

**PERBANDINGAN TEKNIK MACHINE LEARNING UNTUK
DIAGNOSIS DAN PREDIKSI PENYAKIT JANTUNG**

SKRIPSI

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana Program
Studi Informatika



disusun oleh

MUHAMMAD ILHAM AKBAR

20.11.3691

Kepada

FAKULTAS ILMU KOMPUTER

UNIVERSITAS AMIKOM YOGYAKARTA

YOGYAKARTA

2024

**PERBANDINGAN TEKNIK MACHINE LEARNING UNTUK
DIAGNOSIS DAN PREDIKSI PENYAKIT JANTUNG**

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana Program Studi
Informatika



disusun oleh

MUHAMMAD ILHAM AKBAR

20.11.3691

Kepada

FAKULTAS ILMU KOMPUTER

UNIVERSITAS AMIKOM YOGYAKARTA

YOGYAKARTA

2024

HALAMAN PERSETUJUAN

SKRIPSI

PERBANDINGAN TEKNIK MACHINE LEARNING UNTUK DIAGNOSIS
DAN PREDIKSI PENYAKIT JANTUNG

yang disusun dan diajukan oleh

MUHAMMAD ILHAM AKBAR

20.11.3691

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 19 Maret 2024

Dosen Pembimbing,



FERIAN FAUZI ABDULLOH, M.Kom

NIK. 190302276

HALAMAN PENGESAHAN

SKRIPSI

PERBANDINGAN TEKNIK MACHINE LEARNING UNTUK DIAGNOSIS DAN
PREDIKSI PENYAKIT JANTUNG

yang disusun dan diajukan oleh

MUHAMMAD ILHAM AKBAR

20.11.3691

Telah dipertahankan di depan Dewan Penguji
pada tanggal 19 Maret 2024

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

FERIAN FAUZI ABDULLOH, M.Kom
NIK. 190302276

FIRMAN ASHARUDIN S.Kom, M.Kom
NIK. 190302315

RIFDA FATICHA ALFA AZIZA, M.Kom
NIK. 190302392

Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 19 Maret 2024

DEKAN FAKULTAS ILMU KOMPUTER



Hanif Al Fatta, S.Kom., M.Kom., Ph.D.
NIK. 190302096

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Muhammad Ilham Akbar
NIM : 20.11.3691

Menyatakan bahwa Skripsi dengan judul berikut:

PERBANDINGAN TEKNIK MACHINE LEARNING UNTUK DIAGNOSIS DAN PREDIKSI PENYAKIT JANTUNG

Dosen Pembimbing : Ferian Fauzi Abdulloh, M.Kom

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 10 Maret 2024

Yang Menyatakan,



Muhammad Ilham Akbar

HALAMAN PERSEMBAHAN

- Ucapan terima kasih kepada Allah SWT yang atas karunia yang diberikan dan tidak lupa penulis selalu bersyukur atas pemberian yang di telah diberikan olehNya.
- Ucapan terima kasih kepada Alm. Bp Dwi Setyo Nugroho dan Ibu Elly Susanti, sebagai madrasah pertama di dunia ini dan selalu mendidik penulis sampai bisa dalam proses ini, dan juga kepada keluarga yang selalu mensupport, mendoakan, memberi semangat segingga kepenulisan skripsi dapat diselesaikan tepat pada waktunya.
- Ucapan terima kasih kepada Bp Afrig Aminuddin, Bp Ferian Fauzi Abdulloh, Bp Majid Rahardi dan Mas Bima selaku Pembina dalam membimbing, dan memberikan arahan, sehingga Penulis dapat menyelesaikan penulisan naskah ini tepat waktu.
- Ucapan terima kasih kepada Mbak Faiz Tri Meliana yang selalu memberikan saran, masukan, suport yang baik dalam proses penulis mengerjakan Skripsi.
- Ucapan terima kasih kepada semua pihak yang membantu, men-suport, memberikan saran, masukan, semangat, emosional yang baik sehingga penulis dapat menyelesaikan naskah Skripsi yang telah dibuat.

KATA PENGANTAR

Puji syukur kehadiran Allah SWT yang telah memberikan rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan skripsi dengan judul “PERBANDINGAN TEKNIK MACHINE LEARNING UNTUK DIAGNOSIS DAN PREDIKSI PENYAKIT JANTUNG”. Skripsi ini disusun sebagai salah satu syarat untuk memperoleh gelar Sarjana Komputer di Fakultas Ilmu Komputer Universitas Amikom Yogyakarta.

Penelitian ini bertujuan untuk membandingkan Mengaplikasikan bidang ilmu *machine learning* agar bermanfaat bagi ilmu pengetahuan dan pengembangan penelitian. Penelitian ini diharapkan dapat memberikan manfaat bagi tenaga medis supaya lebih cepat dalam pengambilan keputusan dan mendiagnosis lebih dini penyakit jantung.

Penulis menyadari bahwa penulisan skripsi ini tidak lepas dari bantuan dan dukungan dari berbagai pihak. Penulis juga menyadari bahwa skripsi ini masih jauh dari sempurna dan memiliki banyak kekurangan. Oleh karena itu, penulis mengharapkan kritik dan saran yang membangun dari pembaca untuk perbaikan di masa depan. Semoga skripsi ini dapat bermanfaat bagi penulis dan khususnya dan bagi pembaca pada umumnya

Yogyakarta, 04 Maret 2024

Penulis

DAFTAR ISI

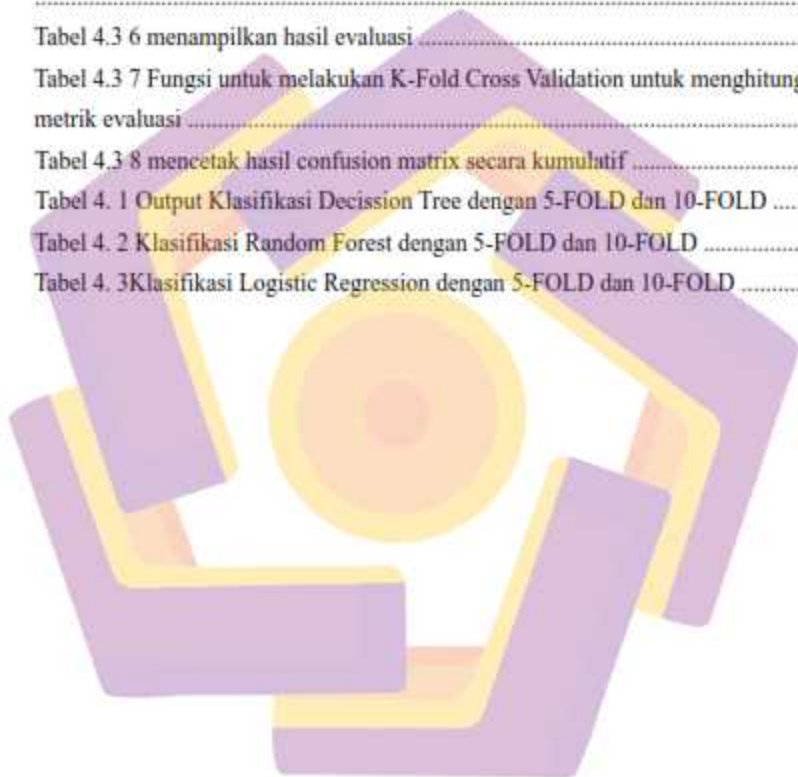
HALAMAN PERSETUJUAN	III
HALAMAN PENGESAHAN	IV
HALAMAN PERNYATAAN KEASLIAN SKRIPSI	V
HALAMAN PERSEMBAHAN	VI
KATA PENGANTAR	VII
DAFTAR ISI	VIII
DAFTAR TABEL	X
DAFTAR GAMBAR	XI
INTISARI	XII
<i>ABSTRACT</i> XIII	
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Batasan Masalah	3
1.4 Tujuan Penelitian	3
1.5 Manfaat Penelitian	3
1.6 Sistematika Penulisan	4
BAB II TINJAUAN PUSTAKA	5
2.1 Studi Literatur	5
2.2 Machine Learning	11
2.3 Algoritma Decision Tree	11
2.4 Algoritma Random Forest	11
2.5 Algoritma Logistic Regression	12
BAB III METODE PENELITIAN	13
3.1 Alur Penelitian	13

3.2	Exploratory Data Analysis (EDA)	14
3.3	Data Preprocessing	16
3.4	Analysis (Train The Model)	18
3.5	Evaluasi Kinerja	21
BAB IV HASIL DAN PEMBAHASAN		22
4.1	Klasifikasi Algoritma Decision Tree	22
4.2	Klasifikasi Algoritma Random Forest	29
4.3	Klasifikasi Logistic Regression	36
4.4	Kurva ROC	44
BAB V PENUTUP		45
5.1	Kesimpulan	45
5.2	Saran	45
REFERENSI		46



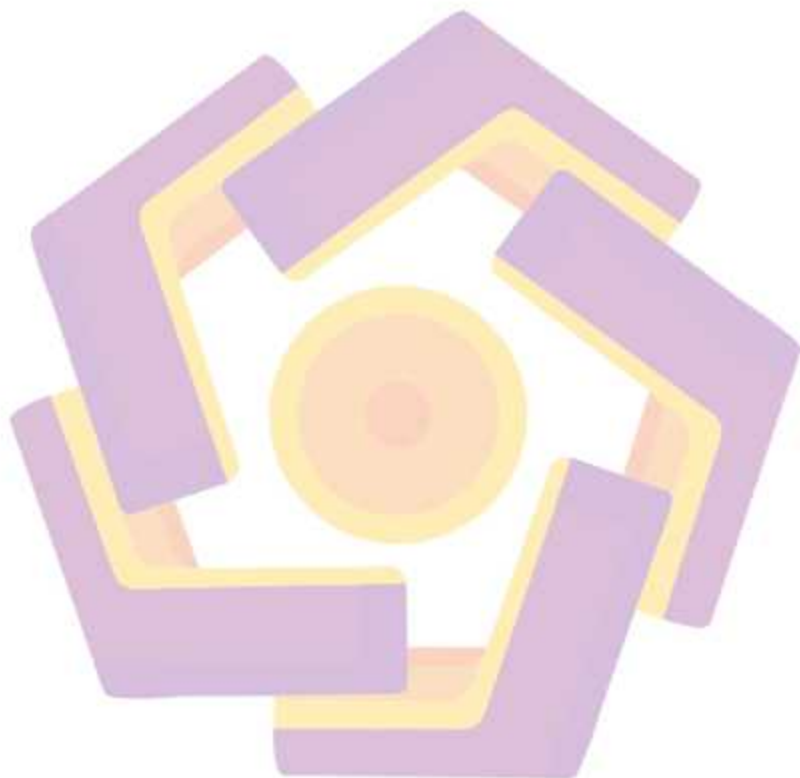
DAFTAR TABEL

Tabel 4.3 1 Memindahkan atribut masukan x dan y	36
Tabel 4.3 2 memecah data latih 70% dan data uji 30%	36
Tabel 4.3 3 Inisiasi model menggunakan algoritma Logistic Regression	37
Tabel 4.3 4 Fungsi untuk melakukan Cross Validation dan Menghitung Metrix Evaluasi.	38
Tabel 4.3 5 Evaluasi model menggunakan K-Fold Cross validation dengan 5 lipatan	38
.....	38
Tabel 4.3 6 menampilkan hasil evaluasi	39
Tabel 4.3 7 Fungsi untuk melakukan K-Fold Cross Validation untuk menghitung metrik evaluasi	40
Tabel 4.3 8 mencetak hasil confusion matrix secara kumulatif	41
Tabel 4. 1 Output Klasifikasi Decission Tree dengan 5-FOLD dan 10-FOLD	28
Tabel 4. 2 Klasifikasi Random Forest dengan 5-FOLD dan 10-FOLD	35
Tabel 4. 3Klasifikasi Logistic Regression dengan 5-FOLD dan 10-FOLD	41



DAFTAR GAMBAR

Gambar 3. 1 Alur Penelitian	13
Gambar 4. 1 Confussion Matrix of 10-FOLD Decision Tree	28
Gambar 4. 2 Confussion Matrix 10-FOLD of Random Forest	35
Gambar 4. 3 Confussion Matrix of 10-FOLD Logistic Regression	43
Gambar 4. 4 Perbandingan Algoritma berdasarkan kurva ROC	44



INTISARI

Penelitian ini membandingkan akurasi prediksi penyakit jantung dengan menerapkan beberapa algoritma seperti Regresi Logistik, Decision Tree, dan Random Forest. Penelitian ini menggunakan 319.795 data yang berisi faktor-faktor penyebab penyakit jantung sebagai dataset, dengan menerapkan metode k-fold cross-validation untuk memilih model terbaik dari ketiga algoritma tersebut berdasarkan nilai pada metrik evaluasi seperti *F1-score*, akurasi, presisi, dan *recall*. Hasil penelitian menunjukkan bahwa Decision Tree dengan kernel linear menghasilkan akurasi tertinggi sebesar 74,59% pada 10-fold. Sementara itu, Random Forest mencapai akurasi tertinggi sebesar 74,75% pada 10-fold, sedangkan Regresi Logistik mencapai akurasi tertinggi sebesar 71,97%. Temuan dari penelitian ini menggambarkan bahwa penggunaan algoritma kecerdasan buatan dapat bermanfaat dalam meningkatkan akurasi prediksi penyakit jantung, hal ini dapat membantu tenaga medis dalam mendiagnosa penyakit dengan lebih cepat, akurat, dan tepat.

Kata kunci: penyakit jantung, machine learning, decision tree, logistic regression, random forest.



ABSTRACT

This study compares heart disease prediction accuracy by applying artificial algorithms, such as Logistic Regression, Decision Tree, and Random Forest. This research uses 319,795 data on factors causing heart disease as a dataset, coupled with the k-fold cross-validation method, to select the best model of the three algorithms based on F1 score, accuracy, precision, and recall. The results showed that Decision Tree with a linear kernel produced the highest accuracy of 74.59% at 10fold. Meanwhile, Random Forest achieved the highest accuracy of 74.75% at 10fold, while Logistic Regression achieved the highest accuracy of 71.97%. The findings of this study illustrate that the use of artificial intelligence algorithms can be beneficial in improving the accuracy of heart disease classification, which will assist medical personnel in diagnosing diseases more quickly, accurately, and precisely.

Keyword: heart disease, machine learning, decision tree, logistic regression, random forest

