

**IMPLEMENTASI DATA MINING DENGAN
ALGORITMA *K – MEANS CLUSTERING*
PENYAKIT MENULAR**

(Di Rumah Sakit Zainal Umar Sidiki)

OLEH

RESKI LATANDA

T3117230

SKRIPSI

Untuk memenuhi salah satu syarat ujian

Guna memperoleh gelar Sarjana



**PROGRAM SARJANA
TEKNIK INFORMATIKA
UNIVERSITAS ICHSAN GORONTALO
GORONTALO
2021**

PENGESAHAN SKRIPSI

IMPLEMENTASI DATA MINING DENGAN ALGORITMA *K – MEANS CLUSTERING* PENYAKIT MENULAR

(Di Rumah Sakit Zainal Umar Sidiki)

OLEH

**RESKI LATANDA
T3117230**

SKRIPSI

Untuk memenuhi salah satu syarat ujian
Guna memperoleh gelar Sarjana
Program Studi Teknik Informatika
Ini telah disetujui oleh Tim Pembimbing

Gorontalo, 04 Juni 2021

Pembimbing Utama



Suhardi Rustam, S.Kom, M.Kom
NIDN : 0915088403

Pembimbing Pendamping



Sumarni S.Kom M.Kom
NIDN : 0926018604

PERSETUJUAN SKRIPSI

IMPLEMENTASI DATA MINING DENGAN ALGORITMA *K – MEANS CLUSTERING* PENYAKIT MENULAR

(Di Rumah Sakit Zainal Umar Sidiki)

OLEH

RESKI LATANDA
T3117230

Diperiksa oleh Panitia Ujian Strata satu (S1)

Universitas Ichsan Gorontalo

1. Ketua Penguji

Irvan Abraham Salihi, M. KOM

2. Anggota

Sudirman S. Panna, M.KOM

3. Anggota

Sarlis Mooduto, M. KOM

4. Anggota

Suhardi Rustam, M. KOM

5. Anggota

Sumarni, M. KOM



Mengetahui

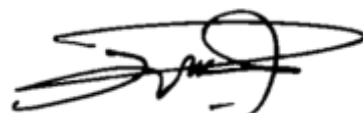
Dekan Fakultas Ilmu Komputer



Zohrahayaty, M. KOM

NIDN. 0912117702

Ketua Program Studi



Irvan Abraham Salihi, M. KOM

NIDN. 0928028101

PERNYATAAN SKRIPSI

Dengan ini saya menyatakan bahwa:

1. Karya tulis (Skripsi) saya ini adalah asli dan belum pernah diajukan untuk mendapatkan gelar akademik (Sarjana) baik di Universitas Ichsan Gorontalo maupun di perguruan tinggi lainnya.
2. Karya tulis (Skripsi) saya ini adalah murni gagasan, rumusan, dan penelitian saya sendiri tanpa bantuan pihak lain, kecuali arahan dari Tim Pembimbing.
3. Dalam karya tulis (Skripsi) saya ini tidak terdapat karya atau pendapat yang telah dipublikasikan orang lain, kecuali secara tertulis dicantumkan sebagai acuan dalam naskah dan dicantumkan pula dalam daftar pustaka.
4. Pernyataan ini saya buat dengan sesungguhnya dan apabila dikemudian hari terdapat penyimpangan dan ketitakbenaran dalam pernyataan ini, maka saya bersedia menerima sanksi akademik berupa pencabutan gelar yang telah diperoleh karya tulis ini, serta sanksi lainnya sesuai dengan norma-norma yang berlaku di Universitas Ichsan Gorontalo.

Kwandang, 04 Juni 2021

Yang Membuat Pernyataan,

Reski Latanda

ABSTRACT

RESKI LATANDA. T3117230. IMPLEMENTATION OF DATA MINING WITH THE K-MEANS CLUSTERING ALGORITHM AT ZAINAL UMAR SIDIKI HOSPITAL

The purposes of this study are 1) to design an infectious disease data processing system using the K-Means Clustering algorithm at Zainal Umar Sidiki Hospital, and 2) to apply the K-Means Clustering algorithm in determining infectious diseases so that it can be implemented. This study employs a qualitative approach with the methods used, namely collecting, analyzing data in the study, and then interpreting it. This study uses the Clustering technique in grouping one activity, namely the classification or grouping of data into several categories or clusters. The data or objects that are grouped into a group have the same characteristics based on certain criteria. In this study, several variables are used in the calculation of the K-Means Clustering algorithm are full name, address, gender, disease, mode of transmission, and hospital rates. From the results of the study, there are 3 clusters obtained. It means that this study is able to help the Hospital in classifying data on patients with infectious diseases.

Keywords: data mining, K-means clustering, infectious disease

ABSTRAK

RESKI LATANDA. T3117230. IMPLEMENTASI DATA MINING DENGAN ALGORITMA *K-MEANS CLUSTERING* PENYAKIT MENULAR DI RUMAH SAKIT ZAINAL UMAR SIDIKI

Tujuan penelitian ini 1) untuk merancang sistem pengolahan data penyakit menular menggunakan algoritma K-Means Clustering di Rumah Sakit Zainal Umar Sidiki. 2) untuk menerapkan algoritma K-Means Clustering dalam menentukan penyakit menular sehingga dapat di implementasikan. Dalam penelitian ini melalui pendekatan kualitatif dengan metode yang digunakan yaitu mengumpulkan, menganalisis data dalam riset kemudian ditafsirkan. Penelitian ini menggunakan teknik Clustering dalam mengelompokkan satu aktivitas yakni klasifikasi atau pengelompokan data ke dalam beberapa kategori atau cluster. Data atau objek yang dikelompokkan kedalam suatu group memiliki karakteristik yang sama berdasarkan kriteria tertentu. Dalam penelitian ini digunakan beberapa variabel dalam perhitungan algoritma K-Means Clustering yaitu nama lengkap, alamat, jenis kelamin, penyakit, cara penularan dan tarif rumah sakit. Dari hasil penelitian diperoleh 3 cluster, sehingga dapat dikatakan penelitian ini dapat membantu pihak Rumah Sakit dalam mengelompokkan data pasien penderita penyakit menular.

Kata kunci: data mining, K-means clustering, penyakit menular

KATA PENGANTAR

Puji syukur saya panjatkan kepada ALLAH S.W.T yang telah memberikan rahmat dan hidayah-nya sehingga penulis dapat menyelesaikan laporan skripsi dengan judul **“Implementasi Data Mining Dengan Algoritma *K-Means Clustering* Penyakit Menular di Rumah Sakit Zainal Umar Sidiki”** Dalam penelitian dan penyusunan skripsi banyak pihak yang telah membantu baik secara langsung ataupun tidak, maka dengan ini penulis mengucapkan banyak terimakasih kepada :

1. Bapak Mohammad Ichsan Gaffar, S.E, M.Ak selaku ketua yayasan pengembangan ilmu pengetahuan dan teknologi (YPIPT) ichsan gorontalo;
2. Bapak Dr. Abdul Gaffar Latjokke, M.Si, selaku Rektor Universitas Ichsan Gorontalo;
3. Ibu Zohrahayaty, M.Kom, selaku dekan Fakultas Ilmu Komputer;
4. Bapak Sudirman S.Panna, M.Kom, selaku wakil dekan I Bidang akademik;
5. Ibu Irma Surya Kumala Idris, M.Kom selaku wakil dekan II bidang administrasi umum dan keuangan;
6. Bapak Sudirman Melangi, M.Kom, selaku wakil dekan III bidang kemahasiswaan;
7. Bapak irvan A. Salihi, M.Kom, selaku ketua program studi Teknik Informatika Fakultas Ilmu Komputer;
8. Bapak Suhardi Rustam S.Kom, M.Kom selaku dosen pembimbing utama, yang juga telah membimbing dan memberikan saran pada pembuatan proposal ini.
9. Ibu Sumarni S.Kom, M.Kom selaku dosen pendamping, yang juga telah memberikan banyak motivasi kepada penulis.
10. Bapak dan ibu dosen yang telah mendidik dan membimbing penulis dalam mengerjakan proposal ini;
11. Kedua orang tua keluarga dan teman-teman yang telah memberikan do’a dorongan dan semangat selama penyusunan proposal ini.

12. Dan semua pihak yang tidak dapat disebut satu per satu yang telah membantu penulisan laporan proposal ini.

Semoga Allah, SWT melimpahkan balasan atas jasa- jasa mereka kepada kami. Penulis ini juga menyadari laporan ini banyak kekurangan, baik didalam penyusunan maupun penulisan.Oleh karena itu saran dan kritik sangat penulis harapkan demi kesempurnaan laporan berikutnya. Akhirnya penulis hanya berharap semoga laporan proposal ini dapat bermanfaat bagi kita semua. Aamin

Gorontalo, 01 juni 2021

Penulis

DAFTAR ISI

HALAMAN SAMPUL.....	
HALAMAN JUDUL	i
PENGESAHAN SKRIPSI.....	ii
PERSETUJUAN SKRIPSI.....	iii
PERNYATAAN SKRIPSI.....	iv
ABSTRACT	v
ABSTRAK	vi
KATA PENGANTAR.....	vii
DAFTAR ISI.....	ix
DAFTAR GAMBAR	xiii
DAFTAR TABEL.....	xv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Identifikasi Masalah	5
1.3 Rumusan Masalah	5
1.4 Tujuan Penelitian	5
1.5 Manfaat Penelitian	6
1.5.1 Pengembang Ilmu.....	6
1.5.2 Praktis.....	6
1.5.3 Peneliti.....	6
BAB II TINJAUAN PUSTAKA.....	7
2.1 Tinjauan Studi	7
2.2 Tinjauan Pustaka	10
2.2.1 Definisi Penyakit	10
2.2.2 Penyakit Menular	11
2.2.3 Data Mining	11
2.2.4 Algoritma <i>K-Means Clustering</i>	12
2.2.5 Penerapan Algoritma <i>K-Means</i>	14
2.2.6 Siklus Hidup Pengembangan Sistem	20

2.2.7 Perencanaan Sistem.....	22
2.2.8 Analisis Sistem.....	22
2.2.9 Desain Sistem.....	24
2.2.10 Desain Sistem Secara Rinci	24
2.2.11 Implementasi Sistem	29
2.2.12 Pemeliharaan Sistem	31
2.2.13 Teknik Pengujian Sistem	32
2.3 Framework	33
2.4 Pengujian Perangkat Lunak.....	33
2.5 Kerangka Pikir.....	35
BAB III METODE PENELITIAN	36
3.1 Objek Penelitian	36
3.2 Metode Penelitian.....	36
3.2.1 Pengumpulan Data	36
3.2.2 Tahap Analisis.....	37
3.2.3 Tahap Desain.....	38
3.2.4 Tahap Pembuatan	39
3.2.5 Pengujian Sistem	39
3.2.6 Tahap Implementasi	40
BAB IV HASIL PENELITIAN.....	41
4.1 Hasil Pengumpulan Data	41
4.2 Hasil Pemodelan	42
4.2.1 Penjelasan Algoritma	42
4.2.2 Perhitungan Algoritma	45
4.3 Hasil Analisis Sistem	51
4.3.1 Proses Sistem Berjalan	51

4.3.2 Proses Sistem Diusulkan	52
4.4 Hasil Pengembangan Sistem	52
4.4.1 Diagram Konteks.....	52
4.4.2 Diagram Berjenjang	53
4.4.3 Diagram Arus Data.....	53
4.4.4 Diagram Arus Data Level 0.....	53
4.4.5 DAD level 1 proses 1	54
4.4.6 DAD level 1 proses 2	54
4.4.7 DAD level 1 proses 3	55
4.5. Kamus Data	55
4.5.1 Kamus Data Terinci.....	55
4.6 Database	58
4.6.1 Daftar Input Yang Didesain	58
4.6.2 Desain File Secara Umum.....	59
4.7 Arsitektur Sistem.....	59
4.8 Interface Design Mekanisme User	59
4.9 Interface Design : Mekanisme Navigasi	60
4.10 Interface Design : Mekanisme Output-Login.....	60
4.11 Intrface Design : Mekanisme Output-Laporan User	61
4.12 Data Design : Struktur Data User.....	61
4.13 Program Design.....	62
4.14 Hasil Kontruksi Sistem.....	64
4.15 Kode Program Pengujian <i>White Box</i>	64
4.16 Flowchart Program Pengujian <i>White Box</i>	65
4.17 Flowgraph Program Pengujian <i>White Box</i>	65

4.18 Perhitungan CC Pada Pengujian <i>White Box</i>	66
4.19 Path Pada Pengujian <i>White Box</i>	66
4.20 Hasil Pengujian <i>Black Box</i>	66
BAB V PEMBAHASAN	68
5.1 Pembahasan Sistem	68
5.1.1 Hasil Tampilan Halaman Login	68
5.1.2 Hasil Tampilan Halaman Dashboard	68
5.1.3 Tampilan Halaman Administrator.....	69
5.1.4 Tampilan Halaman Data Pasien	69
5.1.5 Tampilan Halaman Input Data Centroid	70
5.1.6 Tampilan Halaman Iterasi <i>K-Means</i>	70
5.1.7 Tampilan Hasil <i>Clustering</i>	71
BAB VI PENUTUP	72
6.1 Kesimpulan.....	72
6.2 Saran.....	72
DAFTAR PUSTAKA	73

DAFTAR GAMBAR

Gambar 2.1 Tahapan <i>Knowledge Discovery In Databases</i>	12
Gambar 2.2 Clustering (Sumber : Xu dan Wunsch II, 2009).....	13
Gambar 2.3 Grafik Puskesmas Tiap Cluster.....	15
Gambar 2.4 Grafik Cluster 2.....	15
Gambar 2.5 Grafik Cluster 3.....	17
Gambar 2.6 Grafik Cluster 4.....	18
Gambar 2.7 Grafik Cluster 5.....	19
Gambar 2.8 Grafik Cluster 6.....	20
Gambar 2.9 Siklus Hidup Pengembangan Sistem.....	21
Gambar 2.10 Contoh Notasi Kesatuan Luar	28
Gambar 2.11 Contoh Notasi Arus Data	28
Gambar 2.12 Contoh Notasi Proses	29
Gambar 2.13 Contoh Notasi Smpnan Data	29
Gambar 2.14 Kerangka Pikir.....	35
Gambar 3.1 Sistem Yang Diusulkan.....	38
Gambar 4.1 Bagan Alir Dokumen	51
Gambar 4.2 Sistem Yang Diusulkan.....	52
Gambar 4.3 Diagram Konteks.....	52
Gambar 4.4 Diagram Berjenjang	53
Gambar 4.5 DAD Level 0	53
Gambar 4.6 DAD Level 1 proses 1	54
Gambar 4.7 DAD Level 1 Proses 2.....	54
Gambar 4.8 DAD Level 1 Proses 3.....	55
Gambar 4.9 Database	58
Gambar 4.10 Mekanisme User.....	59
Gambar 4.11 Mekanisme Navigasi	60

Gambar 4.12 Mekanisme <i>Output-Login</i>	60
Gambar 4.13 Mekanisme <i>Output-Laporan User</i>	61
Gambar 5.1 Tampilan <i>Login</i>	68
Gambar 5.2 Tampilan Halaman Dashbord	68
Gambar 5.3 Tampilan Halaman Administrator.....	69
Gambar 5.4 Tampilan Halaman Data Pasien	69
Gambar5.5 Halaman Input Data Pasien	70
Gambar 5.6 Tampilan Iterasi <i>K-Means</i>	70
Gambar 5.7 Tampilan Hasil <i>Cluster</i> Data Pasien	71

DAFTAR TABEL

Tabel 1.1 Data Penyakit Menular.....	3
Tabel 2.1. <i>State Of The Art</i> 6.....	7
Tabel 3.1 Tabel Atribut Data	36
Tabel 4.1 Hasil Pengumpulan Data.....	41
Tabel 4.2 Data Pasien.....	43
Tabel 4.3 Centroid Yang Dipilih.....	44
Tabel 4.4 Tabel Hasil Iterasi 1	48
Tabel 4.5 Tabel Hasil Iterasi 2	49
Tabel 4.6 Tabel Hasil Iterasi 3.....	49
Tabel 4.7 Tabel Hasil Iterasi 4	50
Tabel 4.8 Kamus Data User	55
Tabel 4.9 Kamus Data Pasien	56
Tabel 4.10 Pembobotan	56
Tabel 4.11 Kamus Data Centroid.....	57
Tabel 4.12 Kamus Data Hasil Clustering.....	57

BAB I

PENDAHULUAN

1.1. LATAR BELAKANG

Kesehatan merupakan hal yang sangat penting bagi manusia. Karena jika kesehatan dalam tubuh terganggu maka dapat membuat segala aktivitas yang dilakukan menjadi tidak nyaman. Manusia sangat rentan terhadap penyakit tetapi penyebabnya tidak kita sadari. Penyakit merupakan suatu keadaan dimana terdapat gangguan terhadap bentuk dan fungsi tubuh sehingga berada dalam keadaan tidak normal, salah satu penyakit yang sering menyerang manusia yaitu penyakit menular.

Penyakit menular adalah penyakit yang disebut juga infeksi, yang dapat menular ke manusia dapat disebabkan oleh agen biologi, seperti virus, jamur, bakteri dan parasit, penularanya bisa langsung atau media atau binatang pembawa penyakit. Penyakit menular dapat menyebabkan kesakitan, kematian dan kecacatan yang tinggi sehingga perlu dilakukan penanggulangan melalui upaya pencegahan, pengendalian dan pemberantasan yang efektif dan efisien. Perkembangan teknologi saat ini sangat pesat terutama dalam bidang kesehatan, dalam hal ini dengan memanfaatkan teknologi dapat menghasilkan data informasi yang berlimpah mengenai penyakit menular. Dengan adanya teknologi manusia dapat memanfaatkan teknologi untuk berbagai macam hal yang sifatnya positif. Seperti halnya membantu pemecahan masalah terhadap pengelompokan penyakit-penyakit menular di Rumah Sakit Zainal Umar Sidiki. (Ade Bastian 2018 : 1).

Rumah Sakit Zainal Umar Sidiki merupakan satu dari sekian Rumah Sakit milik Pemkab Gorontalo Utara yang berwujud Rumah Sakit, diurus oleh pemda Kabupaten dan tercatat kedalam Rumah Sakit tipe Belum ditetapkan. Rumah Sakit ini telah terdaftar mulai 27/01/2016 dengan Nomor Surat Izin 197 Tahun 2012 dan Tanggal Surat Izin 00/00/0000 dari bupati Gorontalo Utara dengan sifat perpanjangan, dan berlaku sampai dengan sekarang. Rumah Sakit ini mempunyai layanan unggulan di bagian Trauma Sentrel. Rumah Sakit ini mempunyai jumlah kamar menurut kelas di antaranya 10 kamar untuk ruang kamar kelas III, 1 Ruang kamar PICU, 2 Ruang kamar HCU, 3 Ruang kamar IGD, 2 Ruang kamar bayi baru lahir, dan 3 ruang kamar untuk bersalin serta mempunyai 1 ruang kamar operasi.

Pelayanan pasien pada rumah sakit ini terutama pada pasien penderita penyakit menular, sama seperti Pelayanan pasien pada umumnya, pasien penderita penyakit menular ditempatkan di dalam ruangan bersama pasien tidak menular, hanya ada beberapa pasien penderita penyakit menular diwajibkan memakai masker baik untuk pasien penderita maupun keluarga atau kerabat yang menjenguk, seperti pasien penderita penyakit TB (Tuberkulosis). Sedangkan dokter yang menangani yaitu Dokter spesialis bagian dalam. Pengolahan data penyakit di Rumah Sakit masih berbasis semi komputerisasi artinya pengolahan data menggunakan komputer, tetapi belum ada integritas pengelompokan data yang lebih efisien. Dengan menggunakan Metode *K-Means Clustering*, dapat mengelompokkan beberapa penyakit menular di Rumah Sakit Zainal Umar Sidiki.

Rumah sakit ini memiliki lebih dari 500 pasien setiap Tahunnya, diantaranya adalah pasien penderita penyakit menular. Beberapa penyakit menular seperti, TB (Turbeculosis), Diare, Ispa dan DBD, berikut merupakan tabel data penyakit menular.

Tabel 1.1 data penyakit menular

No	Nama	Tempat Tangga l Lahir	Jenis Kelamin	Penyakit	Cara Penularan	Tarif RS	Tahun
1	Della Dulanimo	Kwand ang/ 21 Januari 1996	Perempu an	Turbekulosi s	Melalui Udara	560,400	2017
2	Yunus Dukalang	Alata/ 07 Juli 1999	Laki- Laki	Turbekulosi s	Melalui Udara	519,250	2017
3	Awaludin I Lihoe	Poso/ 16 Septem ber 2000	Laki- Laki	Ispa	Melalui Udara	501,250	2017
4	Mohamad Noka	Kwand ang/ 22 April 2005	Laki- Laki	Diare	Air,Tanah,Ma kanan Terkontamina si Virus	507,250	2017
5	Feby Sione	Dambal o/ 08 Oktobe r 1980	Perempu an	Ispa	Melalui Udara	393,500	2017
6	Radis Makoi	Leboto/ 18 Juli 2005	Laki- Laki	Dbd	Nyamuk Aedes Aegypty	619,250	2017

K-Means adalah metode penganalisaan data atau metode yang melakukan proses pemodelan tanpa supervise (*unsupervised*) dan merupakan salah satu metode pengelompokan data dengan sistem partisi. Metode ini mengelompokan data yang ada ke dalam beberapa kelompok, dimana yang berbeda dengan data yang ada di dalam kelompok yang lain. Tujuannya untuk meminimalisasikan

objective function yang di set dalam proses *clustering*, yang pada dasarnya berusaha untuk meminimalkan variasi dalam satu cluster dan memaksimalkan variasi antar cluster (Ni Wayan Wardani 2016).

Dalam penelitian Magdalena Simanjuntak, Ediman Manik dan Putri Ratna Sari (2019 : 1) judul “ Penerapan Data Mining Pengelompokan Penyakit Menular Seksual (PMS) Menggunakan Metode Clustering” bahwa penerapan ini bertujuan untuk menentukan konsistensi data kesehatan dapat digunakan teknik data mining yang mampu menggali informasi tersembunyi dari kumpulan data *multidimensi* yang telah diperoleh.

Dalam penelitian ini akan melakukan pengelompokan data penyakit menular menggunakan Algoritma *K-Means*, Algoritma *K-Means* dipilih dalam melakukan pengelompokan karena algoritma tersebut dapat mengelompokkan data yang ada kedalam beberapa kelompok, dimana data dalam satu kelompok mempunyai karakteristik yang sama satu sma lainnya. Sehingga pihak rumah sakit dapat membantu dalam pengelompokan data penyakit menular. Kelebihan dari Algoritma *K-Means* itu sendiri yaitu mudah diimplementasikan. Berikut merupakan variabel yang digunakan dalam penelitian ini yaitu nama penyakit menular, cara penularan, tahun dan jumlah pasien penderita.

Berdasarkan uraian diatas maka dianggap perlu untuk melakukan penelitian mengenai proses yang berjalan di atas, dengan judul “**Implementasi Data Mining Dengan Algoritma *K-Means Clustering* Penyakit Menular Di Rumah Sakit Zainal Umar Sidiki.**

1.2. IDENTIFIKASI MASALAH

Berdasarkan latar belakang di atas dapat diidentifikasi beberapa masalah sebagai berikut :

1. A adanya kesulitan dalam mengelompokan data penyakit menular
2. Belum adanya aplikasi untuk mengelompokan data penyakit menular dengan menggunakan metode *K-Means clustering*.

1.3. RUMUSAN MASALAH

Berdasarkan latar belakang masalah di atas, dapat dirumuskan pokok yang berkaitan yaitu :

1. Bagaimana merancang sistem pengolahan data penyakit menular dengan metode *K-Means Clustering* ?
2. Bagaimana menerapkan metode *K-Means Clustering* dalam mengcluster penyakit menular?

1.4. TUJUAN PENELITIAN

Adapun tujuan dari penelitian ini adalah sebagai berikut :

1. Untuk merancang sistem pengolahan data penyakit menular dengan menggunakan metode *K-Means Clustering*.
2. Untuk penerapan metode *K-Means Clustering* dalam mengcluster penyakit menular sehingga dapat diimplementasikan

1.5 MANFAAT PENELITIAN

Adapun manfaat penelitian sebagai berikut :

1. Pengembangan Ilmu

Dari hasil penelitian ini diharapkan dapat menambah pengetahuan dibidang teknoogi pada umumnya khususnya pada bagaimana cara menerapkan algoritma *K-Means* untuk mengelompokan penyakit menular.

2. Praktis

Sebagai bahan masukan untuk mendukung dalam melaksanakan penelitian tentang Implementasi Data Mining Dengan Algoritma K-Means Clustering Penyakit Menular Di Rumah Sakit Zainal Umar Sidiki.

3. Peneliti

Dapat dilakukan pengembangan lebih lanjut dengan mengcluster penyakit menular berdasarkan jenis penyakitnya.

BAB II

LANDASAN TEORI

2.1. Tinjauan Studi

Ada beberapa peneliti yang terkait tentang implementasi data mining penyakit menular dan menggunakan metode *K-Means Clustering*, seperti di bawah ini

Tabel 2.1 *State of The Art*

NO	PENELITI	JUDUL	TAHUN	METODE	HASIL
1	NI WAYAN WARDANI	Analisis Penerapan <i>K-Means</i> untuk pengelompokan diagnose penyakit kulit dan kelamin berdasarkan rentang usia	2016	K-Means Clustering	Dengan menerapkan dendekatan data mining menggunakan metode atau Aalgoritma <i>K-Means Clustering</i> untuk menemukan data atau informasi tentang penyakit menular pada data rekam medis, Sehingga dapat diketahui hasilnya bahwa

					faktor usia dapat mempengaruhi jenis penyakit seseorang, dan kebanyakan penderita berusia antara 26 – 50 tahun.
NO	PENELITI	JUDUL	TAHUN	METODE	HASIL
2	ADE BASTIAN	Penerapan Algoritma K-Means Clustering Analysis pada penyakit menular manusia (Studi kasus Kabupaten Majalengka)	2018	K-Means Clustering	Penelitian ini menggunakan metode K-Means Clustering yaitu pengelompokan objek cluster bertujuan untuk mempartisi objek yang ada kedalam satu cluster objek atau lebih berdasarkan karakterisiknya. Dalam penelitian ini menggunakan metode K-Means cluster analysis dalam

					<p>penyakit menular pada manusia berdasarkan set variable yang dibentuk per kecamatan tiap puskesmas yang jumlahnya 32 kantor puskesmas di Kabupaten Majalengka</p>
NO	PENELITI	JUDUL	TAHUN	METODE	HASIL
3	MAGDALENA SIMANJUNTAK	<p>Penerapan Data Mining Pengelompokan Penyakit Menular Seksual (PMS) Menggunakan Metode Clustering</p>	2019	K-Means Clustering	<p>Pada penelitian ini dapat diketahui pada cluster 3 dari kriteria usia, jenis penyakit dan faktor penyebab, kelompok mana yang memiliki himpunan / nilai paling tinggi dan paling banyak pasien yang mengalami sakit, yaitu pada</p>

					cluster 1 ber jumlah 247 data pasien yang sakit usia 18 – 35 tahun dengan penyakit yang dialami ialah penyakit sifilis dan faktor penyebabnya adalah bakteri spiroseta. Dari pengujian dapat diketahui juga bahwa antara usia, jenis penyakit, dan faktor penyebab memiliki hubungan
--	--	--	--	--	--

2.2 Tinjauan Pustaka

2.2.1 Definisi Penyakit

Penyakit adalah kondisi di mana tidak normalnya organ tubuh yang menyebabkan ketidakselarasan, disfungsi atau tekanan/stress pada seseorang. Penyakit juga dikenal sebagai keadaan di mana berhubungan dengan gejala dan tanda klinis tertentu. Berbagai jenis penyakit yang mengancam kehidupan manusia, dapat disebabkan oleh kuman, virus, bakteri, dan racun. Kematian karena penyakit disebut kematian oleh sebab alami. Secara umum, ada empat jenis

penyakit utama : penyakit menular, penyakit defisiensi, penyakit keturunan, dan penyakit fisiologis.

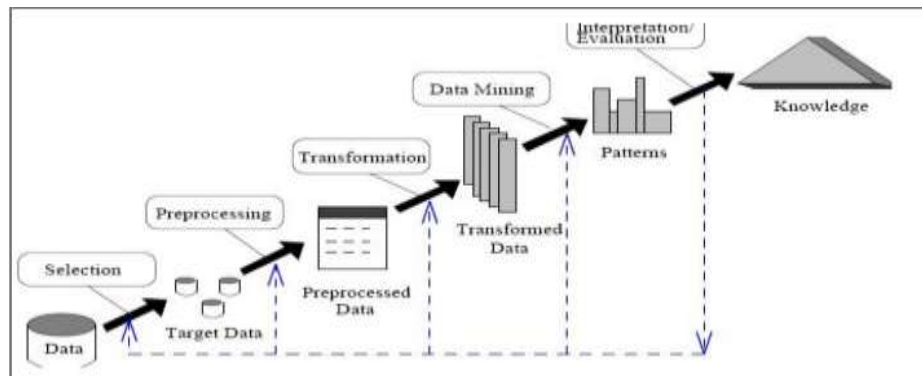
2.2.2 Penyakit Menular

Penyakit menular adalah gangguan pada organ tubuh biasanya disebabkan oleh organisme yang berbahaya bagi tubuh seperti virus, bakteri, dan jamur, penyakit menular juga bisa disebabkan oleh infeksi dari orang ataupun hewan. Penularannya terjadi jika secara langsung yaitu pada saat seseorang melakukan kontak fisik dengan orang yang sakit, seperti lewat sentuhan, melalui udara atau saat bersin dan batuk, biasanya juga melalui kontak dengan cairan tubuh yaitu seperti melalui urin dan darah. Kebanyakan orang yang menularkannya tidak memperlihatkan gejala atau disebut sebagai (carrier) penyakit.

2.2.3 Data Mining

Data mining merupakan suatu proses yang memungkinkan untuk digunakan oleh para pengguna dalam mengakses data secara cepat dengan jumlah yang besar. Sebagai suatu proses penggalian data dan informasi yang belum diketahui sebelumnya, tetapi dapat dipahami oleh pengguna dan untuk membuat suatu keputusan yang sangat penting. Data mining juga merupakan proses interaktif untuk menemukan model baru yang dapat digeneralisasi untuk masa yang akan datang, juga bermanfaat sehingga dapat di mengerti dalam suatu data base yang sangat besar (*massive database*) (D.Firdaus 2017 : 2).

Knowledge discovery data (KDD) adalah keseluruhan proses *non-trivial* untuk mencari dan mengidentifikasi pola (pattern) dalam data, dimana pola yang di temukan bersifat sah, baru dapat bermanfaat dan dapat mengerti. (A. Faliis 2013)



Gambar 2.1 Tahapan *Knowledge Discovery In Databases*

Data mining juga digunakan untuk menemukan data, dan data tersebut akan digunakan pengguna untuk membangun sebuah bentuk ataupun model, sehingga dari model tersebut pengguna dapat mengidentifikasi struktur data yang lain yang berada di luar basis data yang di simpan. Teknik pembelajaran yang digunakan dalam memrediksi suatu nilai dari target variabel kategori, merupakan salah satu dari teknik data mining yaitu teknik klasifikasi.

2.2.4 Algoritma K-Means Clustering

K-Means merupakan algoritma *clustering*, tujuannya yaitu untuk membagi data menjadi beberapa kelompok. Kumpulan data yang sama pada satu kelompok atau data yang berbeda pada kelompok yang lain disebut juga sebagai *cluster*. Sedangkan (Xu & Wunsch II, 2009) mengungkapkan bahwa teknik *clustering* adalah teknik untuk membagi objek atau data menjadi beberapa yaitu bentuk, entitas, contoh, unit, kedalam beberapa jumlah kelompok seperti, grup, bagian ataupun kategori. Tujuan dari proses *clustering* adalah untuk meminimalkan variasi dalam suatu *cluster* serta memaksimalkan variasi antar *cluster*, artinya yaitu untuk mengelompokkan data ke dalam suatu *cluster* sehingga objek tersebut

memiliki kemiripan terhadap objek lain pada *cluster* yang sama, dan tidak memiliki kemiripan terhadap objek lain pada *cluster* yang lain.



Gambar 2.2 Clustering (Sumber : Xu dan Wunsch II, 2009)

Seperti penjelasan diatas bahwa teknik *clustering* yaitu suatu proses yang bertujuan untuk memaksimalkan serta meminimumkan kesamaan antar anggota dalam satu *cluster*, dan juga bertujuan untuk menemukan *cluster* yang berkualitas dalam waktu yang layak. Kegunaan lain dari *clustering* adalah menemukan bentuk atau pola pembagian dalam sebuah data set dan terdapat pada proses analisa data. Berikut merupakan cara untuk melakukan pengelompokan menggunakan algoritma *K-Means Clustering*:

1. Pilih K buah titik *centroid* secara acak
2. Mengelompokan data sehinga terbentuk K atau sebuah *cluster* dengan titik *centroid* dari setiap *cluster* merupakan titik *centroid* yang dipilih sebelumnya
3. Memperbaharui nilai titik *centroid*
4. Ulangi langkah 2 dan 3 sampai nilai dari titik *centroid* tidak lagi berubah

Proses pengelompokan data ke dalam suatu *cluster* dilakukan dengan cara menghitung jarak terdekat dari suatu data ke sebuah titik *centroid* menggunakan

perhitungan jarak minkowski. Adapun rumus perhitungan jarak minkowski adalah sebagai berikut:

$$d(x_i, x_j) = (|x_{i1} - x_{j1}|^g + |x_{i2} - x_{j2}|^g + \dots + |x_{ip} - x_{jp}|^g)^{1/g}$$

Di mana :

$g = 1$, untuk menghitung jarak *Manhattan*

$g = 2$, untuk menghitung jarak *Euclidean*

$g = \infty$, untuk menghitung jarak *Chebychev*

x_i, x_j adalah dua buah data yang akan dihitung jaraknya

p = dimensi dari sebuah data

Berikut adalah pembaharuan suatu titik *centeroid* yang dapat dilakukan dengan rumus berikut :

$$\mu_k = \frac{1}{N_k} \sum_{q=1}^{N_k} x_q$$

Di mana:

μ_k = titik centroid dari cluster ke-K

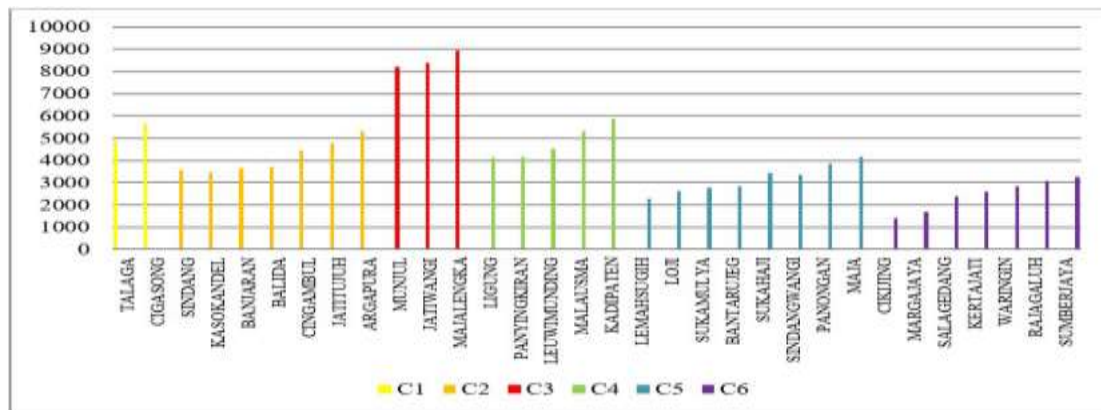
N_k = banyaknya data pada cluster ke-K

x_q = data ke-q pada cluster ke-K

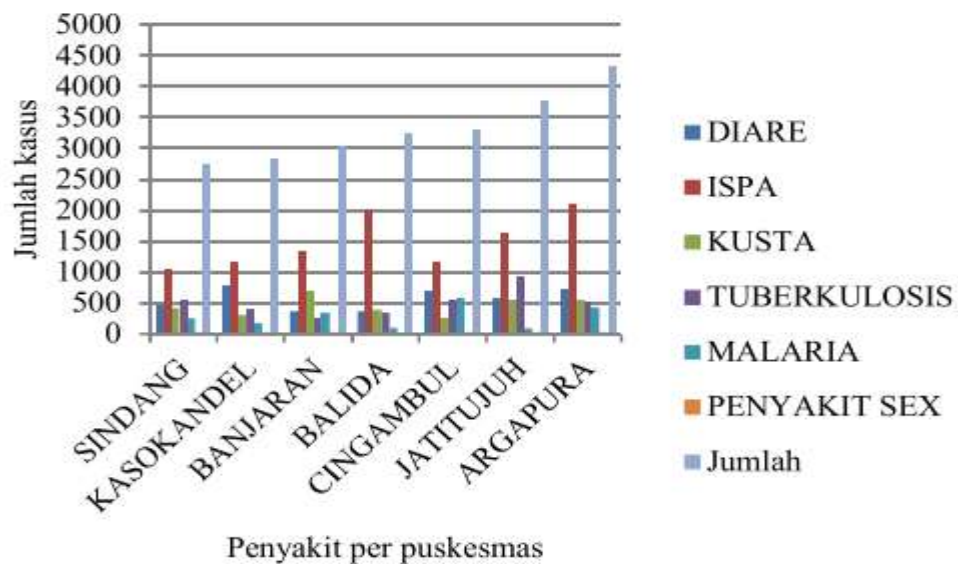
2.2.5 Penerapan Algoritma K-Means

Berikut ini merupakan penelitian oleh (Ade Bastian 2018 : 1, yaitu tentang pengelompokan penyakit menular dengan Metode *K-Means Clusrering*.) Memilih jumlah (*cluster k*), Inisialisasi ke pusat *cluster* ini dilakukan dengan cara random atau acak dan Pusat-pusat *cluster* diberi nilai awal dengan angka-angka

random. Alokasikan semua data atau objek ke *cluster* terdekat dan kedekatan kedua objek ditentukan atas jarak antara kedua objek tersebut.



Gambar 2.3 grafik puskesmas tiap cluster



Gambar 2.4 grafik cluster 2

Pada tahap ini harus menghitung jarak tiap data ke tiap pusat *cluster*. Jarak antara satu data dengan satu *cluster* dapat menentukan suatu data akan masuk ke

cluster yang mana. Berikut merupakan rumus persamaan teori jarak *Euclidean* yaitu untuk menghitung jarak semua data ke titik pusat *cluster*.

$$D(i, j) = \sqrt{(X_{1i} - X_{1j})^2 + (X_{2i} - X_{2j})^2 + \dots + (X_{ki} - X_{kj})^2} \quad (1)$$

Dimana:

Dij diketahui sebagai jarak data ke titik pusat *cluster j*

X_{ki} juga sebagai data ke i pada atribut data ke k

X_{kj} adalah titik pusat ke j pada atribut ke k

Dan jarak pusat *cluster* dihitung kembali dengan anggota *cluster* yang sekarang.

Pusat *cluster* disebut juga sebagai semua data atau objek dalam suatu *cluster* tertentu, jika memungkinkan juga bisa menggunakan media dari *cluster* tersebut.

Setiap objek kemudian ditugaskan kembali menggunakan pusat *cluster* yang baru,

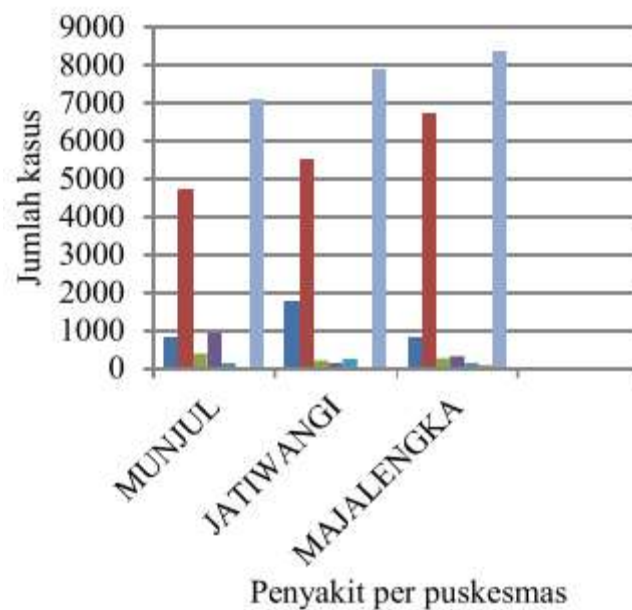
dan proses *clustering* bisa dikatakan selesai jika pusat *cluster* tidak berubah lagi,

tetapi kembali ke langkah nomor 3 bila pusat *cluster* berubah sampai pusat *cluster* benar-benar tidak berubah lagi.

Tahap-tahap yang dilakukan oleh algoritma *K-Means* adalah sebagai berikut:

1. Mulai
2. Menentukan banyaknya *cluster* dari data penyakit menular
3. Pengaturan untuk nilai awal titik tengah atau *centroid*, kemudian memilih pusat *cluster* secara acak pada data awal
4. Melakukan perhitungan data penyakit menular ke *centroid* dengan memakai rumus jarak *Euclid*

5. Melakukan *clustering* pada data memasukan setiap objek ke dalam *cluster* atau grup berdasarkan jarak minimumnya
6. Menghitung pusat cluster baru jika ada data yang harus di pindah. Pusat cluster baru ditentukan menurut pengelompokan anggota tiap-tiap *cluster* baru, dan untuk *cluster* baru yang pertama dihitung berdasarkan rata-rata koordinat. Jika hasil perhitungan menunjukkan adanya angka pusat *cluster* yang sama maka pengulangan dihentikan.
7. Selesai



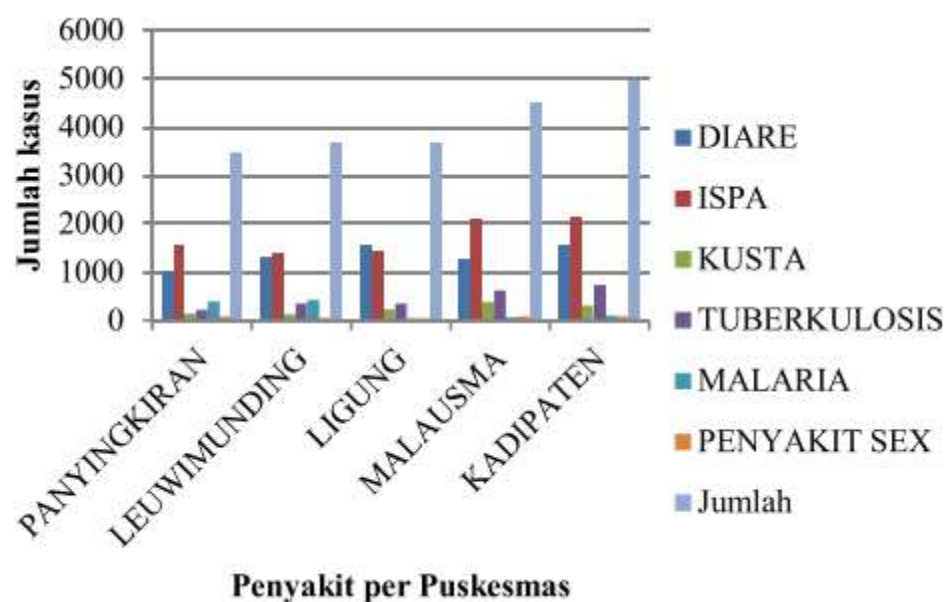
Gambar 2.5 grafik cluster 3

Dalam tahap ini juga harus melakukan perhitungan jarak tiap data ke tiap pusat *cluster*, jarak paling dekat antara suatu data dengan suatu *cluster* tertentu akan menentukan suatu data masuk dalam *cluster* mana. Teori jarak *Euclidean* merupakan rumus yang digunakan untuk menghitung jarak semua data ke tiap

titik pusat *cluster*. Pengulangan dihentikan karena hasil perhitungan adanya pusat *cluster* yang sama pada iterasi ke-5 dan ke-6.

Sementara itu, *cluster* 2 (ditunjukkan pada gambar 2.4) pada puskesmas Cingambul, Banjaran, Argapur, Sindang, Balida, Kasokandel dan Jatitujuh terdapat 23245 jiwa penderita penyakit menular didominasi oleh penyakit ISPA dan Diare sedangkan berdasarkan puskesmas, yaitu berasal dari puskesmas Argapura dan Balida.

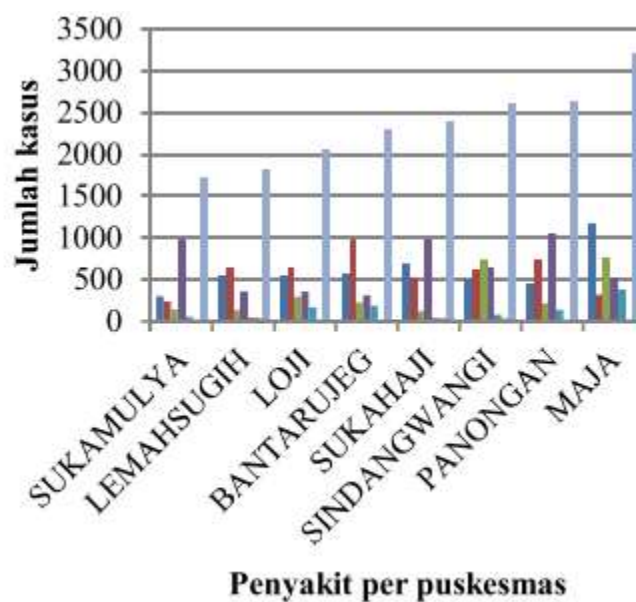
Gambar 2.5 menunjukkan hasil *cluster* 3 *cluster*, pada puskesmas Majalengka, Manjul, dan Jatiwangi terdapat 23415 jiwa penderita penyakit menular, pada *cluster* 3 didominasi oleh penyakit ISPA dan Diare. Sedangkan yang mendominasi berdasarkan puskesmas, yaitu pada puskesmas Majalengka dan Jatiwangi.



Gambar 2.6 grafik cluster 4

Hasil *cluster* 4 seperti pada gambar 2.6, puskesmas Malausma, Leuimunding, Panyingkiran, Kadipaten dan Ligung terdiri dari 20334 jiwa penderita penyakit menular, terlihat bahwa penyakit menular pada *cluster* 4 didominasi oleh penyakit ISPA dan Diare. Sedangkan pada puskesmas, yang mendominasi adalah puskesmas Kadipaten dan Malausma.

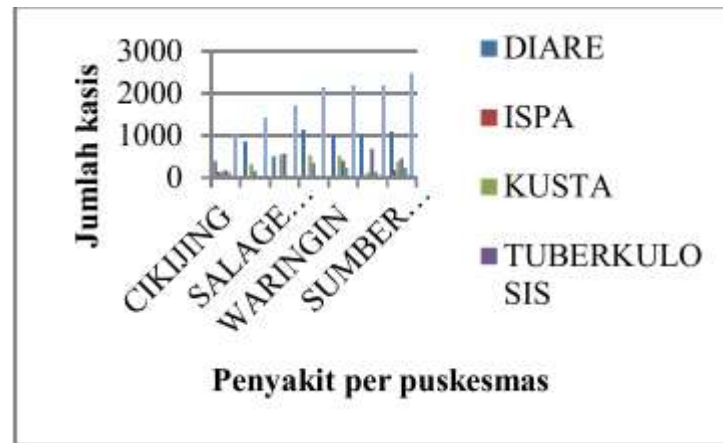
Cluster 5 (gambar 2.7) puskesmas Lemahsugih, Bantarujeg, Maja, Sukahaji, Sindangwangi, Loji tercatat 16384 jiwa penderita penyakit menular, Sukamulya dan Panongan terlihat bahwa penyakit menular pada *cluster* 5 didominasi oleh penyakit Turbekulosis dan Diare. Seadngkan, berdasarkan puskesmas didominasi oleh puskesmas Panongan dan Maja.



Gambar 2.7 grafik cluster 5

Cluster 6 pada puskesmas Margajaya, Cikijing, Salagedang, Rajagaluh, Waringin, Kertajati dan Sumberjaya , terdiri dari 13143 jiwa penderita penyakit menular. Pada gambar 2.8 terlihat bahwa penyakit menular pada *cluster* 6

didominasi oleh penyakit menular Diare dan Turbekulosis. Sedangkan, berdasarkan puskesmas didominasi oleh puskesmas Kertajati dan Sumberjaya.



Gambar 2.8 grafik cluster 6

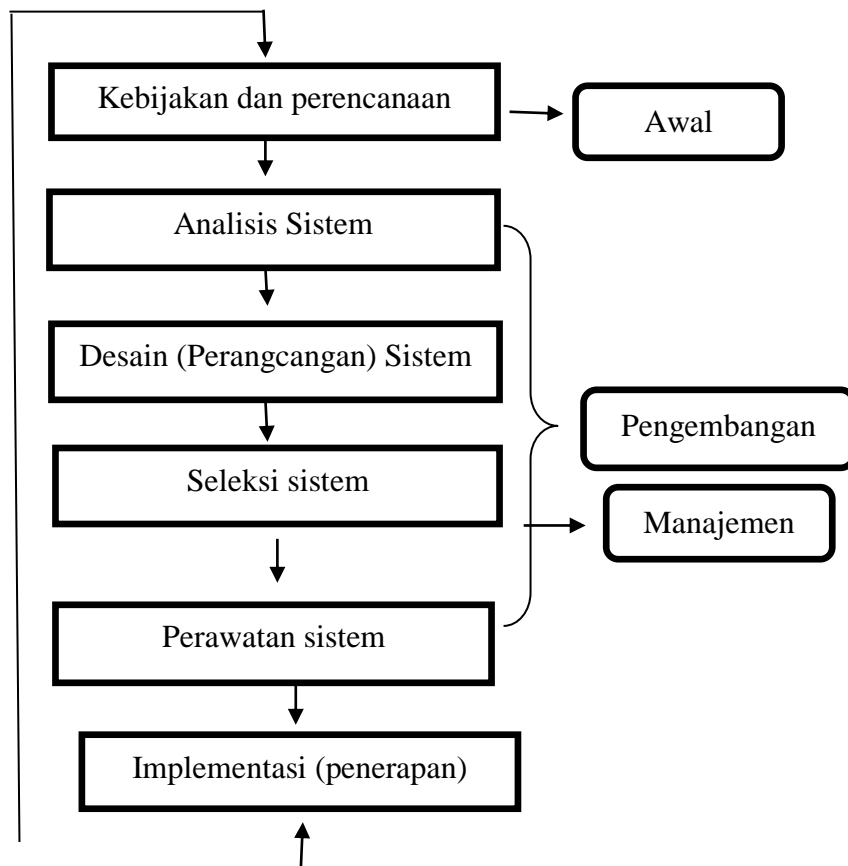
Dari data yang diolah berdasarkan jenis barang diinputkan sampel data sebanyak 32 data dengan jumlah yang diperoleh 6 kelompok data telah tercluster.

2.2.6 Siklus Hidup Pengembangan Sistem

Pada rekayasa sistem dan perangkat lunak siklus hidup pengembangan sistem diartikan sebagai proses pembuatan serta perubahan sistem dan model dengan menggunakan metodologi untuk pengembangan sistem tersebut. Proses ini akan melewati beberapa tahap dari mulai sistem itu direncanakan sampai sistem tersebut diterapkan, dioperasikan dan dipelihara. Jika sistem operasi yang telah dikembangkan masih kelihatan permasalahan-permasalahan yang berat dan tidak dapat diatasi dalam tahap tersebut, maka perlu dikembangkan kembali suatu sistem agar dapat mengatasi proses ini atau kembali ke tahap pertama, yaitu tahap perencanaan sistem. Siklus ini juga disebut dengan siklus hidup suatu sistem

(*systems life cycle*). (Jogianto, HM, 1999 Analisis Dan Desain Sistem Informasi : Pendekatan terstruktur teori dan praktek aplikasi bisnis edisi II Yogyakarta : Andi Yogyakarta). Siklus ini juga disebut dengan siklus hidup suatu sistem (*systems life cycle*). (Jugiyanto HM, 2005 : 52).

Siklus hidup pengembangan sistem dengan langkah-langkah utamanya yang akan digunakan adalah sebagai berikut :



sumber : Jogianto, H M, (2005 : 52)

Gambar 2. 9 Siklus Hidup Pengembangan Sistem

2.2.7 Perencanaan Sistem

Perencanaan sistem merupakan kata dari sebuah konsep, dimana dalam pengembangan suatu sistem/perangkat lunak konseptualisasi ini dilakukan dengan maksud tujuan

Garth N.Jone (2007 : 12) mengungkapkan:

“perencanaan adalah suatu proses pemilihan dan pengambilan dari pada tindakan yang paling baik untuk pencapaian tugas”.

Perencanaan atau *palnning* adalah hal-hal yang berkaitan dengan studi, yaitu tentang kebutuhan pengguna atau (*user's spesification*), studi kelayakan (*feasibility study*) baik secara teknis maupun secara teknologi dan juga penjadwalan pengembangan suatu proyek sistem informasi atau perangkat lunak. Yang mana pada sesi perencanaan ini pengembang melakukan observasi untuk mengenai calon pengguna dari isitem informasi/perangkat lunak yang akan di kembangkan nantinya. Pada pengembangan sistem/perangkat lunak berorientasi objek yang menggunakan DAD sebagai kakas (*tool*), semua permasalahan dimodelkan sebagai *usr case* untuk menggambarkan seluruh kebutuhan-kebutuhan pengguna.

2.2.8 Analisis Sistem

Satzinger, J.W.,Jackson, R.B., dan Burd, S.D. (2010 : p4) mengungkapkan “analisis sistem merupakan penentuan secara rinci atau pemahaman atas apa yang seharusnya dicapai oleh sistem informasi”. *System Analisa* adalah penjabaran suatu sistem informasi dari yang utuh kedalam bagian-bagian elemennnya dengan

maksud untuk dapat mengidentifikasi suatu masalah atau hambatan yang akan timbul pada sistem tersebut sehingga nantinya dapat melakukan penanggulangan atau perbaikan dan juga pengembangan.

Analisis sistem adalah spesialis mempelajari masalah dan kebutuhan sebuah organisasi untuk menentukan bagaimana data proses dan teknologi informasi dapat mencapai kemajuan terbaik untuk bisnis.

Tahap analisis merupakan tahap sangat penting karena jika melakukan kesalahan pada tahap ini akan menyebabkan kesalahan juga ditahap selanjutnya. Pada tahap analisis sistem mencakup studi kelayakan dan analisis kebutuhan, Whitten, et al. (2004 : 33).

Langkah–langkah dasar yang harus dilakukan pada tahap analisis sistem, adalah sebagai berikut :

1. *Identify*, yaitu mengidentifikasi masalah.

Mengidentifikasi masalah adalah langkah awal yang dilakukan dalam tahap analisis sistem. Masalah (*problames*) dapat diartikan sebagai suatu pertanyaan yang di inginkan untuk dipecahkan.

2. *Understand*

Pada tahap ini untuk mengetahui apakah sistem yang ada beroperasi yaitu dengan cara memahami kerja dari sistem yang ada, dan juga data diperlukan untuk mempelajari operasi dari sistem ini yaitu dengan cara dengan cara melakukan penelitian.

3. *Analyze*, yaitu menganalisis sistem tanpa *report*

Langkah ini dilakukan berdasarkan data yang telah diperoleh dari hasil penelitian yang telah dilakukan.

4. Report, yaitu membuat laporan hasil analisis

Tujuan utama dari pembuatan laporan hasil analisis :

- a. Pelapor bahwa analisis telah selesai dilakukan
- b. Meluruskan kesalahan pengertian mengenai apa yang telah ditemukan dan dianalisis oleh analisis sistem tetapi tidak sesuai menurut manajemen.

2.2.9 Desain Sistem

Sebelumnya pada analisis sistem telah mendapatkan gambaran yang jelas apa yang akan dikerjakan selanjutnya. Kemudian yang harus dilakukan yaitu memikirkan bagaimana membentuk sistem tersebut. Tahap tersebut merupakan tahap desain sistem (*system design*).

Whitten, et, al. (2004 : 34) mengungkapkan : “*Syistem design* merupakan spesifikasi atau intruksi solusi yang teknis dan berbasis komputer dan persyaratan bisnis yang didefinisikan dalam analisis sistem.”

2.2.10 Desain sistem secara rinci (*detailed system design*)

1. Desain *Output* terinci

Desain *output* terinci yaitu desain untuk mengetahui bagaimana dan seperti apa bentuk *output – output* dari sistem yang baru. Desain *output* terinci dibagi atas dua yaitu desain *output* berbentuk laporan dimedia kertas dan desain *output* dalam bentuk dialog padal layar terminal.

- 1) Desain *output* dalam bentuk laporan

Desain ini dimaksudkan untuk menghasilkan *output* dalam bentuk laporan dimedia kertas. Bentuk laporan yang paling banyak digunakan adalah dalam bentuk tabel, grafik atau bagan.

2) Desain *output* dalam bentuk dialog layar Terminal

Desain ini merupakan rancang bangun dari percakapan antara pengguna sistem (*user*) dengan komputer. Percakapan ini yaitu terdiri dari proses memasukan data kesistem, menampilkan *output* informasi kepada *user* atau keduanya. Beberapa strategi dalam membuat layar dialog terminal :

- Dialog pertanyaan / jawaban
- Menu

Menu banyak digunakan karena merupakan jalur pemakai yang mudah dipahami dan mudak di gunakan.Menu berisi beberapa alternatif atau option atau pilihan yang disajikan kepada *user*. Pilihan menu akan lebih baik bila dikelompokan sesuai fungsinya.

2. Desain *input* terinci

Masukan adalah awal dimulainya proses informasi. Bahan mentah dari informasi adalah data yang terjadi dari transaksi – transaksi yang dilakukan oleh organisasi. Data hasil transaksi merupakan masukan untuk sistem informasi.Hasil dari sistem informasi tidak lepas dari data yang dimasukan. Desain *input* terinci dimulai dari desain dokumen dasar sebagai penangkap *input* yang pertama kali. Jika dokumen dasar tidak didesain dengan baik, kemungkinan *input* yang tercatat dapat salah bahkan kurang.

Fungsi dokumen dasar untuk pembenahan arus data :

- a. Dapat menunjukan jenis dari data yang harus dikumpulkan.
- b. Data harus dicatat dengan jelas, konsisten dan akurat.
- c. Dapat mendorong lengkapnya data yang dibutuhkan juga disebutkan satu persatu didalam dokumen dasarnya.

3. Desain *database* terinci

Basi data (*database*) dapat diartikan sebagai kumpulan dari data yang saling berhubungan satu dengan yang lain, tersimpan didalam komputer dan digunakan perangkat lunak tertentu untuk memanipulasinya. *Database* merupakan bagian yang sangat penting dalam sistem informasi karena berfungsi sebagai gudang penyimpanan data untuk dapat diolah lebih lanjut.

database system juga merupakan kumpulan dari data yang saling terkait satu sama lain dengan tujuan untuk mengelola data agar lebih *efektif* dan *efisien*, dan juga terdapat beberapa aplikasi yang bermacam – macam didalam suatu organisasi. Dalam sistem basis data, setiap orang atau bagian dapat melihat *database* dari beberapa sudut pandang yang berbeda-beda.

Pada tahap ini, desain *database* yaitu untuk mendefinisikan isi atau struktur dari tiap – tiap file yang telah didefinisikan didesain secara umum.

4. Desain teknologi

Tahapan desain teknologi terbagi atas dua yaitu desain teknologi secara umum dan terinci. Pada tahap ini kita menentukan teknologi yang akan dipergunakan dalam menerima input, menjelaskan model, menyimpan dan mengakses data, menghasilkan dan mengirimkan keluaran dan membantu pegendalian dari sistem secara keseluruhan.

Teknologi yang dimaksud meliputi :

- a. Perangkat keras (*hardware*), yaitu terdiri dari alat masukan, alat pemroses, alat output dan simpana luar.
- b. Perangkat lunak (*software*), yang terdiri dari perangkat lunak sistem operasi (*operating system*), perangkat lunak bahasa (*language software*) dan perangkat lunak aplikasi (*application software*).
- c. Sumber daya manusia (*brainware*), misalnya operator komputer, pemrogram, spesialis telekomunikasi, sistem analisis dan sebagainya.

Desain teknologi sangat diperlukan pada tahap implementasi dan pengujian untuk membuktikan bahwa sistem ini dapat berjalan secara semestinya.

5. Desain model

Tahap desain model dibagi menjadi dua yaitu desain model secara umum dan terinci. Tahap desain model secara umum yaitu berupa desain secara fisik dan logika. Desain fisik dapat digambarkan dengan bagan alir dokumen. Desain secara logika digambarkan dengan diagram arus data (DAD), pada tahap desain model terinci, model akan mendefinisikan secara rinci urutan – urutan langkah dari masing – masing proses yang digambarkan di DAD. Urutan langkah proses ini diwakili oleh suatu program komputer.

Untuk mempermudah penggambaran suatu sistem yang ada atau sistem yang baru yang akan dikembangkan secara logika tanpa memperhatikan lingkungan fisik dimana data tersebut mengalir atau lingkungan fisik dimana data tersebut akan disimpan, maka digunakan Diagram Arus Data (DAD) atau Data Flow

Diagram (DFD). Dalam menggambarkan sistem perlu dilakukan pembentukan simbol, berikut merupakan ini simbol – simbol yang sering digunakan dalam DAD :

1. *Externan entity* (kesatuan luar) atau *boundary* (batas sistem). Setiap sistem pasti mempunyai batas sistem (*boundary*) yng memisahkan suatu sistem dengan lingkungan luarnya. Sistem akan menerima input dan menghasilkan output kepada lingkungan luarnya. Kesatuan luar (*external entity*) adalah kesatuan dilingkungan luar sistem yang dapat berupa orang, organisasi atau sistem lain yang berada dilingkungan luarnya yang akan meberikan *input* serta menerima *output* dari sistem, (Jogiyanto HM, 2005 : 701).



Gambar 2. 10 contoh notasi kesatuan luar

2. Data Flow (arus data)

Arus data ini menunjukkan arus atau aliran data yang dapat berupa masukan atau sistem atau hasil dari proses sistem, (Jogiyanto HM, 2005:702)

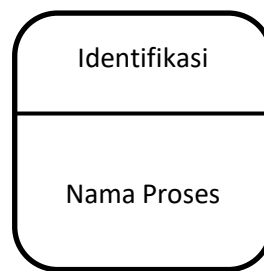
Nama Arus Data



Gambar 2. 11 Contoh Notasi Arus Data

3. *Process* (proses)

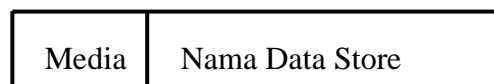
Suatu proses adalah kegiatan atau kerja yang dilakukan orang, mesin atau komputer dari hasil suatu arus data yang masuk kedalam proses untuk dihasilkan arus data yang akan keluar dari proses, (Jogiyanto HM, 2005:705)



Gambar 2.12 Contoh Notasi Process

4. Data Store (simpanan data)

Simpanan data pada DFD dapat disimbolkan dengan sepasang garis horisontal paralel yang tertutup disalah satu ujungnya, (Jogiyanto HM, 2005:707).



Gambar 2. 13 Contoh Notasi Simpanan Data

2.2.11 Implementasi Sistem

Sistem dianalisa dan didesain secara rinci dan teknologi yang diseleksi dan dipilih. Tiba saatnya sekarang sistem untuk diimplementasikan (diterapkan). Tahap implementasi sistem merupakan tahap meletakkan sistem supaya siap untuk dioperasikan. Tahap implementasi sistem dapat terdiri dari langkah-langkah sebagai berikut :

1. Menerapkan rencana implementasi

Rencana implementasi adalah kegiatan awal dari tahap implementasi sistem. Rencana implementasi yaitu untuk mengatur biaya dan waktu yang dibutuhkan selama tahap implementasi.

2. Melakukan kegiatan Implementasi

Kegiatan implementasi dilakukan dengan dasar kegiatan yang telah direncanakan dalam rencana implementasi.

3. Memerlukan tempat instalasi baik untuk perangkat keras maupun perangkat lunak

Jika memiliki peralatan baru, maka perlu juga mempertimbangkan tempat atau ruangan untuk peralatan tersebut. Sistem komputer yang besar juga membutuhkan tempat serta lingkungan yang harus diperhitungkan. Langkah selanjutnya yaitu setelah melakukan persiapan fisik tempat adalah penginstalasi perangkat lunak yang sudah ada.

4. Pemograman dan pengetesan sistem

Pemograman merupakan kegiatan menulis kode program yang akan dieksekusi oleh komputer. Kode program yang ditulis oleh pemogram harus berdasarkan dokumentasi yang harus di sediakan oleh analisis sistem hasil dari desain sistem secara rinci. Sebelum program diterapkan, maka program harus terlebih dahulu terbebas dari kesalahan-kesalahan. Oleh sebab itu, program harus diuji untuk menemukan kesalahan-kesalahan yang mungkin dapat terjadi. Program dites untuk tiap-tiap modul dan dilanjutkan dengan pengetesan untuk semua modul yang dirangkai.

5. Pengetesan sistem

Pengetesan sistem biasanya dilakukan setelah pengetesan program. Pengetesan sistem dilakukan untuk memeriksa kekompakan antar komponen sistem yang diimplementasikan. Tujuan utama dari pengetesan sistem ini adalah untuk memastikan bahwa elemen–elemen atau komponen-komponen dari sistem telah berfungsi sesuai dengan yang diharapkan.

2.2.12 Pemeliharaan Sistem

“Pemeliharaan sistem adalah proses pengubahan sistem setelah beroperasi dan digunakan”. (Hariyanto 2004 : 603)

“Pemeliharaan sistem adalah tahap dimana kita mulai pengoperasian sistem dan, jika diperluakn melakukan perbaikan – perbaikan kecil”. (A. Nugroho, 2010 : 7).

Menurut Zwass dalam Jogyanto (2005) membagi pemeliharaan perangkat lunak menjadi tiga macam, yaitu :

1. Pemeliharaan perfektif

Pemeliharaan perfektif ditunjukan untuk memperbaharui sistem lama sebagai tanggapan atas perubahan kebutuhan pemakai dan kebutuhan organisasi, meningkatkan efisien sistem, dan memperbaiki dokumen.

2. Pemeliharaan adaptif

Pemeliharaan adaptif berupa perubahan aplikasi untuk menyesuaikan diri terhadap lingkungan perangkat keras dan perangkat lunak baru. Sebagai contoh pemeliharaan ini dapat berupa perubahan aplikasi dari *mainframe* ke lingkungan

client/server atau mengkonversi dari sistem berbasis berkas ke lingkungan basis data.

3. Pemeliharaan korektif

Pemeliharaan korektif berupa pembetulan atas kesalahan – kesalahan yang ditemukan pada saat sistem berjalan.

2.2.13 Teknik Pengujian Sistem

1. *White box*

White box merupakan metode pengujian yang menggunakan struktur *control* desain prosedur untuk memperoleh tes *case*. Dengan menggunakan *white box*, perancang sistem dapat melakukan *test case* yang memberikan jaminan bahwa semua jalur independen pada suatu model telah digunakan paling tidak satu kali, menggunakan semua keputusan logis pada sisi *true* dan *false*, mengeksekusi semua *loop* pada batasan mereka dan pada batas operasional mereka, dan menggunakan struktur data internal untuk menjamin validitasnya. Pengujian *basis path* adalah teknik pengujian *white box* yang diusulkan pertama kali oleh Tom McCabe. Metode basis *path* ini memungkinkan desainer *test case* mengukur kompleksitas logis dari desain prosedural dan menggunakan sebagai pedoman untuk menetapkan basis set dari jalur eksekusi. (Roger S. Pressman, 2002 : 536).

Berikut adalah langkah-langkah teknik pengujian *white box*.

1. Menggambar *Flowgraph* yang telah ditransfer dari *flowchart*
2. Menjumlahkan *Cyclomatic Complexity* untuk *Flowgraph* yang telah dibuat
3. Menentukan jalur pengujian dari *flowgraph* yang sesuai dengan jumlah *cyclomatic complexity*.

2 *Black Box*

Black-box testing adalah metode pengujian perangkat lunak yang tes fungsionalitas dari aplikasi yang bertentangan dengan struktur internal atau kerja (lihat pengujian white-box). pengetahuan khusus dari kode aplikasi / struktur internal dan pengetahuan pemrograman pada umumnya tidak diperlukan. Uji kasus dibangun di sekitar spesifikasi dan persyaratan, yakni, aplikasi apa yang seharusnya dilakukan. Menggunakan deskripsi eksternal perangkat lunak, termasuk spesifikasi, persyaratan, dan desain untuk menurunkan uji kasus. Tes ini dapat menjadi fungsional atau non-fungsional, meskipun biasanya fungsional. Perancang uji memilih input yang valid dan tidak valid dan menentukan output yang benar. Tidak ada pengetahuan tentang struktur internal benda uji itu

2.3. *Framework*

Framework adalah sebuah kerangka kerja yang digunakan oleh para pengguna *software* untuk membuat serta mengembangkan aplikasi, juga terdapat perintah dasar yang umum digunakan dalam membangun sebuah *software* atau aplikasi dan diharapkan aplikasi tersebut dapat tersusun serta terstruktur dengan cukup rapi

Framework juga dapat diartikan sebagai komponen pemrograman yang telah jadi dan siap untuk kapan saja digunakan, sehingga pengembang tidak perlu lagi membuat script aplikasi dari awal. Dengan memakai *framework*, aplikasi yang dibuat dapat menjadi lebih terstruktur dan memiliki tingkat keamanan yang jauh lebih baik.

2.4. Pengujian Perangkat Lunak

Perangkat lunak pendukung yang digunakan dalam membangun sistem yaitu PHP dan MySQL.

2.4.1 PHP

PHP merupakan bahasa pemrograman Server Side yang digunakan sebagai bahasa pemrograman umum, dan juga untuk pengembangan sebuah situs web. Php juga biasa digunakan bersamaan dengan HTML.

2.4.2 MySQL

MySQL merupakan software sistem manajemen basis data SQL (*striktire query language*) atau DBMS yang *multi thread* dan *multi user*. PHP dan MySQL adalah pasangan sejati yang tak terpisahkan. Keduanya paling sering digunakan dalam pembuatan aplikasi berbasis web (*web application development*).

2.5 Kerangka Pikir



Gambar 2. 14. Kerangka Pikir

BAB III

METODE PENELITIAN

3.1 Objek Penelitian

Berdasarkan latar belakang dan kerangka pemikiran seperti yang telah diuraikan dalam Bab I dan II, maka menjadi objek penelitian adalah **Pengelompokan Data Penyakit Menular.**

3.2 Metode Penelitian

Pada penelitian ini untuk masalah Pengelompokan Data Penyakit Menular menggunakan metode *deskriptif* dimana metode ini bertujuan untuk mengungkapkan kebenaran yang *Objektif*. Dengan kata lain metode ini menjabarkan suatu keadaan atau fenomena yang terjadi saat ini dengan menggunakan prosedur ilmiah untuk menjawab masalah secara aktual.

3.2.1 Pengumpulan Data

Data primer pada penelitian ini adalah data masuk jumlah pasien penyakit menular dari Januari 2017 s/d Desember 2019 yang dikumpulkan menggunakan teknik dokumentasi dan observasi sedangkan data sekunder dikumpulkan menggunakan teknik dokumentasi. Berikut merupakan variable atau atribut dengan tipe datanya masing-masing.

Tabel 3.1 Atribut Data

NO	NAME	TYPE	VALUE	KETERANGAN
1	Periode	Date	2017	Variable Input
2	Jenis Penyakit Berdasarkan Penularannya	Integer	1 – 4	Variable Input
3	Jumlah Penyakit	Integer	1 – 4	Variable Input
4	Jumlah Pasien Penderita	Integer	50 – 100	Variable Output

Adapun variabel data pasien penderita penyakit menular dalam penelitian ini yaitu meliputi, Periode Merupakan variabel yang menunjukkan Tahun pengumpulan data, Jenis Penyakit Berdasarkan Penularannya Merupakan variabel

yan menunjukkan jenis atau nama penyakit berdasarkan cara penularannya, Jumlah Penyakit Merupakan variabel yang menunjukkan jumlah penyakit menular yang diderita pasien, Jumlah Pasien Penderita Merupakan variabel yang menunjukkan jumlah pasien penderita penyakit menular dari Tahun 2017.

3.2.2 Tahap Analisis

Analisis untuk Pengelompokan Data Penyakit Menular menggunakan metode *K-Means* sebagai berikut :

a) Analisis Sistem Berjalan

Rumah Sakit Zainal Umar Sidiki adalah satu-satunya rumah sakit yang berada di Kabupaten Gorontalo Utara, dalam operasionalnya rumah sakit mengumpulkan informasi melalui sistem komputerisasi menggunakan aplikasi *Microsoft Office* berguna untuk menampung dan mengelola data transaksi, sehingga data tersebut dapat di olah dan dapat diketahui tingkat pada volume suatu data pada suatu waktu tertentu. Terkadang hasil dari pengolahan data dengan menggunakan aplikasi *Microsoft Office* kesulitan untuk menganalisis data.

b) Analisis Sistem yang diusulkan

1. Data sekunder

Data sekunder merupakan data yang diperoleh dari studi literature yang dilakukan. Sumber dari studi literature yaitu jurnal, makalah ilmiah atau buku yang membahas tentang penelitian Algoritma *K-Means*.

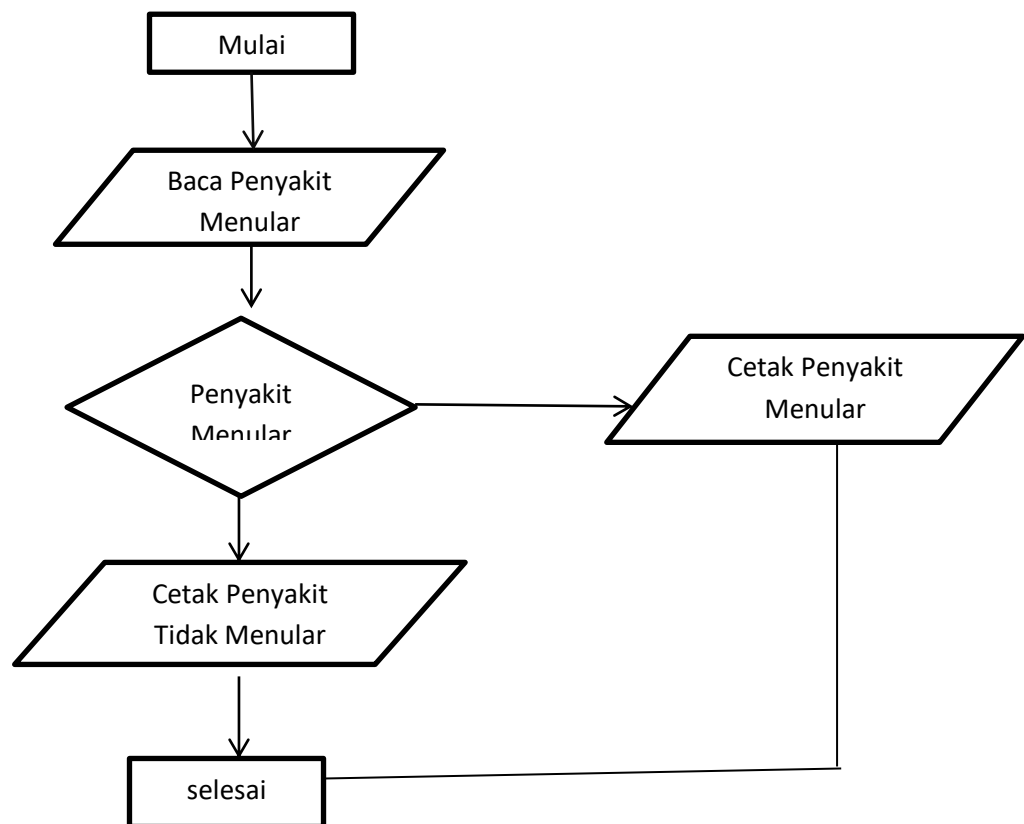
2. Data primer

Data primer adalah data yang diperoleh dari penelitian. Data primer pada penelitian ini yaitu berupa data penyakit menular, seperti TB, Kusta, Diare, ISPA, dan DBD.

Berikut ini merupakan tahap dalam melakukan pengelompokan data penyakit menular

1. Entri data : TB, Diare, ISPA, DBD
2. Proses : Pengelompokan
3. Laporan : Laporan Hasil Pengelompokan

Sistem yang diusulkan dapat digambarkan flowchart dokumen.



Gambar 3.1 Sistem Yang Diusulkan

3.2.3 Tahap Desain

Desain sistem menggunakan pendekatan berorientasi procedural/struktural

1. Desain *Output*

Desain *Output* yaitu untuk mengetahui bagaimana dan seperti apa bentuk *output* dari sistem yang akan dibuat. Desain *Output* terinci terbagi dua, yaitu desain *Output* dalam bentuk laporan media kertas dan *output* dalam bentuk dialog layar terminal (*Monitor*)

2. Desain *input*

Desain *input* atau masukan adalah awal dimulainya proses pengolahan data dan informasi. Bahan mentah dari informasi adalah data yang terjadi dari transaksi-transaksi yang dilakukan oleh konsumen. Data hasil transaksi merupakan data dari hasil dimasukan. Yang pertama kali dilakukan dimulai

dari desain dokumen dasar perangkat *input* jika dokumen dasar tidak di desain dengan baik, adapun kemungkinan *input* yang tercatat salah bahkan kurang.

3. Desain *Database*

Database atau Basis Data adalah kumpulan dari data yang saling berhubungan satu dengan yang lain, tersimpan disimpanan luar komputer dan digunakan perangkat lunak tertentu untuk memanipulasinya. *Database* merupakan salah satu komponen penting disistem informasi, karena berfungsi sebagai absis penyedia informasi bagi para pemakainya, penerapan *database* dalam aplikasi disebut *database system*.

4. Desain Teknologi

Pada tahap ini kita menentukan teknologi yang akan digunakan dalam menerima input, menjelaskan model, menyimpan dan mengakses data, menghasilkan dan mengirimkan keluaran dan membantu pengendalian dari sistem secara keseluruhan.

5. Desain Program

Pada tahap ini menggunakan alat bantu visio dalam bentuk Pseudocode program pada proses penerapan algoritma *K-Means*

3.2.4 Tahap Pembuatan

Adalah tahapan dimana kita melakukan pengembangan, melakukan tahap produksi sistem hasil analisa dan desain sistem sebelumnya. Termasuk didalamnya membangun sebuah sistem Pengelompokan Data Penyakit Menular menggunakan Metode *K-Means* menginstal paket tambahan untuk menjalankan program, menulis *Listing* program dan membangunnya dalam bentuk sebuah formulir, atarmuka atau integrasi sistem-sistem program yang terdiri dari *input*, *proses*, dan *output*, yang tersusun dalam sebuah sistem menu sehingga dapat dijalankan oleh pengguna sistem. Dalam tahapan ini penulis menggunakan bahasa pemrograman PHP, dan *Database MySQL Server*.

3.2.5 Tahap Pengujian

Setelah modul dibuat maka tahap selanjutnya yaitu tahap pengujian. Dalam tahap ini dapat berupa program yang dapat berjalan, dimana seluruh perangkat lunak, program tambahan dan semua program yang terlibat dalam pembangunan sistem yang diuji untuk memastikan sistem dapat berjalan sesuai

dengan rancangan atau belum. Pengujian dapat dilakukan dengan dua teknik, yaitu :

a. *White Box*

Dalam pengujian *White Box* ini dengan membuat bagan alir program, *Liating* program, grafik alir, pengujian *Basis Path* serta perhitungan *Ciclomatic Complexity*.

b. *Black Box*

Pengujian *Black Box*, pada tahap ini melakukan pengujian antarmuka sistem, sehingga pengguna dapat mengetahui apakah sistem tersebut bisa dioperasikan dengan baik atau tidak.

3.2.6 Tahap Implementasi

Tahap implementasi sistem (*System Implementation*) adalah tahap meletakan sistem supaya siap untuk dioperasikan di Rumah Sakit Zainal Umar Sidiki

BAB IV

HASIL PENELITIAN

4.1 Hasil Pengumpulan Data

Penelitian ini dilakukan di Rumah Sakit Zainal Umar Sidiki. dimana, rumah sakit ini merupakan satu dari sekian rumah sakit milik Pemkab Gorontalo utara. Peneliti mengambil judul ini karena masalah yang terjadi di rumah sakit tersebut yaitu kesulitan dalam mengelompokkan penyakit menular, pada anak-anak, remaja, dewasa, dan lansia. Sehingga peneliti memberikan suatu solusi terhadap permasalahan tersebut dengan membangun suatu sistem untuk pengelompokan data penyakit menular berdasarkan usia dengan tujuan merancang sistem pengolahan data penyakit menular menggunakan metode *K-Means Clustering* serta menerapkan hasil pengelompokan data penyakit menular menggunakan Algoritma *K-Means Clustering*.

Tabel 4.1 Hasil Pengumpulan Data

No	Nama	Tempat Tanggal Lahir	Jenis Kelamin	Penyakit	Cara Penularan	Tarif RS	Tahun
1	Della Dulanimo	Kwanda ng/ 21 Januari 1996	Perempu an	Turbekulosi s	Melalui Udara	560,400	2017
2	Yunus Dukalang	Alata/ 07 Juli 1999	Laki- Laki	Turbekulosi s	Melalui Udara	519,250	2017
3	Awaludin I Lihoe	Poso/ 16 Septemb er 2000	Laki- Laki	Ispa	Melalui Udara	501,250	2017
4	Mohamad Noka	Kwanda ng/ 22 April 2005	Laki- Laki	Diare	Air,Tanah,Ma kanan Terkontamina si Virus	507,250	2017
5	Feby Sione	Dambalo / 08 Oktober 1980	Perempu an	Ispa	Melalui Udara	393,500	2017

6	Radis Makoi	Leboto/ 18 Juli 2005	Laki- Laki	Dbd	Nyamuk Aedes Aegypti	619,250	2017
7	Fandi R. Hukkoli	Zuriyati/ 05/02/20 12	Laki- Laki	Ispa	Melalui Udara	851,100	2017
8	Rauf Auna	Molanta du/03/09 /2010	Laki- Laki	Diare	Air,Tanah,Ma kanan Terkontamina si Virus	206,450	2017
9	Reti Giola	Tolango/ 21-06- 1967	Perempu an	Turbekulosi s	Melalui Udara	395,200	2017
10	Rusni Ano	Ibarat/07 -08-1973	Perempu an	Diare	Air,Tanah,Ma kanan Terkontamina si Virus	839,550	2017
11
50	Aisa Arafa	Mootinel o/03/08/ 2011	Perempu an	DBD	Nyamuk Aedes Aegypti	70,300	2017

4.2 Hasil Pemodelan

4.2.1 Penjelasan Algoritma

Pada penelitian ini menggunakan algoritma *K-Means*, *K-Means* adalah metode penganalisaan data atau metode yang melakukan proses pemodelan tanpa supervise (*unsupervised*) dan merupakan salah satu metode pengelompokan data dengan sistem partisi. Metode ini mengelompokan data yang ada ke dalam beberapa kelompok, dimana yang berbeda dengan data yang ada di dalam kelompok yang lain. Tujuannya untuk meminimalisasikan *objective function* yang di set dalam proses *clustering*, yang pada dasarnya berusaha untuk meminimalkan variasi dalam satu cluster dan memaksimalkan variasi antar cluster.

Tahap-tahap yang dilakukan oleh algoritma *K-Means* adalah sebagai berikut:

1. Mulai
2. Menentukan banyaknya *cluster* dari data penyakit menular
3. Pengaturan untuk nilai awal titik tengah atau *centroid*, kemudian memilih pusat *cluster* secara acak pada data awal
4. Melakukan perhitungan data penyakit menular ke *centroid* dengan memakai rumus jarak *Euclid*
5. Melakukan *clustering* pada data memasukan setiap objek ke dalam *cluster* atau grup berdasarkan jarak minimumnya
6. Menghitung pusat cluster baru jika ada data yang harus di pindah. Pusat cluster baru ditentukan menurut pengelompokan anggota tiap-tiap *cluster* baru, dan untuk *cluster* baru yang pertama dihitung berdasarkan rata-rata koordinat. Jika hasil perhitungan menunjukkan adanya angka pusat *cluster* yang sama maka pengulangan dihentikan.
7. Selesai

Tabel 4.2 Data Pasien

No	Nama	Tempat Tanggal Lahir	Jenis Kelamin	Penyakit	Cara Penularan	Tarif RS	Tahun
1	Della Dulanimo	Kwand ang/ 21 Januari 1996	Perempu an	Turbekulosi s	Melalui Udara	560,400	2017
2	Yunus Dukalang	Alata/ 07 Juli 1999	Laki- Laki	Turbekulosi s	Melalui Udara	519,250	2017

3	Awaludin I Lihoe	Poso/ 16 September 2000	Laki-Laki	Ispa	Melalui Udara	501,250	2017
4	Mohamad Noka	Kwandang/ 22 April 2005	Laki-Laki	Diare	Air,Tanah,Makanan Terkontaminasi Virus	507,250	2017
5	Feby Sione	Dambalo/ 08 Oktober 1980	Perempuan	Ispa	Melalui Udara	393,500	2017
6	Radis Makoi	Leboto/ 18 Juli 2005	Laki-Laki	Dbd	Nyamuk Aedes Aegypti	619,250	2017
7	Fandi R. Hukkoli	Zuriyati /05/02/ 2012	Laki-Laki	Ispa	Melalui Udara	851,100	2017
8	Rauf Auna	Molantadu/03/09/2010	Laki-Laki	Diare	Air,Tanah,Makanan Terkontaminasi Virus	206,450	2017
9	Reti Giola	Tolango/ 21-06-1967	Perempuan	Turbekulosis	Melalui Udara	395,200	2017
10	Rusni Ano	Ibarat/07-08-1973	Perempuan	Diare	Air,Tanah,Makanan Terkontaminasi Virus	839,550	2017

Tabel 4.3 Centroid Yang Dipilih

NO	Centroid	Jenis Kelamin	Penyakit	Tarif Rs
1	C 1	1	2	519.250
2	C 2	1	4	507.250
3	C 3	2	4	839.550

4.2.2 Perhitungan Algoritma

- a. Berikut merupakan cara untuk melakukan pengelompokan menggunakan algoritma *K-Means Clustering*:

$$d(x_i, x_j) = (|x_{i1} - x_{j1}|^g + |x_{i2} - x_{j2}|^g + \dots + |x_{ip} - x_{jp}|^g)^{1/g}$$

Di mana :

$g = 1$, untuk menghitung jarak *Manhattan*

$g = 2$, untuk menghitung jarak *Euclidean*

$g = \infty$, untuk menghitung jarak *Chebychev*

x_i, x_j adalah dua buah data yang akan dihitung jaraknya

p = dimensi dari sebuah data

- b. Berikut adalah pembaharuan suatu titik *centroid* yang dapat dilakukan dengan rumus berikut :

$$\mu_k = \frac{1}{N_k} \sum_{q=1}^{N_k} x_q$$

Di mana:

μ_k = titik centroid dari cluster ke-K

N_k = banyaknya data pada cluster ke-K

x_q = data ke-q pada cluster ke-K

- c. Berikut adalah rumus yang di gunakan dengan menggunakan persamaan *Euclidean Distance Space* :

$$C1 = \sqrt{(x1 - y1)^2 + (x2 - y2)^2 + (x3 - y3)^2}$$

$$C2 = \sqrt{(x1 - y1)^2 + (x2 - y2)^2 + (x3 - y3)^2}$$

$$C3 = \sqrt{(x1 - y1)^2 + (x2 - y2)^2 + (x3 - y3)^2}$$

Keterangan :

x= data record

y= data centroid

Berikut cara kerja perhitungan manual sbb :

$$\begin{aligned}
 C1, (1). \quad & \sqrt{(x1 - y1)^2 + (x2 - y2)^2 + (x3 - y3)^2} \\
 & \sqrt{(2 - 1)^2 + (2 - 2)^2 + (560.400 - 519.250)^2} \\
 & \sqrt{(1)^2 + (0)^2 + (41.150)^2} \\
 & \sqrt{1 + 0 + 1.693.322.500} \\
 & \sqrt{1.694.322.501} \\
 & = 41150.00001
 \end{aligned}$$

$$\begin{aligned}
 C1, (2). \quad & \sqrt{(x1 - y1)^2 + (x2 - y2)^2 + (x3 - y3)^2} \\
 & \sqrt{(1 - 1)^2 + (2 - 2)^2 + (519.250 - 519.250)^2} \\
 & \sqrt{(0)^2 + (0)^2 + (0)^2} \\
 & \sqrt{0 + 0 + 0} \\
 & \sqrt{0} \\
 & = 0
 \end{aligned}$$

$$\begin{aligned}
 C1, (3). \quad & \sqrt{(x1 - y1)^2 + (x2 - y2)^2 + (x3 - y3)^2} \\
 & \sqrt{(1 - 1)^2 + (1 - 2)^2 + (501.250 - 519.250)^2} \\
 & \sqrt{(10 + (0 - 1) + (-18.000))^2} \\
 & \sqrt{0 + 1 + 324.000.000} \\
 & \sqrt{324.000.001} \\
 & = 18000.00003
 \end{aligned}$$

$$\begin{aligned}
\text{C2, (1).} \quad & \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2(x_3 - y_3)^2} \\
& \sqrt{(2 - 1)^2 + (2 - 4)^2 + (560.400 - 507.250)^2} \\
& \sqrt{(1)^2 + (-2)^2 + (53.150)^2} \\
& \sqrt{1 + 4 + 2.824.922.500} \\
& \sqrt{2.824.922.505} \\
& = 53150.00005
\end{aligned}$$

$$\begin{aligned}
\text{C2, (2).} \quad & \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2(x_3 - y_3)^2} \\
& \sqrt{(1 - 1)^2 + (2 - 4)^2 + (519.250 - 507.250)^2} \\
& \sqrt{(0) + (-2)^2 + (12.000)^2} \\
& \sqrt{0 + 4 + 144.000.000} \\
& \sqrt{144.000.004} \\
& = 12000.00017
\end{aligned}$$

$$\begin{aligned}
\text{C2, (3).} \quad & \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2(x_3 - y_3)^2} \\
& \sqrt{(1 - 1)^2 + (1 - 4)^2 + (501.250 - 507.250)^2} \\
& \sqrt{(0) + (-3)^2 + (-6.000)^2} \\
& \sqrt{0 + 9 + 36.00.000} \\
& \sqrt{36.00.009} \\
& = 6000.00075
\end{aligned}$$

$$\begin{aligned}
\text{C3, (1).} \quad & \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2(x_3 - y_3)^2} \\
& \sqrt{(2 - 2)^2 + (2 - 24)^2 + (560.400 - 839.550)^2} \\
& \sqrt{(0)^2 + (-2)^2 + (-279.150)^2} \\
& \sqrt{0 + 4 + 77.924.722.500} \\
& \sqrt{77.924.722.504} \\
& = 279.150
\end{aligned}$$

$$\begin{aligned}
C3, (2). \quad & \sqrt{(x1 - y1)^2 + (x2 - y2)^2 + (x3 - y3)^2} \\
& \sqrt{(1 - 2)^2 + (2 - 4)^2 + (519.250 - 839.550)^2} \\
& \sqrt{(-1)^2 + (-2)^2 + (-320300)^2} \\
& \sqrt{1 + 4 + 102.592.090.000} \\
& \sqrt{102.592.090.005} \\
& = 320.300
\end{aligned}$$

$$\begin{aligned}
C3, (3). \quad & \sqrt{(x1 - y1)^2 + (x2 - y2)^2 + (x3 - y3)^2} \\
& \sqrt{(1 - 2)^2 + (1 - 4)^2 + (501.00 - 839.550)^2} \\
& \sqrt{(-1)^2 + (-3)^2 + (338.3000)^2} \\
& \sqrt{1 + 9 + 114.446.890.000} \\
& \sqrt{114.446.890.010} \\
& = 338.300
\end{aligned}$$

Jarak hasil perhitungan akan dilakukan perbandingan dan dipilih jarak terdekat antara data dengan pusat cluster. Jarak ini menunjukkan bahwa data tersebut berada dalam satu kelompok dengan pusat *cluster* terdekat.

Tabel 4.1 Tabel Hasil Iterasi 1

Nama	C1	C2	C3	Jarak Terdekat	Keanggotaan Cluster
Della Dulanimo	41150	53150.00005	279150	41150	C1
Yunus Dukalang	0	12000.00017	320300	0	C1
Awaludin I Lihoe	18000.00003	6000	338300	6000	C2
Mohamad Noka	12000.00017	0	332300	0	C2
Feby Sione	125750	113750	446050	113750	C2

Radis Makoi	100000	112000	220300	100000	C1
Fandi R. Hukkoli	331850	343850	11550	11550	C3
Rauf Auna	312800	300800	633100	300800	C2
Reti Giola	124050	112050	444350	112050	C2
Rusni Ano	320300	332300	0	0	C3

Tabel 4.5 Tabel Hasil Iterasi 2

Nama	C1	C2	C3	Min Jarak	Keanggotaan Cluster
Della Dulanim	5900	159670	284925	5900	C1
Yunus Dukalang	47050	118520	326075	47050	C1
Awaludin I Lihoe	65050	100520	344075	65050	C1
Mohamad Noka	59050	106520	338075	59050	C1
Feby Sione	172800	7230	451825	7230	C2
Radis Makoi	52950	218520	226075	52950	C1
Fandi R. Hukkoli	284800	450370	5775	5775	C3
Rauf Auna	359850	194280	638875	194280	C2
Reti Giola	171100	5530	450125	5530	C2
Rusni Ano	273250	438820	5775	5775	C3

Tabel 4.6 Tabel Hasil Iterasi 3

Nama	C1	C2	C3	Jarak Terdekat	Keanggotaan Cluster
Della Dulanim o	18920	228684	284925	18920	C1
Yunus Dukalan g	22230	187534	326075	22230	C1
Awaludi	40230	169534	344075	40230	C1

n I Lihoe					
Mohama d Noka	34230	175534	338075	34230	C1
Feby Sione	147980	61784	451825	61784	C2
Radis Makoi	77770	287534	226075	77770	C1
Fandi R. Hukkoli	309620	519384	5775	5775	C3
Rauf Auna	335030	125266	638875	125266	C2
Reti Giola	146280	63484	450125	63484	C2
Rusni Ano	298070	507834	5775	5775	C3

Tabel 4.7 Tabel Hasil Iterasi 4

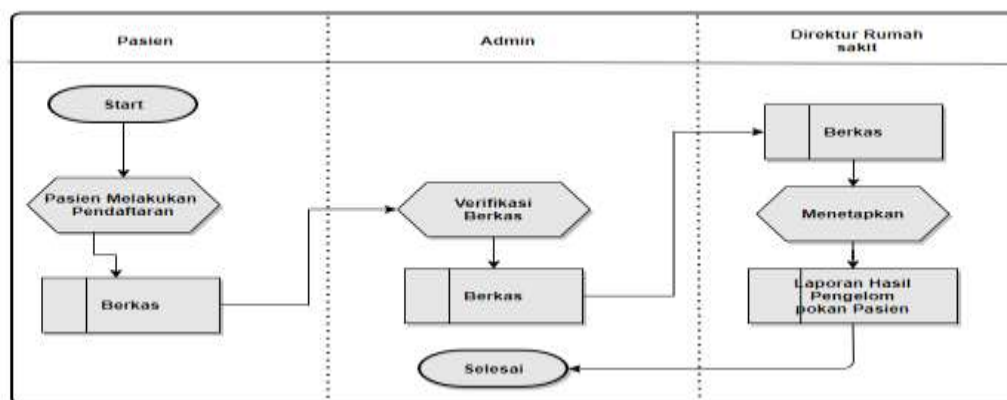
Nama	C1	C2	C3	Jarak Terdekat	Keanggotaan Cluster
Della Dulanim o	18920	228684	284925	18920	C1
Yunus Dukalan g	22230	187534	326075	22230	C1
Awaludi n I Lihoe	40230	169534	344075	40230	C1
Mohama d Noka	34230	175534	338075	34230	C1
Feby Sione	147980	61784	451825	61784	C2
Radis Makoi	77770	287534	226075	77770	C1
Fandi R. Hukkoli	309620	519384	5775	5775	C3
Rauf Auna	335030	125266	638875	125266	C2
Reti Giola	146280	63484	450125	63484	C2
Rusni Ano	298070	507834	5775	5775	C3

Dari hasil iterasi di atas dapat diketahui bahwa, *cluster* 1 didominasi oleh remaja dengan penyakit menular TB, DBD dan Diare dan *cluster* 2 oleh anak-anak dengan penyakit ISPA dan Diare sedangkan *cluster* 3 didominasi oleh Lanjut usia dengan penyakit Tuberkulosis.

4.3 Hasil Analisis Sistem

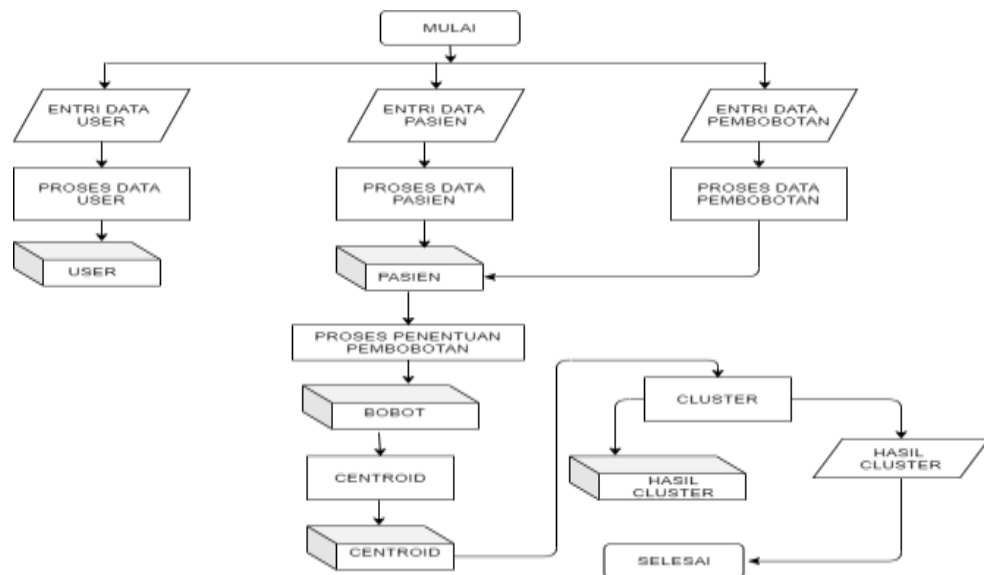
Analisis sistem adalah tahap penguraian komponen-komponen dari suatu sistem pendukung keputusan. Tahap analisis dilakukan setelah tahap perencanaan sistem (*System Planning*) dan sebelum tahap desain sistem (*System Design*). Tahap ini sangat diperlukan dalam mengetahui sejauh mana keputusan yang diambil tersebut digunakan serta mengidentifikasi dan mengevaluasi permasalahan maupun hambatan yang terjadi pada sistem tersebut serta mampu menjelaskan keseluruhan proses yang didukung oleh fakta dan data secara utuh.

4.3.1 Proses Sistem Berjalan



Gambar 4.1 Bagan Alir Dokumen

4.3.2 Proses Sistem Diusulkan

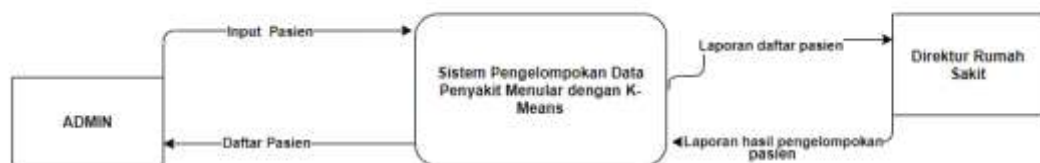


Gambar 4.2 Sistem Yang Diusulkan

4.4 Hasil Pengembangan sistem

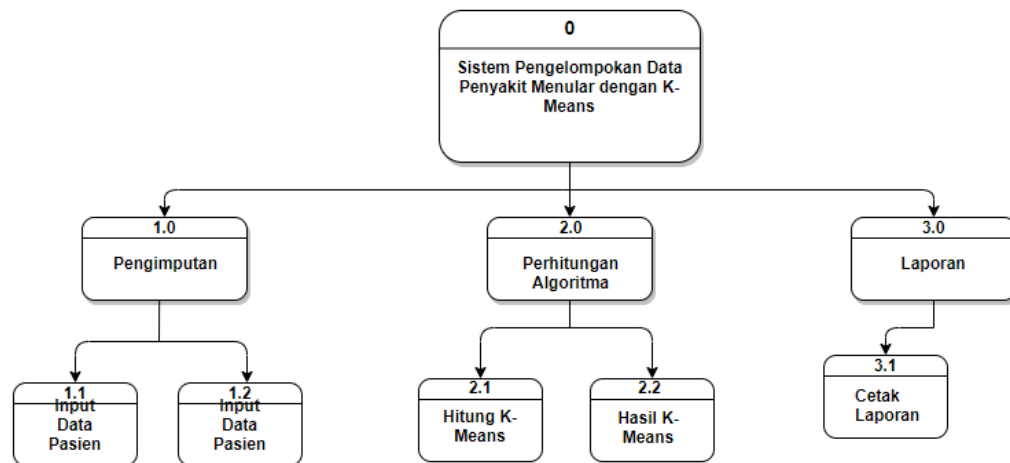
4.4.1. Diagram Konteks

Sistem Pengolahan Data Penyakit Menular Dengan Algoritma *K-Means*



Gambar 4.3 Diagram Konteks

4.4.2. Diagram Berjenjang

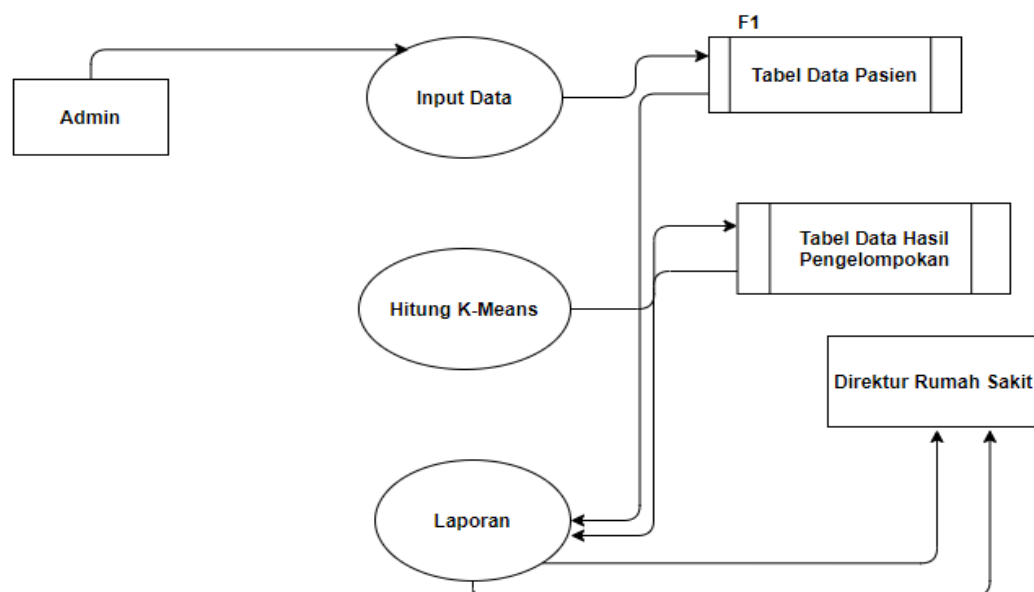


Gambar 4.4 Diagram Berjenjang

4.4.3 Diagram Arus Data

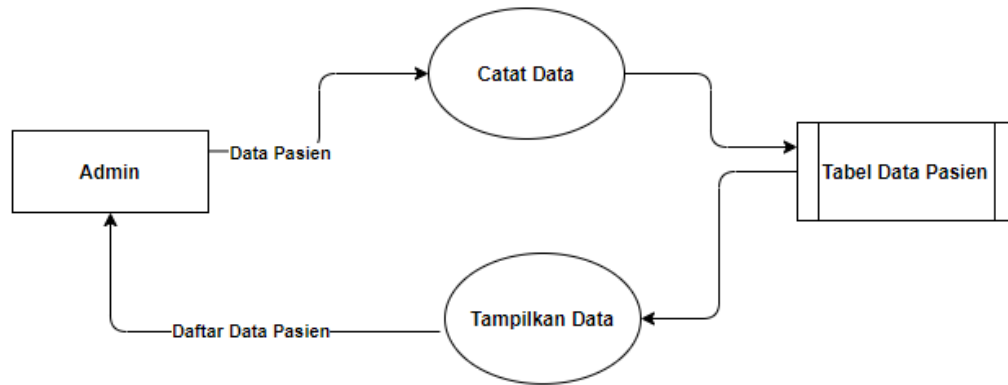
Diagram arus data merupakan suatu proses atau sistem penggambaran aliran data diagram

4.4.4 Diagram Arus Data Level 0



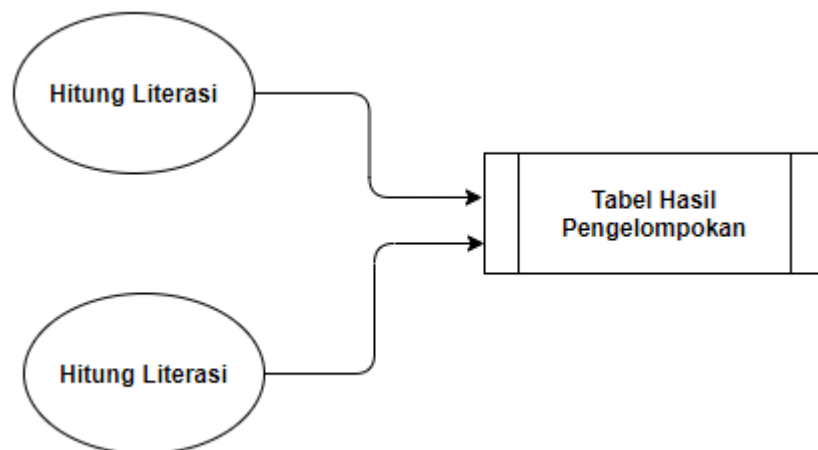
Gambar 4.5 DAD Level 0

4.4.5 DAD Level 1 Proses 1



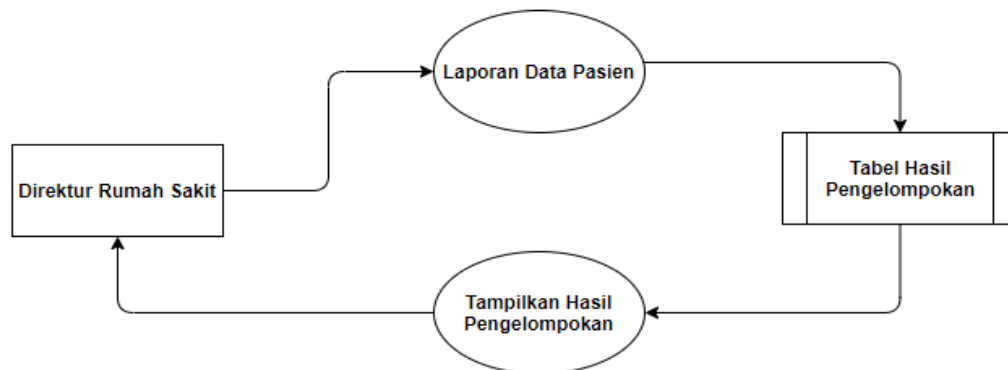
Gambar 4.6 DAD Level 1 Proses 1

4.4.6 DAD Level 1 Proses 2



Gambar 4.7 DAD Level 1 Proses 2

4.4.7 DAD Level 1 Proses 3



Gambar 4.8 DAD Level 1 Proses 3

4.5 Kamus Data

Kamus data merupakan suatu data yang ada di dalam database atau daftar data elemen yang terorganisir dengan definisi sesuai dengan sistem, data digunakan merancang file input dan output database, kamus data dibuat berdasarkan arus data yang mengalir pada DAD, dimana didalamnya tersdapat struktur dari arus data secara detail.

4.5.1 Kamus Data Terinci

Tabel 4.8 Kamus Data User

Kamus Data Pasien				
Nama Arus Data : Pasien			Bentuk Data : D	
Penjelasan : Input Data Pasien			Arus Data : b-1, 1 1-f1,f1,-2	
Periode Struktur : Setiap Ada penambahan data nilai			b-1. 1p,1.p-f1,f1-1,1.2p	
No	Field Name	Type	Size	Ket
1	Id	N	5	Kode Id
2	User Name	C	25	User Name
3	Password	C	20	Password
4	Fulname	C	35	Nama Lengkap

Tabel 4.9 Kamus Data Pasien

Kamus Data Pasien				
Nama Arus Data : Pasien Penjelasan : Input Data Pasien Periode Struktur : Setiap Ada penambahan data nilai			Bentuk Data : Dokumen Arus Data : b-1, 1-f2,f2f-2, b-1,2p, 1 2p-f2,f2-13p,b-2 1p,b-2,2p,f2-3p	
No	Field Name	Type	Size	Ket
1	Pasien	N	5	Kode Pasien
2	Nama Pasien	C	25	Nama Pasien
3	Alamat	C	20	Alamat
4	Penyakit	C	35	Kode Penyakit

Tabel 4.10 Pembobotan

Kamus Data Pembobotan				
Nama Arus Data : Pasien Penjelasan : Input Data Pasien Periode Struktur : Setiap Ada penambahan data nilai			Bentuk Data : D Arus Data : b-1, 1-f3,f3,-2,2- f5. B-2 1p,2 1p-f4,f4-3p	
No	Field Name	Type	Size	Ket
1	Id Pembobotan	N	5	Kode Id Pembobotan
2	Nama Pembobotan	C	25	Nama Pembobotan

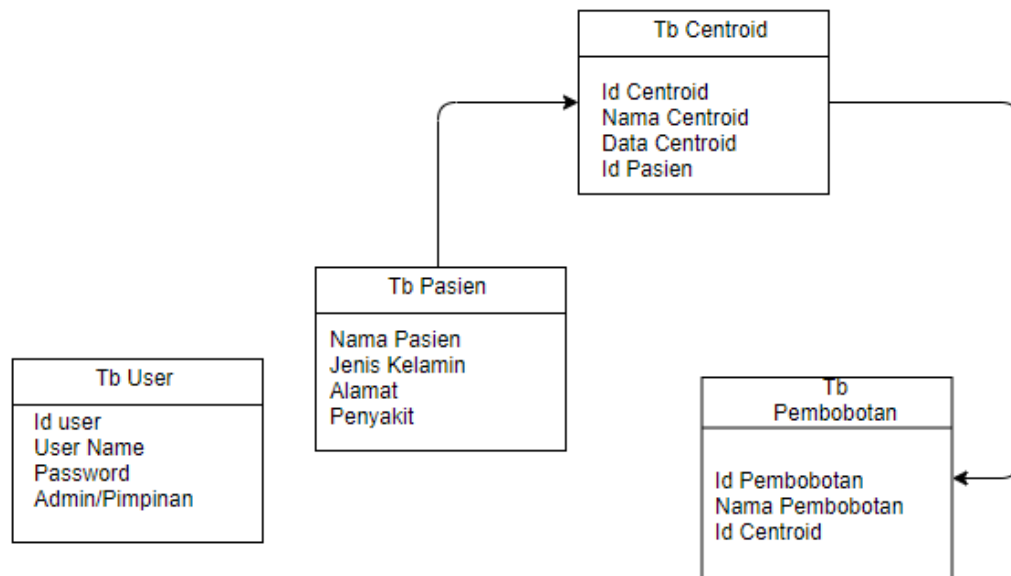
Tabel 4.11 Kamus Data Centroid

Kamus Data Centroid				
Nama Arus Data : Pasien Penjelasan : Input Data Pasien Periode Struktur : Setiap Ada penambahan data nilai			Bentuk Data : D Arus Data : b-1, 1-f3,f3f-2, b-2,2p,2-f5 1 2p-f4,f4-13p,b-2 2p,b-2,2p,f5-3p	
No	Field Name	Type	Size	Ket
1	Id Centroid	N	30	Kode Id Pasien
2	Nama Centroid	N	30	User Name
3	Data Centroid	N	30	Password

Tabel 4.12 Kamus Data Hasil Clustering

Kamus Data Hasil Clustering				
Nama Arus Data : Pasien Penjelasan : Input Data Pasien Periode Struktur : Setiap Ada penambahan data Pasien			Bentuk Data : D Arus Data : b-1, 1-f3,f3,-2, 2-f6. b-2 ,1.p2 1p-f4,f4 -2.2p, 2.2p-f6, f6-3.p	
No	Field Name	Type	Size	Ket
1	Cluster 1	N	25	Kode Id
2	Cluster 2	N	25	User Name
3	Cluster 3	N	25	Password

4.6 Database



Gambar 4.9 Database

4.6.1 Daftar Input Yang Didesain

Daftar Input Yang Didesain Untuk : Rumah Sakit Zainal Umar Sidiki Tahap : Desain Sistem			
Kode Input	Nama Input	Simbol Input	Periodik
1-001	Data User	Admin	Non Periodik
2-002	Data Pasien	Admin	Non Periodik
3-003	Hasil Pembobotan	Admin	Non Periodik
4-004	Centroid	Admin	Non Periodik
5-005	Hasil Clustering	Admin	Non Periodik

4.6.2 Desain File Secara Umum

Kode File	Nama File	Field	Type	File
F1	User	30	C	Id_User
F2	Data Pasien	20	C	Id_Data Pasien
F3	Pembobotan	6	N	Id_Pembobotan
F4	Centroid	30	N	Id_Centroid
F5	Hasil Clustering	20	N	Id_Hasil Clustering

4.7 Arsitektur Sistem

Sistem pengelompokan dengan algoritma *K-Means clustering* sedangkan spesifikasi menggunakan *hardware* dan *software* yang direkomendasikan Yaitu:

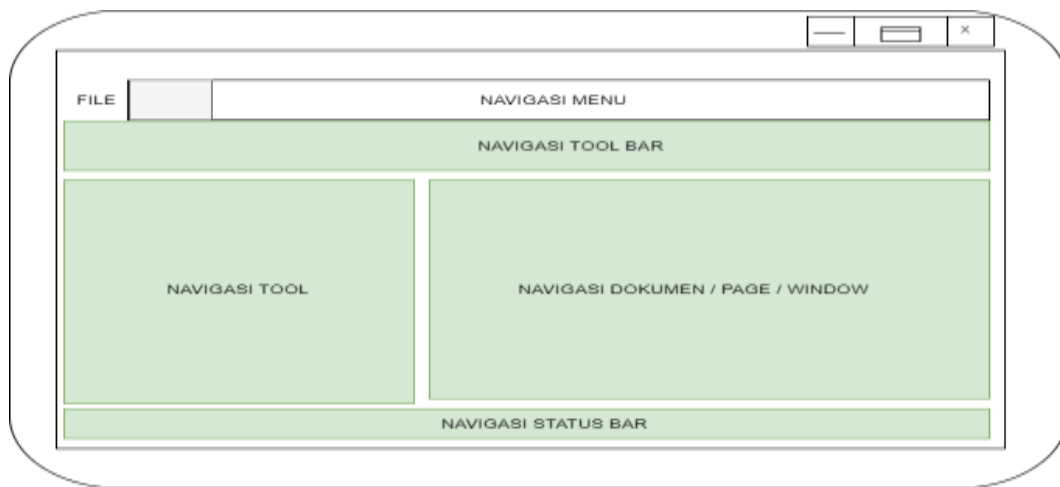
1. Preprocessor : Intel(R) Coleron(R) N400 CPU @ 1.10GHz 1.10 GHz
2. RAM : 4.00 GB 93.83 GB usable)
3. VGA : Intel(R) UHD Graphics 600
4. Hardisk : 64-bit Operating System, x64-baset Processor
5. Operating System : Windows 10
6. Tools : Microsoft Edge

4.8 Interface Design Mekanisme User

User	Kategori	Akses Input	Akses Output
Admin	Administrator	All	All
Pimpinan	Direktur Rumah Sakit	All	All

Gambar 4.10 Mekanisme User

4.9 Interface Design : Mekanisme Navigasi



Gambar 4.11 Mekanisme Navigasi

4.10 Interface Design : Mekanisme Output-Login

The diagram illustrates a login interface design within a window titled **LOGIN**. The window contains the following elements:

- Pick id**: A text input field with a dropdown arrow icon on the right side.
- Enter Password**: A text input field located below the ID field.
- Retry**: A button located below the input fields.
- OK**: A button located to the right of the Retry button.

Gambar 4.12 Mekanisme Output-Login

4.11 Infrface Design : Mekanisme Output-Laporan User

LOGO		LAPORAN AKTIVITAS PT. XXX	
NO	TGL	AKTIFITAS	WAKTU PEMAKAIAN
USER..... (GROUP BY USER)			
00000001	17 FEB 2021	MERUBAH DATA USER XXXX	3 JAM MENIT
SUB TOTAL		XXX	
GRAND TOTAL		XXX	
GORONTALO, dd/MMM/yyyy Mengetahui XXX			
page N of M			

Gambar 4.13 Mekanisme Output-Laporan User

4.12 Data Design : Struktur Data-User

Nama : tblUser.mdf Type : Transaksi Primery Key : usld Forigen Key : - Media : Hardisk Fungsi : Merupakan data pengguna aplikasi Struktur Data :					
NO	FIELD	TYPE	SIZE	RANGE	KETERANGAN
1	Usld	Char	4	3	ID User
2	usNama	Varchar	32	50	Nama User
3	usPasword	Varchar	32	50	Pasword User
4	usEmail	Varchar	32	50	Email User
5	usTglRegis	DataTime	8	-	Tgl registrasi
6	usTglLogin	DataTime	8	-	Tgl terakhir Login

7	usAsk	Varchar	32	50	Pertanyaan
8	usAnswer	Varchar	32	50	Jawaban
9	usAktif	Boolean	4	1	Apakah aktif
10	usValid	Boolean	4	1	Apakah sdh valid

4.13 Program Desing

CLASS/TYPE	ATTRIBUTES [TYPE]	METHODS [EVENT or TYPE]
FrmMain	AktifDokument [String] Home [Menu] Logout [Menu] Prediksi [Menu] Setting [Menu] Report [Menu] Help [Menu] About [Menu] Add [Toolbar] Edit [Toolbar] Delete [Toolbar] Save [Toolbar] Cancel [Toolbar] Print [Toolbar] Search [Toolbar] UserActive [StatusBar] DataTime [StatusBar] SystemStatus [StatusBar]	frmMain [Load] frmMain [Closing] Home [Click] Logout [Click] Prediksi [Click] Setting [Click] Report [Click] Help [Click] About [Click] Add [Click] Edit [Click] Delete [Click] Save [Click] Cancel [Click] Print [Click] Search [Click] LoadData() [Boolean]
frmLogin	ID [TextBox] Pasword [TextBox] OK [Button]	frmMain [Load] frmMain [Closing] OK [Click]

	Retry [Button]	Retry [Click] Password [Enter]
FrmUser	ID [TextBox] Name [TextBox] Password [TextBox] Email [TextBox] TglGabung [DateTimeEditor] TglLogin [DateTimeEditor] Ask [TextBox] Answer [TextBox] Valid [OptionButton] ViewData [GrindView]	frmUser [Load] frmUser [Closing] ViewData [RowsChanged] Add()[Boolean] Edit() [Boolean] Delete ()[Boolean] Cancel()[Boolean] Retry()[Boolean] Save()[Boolean] Print()[Boolean] Search(x : String) [Boolean]
frmAsosiasi	Items [ComboBox] ListItems [ListBox] ViewData [GridView] minSupport [Integer] minConfidence [Integer] Process [Button]	... Process [Click] SetItems(x : Integer) [Integer] ...
frmReport	ViewReport [ReportViewer] Filter [TextBox] Print [Toolbar] Export [Toolbar] Refresh [Toolbar]
Interface1
Module1
Dataset1
....
N

4.14 Hasil Kontruksi Sistem

Hasil analisis dan desain pada kontruksi sistem diterjemahkan kedalam kontruksi sistem/software dengan menggunakan bahasa pemrograman PHP. Berikut merupakan alat bantu yang digunakan pada tahap ini :

1. Programing PHP untuk pemrogramanya
2. Database Mysql untuk tempat penyimpanan database
3. Visual Studio Code untuk editor

4.15 Kode Program Pengujian *WhiteBox*

```
//centroid baru 1.a
$jum = 0;
$arr = array();
for($i=0;$i<count($arr_c1);$i++)
{
    $arr[$i] = $arr_c1_temp[$i]*$arr_c1[$i];
    if($arr_c1[$i]==1)
    {
        $jum++;
    }
}
$c1a_b = array_sum($arr)/$jum;

//centroid baru 1.b
$jum = 0;
$arr = array();
for($i=0;$i<count($arr_c2);$i++)
{
    $arr[$i] = $arr_c2_temp[$i]*$arr_c1[$i];
    if($arr_c1[$i]==1)
    {
        $jum++;
    }
}
$c1b_b = array_sum($arr)/$jum;

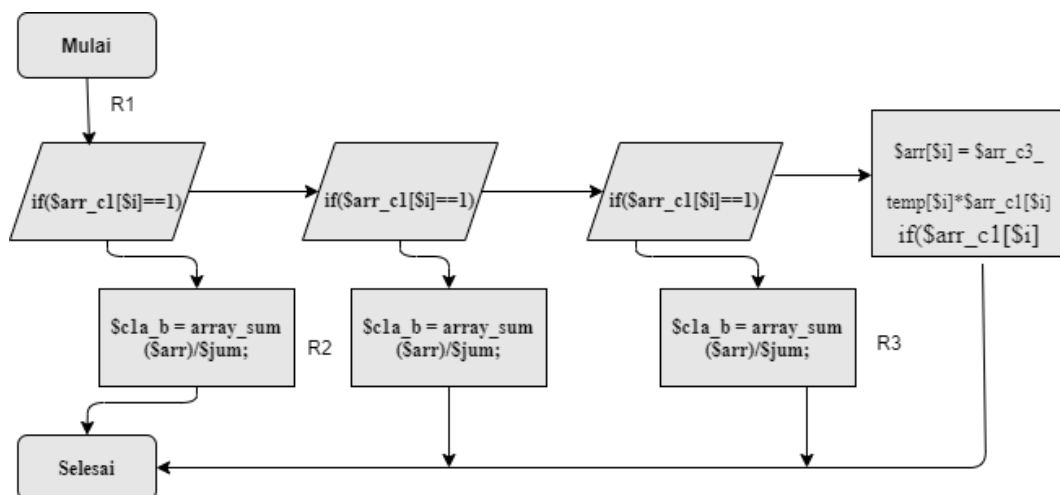
//centroid baru 1.c
$jum = 0;
$arr = array();
for($i=0;$i<count($arr_c3);$i++)
{
    $arr[$i] = $arr_c3_temp[$i]*$arr_c1[$i];
    if($arr_c1[$i]==1)
    {
```

```

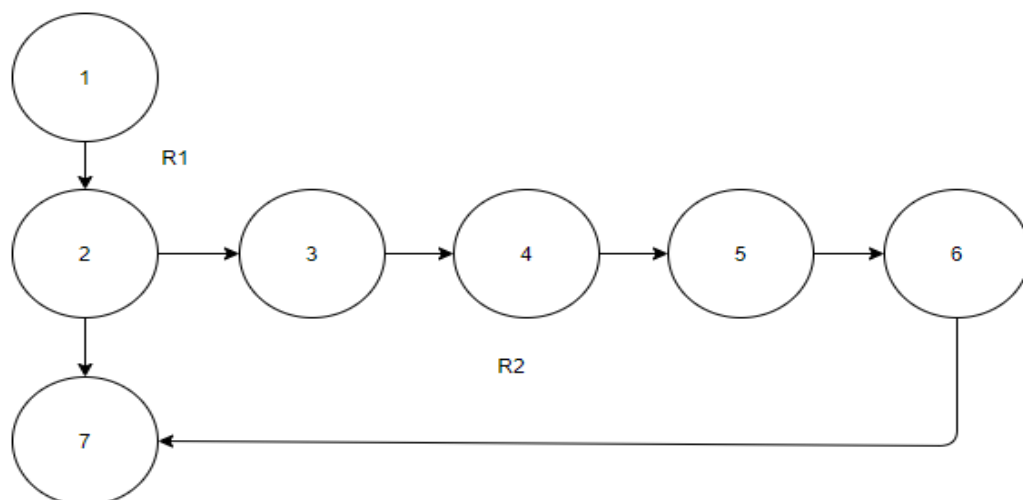
    $jum++;
  }
}
$c1c_b = array_sum($arr)/$jum;

```

4.16 Flowchart Program Pengujian *Whitebox*



4.17 Flowgrap Program Pengujian *Whitebox*



4.18 Perhitungan CC Pada Pengujian *Whitebox*

Pada Flowgraph tersedia *cyclomatic complexity* dari sebuah program dapat dibuat dengan menggunakan rumus di bawah ini :

$$V(G) = E - N + 2$$

Keterangan :

$V(G)$: *cyclomatic complexity*

E : Total Jumlah Edge

N : Total Jumlah Node

P : Predikat Node

Contoh *Flowgraph* diatas dapat dihitung *cyclomatic complexity* sebagai berikut :

Diketahui : Region = 2

Node (N) = 7

Edge (E) = 7

Predikat Node (P) = 4

$$V(G) = (E - N) + 2$$

$$V(G) = P + 1 = 1 + 1 = 2$$

$$V(G) = (7 - 7) + 2 = 2$$

4.19 Path Pada Pengujian *Whitebox*

NO	PATH	KET
1	1-2-3	OK
2	1—2-3.....7	OK

4.20 Hasil Pengujian *Blackbox*

Input/Even	Fungsi	Hasil	Hasil Uji
Login	Login Dengan Input Username Dan Password	-Jika Password Salah Maka Masukan Username Dan Password -Jika Password Benar Akan Masuk Ke Sistem	Selesai
Menu Home	Menampilkan Halaman Admin	Halaman Admin Ditampilkan	Selesai
Pilih Menu Home Page	Menampilkan Halaman Home Page	Halaman Home Page Ditampilkan	Selesai
Pilih Menu Profil	Menampilkan	Halaman Profil Data	Selesai

	Halaman Profil Data	Ditampilkan	
Pilih Menu Semua Data	-Menampilkan Halaman Tabel Data Input - Menampilkan Halaman Input Cluster	-Halaman Tabel Input Data Ditampilkan -Halaman Tabel Input Cluster Ditampilkan	Selesai
Pilih Menu Hasil Clustering	Menampilkan Halaman Hasil Cluster	Halaman Hasil Cluster Ditampilkan	Selesai
Pilih Menu Hasil Diagram	Menampilkan Halaman Hasil Diagram	Halaman Hasil Diagram Ditampilkan	Selesai
Pilih Menu Logout	Kembali Kehalaman Login	Halaman Login Ditampilkan	Selesai
Pilih Menu Edit	Menampilkan Halaman Edit Data	Halaman Edit Data Ditampilkan	Selesai
Pilih Menu Hapus	Menampilkan Halaman Hapus Data	Halaman Hapus Data Ditampilkan	Selesai
Pilih Menu Tambahkan	Menampilkan Halaman Tambah Data	Halaman Tambah Data Ditampilkan	Selesai`

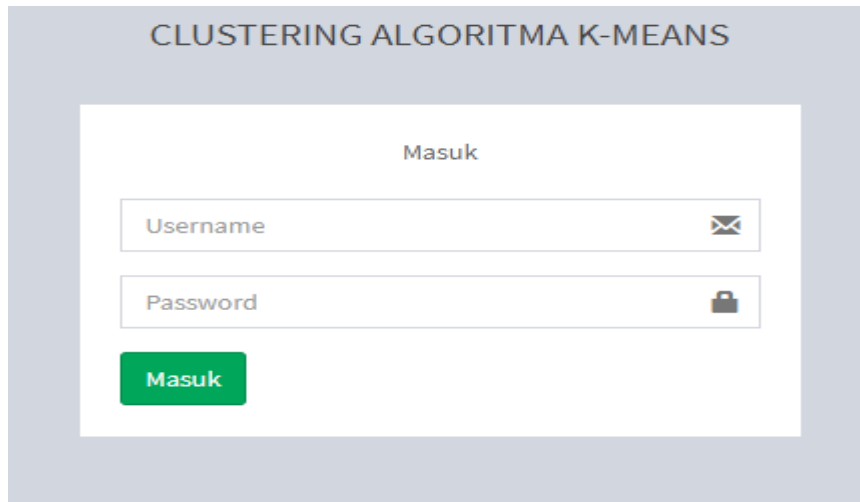
BAB V

PEMBAHASAN

5.1 Pembahasan Sistem

Berikut merupakan tampilan sistem *Clustering* Penerapan Algoritma *K-Means* Dalam Mengelompokkan Data Penyakit Menular

5.1.1 Hasil Tampilan Halaman Login



Gambar 5.1 Tampilan Login

Halaman ini digunakan untuk masuk kehalaman admin. Dengan memasukan Username dan Password, selanjutnya menekan tombol Masuk untuk melanjutkan proses.

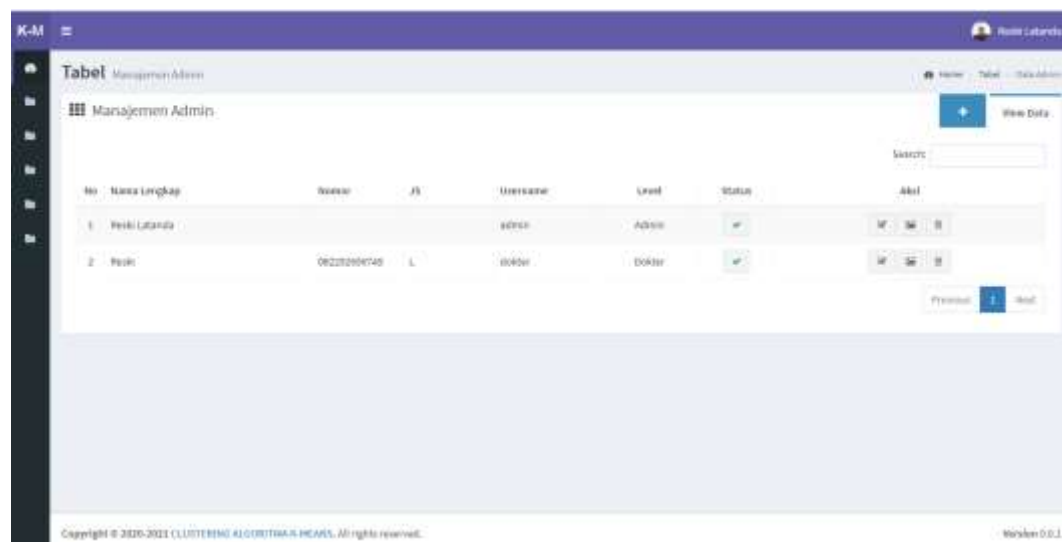
5.1.2 Hasil Tampilan Halaman Dashbord



Gambar 5.2 Tampilan Halaman Dashbord

Halaman beranda ini terdiri dari Navigasi Menu Pasien Penderita Penyakit Menular dan Menu Di samping Kiri terdiri dari Data Referensi, Data Analisa, Data Laporan Dan Menu Administrator.

5.1.3 Tampilan Halaman Administrator



Gambar 5.3 Tampilan Halaman Administrator

Halaman ini merupakan tampilan halaman Manajemen Administrator yaitu profil data Admin dan Pimpinan (Dokter)

5.1.4 Tampilan Halaman Data Pasien

The screenshot shows the 'Tabel Data Pasien' page. It features a table with columns: No, Nama Lengkap, JK, Alamat, Umur, Gejala Penyakit, Diagnosis, and Aksi. There are 10 rows of data. The first row is for 'Della Ouliana' with age 15, address 'Kusambi', symptom 'Malaria', and diagnosis 'Tifus'. The second row is for 'Beti Dinda' with age 24, address 'Tidung', symptom 'Malaria', and diagnosis 'Tifus'. The third row is for 'Nuruz Dzulqarnain' with age 22, address 'Kuta', symptom 'Malaria', and diagnosis 'Tifus'. The fourth row is for 'Anindita Hhor' with age 21, address 'Pesisir', symptom 'Malaria', and diagnosis 'Tifus'. The fifth row is for 'Pendi Alifurrah' with age 9, address 'Zanyah', symptom 'Malaria', and diagnosis 'Tifus'. The sixth row is for 'Muharriz Hake' with age 18, address 'Kusambi', symptom 'Malaria', and diagnosis 'Tifus'. The seventh row is for 'Rach Raha' with age 11, address 'Kusambi', symptom 'Malaria', and diagnosis 'Tifus'. The eighth row is for 'Rafy Sora' with age 41, address 'Kusambi', symptom 'Malaria', and diagnosis 'Tifus'. The ninth row is for 'Rafy Hake' with age 23, address 'Kusambi', symptom 'Malaria', and diagnosis 'Tifus'. The tenth row is for 'Hilalungga' with age 29, address 'Kusambi', symptom 'Malaria', and diagnosis 'Tifus'. The page includes a sidebar with a 'K-M' menu, a top navigation bar with 'Home', 'Tabel', and 'Data Pasien', and a search bar. At the bottom, there is a copyright notice: 'Copyright © 2020-2021 CLUSTERING ALGORITHM IN HEAVS. All rights reserved.' and a version number 'Version 0.0.1'.

No	Nama Lengkap	JK	Alamat	Umur	Gejala Penyakit	Diagnosis	Aksi
1	Della Ouliana	P	Kusambi	15	Malaria	Tifus	✖ ⚙
2	Beti Dinda	P	Tidung	24	Malaria	Tifus	✖ ⚙
3	Nuruz Dzulqarnain	L	Kuta	22	Malaria	Tifus	✖ ⚙
4	Anindita Hhor	L	Pesisir	21	Malaria	Tifus	✖ ⚙
5	Pendi Alifurrah	L	Zanyah	9	Malaria	Tifus	✖ ⚙
6	Muharriz Hake	L	Kusambi	18	Malaria	Tifus	✖ ⚙
7	Rach Raha	L	Kusambi	11	Malaria	Tifus	✖ ⚙
8	Rafy Sora	P	Kusambi	41	Malaria	Tifus	✖ ⚙
9	Rafy Hake	L	Kusambi	23	Malaria	Tifus	✖ ⚙
10	Hilalungga	P	Kusambi	29	Malaria	Tifus	✖ ⚙

Gambar 5.4 Tampilan Halaman Data Pasien

Tampilan halaman data pasien terdiri dari Nama Lengkap, Jenis Kelamin, Alamat, Umur, Gejala Penyakit, Diagnosa dan Aksi. Untuk menambahkan data pasien tekan tombol tambah data, dan tekan edit, simpan, hapus, data pasien.

5.1.5 Tampilan Halaman Input data *Centroid*

The screenshot shows a web form titled "Tambah Data Pasien". It has several input fields: "Nama Lengkap Sesuai KTP", "Jenis Kelamin" with radio buttons for "Pria" and "Wanita", "Alamat Lengkap Sesuai KTP", "Nomor Telpun/HP", "Tanggal Lahir", and "Gejala Penyakit" with a dropdown menu. A blue "Tambah" button is located at the bottom left of the form.

Gambar5.5 Halaman Input Data Pasien

Halaman ini digunakan untuk menambahkan data pasien penderita penyakit menulah dengan memasukkan nama lengkap, alamat, tanggal lahir, jenis kelamin, nomor telpon serta gejala penyakit. Tekan tambah untuk melanjutkan proses

5.1.6 Tampilan Halaman Iterasi *K-Means*

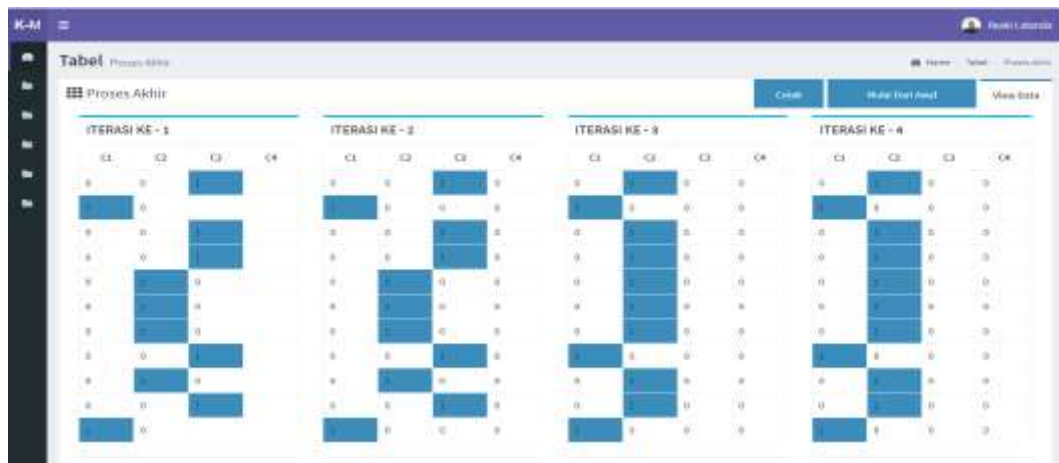
The screenshot shows a table titled "Tabel Data Awal" with 11 rows of patient data. The table has columns for "No", "Nama Lengkap", "Umur", "Gejala", "Diagnosa", and three "Centroid" columns (Centroid 1, Centroid 2, Centroid 3). Each centroid column has sub-columns for x, y, and z coordinates. The last column shows the cluster assignment for each patient, with values 0, 1, or 2.

No	Nama Lengkap	Umur	Gejala	Diagnosa	Centroid 1			Centroid 2			Centroid 3			C1	C2	C3
					x	y	z	x	y	z	x	y	z			
1	Della Dufanren	25	1	1	45,011108778788	7,211102550828	4,2428488871191	0	0	0	0	0	0	0	0	0
2	Reti Gusti	34	1	2	15,902378484289	38,327767049088	32,668818071762	0	0	0	0	0	0	0	0	0
3	Yenni Dufalang	22	2	2	48,962409383782	4,472339540888	2,484897427882	0	0	0	0	0	0	0	0	0
4	Awikahri HPhoe	21	2	3	40,001750834534	3,484310101378	1,454213503772	0	0	0	0	0	0	0	0	0
5	Rendi Runkali	8	3	3	61,096261481591	9,327278588888	11,52988837988	0	0	0	0	0	0	0	0	0
6	Mohamad Nyska	18	4	4	54,203323831471	6	6,489748088879	0	0	0	0	0	0	0	0	0
7	Blud Kanta	11	6	4	50,1881476227841	7,8932480756007	11,269427889582	0	0	0	0	0	0	0	0	0
8	Pelby Siana	41	5	5	25,90340088389	21,400388821439	15,2308481811	0	0	0	0	0	0	0	0	0
9	Rydia Makot	10	7	8	60,336245470001	13,440390588511	11,232644427888	0	0	0	0	0	0	0	0	0
10	Isa Lantongga	26	7	8	41,488703832125	42,267888753698	8,8923332678828	0	0	0	0	0	0	0	0	0
11	Muati Ario	48	10	10	25,4888971987884	32,124788888438	27,802851381588	0	0	0	0	0	0	0	0	0

Gambar 5.6 Tampilan Iterasi K-Means

Halaman ini menampilkan data iterasi k-means yang terdiri dari nama lengkap, umur, gejala, diagnose dan C1,C2,C3 tekan proses iterasi selanjutnya untuk menampilkan hasil iterasi ke 2 dan....”

5.1.7 Tampilan Hasil *Clustering*



Gambar 5.7 Tampilan hasil Cluster Data Pasien

Halaman ini menampilkan hasil clustering, tekan mulai dari awal untuk kembali ketitik awal.

BAB VI

PENUTUP

6.1 KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan sebelumnya, maka dapat ditarik suatu kesimpulan bahwa :

1. Pada perancangan sistem pengolahan data penyakit menular dengan menggunakan metode *K-Means Clustering* di Rumah Sakit Zainal Umar Sidiki dengan hasil penelitian mendapatkan 3 cluster dengan masing-masing cluster berdasarkan usia yaitu, *cluster* 1 didominasi oleh remaja dengan penyakit menular TB, DBD dan Diare, *cluster* 2 oleh anak-anak dengan penyakit ISPA dan Diare sedangkan *cluster* 3 didominasi oleh Lanjut usia dengan penyakit Turbekulosis.
2. Menggunakan metode *K-Means* dengan sistem *clustering* pada penyakit menular yang di uji kinerjanya dengan *white box* dan *Black Box* menghasilkan $V(G) = CC$, sehingga dapat dinyatakan bahwa sistem telah memenuhi syarat logika dan pemrograman, dan menyatakan bahwa sistem telah bebas dari berbagai kesalahan komponen-komponenya.

6.2 SARAN

Setelah melakukan penelitian dengan mengelompokkan penyakit menular menggunakan algoritma *K-Means*, adapun beberapa saran sebagai berikut :

1. Penulis berharap agar penelitian ini menjadi acuan bagi peneliti selanjutnya dalam mengembangkan penelitian untuk mendapatkan hasil yang lebih baik.
2. Diharapkan agar dilakukan penambahan data set penyakit dan pusat cluster agar hasil cluster lebih banyak.

DAFTAR PUSTAKA

- Ade Bastian, dokumen jurnal Sistem Informasi Universitas Majalengka, “Penerapan Algoritma *K-Means Clustering* Analysis Pada Penyakit Menular Manusia”, Majalengka, Vol.14, No.1, April 2018
- NI wayan Wardani Analisis Penerapan *K-Means* Untuk Pengelompokan Diagnosa Penyakit Kulit dan Kelamin Berdasarkan Rentang Usia, Jurnal Teknik Informatika, Denpasar – Bali, 27 Agustus 2016.
- Magdalena Simanjuntak Penerapan Data *Mining* Pengelompokan Penyakit Menular Seksual (PMS) Menggunakan Metode *Clustering*, Jurnal Mahajana Informasi Vol.4 No.1, 2019.
- D.Firdaus, “ Penggunaan Data *Mining* Dalam Kegiatan Sistem Pembelajaran Berbentuk Komputer,” J. Format, Vol. 6,no. 2, pp.91-97, 2017
- Xu & Wunc II, Algoritma *K-Means Clustering*, 2009.
- Ade Bastian Grafik Penerapan Algoritma *K-Means* untuk *clustering* dokumen jurnal Sistem Informasi Universitas Majalengka, Vol.14, No.1, April 2018.
- H. M. Jugiyanto, Analisa dan Desain Sistem Informasi : Pendekatan Terstruktur Teori dan Praktik Aplikasi Bisnis. ANDI, Yokyakarta, 2005
- Garth, N.Jone, Sistem Manajemen Basis Data No.12, 2017
- Satzinger J.W. Jackson, R.B., dan Burd, S.D Metode Analisis Sistem, p4, 2010
- Whiten et al, Jeffery L, Metode Desain & Analisis Sistem, Edisi 6, Edisi International. Yogyakarta : ANDI, Yogyakarta, 2004

A. Nugroho, Rekayasa Perangkat Lunak Berbasis Objek Dengan Metode USOD.
ANDI, No.7 Yogyakarta, 2010

Tom MC Cambe dan Roger S. pressman, 536 Pengujian Sistem, 2002.

KODE PROGRAM

Iterasi K-means

```
<div class="tab-pane active" id="admin">
<div class="box-body">
<table id="tab22" class="table table-bordered table-striped">
<thead>
<tr>
<th rowspan="2" style="width:15px; vertical-align:middle; text-align:center;">No</th>
<th rowspan="2" style="width:120px; vertical-align:middle; text-align:center;">Nama Lengkap</th>
<th rowspan="2" style="width:120px; vertical-align:middle; text-align:center;">Umur</th>
<th rowspan="2" style="width:50px; vertical-align:middle; text-align:center;">Gejala</th>
<th rowspan="2" style="width:80px; vertical-align:middle; text-align:center;">Diagnosa</th>

<th colspan="4" style="width:50px; vertical-align:middle; text-align:center;">Centorid 1</th>
<th colspan="4" style="width:50px; vertical-align:middle; text-align:center;">Centorid 2</th>
<th colspan="4" style="width:50px; vertical-align:middle; text-align:center;">Centorid 3</th>
<!-- <th colspan="4" style="width:50px; vertical-align:middle; text-align:center;">Centorid 4</th> -->

<th rowspan="2" style="width:40px; vertical-align:middle; text-align:center;">C1</th>
<th rowspan="2" style="width:40px; vertical-align:middle; text-align:center;">C2</th>
<th rowspan="2" style="width:40px; vertical-align:middle; text-align:center;">C3</th>
<!-- <th rowspan="2" style="width:40px; vertical-align:middle; text-align:center;">C4</th> -->
</tr>
<tr>
<th>70</th><th>1</th><th>1</th><th>2</th>
<th>27</th><th>2</th><th>2</th><th>2</th>
<th>22</th><th>3</th><th>3</th><th>1</th>
<!-- <th>18</th><th>4</th><th>4</th><th>2</th> -->
```

```

</tr>
</thead>
<tbody>
<?php

$c1a = 70;
$c1b = 1;
$c1c = 1;
// $c1d = 2;

$c2a = 18;
$c2b = 2;
$c2c = 2;
// $c2d = 2;

$c3a = 22;
$c3b = 3;
$c3c = 3;
// $c3d = 1;

<li style="background-color:#3c8dbc; color:#FFF; width:15%; text-align:center;" >
    <a href="home-data-
iterasi_lann&<?php echo $gett; ?>=<?php echo $result; ?>=0017&html" style="background-color:#3c8dbc; color:#FFF; width:100%; text-align:center; border-style:none; height:40px;"
>Proses Iterasi Selanjutnya</a></li>
    <li class="pull-left header"><i class="fa fa-th"></i>Proses Iterasi Selanjutnya</li>
</ul>
<div class="tab-content">

```

DAFTAR RIWAYAT HIDUP



Nama: Reski Latanda

Tempat,Tanggal Lahir: Kwandang, 30-10-1996

Jenis Kelamin: Laki-Laki

Agama: Islam

Status: Belum Menikah

Alamat: Desa Titidu, Kec Kwandang, Kab

Gorontalo Utara

No.Telp/Hp: 082259745502

Email: Reksilatanda@gmail.com

Riwayat Pendidikan:

1. Tahun 2009, Menyelesaikan Pendidikan Di Sekolah Dasar Ponelo, Kecamatan Ponelo Kepulauan, Kabupaten Gorontalo Utara, Provinsi Gorontalo.
2. Tahun 2012, Menyelesaikan Pendidikan Di Sekolah Menengah Pertama Smp Negri 4 Kwandang, Kecamatan Ponelo Kepulauan, Kabupaten Gorontalo Utara, Provinsi Gorontalo.
3. Tahun 2015, Menyelesaikan Pendidikan Di Sekolah Menengah Atas Negri 8 Gorontalo Utara, Kecamatan Ponelo Kepulauan, Kabupaten Gorontalo Utara, Provinsi Gorontalo.
4. Tahun 2017, Telah Di Terima Menjadi Mahasiswa Perguruan Tinggi Swasta Universitas Ichsan Gorontalo Utara, Kecamatan Kwandang, Kabupaten Gorontalo Utara, Provinsi Gorontalo.



PEMERINTAH KABUPATEN GORONTALO UTARA

RSUD dr. ZAINAL UMAR SIDIKI

Jl. Cilomati Desa Bulalo Kec. Kwandang

Yang bertanda tangan di bawah ini :
Nama : **DIAN AFIYANTI ARBIE S. Farm**
NIP : 19840810 200901 2 006
Pangkat/Gol : Penata, III/C
Jabatan : Kepala Ruangan BPJS Center RSUD
Dr. Zainal Umar Sidiki

Dengan ini menerangkan bahwa :

Nama : **RESKI LATANDA**
NIM : T3117230
Program Pendidikan : Sarjana (S1)
Program Studi : Teknik Informatika

Bahwa yang bersangkutan telah selesai melaksanakan kegiatan penelitian tentang ***“IMPLEMENTASI DATA MINING DENGAN ALGORITMA K MEANS CLUSTERING PENYAKIT MENULAR”***, di Rumah Sakit Zainal Umar sidiki.

Demikian surat keterangan ini dibuat untuk dipergunakan sebagai mana mestinya

Gorontalo Utara, Mei 2021


DIAN AFIYANTI ARBIE S. Farm



**KEMENTERIAN PENDIDIKAN DAN KEBUDAYAAN
UNIVERSITAS ICHSAN
(UNISAN) GORONTALO**

SURAT KEPUTUSAN MENDIKNAS RI NOMOR 84/D/O/2001
Jl. Achmad Nadjamuddin No. 17 Telp (0435) 829975 Fax (0435) 829976 Gorontalo

SURAT REKOMENDASI BEBAS PLAGIASI

No. 0711/UNISAN-G/S-BP/V/2021

Yang bertanda tangan di bawah ini:

Nama : Sunarto Taliki, M.Kom
NIDN : 0906058301
Unit Kerja : Pustikom, Universitas Ichsan Gorontalo

Dengan ini Menyatakan bahwa :

Nama Mahasisw : RESKI LATANDA
NIM : T3117230
Program Studi : Teknik Informatika (S1)
Fakultas : Fakultas Ilmu Komputer
Judul Skripsi : IMPLEMENTASI DATA MINING DENGAN
ALGORITMA K-MEANS CLUSTERING PENYAKIT
MENULAR DI RUMAH SAKIT ZAINAL UMAR SIDIKI

Sesuai dengan hasil pengecekan tingkat kemiripan skripsi melalui aplikasi Turnitin untuk judul skripsi di atas diperoleh hasil Similarity sebesar 20%, berdasarkan SK Rektor No. 237/UNISAN-G/SK/IX/2019 tentang Panduan Pencegahan dan Penanggulangan Plagiarisme, bahwa batas kemiripan skripsi maksimal 35% dan sesuai dengan Surat Pernyataan dari kedua Pembimbing yang bersangkutan menyatakan bahwa isi softcopy skripsi yang diolah di Turnitin SAMA ISINYA dengan Skripsi Aslinya serta format penulisannya sudah sesuai dengan Buku Panduan Penulisan Skripsi, untuk itu skripsi tersebut di atas dinyatakan BEBAS PLAGIASI dan layak untuk diujikan.

Demikian surat rekomendasi ini dibuat untuk digunakan sebagaimana mestinya.

Gorontalo, 31 Mei 2021

Tim Verifikasi,



Sunarto Taliki, M.Kom

NIDN. 0906058301

Tembusan :

1. Dekan
2. Ketua Program Studi
3. Pembimbing I dan Pembimbing II
4. Yang bersangkutan
5. Arsip

T3117230 RESKI LATANDA

IMPLEMENTASI DATA MINING DENGAN ALGORITMA K-MEAN...

Sources Overview

20%

OVERALL SIMILARITY

1	www.scribd.com	11%
2	ejournal.catursekti.ac.id	2%
3	simad.unpkediri.ac.id	2%
4	id.123dok.com	<1%
5	repository.its.ac.id	<1%
6	media.neliti.com	<1%
7	fr.scribd.com	<1%
8	irin-halid.blogspot.com	<1%
9	docplayer.info	<1%
10	caraguna.com	<1%
11	repository.ub.ac.id	<1%
12	titonkadir.blogspot.com	<1%
13	id.scribd.com	<1%
14	repository.usd.ac.id	<1%
15	Rismayanti Rismayanti, Fera Damayanti, Khairunnisa Khairunnisa. "Penerapan Data Mining Algoritma C4.5 dalam Menentukan Rekomendasi Produk"	<1%
16	eprints.umm.ac.id	<1%

17 jurnalidigit.org
INTERNET

<1%

18 repository.uin-suska.ac.id
INTERNET

<1%

Excluded search repositories:

- Submitted Works

Excluded from Similarity Report:

- Small Matches (less than 25 words).

Excluded sources:

- None