

**IMPLEMENTASI ALGORITMA C4.5 PADA APLIKASI PREDIKSI
PENYAKIT LIVER BERBASIS WEB**

Tugas Akhir disusun untuk memenuhi syarat
mencapai gelar Kesarjanaan Komputer pada
Program Studi Sistem Informasi
Jenjang Program Strata-1



Oleh :

Mochammad Ibnu Fachruddin

16.01.55.0020

19415

**FAKULTAS TEKNOLOGI INFORMASI
UNIVERSITAS STIKUBANK (UNISBANK)
SEMARANG
[2020]**

PERNYATAAN KESIAPAN UJIAN TUGAS AKHIR

Saya, Mohammad Ibnu Fachruddin, dengan ini menyatakan bahwa Laporan Tugas Akhir yang berjudul:

IMPLEMENTASI ALGORITMA C4.5 PADA APLIKASI PREDIKSI PENYAKIT LIVER BERBASIS WEB

adalah benar hasil karya saya dan belum pernah diajukan sebagai karya ilmiah, sebagian atau seluruhnya, atas nama saya atau pihak lain.

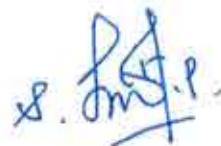


(Mohammad Ibnu Fachruddin)
NIM : 16.01.55.0020

Disetujui oleh Pembimbing

Kami setuju Laporan tersebut diajukan untuk Ujian Tugas Akhir

Semarang, 23 Januari 2020



(Dra. Sulastri, M.Kom)
NIDN : 0608126601



UNIVERSITAS STIKUBANK "UNISBANK" SEMARANG

FAKULTAS TEKNOLOGI INFORMASI

Rokmerot Kampus Maguus
Jl. Lt. 1 Lomba Juring No. 1 Semarang 50241
telp. (024) 8451976, 8311868, 8454746, Fax (024) 8443240
E-mail : Info@unisbank.ac.id

Kampus Bandeng
Jl. Kondong V Boridan Ngisor Semarang
telp. (024) 8414970, Fax (024) 8441738
E-mail : fe@unisbank.ac.id

SURAT PERNYATAAN KEASLIAN TUGAS AKHIR/ SKRIPSI

Yang bertanda tangan di bawah ini, saya menyatakan bahwa TUGAS AKHIR / SKRIPSI dengan Judul :

IMPLEMENTASI ALGORITMA C4.5 PADA APLIKASI PREDIKSI PENYAKIT LIVER BERBASIS WEB

yang telah diuji di depan tim penguji pada tanggal 04 Februari 2020, adalah benar hasil karya saya dan dalam TUGAS AKHIR / SKRIPSI ini tidak terdapat keseluruhan atau sebagian tulisan orang lain yang saya ambil dengan cara menyalin, atau meniru dalam bentuk rangkalan kalimat atau simbol yang saya aku seolah-olah sebagai tulisan saya sendiri dan atau tidak terdapat bagian atau keseluruhan tulisan yang saya salin, tiru atau yang saya ambil dari tulisan orang lain tanpa memberikan pengakuan pada penulis aslinya.

Apabila saya melakukan hal tersebut diatas, baik sengaja maupun tidak, dengan ini saya menyatakan menarik TUGAS AKHIR / SKRIPSI yang saya ajukan sebagai hasil tulisan saya sendiri.

Bila kemudian terbukti bahwa saya ternyata melakukan tindakan menyalin atau meniru tulisan orang lain seolah-olah hasil pemikiran saya sendiri, berarti gelar dan ijazah saya yang telah diberikan oleh Universitas Stikubank (UNISBANK) Semarang batal saya terima.

• Semarang, 4 Februari 2020

Yang Menyatakan



(MOCHAMMAD IBNU FACHRUDDIN)

NIM : 16.01.55.0020

SAKSI 1

Tim Penguji

(SULASTRI, DRA, M.KOM)

SAKSI 2

Tim Penguji

(HERNY FEBRUARIYANTI, ST, M.CS)

SAKSI 3

Tim Penguji

(FELIX ANDREAS SUTANTO, S.KOM., M.CS.)

HALAMAN PENGESAHAN

Tugas Akhir dengan judul:

IMPLEMENTASI ALGORITMA C4.5 PADA APLIKASI PREDIKSI PENYAKIT LIVER BERBASIS WEB

Ditulis oleh :

NIM : 16.01.55.0020
Nama : Mochammad Ibnu Fachruddin

Telah dipertahankan di depan Tim Dosen Penguji Tugas Akhir dan diterima sebagai salah satu syarat guna menyelesaikan Jenjang Program Strata I Program Studi Teknik Informatika pada Fakultas Teknologi Informasi Universitas Stikubank (UNISBANK) Semarang.

Semarang, 10 Februari 2020

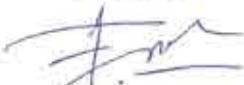
Ketua


(Dra. Sulastri, M.Kom)
NIDN : 0608126601

Sekretaris


(Herny Febriaryanti, ST, M.Cs)
NIDN : 0614027301

Anggota


(Felix Andreas Sutanto, S.Kom., M.Cs.)
NIDN : 0625047801

Mengetahui,

Universitas Stikubank (UNISBANK) Semarang
Fakultas Teknologi Informasi

Dekan


(Kristophorus Hadiono, Ph.D)
NIDN : 0622027601

MOTTO DAN PERSEMPAHAN

MOTTO

- Bekerja keras, Lakukan yang terbaik, Simpan kata-kata Anda Jangan terlalu sombong, Percaya kepada Tuhan, Jangan takut dan jangan pernah lupakan teman.
- Memayu hayuning pribadi, memayu hayuning kulawarga, memayu hayuning sesama, memayu hayuning bawana.

PERSEMPAHAN

- Allah S.W.T yang telah memberikan anugerah hidup, kesehatan dan keselamatan hingga saat ini.
- Kedua orang tua, bapak dan ibu tercinta yang telah melahirkan dan membesarkanku hingga kelak menjadi anak yang berbakti dan berguna.
- kakak, adik beserta keluarga besar saya yang selalu memberi dukungan dan dorongan semangat.
- Bapak dan Ibu Dosen Pengajar Program Studi Sistem Informasi Universitas Stikubank Semarang yang telah mendidik dan membekali ilmu pengetahuan selama masa perkuliahan.
- Seluruh teman-teman Sistem Informasi Angkatan 2016 yang telah membantu, memberi semangat dan masukan hingga terselesaiya Tugas Akhir ini.

**FAKULTAS TEKNOLOGI INFORMASI
UNIVERSITAS STIKUBANK (UNISBANK) SEMARANG**

Program Studi: Sistem Informasi

Tugas Akhir Sarjana Komputer

Semester Ganjil Tahun 2019/2020

**IMPLEMENTASI ALGORITMA C4.5 PADA APLIKASI PREDIKSI
PENYAKIT LIVER BERBASIS WEB**

Mochammad Ibnu Fachruddin

NIM : 16.01.55.0020

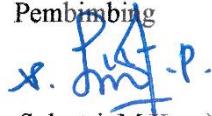
Abstrak

Penyakit liver merupakan peradangan didalam sel-sel hati, yang umumnya diakibatkan oleh infeksi (parasit, bakteri dan virus,), lemak yang berlebih, konsumsi obat-obatan dan alkohol berlebih, serta penyakit autoimmune. Pada stadium awal, penyakit liver cenderung sulit untuk di temukan. Hati dapat menjaga fungsinya bahkan ketika ada bagian yang mengalami kerusakan. Sehingga penaganan penyakit liver pada stadium awal akan memperpanjang usia pasien, untuk itu diperlukan suatu sistem komputerisasi yang mampu membantu dalam memprediksi penyakit liver.

Dalam penelitian ini prediksi dilakukan menggunakan teknik data mining dan metode klasifikasi dengan algoritma *decision tree* C4.5. Serta menggunakan data Indian Liver Patient Dataset (ILPD Dataset) yang diperoleh dari database *UCI Machine Learning Repository*, dengan jumlah data sebanyak 583 record dengan 11 atribut. Metode pengembangan sistem yang digunakan terdiri dari Studi Literatur, Pengumpulan Data, Analisis Model Sistem, Perancangan Sistem, Pembuatan Program serta,Pengujian. Sedangkan perancangan sistem menggunakan UML (*Unified Modeling Language*) dan pembuatan website menggunakan PHP sebagai bahasa pemrograman dan MySql sebagai media database.

Penelitian ini telah menghasilkan aplikasi untuk memprediksi penyakit liver menggunakan metode Decision Tree dengan Algoritma C4.5 berbasis web. Aplikasi dapat mengklasifikasi data pasien penyakit liver untuk menghasilkan keputusan apakah pasien termasuk penderita penyakit liver atau bukan. Pengujian algoritma C4.5 dilakukan menggunakan metode *10 fold cross validation*, dan dihasilkan rata-rata pola prediksi penyakit liver yang terbentuk sebanyak 63 pola (rules) dengan atribut *direct_bilirubin* sebagai atribut berpengaruh dan atribut *sgpt*, *tot_bilirubin*, *albumin* serta *alkphos* sebagai atribut node ke-2. Sedangkan rata-rata nilai akurasi yang dihasilkan sebesar 58,64% dengan akurasi tertinggi 74,14% dan akurasi terendah 41,38%.

Kata Kunci : Penyakit liver, Data Mining, Klasifikasi, C4.5, cross validation

Pembimbing

(Dra. Sulastri, M.Kom)

**FAKULTAS TEKNOLOGI INFORMASI
UNIVERSITAS STIKUBANK (UNISBANK) SEMARANG**

Program Studi: Sistem Informasi

Tugas Akhir Sarjana Komputer

Semester Ganjil Tahun 2019/2020

**IMPLEMENTASI ALGORITMA C4.5 PADA APLIKASI PREDIKSI
PENYAKIT LIVER BERBASIS WEB**

Mochammad Ibnu Fachruddin

NIM : 16.01.55.0020

Abstrak

Liver disease is inflammation in the cells of the liver, which is generally caused by infection (parasites, bacteria and viruses,), excess fat, excessive consumption of drugs and alcohol, and autoimmune disease. In the early stages, liver disease tends to be difficult to find. The liver can maintain its function even when parts are damaged. So that the treatment of liver disease at an early stage will extend the life of the patient, for that we need a computerized system that can help in predicting liver disease.

In this study predictions were made using data mining techniques and classification methods with the decision tree C4.5 algorithm. And using Indian Liver Patient Dataset (ILPD Dataset) data obtained from the UCI Machine Learning Repository database, with a total of 583 records with 11 attributes. System development methods used consist of Literature Studies, Data Collection, Analysis of System Models, System Design, Programming and Testing. While the system design uses UML (Unified Modeling Language) and website creation uses PHP as a programming language and MySql as a media database.

This research has produced applications to predict liver disease using the Decision Tree method with the web-based C4.5 Algorithm. The application can classify the data of liver disease patients to produce a decision whether the patient is a sufferer of liver disease or not. The C4.5 algorithm was tested using the 10 fold cross validation method, and an average of 63 patterns of liver disease prediction were formed with the direct_bilirubin attribute as the influential attribute and the sgpt, tot_bilirubin, albumin and alkphos attributes as the node attribute 2 While the average accuracy value is 58.64% with the highest accuracy of 74.14% and the lowest accuracy of 41.38%.

Kata Kunci : Penyakit liver, Data Mining, Klasifikasi, C4.5, cross validation

Pembimbing



(Dra. Sulastri, M.Kom)

KATA PENGANTAR

Puji syukur kepada Allah SWT atas rahmat dan karunianya sehingga penulis dapat menyelesaikan penelitian dan skripsi yang berjudul : "Implementasi Algoritma C4.5 Pada Aplikasi Prediksi Penyakit Liver Berbasis Web". Skripsi ini disusun untuk memenuhi salah satu syarat kelulusan dalam meraih gelar Sarjana Komputer program strata satu (S-1) Proram Studi Sistem Informasi, Fakultas Teknologi Informasi Universitas Stikubank Semarang.

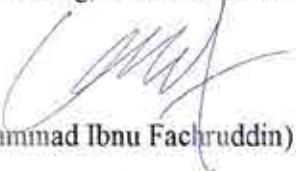
Dalam proses penelitian dan penyusunan skripsi ini tidak luput dari kendala. Kendala tersebut dapat diatasi berkat bantuan, bimbingan dan dukungan dari berbagai pihak. Oleh karena itu, penulis ingin menyampaikan rasa terimakasih sebesar-besarnya kepada :

- 1 Bapak, ibu, kakak, adik beserta keluarga besar saya yang selalu mendukung baik secara moral dan material.
- 2 Bapak Dr. Safik Faozi, S.H., M.Hum selaku Rektor Universitas Stikubank Semarang.
- 3 Bapak Kristophorus Hadiono, Ph.D selaku Dekan Fakultas Teknologi Informasi Universitas Stikubank Semarang.
- 4 Bapak Arief Jananto, S.Kom., M.Cs selaku Ketua Program Studi Sistem Informasi Universitas Stikubank Semarang.
- 5 Bapak Drs. Edy Supriyanto, M.Kom selaku dosen wali yang telah banyak memberikan saran, pengarahan, dan bantuan selama ini.
- 6 Ibu Dra. Sulastri, M.Kom selaku dosen pemimpin yang memberikan banyak pengarahan, bimbingan dan masukan yang sangat berguna bagi penyelesaian skripsi ini.

- 8 Para staf Fakultas Teknologi Informasi Universitas Stikubank Semarang yang telah membantu selama masa perkuliahan berlangsung.
- 9 Seluruh teman-teman Sistem Informasi Angkatan 2016 yang telah membantu, memberi semangat dan masukan hingga terselesainya Tugas Akhir ini.
- 10 Semua pihak yang telah membantu penulis dalam menyelesaikan penulisan skripsi ini yang tidak dapat disebutkan satu-persatu.

Penulis menyadari bahwa dalam penulisan skripsi ini masih banyak kekurangan karena keterbatasan pengetahuan dan pengalaman, untuk itu penulis sangat berharap para pembaca maupun semua pihak terkait memberikan kritik dan saran yang bersifat membangun demi kesempurnaan dalam penelitian selanjutnya.

Semarang, 23 Januari 2020



(Mochammad Ibnu Fachruddin)

DAFTAR ISI

HALAMAN JUDUL.....	i
PERNYATAAN KESIAPAN UJIAN TUGAS AKHIR	ii
SURAT PERNYATAAN KEASLIAN TUGAS AKHIR	iii
HALAMAN PENGESAHAN.....	iv
MOTTO DAN PERSEMBAHAN	v
ABSTRAK	vi
KATA PENGANTAR	vii
DAFTAR ISI	ix
DAFTAR GAMBAR	xv
DAFTAR TABEL.....	xx
DAFTAR RUMUS	xxii
DAFTAR LAMPIRAN.....	xxiii
DAFTAR SINGKATAN	xxiv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Perumusan Masalah	4
1.3 Batasan Masalah	4
1.4 Tujuan dan Manfaat Penelitian	4
1.4.1 Tujuan Penelitian	4
1.4.2 Manfaat Penelitian	5

1.5 Metode Penelitian	5
1.5.1 Studi Literatur.....	5
1.5.2 Pengumpulan Data.....	6
1.5.3 Analisis Model Sistem.....	6
1.5.4 Perancangan Sistem	6
1.5.5 Pembuatan Program.....	6
1.5.6 Pengujian	7
1.5.7 Penarikan Simpulan	7
1.6 Sistematika Penulisan	7
BAB II TINJAUAN PUSTAKA.....	9
2.1 Data Mining	9
2.2 Algoritma C4.5	11
2.3 Penelitian yang Terkait	13
2.4 Perbedaan Penelitian yang dilakukan dengan Penelitian Terdahulu	21
BAB III ANALISIS DAN RANCANGAN SISTEM.....	22
3.1 Analisa Kebutuhan Sistem.....	22
3.1.1 Analisa Kebutuhan Fungsional.....	22
3.1.2 Analisa Kebutuhan Non-fungsional	24
3.1.3 Analisa Kebutuhan Penunjang Sistem.....	24
3.2 Analisa Data Penyakit Liver	25
3.3 Analisa Pengolahan Data	27
3.3.1 Pembersihan Data	28
3.3.2 Transformasi Data	29

3.3.3 Penambangan Data	34
3.4 Perancangan Sistem	35
3.4.1 Diagram Use Case	35
3.4.2 Narasi Usecase.....	36
3.4.3 Diagram Aktivitas.....	45
3.4.4 Class Diagram.....	55
3.4.5 <i>Sequence</i> Diagram	56
3.5 Perancangan Algoritma C4.5	65
3.6 Pengujian Algoritma C4.5	66
3.7 Rancangan <i>Database</i>	67
3.5.1 Tabel User.....	67
3.5.2 Tabel Dataset	67
3.5.3 Tabel Data Training.....	68
3.5.4 Tabel Data Testing.....	69
3.5.5 Tabel Gain Ratio.....	70
3.5.6 Tabel Pohon Keputusan.....	70
3.5.7 Tabel Hasil Prediksi.....	71
3.8 Desain Antarmuka Sistem	72
3.6.1 Desain Menu Daftar.....	72
3.6.2 Desain Menu Login	73
3.6.3 Desain Halaman Dashboard Admin	74
3.6.4 Desain Menu Dataset	75
3.6.5 Desain Menu Data Training.....	76

3.6.6	Desain Menu Data Testing	77
3.6.7	Desain Menu Proses Mining.....	78
3.6.8	Desain Menu Pohon Keputusan.....	79
3.6.9	Desain Menu Hitung Akurasi	80
3.6.10	Desain Menu Prediksi	81
3.6.11	Desain Menu Hasil prediksi.....	82
3.6.12	Desain Dashboard User.....	83
	BAB IV IMPLEMENTASI	84
4.1	Implementasi Database	84
4.1.1	Tabel User.....	84
4.1.2	Tabel Dataset	85
4.1.3	Tabel Data Traning	86
4.1.4	Tabel Data Testing.....	87
4.1.5	Tabel Gain Ratio.....	88
4.1.6	Tabel Pohon Keputusan	88
4.1.7	Tabel Hasil Prediksi.....	89
4.2	Implementasi User Interface	90
4.2.1	Halaman Depan	90
4.2.2	Halaman Daftar.....	91
4.2.3	Halaman Login	92
4.2.4	Halaman Dashboard Admin	93
4.2.5	Halaman Dataset.....	94
4.2.6	Halaman Data Traning.....	95

4.2.7 Halaman Data Testing	96
4.2.8 Halaman Proses Mining.....	97
4.2.9 Halaman Pohon Keputusan.....	98
4.2.10 Halaman Hitung Akurasi	99
4.2.11 Halaman Prediksi	100
4.2.12 Halaman Hasil Prediksi.....	101
4.2.13 Halaman Dashboard User	102
4.3 Implementasi Algoritma C4.5.....	102
4.3.1 Menghitung Entropy Total	103
4.3.2 Menghitung <i>Entropy</i> Setiap Atribut	104
4.3.3 Menghitung Information Gain.....	105
4.3.4 Menghitung Gain Ratio	106
4.3.5 Menghitung Gain Ratio Tertinggi	107
4.3.6 Pembentukan Node Baru	108
4.3.7 Perulangan Proses Perhitungan.....	111
BAB V HASIL PENELITIAN DAN PEMBAHASAN	113
5.1 Hasil Pengujian Aplikasi	113
5.2 Pengujian Program.....	120
5.3 Hasil Pengujian Algoritma C4.5.....	145
BAB VI KESIMPULAN DAN SARAN	178
6.1 Kesimpulan	178
6.2 Saran	178
DAFTAR PUSTAKA	180

LAMPIRAN	182
1. Listing Program	182
2. Lembar Bimbingan	18213