

**KLASTERISASI DATA REKAM MEDIS PADA DIAGNOSA PENYAKIT
BERDASARKAN USIA PASIEN MENGGUNAKAN ALGORITMA K-MEANS DI
PUSKESMAS LUBUK ALUNG**

HENDRA NUSA PUTRA¹, DINDA PUTRI ANISA²

Rekam Medis STIKES Dharma Landbouw Padang^{1,2}
nusahendra@gmail.com¹, ditrisa94@gmail.com²

Abstract : *Clustering is a process of grouping data sets into several groups. K-Mean is an innovation or new knowledge for puskesmas in calculating patient disease data. This study aims to determine the clustering of medical record data using the k-mean algorithm. This type of research uses descriptive quantitative data. The population of medical record data taken is the last 3 months of 2020. The type of data used is secondary data, data collection by observation and data analysis using the algorithm K-Mean. The result of the research is that there are 3 clusters determined. Among them are the Low Suffering Disease Cluster, Medium Suffered Disease and High Suffered Disease Cluster. Low suffering disease which is coded to cluster A there are 97 patients with a percentage (14%), high suffering disease is coded to cluster B there are 318 patients with a percentage (45%), and moderate illness is coded to cluster C there are 292 patients with a percentage (41%). Disease The most common at the Lubuk Alung Health Center are Diabetes, Schizophrenia, Hypertension, Stroke, Refractive Disorders, Rheumatism, Tinea, Gout, Epilepsy and Cataracts. From this pattern, the researcher compared with several sources, that the pattern of the disease is indeed suffered by a lot of people in the age range above 30 years. The K-Mean method can be a new innovation that is expected to make it easier to calculate patient disease data at the Lubuk Alung Health Center.*

Keywords: *Clustering, K-Means, Age, Disease, WEKA.*

Abstrak: Klasterisasi merupakan suatu proses pengelompokkan kumpulan data menjadi beberapa kelompok. K-Means merupakan inovasi atau pengetahuan baru bagi puskesmas dalam menghitung data penyakit pasien. Penelitian ini bertujuan untuk mengetahui klasterisasi data rekam medis menggunakan algoritma k-means. Jenis penelitian ini menggunakan data kuantitatif deskriptif. Populasi data rekam medis yang diambil 3 bulan terakhir tahun 2020. Jenis data yang digunakan adalah data sekunder, pengambilan data secara observasi dan analisa data menggunakan algoritma K-Means. Hasil penelitian yang didapatkan ialah klaster yang ditentukan sebanyak 3 klaster. Diantaranya Klaster Penyakit Diderita Rendah, Penyakit Diderita Sedang dan Penyakit Diderita Tinggi. Penyakit diderita rendah terdapat 97 Pasien dengan persentase (14%), penyakit diderita tinggi terdapat 318 Pasien dengan persentase (45%), dan penyakit diderita sedang terdapat 292 Pasien dengan persentase (41%) Penyakit terbanyak yang ada pada Puskesmas Lubuk Alung ialah Diabetes, Skizoprenia, Hipertensi, Stroke, Kelainan Refraksi, Rematik, Tinea, Asam Urat, Epilepsi dan Katarak. Dari pola tersebut peneliti membandingkan dengan beberapa narasumber, bahwa pola penyakit tersebut memang banyak diderita oleh seseorang pada rentang usia diatas 30 tahun. Metode klasterisasi dapat menjadi metode yang diharapkan dapat mempermudah dalam menghitung data penyakit pasien pada Puskesmas Lubuk Alung.

Kata kunci : Klasterisasi, K-Means, Usia, Penyakit, WEKA

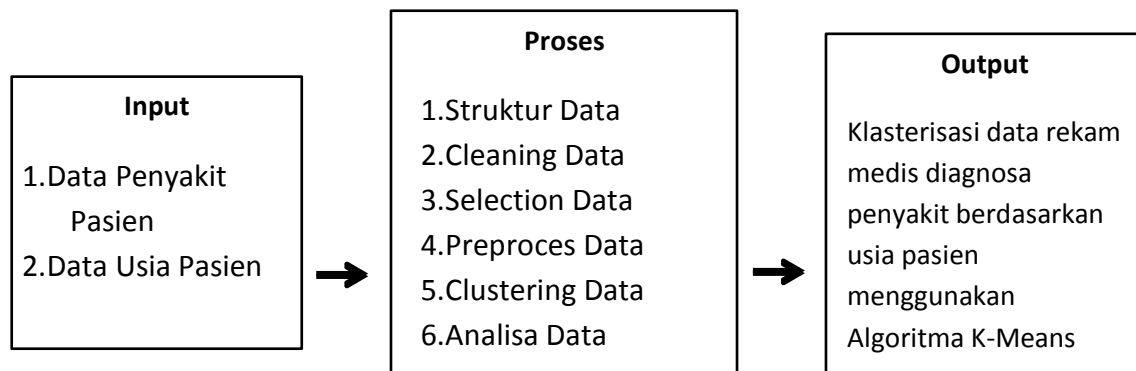
A. Pendahuluan

Rekam medis merupakan berkas yang berisi tentang identitas pasien, catatan pemeriksaan, catatan pengobatan dan catatan pelayanan kesehatan lainnya (Peraturan Menteri Kesehatan No.55 Tahun 2013). Menurut Putri dan Santoso (2016) data mining ialah suatu metode dimana untuk pengolahan data guna menemukan pola yang tersembunyi pada data yang sedang diolah. Dari data yang diolah dengan menggunakan teknik data mining akan ditemukan suatu informasi atau pengetahuan baru, sehingga dapat digunakan untuk pertimbangan dimasa yang akan datang.

Menurut penelitian yang dilakukan oleh Juninda dan Andri (2019) tentang Penerapan Algoritma K-Medoids untuk Pengelompokan Penyakit di Pekanbaru Riau, didapatkan hasil pada *Cluster 1* rentang usia 17 tahun s/d 25 tahun sebesar 44,39% menderita maag, *Cluster 2* rentang usia 17 tahun s/d 25 tahun sebesar 16,98% menderita diare dan sakit perut, *Cluster 3* rentang usia 17 tahun s/d 25 tahun sebesar 65,21% menderita batuk dan pilek, dan *Cluster 4* rentang usia 17 tahun s/d 25 tahun sebesar 54,10% menderita batuk dan pilek. Berdasarkan survei awal yang dilakukan pada Puskesmas Lubuk Alung, ditemukan pengelompokan penyakit belum terinci dikarenakan pencatatan masih dilakukan secara manual yaitu dicatat dalam buku besar. Tujuan penelitian ini untuk mengetahui pengelompokan penyakit dengan menggunakan metode data mining, sehingga dapat mengetahui dari informasi yang dihasilkan apa saja penyakit yang sering diderita oleh pasien yang berobat pada Puskesmas Lubuk Alung.

B. Metodologi Penelitian

Penelitian ini terdiri dari input, proses dan output. Dimana pada input proses memasukkan data penyakit dan usia pasien, selanjutnya pada proses yaitu mulai dari pengambilan data rekam medis, proses pembersihan data dimana data yang tidak lengkap tidak akan dimasukkan, seleksi data, preproses data yaitu memasukkan data ke dalam excel agar dapat memudahkan software dalam penghitungan data, clustering data dan analisa data. Dan proses yang terakhir ialah output, outputnya merupakan hasil akhir dari klasterisasi atau pengelompokan data rekam medis diagnosa berdasarkan usia pasien. Penelitian ini merupakan penelitian kuantitatif deskriptif dengan tujuan utamanya ialah mendeskripsikan dan menggambarkan tentang klasterisasi data rekam medis pada diagnosa penyakit berdasarkan usia pasien menggunakan metode algoritma k-means di Puskesmas Lubuk Alung Tahun 2021. Populasi yang diambil ialah 3 bulan terakhir tahun 2020 dengan data sebanyak 834 data setelah di bersihkan keseluruhan data yang dapat digunakan ialah sebanyak 712 data. Pengumpulan data menggunakan metode observasi, dimana data dikumpulkan secara langsung ke Puskesmas Lubuk Alung. Analisa data dilakukan dengan menggunakan software WEKA.



Gambar. 1 Alur Pikir

1.Input

Proses ini mengambil dan menginputkan data rekam medis. Diantaranya ada Nama, No. Rekam Medis, Alamat, Usia, Diagnosa Penyakit

2.Proses

- Struktur Data** - mengambil data Rekam Medis Puskesmas Lubuk Alung yang berisikan No.Rekam Medis, Nama, Alamat, Usia, Diagnosa Penyakit
- Cleaning Data** - proses ini akan dilakukan pembersihan data dimana informasi data yang tidak lengkap atau kurang jelas akan dibuang
- Selection Data** - proses ini akan diseleksi kembali jika ada data yang kurang lengkap ataupun tidak jelas informasinya
- Preproses Data** - proses ini menyimpan data ke dalam excel dengan format bentuk data csv. pada proses ini data akan di proses menggunakan aplikasi weka dengan

cara memasukkan data excel tadi ke dalam aplikasi weka kemudian, data akan disimpan dalam bentuk .arf agar dapat dibaca oleh aplikasi weka

e. **Clustering Data** - Pada proses ini mengelompokkan data berdasarkan karakteristik yang sama ke dalam satu cluster dengan menggunakan Algoritma K-Means dengan rumus *euclidean distance*.

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

f. **Analisa Data** - Pada proses ini hasil dari perhitungan dari pengelompokkan klaster

3. Output

Proses terakhir ialah merupakan hasil atau informasi dari penghitungan data dalam bentuk persentase.

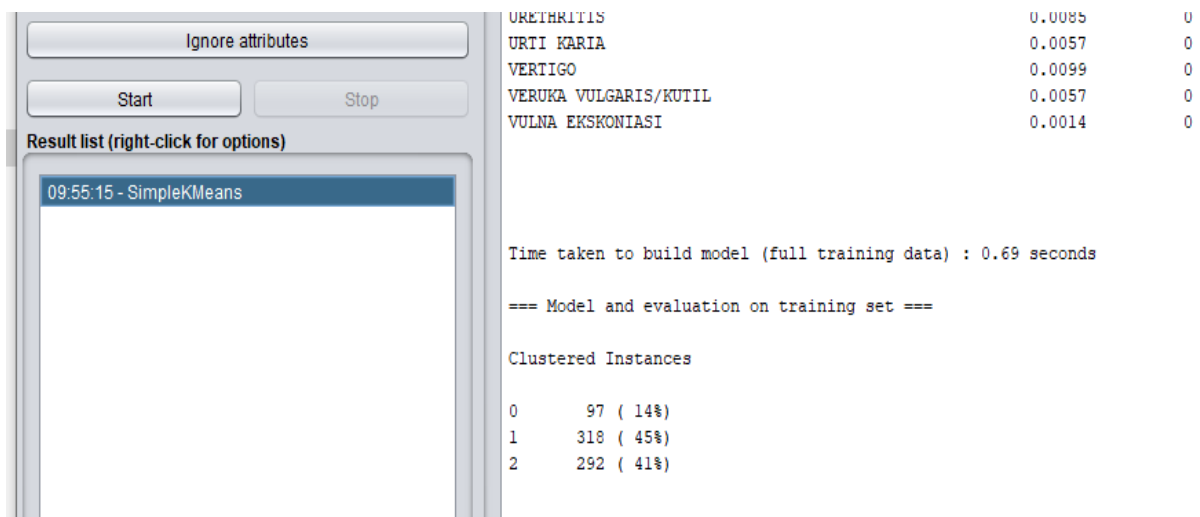
C. Hasil dan Pembahasan

Pengelompokkan usia menurut Departemen Kesehatan Republik Indonesia (2009) sebagai berikut :

Tabel 1. Konversi Usia

Kategori Usia	Usia (Tahun)	Konversi
Masa Anak – Anak	0 – 13	1
Masa Remaja	14 – 25	2
Masa Dewasa	26 – 45	3
Masa Lansia	46 – sampai atas	4

Kategori usia merupakan pengelompokkan usia berdasarkan usia pasien, yaitu ada masa anak – anak, masa remaja, masa dewasa, dan masa lansia, sehingga dalam pembacaan pada sistem perlu dilakukan konversi data menjadi angka, adapun untuk masa anak – anak menjadi angka (1), masa remaja menjadi angka (2), masa dewasa menjadi angka (3), dan masa lansia menjadi angka (4).



Gambar 2. Hasil klasterisasi menggunakan WEKA

Tabel 2. Hasil Pembagian Klaster

Nama Klaster Pada Aplikasi	Jumlah Pasien	Persentase	Dinamakan
Cluster 0	97 Pasien	14%	Klaster A
Cluster 1	318 Pasien	45%	Klaster B
Cluster 2	292 Pasien	41%	Klaster C



Gambar 3. Grafik Pembagian Klaster

Tabel 3. Klaster A

Usia	Jumlah Pasien	Penyakit
1 (0 – 13)	-	-
2 (14 – 25)	-	-
3 (26 – 45)	18	Diabetes
4 (46 – sampai atas)	79	Diabetes
Total	97 Pasien	

Pada tabel diatas dapat dilihat hasil dari penghitungan didominasi rentang usia diatas 30 tahun dengan persentase 14% menderita diabetes.

Tabel 4. Klaster B

Usia	Jumlah Pasien	Penyakit
1 (0 – 13)	17	Influenza
2 (14 – 25)	38	Infeksi Kulit
3 (26 – 45)	81	11 Pasien Epilepsi
4 (46 – sampai atas)	182	10 Pasien Katarak 12 Pasien Asam Urat 16 Pasien Kelainan Refraksi 19 Pasien Hipertensi
0		
Total	318 Pasien	

Pada tabel diatas dapat dilihat hasil dari penghitungan didominasi rentang usia diatas 30 tahun dengan persentase 45% menderita epilepsi, hipertensi, katarak, asam urat, dan kelainan refraksi

Tabel 5. Klaster C

Usia	Jumlah Pasien	Penyakit
1 (0 – 13)	21	Malnutrisi Otitis Media
2 (14 – 25)	45	Skizoprenia
3 (26 – 45)	78	29 Pasien Skizoprenia

4 (46 – sampaiatas)	148	10 Pasien Paru Obstruktif 13 Pasien Tinea 14 Pasien Rematik 16 Pasien Stroke
Total	292 Pasien	

Pada tabel diatas dapat dilihat hasil dari penghitungan menggunakan software WEKA didominasi rentang usia diatas 30 tahun dengan persentase 41% menderita skizoprenia, paru obstruktif, tinea, rematik dan stroke. Berdasarkan hasil penelitian diatas klaster dibagi menjadi 3 yaitu ada klaster penyakit diderita rendah, klaster penyakit diderita sedang dan klaster penyakit diderita tinggi. Alasan penamaan klaster dikarenakan pada saat penelitian didapatkan hasil bahwa usia tidak terartur dan adanya perbedaan penyakit pada setiap klaster. Solusi yang dapat dilakukan petugas ialah dengan mengadakan sosialisasi tentang pencegahan penyakit Diabetes, Hipertensi, Stroke, Epilepsi, Rematik, Penyakit Paru Obstruktif. Sehingga masyarakat mengetahui apa saja yang menjadi faktor resiko penyakit tersebut dan masyarakat dapat meminimalisir penyakit tersebut.

D. Penutup

Berdasarkan hasil pengujian yang telah dilakukan maka dapat ditarik kesimpulan sebagai berikut: 1) Penyakit Diderita Rendah pada Klaster A ialah Diabetes dengan persentase (14%); 2) Penyakit Diderita Sedang pada Klaster C ialah Skizoprenia, Paru Obstruktif, Tinea, Rematik, dan Stroke (41%); dan 3) Penyakit Diderita Tinggi pada Klaster B ialah Epilepsi, Katarak, Asam Urat, Kelainan Refraksi dan Hipertensi (45%)

Daftar Pustaka

- Ali, A. (2019). Klasterisasi Data Rekam Medis Pasien Menggunakan Metode K- Means Clustering di Rumah Sakit Anwar Medika Balong Bendo Sidoarjo. *MATRIK : Jurnal Manajemen, Teknik Informatika Dan Rekayasa Komputer*, 19(1), 186–195. <https://doi.org/10.30812/matrik.v19i1.529>
- Asroni, R. A. (2015). Penerapan Metode K-Means Untuk Clustering Mahasiswa Berdasarkan Nilai Akademik Dengan Weka Interface Studi Kasus Pada Jurusan Teknik Informatika UMM Magelang. *Ilmiah Semesta Teknika*, 18(1), 76–82.
- Iswandy, E., Komputer, D. S. T. M. I., & Padang, S. J. (2015). Sistem Penunjang Keputusan Untuk Menentukan Dan Santunan Sosial Anak Nagari Dan Penyaluran Bagi Mahasiswa Dan Pelajar Kurang Mampu. *Jurnal TEKNOIF*, 3(2). <https://doi.org/2338-2724>
- Juninda, T., & Andri, E. (2019). Penerapan Algoritma K-Medoids untuk Pengelompokan Penyakit di Pekanbaru Riau. November, 42–49.
- Pujiono, S., Amborowati, A., & Suyanto, M. (2013). Analisis Kepuasan Publik Menggunakan Weka Dalam Mewujudkan Good Governance Di Kota Yogyakarta. *Data Manajemen Dan Teknologi Informasi (DASI)*, 14(2), 45.
- Putri, D. L., & Santoso. (2016). Implementasi Algoritma K-Means Untuk Pengelompokan Penyakit Pasien (Studi Kasus : Puskesmas Kajen) K-Means Algorithm Implementation for Classification of Disease Patient (Case Study : Health Centers Kajen Regency Pekalongan). 1.
- Rahmayani, M. T. I. (2018). Analisis Clustering Tingkat Keparahan Penyakit Pasien Menggunakan Algoritma K-Means. *Jurnal Inovasi Teknik Informatika*, 1(2), 40–44.
- Sastika, W. (2018). Analisis Kualitas Layanan dengan Menggunakan E-service Quality untuk Mengetahui Kepuasan Pelanggan Belanja Online Shoppe (Studi Kasus: Pelanggan. *Journals.Upi-Yai.Ac.Id*, 2, 6.
- Silitonga Irene Sri, P. M. (2017). Klusterisasi Pola Penyebaran Penyakit Pasien Berdasarkan Usia Pasien Dengan Menggunakan K-Means Clustering. *Jurnal TIMES*, VI(Vol 6, No 2 (2017)), 22–25. <http://ejournal.stmik-time.ac.id/index.php/jurnalTIMES/article/view/584>

- Siregar, A. M. (2019). Penerapan Algoritma K-Means Untuk Pengelompokan Daerah Rawan Bencana Di Indonesia. *INTERNAL (Information System Journal)*, 1(2), 1–10.
<https://doi.org/10.32627/internal.v1i2.42>
- Wardani, N. W., Murni, N. N., Luka, S. S. P., & G.Indrawan. (2016). Analisis Penerapan K-means Untuk Pengelompokkan Diagnosa Penyakit Kulit dan Kelamin Berdasarkan Rentang Usia. Senapati, Senapati.
- Wardani, N. W., Murni, N. N., Luka, S. S. P., & G.Indrawan. (2016). Analisis Penerapan K-means Untuk Pengelompokkan Diagnosa Penyakit Kulit dan Kelamin Berdasarkan Rentang Usia. Senapati, Senapati.