

**SISTEM PREDIKSI TINGKAT RESIKO PENYAKIT DIABETES
MELLITUS MENGGUNAKAN TEKNIK DATA MINING
DENGAN METODE DECISION TREE ID3**

SKRIPSI



OLEH:
FAHMI DWI NOVIANTO
10.621.042

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS MUHAMMADIYAH GRESIK
2015**

**SISTEM PREDIKSI TINGKAT RESIKO PENYAKIT DIABETES
MELLITUS MENGGUNAKAN TEKNIK DATA MINING
DENGAN METODE DECISION TREE ID3**

SKRIPSI

Diajukan sebagai syarat memperoleh gelar Sarjana Komputer
Program Studi Teknik Informatika jenjang S-1 Fakultas Teknik
Universitas Muhammadiyah Gresik



OLEH:
FAHMI DWI NOVIANTO
10.621.042

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS MUHAMMADIYAH GRESIK
2015**

PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan dibawah ini:

Nama : Fahmi Dwi Novianto
NIM : 10 621 042
Fakultas/Program studi : Teknik/Teknik Informatika

Menyatakan dengan sebenarnya bahwa skripsi yang berjudul "**SISTEM PREDIKSI TINGKAT RESIKO PENYAKIT DIABETES MELLITUS MENGGUNAKAN TEKNIK DATA MINING DEGAN METODE DECISION TREE ID3**" yang saya buat untuk melengkapi sebagian persyaratan menjadi Sarjana Komputer pada Program Studi Teknik Informatika Fakultas Teknik Universitas Muhammadiyah Gresik, bukan merupakan tiruan atau duplikasi dari skripsi yang sudah dipublikasikan dan atau pernah dipakai untuk mendapatkan gelar kesarjanaan di lingkungan Universitas Muhammadiyah Gresik maupun di Perguruan Tinggi atau Instansi manapun, kecuali bagian yang secara tertulis dikutip dalam naskah ini atau disebutkan dalam sumber kutipan dan daftar pustaka.

Apabila ternyata di kemudian hari, saya terbukti melanggar pernyataan saya tersebut di atas, maka saya bersedia untuk mempertanggungjawabkan serta diproses sesuai peraturan yang berlaku.

Gresik, 5 Januari 2015

FAHMI DWI NOVIANTO

NIM 10 621 042

KATA PENGANTAR



Dengan menyebut asma Allah Yang Maha Pengasih dan Maha Penyayang, puji syukur kehadirat Allah SWT yang hanya atas ijin-Nya penulis dapat menyelesaikan proposal skripsi dengan judul “*Sistem Prediksi Tingkat Resiko Penyakit Diabetes Mellitus Menggunakan Teknik Data Mining Dengan Metode Decision Tree ID3*”.

Melalui kesempatan yang sangat berharga ini penulis menyampaikan ucapan terima kasih yang sebesar-besarnya kepada semua pihak yang telah membantu penyelesaian proposal skripsi ini, terutama kepada :

1. Ibu dan ayah yang senantiasa memberikan dorongan dan mendoakan penulis.
2. Bpk. Harunur Rasyid,S.Kom.,M.Kom selaku pembimbing I dan Bpk. Eko Prasetyo,S.Kom.,M.Kom selaku Pembimbing II, yang senantiasa meluangkan waktunya untuk memberikan bimbingan, arahan serta masukan bagi penulis dalam penyelesaian skripsi.
3. Teman – teman Informatika Universitas Muhammadiyah Gresik yang telah banyak membantu, memberi dukungan, semangat dan do'a.

Penulis menyadari bahwa tidak ada sesuatupun yang sempurna kecuali Allah SWT. Oleh karena itu, dengan senang hati penulis menerima kritik dan saran yang membangun dari segala pihak.

Gresik, Januari 2015

Penulis

LEMBAR PERSETUJUAN
SIDANG SKRIPSI

**SISTEM PREDIKSI TINGKAT RESIKO PENYAKIT DIABETES
MELLITUS DENGAN MENGGUNAKAN TEKNIK DATA MINING
DENGAN METODE DECISION TREE ID3**

Oleh

FAHMI DWI NOVIANTO

NIM 10.621.042

Disetujui untuk dipresentasikan dalam sidang skripsi.

Susunan Tim Pembimbing

Pembimbing I

Pembimbing II

Harunur Rosyid, S.T.,M.Kom. Eko Prasetyo, S.Kom.,M.Kom.

NIP 06 210 408 106

NIP 06 240 511 015

Diterima pada tanggal 5 Januari 2015 dan dinyatakan telah memenuhi seluruh
persyaratan sidang skripsi.

Mengetahui,

Ketua Program Studi Teknik Informatika

Universitas Muhammadiyah Gresik

Deni Sutaji,S.kom
NIP. 06 230 909 213

LEMBAR PENGESAHAN SKRIPSI

SISTEM PREDIKSI TINGKAT RESIKO PENYAKIT DIABETES MELLITUS MENGGUNAKAN TEKNIK DATA MINING DENGAN METODE DECISION TREE ID3

Oleh

FAHMI DWI NOVIANTO

NIM 10.621.042

Telah dipertahankan di depan Tim Penguji

Pada tanggal 20 Januari 2015

Susunan Tim Penguji

Penguji I (Ketua)

Soffiana Agustin,S.kom.,M.kom

NIP. 197711292005012001

Penguji III (Anggota)

Harunur Rosyid, S.T.,M.Kom.

NIP 06210408106

Penguji II (Sekretaris)

Misbah, S.T., M.T.

NIP. 197606282005011001

Penguji IV (Anggota)

Eko Prasetyo,S.Kom., M.Kom.

NIP. 06240511015

Diterima dan dinyatakan memenuhi syarat kelulusan pada tanggal 13 Februari 2015

Mengetahui,

Dekan Fakultas Teknik
Universitas Muhammadiyah Gresik

Ketua Program Studi Teknik Informatika
Universitas Muhammadiyah Gresik

Misbah, S.T., M.T.
NIP. 197606282005011001

Deni Sutaji,S.kom
NIP. 06230909213

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN SAMPUL DALAM	ii
LEMBAR PERNYATAAN	iii
LEMBAR PERSETUJUAN	iv
LEMBAR PENGESAHAN	v
KATA PENGANTAR	vi
DAFTAR ISI	vii
DAFTAR GAMBAR	x
DAFTAR TABEL	xii
DAFTAR KODE PROGRAM	xiv
INTISARI.....	xv
ABSTRACT	xvi
BAB I PENDAHULUAN	
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Tujuan Penelitian	3
1.4 Batasan Masalah.....	3
1.5 Metodologi Penelitian	3
1.6 Sistematika Penulisan.....	5
BAB II LANDASAN TEORI	
2.1 Diabetes Mellitus	7
2.2 Data Mining	8
2.2.1 Pengelompokan Data Mining.....	8
2.2.2 Knowledge Discovery in Databases (KDD)	9
2.3 Metode Data Mining Kalsifikasi	11
2.4 Decision Tree (Pohon Keputusan)	11
2.4.1 Model Decision Tree.....	12
2.4.2 Decision Tree ID3	12

2.4.3 Algoritma	12
2.4.4 Contoh Perhitungan.....	14
2.5 Penelitian Sebelumnya	18

BAB III ANALISIS DAN PERANCANGAN SISTEM

3.1 Analisis Sistem.....	21
3.2 Hasil Analisis	21
3.2.1 Data yang diolah	22
3.2.2 Deskripsi Sistem	24
3.3 Representasi Data.....	25
3.4 Perhitungan Decision Tree ID3.....	28
3.5 Perancangan Sistem.....	69
3.5.1 Diagram Context	69
3.5.2 Diagram Berjenjang	70
3.5.3 Data Flow Diagram Level 0.....	71
3.5.4 Data Flow Diagram Level 1	72
3.6 Kebutuhan Pembuatan Sistem.....	72
3.7 Struktur Tabel.....	73
3.8 Desain Antar Muka	76
3.9 Evaluasi Sistem	81
3.10 Skenario Pengujian Sistem.....	84

BAB IV IMPLEMENTASI DAN PENGUJIAN SISTEM

4.1 Implementasi	85
4.1.1 Perangkat keras (Hardware) yang digunakan	85
4.1.2 Perangkat lunak (Software) yang digunakan	85
4.1.3 Batasan Implementasi	86
4.2 Implementasi Sistem Prediksi	86
4.2.1 Perhitungan Gain.....	86
4.2.2 Pemilihan Atribut.....	89
4.2.3 Pemecahan Cabang	90
4.2.4 Cek Kondisi Data	92
4.3 Implementasi Antar Muka.....	92

4.3.1 Antar Muka Login.....	92
4.3.2 Antar Muka Halaman Utama	93
4.3.3 Antar Muka Olah Data	94
4.3.4 Antar Muka Mining	95
4.3.5 Antar Muka Pohon Kpventusan.....	95
4.3.6 Antar Muka Hasil Prediksi.....	96
4.3.7 Antar Muka Data User	97
4.3.8 Antar Muka Prediksi	97
4.4 Pengujian Sistem.....	98
4.4.1 Pengujian menggunakan data uji	98
4.4.2 Pengujian satu data.....	106
4.4.3 Evaluasi hasil pengujian.....	107

BAB V PENUTUP

5.1 Kesimpulan.....	108
5.2 Saran.....	108

DAFTAR PUSTAKA

LAMPIRAN

DAFTAR GAMBAR

Gambar 1.1 Block Diagram	4
Gambar 2.1 Flowchart Algorima Decision Tree ID3.....	13
Gambar 2.2 Contoh pohon keputusan	18
Gambar 3.1 Flowchart Sistem.....	25
Gambar 3.2 Hasil pembentukan cabang pada atribut IMT	35
Gambar 3.3 Percabangan usia pada kasus IMT besar.....	42
Gambar 3.4 Pohon percabangan keturunan kasus IMT sedang	51
Gambar 3.5 Pohon percabangan keturunan kasus IMT kecil	67
Gambar 3.6 Pohon keputusan yang terbentuk.....	68
Gambar 3.7 Diagram Context	69
Gambar 3.8 Diagram Berjenjang	70
Gambar 3.9 DFD Level 0	71
Gambar 3.10 DFD Level 1 proses pembentukan aturan (pohon keputusan)	72
Gambar 3.11 Rancangan halaman login	76
Gambar 3.12 Rancangan halaman utama	76
Gambar 3.13 Rancangan halaman olah data	77
Gambar 3.14 Rancangan halaman mining	77
Gambar 3.15 Rancangan halaman pohon keputusan	78
Gambar 3.16 Rancangan halaman uji pohon keputusan	78
Gambar 3.17 Rancangan halaman hasil	79
Gambar 3.18 Rancangan halaman data user	79
Gambar 3.19 Rancangan halaman prediksi.....	80
Gambar 3.20 Rancangan halaman hasil prediksi	80
Gambar 3.21 Rancangan halaman ubah password.....	81
Gambar 4.1 Antar muka halaman login	93
Gambar 4.2 Antar muka halaman utama admin.....	93
Gambar 4.3 Antar muka halaman utama pengguna atau user.....	94
Gambar 4.4 Antar muka halaman olah data.....	94
Gambar 4.5 Antar muka halaman mining	95

Gambar 4.6 Antar muka halaman pohon keputusan bentuk <i>rule if-then</i>	95
Gambar 4.7 Antar muka halaman uji pohon keputusan	96
Gambar 4.8 Antar muka halaman hasil prediksi	96
Gambar 4.9 Antar muka halaman data user	97
Gambar 4.10 Antar muka halaman prediksi	97
Gambar 4.11 Pohon keputusan pengujian pertama bentuk <i>rule</i> percobaan ke-1 ...	98
Gambar 4.12 Hasil pengujian pertama percobaan ke-1	99
Gambar 4.13 Pohon keputusan pengujian pertama bentuk <i>rule</i> percobaan ke-2.	100
Gambar 4.14 Hasil pengujian pertama percobaan ke-2	100
Gambar 4.15 Pohon keputusan pengujian pertama bentuk <i>rule</i> percobaan ke-3.	101
Gambar 4.16 Hasil pengujian pertama percobaan ke-3	101
Gambar 4.17 Pohon keputusan pengujian pertma bentuk <i>rule</i> percobaan ke-4...	102
Gambar 4.18 Hasil pengujian pertama percobaan ke-4	102
Gambar 4.19 Pohon keputusan pengujian pertama bentuk <i>rule</i> percobaan ke-5.	103
Gambar 4.20 Hasil pengujian pertama percobaan ke-5	103
Gambar 4.21 Pohon keputusan pengujian kedua bentuk <i>rule</i> percobaan ke-1	104
Gambar 4.22 Hasil pengujian kedua percobaan ke-1.....	105
Gambar 4.23 Pohon keputusan pengujian kedua bentuk <i>rule</i> percobaan ke-2	105
Gambar 4.24 Hasil pengujian kedua percobaan ke-2.....	106
Gambar 4.25 Pohon keputusan pengujian kedua bentuk <i>rule</i> percobaan ke-3	106
Gambar 4.26 Hasil pengujian kedua percobaan ke-3.....	107
Gambar 4.27 Pohon keputusan pengujian kedua bentuk <i>rule</i> percobaan ke-4	107
Gambar 4.28 Hasil pengujian kedua percobaan ke-4.....	108
Gambar 4.29 Pohon keputusan pengujian kedua bentuk <i>rule</i> percobaan ke-5	108
Gambar 4.30 Hasil pengujian kedua percobaan ke-5.....	109
Gambar 4.31 Pohon keputusan terbaik pengujian kedua percobaan ke-4	110
Gambar 4.32 Pohon keputusan terbaik pengujian kedua percobaan ke-5	111
Gambar 4.33 Antar muka puskesmas saat melakukan prediksi	114
Gambar 4.34 Antar muka hasil prediksi tingkat resiko penyakit diabetes	114

DAFTAR TABEL

Tabel 2.1 Contoh data set.....	15
Tabel 2.2 Pembagian atribut cuaca cerah.....	17
Tabel 2.3 Pembagian atribut cuaca mendung	17
Tabel 2.4 Pembagian atribut cuaca hujan	18
Tabel 3.1 Pertanyaan kuisioner	23
Tabel 3.2 Tipe data atribut	24
Tabel 3.3 Data keseluruhan.....	26
Tabel 3.4 Data latih	27
Tabel 3.5 Data uji	27
Tabel 3.6 Hasil perhitungan entropy dan gain atribut pada node akar.....	35
Tabel 3.7 Data pada kasus IMT nilai besar.....	35
Tabel 3.8 Data pada kasus IMT nilai sedang	36
Tabel 3.9 Data pada kasus IMT nilai kecil	36
Tabel 3.10 Hasil perhitungan gain pada atribut IMT nilai besar	41
Tabel 3.11 Data pembagian cabang kasus usia.....	41
Tabel 3.12 Hasil perhitungan gain atribut usia di atribut IMT nilai sedang	46
Tabel 3.13 Data pembagian cabang kasus keturunan di atribut IMT sedang	47
Tabel 3.14 Hasil perhitungan gain pemecah cabang kasus keturunan paman bibi	50
Tabel 3.15 Data pemecah cabang kasus olahraga.....	50
Tabel 3.16 Hasil perhitungan gain pada atribut IMT nilai kecil	53
Tabel 3.17 Data pemecah cabang kasus usia	54
Tabel 3.18 Data pemecah cabang kasus keturunan pada IMT kecil	54
Tabel 3.19 Hasil perhitungan gain kasus atribut keturunan di atribut nilai orang tua	59
Tabel 3.20 Data pemecah cabang pada kasus keturunan diabetes	60
Tabel 3.21 Hasil perhitungan gain kasus usia	63
Tabel 3.22 Data pemecah cabang kasus usia	63
Tabel 3.23 Hasil perhitungan gain kasus usia	66
Tabel 3.24 Data pemecah cabang kasus usia	66

Tabel 3.25 Struktur tabel data latih	74
Tabel 3.26 Struktur tabel data uji	74
Tabel 3.27 Struktur tabel hasil prediksi	75
Tabel 3.28 Struktur tabel user	75
Tabel 3.29 Struktur tabel aturan	75
Tabel 3.30 Confusion Matrix	81
Tabel 3.31 Perhitungan Evaluasi Sistem	83
Tabel 4.1 Keterangan pohon keputusan pengujian kedua percobaan ke-4	111
Tabel 4.2 Keterangan pohon keputusan pengujian kedua percobaan ke-5	114
Tabel 4.3 Rincian hasil pengujian pertama	119
Tabel 4.4 Rincian hasil pengujian kedua	121
Tabel 4.5 Evaluasi hasil pengujian pertama	125
Tabel 4.6 Evaluasi hasil pengujian kedua	125

DAFTAR KODE PROGRAM

Kode Program 4.1 Fungsi hitung jumlah data	86
Kode Program 4.2 Perhitungan gain	87
Kode Program 4.3 Pemilihan atribut	89
Kode Program 4.4 Pemecahan cabang	91
Kode Program 4.5 Memanggil fungsi proses_D.....	91
Kode Program 4.6 Cek kondisi data	92

SISTEM PREDIKSI TINGKAT RESIKO PENYAKIT DIABETES MELLITUS MENGGUNAKAN TEKNIK DATA MINING DENGAN METODE DECISION TREE ID3

Oleh
FAHMI DWI NOVIANTO
10 621 042

Diajukan kepada Program Studi Informatika Fakultas Teknik Universitas Muhammadiyah Gresik pada tanggal 20 Januari 2015 untuk memenuhi sebagian persyaratan memperoleh derajat sarjana S-1 Program Studi Teknik Informatika

INTISARI

Diabetes Mellitus adalah penyakit yang mempengaruhi gula darah, hal ini terjadi karena glukosa (gula sederhana) di dalam darah terlalu tinggi. Sehingga tubuh tidak dapat menggunakan insulin dengan benar atau tidak sempurna. Di zaman sekarang ini, kurang sadarnya manusia akan pola hidup teratur yang disibukkan dengan aktivitas sehari-hari yang sangat tinggi, sehingga menyebakan kurangnya perhatian akan pola hidup sehat, dimana ini merupakan penyebab utama manusia terserang berbagai macam penyakit, salah satunya adalah diabetes.

Penelitian ini menerapkan teknik *data mining classification* dengan menggunakan metode *decision tree ID3* untuk memprediksi prestasi tingkat resiko penyakit diabetes. Atribut yang digunakan adalah usia, keturunan diabetes, index masa tubuh (IMT), intensitas olahraga.

Data yang digunakan diambil dari data kuisioner Kusioner dari 50 sampel data yang diambil dari penelitian sebelumnya yaitu tugas akhir Azhuardhi Azhro dengan judul “Aplikasi Perhitungan Tingkat Resiko Penyakit Diabetes Dengan Metode Sugeno” yang digunakan sebagai konstanta nilai dengan kevalidan data yang sudah diuji, tanpa ada pengurangan maupun penambahan data responden. Pengujian sistem dilakukan dengan melakukan sepuluh kali percobaan menggunakan komposisi data yang berbeda-beda untuk mengetahui tingkat akurasi dari masing-masing percobaan. Pohon keputusan yang digunakan adalah hasil dari pembentukan pohon keputusan pada percobaan kesatu dan kedua, karena tingkat akurasinya paling tinggi yaitu 100%.

Kata Kunci: *Data Mining, Classification, Decision Tree ID3.*

Pembimbing I : Harunur Rosyid, S.T.,M.Kom.
Pembimbing II : Eko Prasetyo, S.Kom., M.Kom.

**PREDICTION SYSTEM LEVEL OF RISK DIABETES
MELLITUS USING DATA MINING TECHNIQUES USING
DECISION TREE ID3**

By

FAHMI DWI NOVIANTO

10 621 042

Information submitted to the Faculty of Engineering Program
Muhammadiyah University of Gresik on January 20th 2015 to meet most
requirements of obtaining an undergraduate degree S-1 Engineering Program
Information

ABSTRACT

Diabetes Mellitus is a disease that affects blood sugar, this happens because glucose (a simple sugar) in the blood is too high. So that the body can not use insulin properly or not perfect. In this day and age, less conscious people will regularly being busy lifestyle with daily activities are very high, so that caused a lack of attention to a healthy lifestyle, which is a major cause humans to various diseases, one of which is diabetes.

This study applied the techniques of data mining classification using the ID3 decision tree method to predict diabetes risk level of achievement. Attributes used are age, heredity of diabetes, body mass index (BMI), the intensity of exercise.

The data used are taken from the questionnaire questionnaire data from 50 samples of data taken from previous research that Azhuardhi Azhro thesis entitled "Application Calculation of Risk of Diabetes With Sugeno method" is used as a constant value to the validity of the data that has been tested, with no reduction and the addition of respondent data. System testing is done by ten times the experiment using the composition of different data to determine the accuracy of each experiment. Decision trees used are the result of the formation of a decision tree on the first and second trial, because the highest level of accuracy that is 100%.

Keywords: *Data Mining, Classification, Decision Tree ID3.*

Supervisor : Harunur Rosyid, S.T.,M.Kom.
Co. Supervisor : Eko Prasetyo,S.Kom.,M.Kom.