



## IMPLEMENTASI ALGORITMA K-MEANS UNTUK PENGELOMPOKAN PENYAKIT PADA PUSKESMAS BUKIT KAYU KAPUR

Nur Rubiati<sup>1</sup>, Rahmad Kurniawan<sup>2</sup>, Rahmat Hidayatullah<sup>3</sup>, Ridarmin<sup>4</sup>, Sukri Adrianto<sup>5</sup>, Putri Yunita<sup>6</sup>, Fitri Pratiwi<sup>7</sup>, Adi Arianto<sup>8</sup>, Pulla Pandika Widodo<sup>9</sup>, Mardayulis<sup>10</sup>, Suci Nur Abnita<sup>11</sup>

<sup>1,2,3,4,5,6,7,8,9,10,11</sup>Fakultas Ilmu Komputer, Universitas Dumai

<sup>1,2,3,4,5,6,7,8,9,10,11</sup>Jl. Utama Karya, Bukit Batrem, Kota Dumai, 28811

e-mail : [nurrubiati0412@gmail.com](mailto:nurrubiati0412@gmail.com)

### ABSTRAK

*Seiring dengan bertambahnya jumlah pasien Puskesmas Bukit Kayu Kapur, perlu adanya pendataan dan analisa terhadap data pasien tersebut. Analisa yang perlu diperhatikan yaitu mengenai pengelompokan penyakit akut dan kronis. Pengelompokan tersebut dapat dilakukan dengan menggunakan metode K-Means. Agar proses analisa lebih mudah, maka metode K-Means tersebut dapat diimplementasikan ke dalam sebuah sistem. Pembuatan sistem ini menggunakan Flutter dan berbasisan Web. Perancangan sistem ini terbagi atas use case diagram, activity diagram, flowchart, dan permodelan data. Hasil implementasi dan pengujian menunjukkan bahwa sistem telah berfungsi dengan baik. Sistem yang dibuat dapat digunakan untuk mengelompokkan data dengan mudah dengan visualisasi data yang mudah dipahami.*

**Kata kunci :** K-Means, Penyakit, Puskesmas

### ABSTRACT

*Along with the increasing number of patients at Bukit Kayu Kapur Public Health Center, there is a need for data collection and analysis of these patient data. The analysis that needs to be considered is regarding the grouping of acute and chronic diseases. The grouping can be done using the K-Means method. To make the analysis process easier, the K-Means method can be implemented into a system. Making this system using Flutter and Web-based. The design of this system is divided into use case diagrams, activity diagrams, flowcharts, and data modeling. The results of implementation and testing show that the system has been functioning properly. The system created can be used to group data easily with easy-to-understand data visualizations.*

**Kata kunci :** K-Means, Disease, Health Center

### 1. PENDAHULUAN

Puskesmas merupakan salah satu instansi kesehatan yang berada ditingkat kecamatan, termasuk Puskesmas Bukit Kayu Kapur yang berada di kecamatan Bukit Kayu Kapur Kota Dumai. Dalam rangka menciptakan pelayanan instansi kesehatan yang baik diperlukan tata kerja yang tertib, rapi, dan teliti sehingga menghasilkan informasi yang cepat, akurat, dan tepat waktu sesuai kebutuhan. Seperti instansi kesehatan pada umumnya, Puskesmas Bukit Kayu Kapur juga memberikan pelayanan kesehatan bagi masyarakat setiap harinya.

Seiring dengan bertambahnya jumlah pasien Puskesmas Bukit Kayu Kapur, perlu adanya pendataan dan analisa terhadap data pasien tersebut sehingga hasilnya dapat dijadikan bahan atau dasar penyuluhan kesehatan oleh Puskesmas Bukit Kayu Kapur. Salah satu data yang perlu diperhatikan yaitu mengenai penyakit akut dan kronis yang banyak diderita pasien pada Puskesmas Bukit Kayu Kapur. Dengan adanya data penyakit tersebut dapat membantu pihak pemerintah khususnya Dinas Kesehatan untuk mengetahui penyakit apa yang paling banyak diderita dan juga bisa menjadi bahan acuan dalam

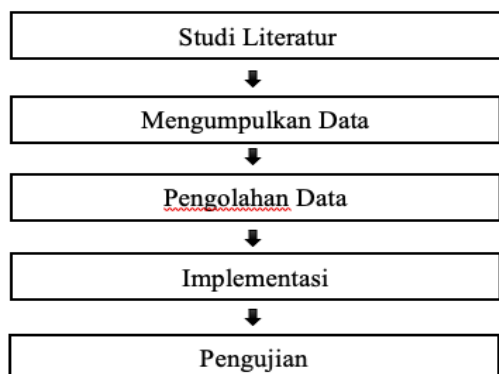


pemberian penyuluhan kesehatan kepada masyarakat sekitar.

Untuk mempermudah proses pengelolaan data pasien tersebut, di perlukan suatu sistem pengelompokan data penyakit berdasarkan penyakit akut dan kronis yang banyak diderita pasien pada Puskesmas Bukit Kayu Kapur. Pengelompokan data pasien tersebut menggunakan *K-Means*. Menurut Buaton *et al.* (2019), *Algoritma K-Means* disusun atas dasar ide yang sederhana. Pada awalnya ditentukan berapa *cluster* yang akan dibentuk. Sebarang obyek atau elemen pertama dalam *cluster* dapat dipilih untuk dijadikan sebagai titik tengah (*centroid point*) *cluster*. *Algoritma K-Means* selanjutnya akan melakukan pengulangan langkah-langkah berikut sampai terjadi kestabilan (tidak ada obyek yang dapat dipindahkan). *Algoritma* ini mampu meminimalkan jarak antara data ke *clusternya*. Pada dasarnya penggunaan *algoritma* ini dalam proses *clustering* tergantung pada data yang didapatkan dan konklusi yang ingin dicapai di akhir proses. Secara umum, klasterisasi pada *algoritma K-Means* dapat dilakukan dengan menentukan jumlah nilai *cluster* (*k*) terlebih dahulu, lalu melakukan inialisasi *k* pusat *cluster* (*centroid*) secara *random* dan menempatkan setiap objek ke *cluster* terdekat berdasarkan perhitungan jarak menggunakan persamaan *Euclidean Distance*. *Euclidean distance* adalah salah satu perhitungan untuk proses *K-Means* dengan cara mengukur jarak antar titik yang berbeda.

## 2. METODOLOGI PENELITIAN

Metode penelitian merupakan urutan rangkaian kegiatan yang akan dilaksanakan dalam sebuah penelitian yang tergambar pada gambar di bawah ini:



Gambar 1. Kerangka Penelitian

1. Studi Literatur  
Studi literatur adalah tahapan mengumpulkan, mempelajari dan memahami *algoritma K-Means clustering* dan bagaimana pengimplementasiannya dari berbagai sumber literatur yang ada.
2. Mengumpulkan Data  
Pada tahap ini dilakukan pengumpulan data penyakit pasien dari Puskesmas Bukit Kayu Kapur.
3. Pengolahan Data  
Pada tahap ini dilakukan pengolahan terhadap data yang telah dikumpulkan.
4. Implementasi  
Pada tahap ini dilakukan pengkodean dan penerapan *algoritma* ke dalam sebuah sistem yang dibuat menggunakan *Flutter*.
5. Pengujian  
Pada tahap ini dilakukan pengujian apakah sistem yang dibuat terdapat kesalahan atau bug program yang harus diperbaiki.

## 3. HASIL DAN PEMBAHASAN

### a. Analisis Sistem

Seiring dengan bertambahnya jumlah pasien Puskesmas Bukit Kayu Kapur perlu diperhatikan data pasien penyakit akut dan kronis yang banyak diderita pasien pada Puskesmas Bukit Kayu Kapur. Untuk mempermudah proses pengelolaan data pasien tersebut, di perlukan suatu sistem yang dapat mengelompokkan data penyakit berdasarkan penyakit akut dan kronis yang banyak diderita pasien pada Puskesmas Bukit Kayu Kapur. Pengelompokan data pasien tersebut menggunakan *algoritma K-Means*. *K-Means* merupakan suatu *algoritma* yang digunakan dalam pengelompokan secara partisi yang memisahkan data ke dalam kelompok yang berbeda-beda. *Algoritma* ini mampu meminimalkan jarak antara data ke *clusternya*. *Algoritma K-Means* berusaha mengelompokkan data yang ada ke dalam beberapa kelompok, yang mana data dalam satu kelompok mempunyai karakteristik yang sama satu sama lainnya dan mempunyai karakteristik yang berbeda dengan data yang ada di dalam kelompok yang lain.

Selain pengelompokan data, sistem juga menampilkan data secara visual dan beberapa visualisasi data lainnya, seperti visualisasi berdasarkan jenis kelamin, umur dan penyakit. Data dapat ditambah, diubah atau dihapus. Sehingga dengan sistem ini dapat memberikan hasil analisa secara cepat dan mudah terhadap data pasien Puskesmas Bukit Kayu Kapur.



## b. Tahapan-Tahapan Metode

Data penelitian yang sedang dilakukan merupakan data penyakit pasien Puskesmas Bukit Kayu Kapur sebanyak 141 data yang merupakan data pasien yang mengidap penyakit yang masih dapat disembuhkan. Data tersebut kemudian dikelompokkan menjadi 2 *cluster* yaitu kelompok penyakit akut (C1) dan kronis (C2). Penyakit akut adalah penyakit dengan lama mengidap yang pendek sedangkan kronis adalah penyakit dengan lama mengidap yang panjang. Pengelompokan tersebut berdasarkan atribut umur dan lama mengidap (hari), yang kemudian atribut tersebut akan diolah menggunakan *algoritma K-Means*. Sampel dari data penyakit pasien Puskesmas Bukit Kayu Kapur dapat dilihat pada tabel 1 di bawah ini.

Tabel 1. Sampel Data

| NO | UMUR | LAMA MENGIDAP |
|----|------|---------------|
| 1  | 30   | 6             |
| 2  | 32   | 12            |
| 3  | 33   | 18            |
| 4  | 27   | 19            |
| 5  | 21   | 6             |
| 6  | 31   | 11            |
| 7  | 30   | 12            |
| 8  | 18   | 182           |
| 9  | 9    | 17            |
| 10 | 17   | 10            |
| 11 | 41   | 20            |
| 12 | 8    | 6             |
| 13 | 9    | 7             |
| 14 | 7    | 14            |
| 15 | 23   | 84            |
| 16 | 52   | 105           |
| 17 | 10   | 4             |
| 18 | 9    | 18            |
| 19 | 11   | 11            |
| 20 | 8    | 16            |

1. Tentukan nilai  $k = 2$ .
2. Tetapkan pusat *cluster*. Dalam penelitian ini pusat *cluster* awal diambil dari median atribut umur yang sudah diurutkan dan lama mengidap tertinggi dan terendah.  
Rumus Median:

$$Me = x \frac{(n+1)}{2}$$

Umur (x) = 7, 8, 8, 9, 9, 9, 10, 11, 17, 18, 21, 23, 27, 30, 30, 31, 32, 33, 41, 52

$$Me = x \frac{(20+1)}{2} = 12$$

Median umur:  $X_{12} = 23$

Lama mengidap terendah = 4

Lama mengidap tertinggi = 182

Maka nilai pusat *cluster* awal adalah:

C1 = (23, 4)

C2 = (23, 182)

3. Hitung jarak setiap data terhadap setiap pusat *cluster*. Untuk melakukan penghitungan jarak data ke- $i$  ( $X_i$ ) pada pusat *cluster* ke- $k$  ( $C_k$ ), diberi nama ( $D_{ij}$ ), dapat digunakan formula *Euclidean*, yaitu:

$$d_{ij} = \sqrt{\sum_{j=1}^m (x_{ij} - c_{kj})^2}$$

$d_{ij}$  : Jarak (*distance*) data ke- $i$  dengan pusat cluster ke- $j$

$x_{ij}$  : Nilai data ke- $i$  pada atribut  $j$

$c_{kj}$  : Nilai pusat *cluster* ke- $k$  pada atribut  $j$

$\sum_{j=1}^m$  : Jumlah, untuk nilai  $j$  dari 1 sampai  $m$

Misalkan untuk menghitung jarak data pertama (No.1) dengan pusat *cluster* pertama adalah:

$$d_{11} = \sqrt{(30 - 23)^2 + (6 - 4)^2} = 7.28$$

Jarak data pertama (No.1) dengan pusat *cluster* kedua adalah:

$$d_{12} = \sqrt{(30 - 23)^2 + (6 - 182)^2} = 176.139$$

Hasil perhitungan jarak selengkapnya adalah:

Tabel 2. Tabel Perhitungan Jarak Iterasi pertama

| NO | UMUR | LAMA MENGIDAP | JARAK C1 | JARAK C2 |
|----|------|---------------|----------|----------|
| 1  | 30   | 6             | 7.28     | 176.139  |
| 2  | 32   | 12            | 12.042   | 170.238  |
| 3  | 33   | 18            | 17.205   | 164.305  |
| 4  | 27   | 19            | 15.524   | 163.049  |
| 5  | 21   | 6             | 2.828    | 176.011  |
| 6  | 31   | 11            | 10.63    | 171.187  |
| 7  | 30   | 12            | 10.63    | 170.144  |
| 8  | 18   | 182           | 178.07   | 5        |
| 9  | 9    | 17            | 19.105   | 165.593  |
| 10 | 17   | 10            | 8.485    | 172.105  |
| 11 | 41   | 20            | 24.083   | 162.997  |
| 12 | 8    | 6             | 15.133   | 176.638  |
| 13 | 9    | 7             | 14.318   | 175.559  |
| 14 | 7    | 14            | 18.868   | 168.76   |
| 15 | 23   | 84            | 80       | 98       |
| 16 | 52   | 105           | 105.081  | 82.28    |
| 17 | 10   | 4             | 13       | 178.474  |



|    |    |    |        |         |
|----|----|----|--------|---------|
| 18 | 9  | 18 | 19.799 | 164.596 |
| 19 | 11 | 11 | 13.892 | 171.421 |
| 20 | 8  | 16 | 19.209 | 166.676 |

Suatu data akan menjadi anggota dari suatu *cluster* yang memiliki jarak terkecil dari pusat *cluster*nya. Misalkan untuk data pertama, jarak terkecil diperoleh pada *cluster* pertama, sehingga data pertama akan menjadi anggota dari *cluster* pertama. Demikian juga untuk data kedua (No.2), jarak terkecil ada pada *cluster* pertama, maka data tersebut akan masuk pada *cluster* pertama. Berikut ini tabel posisi *cluster*, tanda bintang (\*) menunjukkan ke *cluster* mana data tersebut dimasukkan.

Tabel 3. Tabel *Clustering* Iterasi Pertama

| NO | UMUR | LAMA MENGIDAP | C1 | C2 |
|----|------|---------------|----|----|
| 1  | 30   | 6             | *  |    |
| 2  | 32   | 12            | *  |    |
| 3  | 33   | 18            | *  |    |
| 4  | 27   | 19            | *  |    |
| 5  | 21   | 6             | *  |    |
| 6  | 31   | 11            | *  |    |
| 7  | 30   | 12            | *  |    |
| 8  | 18   | 182           |    | *  |
| 9  | 9    | 17            | *  |    |
| 10 | 17   | 10            | *  |    |
| 11 | 41   | 20            | *  |    |
| 12 | 8    | 6             | *  |    |
| 13 | 9    | 7             | *  |    |
| 14 | 7    | 14            | *  |    |
| 15 | 23   | 84            | *  |    |
| 16 | 52   | 105           |    | *  |
| 17 | 10   | 4             | *  |    |
| 18 | 9    | 18            | *  |    |
| 19 | 11   | 11            | *  |    |
| 20 | 8    | 16            | *  |    |

Berdasarkan hasil penggolongan tersebut, diperoleh anggota *cluster* pertama ada 18 data dan *cluster* kedua ada 2 data.

#### 4. Hitung pusat *cluster* baru

Nilai pusat *cluster* yang baru dapat dihitung dengan cara mencari nilai rata-rata dari data yang menjadi anggota pada *cluster* tersebut, dengan rumus:

$$c_{kj} = \frac{\sum_{h=1}^p y_{hj}}{p}$$

$c_{kj}$  : Nilai pusat *cluster* ke- $k$  pada atribut  $j$

$y_{hj}$  : Nilai data ke- $h$  pada atribut  $j$

$p$  : Banyaknya data

$\sum_{j=1}^m$  : Jumlah semua nilai  $y$ , untuk nilai  $h$  dari 1 sampai  $p$

Untuk *cluster* pertama, ada 18 data yaitu data ke-1 sampai data ke-7, data ke-9 sampai data ke-15 dan data ke-17 sampai data ke 20. sehingga:

$$C_{11} = \frac{30 + 32 + 33 + 27 + 21 + 31 + 30 + 9 + 17 + 41 + 8 + 9 + 7 + 23 + 10 + 9 + 11 + 8}{18} = 19.778$$

$$C_{12} = \frac{6 + 12 + 18 + 19 + 6 + 11 + 12 + 17 + 10 + 20 + 6 + 7 + 14 + 84 + 4 + 18 + 11 + 16}{18} = 16.167$$

Untuk *cluster* kedua, ada 2 data yaitu data ke-8 dan ke-16, sehingga:

$$C_{21} = \frac{18 + 52}{2} = 35$$

$$C_{22} = \frac{182 + 105}{2} = 143.5$$

- Ulangi menghitung jarak setiap data terhadap setiap pusat *cluster* yang baru. Hasil perhitungan jarak selengkapnya terlihat pada tabel berikut.

Tabel 4. Tabel Perhitungan Jarak Iterasi Kedua

| NO | UMUR | LAMA MENGIDAP | JARAK C1 | JARAK C2 |
|----|------|---------------|----------|----------|
| 1  | 30   | 6             | 14.417   | 137.591  |
| 2  | 32   | 12            | 12.913   | 131.534  |
| 3  | 33   | 18            | 13.348   | 125.516  |
| 4  | 27   | 19            | 7.758    | 124.757  |
| 5  | 21   | 6             | 10.24    | 138.211  |
| 6  | 31   | 11            | 12.354   | 132.56   |
| 7  | 30   | 12            | 11.039   | 131.595  |
| 8  | 18   | 182           | 165.843  | 42.086   |
| 9  | 9    | 17            | 10.81    | 129.144  |
| 10 | 17   | 10            | 6.764    | 134.708  |
| 11 | 41   | 20            | 21.565   | 123.646  |
| 12 | 8    | 6             | 15.559   | 140.126  |
| 13 | 9    | 7             | 14.149   | 138.954  |
| 14 | 7    | 14            | 12.96    | 132.492  |
| 15 | 23   | 84            | 67.909   | 60.698   |
| 16 | 52   | 105           | 94.496   | 42.086   |
| 17 | 10   | 4             | 15.609   | 141.722  |
| 18 | 9    | 18            | 10.933   | 128.165  |
| 19 | 11   | 11            | 10.186   | 134.656  |
| 20 | 8    | 16            | 11.779   | 130.327  |



6. Posisi *cluster* selengkapnya terlihat pada tabel berikut.

Tabel 5. Tabel *Clustering* Iterasi Kedua

| NO | UMUR | LAMA MENGIDAP | C1 | C2 |
|----|------|---------------|----|----|
| 1  | 30   | 6             | *  |    |
| 2  | 32   | 12            | *  |    |
| 3  | 33   | 18            | *  |    |
| 4  | 27   | 19            | *  |    |
| 5  | 21   | 6             | *  |    |
| 6  | 31   | 11            | *  |    |
| 7  | 30   | 12            | *  |    |
| 8  | 18   | 182           |    | *  |
| 9  | 9    | 17            | *  |    |
| 10 | 17   | 10            | *  |    |
| 11 | 41   | 20            | *  |    |
| 12 | 8    | 6             | *  |    |
| 13 | 9    | 7             | *  |    |
| 14 | 7    | 14            | *  |    |
| 15 | 23   | 84            |    | *  |
| 16 | 52   | 105           |    | *  |
| 17 | 10   | 4             | *  |    |
| 18 | 9    | 18            | *  |    |
| 19 | 11   | 11            | *  |    |
| 20 | 8    | 16            | *  |    |

Terlihat masih ada 1 data yang berubah posisi dari kondisi semula, yaitu data ke-15. Sehingga perlu dihitung pusat *cluster* baru.

7. Hitung pusat *cluster* baru sebagaimana pada langkah ke-4, sehingga diperoleh:

$$C_{11} = 19.558; C_{12} = 12.176$$

$$C_{21} = 31; C_{22} = 123.667$$

8. Ulangi menghitung jarak setiap data terhadap setiap pusat *cluster* yang baru.  
Hasil perhitungan jarak selengkapnya terlihat pada tabel berikut.

Tabel 6. Tabel Perhitungan Jarak Iterasi Ketiga

| NO | UMUR | LAMA MENGIDAP | JARAK C1 | JARAK C2 |
|----|------|---------------|----------|----------|
| 1  | 30   | 6             | 12.106   | 117.671  |
| 2  | 32   | 12            | 12.413   | 111.671  |
| 3  | 33   | 18            | 14.622   | 105.686  |
| 4  | 27   | 19            | 10.075   | 104.743  |
| 5  | 21   | 6             | 6.335    | 118.091  |
| 6  | 31   | 11            | 11.472   | 112.667  |
| 7  | 30   | 12            | 10.413   | 111.671  |
| 8  | 18   | 182           | 169.831  | 59.764   |
| 9  | 9    | 17            | 11.635   | 108.912  |
| 10 | 17   | 10            | 3.381    | 114.526  |
| 11 | 41   | 20            | 22.797   | 104.148  |
| 12 | 8    | 6             | 13.131   | 119.894  |
| 13 | 9    | 7             | 11.785   | 118.723  |
| 14 | 7    | 14            | 12.719   | 112.262  |
| 15 | 23   | 84            | 71.905   | 40.466   |
| 16 | 52   | 105           | 98.32    | 28.097   |
| 17 | 10   | 4             | 12.601   | 121.496  |

|    |    |    |        |         |
|----|----|----|--------|---------|
| 18 | 9  | 18 | 12.084 | 107.933 |
| 19 | 11 | 11 | 8.668  | 114.428 |
| 20 | 8  | 16 | 12.203 | 110.096 |

9. Posisi *cluster* selengkapnya terlihat pada tabel berikut.

Tabel 7. Tabel *Clustering* Iterasi Kedua

| NO | UMUR | LAMA MENGIDAP | C1 | C2 |
|----|------|---------------|----|----|
| 1  | 30   | 6             | *  |    |
| 2  | 32   | 12            | *  |    |
| 3  | 33   | 18            | *  |    |
| 4  | 27   | 19            | *  |    |
| 5  | 21   | 6             | *  |    |
| 6  | 31   | 11            | *  |    |
| 7  | 30   | 12            | *  |    |
| 8  | 18   | 182           |    | *  |
| 9  | 9    | 17            | *  |    |
| 10 | 17   | 10            | *  |    |
| 11 | 41   | 20            | *  |    |
| 12 | 8    | 6             | *  |    |
| 13 | 9    | 7             | *  |    |
| 14 | 7    | 14            | *  |    |
| 15 | 23   | 84            |    | *  |
| 16 | 52   | 105           |    | *  |
| 17 | 10   | 4             | *  |    |
| 18 | 9    | 18            | *  |    |
| 19 | 11   | 11            | *  |    |
| 20 | 8    | 16            | *  |    |

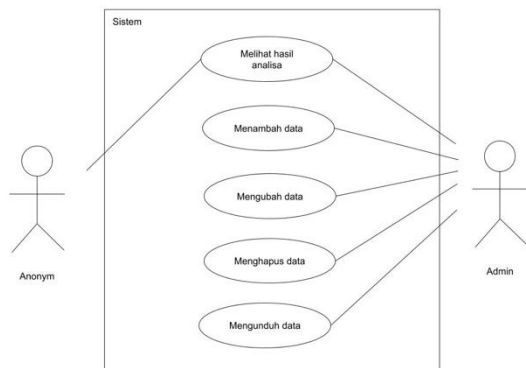
Terlihat bahwa posisi setiap data pada setiap *cluster* sudah tidak terjadi perpindahan, sehingga proses iterasi sudah dapat dihentikan. Hasil akhir yang diperoleh adalah 2 *cluster* dengan:

1. *Cluster* pertama memiliki pusat (19.558, 12.176) yang dapat diartikan sebagai kelompok pasien yang mengidap penyakit akut. Ada 17 pasien dalam kelompok ini dengan rentang umur 7-41 tahun dan lama mengidap 4-20 hari.
2. *Cluster* kedua memiliki pusat (31, 123.667) yang dapat diartikan sebagai kelompok pasien yang mengidap penyakit kronis. Ada 3 pasien dalam kelompok ini dengan rentang umur 18-52 tahun dan lama mengidap 84-182 hari.

## c. Perancangan Sistem

### a. Use Case Diagram

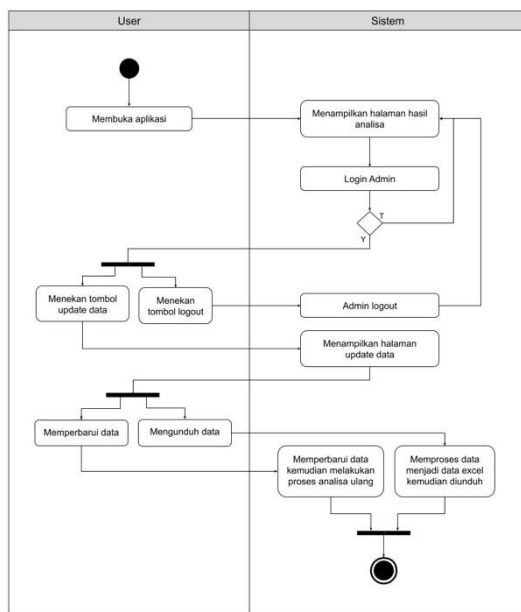
*Use case diagram* merupakan gambaran skenario dari interaksi antara pengguna dengan sistem. *Use case* merupakan fungsionalitas yang disediakan sistem sebagai unit-unit yang saling bertukar pesan antar unit atau aktor (Rosa & Shalahudin, 2016:156).



Gambar 2. Use Case Diagram

## b. Activity Diagram

*Activity Diagram* menggambarkan *workflow* (aliran kerja) atau aktifitas dari sebuah sistem atau proses bisnis atau menu yang ada pada perangkat lunak. perlu diperhatikan bahwa diagram aktivitas menggambarkan aktivitas sistem bukan apa yang dilakukan aktor, jadi aktivitas yang dapat dilakukan oleh sistem (Rosa & Shalahudin, 2016:162).

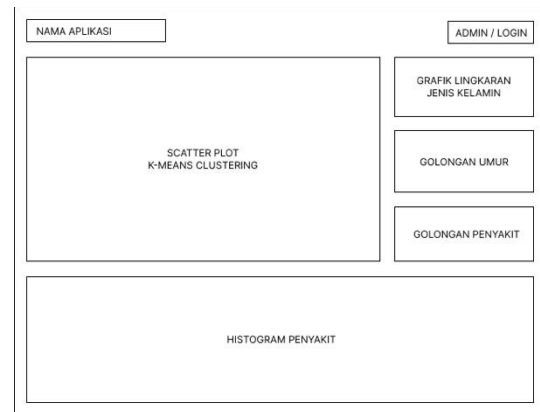


Gambar 3. Activity Diagram

## c. Pemodelan Output

Pemodelan *output* pada penelitian ini adalah menampilkan visualisasi hasil *clustering* dengan grafik *scatter plot*. Grafik *scatter plot* atau yang memiliki nama lain grafik pencar merupakan diagram grafis yang dibangun dari dua sumbu X

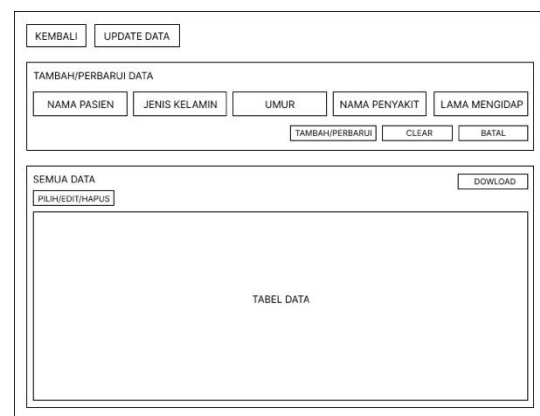
dan Y (variabel X dan variabel Y). Nilai dari sepasang variabel ini digambarkan sebagai titik-titik atau *point*. Selain itu, juga ditampilkan grafik lingkaran untuk pengelompokan jenis kelamin, histogram penyakit, data berdasarkan penggolongan umur dan penyakit.



Gambar 4. Pemodelan Output

## d. Pemodelan Input

Pemodelan *input* pada penelitian ini adalah berupa *form* dengan *field* input nama, jenis kelamin, umur, nama penyakit dan lama mengidap. Selain dapat diinput/ditambah, data juga dapat diubah atau dihapus melalui field yang sama.



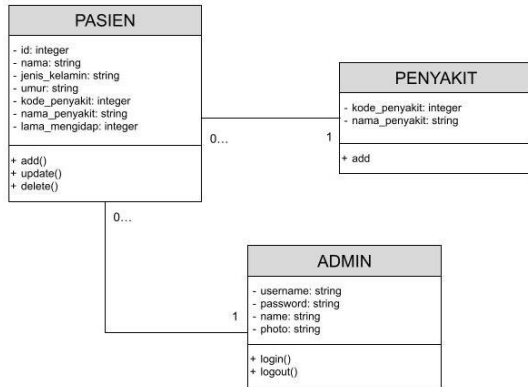
Gambar 5. Pemodelan Input

## e. Pemodelan Data

Pemodelan data pada penelitian ini menggunakan pemodelan UML yaitu pemodelan yang menggunakan *class diagram*. *Class diagram* ini menampilkan beberapa kelas beserta



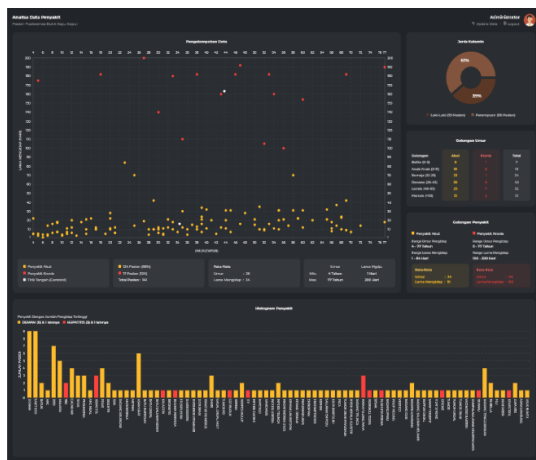
atribut dan operasi yang terjadi di dalam kelas dan interaksi terhadap kelas yang lain



Gambar 6. Pemodelan Data

## d. Hasil Implementasi

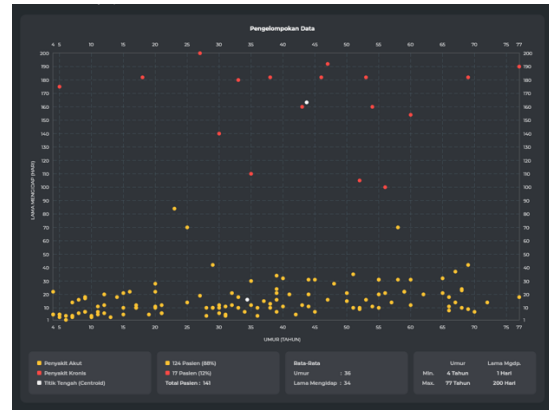
Berikut ini adalah hasil implementasi algoritma K-Means ke dalam sebuah sistem yang dibuat menggunakan Flutter dan berbasis web.



Gambar 7. Hasil Implementasi

Dapat dilihat pada Gambar 4.1 di atas, hasil pengolahan data divisualisasikan ke dalam beberapa bentuk grafik dan bagian. Terdapat satu grafik *scatter plot* yang menampilkan titik-titik data beserta *cluster* dan *centroid* setiap *cluster*. Titik-titik yang berwarna kuning merupakan data yang dikelompokkan ke dalam kelompok penyakit akut (*Cluster 1*). Titik-titik yang berwarna merah dikelompokkan ke dalam kelompok penyakit kronis (*Cluster 2*). Sedangkan titik-titik berwarna putih merupakan *centroid*

setiap *cluster*. Selain itu, juga ditampilkan informasi jumlah keseluruhan data, jumlah data pada setiap *cluster*, rata-rata dan nilai terendah dan tertinggi setiap atribut.



Gambar 8. Grafik Scatter Plot

## e. Pengujian Sistem

Pengujian sistem yang digunakan dalam penelitian ini adalah pengujian yang menggunakan metode *Black Box*. *Black Box* Testing yaitu menguji perangkat lunak dari segi spesifikasi fungsional tanpa menguji desain dan kode program (Rosa & Shalahudin, 2016:275).

Tabel 8. Hasil Pengujian Pada Halaman Utama

| Aksi/Bentuk Pengujian                                | Hasil yang Diharapkan                                                                       | Hasil Pengamatan                    | Kesimpulan |
|------------------------------------------------------|---------------------------------------------------------------------------------------------|-------------------------------------|------------|
| Login Admin                                          | Sistem menampilkan informasi pengguna dan tombol untuk menuju halaman <i>update data</i>    | Sesuai dengan hasil yang diharapkan | Berhasil   |
| Menampilkan visualisasi data hasil <i>clustering</i> | Sistem menampilkan data hasil <i>clustering</i> yang sesuai dengan hasil perhitungan manual | Sesuai dengan hasil yang diharapkan | Berhasil   |
| Menambahkan data baru                                | Data ditambahkan ke dalam sistem dan sesuai dengan <i>inputan</i>                           | Sesuai dengan hasil yang diharapkan | Berhasil   |



|                  |                                                                                            |                                     |          |
|------------------|--------------------------------------------------------------------------------------------|-------------------------------------|----------|
| Memperbarui data | Data berubah sesuai dengan <i>inputan</i>                                                  | Sesuai dengan hasil yang diharapkan | Berhasil |
| Menghapus data   | Data dihapus dari sistem                                                                   | Sesuai dengan hasil yang diharapkan | Berhasil |
| Mengunduh data   | Data diunduh dalam bentuk Excel                                                            | Sesuai dengan hasil yang diharapkan | Berhasil |
| Logout Admin     | Sistem menghilangkan informasi pengguna dan tombol untuk menuju halaman <i>update data</i> | Sesuai dengan hasil yang diharapkan | Berhasil |

## f. Analisis Hasil Pengujian Black Box

Pengujian *Black Box* yang dilakukan pada sistem dibagi ke dalam 7 bentuk pengujian. Pada setiap bentuk pengujian tersebut, hasil pengamatan menunjukkan hasil telah sesuai dengan hasil yang diharapkan.

Dari hasil pengujian tersebut dapat disimpulkan bahwa implementasi *algoritma K-Means clustering* pada sistem telah berhasil dan berfungsi dengan baik dan sesuai sebagaimana yang diharapkan.

## 4. KESIMPULAN

Berdasarkan dari identifikasi masalah, rumusan masalah dan hasil penelitian, maka dapat ditarik kesimpulan sebagai berikut:

1. Dari 141 data penyakit pada Puskesmas Bukit Kayu Kapur yang telah dikelompokkan menjadi dua kelompok penyakit, yaitu kelompok penyakit akut dengan jumlah data 124 data, dan kelompok penyakit kronis dengan jumlah data 17 data.
2. Dengan adanya sistem yang mengimplementasikan *algoritma K-Means clustering*, pengelompokan data penyakit akut dan kronis dapat dilakukan dengan mudah dengan hasil yang mendetail dan divisualisasikan dengan grafik yang menarik dan mudah dipahami.

## 5. REFERENSI

- A.S., Rosa & Shalahuddin, M. (2016). *Rekayasa Perangkat Lunak Terstruktur dan Berorientasi Objek*. Bandung: Informatika.
- Buaton, R., Mawengkang, H., Zarlis, M., Effendi, S., Manaor Hara Pardede, A., Maulita, Y., Fauzi, A., & Novriyenni, N. (2019). Time Series Optimization on Data Mining. *Journal of Physics: Conference Series*, 1235(1). <https://doi.org/10.1088/1742-6596/1235/1/012014>
- Ma'rifat, U. (2020). Implementasi Algoritma K-Means untuk Pengelompokan Penyakit Pasien Pada Puskesmas Warujayeng. *Seminar Nasional Inovasi Teknologi*, 2549–7952, 285–291.
- Sugianto, C. A., Rahayu, A. H., & Gusman, A. (2020). Algoritma K-Means untuk Pengelompokan Penyakit Pasien pada Puskesmas Cigugur Tengah. *Journal of Information Technology*, 2(2), 39–44. <https://doi.org/10.47292/joint.v2i2.30>
- Suntoro, J. (2019). *Data Mining: Algoritma dan Implementasi dengan Pemrograman PHP*. Elex Media Komputindo.
- Suryana, T. (2021). Belajar Bahasa Pemrograman Dart (1). In *Teknik Informatika Unikom* (Issue 1, p. 4). Universitas Komputer Indonesia. <https://repository.unikom.ac.id/68459/>
- Sutoyo, M. N. (2009). *Algoritma K-Means*. 1, 1–7.
- Utomo, W. (2020). Penerapan Algoritma K-Means Untuk Pengelompokan Penyakit Kronis pada Warga Lansia (Studi Kasus Pada: Posyandu Lansia RW 07). *Jurnal Media Informatika*, 4(4), 1153–1161. <https://doi.org/10.30865/mib.v4i4.2410>