

## Penerapan Data Mining Menggunakan Algoritma K-Means Clustering Dalam Pengelompokan Data Penerima Vaksinasi Covid-19

Alfredo Torence<sup>1</sup>, Mukhlis Ramadhan<sup>2</sup>, Erika Fahmi Ginting<sup>3</sup>

<sup>1,2,3</sup>Sistem Informasi, STMIK Triguna Dharma

Email: <sup>1</sup>alfredotorence@gmail.com, <sup>2</sup>mukhlisramadhan.tgd@gmail.com, <sup>3</sup>erikafg04@gmail.com

Email Penulis Korespondensi: alfredotorence@gmail.com

### Abstrak

Masalah yang ditemukan di Puskesmas Bandar Baru cabang Sibolangit adalah pemerintah menetapkan untuk penerima vaksin yaitu berdasarkan prioritas tertentu, seperti PNS, tenaga medis, manusia lanjut usia dan masyarakat biasa. Ini membuat pusat layanan kesehatan masyarakat seperti puskesmas harus dapat mengelompokkan data-data di masyarakat berdasarkan prioritas yang telah ditetapkan oleh pemerintah. Oleh karena itu maka dibangunlah sebuah data mining yang dapat melakukan proses pengelompokan data vaksinasi covid-19 berdasarkan kriteria yang telah ditentukan dan dikombinasikan dengan metode *K-means clustering*. Metode *k-means clustering* adalah sebuah metode *cluster analysis* non hirarki yang berusaha mempartisi objek berdasarkan karakteristiknya, sehingga objek yang mempunyai karakteristik yang sama dikelompokkan dalam satu cluster yang sama dan objek yang mempunyai karakteristik yang berbeda dikelompokkan kedalam cluster yang lain. Hasil yang diperoleh adalah sebuah sistem yang dapat memberikan hasil akhir berupa data-data kelompok penerima vaksinasi perbulannya dimulai dari nilai tertinggi ke terendah. Data ini diharapkan dapat membantu pihak puskesmas Bandar Baru cabang Sibolangit untuk mengelompokkan data penerima vaksinasi covid-19 kriteria vaksinasi I, vaksinasi II dan vaksin *booster*.

**Kata Kunci:** *Data Mining, K Means, Clustering, Vaksinasi, Covid-19.*

### Abstract

The problem found at the Sibolangit branch of the Bandar Baru Health Center is that the government determines vaccine recipients based on certain priorities, such as civil servants, medical personnel, elderly people and ordinary people. This makes community health service centers such as puskesmas must be able to classify data in the community based on priorities set by the government. Therefore, a data mining was developed that can carry out the process of grouping Covid-19 vaccination data based on predetermined criteria and combined with the *K-means clustering* method. The *k-means clustering* method is a non-hierarchical cluster analysis method that seeks to partition objects based on their characteristics, so that objects with the same characteristics are grouped in the same cluster and objects with different characteristics are grouped into other clusters. The result obtained is a system that can provide the final result in the form of data on monthly vaccination recipient groups starting from the highest value to the lowest. This data is expected to help the Sibolangit branch of the Bandar Baru Health Center to classify data on recipients of the Covid-19 vaccination with criteria for vaccination I, vaccination II and booster vaccines.

**Keywords:** *Data Mining, K Means, Clustering, Vaksinasi, Covid-19.*

## 1. PENDAHULUAN

Masalah yang ditemukan di Puskesmas Bandar Baru cabang Sibolangit adalah pemerintah menetapkan untuk penerima vaksin yaitu berdasarkan prioritas tertentu, seperti PNS, tenaga medis, manusia lanjut usia dan masyarakat biasa. Ini membuat pusat layanan kesehatan masyarakat seperti puskesmas harus dapat mengelompokkan data-data di masyarakat berdasarkan prioritas yang telah ditetapkan oleh pemerintah. Oleh karena itu maka dibangunlah sebuah data mining yang dapat melakukan proses pengelompokan data vaksinasi covid-19 berdasarkan kriteria yang telah ditentukan dan dikombinasikan dengan metode *K-means clustering*. Metode *k-means clustering* adalah sebuah metode cluster analysis non hirarki yang berusaha mempartisi objek berdasarkan karakteristiknya, sehingga objek yang mempunyai karakteristik yang sama dikelompokkan dalam satu cluster yang sama dan objek yang mempunyai karakteristik yang berbeda dikelompokkan kedalam *cluster* yang lain. Hasil yang diperoleh adalah sebuah sistem yang dapat memberikan hasil akhir berupa data-data kelompok penerima vaksinasi perbulannya dimulai dari nilai tertinggi ke terendah. Data ini diharapkan dapat membantu pihak puskesmas Bandar Baru cabang Sibolangit untuk mengelompokkan data penerima vaksinasi covid-19 kriteria vaksinasi I, vaksinasi II dan vaksin *booster* [1].

Begitu juga halnya dengan Puskesmas Bandar baru yang menjadi pusat kesehatan yang menyediakan vaksinasi Covid-19 bagi masyarakat khususnya di daerah bandar baru, permasalahannya adalah pemerintah menetapkan untuk penerima vaksinasi yaitu berdasarkan prioritas tertentu, seperti PNS, tenaga medis, Manula, dan masyarakat biasa. Ini membuat puskesmas harus dapat mengelompokkan data – data di masyarakat berdasarkan prioritas yang telah ditetapkan oleh pemerintah. Oleh karena itu dibutuhkan lah suatu sistem yang dapat mengelompokkan data berdasarkan syarat – syarat tertentu yang disebut *Data Mining*. *Data mining* adalah proses pengolahan data ataupun big data yang diolah berdasarkan kebutuhan tertentu dengan menghasilkan informasi yang sangat dibutuhkan.

*Data Mining* adalah penambangan atau penemuan informasi baru dengan mencari pola atau aturan tertentu dari sejumlah data yang sangat besar [2]. *Data Mining* juga disebut sebagai serangkaian proses untuk menggali nilai tambah berupa pengetahuan yang selama ini tidak diketahui secara manual dari suatu kumpulan data. Pada penelitian ini analisa data mining dilakukan dengan menggunakan metode *K-Means*. Dengan menggunakan metode ini, data-data yang telah didapatkan dapat dikelompokkan kedalam beberapa *cluster* berdasarkan kemiripan dari data-data tersebut dan dikelompokkan dalam cluster yang lain yang memiliki karakteristik yang sama [3].

*K-means clustering* merupakan salah satu metode *cluster analysis* non hirarki yang berusaha mempartisi objek yang ada kedalam satu atau lebih cluster atau kelompok objek berdasarkan karakteristiknya, sehingga objek yang mempunyai karakteristik yang sama dikelompokkan dalam satu cluster yang sama dan objek yang mempunyai karakteristik yang berbeda dikelompokkan kedalam *cluster* yang lain. Metode *K-means clustering* berusaha mengelompokkan data yang ada ke dalam beberapa kelompok, dimana 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 [4].

*Data mining* berkembang menjadi alat bantu untuk mencari pola-pola yang berharga dalam suatu basisdata yang sangat besar jumlahnya, sehingga tidak memungkinkan dicari secara manual. Beberapa teknik data mining dapat diklasifikasikan ke dalam kategori berikut, meliputi klasifikasi, *Clustering*, penggalian kaidah asosiasi, analisa pola sekuensial, prediksi, visualisasi data dan lain sebagainya. Teknik *Clustering* adalah teknik yang digunakan untuk menangani data yang besar dengan banyak atribut ke dalam sejumlah kelompok kecil. *Clustering* dilakukan dengan terlebih dahulu menganalisis bagian kecil dari data untuk menentukan klaster. Dalam data mining banyak metode yang sesuai untuk digunakan dalam menganalisa suatu pola penjualan salah satunya adalah algoritma *K-Means* [5]. Algoritma *K-Means* merupakan algoritma pengelompokan iteratif yang melakukan partisi set data ke dalam sejumlah *K cluster* yang sudah ditetapkan di awal. Algoritma *K-Means* sederhana untuk diimplementasikan dan dijalankan, relatif cepat, mudah beradaptasi, umum Penggunaannya dalam praktek. *K-Means* dapat diterapkan pada data yang direpresentasikan dalam *r*-dimensi ruang tempat [6].

## 2. METODOLOGI PENELITIAN

### 2.1 Tahapan Penelitian

Tahapan penelitian dalam kasus pengelompokkan data penerima vaksinasi Covid-19 terdapat beberapa bagian penting, yaitu sebagai berikut ini:

a. Teknik Pengumpulan Data (*Data Collecting*)

*Data Collecting* adalah suatu teknik pengumpulan data yang digunakan untuk mencari informasi yang dibutuhkan dalam penelitian.

1. Pengamatan Langsung (*Observasi*)

2. Wawancara (*Interview*)

b. Studi Kepustakaan (*Study of Literature*)

c. Penerapan Metode *K-Means* dalam pengolahan data.

### 2.2 Data Mining

Data Mining adalah suatu istilah yang digunakan untuk menguraikan penemuan pengetahuan di dalam database. Data mining adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan machine learning untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terikat dari berbagai *database* besar. Berdasarkan pengertian data mining yang telah dijelaskan di atas, maka data mining merupakan pengetahuan yang tersembunyi di dalam *database* yang di proses untuk menemukan pola dan teknik statistik matematika, kecerdasan buatan, dan machine learning untuk mengekstraksi dan mengidentifikasi informasi pengetahuan dari *database* tersebut [7].

*Knowledge Discovery In Database* (KDD) adalah metode yang digunakan untuk mencari pengetahuan atau informasi yang belum diketahui dari sebuah database. KDD merupakan nama lain dari Data Mining walaupun sesungguhnya kedua istilah tersebut memiliki konsep yang tidak sama, namun berkaitan satu sama lain, dan salah satu dari tahapan proses keseluruhan KDD merupakan *Data Mining* yang menjadi inti dari proses KDD [8].

Data Mining merupakan salah satu teknik untuk menemukan, mencari, atau menggali informasi atau pengetahuan baru dari sekumpulan data yang sangat besar, dengan integrasi atau penggabungan dengan disiplin ilmu lain seperti statistika, kecerdasan buatan, serta *machine learning*, menjadikan *Data Mining* sebagai salah satu alat bantu untuk menganalisa data yang kemudian menghasilkan informasi yang berguna [9].

### 2.3 K-Means Clustering

*K-Means* merupakan salah satu metode pengelompokan data nonhierarki (sekatan) yang berusaha mempartisi data yang ada ke dalam bentuk dua atau lebih kelompok. Metode ini mempartisi data ke dalam kelompok sehingga data berkarakteristik sama dimasukkan ke dalam satu kelompok yang sama dan data yang berkarakteristik berbeda dikelompokkan kedalam kelompok yang lain [10]:

*K-Means* ialah algoritma yang digunakan kedalam suatu pengelompokkan secara membagi yang memisahkan data ke dalam kelompok yang berbeda – beda. Algoritma ini mampu memperpendek jarak antara data ke *cluster* nya. Pada dasar penggunaan algoritma ini tergantung dalam proses *clustering* pada data yang dihasilkan dan konklusi yang ingin dicapai di akhir proses. Sehingga dalam penggunaan algoritma *k-means* memiliki beberapa aturan sebagai berikut:

1. Berapa banyak jumlah *cluster* yang perlukan

2. Hanya mempunyai atribut bertipe *numeric*

Algoritma *k-means* hanya mengambil beberapa sampel dari seluruh populasi komponen yang didapatkan agar kemudian dijadikan pusat cluster awal, pada penentuan pusat cluster ini dipilih dengan cara acak dari populasi data.

Kemudian algoritma k-means akan menguji setiap komponennya ke dalam jumlah populasi data tersebut dan menandai komponen-komponen tersebut ke dalam salah satu pusat *cluster* yang telah dideskripsikan sebelumnya tergantung dari jarak minimum antar komponen dengan setiap pusat *cluster* yang ada. Selanjutnya posisi pusat *cluster*nya akan dihitung kembali sampai semua komponen data dikelompokkan ke dalam setiap *cluster* dan terakhir akan terbentuk *cluster* baru [11].

### 3. HASIL DAN PEMBAHASAN

#### 3.1 Penerapan Metode K-Means

Penerapan Metode *K-Means* merupakan langkah penyelesaian terkait pengelompokkan penerima vaksinasi *Covid-19* secara berurutan sesuai dengan referensi yang telah digunakan:

##### 3.1.1 Persiapan Data

Berikut ini merupakan data sampel penjumlahan yang nantinya akan dilakukan proses perhitungan:

Tabel 1. Data Penerima Vaksinasi

| No | Kode | Bulan     | Dosis 1 | Dosis 2 | Booster |
|----|------|-----------|---------|---------|---------|
| 1  | B1   | Januari   | 140     | 65      | 30      |
| 2  | B2   | Februari  | 57      | 65      | 15      |
| 3  | B3   | Maret     | 110     | 60      | 25      |
| 4  | B4   | April     | 180     | 90      | 53      |
| 5  | B5   | Mei       | 101     | 60      | 44      |
| 6  | B6   | Juni      | 284     | 100     | 75      |
| 7  | B7   | Juli      | 205     | 53      | 78      |
| 8  | B8   | Agustus   | 137     | 78      | 68      |
| 9  | B9   | September | 178     | 48      | 40      |
| 10 | B10  | Oktober   | 301     | 30      | 114     |
| 11 | B11  | November  | 257     | 53      | 87      |
| 12 | B12  | Desember  | 473     | 72      | 148     |

##### 3.1.2 Proses Perhitungan Metode K-Means

Berikut ini merupakan proses perhitungan Metode *K-Means*:

$$d(x,y) = \|x-y\| = \sqrt{\sum_{i=1}^n (X_i - Y_i)^2}; i = 1, 2, 3 \dots n$$

Penerapan jumlah *cluster* (K) yaitu 3 *cluster*, Setelah menetapkan jumlah *cluster*, Tentukan titik pusat awal *cluster* (*Centroid*), Berikut ini titik *Centroid* yang telah dipilih:

Tabel 2. Data Centroid

| Centroid   | Kode | Dosis 1 | Dosis 2 | Booster | Keadaan |
|------------|------|---------|---------|---------|---------|
| Centroid 1 | B12  | 473     | 72      | 148     | Tinggi  |
| Centroid 2 | B6   | 284     | 100     | 75      | Sedang  |
| Centroid 3 | B2   | 57      | 65      | 15      | Rendah  |

Hitung jarak data ke *Centroid* menggunakan rumus *Euclidean*, data tersebut akan ditetapkan sebagai anggota dari *cluster* terdekatnya. Menghitung *Distance* (jarak) antara *variable* dari setiap sampel data dengan *Centroid* yaitu:

1. Dengan *Centroid* B12(65;3;6)
  1. Jarak antara B12 dengan titik B12
 
$$= \sqrt{\sum_{i=1}^n (X_i - Y_i)^2}$$

$$= \sqrt{(473 - 473)^2 + (72 - 72)^2 + (148 - 148)^2}$$

$$= 0,00$$
2. Dengan *Centroid* B21 (23;1;5)
  2. Jarak antara B6 dengan titik B2
 
$$= \sqrt{\sum_{i=1}^n (X_i - Y_i)^2}$$

$$= \sqrt{(473 - 284)^2 + (72 - 21)^2 + (148 - 36)^2}$$

$$= 15334$$
3. Dengan *Centroid* B2 (14;3;2)
  3. Jarak antara B12 dengan titik B2

$$\begin{aligned}
 &= \sqrt{\sum_{i=1}^n (X_i - Y_i)^2} \\
 &= \sqrt{(473 - 57)^2 + (72 - 1)^2 + (148 - 3)^2} \\
 &= 26482
 \end{aligned}$$

Lakukan proses perhitungan yang sama dengan objek ke 12, Adapun hasil dari perhitungan literasi 1 dapat dilihat pada table dibawah ini. Dimana jarak terdekat dilihat dari perhitungan yang paling dekat ke pusat cluster. Sementara WCV (*Within Cluster Variatio*) adalah keadaan dari hasil pangkat dari perhitungan jarak terdekat ke pusat cluster.

Tabel 3. Hasil Iterasi I

| No    | Nama | C1    | C2   | C3    | Cluster | JCT  | JCT^2    |
|-------|------|-------|------|-------|---------|------|----------|
| 1     | B1   | 25682 | 1466 | 6894  | C2      | 1466 | 2149156  |
| 2     | B2   | 26482 | 1716 | 0     | C3      | 0    | 0        |
| 3     | B3   | 25573 | 1459 | 2817  | C2      | 1459 | 2128681  |
| 4     | B4   | 21679 | 661  | 15361 | C3      | 661  | 436921   |
| 5     | B5   | 14437 | 283  | 3813  | C3      | 283  | 80089    |
| 6     | B6   | 20707 | 585  | 4329  | C3      | 585  | 342225   |
| 7     | B7   | 22849 | 833  | 22029 | C3      | 833  | 693889   |
| 8     | B8   | 22917 | 901  | 6525  | C3      | 901  | 811801   |
| 9     | B9   | 19080 | 444  | 15326 | C3      | 444  | 197136   |
| 10    | B10  | 3092  | 6182 | 72698 | C1      | 3092 | 9560464  |
| 11    | B11  | 4298  | 3652 | 49760 | C1      | 3652 | 13337104 |
| 12    | B12  | 4753  | 3391 | 37301 | C1      | 3391 | 11498881 |
| TOTAL |      |       |      |       |         |      | 42748476 |

Berdasarkan perhitungan yang dilakukan pada tabel diatas, maka dapat disimpulkan sebagai berikut:

-C1 = {B10,B11,B12}

-C2 = {B1,B3 }

- C3 = {B2,B4,B5,B6,B7,B8,B9,B13,B14,B15,B16,B17,B18,B19,B20,B21,B22 }

Keterangan:

BCV: *Between Cluster Variation*

WCV: *Within Cluster Variation*

Pada langkah ini dihitung pula rasio besaran BCV dan WCV:

Karena *Centroid* m1 = (65;3;6),m2 = (23;1;5), m3 = (4;3;2)

$$\begin{aligned}
 d(m1,m2) &= \sqrt{(472 - 284)^2 + (72 - 21)^2 + (148 - 36)^2} \\
 &= 15334
 \end{aligned}$$

$$\begin{aligned}
 d(m1,m3) &= \sqrt{(472 - 57)^2 + (72 - 1)^2 + (148 - 3)^2} \\
 &= 26482
 \end{aligned}$$

$$\begin{aligned}
 d(m2,m3) &= \sqrt{(284 - 57)^2 + (21 - 1)^2 + (36 - 3)^2} \\
 &= 1716
 \end{aligned}$$

$$BCV = d(m1,m2) + d(m1,m3) + d(m2,m3) = 43532$$

WCV = adalah memilih jumlah jarak terkecil di pangkat dua antara data dengan

*centroid* pada masing-masing cluster = 42748476

Sehingga Besar Rasio = BCV/WCV = 43532/ 42748476= **0,001**

Lakukan pembaruan *Centroid* dari hasil cluster seperti berikut:

C1 = rata-rata { B10,B11,B12} = (261;39;98)

C2 = rata-rata { B1, B3 } = (125;3;5)

C3 = rata-rata { B2,B4,B5,B6,B7,B8,B9,B10 } = (199;9;46)

Nilai *Centroid* berubah dari nilai *Centroid* sebelumnya, maka algoritma. Selanjutnya, dilakukan perhitungan dengan rumus yang sama hingga iterasi selanjutnya. Berikut ini adalah hasil perhitungan setelah dilakukan sebanyak 2 Iterasi menggunakan algoritma K-Means:

Tabel 4. Hasil Iterasi 2

| No    | Nama | C1    | C2    | C3   | Cluster | JCT | JCT^2    |
|-------|------|-------|-------|------|---------|-----|----------|
| 1     | B1   | 10102 | 15    | 2008 | C3      | 15  | 233      |
| 2     | B2   | 10635 | 74    | 2293 | C3      | 74  | 5513     |
| 3     | B3   | 10058 | 15    | 2006 | C2      | 15  | 233      |
| 4     | B4   | 7634  | 219   | 990  | C1      | 219 | 48071    |
| 5     | B5   | 3632  | 1701  | 114  | C2      | 114 | 13054    |
| 6     | B6   | 7064  | 281   | 834  | C1      | 281 | 79102    |
| 7     | B7   | 8336  | 156   | 1248 | C1      | 156 | 24414    |
| 8     | B8   | 8404  | 88    | 1304 | C3      | 88  | 7788     |
| 9     | B9   | 6047  | 631   | 519  | C2      | 519 | 269104   |
| 10    | B10  | 394   | 12813 | 4865 | C1      | 394 | 155148   |
| 11    | B11  | 305   | 9406  | 2933 | C1      | 305 | 92754    |
| 12    | B12  | 87    | 8725  | 2419 | C1      | 87  | 7550     |
| TOTAL |      |       |       |      |         |     | 16049450 |

Pada langkah ini dihitung pula rasio besaran BCV dan WCV:

Karena *Centroid* M1 (201;21;53), m2 = (230;15;35), m3 = (134;9;24)

$$d(m1,m2) = \sqrt{(201 - 230)^2 + (21 - 15)^2 + (53 - 35)^2}$$

$$= 10080$$

$$d(m1,m3) = \sqrt{(201 - 134)^2 + (21 - 9)^2 + (53 - 24)^2}$$

$$= 3170$$

$$d(m2,m3) = \sqrt{(230 - 134)^2 + (15 - 9)^2 + (35 - 24)^2}$$

$$= 2007$$

$$BCV = d(m1,m2) + d(m1,m3) + d(m2,m3) = 15257$$

WCV = adalah memilih jumlah jarak terkecil di pangkat dua antara data dengan *centroid* pada masing-masing cluster = 16049450

$$\text{Sehingga Besar Rasio} = BCV/WCV = 15257/16049450 = \mathbf{0,001}$$

Rasio tidak berubah dari literasi 1 ke literasi 2 maka *clustering* selesai dilakukan.

Tabel 5. Kesimpulan Hasil Iterasi

| NO | KODE BULAN | Bulan     | CLUSTER | KETERANGAN |
|----|------------|-----------|---------|------------|
| 1  | B1         | Januari   | C3      | Rendah     |
| 2  | B2         | Februari  | C3      | Rendah     |
| 4  | B3         | Maret     | C3      | Rendah     |
| 5  | B4         | April     | C2      | Sedang     |
| 6  | B5         | Mei       | C3      | Rendah     |
| 7  | B6         | Juni      | C2      | Sedang     |
| 3  | B7         | Juli      | C2      | Sedang     |
| 8  | B8         | Agustus   | C3      | Rendah     |
| 9  | B9         | September | C2      | Sedang     |
| 10 | B10        | Oktober   | C2      | Sedang     |
| 11 | B11        | November  | C2      | Sedang     |
| 12 | B12        | Desember  | C1      | Tinggi     |

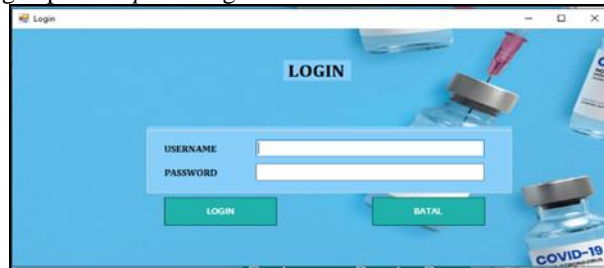
Maka dari hasil *klustering* diatas didapatkan kesimpulan yaitu pengelompokan Data Vaksinasi di bagi atas tiga kategori yaitu Tinggi, Sedang, dan Rendah. Hasil ini bisa menjadi informasi penting untuk membuat sebuah kebijakan-kebijakan pihak Puskesmas Bandar Baru Cabang Sibolangit untuk melakukan sosialisasi terhadap data vaksinasi Covid 19 yang dominan.

### 3.2 Implementasi Sistem

Berikut ini merupakan hasil implementasi sistem yang telah dibangun dengan berbasis *Desktop* menggunakan *Microsoft Visual Studio 2008* dan *database Microsoft Acces*.

a. *Form Login*

*Form login* berfungsi sebagai validasi akses dari admin untuk masuk kedalam sistem, pada *form login* terdapat *username* dan *password* yang dapat di *input* sebagai data validasi.



Gambar 1. Tampilan *Form Login*

b. *Form Menu Utama*

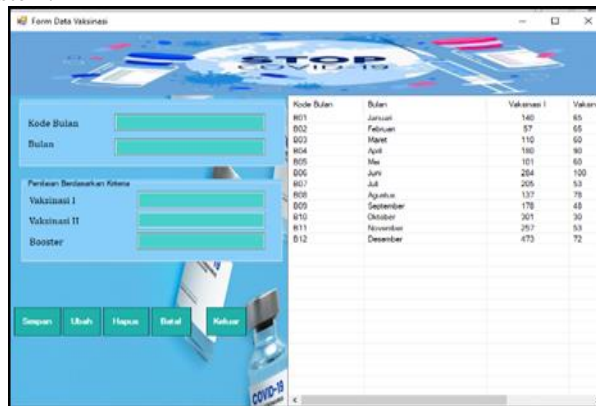
*Form Menu Utama* berfungsi sebagai halaman navigasi untuk membuka menu-menu yang lainnya.



Gambar 2. Tampilan *Form Menu Utama*

c. *Form Data Vaksinasi*

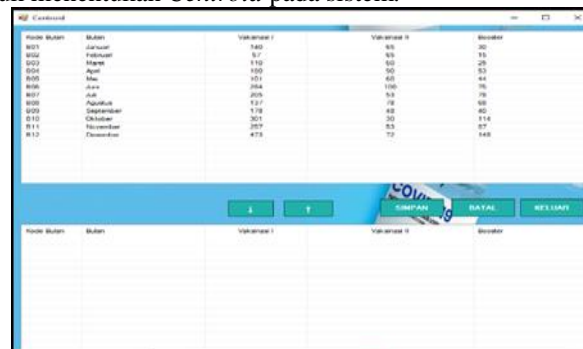
*Form Data Vaksinasi* berfungsi untuk mengelola data Vaksinasi seperti menampilkan, menyimpan, menghapus dan mengubah Vaksinasi pada sistem.



Gambar 3. Tampilan *Form Data Vaksinasi*

d. *Form Centroid*

*Form Centroid* berfungsi untuk menentukan *Centroid* pada sistem.



Gambar 4. Tampilan *Form Centroid*

e. *Form Proses K-Means*

*Form Proses K-Means* berfungsi untuk melakukan proses *Clustering* data menggunakan algoritma *K-Means Clustering*.



Gambar 5. Tampilan Form Proses K-Means

#### 4. KESIMPULAN

Menganalisa masalah dan menerapkan metode K-Means dalam mengelompokkan jumlah data penerima vaksinasi Covid-19 pada Puskesmas Bandar Baru Kecamatan Sibolangit menggunakan data penerima vaksinasi yang diberikan oleh pihak puskesmas dan menggunakan kriteria Vaksinasi I, Vaksinasi II dan Vaksin Booster. Merancang aplikasi dengan metode K-Means clustering dalam pengelompokkan data penerima vaksinasi Covid-19 di Puskesmas Bandar Baru Kecamatan Sibolangit menggunakan pemrograman desktop Microsoft Visual Studio 2010 dan menggunakan database Microsoft Access 2010. Mengimplementasikan metode K-Means clustering dalam pengelompokkan data penerima vaksinasi Covid-19 di Puskesmas Bandar Baru Kecamatan Sibolangit dengan data yang akurat dan perhitungan cluster berdasarkan kriteria Vaksinasi I, Vaksinasi II dan Vaksin Booster

#### UCAPAN TERIMAKASIH

Puji dan Syukur dipanjatkan kepada Tuhan Yang Maha Esa yang memberikan karunia sehingga mampu menyelesaikan jurnal ini. Kemudian kepada Bapak Mukhlis Ramadhan dan Ibu Erika Fahmi Ginting atas segala waktu dan ilmunya yang telah memberikan bimbingan selama masa pengerjaan hingga menyelesaikan jurnal ini dan kepada seluruh dosen serta pegawai kampus STMIK Triguna Dharma yang telah banyak membantu baik dari segi informasi ataupun dukungan lainnya.

#### DAFTAR PUSTAKA

- [1] Rachmadi, T., Rahayu, T. P., Waluyo, A., & Yuliyanto, W. (2021). Pemberian vaksinasi covid-19 bagi masyarakat kelompok petugas pelayanan publik di Kecamatan Buluspesantren. *JURPIKAT (Jurnal Pengabdian Kepada Masyarakat)*, 2(2), 318-333.
- [2] M. Ramadhan, J. Hutagalung, M. Dahria, I. Zulkarnain, and H. Jaya, "Prediksi Penjualan Spare Part Mobil Daihatsu Menggunakan Algoritma Apriori," *Techno.Com*, vol. 22, no. 1, pp. 156-166, 2023, doi: 10.33633/tc.v22i1.7192.
- [3] A. Gusderia, M. Ramadhan, and M. I. Perangin-angin, "Data Mining Untuk Klasifikasi Data Penjualan Alat Teknik Menggunakan Metode Naive Bayesian Clacifier," vol. 21, no. 2, pp. 73-79, 2022.
- [4] Y. Syahra, "Penerapan Data Mining Dalam Pengelompokkan Data Nilai Siswa Untuk Penentuan Jurusan Siswa Pada SMA Tamora Menggunakan Algoritma K-Means Clustering," *J. SAINTIKOM (Jurnal Sains Manaj. Inform. dan Komputer)*, vol. 17, no. 2, p. 228, 2018, doi: 10.53513/jis.v17i2.70.
- [5] A. Fitri, Y. Syahra, and R. Kustini, "Penerapan Data Mining Dalam Mengklusterisasi Location Best Pb Tambahan Pada Regional IV PT Indomarco Prismatama Cab.Medan Dengan Menggunakan Metode K-Means," *J. SAINTIKOM (Jurnal Sains Manaj. Inform. dan Komputer)*, vol. 19, no. 2, p. 11, 2020, doi: 10.53513/jis.v19i2.2330.
- [6] S. D. Prasetiani and N. Rochmawati, "Penerapan Data Mining Untuk Clustering Menu Favorit Menggunakan Algoritma K-Means (Studi Kasus Kedai Expo)," vol. 03, pp. 278-286, 2022.
- [7] R. Girsang, E. F. Ginting, and M. Hutasuhut, "Penerapan Algoritma C4.5 Pada Penentuan Penerima Program Bantuan Pemerintah Daerah," *J. Sist. Inf. Triguna Dharma (JURSI TGD)*, vol. 1, no. 4, p. 449, 2022, doi: 10.53513/jursi.v1i4.5727.
- [8] D. P. Utomo, "Analisis Komparasi Metode Klasifikasi Data Mining dan Reduksi Atribut Pada Data Set Penyakit Jantung," vol. 4, no. April, pp. 437-444, 2020, doi: 10.30865/mib.v4i2.2080.
- [9] R. Aditiya, S. Defit, and G. W. Nurcahyo, "Prediksi Tingkat Ketersediaan Stock Sembako Menggunakan Algoritma FP-Growth dalam Meningkatkan Penjualan," *J. Inform. Ekon. Bisnis*, vol. 2, pp. 67-73, 2020, doi: 10.37034/infv2i3.44.
- [10] F. Yunita, "Penerapan Data Mining Menggunakan Algoritma K-Means Clustering Pada Penerimaan Mahasiswa Baru," *Sistemasi*, vol. 7, no. 3, p. 238, 2018, doi: 10.32520/stmsi.v7i3.388.
- [11] M. Syukri Mustafa, M. Rizky Ramadhan, and A. P. Thenata, "Implementasi Data Mining untuk Evaluasi Kinerja Akademik Mahasiswa Menggunakan Algoritma Naive Bayes Classifier," *Citec J.*, vol. 4, no. 2, pp. 151-162, 2017.