

TUGAS AKHIR

PENERAPAN *DATA MINING* DALAM KLASIFIKASI HUBUNGAN ANTARA KEJADIAN KATARAK DENGAN DIABETES MELITUS MENGGUNAKAN ALGORITMA C4.5



Oleh

**YOSSIE YOLANDA
NIM 09111003043**

**JURUSAN SISTEM INFORMASI
FAKULTAS ILMU KOMPUTER
UNIVERSITAS SRIWIJAYA
2015**

HALAMAN PENGESAHAN

**PENERAPAN *DATA MINING* DALAM KLASIFIKASI
HUBUNGAN ANTARA KEJADIAN KATARAK DENGAN
DIABETES MELITUS MENGGUNAKAN ALGORITMA C4.5**

TUGAS AKHIR

Sebagai salah satu syarat untuk menyelesaikan studi
di Program Studi Sistem Informasi Reguler Jenjang Strata 1

Oleh :

Yossie Yolanda 09111003043

**Mengetahui,
Ketua Jurusan Sistem Informasi**

**Fathoni, MMSI
NIP. 197210182008121001**

Indralaya, September 2015

Pembimbing

**Firdaus, M.Kom
NIP. 197801212008121003**

HALAMAN PERSETUJUAN

Telah diuji dan lulus pada:

Hari : Rabu

Tanggal : 2 September 2015

Tim Penguji :

1. Ketua (Pembimbing) : Firdaus, M.Kom _____
2. Anggota : Jaidan Jauhari, M.T _____
3. Anggota : Yudha Pratomo, Ph.D _____

Mengetahui
Ketua Jurusan Sistem Informasi,

Fathoni, MMSI
NIP. 197210182008121001

HALAMAN PERSEMBAHAN



MOTTO

- ✓ “Dia memberikan hikmah (ilmu yang berguna) kepada siapa yang dikehendaki-Nya. Barang siapa yang mendapat hikmah itu Sesungguhnya ia telah mendapat kebajikan yang banyak. Dan tiadalah yang menerima peringatan melainkan orang-orang yang berakal”. (Q.S. Al-Baqoroh: 269).
- ✓ Jangan menganggap remeh diri sendiri, karena setiap orang memiliki kemungkinan yang tidak terhingga (Portgas D Ace).
- ✓ Jangan kamu ceritakan beban hidupmu kecuali kepada orang-orang yang bisa membantu dengan pikiran dan nasihat untuk mencapai sebuah kebahagiaan.

PERSEMBAHAN

Dengan mengucapkan syukur Alhamdulillah kepada Allah SWT yang telah melimpahkan rahmat dan rezekinya, kupersembahkan karya kecilku ini untuk orang-orang yang kusayangi :

- ✓ Keluargaku tercinta, Ayahku Yusron dan Ibuku Renna Shopiana serta adikku Yoggie Al Hanif.
- ✓ Dosen pembimbing yang telah memberi masukan yang bermanfaat terhadap penulisan tugas akhir ini.
- ✓ Teman-teman seperjuanganku Sistem Informasi 2011 Fakultas Ilmu Komputer Universitas Sriwijaya.

KATA PENGANTAR

Puji dan syukur penulis panjatkan kehadiran Allah SWT karena atas hidayah dan rahmat-Nya jugalah Laporan Tugas Akhir ini yang berjudul **“PENERAPAN *DATA MINING* DALAM KLASIFIKASI HUBUNGAN ANTARA KEJADIAN KATARAK DENGAN DIABETES MELITUS MENGGUNAKAN ALGORITMA C4.5”** dapat diselesaikan sebagai syarat untuk penyelesaian studi di Program Studi Sistem Informasi pada Fakultas Ilmu Komputer Universitas Sriwijaya.

Dalam penulisan ini penulis menyadari bahwa tanpa adanya bimbingan, bantuan, dukungan serta petunjuk dari semua pihak tidak mungkin laporan ini dapat diselesaikan. Oleh karena itu pada kesempatan ini penulis ingin menyampaikan rasa terima kasih setulus-tulusnya kepada :

1. Allah SWT yang telah memberikan nikmat kesehatan dan hidayah kepada penulis hingga dapat menyelesaikan laporan Tugas Akhir ini.
2. Bapak Dr. Darmawijoyo, M.Si., M.Sc. selaku Dekan Fakultas Ilmu Komputer Universitas Sriwijaya.
3. Bapak Firdaus, M.Kom selaku Pembimbing yang senantiasa memberikan bimbingannya kepada penulis.
4. Orangtua yang sangat berjasa kepada penulis ayah penulis Yusron dan ibu penulis Renna Shopiana serta adik penulis Yoggie Al Hanif yang senantiasa mendoakan dan memberikan dukungan kepada penulis.
5. Kurniawan Saputra yang telah banyak memberikan dukungan berupa motivasi dan doa kepada penulis.

6. Selli Apriani dan Utari Amalita yang telah membantu dalam penyelesaian tugas akhir ini.
7. Teman – teman Sistem Informasi Angkatan 2011 yang tidak dapat penulis sebutkan satu persatu.
8. Dan semua pihak yang telah membantu dalam penyelesaian Tugas Akhir yang tidak dapat penulis sebutkan satu persatu.

Akhir kata penulis menyadari sepenuhnya bahwa laporan ini masih terdapat banyak kekurangannya. Namun demikian harapan penulis kiranya dapat bermanfaat bagi kita semua.

Semoga Allah SWT melimpahkan taufik dan hidayah-Nya bagi kita semua. Aamiin.

Indralaya, September 2015

Yossie Yolanda
NIM 09111003043

PENERAPAN *DATA MINING* DALAM KLASIFIKASI HUBUNGAN ANTARA KEJADIAN KATARAK DENGAN DIABETES MELITUS MENGUNAKAN ALGORITMA C4.5

Oleh

**Yossie Yolanda
09111003043**

ABSTRAK

Kegiatan operasional yang terjadi di RS Khusus Mata Provinsi Sumatera Selatan dapat menghasilkan dan mengumpulkan banyaknya data rekam medis setiap hari. Berdasarkan laporan rekam medis tahun 2014 tercatat penyakit katarak menjadi salah satu dari 10 penyakit terbesar rawat jalan dan belum diketahui penyebabnya. Dengan menerapkan *data warehouse* sebagai persiapan data awal yang akan digunakan pada *data mining*, penelitian ini bertujuan untuk mengklasifikasikan penyebab katarak berdasarkan diabetes melitus menggunakan algoritma C4.5. Metodologi yang digunakan adalah *Nine Methodogy* Kimball untuk *data warehouse* dan *decision tree* mengikuti tahapan *Cross Industry Standar Process for Data Mining* (CRISP-DM) untuk *data mining*. Sedangkan pengujian sistem menggunakan *Black Box Testing* untuk mengukur apakah sistem berjalan dengan baik sesuai dengan yang diharapkan. Hasil penelitian ini akan memberikan informasi dalam menentukan apakah pasien katarak disebabkan karena diabetes melitus.

Kata Kunci: *data mining*, algoritma C4.5, katarak, diabetes melitus

**DATA MINING IMPLEMENTATION ON RELATION CLASSIFYING
BETWEEN CATARACT WITH DIABETES MELITUS USING
C4.5 ALGORITM**

Oleh

**Yossie Yolanda
09111003043**

ABSTRACT

Operations that occur in the Special Eye Hospital of South Sumatra Province can produce and collect the number of medical record data every day. Based on the medical records of 2014 report noted cataract disease to be one of the 10 largest disease outpatient and unknown causes. In preparation of the initial data used in data mining by implementing data warehouse, this study aims to classify methodology used Methodology Kimball for data warehouse and decision tree to follow the stages Standard Cross Industry Process for Data Mining (CRISP-DM) for data mining. While testing the system using the Black Box Testing to measure whether the system is running properly in accordance with the expected. Results of this research will provide information to determine whether the patient's cataract caused by diabetes mellitus.

Keyword: *data mining, C4.5 algoritm, cataract, diabetes melitus.*

DAFTAR ISI

	Halaman
HALAMAN JUDUL	i
HALAMAN PENGESAHAN	ii
HALAMAN PERSETUJUAN	iii
HALAMAN PERSEMBAHAN	iv
KATA PENGANTAR	v
ABSTRAK	vii
DAFTAR ISI	ix
DAFTAR GAMBAR	xii
DAFTAR TABEL	xiv
DAFTAR PERSAMAAN	xvi
DAFTAR LAMPIRAN	xvii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Tujuan dan Manfaat	4
1.2.1 Tujuan Penelitian	4
1.2.2 Manfaat Penelitian	4
1.3 Batasan Masalah	5
BAB II TINJAUAN PUSTAKA	6
2.1 Hubungan Katarak dengan Diabetes Melitus	6
2.2 <i>Data Warehouse</i>	8
2.2.1 Pengertian <i>Data Warehouse</i>	8
2.2.2 Karakteristik <i>Data Warehouse</i>	9
2.2.3 Bentuk <i>Data Warehouse</i>	10
2.2.3.1 <i>Data Warehouse</i> Fungsional	10
2.2.3.2 <i>Data Warehouse</i> Terpusat	10
2.2.3.3 <i>Data Warehouse</i> Terdistribusi	11
2.2.4 Metodologi <i>Data Warehouse</i>	12
2.2.5 Skema <i>Data Warehouse</i>	15
2.2.5.1 Tabel Fakta	15
2.2.5.2 Tabel Dimensi	15
2.2.5.3 Skema Bintang	15
2.2.5.4 Skema Bola Salju	17
2.2.5.5 Skema Galaksi	18
2.3 <i>Data Mining</i>	19
2.3.1 Pengertian <i>Data Mining</i>	19
2.3.2 Arsitektur <i>Data Mining</i>	19
2.3.3 Metodologi <i>Data Mining</i>	20
2.3.4 Pengelompokan <i>Data Mining</i>	23
2.4 Klasifikasi	24
2.4.1 Pengertian Klasifikasi	24
2.4.2 Model Klasifikasi	24

2.5	<i>Decision Tree</i>	25
2.5.1	Pengertian <i>Decision Tree</i>	25
2.5.2	Konsep <i>Decision Tree</i>	26
2.6	Algoritma C4.5	26
BAB III	METODE PENELITIAN	29
3.1	Waktu dan Pelaksanaan	29
3.2	Kerangka Pikir Penelitian	29
3.3	Teknik Pengumpulan Data	29
3.3.1	Jenis Data	29
3.3.2	Sumber Data	30
3.3.3	Metode Pengumpulan Data	30
3.4	Metode Analisis Data	31
3.4.1	Metodologi <i>Data Warehouse</i>	31
3.4.2	Skema <i>Data Warehouse</i>	35
3.4.3	Metodologi <i>Data Mining</i>	36
BAB IV	ANALISIS DATA	39
4.1	Analisis Pengumpulan Data	39
4.1.1	Wawancara	39
4.1.2	Observasi	39
4.1.3	Pengambilan Data	39
4.2	Metodologi <i>Data Warehouse</i>	39
4.2.1	Pemilihan Proses	40
4.2.2	Pemilihan <i>Grain</i>	40
4.2.3	Identifikasi dan Penyesuaian Dimensi	40
4.2.4	Pemilihan Fakta	42
4.2.5	Menyimpan Pre-Kalkulasi dalam Tabel Fakta	42
4.2.6	Melengkapi Tabel Dimensi	42
4.2.7	Memilih Durasi dari <i>Database</i>	46
4.2.8	Melacak Perubahan dari Dimensi Secara Perlahan	46
4.2.9	Memilih Prioritas <i>Query</i> dan Mode <i>Query</i>	46
4.2.9.1	<i>ETL Design</i>	46
4.2.9.2	Analisis Kapasitas Media Penyimpanan	59
4.3	Skema <i>Data Warehouse</i>	63
4.3.1	Skema Bintang	63
4.3.2	Skema Galaksi	65
4.3.3	Metadata	66
4.4	Metodologi <i>Data Mining</i>	69
4.4.1	<i>Business Understanding Phase</i>	69
4.4.2	<i>Data Understanding Phase</i>	69
4.4.3	<i>Data Preparation Phase</i>	72
4.4.4	<i>Modelling Phase</i>	76
4.4.5	<i>Evaluation</i>	97
BAB V	HASIL DAN PEMBAHASAN	99
5.1	Hasil	99
5.1.1	Hasil <i>Data Warehouse</i>	99
5.1.2	Hasil <i>Data Mining</i>	100

5.2	Pembahasan.....	101
5.2.1	Halaman Login	101
5.2.2	Halaman Data Pasien	102
5.2.3	Halaman OLAP Rawat Jalan	107
5.2.4	Halaman OLAP Laboratorium.....	111
5.2.5	Halaman Proses Mining.....	116
5.2.6	Halaman Perhitungan C4.5	116
5.2.7	Halaman Pohon Keputusan.....	117
5.2.8	Halaman Penentu Keputusan	118
5.3	Pengujian Sistem.....	119
BAB VI KESIMPULAN DAN SARAN		123
6.1	Kesimpulan	123
6.2	Saran.....	124
DAFTAR PUSTAKA		125

DAFTAR GAMBAR

	Halaman
Gambar 2.1 Diagram Proporsi Jumlah Pasien Katarak.....	7
Gambar 2.2 <i>Data Warehouse</i> Fungsional	10
Gambar 2.3 <i>Data Warehouse</i> Terpusat	11
Gambar 2.4 <i>Data Warehouse</i> Terdistribusi	12
Gambar 2.5 Skema Bintang	16
Gambar 2.6 Skema Bola Salju	17
Gambar 2.7 Skema Galaksi.....	18
Gambar 2.8 Ilmu <i>Data Mining</i>	19
Gambar 2.9 Arsitektur <i>Data Mining</i>	20
Gambar 2.10 Siklus CRISP-DM.....	21
Gambar 2.11 Representasi Pemodelan dengan Klasifikasi.....	25
Gambar 2.12 Pohon Keputusan / <i>Decision Tree</i>	25
Gambar 3.1 Kerangka Pikir Penelitian.....	29
Gambar 4.1 Data Pasien.....	47
Gambar 4.2 ETL Dimensi Pasien	48
Gambar 4.3 ETL Dimensi Umur.....	49
Gambar 4.4 Data Riwayat Pasien.....	51
Gambar 4.5 ETL Dimensi Riwayat Pasien	51
Gambar 4.6 ETL Dimensi Riwayat Keluarga Pasien.....	52
Gambar 4.7 Data List Nama Penyakit.....	54
Gambar 4.8 ETL Dimensi Diagnosa	54
Gambar 4.9 ETL Dimensi Pemeriksaan	55
Gambar 4.10 ETL Dimensi Waktu	57
Gambar 4.11 ETL Fakta Rawat Jalan	58
Gambar 4.12 ETL Fakta Laboratorium.....	59
Gambar 4.13 Skema Bintang <i>Data Mart</i> Rawat Jalan.....	64
Gambar 4.14 Skema Bintang <i>Data Mart</i> Laboratorium	65
Gambar 4.15 Skema Galaksi OLAP RSKMATA.....	65
Gambar 4.16 <i>Data Selection</i>	70
Gambar 4.17 ETL Tabel Kasus Katarak	70
Gambar 4.18 Konsep Data <i>Decision Tree</i> Kasus Katarak	76
Gambar 4.19 Pohon Keputusan Hasil Perhitungan Node 1	82
Gambar 4.20 Pohon Keputusan Hasil Perhitungan Node 1.1	87
Gambar 4.21 Pohon Keputusan Hasil Perhitungan Node 1.1.3	91
Gambar 4.22 <i>Decision Tree</i> Node 1.2.....	96
Gambar 4.23 Pohon Keputusan Akhir	97
Gambar 4.24 Hasil <i>Supplied Test Set</i> dengan RapidMiner	98
Gambar 5.1 Halaman Login.....	101
Gambar 5.2 Halaman Gagal Login	102
Gambar 5.3 Halaman Gagal Akses	102
Gambar 5.4 Halaman Data Pasien	103
Gambar 5.5 Form Input Data Pasien.....	104
Gambar 5.6 Tabel View Data Pasien	106
Gambar 5.7 Halaman Edit Data Pasien.....	106

Gambar 5.8	Total Pasien RJ berdasarkan Diagnosa	107
Gambar 5.9	Total Pasien RJ berdasarkan Jenis Kelamin.....	108
Gambar 5.10	Total Pasien RJ berdasarkan Kelompok Umur	109
Gambar 5.11	Total Pasien RJ berdasarkan Diagnosa	110
Gambar 5.12	Total Pasien Lab berdasarkan Waktu.....	111
Gambar 5.13	Total Pasien Lab berdasarkan Jenis Kelamin.....	112
Gambar 5.14	Total Pasien Lab berdasarkan Kelompok Umur	113
Gambar 5.15	Total Pasien Lab berdasarkan Riwayat Pasien.....	114
Gambar 5.16	Total Pasien Lab berdasarkan Jenis Pemeriksaan	115
Gambar 5.17	Halaman Proses Mining	116
Gambar 5.18	Halaman Hasil Perhitungan Proses Mining	116
Gambar 5.19	Halaman Pohon Keputusan	117
Gambar 5.20	Halaman Penentu Keputusan	118

DAFTAR TABEL

	Halaman
Tabel 4.1	Matrik <i>Gain</i> dan Dimensi pada <i>Data Mart</i> Rawat Jalan.....41
Tabel 4.2	Matrik <i>Gain</i> dan Dimensi pada <i>Data Mart</i> Laboratorium41
Tabel 4.3	<i>Rounding Out Dimension</i>42
Tabel 4.4	Dimensi Waktu43
Tabel 4.5	Dimensi Pasien44
Tabel 4.6	Dimensi Umur44
Tabel 4.7	Dimensi Riwayat Pasien.....44
Tabel 4.8	Dimensi Riwayat Keluarga.....45
Tabel 4.9	Dimensi Diagnosa45
Tabel 4.10	Dimensi Pemeriksaan45
Tabel 4.11	Durasi <i>Database</i>46
Tabel 4.12	Contoh Perubahan Dimensi Umur.....46
Tabel 4.13	Keterangan Atribut pada Data Pasien.....48
Tabel 4.14	Klasifikasi Umur50
Tabel 4.15	Keterangan Atribut pada Data Riwayat Pasien51
Tabel 4.16	Keterangan Atribut pada Data Diagnosa Rawat Jalan54
Tabel 4.17	Klasifikasi Hasil Pemeriksaan56
Tabel 4.18	Analisis Kapasitas Media Penyimpanan Tabel Fakta61
Tabel 4.19	Analisis Kapasitas Media Penyimpanan Tabel Dimensi.....63
Tabel 4.20	Metadata Dimensi Waktu67
Tabel 4.21	Metadata Dimensi Pasien67
Tabel 4.22	Metadata Dimensi Umur67
Tabel 4.23	Metadata Dimensi Riwayat Pasien.....67
Tabel 4.24	Metadata Dimensi Riwayat Keluarga.....68
Tabel 4.25	Metadata Dimensi Diagnosa.....68
Tabel 4.26	Metadata Dimensi Pemeriksaan68
Tabel 4.27	Metadata Fakta Rawat Jalan.....69
Tabel 4.28	Metadata Fakta Laboratorium69
Tabel 4.29	Atribut pada Data Target71
Tabel 4.30	Hasil Perhitungan <i>Chi Square Person</i>71
Tabel 4.31	Hasil Proses <i>Data Preparation</i>73
Tabel 4.32	Klasifikasi Glukosa Darah.....75
Tabel 4.33	Perhitungan <i>Node 1</i>81
Tabel 4.34	Perhitungan <i>Node 1.1</i>86
Tabel 4.35	Perhitungan <i>Node 1.1.3</i>90
Tabel 4.36	Perhitungan <i>Node 1.2</i>95
Tabel 4.37	<i>Rules</i> dari Hasil Pohon Keputusan98
Tabel 5.1	Pengujian <i>Black Box</i>119
Tabel 5.2	Halaman Login120
Tabel 5.3	Halaman Data Pasien.....120
Tabel 5.4	Halaman OLAP Rawat Jalan.....121
Tabel 5.5	Halaman OLAP Laboratorium121
Tabel 5.6	Halaman Proses Mining121
Tabel 5.7	Halaman Perhitungan122

Tabel 5.8	Halaman Pohon Keputusan	122
Tabel 5.9	Halaman Penentu Keputusan.....	122
Tabel 5.10	Halaman Logout	122

DAFTAR PERSAMAAN

	Halaman
Persamaan 2.1 Perhitungan Kebutuhan Penyimpanan <i>Record</i> Tabel Fakta	14
Persamaan 2.2 Perhitungan Kebutuhan Penyimpanan <i>Record</i> Tabel Dimensi....	14
Persamaan 2.3 Perhitungan <i>Num Rows</i>	14
Persamaan 2.4 Perhitungan <i>Num Col</i>	14
Persamaan 2.5 Perhitungan <i>Fixed Data Size</i>	14
Persamaan 2.6 Perhitungan <i>Null Bitmap</i>	15
Persamaan 2.7 Perhitungan <i>Row Size</i>	15
Persamaan 2.8 Perhitungan <i>Row Per Page</i>	15
Persamaan 2.9 Perhitungan <i>Num Of Pages</i>	15
Persamaan 2.10 Perhitungan <i>Num Of Bytes</i>	15
Persamaan 2.11 Perhitungan <i>Num Of Kbytes</i>	15
Persamaan 2.12 Perhitungan Nilai <i>Gain</i>	27
Persamaan 2.13 Perhitungan Nilai <i>Entropy</i>	27
Persamaan 2.14 Perhitungan Nilai <i>Gain Ratio</i>	28
Persamaan 2.15 Perhitungan Nilai <i>Split Info</i>	28

DAFTAR LAMPIRAN

	Halaman
Lampiran 1	Form Kartu Konsultasi A-1
Lampiran 2	Form Rekomendasi Sidang Komprehensif.....B-1
Lampiran 3	SK Pembimbing.....C-1
Lampiran 4	Surat Balasan Penelitian D-1
Lampiran 5	Form Wawancara.....E-1
Lampiran 6	Hasil Observasi Aplikasi PROSIRS F-1
Lampiran 7	Data 10 Penyakit Terbesar Rawat Jalan G-1
Lampiran 8	Data Pasien H-1
Lampiran 9	Data Riwayat Diabetes Pasien..... I-1
Lampiran 10	Data Diagnosa Rawat JalanJ-1
Lampiran 11	Data Pemeriksaan Laboratorium K-1
Lampiran 12	Data Kasus Katarak L-1
Lampiran 13	Hasil Perhitungan Chi Square Person dengan SPSSM-1
Lampiran 14	Hasil <i>Data Mining</i> menggunakan RapidMiner N-1
Lampiran 15	Dokumentasi Pengujian <i>Black Box</i> O-1

BAB I

PENDAHULUAN

1.1 Latar Belakang

RS Khusus Mata Provinsi Sumatera Selatan merupakan Unit Pelaksana Teknis Dinas Kesehatan Provinsi Sumatera Selatan yang bergerak di bidang jasa untuk melayani kesehatan khususnya mata bagi masyarakat Palembang. RS Khusus Mata Provinsi Sumatera Selatan memiliki tugas untuk memberikan laporan hasil rekam medis ke Dinas Kesehatan Provinsi Sumatera Selatan yang dilakukan setiap triwulan. Data rekam medis yang diberikan oleh pihak rumah sakit merupakan data yang perlu diolah untuk mengetahui informasi yang terkandung di dalamnya

Laporan rekam medis tersebut bisa memberikan informasi bagi dinas kesehatan. Laporan dari data rekam medis tersebut yang saat ini dijadikan oleh Dinas Kesehatan Provinsi Sumatera Selatan untuk melakukan kebijakan-kebijakan apabila akan memberikan penyuluhan kepada masyarakat. Mengenai penyebab penyakit yang diderita oleh masyarakat masih belum digali untuk dijadikan acuan apabila melakukan penyuluhan atau pencegahan penyakit.

Data rekam medis yang ada di RS Khusus Mata Provinsi Sumatera Selatan saat ini hanya tersimpan di *database* dan sebatas memberikan statistik jumlah pasien yang berobat dengan penyakit yang dideritanya. Laporan yang dihasilkan dari aplikasi PROSIRS (formulir laporan terlampir) belum dapat menampilkan informasi secara ringkas, sehingga menyulitkan pihak (rumah sakit, dinas kesehatan) untuk menemukan informasi dari laporan tersebut dengan cepat.

Dengan adanya *data warehouse* diharapkan data dalam aplikasi PROSIRS yang telah ditransformasikan ke dalam *data warehouse* dapat mempermudah dalam mengolah data menjadi informasi. Penelitian yang dilakukan oleh (Triesia, Pramiarta, & Rahman, 2013), mengenai pemanfaatan *data warehouse* yang dapat menyimpan data secara tersusun dan mempermudah dalam menganalisa data untuk mengambil sebuah keputusan. Penelitian ini bertujuan untuk memberikan informasi yang jelas mudah dimengerti dan sesuai dengan kebutuhan, seperti menganalisa data pasien rawat inap dan rawat jalan RS. Muhamadyah Palembang.

Berdasarkan hasil survei awal berdasarkan laporan formulir RL54, bahwa penyakit katarak merupakan salah satu dari 10 penyakit terbesar rawat jalan di RS Khusus Mata Provinsi Sumatera Selatan. Dalam penelitian (Rizkawati, 2012), menyatakan katarak bisa disebabkan oleh usia, komplikasi penyakit mata, pasca-operasi, trauma, herediter, infeksi intrauterin dan penyakit sistemik seperti diabetes melitus. Masih dalam penelitian (Rizkawati, 2012), menyebutkan penelitian dan studi klinik mengenai katarak telah banyak dilakukan diantaranya adalah penelitian dan studi mengenai hubungan antara penyakit diabetes melitus dengan katarak. Seong Il Kim dan Sung Jin Kim dari *Department of Ophthalmology, Eulji University School of Medicine, Seoul, Korea* telah melakukan penelitian yang berjudul Prevalensi dan Faktor Risiko Katarak pada Penderita Diabetes Melitus Tipe 2. Hasil penelitian tersebut menyatakan bahwa diabetes melitus merupakan faktor resiko yang paling signifikan dalam menimbulkan katarak. Selain itu, *Wisconsin Epidemiologic Study of Diabetic Retinopathy* juga telah melakukan penelitian mengenai katarak dan menemukan insiden katarak pada penderita diabetes melitus.

Pada bidang kesehatan, *data mining* dapat dimanfaatkan untuk memprediksi suatu penyakit dari data rekam medis pasien. Beberapa penelitian sudah dilakukan terkait penerapan *data mining* untuk memprediksi suatu penyakit. Penelitian mengenai implementasi teknik *data mining* di bidang kesehatan yang dilakukan (Rodiyatul & Tama, 2013), bertujuan untuk menemukan informasi yang berharga menggunakan teknik *data mining* untuk membantu pihak pengambil keputusan di bidang kesehatan dalam memahami *rules* yang mungkin terjadi didalam diagnosa DM tipe II. Penelitian ini mengklasifikasikan atribut yang menjadi prediksi diagnosa penyakit diabetes melitus tipe II yaitu plasmainsulin, berat badan, riwayat keturunan, dan hasil pemeriksaan gula darah puasa pada pasien RSMH Palembang. Penelitian lain mengenai *data mining* di bidang kesehatan juga yang dilakukan oleh (Wulandari, 2013), yang bertujuan untuk menerapkan *data mining* dalam memprediksi penyakit hipertensi dalam kehamilan dengan metode *decision tree* menggunakan C4.5 berdasarkan data hasil pemeriksaan kehamilan pada pasien RSIA YK Madira Palembang. Adapun atribut yang digunakan dalam memprediksi penyakit hipertensi pada kehamilan yaitu usia, riwayat hipertensi, berat badan dan paritas.

Dari beberapa penelitian sebelumnya mengenai hubungan penyakit katarak dengan diabetes melitus dan penerepan *data mining* dalam memprediksi penyakit, penulis tertarik untuk menuangkan penelitiannya dalam penulisan tugas akhir yang berjudul **“Penerapan *Data Mining* dalam Klasifikasi Hubungan antara Kejadian Katarak dengan Diabetes Melitus menggunakan Algoritma C4.5”**.

1.2 Tujuan dan Manfaat

1.2.1 Tujuan Penelitian

Tujuan dari tugas akhir ini adalah :

1. Menerapkan *data warehouse* sebagai tahapan awal dalam mempersiapkan data yang akan digunakan pada proses data mining.
2. Menerapkan *data mining* dalam mengklasifikasikan penyebab katarak berdasarkan diabetes melitus pada data pasien Rumah Sakit Khusus Mata Provinsi Sumatera Selatan dengan teknik decision tree menggunakan algoritma C4.5.

1.2.2 Manfaat Penelitian

Adapun manfaat penelitian yang dapat diperoleh dalam tugas akhir ini adalah memberikan pengetahuan baru yang sebelumnya tersembunyi dari *data warehouse* menjadi informasi yang dapat membantu proses *data mining* dalam mengklasifikasikan penyebab penyakit katarak yang disebabkan oleh diabetes melitus sehingga informasi tersebut dapat digunakan pihak Rumah Sakit Khusus Mata Provinsi Sumatera Selatan dan Dinas Kesehatan Provinsi Sumatera Selatan dapat melakukan tindakan medis guna membantu masyarakat dalam pencegahan penyakit katarak yang disebabkan oleh diabetes melitus.

1.3 Batasan Masalah

Berikut ini adalah batasan masalah yang dibuat agar pengerjaan tugas akhir mencapai tujuan yang ditetapkan :

1. Sumber data yang digunakan adalah data pasien RS Khusus Mata Provinsi Sumatera Selatan selama 1 tahun (2014).
2. Pengolahan data ini terpusat pada data rawat jalan dan pemeriksaan laboratorium di RS Khusus Mata Provinsi Sumatera Selatan.
3. Menggunakan metode *Nine Methodology* pendekatan Kimball untuk penerapan *data warehouse* dan menggunakan teknik *decision tree* (pohon keputusan) dengan algoritma C4.5 mengikuti tahapan metode *Cross-Industry Standard process for Data Mining* (CRISP-DM) untuk *data mining*.
4. Penyakit yang akan diprediksi pada *data mining* adalah penyakit katarak berdasarkan hasil pemeriksaan glukosa pada data pasien RS Khusus Mata Provinsi Sumatera Selatan.
5. Menggunakan *tools* Pentaho yang merupakan aplikasi *data warehouse* untuk menampilkan informasi laporan dan aplikasi *data mining* berbasis website untuk menampilkan keputusan dalam memprediksi apakah pasien termasuk penderita katarak karena diabetes melitus atau bukan.

DAFTAR PUSTAKA

- Connolly, T., & Begg, C. (2005). *Database System A Practical Approach to Design, Implementation, and Management*. United States of America: Addison Wesley.
- Han, J., & Kamber, M. (2006). *Data Mining Concepts and Techniques*. United States of America: Morgan Kaufmann Publisher.
- Hidayanto, A. N., Hapsari, I. C., Jiwanggi, M. A., & Fitria, D. (2013). *Perbandingan Algoritma Klasifikasi Dalam Pendeteksian Penyakit Kanker*. Depok: Fasilkom Universitas Indonesia.
- Indonesia, M. K. (2008). *Peraturan Menteri Kesehatan Republik Indonesia tentang Rekam Medis*. Departemen Kesehatan.
- Indonesia, M. K. (2011). *Peraturan Menteri Kesehatan Republik Indonesia tentang Sistem Informasi Rumah Sakit*. Departemen Kesehatan.
- Inmon, W. H. (2005). *Building the Data Warehouse, Fourth Edition*. Canada: Wiley Publishing, Inc.
- Kimball, R., & Ross, M. (2002). *The Data Warehouse Toolkit The Complete Guide to Dimensional Modeling* (2nd Edition ed.). Canada: John Wiley and Sons, Inc.
- Lane, P. (2002). *Oracle9i Data Warehousing Guide*. Oracle Corporation.
- Larose, D. T. (2005). *Discovering Knowledge in Data*. Canada: John Wiley and Sons, Inc.
- Manaf, A. (2007). *Prediabetes*. Padang: Fakultas Kedokteran Universitas Andalas.
- Paramita, D. (2012). *Penerapan Data Mart Penjualan Hypermarket XYZ menggunakan Metode From Enterprise Models to Dimensional Models*. Palembang: Tugas Akhir Tidak Diterbitkan.
- Perkeni. (2011). *Konsensus Pengendalian dan Pencegahan Diabetes Melitus Tipe 2 di Indonesia 2011*. Perkeni.
- Pratama, F. F. (2014). *Analisis Faktor Risiko Klaim Asuransi Kendaraan menggunakan Algoritma C4.5*. Palembang: Tugas Akhir Tidak Diterbitkan.
- Pressman, R. S. (2001). *Software Engineering: A Practioner's Approach*. McGraw-Hill.

- Raharja, Y. P. (2014). *Rancang Bangun Sistem Rekomendasi Beasiswa menggunakan Algoritma Klasifikasi C4.5 pada Universitas Dian Nuswantoro*. Semarang: Universitas Dian Nuswantoro.
- Rizkawati. (2012). *Hubungan Antara Kejadian Katarak dengan Diabetes Melitus di Poli Mata RSUD Dr. Soedarso Pontianak*. Pontianak: Universitas Tanjung Pura.
- Rodiyatul, & Tama, B. A. (2013). *Implementasi Teknik Data Mining Didalam Analisis Penyakit Diabetes Melitus Tipe II menggunakan Decision Tree*. Palembang: Universitas Sriwijaya.
- Suprajitno. (2004). *Asuhan Keperawatan Keluarga Aplikasi Dalam Praktik*. Jakarta: EGC.
- Tandra, H. (2007). *Segala Sesuatu yang harus Anda Ketahui tentang Diabetes*. Jakarta: PT Gramedia Pustaka Utama.
- Triesia, D., Pramiarta, D., & Rahman, A. (2013). *Perancangan Data Warehouse Pasien pada RS. Muhammadiyah Palembang*. Palembang: STMIK GI MDP.
- Wulandari, R. A. (2013). *Implementasi Data Mining untuk Prediksi Penyakit Hipertensi dalam Kehamilan dengan Teknik Decision Tree Menggunakan Algoritma C4.5*. Palembang: Tugas Akhir Tidak Diterbitkan.
- Zaki, M. J., & Jr, W. M. (2013). *Data Mining and Analysis: Fundamental Concepts and Algorithms*. Cambridge University Press.