

**IMPLEMENTASI ALGORITMA RANDOM FOREST UNTUK
PREDIKSI PENYAKIT JANTUNG MENGGUNAKAN
DATASET HEARTS DISEASE UCI**

SKRIPSI

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



disusun oleh

PARNO HUSI

21.11.4035

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2025

**IMPELEMENTASI ALGORITMA RANDOM FOREST UNTUK
PREDIKSI PENYAKIT JANTUNG MENGGUNAKAN
DATASET HEARTS DISEASE UCI**

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



disusun oleh

PARNO HUSI

21.11.4035

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2025

HALAMAN PERSETUJUAN

SKRIPSI

**IMPELEMENTASI ALGORITMA RANDOM FOREST UNTUK
PREDIKSI PENYAKIT JANTUNG MENGGUNAKAN
DATASET HEARTS DISEASE UCI**

yang disusun dan diajukan oleh

Parno Husi

21.11.4035

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 22 Mei 2025

Dosen Pembimbing,



Lukman, M. Kom
NIK. 190302151

HALAMAN PENGESAHAN
SKRIPSI
IMPELEMENTASI ALGORITMA RANDOM FOREST UNTUK
PREDIKSI PENYAKIT JANTUNG MENGGUNAKAN
DATASET HEARTS DISEASE UCI

yang disusun dan diajukan oleh

Parno Husi

21.11.4035

Telah dipertahankan di depan Dewan Penguji
pada tanggal 22 Mei 2025

Susunan Dewan Penguji

Nama Penguji

Heri Sismoro, S.Kom., M.Kom
NIK. 190302057

Novi Prisma Yunita, M.Kom
NIK. 190302526

Lukman, S.Kom., M.Kom
NIK. 190302151

Tanda Tangan



Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 22 Mei 2025

DEKAN FAKULTAS ILMU KOMPUTER



Prof. Dr. Kusriani, M.Kom.

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Parno Husi
NIM : 21.11.4035

Menyatakan bahwa Skripsi dengan judul berikut:

IMPELEMENTASI ALGORITMA RANDOM FOREST UNTUK PREDIKSI PENYAKIT JANTUNG MENGGUNAKAN DATASET HEARTS DISEASE UCI

Dosen Pembimbing : Lukman, 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, 22 Mei 2025

Yang Menyatakan,



1CAMX303596548

Parno Husi

HALAMAN PERSEMBAHAN

Segala puji dan syukur saya panjatkan kepada Tuhan Yang Maha Esa, Allah SWT, atas segala nikmat, rahmat, dan hidayah-Nya yang senantiasa menyertai setiap langkah dalam hidup saya. Berkat pertolongan-Nya, saya diberikan kekuatan, kesabaran, kesehatan, serta kelancaran dalam menyelesaikan salah satu proses penting dalam kehidupan akademik ini, yakni penulisan dan penyusunan skripsi. Tanpa kasih sayang dan izin-Nya, saya tidak akan mampu melewati setiap tantangan, hambatan, dan keraguan yang muncul selama proses ini. Dengan segala kerendahan hati dan rasa syukur yang mendalam, saya mempersembahkan karya ini kepada:

1. Allah SWT, Tuhan yang Maha Kuasa, yang telah memberikan saya napas kehidupan, petunjuk dalam kegelapan, kekuatan saat lelah, serta keyakinan ketika ragu menghampiri. Tiada daya dan upaya selain atas izin-Nya. Segala pencapaian ini adalah buah dari kasih-Nya yang tak ternilai.
2. Orang tua tercinta, Bapak La Husi dan Ibu Erna Kurniati, sosok luar biasa yang tidak pernah lelah mendoakan, membimbing, dan mendukung setiap langkah saya. Terima kasih atas kasih sayang yang tulus, kerja keras, pengorbanan, dan segala perjuangan yang tidak pernah saya mampu balas sepenuhnya. Kalian adalah alasan utama saya untuk terus melangkah dan menjadi pribadi yang lebih baik setiap harinya.
3. Bapak Lukman, M.Kom, selaku dosen pembimbing yang telah dengan sabar dan penuh dedikasi membimbing saya selama proses penyusunan skripsi ini. Terima kasih atas segala ilmu, arahan, kritik yang membangun, serta waktu yang telah Bapak luangkan untuk membimbing saya. Kehadiran dan bimbingan Bapak telah memberi pengaruh besar terhadap kelancaran dan hasil dari karya ini.
4. Teman-teman terdekat saya, Akbar, Yuza, Eker, Ando, Ndaru, Reza, Fayza dan Wildan, yang telah menjadi tempat berbagi cerita, tawa, dan semangat selama proses perkuliahan hingga penyusunan skripsi ini. Terima kasih atas

kebersamaan, bantuan, dan dukungan yang begitu berarti. Persahabatan kalian adalah anugerah luar biasa yang menguatkan saya dalam berbagai situasi.

5. Dan tak lupa, untuk diri saya sendiri, Parno Husi, terima kasih telah bertahan sejauh ini. Terima kasih atas setiap usaha, air mata, malam tanpa tidur, dan semangat yang tidak pernah padam walaupun lelah datang silih berganti. Skripsi ini adalah bukti nyata dari ketekunan, dedikasi, dan kerja kerasmu. Teruslah percaya pada diri sendiri dan jangan pernah berhenti untuk bermimpi dan berjuang.



KATA PENGANTAR

Puji syukur ke hadirat Tuhan Yang Maha Esa atas rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan skripsi ini dengan judul **“IMPLEMENTASI ALGORITMA RANDOM FOREST UNTUK PREDIKSI PENYAKIT JANTUNG MENGGUNAKAN DATASET HEART DISEASE UCI”**. Skripsi ini disusun sebagai salah satu syarat untuk memperoleh gelar Sarjana Ilmu Komputer di Universitas AMIKOM Yogyakarta.

1. Bapak Prof. Dr. M. Suyanto, M.M., selaku Rektor Universitas AMIKOM Yogyakarta yang telah menyediakan berbagai fasilitas dan peluang yang memungkinkan penulis untuk menempuh pendidikan dengan baik di universitas ini.
2. Ibu Prof. Dr. Kusri, S.Kom., M.Kom., selaku Dekan Fakultas Ilmu Komputer Universitas AMIKOM Yogyakarta.
3. Ibu Eli Pujastuti, M.Kom., selaku Ketua Program Studi S1 Informatika Universitas AMIKOM Yogyakarta.
4. Bapak Lukman, M.Kom., selaku Dosen Pembimbing, yang telah dengan sabar meluangkan waktu, memberikan bimbingan, saran, kritik, serta motivasi yang sangat berarti selama proses penyusunan skripsi ini, sehingga dapat diselesaikan dengan baik.

Penulis menyadari bahwa skripsi ini masih jauh dari kata sempurna. Oleh karena itu, kritik dan saran yang membangun sangat diharapkan demi penyempurnaan karya ini di masa mendatang. Semoga skripsi ini dapat memberikan manfaat bagi pembaca dan dunia akademik.

Yogyakarta, 12 Mei 2025



Parno Husi

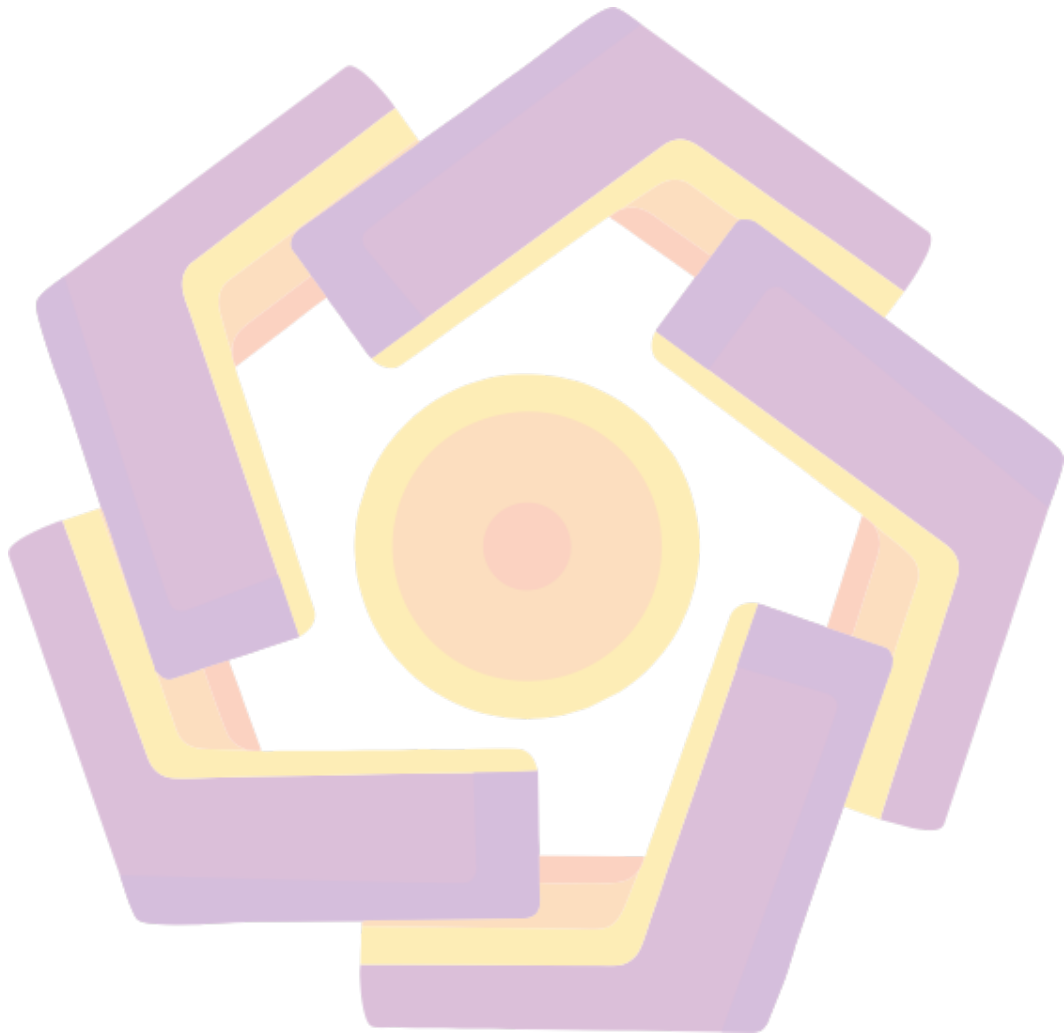
DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERSETUJUAN	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERNYATAAN KEASLIAN SKRIPSI	iv
HALAMAN PERSEMBAHAN	v
KATA PENGANTAR	vii
DAFTAR ISI	viii
DAFTAR TABEL	x
DAFTAR GAMBAR	xi
DAFTAR LAMPIRAN	xii
DAFTAR LAMBANG DAN SINGKATAN	xiii
DAFTAR ISTILAH	xiv
INTISARI	xvi
ABSTRACT	xvii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Batasan Masalah	2
1.4 Tujuan Penelitian	2
1.5 Manfaat Penelitian	3
1.6 Sistematika Penulisan	3
BAB II TINJAUAN PUSTAKA	5
2.1 Studi Literatur	5
2.2 Dasar Teori	15
2.2.1 Penyakit Jantung	15
2.2.2 Data Mining	15
2.2.3 Random Forest	15
2.2.4 Dataset Heart Disease UCI	16
2.2.5 SHAP	16
2.2.6 Confusion Matrix	16
BAB III METODE PENELITIAN	18

3.1	Objek Penelitian	18
3.2	Alur Penelitian.....	20
3.2.1	Pengumpulan Data	21
3.2.2	Pra-pemrosesan Data.....	22
3.2.3	Pembagian Data	24
BAB IV HASIL DAN PEMBAHASAN		27
4.1	Inisiasi Model <i>Random Forest</i>	27
4.2	Evaluasi Model.....	28
4.2.1	<i>Confusion Matrix</i>	29
4.3	Analisis Hasil	31
4.3.1	Analisis Interpretasi Model dengan SHAP	31
BAB V PENUTUP		33
5.1	Kesimpulan.....	33
5.2	Saran.....	33
REFERENSI.....		35
LAMPIRAN		38

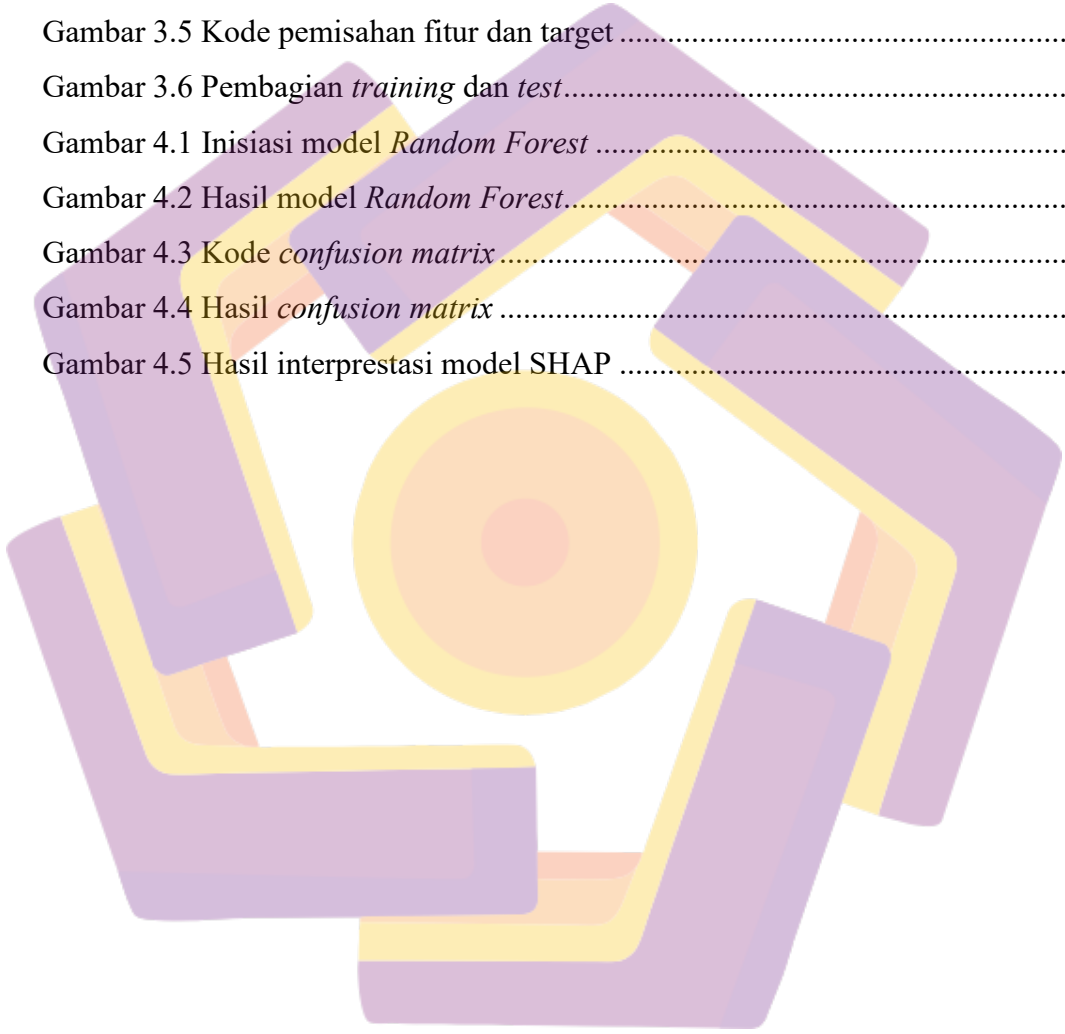
DAFTAR TABEL

Tabel 2.1 Keaslian Penelitian	9
Tabel 2.2 <i>Confusion Matrix</i>	17



DAFTAR GAMBAR

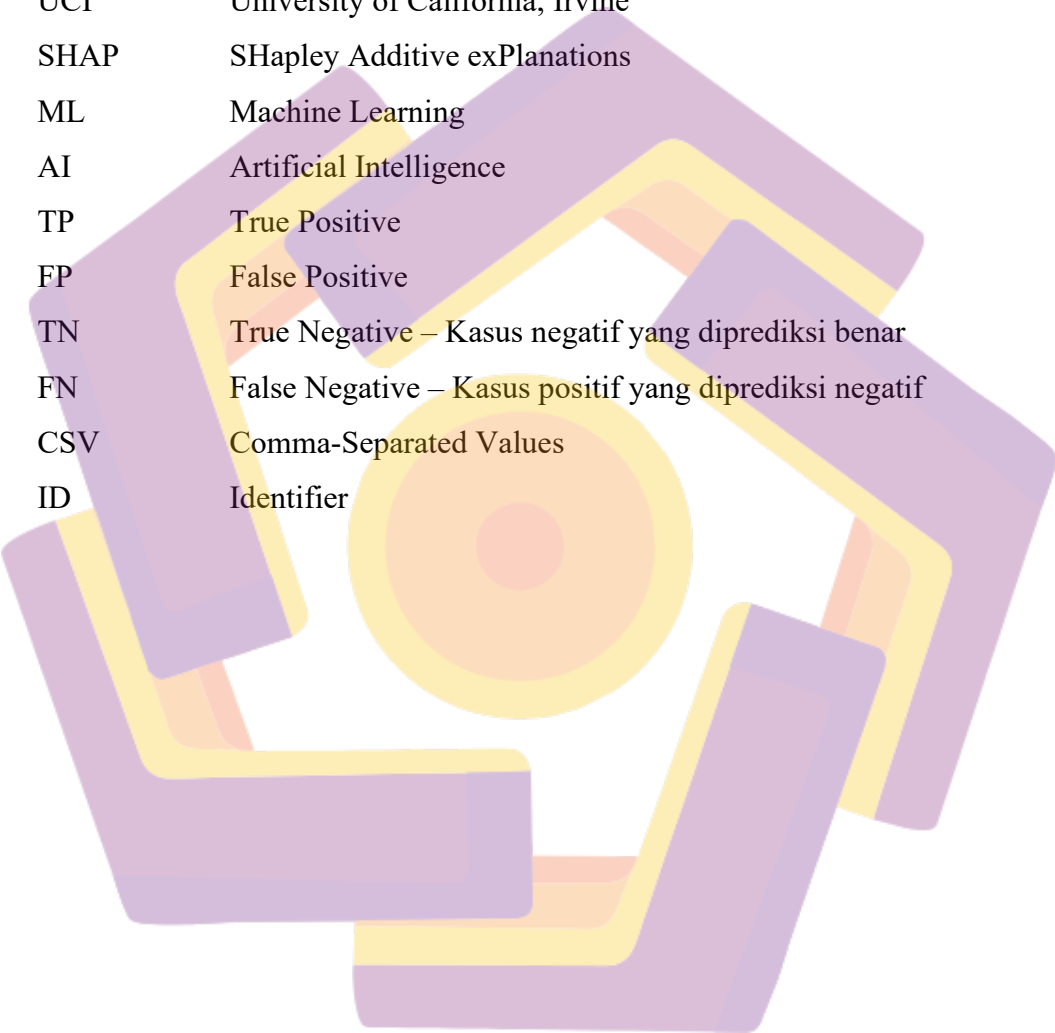
Gambar 3.1 Distribusi Jenis Kelamin	19
Gambar 3.2 Alur penelitian.....	21
Gambar 3.3 Kode pengecekan nilai yang hilang	22
Gambar 3.4 Hasil pengecekan nilai yang hilang.....	23
Gambar 3.5 Kode pemisahan fitur dan target	24
Gambar 3.6 Pembagian <i>training</i> dan <i>test</i>	25
Gambar 4.1 Inisiasi model <i>Random Forest</i>	27
Gambar 4.2 Hasil model <i>Random Forest</i>	27
Gambar 4.3 Kode <i>confusion matrix</i>	29
Gambar 4.4 Hasil <i>confusion matrix</i>	30
Gambar 4.5 Hasil interpretasi model SHAP	32



DAFTAR LAMPIRAN

Lampiran 1. Kode daftar <i>library</i>	38
Lampiran 2. Import dataset	38
Lampiran 3. Mengecek informasi dataset	38
Lampiran 4. Mengecek data yang hilang	38
Lampiran 5. Statistik deskriptif dataset	38
Lampiran 6. Pembagian fitur	39
Lampiran 7. Inisiasi model <i>Random Forest</i>	39
Lampiran 8. Laporan klasifikasi	39
Lampiran 9. <i>Confusion Matrix</i>	39
Lampiran 10. Interpretasi Model SHAP	39

DAFTAR LAMBANG DAN SINGKATAN



RF	Random Forest keputusan
SVM	Support Vector Machines
KNN	K-Nearest Neighbor
UCI	University of California, Irvine
SHAP	SHapley Additive exPlanations
ML	Machine Learning
AI	Artificial Intelligence
TP	True Positive
FP	False Positive
TN	True Negative – Kasus negatif yang diprediksi benar
FN	False Negative – Kasus positif yang diprediksi negatif
CSV	Comma-Separated Values
ID	Identifier

DAFTAR ISTILAH



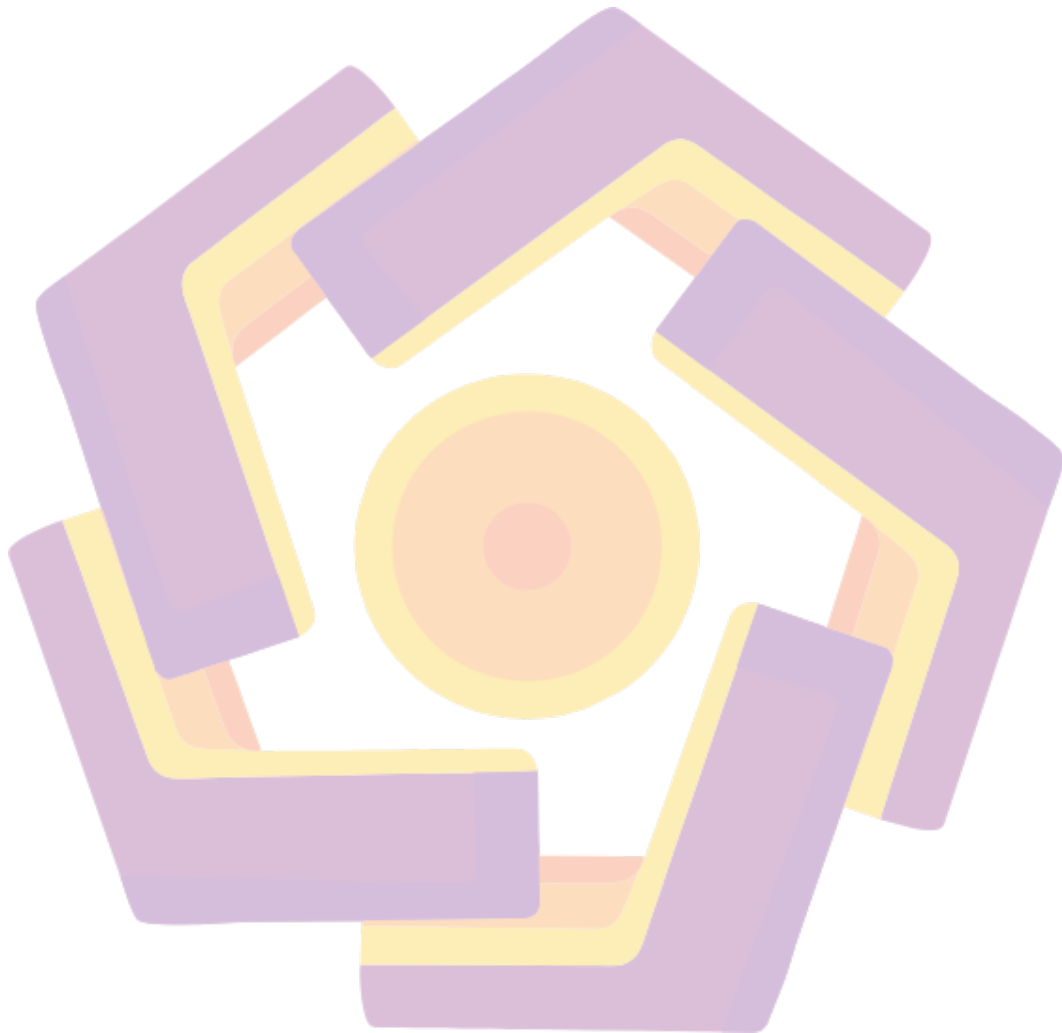
Dataset	Kumpulan data terstruktur yang digunakan sebagai sumber pelatihan dan pengujian model machine learning.
Random Forest	Algoritma ensemble learning yang membangun banyak pohon keputusan untuk meningkatkan akurasi prediksi.
Algoritma Klasifikasi	Prosedur sistematis dalam machine learning untuk mengelompokkan data ke dalam kelas tertentu.
Train-Test Split	Teknik pembagian data menjadi dua bagian: data pelatihan (train) dan data pengujian (test).
Overfitting	Kondisi di mana model terlalu sesuai dengan data latih dan tidak mampu menggeneralisasi data baru.
SHAP	Metode untuk menjelaskan kontribusi fitur terhadap prediksi model berbasis nilai Shapley dari teori permainan.
Confusion Matrix	Matriks evaluasi model klasifikasi berdasarkan nilai prediksi benar dan salah terhadap data aktual.
Akurasi	Rasio prediksi yang benar terhadap seluruh jumlah prediksi yang dilakukan.
Precision	Proporsi prediksi positif yang benar dari seluruh prediksi positif yang dibuat model.
Recall	Proporsi kasus positif yang berhasil dikenali dengan benar oleh model.
F1-Score	Rata-rata harmonis antara precision dan recall, digunakan untuk evaluasi menyeluruh.
Interpretabilitas Model	Kemampuan untuk menjelaskan atau memahami hasil keluaran dari suatu model machine learning.

UCI Heart Disease
Dataset

Dataset yang disediakan oleh University of California
Irvine yang memuat parameter klinis pasien terkait
penyakit jantung.

Preprocessing

Tahapan awal pengolahan data, seperti pengecekan
nilai kosong, seleksi fitur, dan transformasi data.



INTISARI

Penyakit jantung merupakan salah satu penyebab kematian tertinggi di dunia dengan angka mortalitas yang terus meningkat setiap tahunnya. Keterlambatan dalam mendeteksi penyakit jantung menjadi faktor kritis yang mempengaruhi tingkat keberhasilan penanganan pasien. Penelitian ini mengimplementasikan algoritma *Random Forest* untuk mengembangkan model prediksi penyakit jantung menggunakan dataset UCI Heart Disease yang terdiri dari 1026 records dengan 13 parameter medis. Metodologi penelitian meliputi tahap preprocessing data seperti pengecekan data kosong dan pemilihan fitur, pembagian data training-testing dengan rasio 80:20, implementasi algoritma *Random Forest* menggunakan bahasa pemrograman Python, serta evaluasi performa model menggunakan *metrics accuracy, precision, recall, dan F1-score*. Hasil penelitian menunjukkan bahwa model *Random Forest* mencapai akurasi 99% dalam memprediksi penyakit jantung. Model yang dikembangkan dapat dimanfaatkan oleh tenaga medis sebagai alat bantu screening awal pasien penyakit jantung, sehingga dapat mengoptimalkan proses diagnosa dan mempercepat penanganan pasien. Penelitian lebih lanjut dapat dilakukan dengan mengintegrasikan deep learning dan menambahkan parameter medis lainnya untuk meningkatkan akurasi prediksi.

Kata kunci: *Machine Learning, Random Forest, Prediksi Penyakit Jantung, UCI Dataset, Python*

ABSTRACT

Heart disease is one of the leading causes of death in the world with mortality rates that continue to increase every year. Delays in detecting heart disease are a critical factor that affects the success rate of patient treatment. This study implements the Random Forest algorithm to develop a heart disease prediction model using the UCI Heart Disease dataset consisting of 1026 records with 13 medical parameters. The research methodology includes data preprocessing stages such as handling missing values and feature selection, dividing training-testing data with a ratio of 80:20, implementing the Random Forest algorithm using the Python programming language, and evaluating model performance using metrics accuracy, precision, recall, and F1-score. The results showed that the Random Forest model achieved 99% accuracy in predicting heart disease. The developed model can be used by medical personnel as an early screening tool for heart disease patients, so that it can optimize the diagnostic process and accelerate patient treatment. Further research can be done by integrating deep learning and adding other medical parameters to improve prediction accuracy.

Keywords: *Machine Learning, Random Forest, Heart Disease Prediction, UCI Dataset, Python*