

**SISTEM PENDETEKSI PENYAKIT
KARDIOVASKULAR BERBASIS *WEB* DENGAN
FRAMEWORK FLASK MENGGUNAKAN METODE
*RANDOM FOREST***

SKRIPSI

Oleh:

Resty Nandya

201810225046



**PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS BHAYANGKARA JAKARTA RAYA
2022**

LEMBAR PERSETUJUAN PEMBIMBING

Judul Skripsi : Sistem Pendeteksi Penyakit Kardiovaskular berbasis
*Web dengan Framework Flask menggunakan Metode
Random Forest*

Nama Mahasiswa : Resty Nandya

Nomor Pokok Mahasiswa : 201810225046

Program Studi/Fakultas : Informatika/Ilmu Komputer

Tanggal Lulus Ujian Skripsi : 7 Juli 2022

Bekasi, 18 Juli 2020

MENYETUJUI,

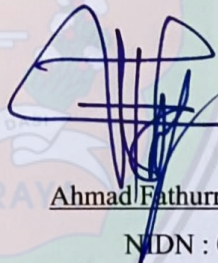
Pembimbing I

Pembimbing II



Prima Dina Atika, S.Kom., M.Kom.

NIDN : 0311037107



Ahmad Fathurrozi, S.E., M.M.S.I.

NIDN : 0327117402

LEMBAR PENGESAHAN

Judul Skripsi : Sistem Pendeteksi Penyakit Kardiovaskular berbasis
Web dengan Framework Flask menggunakan
Metode Random Forest
Nama Mahasiswa : Resty Nandya
Nomor Pokok Mahasiswa : 201810225046
Program Studi / Fakultas : Informatika / Ilmu Komputer
Tanggal Lulus Ujian Skripsi : 7 Juli 2022

Bekasi, 18 Juli 2022

MENGESAHKAN,

Ketua Tim Penguji : Sugiyatno, S.Kom., M.Kom.

NIDN: 0313077206

Penguji II : Rakhmat Purnomo, S.Pd., S.Kom., M.Kom.

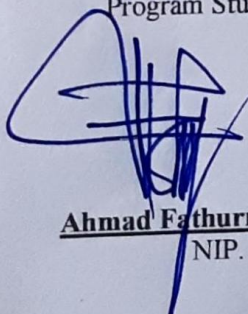
NIDN: 0322108201

Penguji III : Prima Dina Atika, S.Kom., M.Kom.

NIDN: 0311037107

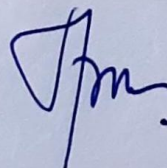
MENGETAHUI,

Ketua
Program Studi Informatika



Ahmad Fathurrozi, S.E., M.M.S.I.
NIP. 2012486

Dekan
Fakultas Ilmu Komputer



Dr. Dra. Tyastuti Sri Lestari, M.M.
NIP. 1408206



LEMBAR PERNYATAAN BUKAN PLAGIASI

Yang bertanda tangan dibawah ini :

Nama : Resty Nandya
NPM : 201810225046
Program Studi : Informatika
Fakultas : Ilmu Komputer
Judul Tugas Akhir : Sistem Pendeteksi Penyakit Kardiovaskular berbasis *Web*
dengan *Framework* Flask menggunakan Metode *Random Forest*

Dengan ini menyatakan bahwa hasil penulisan skripsi yang telah saya buat ini merupakan **hasil karya saya sendiri dan benar keasliannya**. Apabila dikemudian hari penulisan skripsi ini merupakan plagiat atau penjiplakan terhadap karya orang lain, maka saya bersedia mempertanggungjawabkan sekaligus bersedia menerima sanksi berdasarkan tata tertib di Universitas Bhayangkara Jakarta Raya.

Demikian pernyataan ini saya buat dalam keadaan sadar dan tidak dipaksakan dari pihak manapun.

Bekasi, 22 Juli 2022

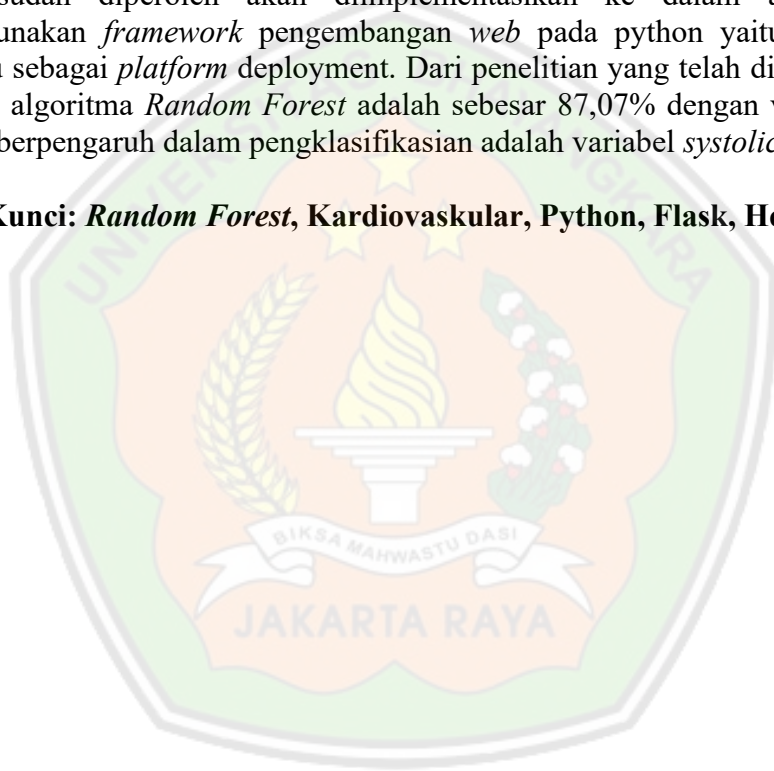
Penulis



ABSTRAK

Resty. 2022. Penyakit Kardiovaskular menurut definisi WHO (*World Health Organization*) adalah sekelompok gangguan jantung dan pembuluh darah, seperti penyakit jantung koroner, penyakit serebrovaskular, penyakit arteri perifer, penyakit jantung kongenital, penyakit jantung rematik, dan trombosis vena dalam. Menurut data yang diberikan oleh WHO (*World Health Organization*), Penyakit Kardiovaskular menjadi penyebab kematian nomor satu secara global dan diperkirakan 17,9 juta orang meninggal disetiap tahunnya. Tujuan dari penelitian ini adalah membangun sistem pendeteksi penyakit Kardiovaskular menggunakan metode *Random Forest* dengan tingkat akurasi yang lebih baik. Penelitian ini berbasis *web* dan dibangun dengan bahasa pemrograman Python. Adapun model yang sudah diperoleh akan diimplementasikan ke dalam aplikasi *web* menggunakan *framework* pengembangan *web* pada python yaitu Flask serta Heroku sebagai *platform* deployment. Dari penelitian yang telah dilakukan hasil akurasi algoritma *Random Forest* adalah sebesar 87,07% dengan variabel yang paling berpengaruh dalam pengklasifikasian adalah variabel *systolic*.

Kata Kunci: *Random Forest*, Kardiovaskular, Python, Flask, Heroku



ABSTRACT

Resty. 2022. Cardiovascular disease (CVD) according to the WHO (World Health Organization) definition is a group of heart and blood vessel disorders, such as coronary heart disease, cerebrovascular disease, peripheral arterial disease, congenital heart disease, rheumatic heart disease, and deep vein thrombosis. According to data provided by the WHO (World Health Organization), cardiovascular disease is the number one cause of death globally and an estimated 17.9 million people die each year. The purpose of this study is to build a cardiovascular disease detection system using the Random Forest method with a better level of accuracy. This research is web-based and built with the Python programming language. The model that has been obtained will be implemented into a web application using a web development framework in python, namely Flask and Heroku as a deployment platform. From the research that has been done, the accuracy of the Random Forest algorithm is 87.07% with the most influential variable in the classification is the systolic variable.

Keywords: *Random Forest, Cardiovascular, Python, Flask, Heroku*



LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIK

Sebagai sivitas akademik Universitas Bhayangkara Jakarta Raya, saya yang bertanda tangan di bawah ini :

Nama : Resty Nandya
NPM : 201810225046
Program Studi : Informatika
Fakultas : Ilmu Komputer
Jenis Karya : Skripsi

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Bhayangkara Jakarta Raya **Hak Bebas Royalti Non-Eksklusif (*Non-Exclusive Royalty-Free Right*)**, atas karya ilmiah saya yang berjudul :

Sistem Pendeteksi Penyakit Kardiovaskular berbasis *Web* dengan *Framework* Flask menggunakan metode *Random Forest*

beserta perangkat yang ada (bila diperlukan). Dengan hak bebas royalti non-ekslusif ini, Universitas Bhayangkara Jakarta Raya berhak menyimpan, mengalihmediakan, mengelolanya dalam bentuk pangkalan data (*database*), mendistribusikannya dan mempublikasikannya di Internet atau media lain untuk kepentingan akademis tanpa perlu meminta ijin dari saya selama tetap mencantumkan nama saya sebagai penulis dan sebagai pemilik hak cipta.

Segala bentuk tuntutan hukum yang timbul atas pelanggaran hak cipta dalam karya ilmiah ini menjadi tanggung jawab saya pribadi

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Bekasi
Pada tanggal : 22 Juli 2022
Yang Menyatakan



Resty Nandya

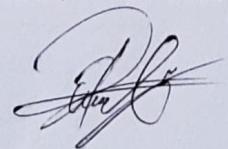
KATA PENGANTAR

Puji syukur kehadiran Allah SWT karena telah melimpahkan rahmat dan karunia-Nya sehingga peneliti dapat menyelesaikan Skripsi ini dengan judul **“Sistem Pendeteksi Penyakit Kardiovaskular berbasis Web dengan Framework Flask menggunakan Metode Random Forest”**. Skripsi ini disusun sebagai salah satu syarat untuk melanjutkan Skripsi pada program Strata-1 Program Studi Informatika Fakultas Ilmu Komputer Universitas Bhayangkara Jakarta Raya. Pada kesempatan ini peneliti menyampaikan rasa terima kasih kepada:

1. Bapak Irjen Polisi (P) Dr. Drs. Bambang Karsono, S.H., M.M. selaku Rektor Universitas Bhayangkara Jakarta Raya Bekasi.
2. Ibu Dr. Dra. Tyastuti Sri Lestari, M.M. selaku Dekan Fakultas Ilmu Komputer Universitas Bhayangkara Jakarta Raya Bekasi.
3. Bapak Ahmad Fathurrozi, S.E., M.M.S.I. selaku Ketua Program Studi Informatika Fakultas Ilmu Komputer Universitas Bhayangkara Jakarta Raya dan pembimbing dua dalam penyusunan Skripsi yang senantiasa memberikan arahan.
4. Ibu Prima Dina Atika, S.Kom., M.Kom. selaku pembimbing satu dalam penyusunan Skripsi yang telah membimbing dan memberikan arahan kepada peneliti selama penyusunan.
5. Orang tua dan keluarga tercinta atas motivasi dan do'anya sehingga semua yang dikerjakan dapat berjalan lancar.
6. Teman seperjuangan, Villa, terima kasih atas do'a dan segala dukungan.

Peneliti menyadari dalam penulisan Skripsi ini tidak lepas dari keterbatasan dan kekurangan peneliti, semoga hasil Skripsi ini dapat memberikan manfaat bagi pembaca.

Bekasi, 30 Juni 2022



Resty Nandya

DAFTAR ISI

Halaman

LEMBAR PERSETUJUAN PEMBIMBING	i
LEMBAR PENGESAHAN	ii
LEMBAR PERNYATAAN BUKAN PLAGIASI	iii
ABSTRAK	iv
ABSTRACT	v
LEMBAR PERNYATAAN PUBLIKASI	vi
KATA PENGANTAR.....	vii
DAFTAR ISI.....	viii
DAFTAR TABEL	xi
DAFTAR GAMBAR.....	xii
DAFTAR LAMPIRAN	xiv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Identifikasi Masalah.....	3
1.3 Rumusan Masalah.....	3
1.4 Batasan Masalah	3
1.5 Tujuan dan Manfaat Penelitian.....	4
1.5.1 Tujuan Penelitian	4
1.5.2 Manfaat Penelitian	4
1.6 Sistematika Penulisan	4
BAB II LANDASAN TEORI	6
2.1 Tinjauan Pustaka.....	6
2.2 Landasan Teori	9
2.2.1 <i>Web</i>	9
2.2.2 <i>Data Mining</i>	9
2.2.3 <i>Cross-Industry Standard Process for Data Mining (CRISP-DM)</i>	9
2.2.4 Klasifikasi	10
2.2.5 Penyakit Kardiovaskular	11

2.2.6	<i>Random Forest</i>	11
2.2.7	Evaluasi Model.....	12
2.2.8	XP (<i>Extreme Programming</i>)	14
2.2.9	UML (<i>Unified Modelling Language</i>).....	15
2.2.10	Python	18
2.2.11	HTML (<i>Hypertext Markup Language</i>)	18
2.2.12	CSS (<i>Cascading Style Sheet</i>)	18
2.2.13	Flask	19
2.2.14	Heroku.....	19
BAB III METODOLOGI PENELITIAN		20
3.1	Waktu dan Tempat Penelitian.....	20
3.2	Kerangka Pikir	20
3.3	Metode Analisis	21
3.3.1	Analisis Sistem Berjalan	21
3.3.2	Analisis Permasalahan	21
3.3.3	Analisis Sistem Usulan	22
3.4	Metode Pengumpulan Data.....	22
3.5	Metode <i>Data Mining</i>	23
3.5.1	<i>Business Understanding</i>	23
3.5.2	<i>Data Understanding</i>	23
3.5.3	<i>Data Preparation</i>	25
3.6	Analisis Kebutuhan.....	33
3.6.1	Kebutuhan Perangkat Keras	33
3.6.2	Kebutuhan Perangkat Lunak	33
BAB IV PERANCANGAN SISTEM DAN IMPLEMENTASI		34
4.1	<i>Modelling</i>	34
4.2	<i>Evaluation</i>	35
4.3	<i>Deployment</i>	36
4.3.1	Perancangan Sistem	36

4.3.2 Implementasi.....	40
4.3.3 Pengujian.....	43
BAB V PENUTUP.....	45
5.1 Kesimpulan.....	45
5.2 Saran	45
DAFTAR PUSTAKA.....	46
LAMPIRAN.....	50



DAFTAR TABEL

	Halaman
Tabel 2.1 <i>State of The Art</i> tentang Prediksi Penyakit Jantung	7
Tabel 2.2 Kelebihan dan Kekurangan <i>Random Forest</i>	12
Tabel 2.3 Simbol <i>Use Case Diagram</i>	15
Tabel 2.4 Simbol <i>Activity Diagram</i>	16
Tabel 2.5 Simbol <i>Sequence Diagram</i>	17
Tabel 2.6 Simbol <i>Class Diagram</i>	17
Tabel 3.1 <i>Dataset</i> Penyakit Kardiovaskular	23
Tabel 4.1 Deskripsi <i>Use Case</i>	37
Tabel 4.2 Pengujian Sistem	43



DAFTAR GAMBAR

	Halaman
Gambar 2.1 CRISP-DM <i>Phase</i>	9
Gambar 2.2 <i>Workflow Random Forest Classifier</i>	12
Gambar 2.3 <i>Confusion Matrix</i>	12
Gambar 2.4 Siklus Metode <i>Extreme Programming (XP)</i>	14
Gambar 3.1 Kerangka Pikir.....	20
Gambar 3.2 Sistem Berjalan	21
Gambar 3.3 Sistem Usulan.....	22
Gambar 3.4 <i>Correlation Heatmap</i>	24
Gambar 3.5 Deskriptif Statistik Data	25
Gambar 3.6 <i>Source Code Drop</i> Variabel “id”	25
Gambar 3.7 <i>Source Code</i> Memeriksa Data Duplikat.....	26
Gambar 3.8 <i>Source Code Drop</i> Data Duplikat	26
Gambar 3.9 <i>Source Code</i> Memeriksa <i>Missing Values</i>	26
Gambar 3.10 <i>Source Code Rename Columns</i>	27
Gambar 3.11 <i>Source Code</i> Konversi variabel “age”	27
Gambar 3.12 <i>Source Code</i> konversi variabel “gender”	27
Gambar 3.13 <i>Source Code Create</i> Variabel “bmi”	28
Gambar 3.14 <i>Boxplot</i> Variabel “bmi”	28
Gambar 3.15 <i>Source Code Remove Outlier</i> “bmi”	29
Gambar 3.16 <i>Boxplot</i> variabel “systolic” dan “diastolic”	29
Gambar 3.17 <i>Source Code Remove Outlier</i> “systolic” dan “diastolic”	29
Gambar 3.18 <i>Source Code Remove Outlier</i> “systolic”	30
Gambar 3.19 <i>Source Code Remove Outlier</i> “diastolic”	30
Gambar 3.20 <i>Source Code Data Splitting</i>	31
Gambar 3.21 Standarisasi Data	31
Gambar 3.22 <i>Create df_scaled</i>	32
Gambar 3.23 Ringkasan Data Setelah Tahap <i>Preprocessing</i>	32
Gambar 4.1 Desain Penelitian.....	34
Gambar 4.2 <i>Source Code Train Model</i>	34
Gambar 4.3 <i>Confusion Matrix</i>	35

Gambar 4.4 Evaluasi Model.....	36
Gambar 4.5 <i>Use Case Diagram</i>	37
Gambar 4.6 <i>Activity Diagram</i> (Mengisi <i>Form</i> Prediksi)	38
Gambar 4.7 <i>Activity Diagram</i> (Memperoleh Hasil Prediksi)	38
Gambar 4.8 <i>Sequence Diagram</i> (Mengisi <i>Form</i> Prediksi).....	39
Gambar 4.9 <i>Sequence Diagram</i> (Memperoleh Hasil Prediksi).....	39
Gambar 4.10 <i>Class Diagram</i>	40
Gambar 4.11 Halaman <i>Home</i>	41
Gambar 4.12 Halaman Mulai Prediksi.....	41
Gambar 4.13 Hasil Prediksi (Normal)	42
Gambar 4.14 Halaman Hasil Prediksi (Kardiovaskular).....	42



DAFTAR LAMPIRAN

Lampiran 1 Perhitungan <i>Feature Importance</i>	50
Lampiran 2 Diagram pohon Keputusan (<i>Tree</i>)	52
Lampiran 3 <i>Source Code</i>	53
Lampiran 4 <i>Deployment Aplikasi</i>	62

