

**ANALISIS POLA PENYAKIT PASIEN
PADA RSUD EMBUNG FATIMAH KOTA BATAM
MENGUNAKAN DATA MINING**

SKRIPSI



**Oleh:
Eltarianingsih N
141510133**

**PROGRAM STUDI SISTEM INFORMASI
UNIVERSITAS PUTERA BATAM
2018**

**ANALISIS POLA PENYAKIT PASIEN
PADA RSUD EMBUNG FATIMAH KOTA BATAM
MENGUNAKAN DATA MINING**

SKRIPSI

**Untuk memenuhi salah satu syarat
guna memperoleh gelar Sarjana**



**Oleh
Eltarianingsih N
141510133**

**PROGRAM STUDI SISTEM INFORMASI
UNIVERSITAS PUTERA BATAM
2018**

SURAT PERNYATAAN ORISINALITAS

Yang bertanda tangan di bawah ini saya:

Nama : Eltarianingsih N
NPM/NIP : 141510133
Fakultas : Teknik dan Komputer
Program Studi : Sistem Informasi

Menyatakan bahwa “**Skripsi**” yang saya buat dengan judul:

ANALISIS POLA PENYAKIT PASIEN PADA RSUD EMBUNG FATIMAH KOTA BATAM MENGGUNAKAN DATA MINING

Adalah hasil karya sendiri dan bukan “duplikasi” dari karya orang lain. Sepengetahuan saya, didalam naskah Skripsi ini tidak terdapat karya ilmiah atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara tertulis dikutip didalam naskah ini dan disebutkan dalam sumber kutipan dan daftar pustaka.

Apabila ternyata di dalam naskah Skripsi ini dapat dibuktikan terdapat unsur-unsur PLAGIASI, saya bersedia naskah Skripsi ini digugurkan dan gelar akademik yang saya peroleh dibatalkan, serta diproses sesuai dengan peraturan perundang-undangan yang berlaku.

Demikian pernyataan ini saya buat dengan sebenarnya tanpa ada paksaan dari siapapun

Batam, 03 Agustus 2018

Materai 6000

Eltarianingsih N

141510133

**ANALISIS POLA PENYAKIT PASIEN
PADA RSUD EMBUNG FATIMAH KOTA BATAM
MENGUNAKAN DATA MINING**

SKRIPSI

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

Oleh

**Eltarianingsih N
141510133**

**Telah disetujui oleh Pembimbing pada tanggal
Seperti tertera dibawah ini**

Batam, 03 Agustus 2018

Nurul Azwanti, S.Kom., M.Kom.

Pembimbing

ABSTRAK

Semua manusia yang hidup di bumi tidak lepas dari penyakit yang bersarang didalam tubuhnya, baik penyakit yang sudah terbawa sejak lahir maupun penyakit yang muncul setelah beberapa tahun manusia tersebut menjalani kehidupannya. Hipertensi merupakan salah satu jenis penyakit yang menyerang organ dalam tubuh yang dapat membahayakan hidup seseorang. Terdapat sekitar 95% kasus hipertensi yang tidak diketahui penyebabnya, sedangkan sisanya ditimbulkan akibat adanya penyakit lain seperti penyakit jantung koroner, gangguan fungsi ginjal, dan gangguan fungsi kognitif atau stroke. RSUD Embung Fatimah adalah rumah sakit Indonesia yang terletak di Provinsi Kepulauan Riau Kota Batam. Di tahun 2015, jumlah total pasien untuk rawat inap mencapai 10.317 jiwa. Untuk mengatasi masalah dalam menanggulangi penderita penyakit Hipertensi, diperlukan analisa terhadap data penyakit yang ada, untuk diprediksi penyakit pasien yang harus ditangani berdasarkan pola penyakit. Dalam data mining ada sebuah model yang dapat digunakan untuk memprediksikan sebuah pola dalam suatu kondisi yaitu model prediktif atau prediksi. Salah satu algoritma yang dapat digunakan untuk membuat pohon keputusan (*decission tree*) adalah algoritma C4.5. Algoritma C4.5 merupakan sebuah metode yang digunakan untuk klasifikasi yang bersifat prediktif. Dengan menggunakan metode algoritma C4.5 peneliti dapat mengklasifikasikan pola penyakit hipertensi sebagai penyakit penyerta dari penyakit gagal jantung, gagal ginjal, diabetes, stroke dan hipoglikemia. Dalam penelitian ini, peneliti menggunakan *software* WEKA (*Waikato Environment for Knowledge Analysis*) sebagai *tools* atau alat yang digunakan untuk melakukan pengujian dalam rangka untuk mendapatkan pola penyakit dari hipertensi. Dari hasil penelitian di dapatkan informasi bahwa dalam prediksi penyakit hipertensi sebagai penyakit penyerta ,atribut yang sangat berpengaruh terhadap hipertensi adalah gagal jantung.

Kata kunci: *Data mining*, Algoritma C4.5, WEKA.

ABSTRACT

All humans who live on earth can not be separated from the disease that lodged in the body, both the disease that has been brought from birth and the disease that emerged after several years of human life to live his life. Hypertension is one type of disease that attacks the internal organs that can endanger a person's life. There are approximately 95% of cases of unknown cause of hypertension, while the rest caused by other diseases such as coronary heart disease, impaired kidney function, and impaired cognitive function or stroke. RSUD Embung Fatimah is an Indonesian hospital located in Batam Island Riau Province. In 2015, the total number of inpatients for hospitalization reaches 10,317 inhabitants. With the large number of patients per year it causes patient data is increasing. To overcome the problem in tackling people with hypertension disease, it is necessary to analyze the existing disease data, to predict the patient's illness which must be handled based on the pattern of the disease. In data mining there is a model that can be used to predict a pattern in a condition that is predictive or prediction model. One of the algorithms that can be used to create a decision tree (decission tree) is the C4.5 algorithm. The C4.5 algorithm is a method used for predictive classification. Using C4.5 algorithm method, the researcher can classify the pattern of hypertension as a comorbid illness of heart failure, kidney failure, diabetes, stroke and hypoglycemia. In this study, researchers used WEKA (Waikato Environment for Knowledge Analysis) software as tools or tools used to perform testing in order to obtain the pattern of disease from hypertension. From the research findings in the find that in the prediction of hypertension disease as a disease, the attributes that are very influential to hypertension are heart failure.

Keyword: Data mining, C4.5 algorithm, WEKA

KATA PENGANTAR

Segala puji dan sukur penulis panjatkan kepada Tuhan Yang Maha Esa yang telah melimpahkan segala rahmat dan karuniaNya, sehingga penulis dapat menyelesaikan laporan tugas akhir yang merupakan salah satu persyaratan untuk menyelesaikan program studi strata satu (SI) pada Program Studi Sistem Informasi Universitas Putera Batam.

Penulis menyadari bahwa skripsi ini masih jauh dari sempurna. Karena itu, kritik dan saran akan senantiasa penulis terima dengan senang hati.

Dengan segala keterbatasan, penulis menyadari pula bahwa skripsi ini takkan terwujud tanpa bantuan, bimbingan, dan dorongan dari berbagai pihak. Untuk itu, dengan segala kerendahan hati, penulis menyampaikan ucapan terima kasih kepada:

1. Ibu Dr. Nur Elfi Husda, S.Kom., M.SI, selaku Rektor Universitas Putera Batam.
2. Bapak Muhammad Rasid Ridho, S.Kom., M.SI. selaku Ketua Program Studi Sistem Informasi.
3. Ibu Nurul Azwanti, S.Kom., M.Kom selaku pembimbing skripsi pada Program Studi Sistem Informasi Universitas Putera Batam.
4. Dosen dan Staff Universitas Putera Batam
5. Rekan – rekan mahasiswa sistem informasi
6. Rekan – rekan RSUD Embung Fatimah
7. Management RSUD Embung Fatimah.
8. Kedua Orang tua, dan adik serta seluruh keluarga besar peneliti
9. Sahabat terkasih

Semoga Tuhan membalas kebaikan dan selalu mencurahkan kasih serta berkah-NYA, amin.

Batam, 3 Agustus 2018

Penulis

DAFTAR ISI

HALAMAN SAMPUL DEPAN.....	i
HALAMAN JUDUL.....	ii
SURAT PERNYATAAN ORISINALITAS	iii
HALAMAN PENGESAH	iv
ABSTRAK.....	v
<i>ABSTRACT</i>	vi
KATA PENGANTAR	vii
DAFTAR ISI.....	viii
DAFTAR GAMBAR	xi
DAFTAR TABEL.....	xii
DAFTAR RUMUS	xiii
BAB I PENDAHULUAN	
1.1 Latar belakang	1
1.2 Identifikasi Masalah	4
1.3 Batasan Masalah	4
1.4 Rumusan Masalah	5
1.5 Tujuan Penelitian.....	5
1.6 Manfaat Penelitian.....	6
BAB II TINJAUAN PUSTAKA	
2.1 Teori Dasar	8
2.1.1 Penyakit	8
2.1.1.1 Hipertensi	9
2.1.1.2 Hipoglikemia	10
2.1.1.3 Diabetes	10
2.1.1.4 Gagal Jantung.....	11
2.1.1.5 <i>Stroke</i>	11
2.1.1.6 Gagal ginjal kronik	12
2.1.2 Data.....	12
2.1.3 Data Mining.....	14
2.1.4 Pohon keputusan (<i>Decission Tree</i>).....	17

2.1.5	Algoritma C4.5	18
2.1.6	<i>Waikato Environment for Knowledge Analysis (WEKA)</i>	19
2.2	Penelitian Terdahulu.....	23
2.3	Kerangka Pemikiran.	32
2.4	Hipotesis	32
BAB III METODE PENELITIAN.....		33
3.1	Desain Penelitian	33
3.1.1	Pemilihan variabel	35
3.1.2	Klasifikasi Data	35
3.1.3	Pilih atribut menjadi akar	38
3.1.4	Buat cabang	39
3.1.5	Bagi kasus.....	39
3.1.6	Ulangi	40
3.1.7	Membuat pohon keputusan.....	40
3.2	Operasional Variabel	40
3.3	Populasi dan Sampel.....	41
3.3.1	Populasi	41
3.3.2	Sampel	41
3.4	Teknik Pengumpulan Data	41
3.5	Metode Analisis Data	42
3.6.1	Dokumentasi.....	43
3.6.2	Pengumpulan Data.....	43
3.6.3	Pemilihan Data	43
3.6.4	Evaluasi Data.....	43
3.6.5	Analisis data dengan menggunakan Weka C4.5	43
3.6.6	Pengujian data pada aplikasi WEKA.....	44
3.6.7	Hasil data pada aplikasi WEKA 3.7	44
3.6	RSUD Embung Fatimah.....	44
3.6.1	Visi Rumah Sakit.....	46
3.6.2	Misi Rumah Sakit.....	46
3.6.3	Moto Rumah Sakit.....	46

BAB IV HASIL PENELITIAN DAN PEMBAHASAN

4.1	Hasil Penelitian.....	48
4.1.1	Melakukan Pra – proses.....	49
4.1.2	Menghitung Atribut Akar	50
4.1.2.1	Menghitung Nilai Entropy tiap – tiap atribut	50
4.1.2.2	Menghitung Nilai Gain tiap – tiap atribut.	52
4.1.2.3	Menghitung Tree dengan perhitungan Node 1.1	55
4.1.2.4	Menghitung Nilai Entropy tiap – tiap atribut	55
4.1.2.5	Menghitung Nilai Gain tiap – tiap atribut.	57
4.1.2.6	<i>Rule</i> pohon keputusan.....	60
4.2.	Pengujian <i>software Weka</i>	60
4.2.1.	Langkah – langkah pengujian pada aplikasi <i>Weka</i>	60
4.2.2.	<i>Rule</i> Hasil dari pengujian <i>Weka</i>	65

BAB V SIMPULAN DAN SARAN

5.1	Simpulan.....	66
5.2	Saran	67

DAFTAR PUSTAKA

DAFTAR RIWAYAT HIDUP

DAFTAR GAMBAR

Gambar 2.1 <i>Data Mining</i> sebagai Salah Satu Tahap KDD	15
Gambar 2.2 Tampilan depan weka	21
Gambar 2.3 Eksekusi di alat weka.....	32
Gambar 2.4 Kerangka pemikiran.....	32
Gambar 3.1 Desain Penelitian	34
Gambar 3.2 Struktur Organisasi	47
Gambar 4.1 Pohon keputusan hasil perhitungan node 1	55
Gambar 4.2 pohon keputusan	60
Gambar 4.3 Tampilan Awal Weka.....	61
Gambar 4.4 Memulai proses Weka	61
Gambar 4.5 Mencari <i>type</i> .CSV.....	62
Gambar 4.6 Membaca file CSV	62
Gambar 4.7 Pemilihan pohon yang akan diolah.....	63
Gambar 4.8 Manampilkan <i>result list</i>	63
Gambar 4.9 <i>visualize tree</i>	64
Gambar 4.10 Pohon Keputusan	64

DAFTAR TABEL

Tabel 2.1	Penelitian terdahulu	23
Tabel 3.1	Kategori usia.....	36
Tabel 3.2	Tabel lama rawatan	37
Tabel 3.3	Jadwal penelitian	42
Tabel 4.1	Keputusan Hipertensi	49
Tabel 4.2	Perhitungan Node 1	54
Tabel 4.3	Tabel Node 1.1	59

DAFTAR RUMUS

Rumus 3.1 Rumus Entropy	38
Rumus 3.2 Rumus Gain	39

BAB I

PENDAHULUAN

1.1 Latar belakang

Semua manusia yang hidup di bumi tidak lepas dari penyakit yang bersarang didalam tubuhnya, baik penyakit yang sudah terbawa sejak lahir maupun penyakit yang muncul setelah beberapa tahun manusia tersebut menjalani kehidupannya. Pada umumnya penyakit tidak mengenal usia, karena penyakit bisa menyerang siapa saja, dimanapun, dan kapanpun. Penyakit yang menyerang manusia ada yang dapat dilihat secara langsung maupun secara tidak langsung, dan ada juga penyakit yang mudah teridentifikasi maupun sulit untuk teridentifikasi berhubungan erat dari berbagai faktor penyakit tersebut. Banyak proses yang harus dilewati agar jenis penyakit tersebut dapat diketahui dan diberikan solusi pengobatannya.

Hipertensi merupakan salah satu jenis penyakit yang menyerang organ dalam tubuh yang dapat membahayakan hidup seseorang. Di Negara maju prevalensi mencapai 37,3% (M. Nizar Syarif Hamidi, 2014: 29). Terdapat sekitar 95% kasus hipertensi yang tidak diketahui penyebabnya, sedangkan sisanya ditimbulkan akibat adanya penyakit lain seperti penyakit jantung koroner, gangguan fungsi ginjal, dan gangguan fungsi kognitif atau stroke (Hartama, 2017). Di Indonesia masalah hipertensi cenderung meningkat (Rahajeng &

Tuminah, 2009: 581). Prevalensi penyakit hipertensi di Indonesia rata-rata pada penduduk usia 25-70 tahun sebesar 25,8%, namun cakupan kasus oleh tenaga kesehatan hanya sebesar 36,8% dari total penderita yang diperkirakan dan sisanya tidak terdiagnosis dengan baik di masyarakat (Wahyuddin & Andajani, 2016: 120).

RSUD Embung Fatimah adalah rumah sakit Indonesia yang terletak di Provinsi Kepulauan Riau Kota Batam. Di tahun 2015, jumlah total pasien untuk rawat inap mencapai 10.317 jiwa (Profil kesehatan Kota Batam, 2015: 231). Dengan komplikasi beberapa penyakit yang muncul secara bersama, Dengan jumlah pasien penderita hipertensi yang cukup banyak, tidak menutup kemungkinan dapat mengakibatkan bertambahnya angka kematian akibat hipertensi. Hal tersebut dikarenakan hipertensi dapat menjadi faktor penyerta penyakit – penyakit lainnya seperti, gagal ginjal, gagal jantung, diabetes, *stroke* dan hipoglikemia. Dengan banyaknya jumlah pasien pertahun nya menyebabkan data pasien semakin bertambah. Saat ini data – data tersebut disimpan pada aplikasi SIMRS (Sistem Informasi Manajemen Rumah Sakit) dan belum dimanfaatkan secara maksimal. Kurangnya perhatian terhadap tumpukan data pasien ini dapat menyebabkan banyaknya kendala atau permasalahan yang seharusnya dapat diantisipasi jauh sebelum permasalahan tersebut muncul. Semakin lama data pasien yang disimpan tersebut dibiarkan tentunya akan semakin sulit memprediksikan sebuah solusi untuk memecahkan permasalahan apabila permasalahan tersebut muncul. Selain itu, data hanya akan menjadi sebuah tumpukan arsip atau file yang disimpan sebagai formalitas kerja manajemen

rumah sakit saja. Oleh sebab itu peneliti melakukan penelitian dengan memanfaatkan tumpukan data tersebut agar dapat dimanfaatkan secara maksimal.

Untuk mengatasi masalah dalam menanggulangi penderita penyakit Hipertensi, diperlukan analisa terhadap data penyakit yang ada, untuk diprediksi penyakit pasien yang harus ditangani berdasarkan pola penyakit. Data mining digunakan untuk mengetahui pola data, mengatur hubungan informasi tersembunyi, struktur aturan Asosiasi, memperkirakan item tidak diketahui, nilai-nilai untuk mengklasifikasikan benda, menyusun cluster objek homogen, dan mengungkap berbagai temuan yang tidak mudah diproduksi oleh CBIS (*computer based information system*) klasik. Dengan demikian, hasil DM mewakili dukungan yang berharga untuk membuat keputusan (Wang et al., 2018: 1). Dalam data mining ada sebuah model yang dapat digunakan untuk memprediksikan sebuah pola dalam suatu kondisi yaitu model prediktif atau prediksi. Model prediktif sering menerapkan pembelajaran diawasi fungsi untuk memperkirakan nilai-nilai tidak diketahui atau masa depan tergantung variables berdasarkan fitur terkait variabel independen. (Peña-Ayala, 2014). Untuk memprediksikan suatu kondisi dalam data mining, diperlukan sebuah alat bantu yang mampu mengolah data-data menjadi informasi yang akan digunakan untuk memprediksikan dan memuatnya menjadi sebuah pola.

Pohon keputusan dibangun sebagai seperangkat aturan keputusan, ditata secara visual sebagai pohon (Yeo & Grant, 2018). Pohon keputusan yang memperlihatkan faktor-faktor kemungkinan atau probabilitas yang akan mempengaruhi alternatif-alternatif keputusan tersebut, disertai dengan estimasi

hasil akhir yang akan didapat bila kita mengambil alternatif keputusan tersebut (Pramadhani & Setiadi, 2014). Salah satu algoritma yang dapat digunakan untuk membuat pohon keputusan (*decission tree*) adalah algoritma C4.5 (Yuli, 2017: 216). Algoritma C4.5 merupakan sebuah metode yang digunakan untuk klasifikasi yang bersifat prediktif (Febriana, Furqon, & Rahayudi, 2018).

Berdasarkan latar belakang tersebut maka penulis tertarik melakukan penelitian dengan judul “ **Analisis Pola Penyakit Pasien Pada RSUD Embung Fatimah Kota Batam Menggunakan Metode Data Mining**”.

1.2 Identifikasi Masalah

Dari latar belakang di atas, di dapati permasalahan yakni :

1. Jumlah pasien penderita hipertensi yang cukup banyak, tidak menutup kemungkinan dapat mengakibatkan bertambahnya angka kematian akibat hipertensi.
2. Banyaknya jumlah pasien pertahun nya menyebabkan data pasien semakin bertambah. Data – data tersebut disimpan pada aplikasi SIMRS (Sistem Informasi Manajemen Rumah Sakit) namun belum dikelola secara maksimal.

1.3 Batasan Masalah

Batasan masalah pada penelitian ini adalah sebagai berikut:

1. Pada penelitian ini analisis data mining menggunakan algoritma C4.5

2. Sumber data dihimpun dari data pasien rawat inap pada RSUD Embung Fatimah dengan jaminan BPJS.
3. Data yang digunakan pada penelitian ini adalah data yang teridentifikasi penyakit hipertensi.
4. Rentang data yang diambil untuk penelitian ini mulai dari bulan januari 2016 – desember 2016 (12 bulan).
5. Hasil output dari penelitian ini adalah menghasilkan pohon keputusan guna memprediksi pola penyakit hipertensi.

1.4 Rumusan Masalah

Perumusan masalah dari penelitian ini adalah:

1. Bagaimana penelitian ini dapat menghasilkan sebuah solusi yang bisa digunakan dalam menanggulangi dan mengurangi angka kematian akibat hipertensi di RS Embung Fatimah.
2. Bagaimana pengolahan data mining terhadap data - data yang saat ini belum dimanfaatkan secara maksimal di RS Embung Fatimah dengan menggunakan metode algoritma C4.5.

1.5 Tujuan Penelitian

Tujuan penelitian ini adalah:

1. Untuk menanggulangi dan mengurangi jumlah angka kematian akibat penyakit hipertensi.

2. Untuk menerapkan metode data mining dengan algoritma C4.5 dalam memprediksi pola penyakit hipertensi.

1.6 Manfaat Penelitian

Pada penelitian ini, manfaat yang diharapkan adalah dapat berguna dalam hal-hal berikut ini:

1.6.1 Peneliti

Dengan dilakukannya penelitian ini, peneliti dapat menyelesaikan tugas akhir masa studi (Skripsi) dan peneliti mendapat kesempatan untuk mengimplementasikan pengetahuan mengenai teknik analisis data mining yang selama masa pendidikan di Universitas Putera Batam telah diajarkan oleh para dosen pengajar. Selain itu, peneliti juga dapat berbagi pengetahuan kepada rekan-rekan mahasiswa lain, baik senior maupun junior, bahwa tumpukan data yang sudah lama disimpan dalam database masih dapat di olah untuk menghasilkan informasi yang berguna.

1.6.2 Universitas

Penelitian ini dapat dijadikan sebagai bahan masukan guna meningkatkan kualitas pendidikan di Universitas Putera Batam, termasuk bagi mahasiswa/i dapat dijadikan sebagai salah satu ilmu pengetahuan tentang analisis data mining. Penelitian ini juga dapat digunakan sebagai referensi pada perpustakaan bagi mahasiswa lain yang melakukan penelitian dengan tema analisis data mining.

1.6.3 Instansi

Bagi obyek penelitian atau RSUD Embung fatimah penelitian ini dapat digunakan dalam mengidentifikasi data penyakit pasien yang menumpuk guna mengidentifikasi hubungan yang terjadi pada pola penyakit pasien sehingga dapat bermanfaat bagi Rumah sakit.

BAB II

TINJAUAN PUSTAKA

2.1 Teori Dasar

2.1.1 Penyakit

Pengertian Sakit Menurut Revelly dalam Effendy (1998), sakit adalah tidak adanya keselarasan antara lingkungan dengan individu. Menurut Perkins dalam Notosoedirjo & Latipun (2002), sakit adalah suatu keadaan yang tidak menyenangkan yang menimpa seseorang sehingga menimbulkan gangguan aktifitas sehari-hari baik aktifitas jasmani, rohani, dan sosial (Sani, 2011)

Sementara itu penyakit adalah gangguan kesehatan yang disebabkan oleh bakteri, virus, atau kelainan sistem faal atau jaringan pada organ tubuh (pada makhluk hidup) (Kbbi.kemdikbud.go.id, 2016). ketika seseorang menderita sakit maka akan mencari pengobatan yang relative lebih murah sesuai dengan penghasilannya (Siti Uswatun Chasanah, 2017: 7). Semakin bertambahnya tingkat populasi hidup dari tahun ke tahun tidak menutup kemungkinan semakin meningkat pula resiko untuk seseorang tertular sebuah penyakit (World Health Organization, 2017: 2).

Pada penelitian ini variabel penyakit adalah hipertensi, gagal ginjal, diabetes, hipoglikemia, gagal jantung, *stroke*.

2.1.1.1 Hipertensi

Hipertensi merupakan salah satu penyakit degeneratif yang perlu diwaspadai karena merupakan penyebab kematian utama ketiga di Indonesia untuk semua umur (6,8%), setelah *stroke* (15,4%) dan tuberculosis (7,5%) (Anisa, Bangun, & Sinulingga, 2014:). Pada penelitian (Khotimah, 2013) disebutkan, bahwa hipertensi merupakan gangguan sistem peredaran darah yang ditandai oleh meningkatnya tekanan darah dalam tubuh. Seseorang yang menderita hipertensi dapat berpotensi mengalami komplikasi, seperti *stroke* dan penyakit jantung lainnya yang dapat berakibat fatal. Kemudian dalam penelitian (M. Nizar Syarif Hamidi, 2014) menurut WHO, Hipertensi atau "tekanan darah tinggi" adalah tekanan darah sistole ≥ 140 mmHg dan tekanan darah diastole ≥ 90 mmHg. Tekanan darah sistole adalah suatu kondisi penguncupan jantung akibat kontraksi otot jantung, sedangkan tekanan darah diastole adalah sebuah kondisi pengenduran otot jantung pada waktu pengisian kembali jantung oleh darah (Kementerian Pendidikan dan Kebudayaan Republik Indonesia, 2016).

Hipertensi 90% tidak diketahui secara pasti faktor penyebabnya, namun dari berbagai penelitian telah di temukan beberapa faktor yang sering menyebabkan terjadinya hipertensi (Khotimah, 2013). Namun dalam penelitian (Rahajeng & Tuminah, 2009: 586) menyebutkan, faktor risiko hipertensi di Indonesia adalah umur, pria, pendidikan rendah, kebiasaan merokok, konsumsi minuman berkafein >1 kali per hari, konsumsi alkohol, kurang aktivitas fisik, obesitas dan obesitas abdominal.

2.1.1.2 Hipoglikemia

Hipoglikemia adalah komplikasi akut diabetes melitus yang seringkali terjadi secara berulang yang ditandai dengan gula darah kurang dari 70 mg/dl. Penyandang diabetes melitus akan menghadapi situasi dilematik dimana mereka diharuskan memperoleh terapi obat penurun gula darah untuk mengontrol kadar gula darah tetap normal, namun juga menghadapi kekhawatiran akan efek samping terapi yang dapat menyebabkan komplikasi hipoglikemia (Jon Hafan Sutawardana, Yulia, 2016).

2.1.1.3 Diabetes

Diabetes Melitus (DM) merupakan penyakit metabolik yang ditandai dengan hiperglikemia yang terjadi karena kelainan sekresi insulin, kerja insulin, atau keduanya. World Health Organization (WHO) memperkirakan jumlah penduduk dunia yang menderita DM pada Tahun 2030 akan meningkat paling sedikit menjadi 366 juta. Menurut survei yang telah dilakukan oleh WHO, Indonesia menempati urutan ke-4 terbesar dalam jumlah penderita DM setelah India, China, dan Amerika Serikat. Sedangkan berdasarkan hasil Riset Kesehatan Dasar (Riskesdas) pada Tahun 2013, didapatkan hasil bahwa proporsi DM meningkat dua kali lipat dibandingkan Tahun 2007. Proporsi DM di Indonesia adalah sebesar 6,9% dengan estimasi sebanyak 176.689.336 penduduk usia 15 tahun ke atas sebagai penderitanya. Hal ini menunjukkan bahwa di Indonesia, penyakit DM merupakan masalah kesehatan masyarakat yang sangat serius (Kemenkes.RI, 2014).

2.1.1.4 Gagal Jantung

Congestive Heart Failure (CHF) merupakan suatu keadaan patologis di mana kelainan fungsi jantung menyebabkan kegagalan jantung memompa darah untuk memenuhi kebutuhan jaringan, atau hanya dapat memenuhi kebutuhan jaringan dengan meningkatkan tekanan pengisian (McPhee & Ganong, 2010). Penyakit hipertensi menjadi penyakit yang paling banyak dialami oleh responden selain penyakit CHF yang dimiliki. Prosentase mencapai 43,8% pada kelompok intervensi dan 62,5% pada kelompok kontrol. Hal ini karena peningkatan tekanan darah yang bersifat kronis membuat jantung memompa dengan sangat kuat untuk mendorong darah ke dalam arteri sehingga otot-otot jantung menebal dan membesar. Hal ini mengakibatkan irama jantung menjadi kaku sehingga irama denyut nadi tidak teratur. Pemompaan yang kurang efektif ini dapat mengakibatkan gagal jantung (Jon Hafan Sutawardana, Yulia, 2016).

2.1.1.5 Stroke

Stroke adalah penyakit serebrovaskular yang sering ditemukan di negara maju, saat ini juga terdapat di negara berkembang salah satunya di Indonesia. Prevalensi stroke yang tinggi di Indonesia dipengaruhi oleh beberapa faktor resiko antara lain obesitas, kurang aktifitas fisik, diet tidak sehat, merokok, tekanan darah tinggi, peningkatan gula darah, dan peningkatan profil lipid darah.

Masalah stroke di Indonesia menjadi semakin penting karena di Asia menduduki urutan pertama dengan jumlah kasusnya yang semakin banyak. Penyakit stroke merupakan salah satu dari penyakit tidak menular yang masih

menjadi masalah kesehatan yang penting di Indonesia. Seiring dengan semakin meningkatnya morbiditas dan mortalitas dalam waktu yang bersamaan, dimana di Indonesia peningkatan kasus dapat berdampak negatif terhadap ekonomi dan produktivitas bangsa, karena pengobatan stroke membutuhkan waktu lama dan memerlukan biaya yang besar (Rahmawati, 2017).

2.1.1.6 Gagal ginjal kronik

Penyakit Gagal Ginjal Kronik (GGK) merupakan salah satu penyakit yang menjadi masalah besar di dunia. Gagal ginjal kronik merupakan suatu penyakit yang menyebabkan fungsi organ ginjal mengalami penurunan hingga akhirnya tidak mampu melakukan fungsinya dengan baik . Gangguan fungsi ginjal ini terjadi ketika tubuh gagal untuk mempertahankan metabolisme dan keseimbangan cairan dan elektrolit sehingga menyebabkan retensi urea dan sampah nitrogen lain dalam darah. Kerusakan ginjal ini mengakibatkan masalah pada kemampuan dan kekuatan tubuh yang menyebabkan aktivitas kerja terganggu, tubuh jadi mudah lelah dan lemas sehingga kualitas hidup pasien menurun (Ali, Masi, & Kallo, 2017).

2.1.2 Data

Data merupakan bentuk jamak dari bentuk tunggal datum atau data-item. data adalah kenyataan yang menggambarkan suatu kejadian-kejadaian dan kesatuan nyata. Kejadian-kejadian (event) adalah suatu yang terjadi pada saat yang tertentu (Iswany & Informasi, 2016:108).

Dimana:

1. *Characters*

Merupakan bagian data yang terkecil, dapat berupa karakter numerik, huruf ataupun karakter-karakter khusus (*special characters*) yang membentuk suatu *item data/field*.

2. *Field*

merepresentasikan suatu atribut dari *record* yang menunjukkan suatu item dari data, seperti misalnya nama, alamat dan lain sebagainya. Kumpulan dari *field* membentuk suatu *record*.

a. *field name* : harus diberi nama untuk membedakan *field* yang satu dengan lainnya.

b. *field representation* : Tipe *field* (karakter, teks, tanggal, angka, dsb), lebar *field* (ruang maksimum yang dapat diisi dengan karakter-karakter data).

c. *field value* : isi dari *field* untuk masing-masing *record*.

3. *Record*

Kumpulan dari *field* membentuk suatu *record*. *Record* menggambarkan suatu unit data individu yang tertentu. Kumpulan dari *record* membentuk suatu *file*. Misalnya *file* personalia, tiap-tiap record dapat mewakili data tiap-tiap karyawan.

4. *File*

File terdiri dari *record-record* yang menggambarkan satu kesatuan data yang sejenis. Misalnya *file* mata pelajaran berisi data tentang semua mata pelajaran yang ada.

5. *Database*

Kumpulan dari *file* / tabel yang membentuk suatu *database* (Minarni & Susanti, 2014).

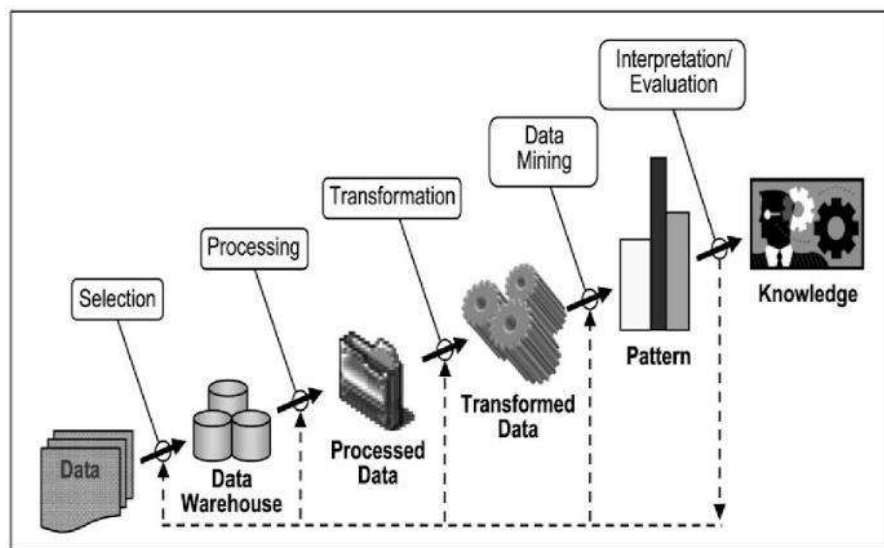
Dapat disimpulkan bahwa data merupakan bentuk yang masih mentah yang belum dapat bercerita banyak, sehingga perlu diolah lebih lanjut. Data diolah melalui suatu model untuk menghasilkan informasi (Lipursari, 2013).

2.1.3 Data Mining

Data mining adalah serangkaian proses untuk menggali nilai tambah dari suatu kumpulan data berupa pengetahuan yang selama ini tidak diketahui (Pramadhani & Setiadi, 2014: 832). Sedangkan dalam Penelitian (Yuli, 2017: 214), menurut Gartner Group, data mining adalah proses menemukan hubungan baru yang mempunyai arti, pola dan kebiasaan dengan memilah-milah sebagian besar data yang disimpan dalam media penyimpanan dengan menggunakan teknologi pengenalan pola seperti teknik statistik dan matematika.

Dapat disimpulkan bahwa *data mining* adalah penambangan atau penemuan informasi baru dengan mencari pola atau aturan tertentu dari sejumlah data yang sangat besar. Data mining juga disebut sebagai serangkaian proses untuk menggali nilai tambah berupa pengetahuan yang selama ini tidak diketahui secara manual dari suatu kumpulan data. Penggalan data merupakan salah satu cara

yang cukup efektif untuk mengetahui adanya serangkaian pola informasi dari sejumlah besar data yang ada. Data mining, sering juga disebut sebagai *knowledge discovery in database* (KDD). KDD adalah kegiatan yang meliputi pengumpulan, pemakaian data, historis untuk menemukan keteraturan, pola atau hubungan dalam set data berukuran besar (Sari, 2016:49).



Gambar 2. 1 *Data Mining* sebagai Salah Satu Tahap KDD

(Sumber: Syahfitri, 2017: 19)

Alur kerja dari Data Mining berisi beberapa tahap, yaitu :

1. *Data Collection*

Pada tahap pengumpulan, data sering disimpan di database atau data *warehouse* untuk selanjutnya di proses.

2. *Feature extraction and data cleaning*

Ketika data sudah terkumpul, banyak bentuk data yang tidak cocok digunakan untuk diolah. Untuk membuat data sesuai dengan format agar data dapat diolah, penting untuk mengubah data menjadi format yang sesuai

dengan algoritma data mining. Hasil akhir dari prosedur ini adalah sebuah kumpulan data set yang terstruktur, yang dapat secara efektif digunakan oleh program komputer.

3. *Analytical processing and algorithms*

Bagian akhir dari proses Data Mining adalah untuk merancang metode analisis yang efektif dari data yang diolah.

4. *Preprocessing*

Preprocessing adalah suatu proses atau langkah yang dilakukan untuk membuat data mentah menjadi data yang berkualitas serta menjadi input yang baik bagi data mining tools. Tahap *preprocessing* terbagi ke dalam beberapa tahap, yaitu :

a. *Data Cleaning*

Data Cleaning berfungsi untuk membersihkan data dengan mengisi nilai – nilai yang hilang, *smoothing noise* data, mengidentifikasi dan menghapus *outlier*, serta menyelesaikan inkonsisten

b. *Data Integration*

Data Integration merupakan penggabungan data dari berbagai *database* ke dalam satu *database* baru. *Data integration* dapat membantu mengurangi dan menghindari redundansi data dan inkonsistensi dalam kumpulan data yang dihasilkan. Ini dapat membantu meningkatkan akurasi dan kecepatan proses data mining berikutnya.

c. *Data Reduction*

Data Reduction dapat diterapkan untuk mengurangi ukuran data dengan menggabungkan atau menghilangkan data yang tidak dibutuhkan.

d. *Data Transformation*

Data Transformation berfungsi mengubah data ke dalam format yang sesuai untuk diproses dalam *data mining*. Beberapa metode *data mining* membutuhkan format data yang khusus sebelum bisa diaplikasikan (Syahfitri, 2017)

2.1.4 Pohon keputusan (*Decision Tree*)

Sebuah pohon keputusan adalah sebuah struktur yang dapat digunakan untuk membagi kumpulan-kumpulan data yang besar menjadi himpunan-himpunan record yang lebih kecil dengan menerapkan serangkaian aturan keputusan. Dengan masing-masing rangkaian pembagi, anggota himpunan hasil menjadi mirip satu dengan yang lainnya. Data dalam pohon keputusan biasanya dinyatakan dalam bentuk tabel dengan *atribut* dan *record*. Atribut menyatakan suatu parameter yang dibuat sebagai kriteria dalam pembentukan pohon. Misalkan untuk menentukan main tenis, kriteria yang diperhatikan adalah cuaca, angin dan temperatur. Salah satu atribut merupakan atribut yang menyatakan data solusi per item data yang disebut target atribut. Atribut memiliki nilai-nilai yang dinamakan dengan *instance*. Misalkan atribut cuaca mempunyai *instance* berupa cerah, berawan dan hujan. Proses pada pohon keputusan adalah mengubah bentuk data (tabel) menjadi model pohon, mengubah model pohon menjadi rule, dan menyederhanakan rule (Yuli, 2017:216).

pohon keputusan yang memperlihatkan faktor-faktor kemungkinan atau probabilitas yang akan mempengaruhi alternatif-alternatif keputusan tersebut, disertai dengan estimasi hasil akhir yang akan didapat bila kita mengambil alternatif keputusan tersebut (Pramadhani & Setiadi, 2014: 832). Pohon keputusan (*Decision tree*) terdapat 3 jenis node yaitu:

- a. *Root Node*, merupakan node paling atas, pada node ini tidak ada input dan bisa tidak mempunyai *output* atau mempunyai *output* lebih dari satu.
- b. *Internal Node*, merupakan node percabangan, pada node ini hanya terdapat satu *input* dan mempunyai *output* minimal dua.
- c. *Leaf node* atau terminal node, merupakan node akhir, pada node ini hanya terdapat satu *input* dan tidak mempunyai *output* (Andri, Kunang, & Murniati, 2013).

2.1.5 Algoritma C4.5

Algoritma C4.5 merupakan algoritma yang sangat populer yang digunakan oleh banyak peneliti di dunia, hal ini dijelaskan oleh Xindong Wu dan Vipin Kumar dalam bukunya yang berjudul *The Top Ten Algorithms in Data Mining*. Algoritma C4.5 merupakan pengembangan dari algoritma ID3 yang di ciptakan oleh J. Rose Quinlan (Yuli, 2017: 216). Algoritma C4.5 mempunyai kelebihan karena dapat menghasilkan model berupa pohon (Andri, Kunang, & Murniati, 2013: 58).

Secara umum Algoritma C4.5 untuk membangun pohon keputusan adalah sebagai berikut :

- a. Pilih atribut sebagai akar
- b. Buat cabang untuk masing-masing nilai
- c. Bagi kasus dalam cabang
- d. Ulangi proses untuk masing-masing cabang sampai semua kasus pada cabang memiliki kelas yang sama (Kamagi & Hansun, 2014).

Algoritme C4.5 digunakan untuk membangun sebuah pohon keputusan yang mudah dimengerti, fleksibel, dan menarik karena dapat divisualisasikan dalam bentuk gambar pohon keputusan. Pohon keputusan adalah model prediksi menggunakan struktur pohon atau struktur berhirarki. Konsep dari pohon keputusan adalah mengubah data menjadi pohon keputusan dan aturan-aturan keputusan (Iswanto, Permanasari, Nugroho, Grafika, & Bulaksumur, 2015).

2.1.6 Waikato Environment for Knowledge Analysis (WEKA)

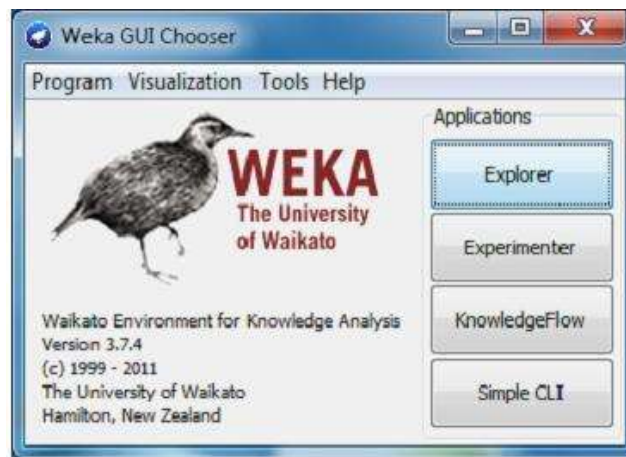
WEKA adalah sebuah paket tools machine learning praktis. *WEKA* merupakan singkatan dari *Waikato Environment for Knowledge Analysis*, yang dibuat di Universitas *Waikato, New Zealand* untuk penelitian, pendidikan dan berbagai aplikasi. *WEKA* mampu menyelesaikan masalah-masalah data mining di dunia nyata, khususnya klasifikasi yang mendasari pendekatan-pendekatan machine learning. Perangkat lunak ini ditulis dalam hirarki class Java dengan metode berorientasi objek dan dapat berjalan hampir di semua platform. *WEKA* mudah digunakan dan diterapkan pada beberapa tingkatan yang berbeda. Tersedia implementasi algoritma-algoritma pembelajaran *state-of-the-art* yang dapat diterapkan pada dataset dari command line. *WEKA* mengandung tools untuk pre-processing data, klasifikasi, regresi, clustering, aturan asosiasi, dan visualisasi. *User* dapat melakukan *preprocess* pada data, memasukkannya dalam sebuah

skema pembelajaran, dan menganalisa classifier yang dihasilkan dan performansinya semua itu tanpa menulis kode program sama sekali. Ada banyak metodologi Data Mining yang dapat dilakukan pada WEKA, salah satu yang populer adalah pohon keputusan (*decision tree*). Pohon keputusan merupakan salah satu metode klasifikasi yang sangat menarik yang melibatkan konstruksi pohon keputusan yang terdiri dari node keputusan yang dihubungkan dengan cabang-cabang dari simpul akar sampai ke node daun (akhir). Pada node keputusan atribut akan diuji, dan setiap hasil akan menghasilkan cabang. Setiap cabang akan diarahkan ke node lain atau ke node akhir untuk menghasilkan suatu keputusan (Pujiono, Amborowati, Suyanto, & Kunci, 2013).

Pengetahuan mendalam tentang penambangan data itu alasannya alat penambangan data yang sangat populer. Weka juga menyediakan antarmuka pengguna grafis dari pengguna dan menyediakan banyak fasilitas. *The GUI Chooser* terdiri dari empat tombol — satu untuk masing-masing dari empat aplikasi utama Weka.

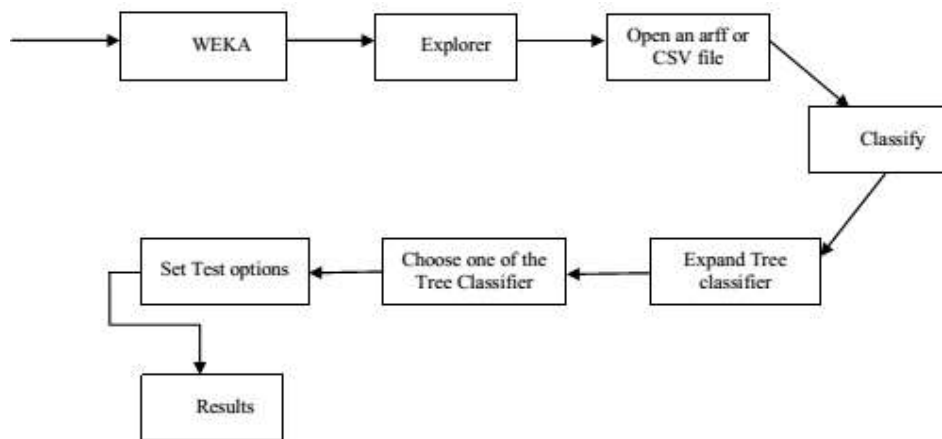
1. *Explorer*: Ini adalah antarmuka utama di Weka. Ini memiliki satu set panel, yang masing-masing dapat digunakan untuk melakukan tugas tertentu. Setelah dataset sudah dimuat, salah satu panel lain di *Explorer* dapat digunakan untuk melakukan analisis lebih lanjut.
2. *Experimenter*: Lingkungan untuk melakukan eksperimen dan melakukan uji statistik di antara skema pembelajaran.

3. Aliran Pengetahuan: Lingkungan ini mendukung fungsi yang sama dengan *Explorer*, tetapi dengan antarmuka tarik -dan jatuhkan. Salah satu keunggulannya adalah mendukung pembelajaran tambahan.
4. CLI Sederhana: Menyediakan antarmuka baris perintah sederhana yang memungkinkan pelaksanaan langsung perintah WEKA untuk sistem operasi yang tidak menyediakan antarmuka baris perintahnya sendiri.



Gambar 2.2 Tampilan depan weka

Eksekusi dalam weka- Eksekusi di weka adalah proses selangkah demi selangkah. Pertama adalah pemuatan data, Data dapat dimuat dari berbagai sumber, termasuk file, URL, dan basis data. WEKA memiliki kapasitas untuk membaca dalam format ".csv". pertama kita mengambil excel datasheet dari dunia nyata, baris pertama berisi nama atribut (dipisahkan oleh koma) diikuti oleh setiap baris data dengan nilai atribut yang tercantum dalam urutan yang sama (juga dipisahkan oleh koma), mengkonversi dalam format *file* .csv. Dari pada pergi ke tombol jelajahi di weka dan simpan file .csv ini. Setelah data dimuat ke WEKA, kumpulan data secara otomatis disimpan ke dalam format ARFF.



Gambar 2.3 Eksekusi di alat weka

Memilih data dari file, setelah data dimuat, WEKA akan mengenali atribut dan selama pemindaian data akan menghitung beberapa statistik dasar pada setiap atribut. daftar atribut yang diakui, sedangkan panel atas menunjukkan nama-nama relasi dasar (tabel) dan relasi kerja saat ini. panel kiri akan menampilkan statistik dasar pada atribut tersebut. Klik pada atribut apa pun. Untuk atribut kategori, frekuensi untuk setiap nilai atribut ditampilkan; sedangkan untuk atribut kontinyu kita dapat memperoleh min, max, mean, standar deviasi, dll.

Siapkan data yang akan ditambang, memilih atribut dari *file* data sampel, setiap *record* secara individu diidentifikasi oleh atribut dan menggunakan *filter* Atribut di WEKA. Pada panel "*Filter*", klik tombol filter (di sebelah kiri tombol "Tambah"). Ini akan menampilkan jendela popup dengan daftar filter yang tersedia. Gulir ke bawah daftar dan pilih `weka.filters.AttributeFilter` "

Setelah menyetel filter, buka tombol klasifikasi dan klik di atasnya. Ini akan menampilkan jendela munculan dengan daftar algoritme klasifikasi, perluas pohon keputusan ini dan pilih pohon mana yang ingin Anda eksperimen.

Sebelum implementasi database beberapa terminologi penting adalah :

1. N - Jumlah total instance rahasia.
2. *True Positive* (TP) - benar memprediksi kelas positif.
3. *True Negative* (TN) - prediksi kelas negatif dengan benar.
4. *False Positive* (FP) - salah prediksi sebagai kelas positif.
5. *False Negative* (FN) - total salah prediksi sebagai kelas negatif.
6. *False Positive Rate* (FPR) - negatif dengan benar diklasifikasikan / total negatif.
7. *True Positive Rate* (TPR) - positif benar diklasifikasikan / total positif
(Sewaiwar & Verma, 2015).

2.2 Penelitian Terdahulu

Berikut ini daftar beberapa penelitian terdahulu yang digunakan sebagai bahan referensi dalam penelitian ini.

Tabel 2.1 Penelitian terdahulu

No	Peneliti Dan Tahun	Judul	Hasil
1	Yuli, mardi. (Yuli, 2017: 214) Vol. 2 Issn : 2407 - 0491	Data mining : klasifikasi menggunakan algoritma c4.5	Proses data mining hingga menghasilkan sebuah pohon Keputusan adalah sebagai berikut : 1. Sumber data, merupakan database yang didalamnya terdapat informasi yang bisa

			<p>diambil dan dimanfaatkan untuk kepentingan bisnis dan penelitian.</p> <p>2. Proses KDD, merupakan proses yang dilakukan untuk mengambil informasi yang terdapat dalam database, di antara proses tersebut terdapat proses data mining.</p> <p>3. Data mining, data mining merupakan bagian dari proses kdd,</p>
2	<p>Andri, Yesi Novaria Kunang , srimurniati. Seminar Nasional Informatika. (Andri, Kunang, & Murniati, 2013: 58). UPN "Veteran" ISSN: 1979-2328</p>	<p>Implementasi teknik datamining untukmemprediksi tingkat kelulusanmahasiswa pada universitas bina darma palembang Andri1),</p>	<p>Berdasarkan hasil penelitian dan pembahasan yang telah dijelaskan maka dapat disimpulkan bahwa atribut tempat lahir adalah atribut yang memiliki nilai gain tertinggi, sehingga atribut ini menjadi root pada decision tree. Rata-rata klasifikasi algoritma j48</p>

			<p>dalam melakukan klasifikasi data mencapai akurasi diatas 90%. Hal ini menunjukkan bahwa algoritma tersebut memiliki performa yang handal dalam melakukan klasifikasi.</p> <p>A-62</p>
3	<p>Aline Embun Pramadhani (08018360), Tedy Setiadi (0407016801) Volume 2 Nomor 1, Februari 2014 E-ISSN: 2338- 5197</p>	<p>Penerapan data mining untuk klasifikasi prediksi penyakit ispa (infeksi saluran pernapasan akut) dengan algoritma decision tree (id3)</p>	<p>Perangkat lunak yang dihasilkan mampu menerapkan proses data mining untuk mengklasifikasikan penyakit ispa dengan atribut yang paling berpengaruh yang pertama adalah gizi sebesar 240.9 kemudian atribut yang berpengaruh kedua adalah imunisasi sebesar 18.7 dan yang terakhir adalah atribut usia sebesar 10.7. Dengan demikian, atribut yang paling tidak terpengaruh</p>

			seseorang terkena ispa adalah jenis kelamin.
4	Hidayatulah himawan, oliver s. Simanjuntak, agus triawan Upn "veteran" yogyakarta, 14 november 2015 ISSN: 1979-2328	Diagnosa tingkat kesehatan pasien menggunakan metode decision tree	Dari hasil penelitian ini, maka diambil kesimpulan sebagai berikut : 1. Jumlah rokok perhari di dapat $q_1 = 0,97$, $q_2 = 0,94$, $q_3 = -0,05$ 2. Dengan entropy atribut yang dihasilkan sebesar -0,2 maka proses penghitungan jumlah rokok perhari dapat dihasilkan dengan entropy tersebut. 3. Aplikasi untuk mendiagnosa tingkat kesehatan pasien menghasilkan information gain atribut panas sebesar 0,79. Sehingga proses analisa tingkat kesehatan diakibatkan oleh penyakit dahak yang lebih mengarah kepada bronkhitis.

			Daftar
5	Ulva Febriana, M. Tanzil Furqon, Bayu Rahayudi Vol. 2, No. 3, Maret 2018, e- ISSN: 2548-964X http://j-ptiik.ub.ac.id	Klasifikasi Penyakit Typhoid Fever (TF) dan Dengue Haemorrhagic Fever (DHF) dengan Menerapkan Algoritma Decision Tree C4.5 (Studi Kasus : Rumah Sakit Wilujeng Kediri) Ulva	Algoritma decision tree C4.5 dapat diterapkan dalam proses klasifikasi penyakit Typhoid Fever (TF) dan Dengue Haemorrhagic Fever (DHF) dimana dilakukan beberapa tahapan proses pembentukan decision tree yaitu dengan menghitung nilai entropy dan gain.
6	M.Nizar Syarif Hamidi Vol 5, Ed 2, Oktober 2014 Issn :20880030	Hubungan Pola Makan Dengan Kejadian Penyakit Hipertensi Di Puskesmas Kuok Tahun 2014	Dari hasil penelitian maka dapat ditarik kesimpulan pada penelitian ini sebagian besar pola makan responden berada pada kategori tidak sehat, sebagian besar responden menderita hipertensi, ada hubungan yang signifikan antara pola makan dengan kejadian penyakit hipertensi
7	Rajeev	Hipertensi: Faktor	Peninjauan narasi kami

	<p>Guptaa,Denis Xavierb</p> <p>G Model IHJ 1409 No. Of Pages 8 Indian 2018 ISSN : 00194832</p>	<p>risiko penyakit tidak menular yang paling penting di India</p>	<p>menyoroti data dari registrar general of india, 5,6 who, 18,19 gbd, 20 dan indian gbd27 penelitian menunjukkan bahwa penyakit kardiovaskular adalah penyebab kematian dan kecacatan yang paling penting di negara ini.</p>
8	<p>WHO, 2017 ISBN 978-92-4- 156548-6</p>	<p>World health statistics 2017: monitoring health for the sdgs, Sustainable Development Goals</p>	<p>World Health Statistics 2017 berfokus pada Tujuan Pembangunan Berkelanjutan (sdgs) yang terkait dengan kesehatan dan kesehatan dan target terkait dengan menyatukan data pada berbagai indikator SDG yang relevan.</p> <p>Selain itu, sejumlah statistik telah diperoleh dari data yang dihasilkan dan dipelihara oleh organisasi internasional lainnya,</p>

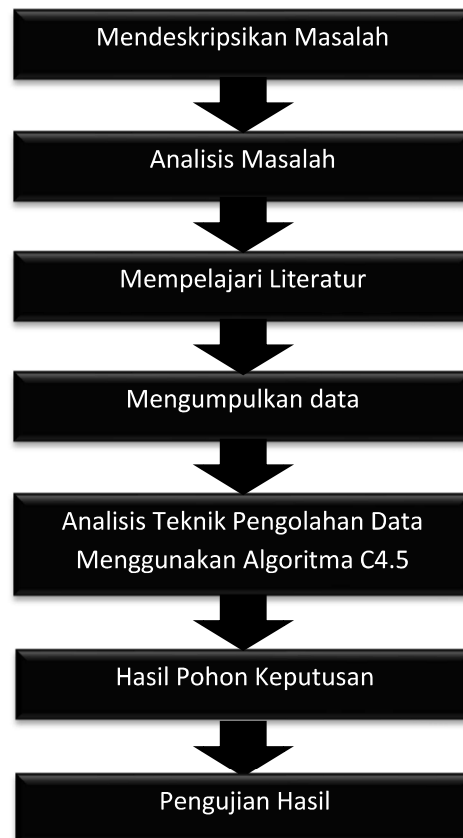
			seperti Departemen Urusan Ekonomi dan Sosial PBB (UNDESA) dan Divisi Populasinya.
9	Rajeev Gupta, Denis Xavier Eternal Heart Care Centre & Research Institute, Jagatpura Road, Jawahar Circle, Jaipur 302017 India. 2018	Hypertension: The most important non communicable disease risk factor in India	Peninjauan narasi kami menyoroti data dari Registrar General of India, 5,6 WHO, 18,19 GBD, 20 dan Indian GBD27 Penelitian menunjukkan bahwa penyakit kardiovaskular adalah penyebab kematian dan kecacatan yang paling penting di negara ini. Tujuan SDG nomor 3 difokuskan pada kesehatan dan Bagian 3.4 telah merekomendasikan pengurangan penyakit tidak terkait kematian oleh sepertiga melalui strategi

			<p>pencegahan dan pengobatan.⁷² Komisi Lancet pada Hipertensi¹⁶ dan World Heart Federation⁶⁰ telah memberikan peta jalan untuk mencapai tujuan ini . Kami percaya bahwa untuk mencapai target SDG PBB dan WHO, Ini juga akan membantu pembuat kebijakan India mencapai target Tujuan Pembangunan Berkelanjutan PBB.</p>
10	<p>M.nizar syarif hamidi Vol 5, ed 2, oktober 2014 Issn 20880030</p>	<p>Hubungan pola makan dengan kejadian penyakit hipertensi di puskesmas kuok tahun 2014</p>	<p>Dari hasil penelitian maka Dapat ditarik kesimpulan pada penelitian ini sebagian besar pola makan responden berada pada kategori tidak sehat, sebagian besar Responden menderita</p>

			Hipertensi, ada hubungan yang signifikan antara pola makan dengan kejadian penyakit hipertensi.
11	Anik Andriani Seminar Nasional Teknologi Informasi dan Komunikasi 2013 (SENTIKA 2013) Yogyakarta, 9Maret 2013 ISSN: 2089-9815 SISTEM	Sistem pendukung keputusan berbasis decision tree dalam pemberian beasiswa studi kasus: amik “bsi yogyakarta”	Penerapan sistem pendukung keputusan dapat mempercepat Waktu Pengambilan keputusan Dalam menentukan pemberian beasiswa kepada mahasiswa yang mengajukan Beasiswa Di Kampus AMIK “BSI Yogyakarta” Gambar

2.3 Kerangka Pemikiran.

Adapun kerangka pemikiran pada penelitian ini adalah :



Gambar 2.4 Kerangka pemikiran

2.4 Hipotesis

Berdasarkan pada penjelasan di bab sebelumnya, maka penulis dalam penelitian ini mendapati hipotesis yaitu, pengelolaan data dengan menggunakan data mining C4.5 dapat memprediksikan pola penyakit hipertensi.

usia, jenis kelamin, lama dirawat, diabetes, hipoglikemia, gagal jantung, stroke, gagal ginjal.

Metodelogi Penelitian dilakukan oleh peneliti secara sistematis sehingga menjadi pedoman dalam meneliti agar hasil yang dihasilkan dapat sesuai dengan yang telah ditetapkan.

Berikut adalah desain penelitian yang merupakan langkah – langkah dalam penyelesaian masalah dalam penelitian.



Gambar 3.1 Desain Penelitian

3.1.1 Pemilihan variabel

Langkah pertama adalah langkah terpenting dalam penelitian yaitu penentuan variabel yang akan diteliti dengan mendefinisikan variabel yang akan diolah menjadi sebuah pohon keputusan.

Peneliti mencoba dan mencari cara untuk menggambarkan pola dan kecenderungan dalam data sehingga mendapatkan pohon keputusan untuk mengklasifikasi pola penyakit Hipertensi kedalam kategori hipertensi ringan atau hipertensi berat berdasarkan variabel yang akan diolah yaitu :

1. Usia
2. Jenis kelamin
3. Lama rawatan
4. Diabetes
5. Hipoglikemia
6. Gagal jantung
7. *Stroke*
8. Gagal ginjal

3.1.2 Klasifikasi Data

Klasifikasi data adalah tahap yang dilakukan untuk pengklasifikasian data sehingga dapat mengklasifikasikan penyakit hipertensi sesuai dengan katagori keputusannya yaitu, hipertensi ringan atau hipertensi berat.

Berikut pengklasifikasian data berdasarkan variabel dalam penelitian adalah sebagai berikut :

1. Usia dikategorikan berdasarkan range dinas kesehatan menurut Depkes RI (2009) dapat diklasifikasikan sebagai berikut:

Tabel 3.1 Kategori usia

Umur	Klasifikasi
0 – 5 Tahun	Balita
6 – 11 Tahun	Kanak – kanak
12 – 16 Tahun	Remaja Awal
17 – 25 Tahun	Remaja Akhir
26 – 35 Tahun	Dewasa Awal
36 – 45 Tahun	Dewasa Akhir
46 – 55 Tahun	Lansia Awal
56 – 65 Tahun	Lansia Akhir
>65 Tahun	Manula

2. Jenis Kelamin dapat diklasifikasi menjadi Pria dan Wanita.
3. Lama dirawat dirumah sakit di simbolkan dengan A, B, C, D, E, F dan Pengelompokan lama rawatan di klasifikasikan dengan rumus Sturgess (Biologis et al., 2018).

$$\begin{aligned}
 \text{Jumlah Kelas} &= 1 + 3.3 \log n \text{ (jumlah data)} \\
 &= 1 + 3.3 \log (17) \\
 &= 1 + 3.3 (1.53147) \\
 &= 6.0538 \text{ atau } 6
 \end{aligned}$$

Tabel 3.2 Tabel lama rawatan

Lama Rawatan	Klasifikasi
1 – 6 Hari	A
7 – 12 Hari	B
13 – 18 Hari	C
19 - 24 Hari	D
25 – 30 Hari	E
31 – 36 > Hari	F

4. Diabetes diklasifikasikan dengan apakah pasien mengidap penyakit diabetes, dan dibatasi dengan Ya atau Tidak
5. hipogilkemia diklasifikasikan dengan apakah pasien mengidap penyakit hipogilkemia, dan dibatasi dengan Ya atau Tidak
6. Gagal Jantung diklasifikasikan dengan apakah pasien mengidap penyakit gagal jantung, dan dibatasi dengan Ya atau Tidak
7. Stroke diklasifikasikan dengan apakah pasien mengidap penyakit stroke, dan dibatasi dengan Ya atau Tidak
8. Gagal ginjal diklasifikasikan dengan apakah pasien mengidap penyakit gagal ginjal, dan dibatasi dengan Ya atau Tidak.
9. Pada Keputusan diklasifikasikan menjadi Hipertensi berat atau hipertensi ringan dan dibatasi dengan ya atau tidak.

3.1.3 Pilih atribut menjadi akar

Menentukan atribut sebagai akar menghitung nilai informasi gain atribut. Untuk memilih atribut sebagai akar, didasarkan pada nilai gain tertinggi dari atribut – atribut yang ada. Dibutuhkan nilai Entropy untuk menentukan gain tertinggi.

Langkah dalam pembentukan pohon keputusan menggunakan Algoritma C4.5 untuk menyelesaikan masalah pada penelitian ini. Yaitu dengan merumuskan :

1. Menghitung jumlah kasus untuk keputusan Hipertensi Ringan dan jumlah keputusan untuk keputusan Hipertensi Berat.
2. Menghitung Entropy dari semua kasus berdasarkan atribut : Usia, jenis kelamin, lama dirawat dan klasifikasi penyakit Diabetes, hipoglikemia, gagal jantung, stroke, gagal ginjal, hipovolemia, dengan rumus sebagai berikut :

$$Entropy(S) = \sum_{i=1}^n - p_i * \log_2 p_i \quad (1)$$

Rumus 3.1 Rumus *Entropy*

Keterangan :

S : Himpunan kasus

A: fitur

n : Jumlah partisi ke – i

pi : Proporsi dari Si Terhadap S

3. Setelah itu dilakukan perhitungan Gain untuk setiap atribut

$$Gain(S, A) = Entropy(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} * Entropy(S_i)$$

(2)

Rumus 3.2 Rumus *Gain*

Keterangan :

S : Himpunan kasus

A: Atribut

n : Jumlah partisi ke – i

|S_i| : jumlah kasus pada partisi ke – i

|S| : Jumlah kasus dalam S

3.1.4 Buat cabang

Dengan demikian hasil perhitungan nilai *gain* untuk tiap atribut yang tertinggi akan menjadi akar dari pohon keputusan

3.1.5 Bagi kasus

Dari nilai atribut nilai yang tertinggi kemudian harus dibagi kasus berdasarkan keputusan, dan harus dilakukan perhitungan mendalam terhadap atribut sehingga menjadi akar.

3.1.6 Ulangi

Dari hasil perhitungan atribut dengan gain tertinggi maka akan menjadi node cabang dari nilai atribut tersebut. Dan nilai atribut yang belum menjadi menjadi node maka masih perlu dilakukan perhitungan lagi.

Proses partisi pohon keputusan akan berhenti disaat :

1. Semua tupelo dalam record dalam simpul m mendapat kelas yang sama
2. Tidak ada atribut dalam record yang dipartisi lagi
3. Tidak ada record didalam cabang yang kosong. (Iswanto et al., 2015)

3.1.7 Membuat pohon keputusan

Sumber data, merupakan database yang didalamnya terdapat informasi yang bisa diambil dan dimanfaatkan untuk kepentingan bisnis dan penelitian. Dapat diketahui bahwa atribut dengan gain tertinggi dapat menjadi node cabang dari nilai atribut, setelah itu dilakukan perhitungan untuk tiap – tiap atribut dengan mengklasifikasikan kasus berdasarkan keputusan yang ada.

3.2 Operasional Variabel

Variabel dapat didefinisikan sebagai atribut seseorang atau obyek, yang mempunyai variasi antara satu objek dengan objek lainnya. (Sugiono, 2014). Adapun variabel yang akan di olah dalam penelitian ini telah dijelaskan pada sub bab pemilihan variabel diatas.

3.3 Populasi dan Sampel

3.3.1 Populasi

Populasi adalah wilayah generalisasi yang terdiri atas obyek atau subyek yang mempunyai kualitas dan karakteristik tertentu yang ditetapkan oleh peneliti untuk dipelajari dan kemudian ditarik kesimpulannya (Sugiono, 2014) populasi pada penelitian ini adalah data penyakit pasien hipertensi dari instalasi rawat inap yang diambil dari data rawatan BPJS (Badan penyelenggara Jaminan Sosial) yang ada pada system aplikasi INACBG pada tahun 2016.

3.3.2 Sampel

Sampel adalah sebagian objek yang diambil dari keseluruhan objek yang diteliti dan dianggap mewakili seluruh bagian dari jumlah populasi(Sugiono, 2014).

3.4 Teknik Pengumpulan Data

Teknik pengumpulan data yang digunakan peneliti adalah teknik Dokumentasi untuk mencari data mengenai variabel yang berupa catatan, transkrip, buku dan sebagainya. (Sugiono, 2014) data di gali dilakukan guna mencari informasi tentang pola penyakit pasien. Hasil yang didapat dari teknik pengumpulan data adalah data penyakit pasien tahun 2016.

3.5 Metode Analisis Data

Data yang diperoleh kemudian diolah menggunakan Algoritma C4.5 sehingga mendapatkan pohon keputusan (decision tree) dan menghasilkan rules pola penyakit hipertensi.

3.6 Lokasi dan Jadwal Penelitian

Penelitian ini dilakukan pada RSUD Embung Fatimah Kota Batam yang beralamatkan di jalan R. Soeprpto Blok D 1 – 9 Batuaji, Kota Batam Kepulauan Riau pada Bulan april 2018. Berikut adalah jadwal penelitian yang peneliti lakukan pada saat melakukan kegiatan penelitian.

Tabel 3.3 Jadwal penelitian

No	KEGIATAN	BULAN															
		April				Mei				Juni				Juli			
		Minggu ke-				Minggu ke-				Minggu ke-				Minggu ke-			
		1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
1	Dokumentasi	■	■														
2	Pengumpulan Data			■	■												
3	Pemilihan Data					■	■										
4	Evaluasi data							■	■								
5	Analisis C4.5									■	■	■	■				
6	Pengujian data ke WEKA													■	■		
7	Hasil																■

3.6.1 Dokumentasi

Dilakukan pada bulan april 2018 pada minggu pertama dan kedua adapun tahapan yang dilakukan pada saat dokumentasi adalah mengadakan study kelayakan dengan meminta ijin pada pihak rumah sakit untuk mengadakan penelitian.

3.6.2 Pengumpulan Data

Dilakukan pada minggu ketiga dan keempat pada bulan april 2018 dengan melakukan pengumpulan data dari pihak Rumah Sakit RSUD Embung Fatimah Kota Batam.

3.6.3 Pemilihan Data

Dilakukan pada dilakukan pada bulan Mei pada minggu pertama dan kedua dengan melakukan pemilihan data yang akan dipakai dalam penelitian.

3.6.4 Evaluasi Data

Dilakukan pada bulan Mei minggu ketiga dan keempat dengan melakukan evaluasi data yang sudah dipilih pada tahapan pemilihan data.

3.6.5 Analisis data dengan menggunakan Weka C4.5

Dilakukan pada minggu pertama, kedua, ketiga dan keempat pada bulan juni 2018.

3.6.6 Pengujian data pada aplikasi WEKA

Dilakukan pada minggu pertama dan kedua pada bulan juli 2018 dengan menginput data pada aplikasi WEKA supaya mendapat pohon keputusan dari klasifikasi pola penyakit berdasarkan nama, usia, jenis kelamin, berat, jumlah keluhan dan lama dirawat terhadap penyakit hipertensi.

3.6.7 Hasil data pada aplikasi WEKA 3.7

Untuk mendapatkan pohon keputusan pada bulan juli di minggu ketiga dan keempat.

3.6 RSUD Embung Fatimah

Rumah Sakit Umum Daerah Embung Fatimah Kota Batam adalah Rumah Sakit milik Pemerintah Kota Batam Provinsi Kepulauan Riau, berada di lokasi strategis kota Batam, tepatnya di Jalan R. Soeprpto Blok D 1-9 Batu Aji Kota Batam, telpon 0778-364446, Fax 0778-361363, Email : rsud_batam@yahoo.co.id.

Rumah Sakit Umum Daerah Embung Fatimah berangkat dari sebuah Puskesmas Rawat Jalan yang dioperasikan pada tanggal 8 Oktober 1986 bernama Puskesmas Batu Aji. Pada tahun 1988 statusnya berubah menjadi Puskesmas Rawat Inap dengan 6 tempat tidur. Pada tanggal 11 oktober 2004 diresmikan menjadi rumah sakit umum daerah batu aji kelas D berdasarkan SK Menkes RI No. 799/Menkes/SK/VII/2004 dan disahkan oleh Walikota Batam.

Pada tanggal 31 Desember 2009 berubah menjadi Badan Layanan Umum Daerah (BLUD) berdasarkan Surat Keputusan Walikota Batam Nomor KPTS

351/HK/XII/2009 tentang Penetapan Pola Pengelolaan Keuangan. Status rumah sakit umum daerah Batu aji Kota Batam tanggal 30 Mei 2011 berubah nama menjadi Rumah Sakit Umum Daerah Embung Fatimah Kota Batam. Pada tanggal 21 November 2011 RSUD Embung Fatimah Kota Batam berubah Kelas dari Kelas D menjadi Kelas C berdasarkan Keputusan Menteri Kesehatan Republik Indonesia No: Hk.C3.05/I/2615/11, dan telah terakreditasi lima (5) Pelayanan oleh KARS sesuai Sertifikat No: KARS-SERT/193/XII/2011 pada tanggal 09 Desember 2011.

Seiring dengan makin beragamnya jenis pelayanan dan dokter spesialis yang dimiliki maka pada tanggal 21 Januari 2013 berdasarkan Keputusan Menteri Kesehatan Republik Indonesia No : Hk.02.03/I/0154/2013 Rumah Sakit Umum Daerah Embung Fatimah Kota Batam berubah dari Kelas C menjadi kelas B. Kemudian rumah sakit ini terus meningkatkan pelayanannya dan pada tanggal 5 Januari 2016 rumah sakit ini telah memenuhi standar akreditasi dan dinyatakan Lulus Tingkat PARIPURNA oleh KARS sesuai Sertifikat No. KARS-SERT/185/I/2016. Rumah Sakit Umum Daerah Embung Fatimah Kota Batam dibangun diatas lahan seluas 38.000m² dengan luas bangunan 11.392 m². Luas lahan yang ada sangat memadai untuk pengembangan rumah sakit dan peningkatan layanan. Bangunan Rumah Sakit Umum Daerah Embung Fatimah Kota Batam dibagi kedalam tiga Blok yaitu: blok 1 gedung administrasi yang terdiri dari 2 lantai, Blok 2 gedung Rawat inap yang terdiri dari 4 lantai dan blok 3 gedung IGD, rawat jalan dan penunjang yang terdiri dari 2 lantai serta penambahan pengembangan gedung baru

3.6.1 Visi Rumah Sakit

“ Menjadi Rumah Sakit Unggulan Dengan Pelayanan Prima “

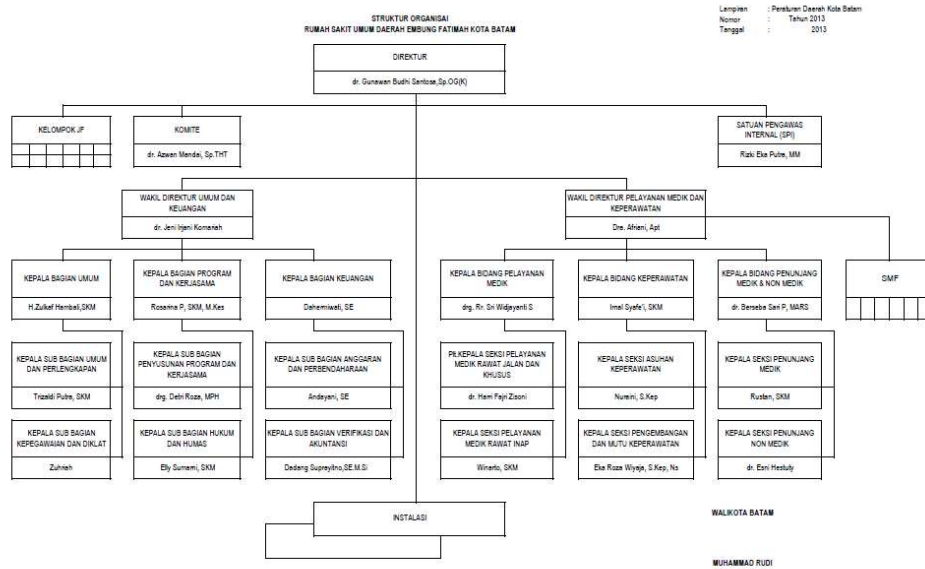
3.6.2 Misi Rumah Sakit

- a. Meningkatkan Profesionalisme dan Kualitas SDM
- b. Meningkatkan dan Mengembangkan Sarana dan Prasarana BLUD / RSUD Sesuai Standart
- c. Meningkatkan Pelayanan yang Cepat , Tepat , Ramah dan Terjangkau
- d. Mengembangkan BLUD / RSUD Sebagai Pusat Penanganan Kecelakaan Lalu Lintas

3.6.3 Moto Rumah Sakit

“ Melayani Dengan Kasih “

3.6.4 Struktur Organisasi



Gambar 3.2 Struktur Organisasi