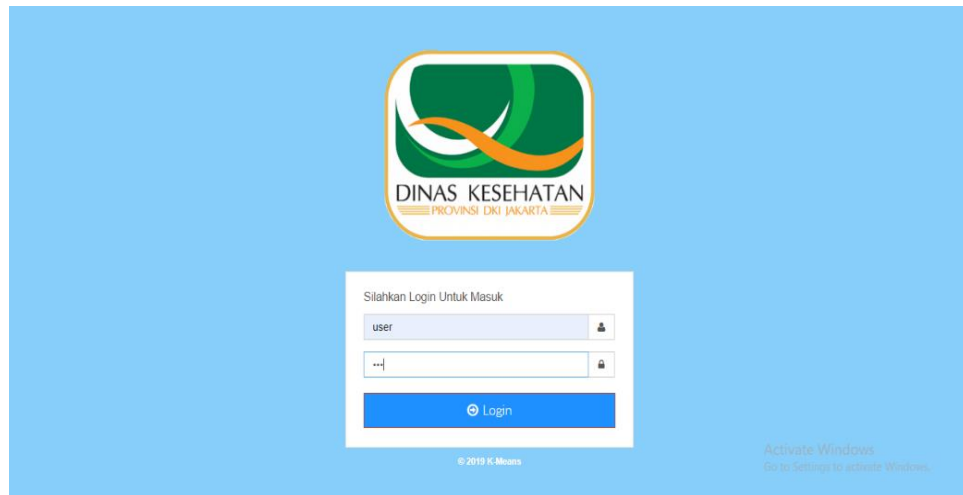
Pada tampilan menu hasil Iterasi umum user dapat melihat hasil laporan yang sudah diProses menggunakan algoritma K-Means.



Gambar 5. Tampilan Hasil Iterasi Umum

**c. Tampilan Halaman Login User**

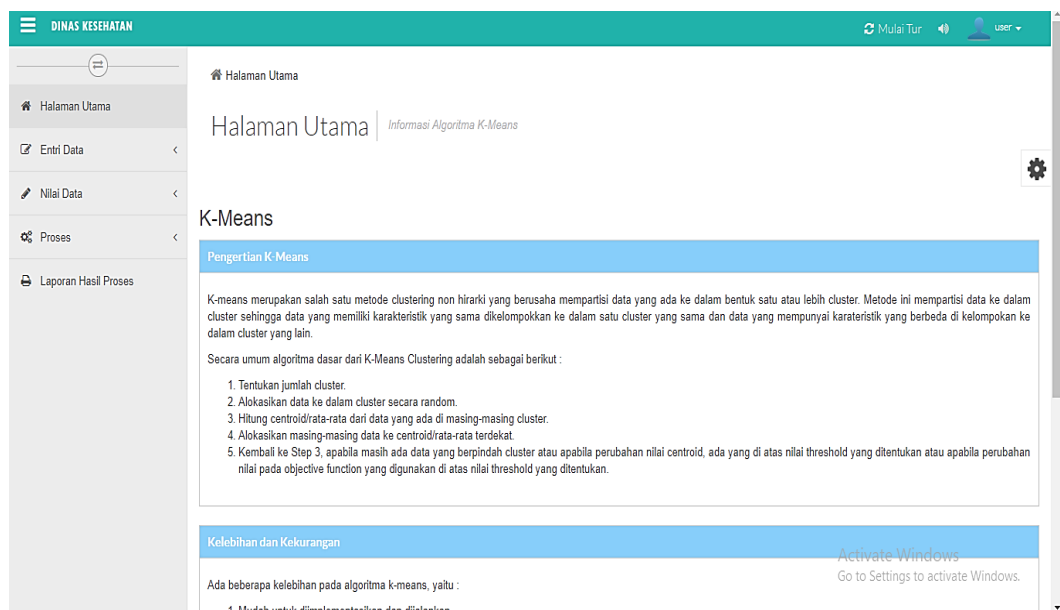
Pada gambar 6 tampilan halaman Login User untuk memulai proses data pada penerapan algoritma K-Means.



Gambar 6. Tampilan Menu Login

d. **Tampilan Utama User**

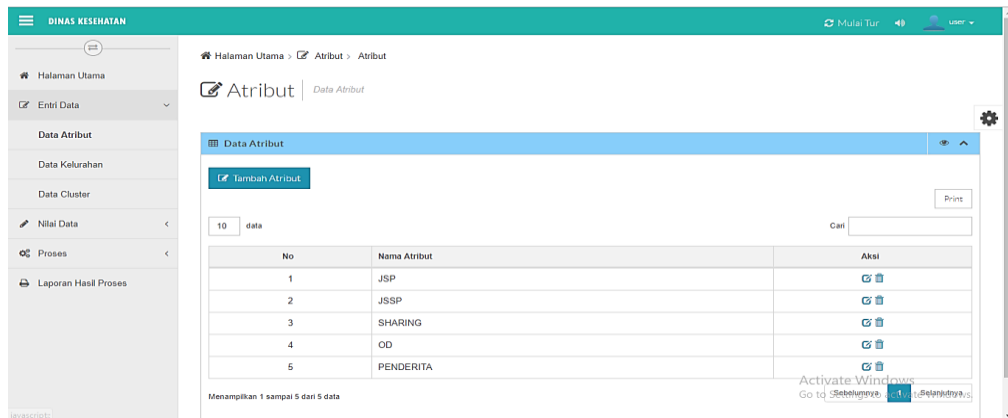
Pada gambar 7 merupakan gambar tampilan menu yang ada di user, setelah melakukan login.



Gambar 7. Tampilan Menu Utama User

e. **Tampilan Entry Data Atribut**

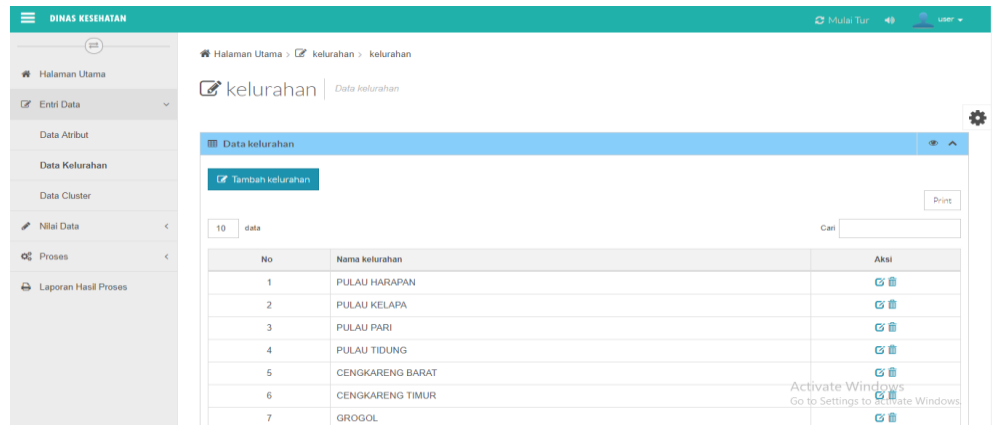
Pada gambar 8 tampilan menu entry data atribut yang berfungsi untuk memasukan data-data atribut dalam memproses klastering berdasarkan atribut atau himpunan yang sudah ditentukan dalam penerapan algoritma K-Means.



Gambar 8. Tampilan Entry Data Atribut

f. Tampilan Menu Data Kelurahan

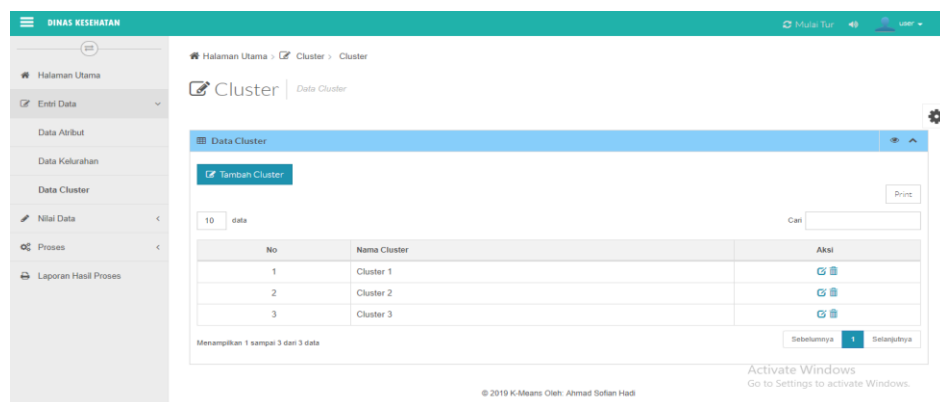
Pada gambar 9 tampilan menu untuk mengentry data kelurahan Ditunjukkan pada gambar dibawah ini.



Gambar 9. Tampilan Entry Data Kelurahan

g. Tampilan Entry Data Cluster

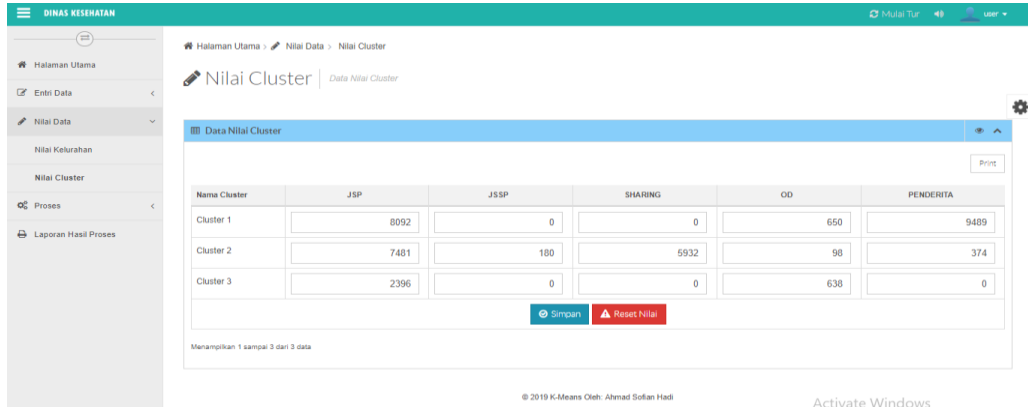
Pada gambar 10 menu data cluster menjelaskan proses entry data cluster. Ditunjukkan pada gambar 10



Gambar 10. Tampilan Data Entry Cluster

**h. Tampilan Entry Data Nilai Cluster**

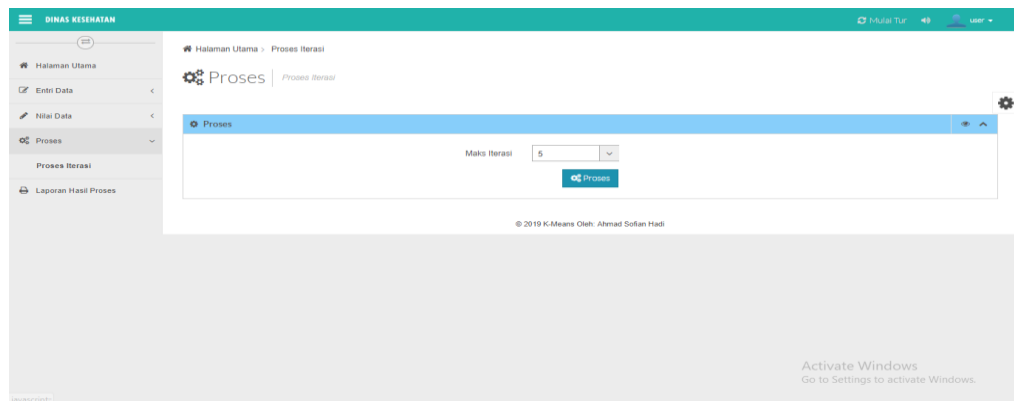
Pada gambar 11 tampilan entry data nilai yang akan dimasukkan dan dapat diproses dengan algoritma k-means ditunjukkan pada gambar 11.



**Gambar 11. Tampilan Entry Data Nilai Cluster**

**i. Tampilan Menu Proses Iterasi**

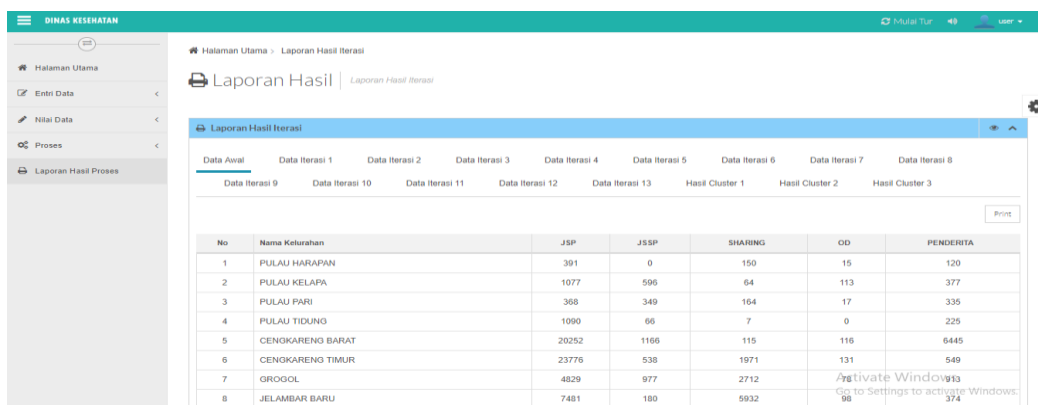
Pada gambar 12 tampilan menu proses iterasi menjelaskan halaman untuk proses data dengan memproses data dengan memasukkan maksimal iterasi lalu klik proses.



**Gambar 12. Tampilan Menu Proses Iterasi**

**j. Tampilan Hasil Proses**

Terlihat pada gambar 13 tampilan hasil proses, dimana menjelaskan halaman laporan hasil setelah sudah melakukan proses sebelumnya.



No	Nama Kelurahan	JSP	JSSP	SHARING	OD	PENDERITA
1	PULAU HARAPAN	391	0	150	15	120
2	PULAU KELAPA	1077	596	64	113	377
3	PULAU PARI	368	349	164	17	335
4	PULAU TIDUNG	1090	66	7	0	225
5	CENOKARENG BARAT	20252	1166	115	116	6445
6	CENOKARENG TIMUR	23776	538	1971	131	549
7	GROGOL	4829	977	2712		
8	JELAMBAR BARU	7481	180	5932		

Gambar 13. Tampilan Menu Laporan Hasil Proses

## V. KESIMPULAN

Berdasarkan proses yang telah dilakukan dalam menerapkan algoritma k-means, dengan himpunan atribut JSP (Jamban Sehat Permanen), JSSP (Akses Jamban Sehat Semi Pemanen), SHARING (Masih Menumpang ke Jamban ke Jamban Sehat), OD (Open Defecation), Penderita Diare. Pada penelitian ini berhasil menentukan kluster yang terdiri dari C1, C2 dan C3, dimana C1 berjumlah 11 Kelurahan, C2 berjumlah 34 dan C3 berjumlah 43 kelurahan dengan data awal yang digunakan berjumlah 88 data kelurahan. Dimana berdasarkan klustering bahwa C1 dikategorikan sebagai kelurahan yang memiliki titik rawan diare, sedangkan c2 dikategorikan sebagai kelurahan yang daerah kemungkinan rawan penyebaran diare dan sedang c3 dikategorikan sebagai kelurahan yang dikategorikan aman dari penyebaran diare. Pada penelitian ini sudah dibuat aplikasi berbasis web yang dapat dijadikan proses dalam penerapan algoritma k-means dalam mengetahui kluster atau pengelompokan data daerah-daerah penyebaran penyakit diare di daerah DKI Jakarta dan memberikan informasi kesehatan kepada masyarakat dalam menjaga lingkungan yang bersih, karena hasil berpusat pada titik rawan penyebaran diare.

## DAFTAR PUSTAKA

- [1] Atthina, N., & Iswari, L. Klasterisasi Data Kesehatan Penduduk Untuk Menentukan Rentang Derajat Kesehatan Dengan Metode K-Means. *Seminar Nasional Aplikasi Teknologi Informasi (SNATI)*. 2014
- [2] BPS PROVINSI DKI JAKARTA. *jakarta.bps.go.id*. Retrieved from BPS Provinsi DKI Jakarta: <https://jakarta.bps.go.id/statictable/2017/01/30/142/jumlah-penduduk-menurut-kelompok-umur-dan-jenis-kelamin-di-provinsi-dki-jakarta-2015.html>. Januari 30 2017
- [3] Fauzi, M., & Yudi. Penerapan Algoritma K-Means Clustering Untuk Mendeteksi Penyebaran Penyakit TBC. *Teknik Informatika KAPUTAMA*, 1(2), 1-7. 2017.
- [4] KEMKES RI. *Data dan Informasi*. Jakarta: Kementerian Kesehatan Republik Indonesia. 2016
- [5] KEMKES RI. *Data dan Informasi*. Jakarta: Kementerian Kesehatan Republik Indonesia. 2017.
- [6] Larose, D.T. *Discovering Knowledge in Data*. . 2005.
- [7] Munawar. *Pemodelan Visual dengan UML*. Yogyakarta: Graha Ilmu. 2005
- [8] Munggarah, T. P., & T. H. Penerapan Algoritma C.45 Untuk Diagnosa Penyakit Diare Pada Anak Balita Berbasis Mobile. *Swabumi*, vol : II No.1.
- [9] Nasari, F., & Sinaturi, C. J. Penerapan Algoritma K-Means Clustering Untuk Pengelompokan Penyebaran Diare di Kabupaten. *Cogito Smart Jurnal*, Vol 2, No 2, 108, 108. 2016.
- [10] Santoso, T. B. Analisa dan Penerapan Metode C4.5 untuk Prediksi Loyalitas Pelanggan. *Jurnal Ilmiah Fakultas Teknik - LIMIT'S Vol.10 (1) Hal.22-31*. 2014.
- [11] Tan P, Steinbach M, Kumar V. Introduction to Data Mining. In E. Prasetyo, *Data Mining - Konsep dan Aplikasi Menggunakan MATLAB* (p. 2). Gresik: ANDI Yogyakarta. 2006.
- [12] UNICEF. *Diarrhoea: Why children are still dying and what can be done*. United State: The United Nations Children's Fund (UNICEF)/ World Health Organization (WHO). 2009.
- [13] Widoyono. *Penyakit Tropis Epidemiologi, Penularan, Pencegahan dan Pemberantasannya*. Jakarta: Erlanga. 2008.