

Penerapan Metode Clustering Dalam Menentukan Persentase Penyakit Pada Pasien Dengan Menggunakan Algoritma K-Means Pada Puskesmas Jawa Maraja, Kab. Simalungun

Ughie Yulika**, Ahmad Fitri Boy, S.Kom., M. Kom.**, Guntur Syahputra, S.Kom., M.Kom.**

* Program Studi Mahasiswa, STMIK Triguna Dharma

** Program Studi Dosen Pembimbing, STMIK Triguna Dharma

Article Info

Article history:

Received Jun 12th, 2020

Revised Aug 20th, 2020

Accepted Aug 26th, 2020

Keyword:

Data Mining, Metode Clustering, Algoritma K-Means Pengelompokan, Penyakit Pada Pasien Puskesmas Jawa Maraja.

ABSTRACT

Pada dasarnya setiap orang pasti pernah terserang penyakit, baik itu penyakit biasa ataupun kronis. Tubuh manusia tersusun atas berbagai macam organ dan juga sel jaringan yang rentan terserang penyakit. Persentase penyakit di puskesmas Jawa Maraja tidak dapat diprediksi. Karena banyak pasien yang mengeluh tentang penyakit yang berbeda-beda. Kondisi tersebut mengakibatkan karyawan di Puskesmas Jawa Maraja kesulitan untuk membuat laporan data penyakit pasien. Disini menggunakan metode Clustering untuk mengelompokan penyakit pasien yang sering diperiksa oleh karyawan Puskesmas Jawa Maraja. Data penyakit pasien yang masih ditulis secara manual mempersulit karyawan yang bekerja di puskesmas. Dan juga pegawai kesulitan untuk menghitung persentase penyakit pada pasien di Puskesmas Jawa Maraja karena dihitung secara manual. Dengan adanya pengelompokan penyakit pada pasien di Puskesmas Jawa Maraja berbasis desktop akan membantu karyawan yang bertanggung jawab untuk membuat laporan penyakit pasien menjadi lebih mudah. Penelitian ini juga diharapkan dapat memberikan solusi kepada Puskesmas Jawa Maraja untuk dapat menjadi Puskesmas yang lebih baik kedepannya.

Copyright © 2020 STMIK Triguna Dharma.
All rights reserved.

Corresponding Author:

Nama : Ughie Yulika

Kampus : STMIK Triguna Dharma

Program Studi : Sistem Informasi

E-Mail : ughieyulika12345@gmail.com

1. PENDAHULUAN

Pada dasarnya setiap orang pasti pernah terserang penyakit, baik itu penyakit biasa ataupun kronis. Tubuh manusia tersusun atas berbagai macam organ dan juga sel jaringan yang rentan terserang penyakit. Persentase penyakit di puskesmas Jawa Maraja tidak dapat diprediksi. Karena banyak pasien yang mengeluh tentang penyakit yang berbeda-beda. Kondisi tersebut mengakibatkan karyawan di Puskesmas Jawa Maraja kesulitan untuk membuat laporan data penyakit pasien. Disini menggunakan metode Clustering untuk mengelompokan penyakit pasien yang sering diperiksa oleh karyawan Puskesmas Jawa Maraja. Data penyakit pasien yang masih ditulis secara manual mempersulit karyawan yang bekerja di puskesmas. Dan juga pegawai kesulitan untuk menghitung persentase penyakit pada pasien di Puskesmas Jawa Maraja karena dihitung secara manual. Dengan adanya pengelompokan penyakit pada pasien di Puskesmas Jawa Maraja berbasis desktop akan membantu karyawan yang bertanggung jawab untuk membuat laporan penyakit pasien menjadi lebih mudah.

Data mining dapat diartikan sebagai proses penambangan data yang menghasilkan sebuah *output* (keluaran) berupa pengetahuan[3]. *Cluster* adalah sekumpulan objek-objek data yang mirip satu sama lain. Objek akan dikelompokkan kedalam satu atau lebih *cluster* sehingga objek-objek yang berada dalam satu *cluster* akan mempunyai kesamaan yang tinggi antara satu sama lain.[7]

Pengertian dari K-Means Clustering adalah, *K* dimaksudkan sebagai konstanta jumlah *cluster* yang diinginkan, *Means* dalam hal ini berarti nilai suatu rata-rata dari suatu grup data yang dalam hal ini didefenisikan sebagai *cluster*. Jadi, *K-Means* merupakan algoritma *clustering* yang berulang-ulang.[8] Algoritma *K-Means* menetapkan nilai-nilai *cluster* (*K*) secara random, untuk sementara nilai tersebut menjadi pusat dari *cluster* atau biasa disebut dengan *centroid*, *mean*, atau "*means*". Kemudian menghitung jarak

setiap data yang ada terhadap masing-masing *centroid* menggunakan rumus *Euclidian* hingga ditemukan jarak yang paling dekat dari setiap data dengan *centroid*. Klasifikasikan setiap data berdasarkan kedekatannya dengan *centroid*[9].

2. METODE PENELITIAN

2.1 Data Mining

Data mining adalah proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan teknik atau metode tertentu. Teknik, metode atau algoritma dalam *data mining* sangat bervariasi. Disini, penulis memilih menggunakan metode *Clustering* dengan algoritma *K-Means*.

2.2 Metode Clustering

Menurut Widodo, *Clustering* atau klasifikasi adalah metode yang digunakan untuk membagi rangkaian data menjadi beberapa group berdasarkan kesamaan yang telah ditentukan sebelumnya. *Cluster* adalah sekumpulan objek-objek data yang mirip satu sama lain. Objek akan dikelompokkan kedalam satu atau lebih *cluster* sehingga objek-objek yang berada dalam satu *cluster* akan mempunyai kesamaan yang tinggi antara satu sama lain.[7]

Analisis cluster merupakan salah satu teknik data mining yang bertujuan untuk mengidentifikasi sekelompok objek yang mempunyai kemiripan karakteristik tertentu yang dapat dipisahkan dengan kelompok objek lainnya, sehingga objek yang berada dalam kelompok yang sama relatif lebih homogeni daripada objek yang berada pada kelompok berbeda[3].

2.3 Algoritma K-Means

Pengertian dari *K-Means Clustering* adalah, *K* dimaksudkan sebagai konstanta jumlah *cluster* yang diinginkan, *Means* dalam hal ini berarti nilai suatu rata-rata dari suatu grup data yang dalam hal ini didefinisikan sebagai *cluster*. Jadi, *K-Means* merupakan algoritma *clustering* yang berulang-ulang.[8] Algoritma *K-Means* menetapkan nilai-nilai *cluster* (*K*) secara random, untuk sementara nilai tersebut menjadi pusat dari *cluster* atau biasa disebut dengan *centroid*, *mean*, atau "*means*". Kemudian menghitung jarak setiap data yang ada terhadap masing-masing *centroid* menggunakan rumus *Euclidian* hingga ditemukan jarak yang paling dekat dari setiap data dengan *centroid*. Klasifikasikan setiap data berdasarkan kedekatannya dengan *centroid*[9].

Berikut adalah rumus *Algoritma K-Means* :

1. Tentukan k sebagai jumlah cluster yang akan dibentuk
2. Tentukan k Centroid (titik pusat cluster) awal secara random/acak.

$$v = \frac{\sum_{i=1}^n x_i}{n} ; i = 1, 2, 3, , \dots, n \dots \dots \dots (2.1)$$

Dimana :

v : centroid pada cluster

x_1 : objek ke-i

n : banyaknya jumlah objek yang menjadi anggota *cluster*

3. Hitung jarak setiap objek ke masing-masing *centroid* dari masing-masing *cluster*. Untuk menghitung jarak antara objek dengan *centroid* dapat menggunakan *Euclidian Distance*.
- 4.

$$d(x, y) = |x - y| = \sqrt{\sum_{i=1}^n (x^i - y^i)^2} ; i = 1, 2, 3, \dots, n \dots \dots \dots (2.2)$$

Dimana :

x_i : objek x ke-i

y_i : daya y ke-i

n : banyaknya objek

5. Alokasikan masing-masing objek kedalam *centroid* yang paling dekat.
6. Lakukan literasi, kemudian tentukan posisi *centroid* baru dengan menggunakan persamaan.

3. ANALISA DAN HASIL

3.1 Algoritma Sistem

Algoritma merupakan langkah-langkah sistematis yang dijalankan untuk menyelesaikan suatu permasalahan tertentu. Seperti masalah dalam mengelompokkan penyakit yang sering diperiksa oleh pasien di Puskesmas Jawa Maraja, Kab.Simalungun dengan menggunakan algoritma *K-Means Clustering*.

3.2 Perhitungan Metode *Clustering Menggunakan Algoritma K-Means*

Menghitung jarak antara variabel dari setiap sampel data dengan *centroid* :

- a. Jarak antara penyakit CHO dengan titik m1

$$\begin{aligned} &= \sqrt{\sum_{i=1}^n (xi - yi)^2} \\ &= \sqrt{(5 - 4)^2 + (4 - 2)^2 + (2 - 5)^2} \\ &= 3.74 \end{aligned}$$

Jarak antara penyakit CHO dengan titik m2

$$\begin{aligned} &= \sqrt{\sum_{i=1}^n (xi - yi)^2} \\ &= \sqrt{(5 - 5)^2 + (4 - 4)^2 + (2 - 4)^2} \\ &= 2 \end{aligned}$$

Jarak antara penyakit CHO dengan titik m3

$$\begin{aligned} &= \sqrt{\sum_{i=1}^n (xi - yi)^2} \\ &= \sqrt{(5 - 4)^2 + (4 - 5)^2 + (2 - 5)^2} \\ &= 3.32 \end{aligned}$$

- b. Jarak antara penyakit AU dengan titik m1

$$\begin{aligned} &= \sqrt{\sum_{i=1}^n (xi - yi)^2} \\ &= \sqrt{(2 - 4)^2 + (5 - 2)^2 + (3 - 5)^2} \\ &= 4.12 \end{aligned}$$

Jarak antara penyakit AU dengan titik m2

$$\begin{aligned} &= \sqrt{\sum_{i=1}^n (xi - yi)^2} \\ &= \sqrt{(2 - 5)^2 + (5 - 4)^2 + (3 - 4)^2} \\ &= 3.32 \end{aligned}$$

Jarak antara penyakit AU dengan titik m3

$$\begin{aligned} &= \sqrt{\sum_{i=1}^n (xi - yi)^2} \\ &= \sqrt{(2 - 4)^2 + (5 - 5)^2 + (3 - 5)^2} \\ &= 2.83 \end{aligned}$$

1. Hitung Nilai *WCV (Within Cluster Variation)* dengan cara memangkatkan jarak terdekat *Cluster* dan menjumlahkan setiap nilai *WCV*

$$WCV = 1.41^2 + 0^2 + 1^2 + 2^2 + 0^2 + \dots + 1.41^2$$

$$WCV = 43.00$$

2. Hitung Nilai *BCV (Between Cluster Variation)* dengan cara menjumlahkan hasil dari jarak diantara setiap *centroid*.

$$\begin{aligned} \text{a. } d(m1, m2) &= \sqrt{(m1 - m2)^2} \\ &= \sqrt{(4 - 5)^2 + (2 - 4)^2 + (5 - 4)^2} \\ &= 2.45 \end{aligned}$$

$$\begin{aligned} \text{b. } d(m1, m3) &= \sqrt{(m1 - m3)^2} \\ &= \sqrt{(4 - 4)^2 + (2 - 5)^2 + (5 - 5)^2} \\ &= 3 \end{aligned}$$

$$\begin{aligned} \text{c. } d(m2, m3) &= \sqrt{(m2 - m3)^2} \\ &= \sqrt{(5 - 4)^2 + (4 - 5)^2 + (4 - 5)^2} \\ &= 1.73 \end{aligned}$$

$$\text{Nilai BCV} = d(m1, m2) + (m1, m3) + (m2, m3)$$

$$= 2.45 + 3 + 1.73$$

$$= 7.18$$

3. Menghitung nilai besar rasio dengan membandingkan nilai BCV dan WCV

$$BCV/WCV = 7.18/43.00$$

$$= 0.16$$

Seluruh proses perhitungan seperti diatas dilanjutkan terhadap iterasi berikutnya, sampai menemukan nilai yg sama pada sebuah iterasi.

Setelah dilakukan semua perhitungan dengan centroid yg baru disetiap iterasi, maka ditentukan penyakit apa yg sering diperiksa oleh pegawai puskesmas Jawa Maraja.

4 PENGUJIAN DAN IMPLEMENTASI

4.1 Form Login

Berikut ini merupakan tampilan dari *Form* login yang berfungsi untuk melakukan proses validasi username dan password pengguna sebelum masuk kedalam Menu Utama:



Gambar 4.1 Tampilan *Form Login*

4.2 Form Menu Utama

Halaman ini berfungsi sebagai halaman awal program yaitu sebagai berikut :



Gambar 4.2 Tampilan *Form Menu Utama*

4.3 Form Data Penyakit

Halaman ini memiliki fungsi sebagai tempat untuk menginput data penyakit yaitu sebagai berikut :

No	Kode Penyakit	Nama Penyakit	Jenis Penyakit	Bulan September	Bulan Okt - Nov	Bulan Desember
1	CHO	CHOLESTROL	Akut	18	14	9
2	AU	ASAM URAT	Akut	7	28	8
3	DIA	DIABETES	Akut	12	14	25
4	DEM	DEMAM	Akut	34	17	43
5	TBC	MXKING 150	Akut	8	16	12
6	MUM	NEW R15 GP	Akut	13	26	27
7	HIP	NEW R15	Akut	12	20	15
8	ALE	NEW VIXION	Akut	11	18	15
9	ISPA	NEW VIXION R	Akut	8	71	23
10	DIAR	DIARE	Akut	116	155	155
11	INFLU	INFLUENZA	Akut	20	17	12
12	DYS	DYSPEPSIA	Akut	20	23	26

Gambar 4.3 Tampilan *Form* Data Penyakit

4.4 Form Data Gejala

Halaman ini memiliki fungsi sebagai tempat untuk menginput data titik pusat awal cluster yaitu sebagai berikut :

Kode Penyakit	Nama Penyakit	Jenis Penyakit	Bulan September	Bulan Okt - Nov	Bulan Desember
CHO	CHOLESTROL	Akut	5	4	2
AU	ASAM URAT	Akut	2	5	3
DIA	DIABETES	Akut	3	3	5
DEM	DEMAM	Akut	4	2	5
TBC	MXKING 150	Akut	2	5	4
MUM	NEW R15 GP	Akut	3	5	5
HIP	NEW R15	Akut	3	5	4
ALE	NEW VIXION	Akut	3	5	4
ISPA	NEW VIXION R	Akut	1	5	3
DIAR	DIARE	Akut	4	5	5
INFLU	INFLUENZA	Akut	5	4	4
DYS	DYSPEPSIA	Akut	4	5	5
PK	PENYAKIT KULIT	Akut	4	5	3

Gambar 4.4 Tampilan *Form* Data Gejala

4.5 Form Proses Clustering

Halaman ini memiliki fungsi sebagai tempat untuk memproses data yaitu sebagai berikut :

Gambar 4.5 Tampilan *Form* Proses Clustering

4.6 Form Laporan

Halaman ini memiliki fungsi sebagai tempat menampilkan laporan yaitu sebagai berikut :

No	Kode Penyakit	Nama Penyakit	M1	M2	M3	Hasil Cluster
1	DEM	DEMAM	0,75	3,14	3,36	1
2	DER	DERMATITIS	0,47	2,49	2,51	1
3	DIA	DIABETES	0,75	2,93	2,20	1
4	CHO	CHOLESTROL	3,54	1,10	3,22	2
5	INFLU	INFLUENZA	2,13	1,10	2,44	2
6	PK	PENYAKIT KULIT	3,09	0,95	1,65	2
7	ALE	NEW VIKON	2,62	2,06	0,25	3
8	AU	ASAM URAT	3,50	2,75	1,36	3
9	DIAR	DIARE	2,35	2,21	1,51	3
10	DYS	DYSPEPSIA	2,35	2,21	1,51	3
11	HIP	NEW R15	2,62	2,06	0,25	3

Gambar 4.6 Tampilan *Form* Laporan

5. KESIMPULAN

Berdasarkan analisa pada permasalahan yang terjadi dalam kasus yang diangkat tentang Penerapan Metode Clustering Dalam Menentukan Persentase Penyakit Pada Pasien Dengan Menggunakan Algoritma K-Means Pada Puskesmas Jawa Maraja, Kab. Simalungun, maka ditarik beberapa kesimpulan, sebagai berikut :

1. Penerapan Metode Clustering Menggunakan Algoritma K-Means dapat diterapkan dalam pemecahan masalah pada Puskesmas Jawa Maraja terkait menentukan persentase penyakit di Puskesmas Jawa Maraja.
2. Memodelkan Data Mining dapat dirancang dengan baik menggunakan UML, dan *Flowchart* untuk menggambarkan sistem yang akan dibangun. UML, terdiri dari *Usecase Diagram*, *Activity Diagram*, dan *Class Diagram*.
3. Data Mining yang dibangun menggunakan *Microsoft Visual Studio 2008*, *Microsoft Access 2007*, dan *Crystal Report 9.ink*.
4. Pengujian dan Pengimplementasian sistem yang dirancang untuk Puskesmas Jawa Maraja dalam menentukan persentasi penyakit dilakukan dengan cara menginput nama-nama penyakit yang telah

diperiksa oleh pegawai Puskesmas dan melakukan proses perhitungan dengan *Algoritma K-Means* kedalam sistem. Agar mengetahui apakah sistem yang dibangun telah sesuai atau belum.

5. Pengimplementasian sistem yang telah teruji merupakan upaya sistematis di Puskesmas Jawa Maraja Kab. Simalungun terkait penentuan persentase Penyakit.




UCAPAN TERIMA KASIH

Puji syukur kehadiran Allah SWT atas izin-Nya yang telah melimpahkan rahmat dan karunia-Nya sehingga dapat menyelesaikan jurnal ilmiah ini. Pada kesempatan ini diucapkan terima kasih yang sebesar-besarnya kepada kedua Orang Tua tercinta yang selama ini memberikan do'a dan dorongan baik secara moril maupun materi sehingga dapat terselesaikan pendidikan dari tingkat dasar sampai bangku perkuliahan dan terselesaikannya jurnal ini. Di dalam penyusunan jurnal ini, banyak sekali bimbingan yang didapatkan serta arahan dan bantuan dari pihak yang sangat mendukung. Oleh karena itu dengan segala kerendahan hati, diucapkan terima kasih yang sebesar-besarnya kepada Bapak Rudi Gunawan, SE., M.Si., selaku Ketua Sekolah Tinggi Manajemen InFormatika Dan Komputer (STMIK) Triguna Dharma Medan. Bapak Dr. Zulfian Azmi, ST., M.Kom., selaku Wakil Ketua I Bidang Akademik STMIK Triguna Dharma Medan. Bapak Marsono, S.Kom., M.Kom., selaku Ketua Program Studi Sistem Informasi STMIK Triguna Dharma Medan. Bapak Ahmad Fitri Boy, S.Kom., M.Kom., selaku Dosen Pembimbing I yang telah memberikan saran, arahan dan dukungannya serta motivasi, sehingga penelitian ini dapat terselesaikan dengan baik dan tepat waktu. Bapak Guntur Syahputra, S.Kom., M.Kom., selaku Dosen Pembimbing II yang telah memberikan bimbingan tata cara penulisan, saran dan motivasi sehingga penelitian ini dapat terselesaikan dengan baik dan tepat waktu. Seluruh Dosen, Staff dan Pegawai di STMIK Triguna Dharma Medan.

REFERENSI

- [1] Adzikra Ibrahim, "pengertian Penyakit Menurut Para Ahli." [Online]. Available: <https://pengertiandefinisi.com/pengertian-penyakit-menurut-para-ahli/>.
- [2] M. R. L. Iin Parlina, Agus Perdana Windarto, Anjar Wanto, "Memanfaatkan Algoritma K-Means Dalam Menentukan Pegawai Yang Layak Mengikuti Aseessment Center," *Memanfaatkan Algoritma K-Means Dalam Menentukan Pegawai Yang Layak Mengikuti Aseessment Cent. Untuk Clust. Progr. Sdp*, vol. 3, no. 1, pp. 87–93, 2018.
- [3] M. K. Dicky Nofriansyah, S.Kom. and M. S. DR. Ir. Gunadi Widi Nurcahyo, *Algoritma Data Mining Dan Pengujian*. 2015.
- [4] H. L. S. Benri Melpa Metisen, "Analisis Clustering Menggunakan Metode K-Means Dalam Pengelompokkan Penjualan Produk Pada Swalayan Fadhila," *Anal. Clust. Menggunakan Metod. K-Means Dalam Pengelompokkan Penjualan Prod. Pada Swalayan Fadhila*, 2015.
- [5] C. W. Randi Rian Putra, "IMPLEMENTASI DATA MINING PEMILIHAN PELANGGAN POTENSIAL MENGGUNAKAN ALGORITMA K-MEANS," 2018.
- [6] C. J. Ma. S. Fina Nasari, "Penerapan Algoritma K-Means Clustering Untuk Pengelompokkan Penyebaran Diare Di Kabupaten Langkat," 2016.
- [7] T. Khotimah, "Pengelompokan Surat Dalam Al Qur'an Menggunakan Algoritma K-Means," *Simetris J. Tek. Mesin, Elektro dan Ilmu Komput.*, vol. 5, no. 1, pp. 83–88, 2014, doi: 10.24176/simet.v5i1.141.
- [8] M. J. Nurul Rohmawati, Sofi Defiyanti, "IMPLEMENTASI ALGORITMA K-MEANS DALAM PENGKLASTERAN MAHASISWA PELAMAR BEASISWA," 2016.

BIOGRAFI PENULIS

	<p>Ughie Yulika, Perempuan kelahiran Baja Dolok, 11 Juli 1998, anak kedua dari empat bersaudara ini merupakan seorang mahasiswa STMIK Triguna Dharma yang sedang dalam proses menyelesaikan skripsi.</p>
	<p>Ahmad Fitri Boy, S.Kom., M. Kom, Beliau merupakan dosen STMIK Triguna Dharma Medan.</p>
	<p>Guntur Syahputra, S.Kom., M.Kom, Beliau merupakan dosen STMIK Triguna Dharma Medan.</p>