

**PENERAPAN DATA MINING UNTUK PENGELOMPOKKAN DATA
REKAM MEDIS PASIEN DENGAN ALGORITMA *K-MEANS*
*CLUSTERING***

(Studi Kasus : RSUD M. Natsir Kota Solok)

SKRIPSI

Disusun Untuk Memenuhi Syarat Menyelesaikan Pendidikan Strata 1 (S1)

Departemen Teknik Elektronika

Program Studi Pendidikan Teknik Informatika



Oleh :

TEGAR WIBOWO

18076094 / 2018

**PROGRAM STUDI PENDIDIKAN TEKNIK INFORMATIKA
DEPARTEMEN TEKNIK ELEKTRONIKA
FAKULTAS TEKNIK
UNIVERSITAS NEGERI PADANG**

2023

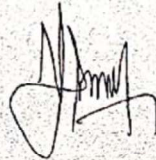
HALAMAN PERSETUJUAN TUGAS AKHIR

Penerapan Data Mining untuk Pengelompokan Data Rekam Medis Pasien dengan Algoritma K-Means Clustering (Studi Kasus RSUD M.Natsir Solok)

Nama : Tegar Wibowo
NIM / TM : 18076094 / 2018
Program Studi : Pendidikan Teknik Informatika
Departemen : Teknik Elektronika
Fakultas : Teknik

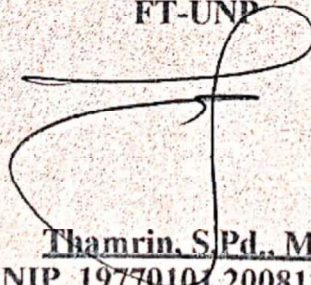
Padang, Oktober 2023

**Disetujui Oleh:
Pembimbing**



Dony Novaliendry, M.Kom
NIP. 19751104 200604 1 002

**Mengetahui,
Ketua Departemen Teknik Elektronika
FT-UNP**



Thamrin, S.Pd., M.T
NIP. 19770101 200812 1 001

HALAMAN PENGESAHAN TUGAS AKHIR

Dinyatakan lulus setelah dipertahankan di depan Tim Penguji Tugas Akhir
Program Studi Pendidikan Teknik Informatika
Departemen Teknik Elektronika, Fakultas Teknik
Universitas Negeri Padang

Judul : Penerapan Data Mining untuk Pengelompokkan Data Rekam
Medis Pasien dengan Algoritma K-Means Clustering (Studi Kasus
RSUD M.Natsir Solok)

Nama : Tegar Wibowo

NIM / TM : 18076094 / 2018

Program Studi : Pendidikan Teknik Informatika

Departemen : Teknik Elektronika

Fakultas : Teknik

Padang, Oktober 2023


Tim Penguji

Tanda Tangan

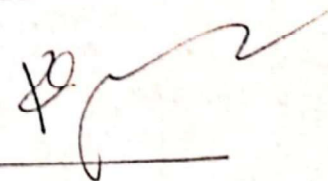
1. Ketua : Dr. Yeka Hendriyani, S.Kom., M.Kom

1. 

2. Anggota : Dony Novaliendry, M.Kom

2. 

3. Anggota : Khairi Budayawan S.Pd, M.Kom

3. 

SURAT PERNYATAAN

Saya yang bertanda tangan di bawah ini :

Nama : Tegar Wibowo
TM/NIM : 2018 / 18076094
Program Studi : Pendidikan Teknik Informatika
Departemen : Teknik Elektronika
Fakultas : Teknik

Dengan ini menyatakan bahwa Skripsi saya yang berjudul **“PENERAPAN DATA MINING UNTUK PENGELOMPOKAN DATA REKAM MEDIS PASIEN DENGAN ALGORITMA *K-MEANS CLUSTERING* (Studi Kasus : RSUD M. Natsir Kota Solok)”** adalah benar merupakan hasil karya saya sendiri. Tidak ada bagian di dalamnya yang merupakan karya atau pendapat yang ditulis atau diterbitkan orang lain kecuali sebagai acuan atau kutipan dengan mengikuti tata penulisan yang lazim. Demikian pernyataan ini saya buat dengan penuh kesadaran dan rasa tanggung jawab sebagai anggota masyarakat ilmiah.

Padang, Oktober 2023



Tegar Wibowo

18076094

ABSTRAK

**Tegar Wibowo : Penerapan Data Mining untuk Pengelompokan data
Rekam Medis Pasien dengan Algoritma *K-Means Clustering*
(Studi Kasus : RSUD M. Natsir Solok)**

RSUD M. Natsir Solok, Solok City menyediakan layanan kesehatan individu berupa kegiatan di dalam gedung meliputi pelayanan rawat inap dan rawat jalan dengan jam pelayanan 24 jam. Layanan rawat inap di termasuk layanan darurat dan layanan kesehatan dasar. Rekam medis adalah file yang berisi catatan dan dokumen mengenai identitas pasien, pemeriksaan, perawatan, tindakan dan layanan lain yang diberikan kepada pasien. Rekam medis harus dibuat secara tertulis, lengkap dan jelas atau secara elektronik. Masalah yang sering terjadi pada rekam medis adalah masalah tumpang tindih data rekam medis di rumah sakit. Salah satu cara untuk mengatasinya adalah dengan menggunakan Data Mining. metode yang digunakan adalah *Clustering*, *Clustering* dilakukan dengan menggunakan *K-Means* Algorithm. Pengelompokan data rekam medis pasien dari proses Data Mining di atas adalah untuk menghasilkan informasi baru tentang pola pengelompokan penyebaran penyakit di setiap desa. Setelah pengelompokan data rekam medis pasien menggunakan metode *K-Means* clustering terbentuk empat klaster, yaitu klaster pertama dengan total 562 item, *Cluster* dua dengan total 406 item, *Cluster* tiga dengan total 791 item, dan *Cluster* empat dengan total 279 item. Hal ini dapat dijadikan acuan bagi pemerintah daerah, khususnya Dinas Kesehatan Kota Solok, untuk melakukan sosialisasi dan pencegahan sumber penyakit berdasarkan diagnosis penyakit, usia dan jenis kelamin yang banyak diderita pasien.

Kata Kunci : Rekam Medis, Data Mining, Klustering, Algoritma *K-Means Clustering*

KATA PENGANTAR

Puji syukur senantiasa tercurahkan kepada Allah SWT, atas segala rahmat dan karunia-Nya, serta tidak lupa shalawat beriring salam semoga dilimpahkan kepada baginda Nabi Muhammad SAW, sehingga penulis dapat menyelesaikan Skripsi yang berjudul **“Penerapan Data Mining untuk Pengelompokan data Rekam Medis Pasien dengan Algoritma *K-Means Clustering* (Studi Kasus : RSUD M. Natsir Solok)”**.

Skripsi ini merupakan salah satu persyaratan dalam memperoleh gelar Sarjana Pendidikan pada Program Studi Pendidikan Teknik Informatika Jurusan Teknik Elektronika Fakultas Teknik Universitas Negeri Padang. Dalam penulisan skripsi ini, penulis banyak diberi motivasi, arahan bimbingan serta nasehat dari berbagai pihak.

Ungkapan terima kasih yang tulus dan tak terhingga, penulis ucapkan kepada :

1. Kedua orangtua tersayang, yang senantiasa mendukung dan memberikan motivasi baik secara moril maupun materil serta do'a yang sangat tulus demi penyelesaian skripsi ini.
2. Bapak Dony Novaliendry, M.Kom. selaku Pembimbing Skripsi dan Pembimbing Akademik yang telah memberikan nasehat dan motivasi selama perkuliahan dan membimbing dari awal penulisan sampai penyelesaian skripsi ini.
3. Ibu Yeka Hendriyani, S.Kom., M.Kom dan Bapak Khairi Budayawan, S.Pd., M.Kom. selaku Dosen Penguji yang telah

banyak memberikan saran dan masukan dalam penyelesaian skripsi ini.

4. Bapak Thamrin, S.Pd., M.T. selaku Ketua Jurusan Teknik Elektronika Fakultas Teknik Universitas Negeri Padang.
5. Bapak / Ibu Dosen Jurusan Teknik Elektronika Fakultas Teknik Universitas Negeri Padang yang telah banyak membantu penulis dalam perkuliahan sehingga dapat menyelesaikan studi dengan baik.
6. Staf Jurusan Teknik Elektronika yang telah membantu dalam kegiatan administrasi dalam penyelesaian skripsi ini.
7. Terima kasih kepada teman-teman f56 yang telah membantu dan memotivasi penulis dalam proses menyelesaikan skripsi ini.
8. Rekan – rekan jurusan mahasiswa Prodi Pendidikan Teknik Informatika 2018 yang telah membantu dan memotivasi penulis dalam menyelesaikan skripsi ini.
9. Serta semua pihak yang tidak dapat disebutkan satu persatu yang telah membantu penulis dalam penyusunan skripsi ini.

Penulis menyadari bahwa dalam penulisan skripsi ini masih terdapat banyak kekurangan dan kelemahannya. Untuk itu penulis sangat mengharapkan masukan berupa kritikan dan saran yang bersifat membangun. Penulis berharap semoga skripsi ini dapat bermanfaat bagi pembaca dan semua pihak yang berkepentingan dengan skripsi ini.

Akhir kata hanya kepada Allah SWT tempat berserah diri, semoga bimbingan dan bantuan yang telah diberikan kepada penulis mendapatkan balasan yang setimpal dari-Nya. Aamiin Yaa Robbal'alamiin.

Padang, Mei 2023
Penulis,

Tegar Wibowo

DAFTAR ISI

Halaman

HALAMAN PERSETUJUAN	
HALAMAN PENGESAHAN	
HALAMAN PERNYATAAN	
ABSTRAK	
KATA PENGANTAR	i
DAFTAR ISI.....	iv
DAFTAR TABEL.....	vi
DAFTAR GAMBAR.....	vii
BAB I PENDAHULUAN	1
A. Latar Belakang	1
B. Identifikasi Masalah.....	5
C. Batasan Masalah.....	6
D. Rumusan Masalah	6
E. Tujuan Penelitian.....	6
F. Manfaat Penelitian.....	7
BAB II LANDASAN TEORI.....	8
A. Data Mining	8
B. Teknik <i>Clustering</i> Data Mining	15
C. Algoritma K- Means <i>Clustering</i>	17
D. Rekam Medis	19

E. <i>RapidMiner</i>	21
F. Penelitian Terdahulu	22
BAB III METODOLOGI PENELITIAN.....	25
A. Jenis Penelitian.....	25
B. Desain Penelitian.....	26
C. Populasi dan Sampel Penelitian	29
D. Teknik Pengumpulan Data	30
E. Teknik Analisis Data.....	30
BAB IV HASIL PENELITIAN DAN PEMBAHASAN	34
A. Analisis Klustering Data Rekam medis dengan Algoritma <i>K-Means</i> <i>Clustering</i>	34
B. Klustering dengan <i>K-Means Clustering</i>	37
C. Pengujian dengan tool <i>RapidMiner</i>	47
D. Pengujian.....	47
E. Hasil	49
BAB V KESIMPULAN DAN SARAN.....	55
A. Kesimpulan.....	55
B. Saran.....	56
DAFTAR PUSTAKA	57
LAMPIRAN	

DAFTAR TABEL

	Halaman
Tabel 1. Penelitian Terdahulu.....	22
Tabel 2. Data set.....	34
Tabel 3. Tabel Inisialisasi Kode ICD-10.....	36
Tabel 4. Tabel Inisialisasi Umur.....	36
Tabel 5. Tabel Inisialisasi Kelurahan	36
Tabel 6. Hasil Perhitungan Jarak Iterasi Pertama.....	39
Tabel 7. Hasil Pengelompokkan Data Iterasi Pertama	40
Tabel 8. Hasil Perhitungan Iterasi Kedua.....	43
Tabel 9. Hasil <i>Cluster</i> Jarak Iterasi Kedua.....	44
Tabel 10. Hasil <i>Cluster</i> Pertama.....	45
Tabel 11. Hasil <i>Cluster</i> Kedua	45
Tabel 12. Hasil <i>Cluster</i> Ketiga	45
Tabel 13. Hasil <i>Cluster</i> Keempat	46
Tabel 14. Nilai DBI.....	54

DAFTAR GAMBAR

	Halaman
Gambar 1. Diagram Venn Data Mining (Sumber. https://algorit.ma)	9
Gambar 2. Flowchart K-Means Clustering.....	19
Gambar 3. Diagram Alur Penelitian	26
Gambar 4. Proses Analisis Data	30
Gambar 5. Rekap Data Diagnosa Penyakit	47
Gambar 6. Tampilan Awal <i>RapidMiner</i>	47
Gambar 7. Import Data.....	48
Gambar 8. Proses Data	49
Gambar 9. Hasil Cluster Model.....	49
Gambar 10. Folder View pad Cluster 1 dan 2.....	50
Gambar 11. Folder View pad Cluster 3 dan 4.....	51
Gambar 12. Data View Hasil permodelan Cluster	52
Gambar 13. Plot View	53
Gambar 14. Proses Hitung nilai DBI.....	53

BAB I

PENDAHULUAN

A. Latar Belakang

Kesehatan merupakan hal yang berharga bagi manusia karena siapa saja dapat mengalami gangguan kesehatan, begitu pula pada manusia yang sangat rentan terhadap berbagai macam penyakit namun penyebabnya tidak kita sadari. Berdasarkan karakteristiknya penyakit dapat digolongkan menjadi dua, yaitu penyakit menular dan penyakit tidak menular. Penyakit menular mendapatkan perhatian yang lebih dari pemerintah dibanding dengan penyakit tidak menular. Penyakit menular adalah penyakit yang disebabkan oleh bakteri, virus, atau parasit yang dapat ditularkan melalui media tertentu. Penyakit menular sering disebut juga penyakit infeksi, karena penyakit ini diderita melalui infeksi virus, bakteri, atau parasit yang ditularkan melalui berbagai macam media, seperti udara, jarum suntik, tranfusi darah, tempat makan atau minum, dan lain sebagainya.

Sedangkan, penyakit tidak menular (PTM) merupakan penyakit yang tidak ditularkan dan tidak ditransmisikan kepada orang lain dengan bentuk kontak apapun. Penyakit tidak menular, khususnya penyakit kardiovaskuler, kanker, penyakit pernapasan kronis, dan diabetes merupakan ancaman utama bagi kesehatan dan perkembangan manusia saat ini. WHO memprediksikan total kematian yang disebabkan oleh penyakit tidak menular akan meningkat sampai 17% dalam 10 tahun. Penyakit tidak menular berhubungan dengan genetic, lingkungan, dan yang paling penting adalah gaya hidup seperti merokok, konsumsi alkohol, pola diet yang buruk,

dan kurangnya aktifitas. Dalam mengatasi hal tersebut, Pemerintah nasional, organisasi multilateral seperti WHO dan LSM nasional dan internasional bekerjasama mengembangkan strategi umum yang berfokus pada pencegahan dan pengendalian penyakit kardiovaskuler, pencegahan dan pengendalian kanker, pencegahan dan pengendalian penyakit pernapasan kronis, dan kontrol diabetes mellitus.

Orang yang sedang menderita sakit biasanya akan berusaha untuk mengatasi dan mengobati penyakit yang dideritanya hingga dia sembuh. Tubuh yang sehat merupakan kehendak setiap orang. Untuk dapat mewujudkan keadaan sehat tersebut banyak hal yang perlu dilakukan. Salah satu di antaranya yang dinilai mempunyai peranan yang cukup penting adalah menyelenggarakan pelayanan kesehatan seperti rumah sakit (Adisasmito, 2009).

RSUD Mohammad Natsir adalah rumah sakit milik Provinsi Sumatera Barat yang terletak di wilayah administratif Kota Solok. Posisi RSUD Mohammad Natsir berada di kelurahan Simpang Rumbio Kecamatan Lubuk Sikarah Kota Solok. Posisi ini terletak di area strategis dengan dikelilingi wilayah kota dan kabupaten sekitarnya. Sebelah utara, berbatasan dengan kabupaten Solok dan Tanah Datar. Sebelah selatan dengan kabupaten Solok dan Solok Selatan. Sebelah barat dengan kabupaten Solok, kota Sawah Lunto, kabupaten Sijunjung dan Dharmasraya. Dengan letak yang strategis ini RSUD Mohammad Natsir banyak melayani pasien dari luar kota Solok seperti kabupaten Solok, kabupaten Sijunjung, kabupaten Dharmasraya kota Sawah Lunto dan

daerah lainnya.

RSUD M. Natsir Solok Kota Solok memberikan pelayanan kesehatan perorangan dalam bentuk kegiatan di dalam gedung meliputi pelayanan rawat inap dan rawat jalan dengan jam pelayanan 24 jam. Pelayanan rawat inap meliputi diantaranya pelayanan gawat darurat dan pelayanan kesehatan dasar. Pelayanan kesehatan masyarakat dalam bentuk kegiatan di dalam dan luar gedung. Pelayanan kesehatan masyarakat tersebut antara lain promosi kesehatan, pemberantasan penyakit, penyehatan lingkungan, perbaikan gizi, peningkatan kesehatan keluarga, keluarga berencana, kesehatan jiwa serta berbagai program kesehatan masyarakat lainnya. Semua catatan dan dokumen mengenai penyakit tersebut tersimpan di rekam medis

Rekam medis adalah berkas yang berisikan catatan dan dokumen tentang identitas pasien, pemeriksaan, pengobatan, tindakan dan pelayanan lain yang diberikan kepada pasien. Rekam medis harus dibuat secara tertulis, lengkap, dan jelas atau secara elektronik. Penyelenggaraan rekam medis dengan menggunakan teknologi informasi elektronik diatur oleh peraturan tersendiri. Informasi dalam rekam medis dijaga kerahasiaannya oleh dokter, tenaga kesehatan dan petugas pengelola serta pimpinan sarana pelayanan kesehatan. Data rekam medis terus terakumulasi setiap hari seiring dengan aktivitas rumah sakit. Pemanfaatan rekam medis dapat dipakai sebagai: (1) pemeliharaan kesehatan dan pengobatan pasien; (2) alat bukti dalam proses penegakkan hukum, disiplin kedokteran dan kedokteran gigi, dan penegakkan etika kedokteran dan kedokteran gigi; (3) keperluan

pendidikan dan penelitian; (4) dasar pembayaran biaya pelayanan kesehatan; (5) data statistik kesehatan. Permasalahan yang sering terjadi di rekam medis yaitu permasalahan penumpukan data rekam medis di rumah sakit.

Salah satu cara yang digunakan untuk mengatasi hal tersebut yaitu dengan menggunakan Data Mining. Perkembangan Ilmu Data Mining memberikan inovasi baru dalam hal pendayagunaan kumpulan data yang banyak sehingga dapat bermanfaat bagi pengembangan pengetahuan, baik secara khusus pada bidang yang berkaitan dengan data tersebut maupun secara global. Banyak fungsi yang dapat diterapkan dari ilmu Data Mining antara lain, estimasi, prediksi, klusterisasi, klasifikasi dan asosiasi. Untuk mencapai fungsi-fungsi tersebut dilakukan dengan berbagai metode (algoritma) seperti regresi untuk estimasi, Support Vector Machine (SVM) untuk prediksi, *K-Means* untuk klusterisasi, C4.5 untuk klasifikasi, apriori untuk asosiasi.

Berdasarkan permasalahan yang telah dijelaskan diatas, maka dibuat penelitian untuk klusterisasi data rekam medis. Dalam hal ini menggunakan salah satu metode Data Mining, yaitu algoritma *K-Means* dalam fungsi Klusterisasi/*Clustering*.

Metode *K-Means Clustering* memiliki konsep pengelompokan berdasarkan ukuran kedekatan dan kemudian menggunakan karakteristik tersebut sebagai vektor karakteristik atau centroid, sehingga Metode *KMeans Clustering* bisa menjadi solusi untuk pengklasifikasian karakteristik dari objek dan memiliki ketelitian yang cukup tinggi terhadap

ukuran objek. *Clustering* yang merupakan metode pengelompokan data. Tujuan dari pengelompokan *Cluster* ini adalah untuk menemukan pengelompokan dari serangkaian pola, titik, objek maupun dokumen. Objek yang berada didalam pengelompokan *Cluster* yang sama memiliki kemiripan antar satu kelompok dan memiliki perbedaan dengan objek oleh kelompok *Cluster* lain. *K-Means* adalah sebuah metode *Clustering* yang termasuk dalam pendekatan *partitioning*. Algoritma *K-Means* merupakan model *Centroid*. Model *Centroid* adalah model yang menggunakan *Centroid* untuk membuat *Cluster*. *Centroid* adalah titik tengah suatu *Cluster*. *Centroid* berupa *Centroid* nilai. *Centroid* digunakan untuk menghitung jarak suatu objek data terhadap *Centroid*. Suatu objek data termasuk dalam *Cluster* jika memiliki jarak terpendek terhadap *Centroid Cluster* tersebut. Algoritma *K-Means* dapat diartikan sebagai algoritma pembelajaran yang sederhana untuk memecahkan suatu permasalahan pengelompokan yang bertujuan untuk meminimalkan kesalahan ganda.

Maka dari itu penulis tertarik ingin mengetahui bagaimana pola atau aturan yang dapat digunakan untuk mengelompokan data rekam medis pasien menggunakan algoritma *K-Means Clustering*

B. Identifikasi Masalah

Berdasarkan latar belakang masalah yang ditemukan, dapat diidentifikasi masalah sebagai berikut :

1. Data Rekam Medis Pasien di RSUD M. Natsir Solok masih dijadikan sebagai arsip dan belum dilakukan analisis sepenuhnya.
2. Belum di ketahui apa penyakit endemic di masing masing kelurahan di Kota Solok

3. Pengelompokkan bertujuan untuk nantinya memudahkan Pemerintah atau Dinas Kesehatan setempat dalam memberikan penyuluhan mengenai penyakit.

C. Batasan Masalah

Adapun batasan masalah dalam penelitian ini adalah sebagai berikut :

1. Ruang lingkup penelitian dibatasi pada data rekam medis pasien di RSUD M. Natsir Solok.
2. Data rekam medis pasien yang dipakai berupa data pasien yang rawat inap di RSUD M. Natsir Kota Solok
3. Data yang digunakan yaitu data Rekam Medis Pasien RSUD M. Natsir Solok tahun 2020
4. Metode yang digunakan yaitu Metode Klustering dengan Algoritma *K-Means Clustering*
5. Program analisis yang digunakan adalah aplikasi data *RapidMiner*.

D. Rumusan Masalah

Adapun rumusan masalah yang menjadi pokok penelitian adalah Bagaimana Menggelompokkan data rekam medis pasien di RSUD M. Natsir Solok menggunakan Algoritma *K-Means Clustering*.

E. Tujuan Penelitian

Dari rumusan masalah yang telah diuraikan di atas, ada beberapatujuan dari penelitian ini, yaitu:

1. Menerapkan metode data mining dengan algoritma *K-Means Clustering* ke dalam analisis pola penyakit pasien.
2. Mengimplementasikan algoritma *K-Means Clustering* untuk melihat bagaimana pengelompokkan data rekam medis pasien.

3. Membuat model prediksi menggunakan *K-Means Clustering* untuk mengelompokkan data rekam medis pasien tahun 2020.
4. Menguji hasil analisis penerapan algoritma *K-Means Clustering* untuk mengelompokkan data rekam medis pasien.

F. Manfaat Penelitian

Hasil penelitian ini diharapkan dapat memberikan manfaat sebagai berikut:

1. Manfaat Teoritis

Penelitian ini diharapkan dapat memberikan sumbangan pemikiran yang cukup signifikan sebagai masukan pengetahuan atau literature ilmiah yang dapat dijadikan bahan kajian bagi para peneliti lain dalam melakukan penelitian lain yang sejenis.

2. Manfaat Praktis

- a. Sebagai sarana untuk menerapkan pengetahuan yang diperoleh selama menempuh studi, khususnya di dalam ilmu data *science*.
- b. Bagi Rumah sakit dan Dinas kesehatan setempat laporan Tugas Akhir ini dapat dijadikan sebagai sarana referensi tambahan mengenai permasalahan yang terkait dengan penulisan Skripsi ini.