

**PENERAPAN *DATA MINING* ANALISA PENYAKIT
MENULAR PADA MANUSIA
SKRIPSI**



**Oleh:
Susi Susanti Tampubolon
170210152**

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK DAN KOMPUTER
UNIVERSITAS PUTERA BATAM
2021**

PENERAPAN *DATA MINING* ANALISA PENYAKIT

MENULAR PADA MANUSIA

SKRIPSI

**Untuk memenuhi salah satu syarat
memperoleh gelar sarjana**



Oleh:

Susi Susanti Tampubolon

170210152

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK DAN KOMPUTER
UNIVERSITAS PUTERA BATAM
2021**

SURAT PERNYATAAN ORSINILITAS

Yang bertanda tangan di bawah ini Penulis:

Nama : Susi Susanti Tampubolon

NPM : 170210152

Fakultas : Teknik Dan Komputer

Program Studi : Teknik Informatika

Menyatakan bahwa “Skripsi” yang penulis buat dengan judul:

PENERAPAN *DATA MINING* ANALISA PENYAKIT MENULAR PADA MANUSIA.

Ini adalah karya sendiri dan bukan “duplikasi” dari karya orang lain. Sejauh yang penulis tahu, dalam teks skripsi ini tidak ada karya ilmiah atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang disebutkan dalam teks ini dan disebutkan dalam sumber dan referensi kutipan. Jika ternyata dalam naskah ini, dapat ditunjukkan bahwa ada elemen PLAGIASI, Penulis siap untuk menghentikan naskah ini dan judul Penulis dibatalkan dan diproses sesuai dengan hukum dan peraturan yang berlaku. Jadi pernyataan yang Penulis buat ini sebenarnya tanpa paksaan dari siapa pun.

Batam, 24 Juli 2021



Susi Susanti Tampubolon
170210152

PENERAPAN *DATA MINING* ANALISA PENYAKIT

MENULAR PADA MANUSIA

SKRIPSI

**Untuk memenuhi salah satu syarat
memperoleh gelar sarjana**

**Oleh:
Susanti Susanti Tampubolon
170210152**

**Telah disetujui Pembimbing pada tanggal
seperti tertera di bawah ini**

Batam, 24 Juli 2021

A handwritten signature in blue ink, appearing to read 'Koko Handoko', is centered within a white rectangular box.

**Koko Handoko, S.Kom., M.Kom.
Pembimbing**

ABSTRAK

Pemanfaatan *data mining* dalam teknologi kini semakin hari semakin beragam. Pada penelitian ini penulis membahas tentang penerapan *data mining* dibidang medis, yaitu analisa data penyakit menular pada manusia serta pemanfaatan data penyakit menular di UPT Puskesmas Sei Langkai. Penyakit menular pada manusia merupakan salah satu jenis penyakit yang memiliki jumlah data yang banyak dan bertumpuk karena pada dasarnya penyakit menular memiliki sebab akibat yang beragam sehingga pada penelitian ini penulis menerapkan *data mining* yang berguna untuk dapat mengelola data yang bertumpuk menjadi sebuah informasi yang bermanfaat. Tujuan dari penelitian ini adalah untuk mengidentifikasi kemudian menganalisis tingkat tertinggi sampai terendah dari 7 jenis data penyakit menular pada manusia dengan total 1.212 pasien ditahun 2019 dan 2020 dengan menggunakan metode *Algoritma K-Means clustering*. Dari data yang telah diolah sehingga memperoleh hasil bahwa penyakit ISPA dan COVID-19 memiliki jumlah data tertinggi diwilayah Tembesi dengan jumlah 419 pasien, penyakit kusta, DBD, dan campak memiliki jumlah data tertinggi di wilayah Sei Langkai dengan jumlah 43 pasien, sedangkan penyakit HIV dan TBC memiliki jumlah data tertinggi di desa Sei Pelunggut dengan jumlah 36 pasien. Kesimpulan dari penelitian ini adalah dengan menggunakan *Metode K-Means* dan pengujian aplikasi *RapidMiner* maka dapat mempermudah dalam pengolahan data dan memiliki nilai akhir yang akurat serta efektif digunakan dalam pengolahan data besar.

Kata Kunci : *Data mining*, Penyakit Menular, *K-Means*, *RapidMiner*

ABSTRACT

The use of data mining in technology is growing day by day. In this research, the author discusses the application of data mining in the medical field, that is data analysis of infectious diseases in humans and the use of infectious disease data in UPT Puskesmas Sei Langkai. Infectious diseases in humans are one type of disease that has a large amount of data and accumulates because basically infectious diseases have various causes and effects so that in this research the authors apply data mining which is useful to be able to manage stacked data into useful information. The purpose of this research was to identify and then analyze the highest to lowest levels of 7 types of data on infectious diseases in humans with a total of 1,212 patients in 2019 and 2020 using the K-Means clustering algorithm. From the data that has been processed get results that Acute Respiratory Infections and COVID-19 have the highest number of data in Tembesi village with 419 patients, leprosy, dengue fever, and measles have the highest number of data in Sei Langkai village with 43 patients, while HIV and TB has the highest number of data in Sei Pelunggut village with 36 patients. The conclusion of this research is using the K-Means method and testing the RapidMiner application, it can facilitate data processing and has an accurate final value and is effectively used in big data processing.

Keywords: Data mining, Infectious disease, K-Means , RapidMiner

KATA PENGANTAR

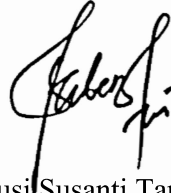
Puji dan Syukur penulis panjatkan kepada Tuhan Yang Maha Esa sehingga penulis dapat menyelesaikan Skripsi, yang menjadi salah satu persyaratan untuk menyelesaikan program studi sarjana (S1) dalam program Studi Teknik Informatika di Universitas Putera Batam.

Penulis menyadari bahwa Skripsi ini jauh dari sempurna. Karena itu kritik dan saran akan senantiasa penulis terima dengan senang hati. Dengan segala keterbatasan, dan dorongan dari berbagai pihak. Untuk itu, dengan segala kerendahan hati, penulis menyampaikan ucapan terima kasih kepada:

1. Rektor Universitas Putera Batam, Ibu Dr. Nur Elfi Husda, S.kom.,M.SI.
2. Dekan Fakultas Teknik dan Komputer, Bapak Welly Sugianto, S.T.,M.M.
3. Ketua Program Studi Teknik Informatika Bapak Andi Maslan,.S.T.,M.SI.
4. Bapak Koko Handoko,S.Kom., M.Kom. Sebagai Pembimbing Skripsi.
5. Dosen dan Staf Universitas Putera Batam.
6. Pihak Tata Usaha dan Tim Medis di Puskesmas Sei Langkai.
7. Orang Tua dan keluarga terkasih yang selalu berdoa dan memberi semangat kepada penulis sehingga dapat menyelesaikan skripsi ini.
8. Rekan kerja yang mau membagikan ilmunya dan berbagi pendapat dalam pembuatan skripsi ini
9. Rekan-Rekan mahasiswa Universitas Putera Batam yang juga memberikan doa dan dukungan mereka.
10. Teman-Teman Kampus yang selalu memberikan semangat dan dukungan kepada penulis dalam menyelesaikan skripsi ini.

11. Serta pihak yang tidak dapat disebutkan satu per satu oleh penulis, semoga senantiasanya diberkati Oleh Tuhan Yang Maha Esa, amin.

Batam, 24 Juli 2021

A handwritten signature in black ink, appearing to read 'Susi Susanti' with a stylized flourish at the end.

Susi Susanti Tampubolon

DAFTAR ISI

HALAMAN SAMPUL.....	i
HALAMAN JUDUL	ii
HALAMAN PERNYATAAN.....	iii
HALAMAN PENGESAHAN.....	iv
ABSTRAK	v
<i>ABSTRACT</i>	vi
KATA PENGANTAR.....	vii
DAFTAR ISI.....	ix
DAFTAR GAMBAR.....	xii
DAFTAR TABEL	xiii
DAFTAR RUMUS	xiv
BAB I.....	1
PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Identifikasi Masalah	2
1.3 Batasan Masalah	3
1.4 Rumusan Masalah	3
1.5 Tujuan Penelitian.....	4
1.6 Mafaat Penelitian.....	4
BAB II	6

KAJIAN PUSTAKA	6
2.1 Teori Dasar	6
2.2 <i>Knowledge Discovery In Database (KDD)</i>	6
2.3 <i>Data Mining</i>	8
2.4 <i>K-Means Clustering</i>	11
2.5 Software Pendukung.....	13
2.5.1 <i>RapidMiner</i>	13
2.5.2 <i>Tanagra</i>	14
2.6 Penelitian Terdahulu.....	15
2.7 Kerangka Pemikiran	19
BAB III.....	21
METODE PENELITIAN	21
3.1 Desain Penelitian	21
3.2 Teknik Pengumpulan Data	23
3.2.1 Observasi	23
3.2.2 Wawancara.....	23
3.2.3 Studi Pustaka.....	23
3.3 Operasional Variabel	24
3.3.1 Kusta	24
3.3.2 <i>Demam Berdarah Dengue (DBD)</i>	24
3.3.3 <i>HIV (Human Immunodeficiency Virus)</i>	25
3.3.4 <i>TBC (Tuberculosis)</i>	25
3.3.5 Campak/Rubeola.....	26
3.3.6 <i>ISPA (Infeksi Saluran Pernafasan Akut)</i>	26
3.3.7 <i>Coronavirus (COVID-19)</i>	26
3.4 Metode Perancangan Sistem.....	27

3.5 Lokasi dan Jadwal Penelitian	29
BAB IV	30
HASIL DAN PEMBAHASAN	30
4.1 Analisa Data	30
4.2 Penyeleksian Data	30
4.3 Analisa Proses Algoritma	31
4.4 Pengujian Data menggunakan Aplikasi <i>RapidMiner 5.3</i>	37
BAB V	41
KESIMPULAN DAN SARAN	41
5.1 Kesimpulan.....	41
5.2 Saran	42
DAFTAR PUSTAKA	43
Lampiran 1. Pendukung Penelitian.....	45
Lampiran 2. Daftar Riwayat Hidup	46
Lampiran 3. Surat Keterangan Penelitian.....	47

DAFTAR GAMBAR

Gambar 2. 1 Aplikasi <i>RapidMiner</i>	13
Gambar 2. 2 Aplikasi <i>Tanagra</i>	14
Gambar 4 3 Tampilan awal Aplikasi <i>RapidMiner</i>	37
Gambar 4 4 Tampilan Menu Operasional Data pada <i>RapidMiner</i>	37
Gambar 4 5 Tampilan menu processing import data.....	38
Gambar 4 6 Tampilan menu pilih file dari pc/laptop	38
Gambar 4 7 Tampilan Reload data	39
Gambar 4 8 Tampilan connecting data excel ke <i>K-Means</i> pada <i>RapidMiner</i>	39
Gambar 4 9 Tampilan Output pada <i>RapidMiner</i>	40
Gambar 4 10 Tampilan output <i>RapidMiner</i> dengan Centroid Plot View	40

DAFTAR TABEL

Tabel 4 1 Analisa data Penyakit menular di Puskesmas Sei Langkai	30
Tabel 4 2 Data Penyakit menular di Puskesmas Sei Langkai	31
Tabel 4 3 Nilai Rata-rata	33
Tabel 4 4 Hasil Literasi pertama	33
Tabel 4 5 Perhitungan data menggunakan Microsoft Excel.....	34
Tabel 4 6 Hasil Literasi Kedua	34
Tabel 4 7 Perbandingan hasil literasi pertama dan kedua	35
Tabel 4 8 Nilai Rata-rata dari Literasi Kedua	35
Tabel 4 9 Hasil Perhitungan data Literasi ketiga	36
Tabel 4 10 Perbandingan Hasil Literasi pertama, kedua,dan ketiga	36

DAFTAR RUMUS

Rumus 2 1 Euclidean Distance.....	12
Rumus 2 2 Nilai rata-rata	12
Rumus 4 1 Mencari jarak Euclidean	31
Rumus 4 2 Mencari nilai rata-rata.....	32