

**PENERAPAN ALGORITMA ADABOOST PADA C50 UNTUK
PENINGKATAN KINERJA KLASIFIKASI PENYAKIT LIVER**

SKRIPSI



Oleh :

ANWAR SANUSI

NPM. 19081010068

PROGRAM STUDI INFORMATIKA

FAKULTAS ILMU KOMPUTER

UNIVERSITAS PEMBANGUNAN NASIONAL "VETERAN"

JAWA TIMUR

2023

**PENERAPAN ALGORITMA ADABOOST PADA C50 UNTUK
PENINGKATAN KINERJA KLASIFIKASI PENYAKIT LIVER**

SKRIPSI

Diajukan Untuk Memenuhi Sebagai Persyaratan Dalam Menempuh Gelar Sarjana
Komputer Program Studi Informatika



Oleh :

ANWAR SANUSI

NPM. 19081010068

**PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS PEMBANGUNAN NASIONAL "VETERAN"
JAWA TIMUR
2023**


LEMBAR PENGESAHAN SKRIPSI

Judul : PENERAPAN ALGORITMA ADABOOST PADA C50 UNTUK
PENINGKATAN KINERJA KLASIFIKASI PENYAKIT LIVER
Oleh : ANWAR SANUSI
NPM : 19081010068

Telah Diseminarkan Dalam Ujian Skripsi Pada :
Hari Senin, Tanggal 10 Juli 2023


Mengetahui

1. **Dosen Pembimbing**

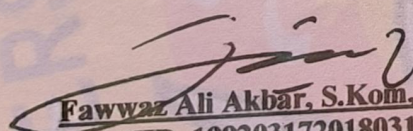

Chrystia Aji Putra, S.Kom, M.T
NIP: 19861008 2021211001

1.

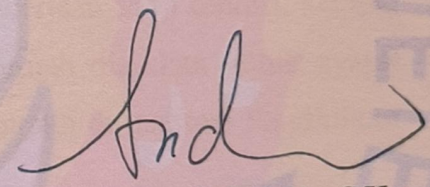
Dosen Penguji


Wahyu SJ Saputra, S.Kom, M.Kom
NIP: 198608252021211003

2.

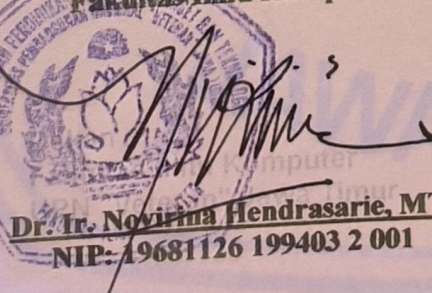

Fawwaz Ali Akbar, S.Kom, M.Kom
NIP: 199203172018031002

2.

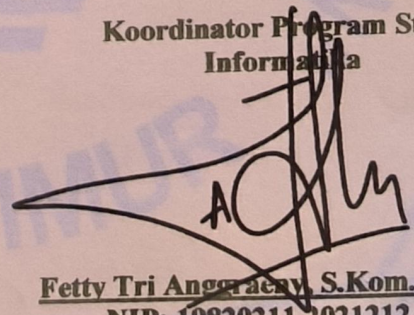

Andreas Nugroho Sihananto, S.Kom, M.Kom
NPT: 21119900412271

Menyetujui

Dekan
Fakultas Ilmu Komputer


Dr. Ir. Novirna Hendrasarie, MT.
NIP: 19681126 199403 2 001

Koordinator Program Studi
Informatika


Fetty Tri Anggraeny, S.Kom, M.Kom.
NIP: 19820211 2021212 005

SURAT PERNYATAAN ANTI PLAGIAT

Saya mahasiswa program studi Informatika UPN “Veteran” Jawa Timur, yang bertandatangan dibawah ini:

Nama : Anwar Sanusi

NPM : 19081010068

Dengan ini menyatakan bahwa Judul Skripsi/Tugas Akhir yang Saya ajukan dan akan dikerjakan, yang berjudul:

“PENERAPAN ALGORITMA ADABOOST PADA C50 UNTUK PENINGKATAN KINERJA KLASIFIKASI PENYAKIT LIVER”

Bukan merupakan plagiat dari Skripsi/Tugas Akhir/Penelitian orang lain juga bukan merupakan kode program yang saya beli dari pihak lain. Saya juga menyatakan bahwa Skripsi/Tugas Akhir ini adalah pekerjaan Saya sendiri, kecuali yang dinyatakan dalam Daftar Pustaka dan tidak pernah diajukan untuk syarat memperoleh gelar di UPN “Veteran” Jawa Timur maupun di institusi Pendidikan lain.

Pernyataan ini saya buat dengan sebenar-benarnya, jika ternyata di kemudian hari pernyataan ini terbukti tidak benar, maka Saya siap menerima segala konsekuensinya.

Surabaya, 10 Juli 2023

Hormat Saya,



ANWAR SANUSI

NPM. 19081010068

PENERAPAN ALGORITMA ADABOOST PADA C50 UNTUK PENINGKATAN KINERJA KLASIFIKASI PENYAKIT LIVER

Nama Mahasiswa : Anwar Sanusi

NPM : 19081010068

Program Studi : Informatika

Dosen Pembimbing : Chrystia Aji Putra, S.Kom, M.T

Fawwaz Ali Akbar, S.Kom, M.Kom

ABSTRAK

Hati merupakan salah satu organ vital yang penting bagi manusia. Hati memegang peran penting dalam metabolisme tubuh seperti proses pengaturan siklus hormone, menangkal dan menetralkan racun. Pengaturan struktur komposisi darah dan sebagainya. Penyakit hati merupakan penyakit umum yang sering terjadi di dunia. Penyakit ini seringkali terjadi tanpa didahului dengan gejala-gejala tertentu, istilah ini sering disebut dengan asimtomatik. Oleh karena itu penyakit hati kerap disebut (*silent killer*). Oleh karena itu proses diagnosis dan penyembuhan penyakit liver harus dilakukan dengan cepat dan tepat. Teknologi data mining bisa bermanfaat untuk mendeteksi penyakit liver dari hasil diagnose laboratorium dengan sangat cepat. Salah satu algoritma klasifikasi yang tepat bisa digunakan yaitu algoritma C50. Namun algoritma C50 bisa mengalami *overfitting* pada data medis yang kompleks sehingga perlu dilakukan proses boosting menggunakan adaboost. Algoritma adaboost bisa membuat algoritma C50 menjadi lebih rentan terhadap *overfitting*. Kelebihan lain dari algoritma adaboost yakni bisa menangani *dataset* yang tidak seimbang pada target labelnya. Setelah dilakukan penelitian didapatkan hasil bahwa algoritma C50 mampu melakukan klasifikasi dengan hasil akurasi sebesar 71,72% kemudian algoritma C50 bisa dimaksimalkan akurasi oleh adaboost menjadi 76,55%.

Kata kunci: Data Mining, Algoritma C50, Algoritma Adaboost, Klasifikasi Penyakit Liver

KATA PENGANTAR

Puji dan syukur penulis panjatkan kepada Allah S.W.T yang telah memberikan kesehatan jasmani dan rohani serta memberikan petunjuk dan kekuatan sehingga penulis berhasil menyelesaikan skripsi yang berjudul “PENERAPAN ALGORITMA ADABOOST PADA C50 UNTUK PENINGKATAN KINERJA KLASIFIKASI PENYAKIT LIVER”.

Penyusunan laporan ini adalah untuk menyelesaikan tugas akhir mata kuliah Skripsi serta memenuhi persyaratan untuk meraih gelar Sarjana Komputer di Program Studi Informatika Fakultas Ilmu Komputer Universitas Pembangunan Nasional “Veteran” Jawa Timur. Penulis ingin mengucapkan terima kasih yang sebesar-besarnya kepada semua pihak yang telah memberikan bantuan dan bimbingan baik secara moral maupun spiritual dalam proses penulisan skripsi ini.

Penulis sebagai seorang manusia yang tidak luput dari kekurangan, menyadari bahwa dalam penyusunan skripsi ini masih terdapat kekurangan dan kesalahan dikarenakan keterbatasan pengetahuan dan kemampuan Penulis. Oleh karena itu, Penulis memohon maaf dan siap menerima kritik serta saran yang membangun demi perbaikan dan menyempurnakan laporan skripsi ini. Penulis berharap semoga skripsi ini dapat memberikan manfaat bagi pembaca yang membacanya.

Surabaya, 10 Juli 2023

Penulis,

Anwar Sanusi

UCAPAN TERIMA KASIH

Segala puji syukur atas kehadiran Allah S.W.T atas berkah dan karunia-Nya yang telah membimbing penulis dalam menyelesaikan penulisan skripsi ini. Penulis juga merasa sangat beruntung karena telah menerima bantuan, kritik dan saran serta dukungan dari berbagai pihak baik secara langsung maupun tidak langsung. Oleh karena itu, pada kesempatan ini, penulis ingin menyampaikan ucapan terima kasih yang tulus kepada:

1. Allah S.W.T atas berkah dan karunia-Nya yang selalu mengiringi penulis dalam menyelesaikan penelitian dan penyusunan tugas akhir ini hingga selesai.
2. Prof. Dr. Ir. Akhmad Fauzi, MMT selaku Rektor Universitas Pembangunan Nasional Veteran Jawa Timur.
3. Ibu Dr. Ir. Novirina Hendrasarie, MT. selaku Dekan Fakultas Ilmu Komputer Universitas Pembangunan Nasional “Veteran” Jawa Timur.
4. Ibu Fetty Tri Anggraeny, S.Kom. M.Kom. selaku Ketua Jurusan Informatika Universitas Pembangunan Nasional “Veteran” Jawa Timur dan Dosen Wali .
5. Bapak Chrystia Aji Putra, S.Kom, M.T selaku Dosen Pembimbing Pertama dan Fawwaz Ali Akbar, S.Kom, M.Kom selaku Dosen Pembimbing Kedua yang telah meluangkan waktunya untuk memberikan bimbingan dan arahan kepada penulis selama proses penyusunan skripsi.
6. Kedua orang tua tercinta penulis, yang dengan penuh kasih sayang telah memberikan dukungan dan bantuan dalam penyusunan skripsi ini. Penulis menyadari bahwa tanpa perjuangan dan usaha keras dari kedua orang tua, penulis tidak akan mampu meraih mimpi dan cita-cita untuk menyelesaikan pendidikan hingga tingkat sarjana. Oleh karena itu, penulis sungguh-sungguh mengucapkan terima kasih atas segala bentuk dukungan, baik itu moral, material, maupun spiritual yang telah diberikan kepada penulis.
7. Kakak tercinta penulis, yang selalu yang memberikan semangat dan dukungan moral yang luar biasa selama penulis menyelesaikan skripsi ini.

8. Kepada NPM 19012010123 yang telah menemani, memberikan semangat, motivasi, perhatian, pencerahan saat mengalami kendala, hiburan saat jenuh dan bosan sehingga penulis dapat menyelesaikan tugas akhir ini.
9. Teman-teman tercinta penulis, Aditya Rizky Ardhana, Alfina Damayanti, Almira Auora Prameswaty, Alya Safira Fatikah, Danil Hendra Suryawan, Dimas Seno Herlambang, Eka Rizqi, Hamdan Yuwafi Mastu Wijaya, Maurisa Arimbi Putri, Mochammad Arya Salsabila, Nadhif Mahardika Awandi, Puteri Aulia Fahlia, Wariyanti Nugroho Putri yang telah memberikan semangat, wawasan, motivasi, perhatian, pencerahan saat mengalami kendala, hiburan saat jenuh dan bosan sehingga penulis dapat menyelesaikan penyusunan skripsi ini serta berbagi cerita suka dan duka selama masa perkuliahan.
10. Seluruh teman-teman dari program studi Informatika Angkatan tahun 2019 lainnya yang tentunya tidak dapat penulis sebutkan satu persatu yang telah membantu dan menjadi bagian dari kisah dan cerita di perkuliahan.

DAFTAR ISI

| | |
|-------------------------------------|------|
| LEMBAR PENGESAHAN SKRIPSI..... | .ii |
| SURAT PERNYATAAN ANTI PLAGIAT | iii |
| ABSTRAK | iv |
| KATA PENGANTAR..... | v |
| UCAPAN TERIMA KASIH..... | vi |
| DAFTAR ISI..... | viii |
| DAFTAR GAMBAR | x |
| DAFTAR TABEL | xi |
| BAB I PENDAHULUAN | 1 |
| 1.1 Latar Belakang | 1 |
| 1.2 Rumusan Masalah | 3 |
| 1.3 Batasan Masalah..... | 3 |
| 1.4 Tujuan Penelitian | 4 |
| 1.5 Manfaat Penelitian | 4 |
| BAB II TINJAUAN PUSTAKA..... | 5 |
| 2.1 Penelitian Terdahulu | 5 |
| 2.2 <i>Data Mining</i> | 6 |
| 2.3 <i>Decision Tree</i> | 7 |
| 2.4 <i>Binary Splitting</i> | 7 |
| 2.5 Algoritma C5.0..... | 8 |
| 2.6 Adaboost | 10 |
| 2.7 <i>Confusion Matrix</i> | 11 |
| BAB III METODOLOGI PENELITIAN..... | 13 |

| | | |
|-----------------------------------|--|----|
| 3.1 | Langkah-langkah Penelitian..... | 13 |
| 3.2 | Pengumpulan Data | 13 |
| 3.3 | <i>Preprocessing Data</i> | 15 |
| 3.4 | Pembagian Data | 18 |
| 3.5 | Algoritma C5.0..... | 18 |
| 3.6 | Adaboost | 31 |
| 3.7 | Uji Coba Sistem | 36 |
| BAB IV HASIL DAN PEMBAHASAN | | 38 |
| 4.1 | Inisialisasi Data | 38 |
| 4.2 | <i>Preprocessing data</i> | 39 |
| 4.3 | Pembagian Data | 41 |
| 4.4 | Klasifikasi Data..... | 43 |
| 4.5 | Hasil Uji Coba Sistem..... | 59 |
| 4.5.1. | Hasil Uji Coba Perbandingan Menggunakan Data Latih 90% dan Data Uji 10% Sebelum Dan Sesudah Boosting. | 59 |
| 4.5.2. | Hasil Uji Coba Perbandingan Menggunakan Data Latih 80% dan Data Uji 20% Sebelum Dan Sesudah Boosting. | 60 |
| 4.5.3. | Hasil Uji Coba Perbandingan Menggunakan Data Latih 75% dan Data Uji 25% Sebelum Dan Sesudah Boosting. | 61 |
| 4.5.4. | Perbandingan Ketiga Skenario Uji Coba..... | 63 |
| BAB V PENUTUP | | 64 |
| 5.1 | Kesimpulan..... | 64 |
| 5.2 | Saran..... | 64 |
| DAFTAR PUSTAKA | | 66 |
| BIODATA PENULIS | | 68 |

DAFTAR GAMBAR

| | |
|---|----|
| Gambar 3. 1 Diagram Alur Pengerjaan Penelitian..... | 13 |
| Gambar 3. 2 Scatter Plot Direct Bilirubin dan Total Bilirubin | 17 |
| Gambar 3. 3 Scatter Plot Atribut Albumin dan Total Protiens | 17 |
| Gambar 3. 4 Alur Implementasi C50 | 18 |
| Gambar 3. 5 Hasil Pembagian Node | 30 |
| Gambar 3. 6 Alur Proses Adaboost..... | 31 |
| Gambar 4. 1 Tabel Dataset..... | 39 |
| Gambar 4. 2 Tabel Setelah Praproses Data..... | 41 |
| Gambar 4. 3 Tabel Data Testing | 42 |
| Gambar 4. 4 Tabel Dataset Dengan Kolom Weight | 45 |
| Gambar 4. 5 Tampilan Efek Visual Loading | 45 |
| Gambar 4. 6 Grafik perbandingan akurasi ketiga skenario..... | 63 |

DAFTAR TABEL

| | |
|---|----|
| Tabel 2. 1 Confusion Matrix | 12 |
| Tabel 3. 1 Daftar Atribut Pada Dataset | 14 |
| Tabel 3. 2 Dataset Sebelum Dilakukan Preprocessing | 14 |
| Tabel 3. 3 Dataset Setelah Dilakukan Data <i>Selection</i> | 16 |
| Tabel 3. 4 Dataset Setelah Diurutkan Berdasarkan Total Bilirubin..... | 19 |
| Tabel 3. 5 Daftar Kandidat Threshold Terbaik | 20 |
| Tabel 3. 6 Hasil Perhitungan Entropy Tiap Kandidat Treshold..... | 21 |
| Tabel 3. 7 Hasil Perhitungan Gain Ratio Tiap Kandidat Treshold | 23 |
| Tabel 3. 8 Daftar Threshold Terbaik Dari Tiap Atribut..... | 24 |
| Tabel 3. 9 Jumlah Data Setelah Dilakukan Pembagian Treshold | 25 |
| Tabel 3. 10 Hasil Perhitungan Entropy Tiap Atribut | 26 |
| Tabel 3. 11 Data Latih Setelah Ditambahkan Bobot..... | 32 |
| Tabel 3. 12 Data Latih Setelah Dihitung Nilai Errornya | 33 |
| Tabel 3. 13 Dataset Setelah Dilakukan Update Bobot..... | 34 |
| Tabel 3. 14 Dataset Setelah Dilakukan Normalisasi | 35 |
| Tabel 3. 15 Tabel Uji Coba Sistem | 36 |
| Tabel 4. 1 Hasil Pengujian Sistem Pada Skenario Pertama | 60 |
| Tabel 4. 2 Hasil Pengujian Sistem Pada Skenario Kedua..... | 61 |
| Tabel 4. 3 Hasil Pengujian Sistem Pada Skenario Ketiga..... | 62 |