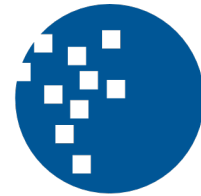


**IMPLEMENTASI ALGORITMA RANDOM FOREST DALAM
DETEKSI PENYAKIT JANTUNG**



UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

SKRIPSI

**NAYASHA CLARISA DWISUTRISNA
00000056883**

**PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA
TANGERANG
2025**

IMPLEMENTASI ALGORITMA RANDOM FOREST DALAM DETEKSI PENYAKIT JANTUNG



NAYASHA CLARISA DWISUTRISNA

00000056883

UMN

UNIVERSITAS

MULTIMEDIA

NUSANTARA

PROGRAM STUDI INFORMATIKA

FAKULTAS TEKNIK DAN INFORMATIKA

UNIVERSITAS MULTIMEDIA NUSANTARA

TANGERANG

2025

HALAMAN PERNYATAAN TIDAK PLAGIAT

Dengan ini saya,

Nama : Nayasha Clarisa Dwisutrisna

Nomor Induk Mahasiswa : 00000056883

Program Studi : Informatika

Skripsi dengan judul:

Implementasi Algoritma Random Forest dalam Deteksi Penyakit Jantung

merupakan hasil karya saya sendiri bukan plagiat dari laporan karya tulis ilmiah yang ditulis oleh orang lain, dan semua sumber, baik yang dikutip maupun dirujuk, telah saya nyatakan dengan benar serta dicantumkan di Daftar Pustaka.

Jika di kemudian hari terbukti ditemukan kecurangan/penyimpangan, baik dalam pelaksanaan maupun dalam penulisan laporan karya tulis ilmiah, saya bersedia menerima konsekuensi dinyatakan **TIDAK LULUS** untuk mata kuliah yang telah saya tempuh.

Tangerang, 3 Juli 2025



(Nayasha Clarisa Dwisutrisna)

UNIVERSITAS
MULTIMEDIA
NUSANTARA

HALAMAN PENGESAHAN

Skripsi dengan judul

IMPLEMENTASI ALGORITMA RANDOM FOREST DALAM DETEKSI PENYAKIT JANTUNG

oleh

Nama : Nayasha Clarisa Dwisutrisna
NIM : 00000056883
Program Studi : Informatika
Fakultas : Fakultas Teknik dan Informatika

Telah diujikan pada hari Senin, 14 Juli 2025

Pukul 08.00 s/s 10.00 dan dinyatakan

LULUS

Dengan susunan penguji sebagai berikut

Ketua Sidang



(Dennis Gunawan, S.Kom., M.Sc.)

NIDN: 0320059001

Penguji



(Aditiyawan, S.Komp., M.Si.)

NIDK : 8994550022

Pembimbing



(Dr. Maria Irmira Prastiyowati, S.Kom., M.T.)

NIDN: 0725057201

Ketua Program Studi Informatika,



(Arya Wicaksana, S.Kom., M.Eng.Sc., OCA)

NIDN: 0315109103

HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS

Yang bertanda tangan di bawah ini:

Nama : Nayasha Clarisa Dwisutrisna
NIM : 00000056883
Program Studi : Informatika
Jenjang : S1
Judul Karya Ilmiah : Implementasi Algoritma Random
Forest dalam Deteksi Penyakit
Jantung

Menyatakan dengan sesungguhnya bahwa saya bersedia (**pilih salah satu**):

- ☒ Saya bersedia memberikan izin sepenuhnya kepada Universitas Multimedia Nusantara untuk mempublikasikan hasil karya ilmiah saya ke dalam repositori Knowledge Center sehingga dapat diakses oleh Sivitas Akademika UMN/Publik. Saya menyatakan bahwa karya ilmiah yang saya buat tidak mengandung data yang bersifat konfidensial.
- ☐ Saya tidak bersedia mempublikasikan hasil karya ilmiah ini ke dalam repositori Knowledge Center, dikarenakan: dalam proses pengajuan publikasi ke jurnal/konferensi nasional/internasional (dibuktikan dengan *letter of acceptance*) **.
- ☐ Lainnya, pilih salah satu:
 - Hanya dapat diakses secara internal Universitas Multimedia Nusantara
 - Embargo publikasi karya ilmiah dalam kurun waktu tiga tahun.

Tangerang, 3 Juli 2025

Yang menyatakan

N. 

Nayasha Clarisa Dwisutrisna

**Jika tidak bisa membuktikan LoA jurnal/HKI, saya bersedia mengizinkan penuh karya ilmiah saya untuk dipublikasikan ke KC UMN dan menjadi hak institusi UMN.

HALAMAN PERSEMBAHAN / MOTTO

"Under this pressure, under this weight. We are diamonds taking shape."

Coldplay

UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

KATA PENGANTAR

Puji Syukur atas berkat dan rahmat kepada Tuhan Yang Maha Esa, atas selesainya penulisan laporan Skripsi ini dengan judul: Implementasi Algoritma Random Forest dalam Deteksi Penyakit Jantung dilakukan untuk memenuhi salah satu syarat untuk mencapai gelar Sarjana Komputer Jurusan Informatika Pada Fakultas Teknik dan Informatika Universitas Multimedia Nusantara. Saya menyadari bahwa, tanpa bantuan dan bimbingan dari berbagai pihak, dari masa perkuliahan sampai pada penyusunan laporan Skripsi ini, sangatlah sulit bagi saya untuk menyelesaikan laporan Skripsi ini. Oleh karena itu, saya mengucapkan terima kasih kepada:

1. Bapak Dr. Ir. Andrey Andoko, M.Sc., selaku Rektor Universitas Multimedia Nusantara.
2. Bapak Dr. Eng. Niki Prastomo, S.T., M.Sc., selaku Dekan Fakultas Teknik dan Informatika Universitas Multimedia Nusantara.
3. Bapak Arya Wicaksana, S.Kom., M.Eng.Sc., OCA, selaku Ketua Program Studi Informatika Universitas Multimedia Nusantara.
4. Ibu Dr. Maria Irminda Prasetyowati, S.Kom., M.T, sebagai Pembimbing yang telah memberikan bimbingan, arahan, dan motivasi atas terselesainya penelitian ini.
5. Ayah, Ibu, dan Kakak yang telah memberikan bantuan dukungan secara material, moral, dan spiritual sehingga penulis dapat menyelesaikan dengan lancar untuk penelitian ini.
6. Dr. Hamzah Muhammad Zein Sp.JP, FIHA yang telah bersedia menjadi narasumber terkait topik penyakit jantung untuk memenuhi kelengkapan informasi dalam penelitian ini.
7. Atthar Alifiano Ibrahim yang telah memberikan bantuan dukungan moral dan spiritual, sehingga penulis dapat menyelesaikan penelitian ini.

8. Mishel, Ayu, Dilla, Nisaul, Salwa, Abelia, Pniel, Nabiilah, Pasha, Indah, Laras, Nuke, Abigail, Maysha, dan Vanesa yang memberi semangat penulis dalam menyelesaikan penelitian ini.

Semoga penelitian ini bermanfaat, baik sebagai sumber informasi maupun sumber inspirasi, bagi para pembaca.

Tangerang, 3 Juli 2025

N.



Nayasha Clarisa Dwisutrisna

UMN

UNIVERSITAS
MULTIMEDIA
NUSANTARA

IMPLEMENTASI ALGORITMA RANDOM FOREST DALAM DETEKSI PENYAKIT JANTUNG

Nayasha Clarisa Dwisutrisna

ABSTRAK

Dilansir dari Organisasi Kesehatan Dunia (WHO), penyebab kematian secara global dengan peringkat paling atas berkaitan dengan Penyakit Kardiovaskular (penyakit jantung iskemik dan *stroke*). Penyakit jantung iskemik atau lebih dikenal dengan penyakit jantung koroner menjadi pembunuh terbesar di dunia yang menyebabkan 13% penyebab kematian. Maka, seiring berkembangnya teknologi hal ini dapat dijadikan solusi potensial berupa *software* dalam membantu deteksi risiko penyakit jantung dengan menganalisis beberapa parameter kesehatan menggunakan algoritma *Random Forest*. Pembangunan model menggunakan *dataset framingham* melibatkan 11 fitur medis yang relevan dengan penyakit jantung koroner dan kelas dengan klasifikasi positif dan negatif. Kemudian, uji coba model dilakukan dengan beberapa skenario seperti menggunakan *Hyperparameter Tuning* dan teknik SMOTE. Hasil terbaik diperoleh pada skenario dengan kombinasi *hyperparameter tuning* dan SMOTE dengan hasil akurasi 73.89%, *Precision* 25.89%, *Recall* 38.67%, dan *F1-Score* 31.02%, karena kombinasi ini mampu mengoptimalkan *parameter* model dan menyeimbangkan distribusi kelas, sehingga meningkatkan kemampuan model dalam mendeteksi kasus positif dengan *Recall* dan *F1-Score* yang lebih baik. Hasil tersebut menunjukkan bahwa akurasi tinggi, belum menjamin performa yang seimbang dalam mendeteksi kasus positif dan negatif dalam kondisi *imbalanced dataset* seperti dalam penelitian. Sehingga, penelitian perlu dilakukan eksplorasi lebih lanjut dengan teknik lain dalam penyeimbangan data dan penggunaan seleksi fitur.

Kata kunci: *Hyperparameter Tuning*, *Imbalanced Dataset*, Penyakit Jantung Koroner, *Random Forest*, SMOTE

U N I V E R S I T A S
M U L T I M E D I A
N U S A N T A R A

IMPLEMENTATION OF RANDOM FOREST ALGORITHM IN HEART DISEASE DETECTION

Nayasha Clarisa Dwisutrisna

ABSTRACT

According to the World Health Organization (WHO), the leading cause of death globally is cardiovascular disease (including ischemic heart disease and stroke). Ischemic heart disease, also known as coronary heart disease, is the world's top killer, accounting for 13% of all deaths. With the advancement of technology, this issue presents a potential solution in the form of software that can assist in detecting heart disease risk by analyzing several health parameters using the Random Forest algorithm. The model was built using the Framingham dataset, which includes 11 medical features relevant to coronary heart disease and a binary classification (positive and negative). The model was tested under several scenarios, including hyperparameter tuning and the application of the SMOTE technique. The best results were obtained using a combination of hyperparameter tuning and SMOTE, achieving an accuracy of 73.89%, precision of 25.89%, recall of 38.67%, and F1-score of 31.02%. This combination optimized the model's parameters and balanced the class distribution, leading to improved ability in detecting positive cases, particularly in terms of recall and F1-score. These results indicate that a high accuracy alone does not guarantee balanced performance in detecting both positive and negative cases in imbalanced datasets. Therefore, further exploration is needed using other data balancing techniques and feature selection methods.

Keywords: *Coronary Heart Disease, Hyperparameter Tuning, Imbalanced Dataset, Random Forest, SMOTE*

U N I V E R S I T A S
M U L T I M E D I A
N U S A N T A R A

DAFTAR ISI

HALAMAN JUDUL	i
PERNYATAAN TIDAK MELAKUKAN PLAGIAT	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH	iv
HALAMAN PERSEMBAHAN/MOTO	v
KATA PENGANTAR	vi
ABSTRAK	viii
ABSTRACT	ix
DAFTAR ISI	x
DAFTAR TABEL	xii
DAFTAR GAMBAR	xiii
DAFTAR KODE	xiv
DAFTAR RUMUS	xv
DAFTAR LAMPIRAN	xvi
BAB 1 PENDAHULUAN	1
1.1 Latar Belakang Masalah	1
1.2 Rumusan Masalah	3
1.3 Batasan Permasalahan	3
1.4 Tujuan Penelitian	3
1.5 Manfaat Penelitian	3
1.6 Sistematika Penulisan	4
BAB 2 LANDASAN TEORI	6
2.1 Penelitian Terdahulu	6
2.2 Penyakit Jantung	12
2.3 Algoritma Random Forest	13
2.4 Confusion Matrix	15
2.5 SMOTE	17
2.6 Random Search	17
2.7 Log Loss Validation	17
2.8 Splitting Rasio Data	18
BAB 3 METODOLOGI PENELITIAN	20
3.1 Tahapan Penelitian	20
3.2 Pre-Processing Data	22
3.3 Pembangunan Model hingga Evaluasi	26
3.4 Perancangan	29
3.4.1 Flowchart	29
3.4.2 Wireframe	32
BAB 4 HASIL DAN DISKUSI	34
4.1 Spesifikasi Perangkat	34
4.2 Implementasi Model Algoritma Random Forest	34
4.2.1 Import Library	35
4.2.2 Load Dataset	36
4.2.3 Pre-Processing Dataset	36
4.2.4 Konfigurasi Hyperparameter Tuning	38
4.2.5 Training Model	40
4.2.6 Testing Model	41
4.2.7 Evaluasi Model	42
4.3 Uji Coba Model	43

4.4	Diskusi	47
4.5	Evaluasi Hasil Uji Coba Model	50
4.5.1	Evaluasi dengan Grafik Log Loss	50
4.5.2	Evaluasi dengan Validation Set	54
4.5.3	Evaluasi dengan Metrik Evaluasi	55
4.6	Hasil UI Input Data Manual	58
BAB 5	SIMPULAN DAN SARAN	60
5.1	Simpulan	60
5.2	Saran	61
DAFTAR PUSTAKA		62



DAFTAR TABEL

Tabel 2.1	Penelitian Terdahulu	6
Tabel 3.1	Tabel Fitur dan Tipe Data	23
Tabel 4.1	Ruang Hyperparameter Random Forest	38
Tabel 4.2	Tanpa Tuning dan Tanpa SMOTE	44
Tabel 4.3	Tanpa Tuning dan Dengan SMOTE	44
Tabel 4.4	Dengan Tuning dan Tanpa SMOTE	45
Tabel 4.5	Dengan Tuning dan Dengan SMOTE	46
Tabel 4.6	Perbandingan Hasil Antar Skenario	48
Tabel 4.7	Confusion Matrix Skenario Pertama	56
Tabel 4.8	Confusion Matrix Skenario Kedua	56
Tabel 4.9	Confusion Matrix Skenario Ketiga	57
Tabel 4.10	Confusion Matrix Skenario Keempat	57



DAFTAR GAMBAR

Gambar 2.1	Diagram Alur Cara Kerja Algoritma Random Forest	13
Gambar 3.1	Diagram Alur Kerja Penelitian	20
Gambar 3.2	Diagram Alur Pre-Processing Data	25
Gambar 3.3	Diagram Alur Pembangunan Model	27
Gambar 3.4	Diagram Alur Program	30
Gambar 3.5	Diagram Alur Modul Algoritma Random Forest	31
Gambar 3.6	Wireframe Input Data Manual	33
Gambar 4.1	Bar Chart Accuracy dan Metrik Evaluasi Keempat Skenario	49
Gambar 4.2	Learning Curve Log Loss Skenario Pertama	51
Gambar 4.3	Learning Curve Log Loss Skenario Kedua	52
Gambar 4.4	Learning Curve Log Loss Skenario Ketiga	53
Gambar 4.5	Learning Curve Log Loss Skenario Keempat	54
Gambar 4.6	Bar Chart Accuracy Validasi Keempat Skenario	55
Gambar 4.7	Bar Chart Metrik Evaluasi Keempat Skenario	58
Gambar 4.8	Tampilan Input Data Manual	59



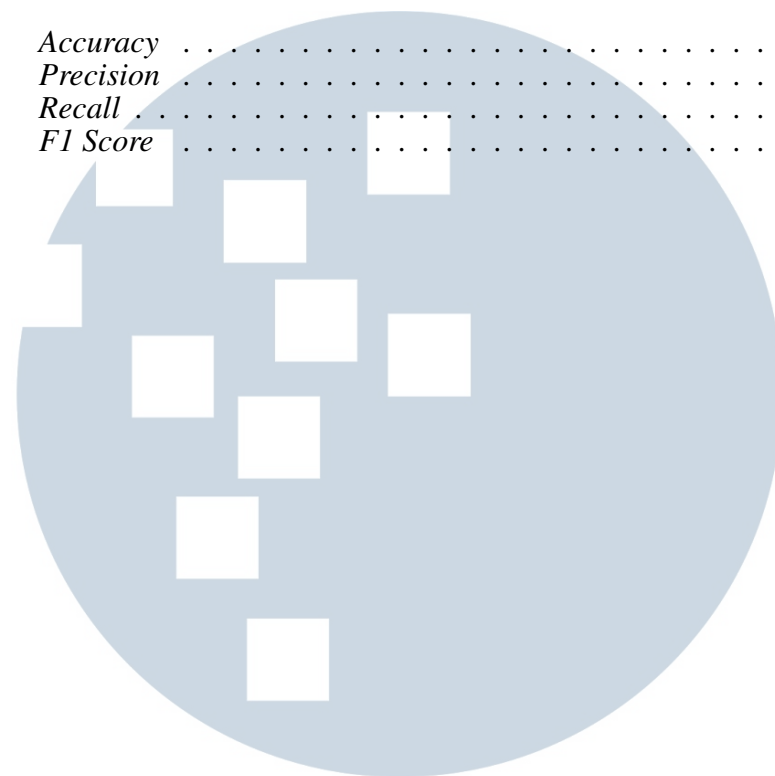
DAFTAR KODE

Kode 4.1	Potongan Kode <i>Import Library</i>	35
Kode 4.2	Potongan Kode <i>Load Dataset</i>	36
Kode 4.3	Potongan Kode <i>Pre-Processing</i>	36
Kode 4.4	Potongan Kode <i>Hyperparameter Tuning</i>	38
Kode 4.5	Potongan Kode <i>Training Model</i>	40
Kode 4.6	Potongan Kode <i>Testing Model</i>	41
Kode 4.7	Potongan Kode Evaluasi Model	42



DAFTAR RUMUS

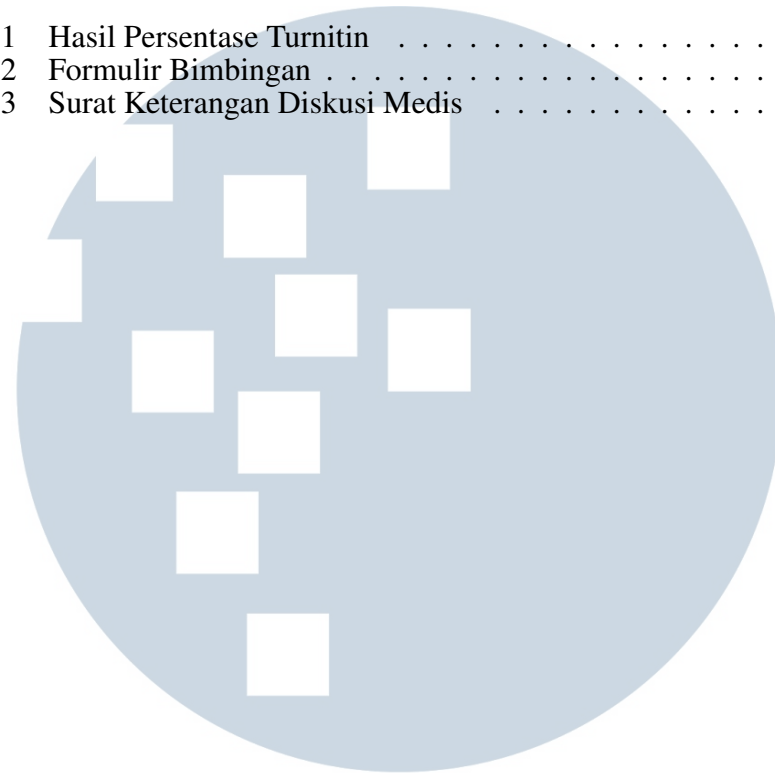
Rumus 2.1	<i>Accuracy</i>	16
Rumus 2.2	<i>Precision</i>	16
Rumus 2.3	<i>Recall</i>	16
Rumus 2.4	<i>F1 Score</i>	16



UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

DAFTAR LAMPIRAN

Lampiran 1	Hasil Persentase Turnitin	66
Lampiran 2	Formulir Bimbingan	67
Lampiran 3	Surat Keterangan Diskusi Medis	69



UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA