

Penerapan Data Mining Pengelompokan Data Vaksinasi Covid-19 Menggunakan Metode Clustering

Leni Tri Ramadhayanti, Relita Buaton, Yani Maulita

Program Studi Sistem Informasi, STMIK Kaputama, Binjai, Indonesia

Email: ¹lenitri@gmail.com

Abstrak—Vaksin adalah produk biologi yang berisi antigen berupa mikroorganisme atau bagiannya atau zat yang dihasilkannya yang telah diolah sedemikian rupa sehingga aman, yang apabila diberikan kepada seseorang akan menimbulkan kekebalan spesifik secara aktif terhadap penyakit tertentu. Vaksinasi merupakan proses di dalam tubuh sehingga seseorang menjadi kebal atau terlindungi dari suatu penyakit. Banyaknya jumlah pendaftar yang ingin melakukan vaksinasi Covid-19 di Puskesmas Kebun Lada, hal ini membuat banyaknya tumpukan data vaksinasi yang ingin melakukan vaksin. Data-data tersebut selain diinputkan secara langsung di website pemerintah, file hard copy hanya disimpan dalam sebuah file yang kemudian disimpan kedalam sebuah map. Melihat situasi ini tentu dari data-data tersebut dapat diambil sebuah informasi yang baru yang diolah dengan menggunakan teknik data mining untuk menggali sebuah informasi yang berguna terkait data vaksinasi. Data mining dapat membantu perusahaan dalam menggali pengetahuan baru dengan cara memproses data yang ada dengan metode clustering dan menggunakan algoritma K-Means. Diperoleh hasil yang serupa pada cluster 1 dan 2 yaitu alamat yang berasal dari binjai utara sementara pada cluster 3 terdapat kesamaan hasil dengan cluster 1 yaitu pada kategori penerima vaksin yaitu masyarakat umum.

Kata Kunci: Data Mining; Clustering; Vaksinasi, Covid 19; Matlab

Abstract— Vaccines are biological products containing antigens in the form of microorganisms or parts thereof or substances they produce which have been processed in such a way that they are safe, which when given to a person will cause active specific immunity against certain diseases. Vaccination is a process in the body so that a person becomes immune or protected from a disease. The large number of registrants who want to vaccinate against Covid-19 at the Kebun Lada Health Center has created a large pile of vaccination data that wants to vaccinate. These data are not only inputted directly on the government website, hard copy files are only stored in a file which is then stored in a folder. Seeing this situation, of course, from these data new information can be taken which is processed using data mining techniques to dig up useful information related to vaccination data. Data mining can help companies explore new knowledge by processing existing data with clustering methods and using the K-Means algorithm. Similar results were obtained in clusters 1 and 2, namely addresses originating from north binjai, while in cluster 3 there were similar results with cluster 1, namely in the category of vaccine recipients, namely the general public.

Keywords: Data Mining; Clustering; Vaccination; Covid 19; Matlab

1. PENDAHULUAN

Vaksin adalah produk biologi yang berisi antigen berupa mikroorganisme atau bagiannya atau zat yang dihasilkannya yang telah diolah sedemikian rupa sehingga aman, yang apabila diberikan kepada seseorang akan menimbulkan kekebalan spesifik secara aktif terhadap penyakit tertentu. Puskesmas Kebun Lada merupakan salah satu tempat pelayanan kesehatan di Kota Binjai yang menangani proses vaksinasi covid-19 kepada masyarakat yang berada di Kota Binjai khususnya di Kelurahan Kebun Lada. Alur Proses vaksinasi Covid-19 yaitu masyarakat datang dengan membawa KTP Elektronik, kemudian mendaftar ke bagian pendaftaran vaksin untuk di data, setelah mendaftar maka menunggu untuk di periksa kesehatannya dan kemudian dilakukan vaksinasi.

Banyaknya jumlah pendaftar yang ingin melakukan vaksinasi Covid-19 di Puskesmas Kebun Lada, hal ini membuat banyaknya tumpukan data vaksinasi yang ingin melakukan vaksin. Data-data tersebut selain diinputkan secara langsung di website pemerintah, file *hard copy* hanya disimpan dalam sebuah file yang kemudian disimpan kedalam sebuah map. Melihat situasi ini tentu dari data-data tersebut dapat diambil sebuah informasi yang baru yang diolah dengan menggunakan teknik data mining untuk menggali sebuah informasi yang berguna terkait data vaksinasi. Data baru tersebut nantinya diharapkan dapat menjadi masukan demi meningkatkan pelayanan kesehatan pada Puskesmas Kebun Lada Kota Binjai.

2. METODOLOGI PENELITIAN

2.1 Vaksin

Vaksinasi adalah pemberian vaksin (antigen) yang dapat merangsang pembentukan imunitas (antibodi) sistem imun di dalam tubuh. Vaksinasi sebagai upaya pencegahan primer yang sangat handal mencegah penyakit yang dapat dicegah dengan vaksinasi. Vaksinasi Covid-19 dilakukan setelah kepastian keamanan dan keampuhannya ada, merupakan upaya untuk menurunkan kesakitan dan kematian dan mendorong terbentuknya kekebalan kelompok (herd imunity). Selain itu, vaksinasi Covid-19 bertujuan untuk melindungi dan memperkuat sistem kesehatan secara menyeluruh, juga menjaga produktivitas dan mengurangi dampak sosial dan ekonomi masyarakat.

2.2 Data Mining

Data mining merupakan langkah analisis terhadap sekumpulan data yang umumnya berukuran besar untuk mendapatkan hubungan antar data tersebut dan merangkumnya dalam bentuk yang mudah dipahami dan digunakan. Hubungan dan rangkuman yang dihasilkan tersebut pada umumnya berupa model (*models*) atau pola (*pattern*).

2.3 Algoritma K-Means

mengelompokkan sejumlah besar obyek dengan atribut tertentu ke dalam kelompok-kelompok (*cluster*) sebanyak K. Pada algoritma *K-Means*, jumlah *cluster* K sudah ditentukan lebih dahulu. *K-Means* adalah salah satu metode pengelompokan non hirarki (sekatan) yang berusaha mempartisi data ke dalam *cluster*/kelompok sehingga data yang memiliki karakteristik yang sama akan dimasukkan ke dalam satu *cluster* yang sama dan data yang memiliki karakteristik yang berbeda dikelompokkan ke dalam kelompok yang lain.

3. HASIL DAN PEMBAHASAN

3.1 Analisa Masalah

Dalam penelitian ini telah dilakukan analisis kelemahan sistem yang berjalan, bahwa pada Puskesmas Kebun Lada Binjai belum memiliki sistem pengelompokan data mining untuk memproses sebuah data yang ada menjadi informasi yang baru dalam hal ini data pengelompokan data vaksinasi Covid-19. Proses yang terjadi yaitu data-data pengelompokan penerima vaksinasi Covid-19 hanya tersimpan dalam bentuk file buku dan disimpan pada lemari penyimpanan berkas. Dilakukan analisis sistem terhadap teknik data mining pengelompokan data mining data vaksinasi Covid-19 dengan menggunakan metode *clustering* yang diproses menggunakan *software matlab* dan *Microsoft excel* sebagai database. Menentukan kebutuhan sistem sehingga dapat dilakukan penganalisaan data mining pengelompokan data vaksinasi Covid-19 dengan menggunakan metode *clustering*. Data data vaksinasi Covid-19 seperti, usia, kategori penerima vaksin dan alamat yang digunakan, yang kemudian diproses kedalam data mining dengan menggunakan *clustering*. Sistem yang baru yang akan dibangun dalam penelitian ini diharapkan dapat diterapkan dan memproses pengelompokan data data vaksinasi Covid-19 dengan menggunakan metode *clustering*. Selanjutnya data tersebut di proses kedalam sistem dengan menggunakan *Software Matlab* dan GUI sebagai interface yang dapat memudahkan user dalam proses pengaplikasiannya.

3.1.1 Penerapan Algoritma K-Means

Pada perancangan sistem pengelompokan data mining data vaksinasi Covid-19 dengan menggunakan metode *clustering*. Akan digambarkan dalam perancangan proses, yaitu bagaimana alur dan *step by step* sistem ini akan dijalankan dan digunakan sehingga dapat menghasilkan suatu informasi data yang baru dari pengelompokan data vaksinasi Covid-19.

Tabel 1. Data Vaksinasi Covid-19

No	Object	Usia	Kategori Penerima Vaksin	Alamat
1	A	31	Masyarakat Umum	Binjai Utara
2	B	28	Pelayanan Publik	Luar Binjai
3	C	26	Pelayanan Publik	Luar Binjai
4	D	43	Pelayanan Publik	Binjai Utara
5	E	19	Masyarakat Umum	Luar Binjai
6	F	33	Pelayanan Publik	Binjai Selatan
7	G	27	Masyarakat Umum	Luar Binjai
8	H	24	Masyarakat Umum	Luar Binjai
9	I	33	Pelayanan Publik	Binjai Utara
10	J	59	Masyarakat Renta	Binjai Timur
11	K	50	Masyarakat Renta	Binjai Utara
12	L	51	Masyarakat Renta	Binjai Utara
13	M	25	Masyarakat Umum	Binjai Timur
14	N	40	Pelayanan Publik	Luar Binjai
15	O	20	Masyarakat Umum	Binjai Timur
16	P	49	Masyarakat Umum	Binjai Barat
17	Q	47	Pelayanan Publik	Binjai Utara
18	R	26	Masyarakat Umum	Binjai Utara
19	S	24	Masyarakat Umum	Binjai Selatan
20	T	24	Masyarakat Umum	Binjai Selatan

Kemudian data perlu ditransformasikan dahulu agar mempermudah proses perhitungan menggunakan rumus *euclidean distance*.

Tabel 2. Hasil Perhitungan Iterasi 1

No	Object	X	Y	Z	C ₁	C ₂	C ₃	Group
1	A	2	2	5	2,24	2,24	1,00	3
2	B	2	1	6	2,24	3,00	1,00	3
3	C	2	1	6	2,24	3,00	1,00	3
4	D	3	1	5	1,00	2,24	1,73	1
5	E	1	2	6	3,32	0,00	1,00	2
6	F	2	1	3	2,83	3,46	3,16	1
7	G	2	2	6	2,45	2,45	0,00	3

No	Object	X	Y	Z	C ₁	C ₂	C ₃	Group
8	H	1	2	6	3,32	3,32	1,00	3
9	I	2	1	5	2,00	2,83	1,41	3
10	J	4	3	4	2,24	1,00	3,00	2
11	K	4	3	5	2,00	0,00	2,45	2
12	L	4	3	5	2,00	0,00	2,45	2
13	M	1	2	4	3,32	3,32	2,24	3
14	N	3	1	6	1,41	2,45	1,41	1
15	O	1	2	4	3,32	3,32	2,24	3
16	P	4	2	1	4,12	4,12	5,39	1
17	Q	4	1	5	0,00	2,00	2,45	1
18	R	2	2	5	2,24	2,24	1,00	3
19	S	1	2	3	3,74	3,74	3,16	3
20	T	1	2	3	3,74	3,74	3,16	3

Setelah dilakukan perhitungan menggunakan rumus *cluster* yang ada terjadi perubahan group dilanjutkan ke iterasi berikut :

Iterasi 2 :

K = 3

Centroid 1 Group 1

$$C_1 = \left(\frac{3+2+3+4+4}{5} \right) = 3,20 \quad \left(\frac{1+1+1+2+1}{6} \right) = 1,20 \quad \left(\frac{5+4+6+1+5}{6} \right) = 4,20$$

$$= (3,20, 1,20, 4,20)$$

Centroid 2 Group 2

$$C_2 = \left(\frac{1+4+4+4}{4} \right) = 3,25 \quad \left(\frac{2+3+3+3}{4} \right) = 2,75 \quad \left(\frac{6+4+5+5}{4} \right) = 5$$

$$= (3,25, 2,7, 5)$$

Centroid 3 Group 3

$$C_3 = \left(\frac{2+2+2+2+1+2+1+1+2+1+1}{11} \right) = 1,55$$

$$\left(\frac{2+1+1+2+2+1+2+2+2+2+2}{11} \right) = 1,73$$

$$\left(\frac{5+6+6+6+6+5+4+4+5+3+3}{11} \right) = 4,82$$

$$= (1,55, 1,73, 4,33)$$

Jadi K=3 Centroid

$$C_1 = (3,20, 1,20, 4,20)$$

$$C_2 = (3,25, 2,75, 5)$$

$$C_3 = (1,55, 1,73, 4,33)$$

Tabel 3. Hasil Perhitungan Iterasi 2

No	Object	X	Y	Z	C ₁	C ₂	C ₃	Group
1	A	2	2	5	1,65	1,46	0,56	3
2	B	2	1	6	2,17	2,37	1,46	3
3	C	2	1	6	2,17	2,37	1,46	3
4	D	3	1	5	0,85	1,77	1,64	1
5	E	1	2	6	2,95	0,00	1,33	2
6	F	2	1	3	1,71	2,94	2,01	1
7	G	2	2	6	2,31	1,77	1,30	3
8	H	1	2	6	2,95	2,57	1,33	3
9	I	2	1	5	1,46	2,15	0,88	3
10	J	4	3	4	1,98	1,27	2,88	2
11	K	4	3	5	2,13	0,79	2,77	2
12	L	4	3	5	2,13	0,79	2,77	2
13	M	1	2	4	2,35	2,57	1,02	3
14	N	3	1	6	1,82	2,03	2,01	1
15	O	1	2	4	2,35	2,57	1,02	3
16	P	4	2	1	3,39	4,14	4,55	1
17	Q	4	1	5	1,15	1,90	2,57	1
18	R	2	2	5	1,65	1,46	0,56	3
19	S	1	2	3	2,63	3,10	1,92	3
20	T	1	2	3	2,63	3,10	1,92	3

Setelah dilakukan perhitungan menggunakan rumus *cluster* yang ada, pada iterasi 2 terjadi sama seperti iterasi pertama dan tidak ada data yang berpindah group lagi sehingga perhitungan dapat dihentikan. Sehingga dapat dibuat grafik cluster data vaksinasi Covid-19. Dari 20 data diperoleh 3 *group*, Cluster 1 terdapat 5 data, Cluster 2 terdapat 4 data, dan Cluster 3 terdapat 11 data.

1. Cluster 1 Terdapat 5 Data (3 1 4)
Dapat diketahui pada *cluster* 1 terdapat kriteria usia yaitu 36-45 tahun (Dewasa Akhir) dengan kategori penerima vaksin pelayanan publik dan alamat Binjai Timur.
2. Cluster 2 Terdapat 4 Data (3 3 5)
Dapat diketahui pada *cluster* 2 terdapat kriteria usia yaitu 36-45 tahun (Dewasa Akhir) dengan kategori penerima vaksin masyarakat renta dan alamat Binjai Utara.
3. Cluster 3 Terdapat 11 Data (1 2 4)
Dapat diketahui pada *cluster* 2 terdapat kriteria usia yaitu 17-25 tahun (Remaja) dengan kategori penerima vaksin masyarakat umum dan alamat Binjai Timur.

4. KESIMPULAN

Dari hasil analisa berdasarkan *clustering* data pengelompokan vaksinasi covid-19, maka dapat diambil suatu kesimpulan, dari data 509 data terdapat 3 cluster yang dikelompokkan Berdasarkan data kriteria vaksinasi covid-19, seperti usia, kategori penerima vaksin dan alamat dapat dikelompokkan dengan menggunakan metode *clustering*. Dengan data vaksinasi covid-19, dapat membangun aplikasi data mining dengan menggunakan metode *clustering* yang dapat mengelompokkan data vaksinasi covid-19. Berdasarkan analisis yang telah dilakukan pada pengelompokan data vaksinasi covid-19 dengan menggunakan metode *clustering*, maka perlu dilakukan proses cluster beberapa kali untuk mendapatkan hasil yang sama sesuai dengan proses yang pertama dikali dilakukan. diperoleh hasil yang serupa pada cluster 1 dan 2 yaitu alamat yang berasal dari Binjai Utara, Sementara pada cluster 3 terdapat kesamaan hasil dengan cluster 1 yaitu pada kategori penerima vaksin yaitu masyarakat umum.

REFERENCES

- [1] Away, G. A. (2014). *The Shortcut of MATLAB Programming*. Informatika. Bandung.
- [2] Handoko, S., Tri, E., & Handayani, E. (2020). IMPLEMENTASI DATA MINING UNTUK MENENTUKAN TINGKAT PENJUALAN PAKET DATA TELKOMSEL MENGGUNAKAN METODE K-MEANS CLUSTERING. *Jurnal Ilmiah Teknologi Dan Rekayasa*, 25(1).
- [3] Hermawati, F. A. (2013). *Data Mining*. CV. Andi Offset, Yogyakarta.
- [4] Indra Yatini, B. (2010). *Flowchart, Algoritma dan Pemrograman menggunakan Bahasa C++ Builder*. Graha Ilmu. Yogyakarta.
- [5] Indraputra, R. A., & Fitriana, R. (2020). K-Means Clustering Data COVID-19. *Jurnal Teknik Industri*, 10(3), 275–282.
- [6] Noviyanto. (2020). Penerapan Data Mining Dalam Mengelompokkan Jumlah Kematian. *Paradigma-Jurnal Informatika Dan Komputer*, 22(2).
- [7] Prasetyo, E. (2012). *Data Mining: Konsep dan Aplikasi menggunakan MATLAB*. CV. Andi Offset. Yogyakarta.
- [8] Prasetyowati, E. (2017). *Data Mining Pengelompokan Data Informasi dan Evaluasi* (M. Afandi (ed.)). Duta Media , Pemekasan.
- [9] Sugiarti, Y. (2013). *Analisis dan Perancangan UML (United Modeling Language) Generated VB.6*. Graha Ilmu. Yogyakarta.