



**MODEL DATA MINING UNTUK PREDIKSI PENYAKIT GAGAL
JANTUNG PADA USIA DEWASA HINGGA LANSIA DENGAN
MENGUNAKAN METODE K-NEAREST NEIGHBOR**

LAPORAN SKRIPSI

MUHAMAD RIZKI

(41819010005)

AHMAD WILDAN

(41819010006)

U RAKA SHAKURALIM KUSMANA

(41819010056)

MERCU BUANA

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA
JAKARTA**

2023



**MODEL DATA MINING UNTUK PREDIKSI PENYAKIT GAGAL
JANTUNG PADA USIA DEWASA HINGGA LANSIA DENGAN
MENGUNAKAN METODE K-NEAREST NEIGHBOR**

LAPORAN SKRIPSI

Diajukan Untuk Melengkapi Salah Satu Syarat
Memperoleh Gelar Sarjana Komputer

Oleh:

MUHAMAD RIZKI
(41819010005)

AHMAD WILDAN
(41819010006)

RAKA SHAKURALIM KUSMANA
(41819010056)

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA
JAKARTA**

2023

HALAMAN PERNYATAAN KARYA SENDIRI

Saya yang bertanda tangan dibawah ini:

Nama Mahasiswa (1) : Muhamad Rizki
NIM : 41819010005
Nama Mahasiswa (2) : Raka Shakuralim Kusmana
NIM : 41819010056
Nama Mahasiswa (3) : Ahmad Wildan
NIM : 41819010006
Program Studi : Sistem Informasi
Judul Laporan Skripsi : Model Data Mining Untuk Prediksi Penyakit Gagal Jantung Pada Usia Dewasa Hngga Lansia Dengan Menggunakan Metode K-Nearest Neighbor

Menyatakan bahwa Laporan Skripsi ini adalah hasil karya saya sendiri dan bukan plagiat, serta semua sumber baik yang dikutip maupun dirujuk telah saya nyatakan dengan benar. Apabila ternyata ditemukan di dalam Laporan Skripsi saya terdapat unsur plagiat, maka saya siap mendapatkan sanksi akademis yang berlaku di Universitas Mercu Buana.

UNIVERSITAS
MERCU BUANA

Jakarta, 01 Agustus 2023

Yang menyatakan,



(Muhamad Rizki)

HALAMAN PENGESAHAN

Nama Mahasiswa (3) : Muhamad Rizki
NIM (41819010005)
Nama Mahasiswa (1) : Ahmad Wildan
NIM (41819010006)
Nama Mahasiswa (2) : Raka Shakuralim Kusmana
NIM (41819010056)

Judul Tugas Akhir : Model Data Mining Untuk Prediksi Penyakit Gagal Jantung Pada Usia Dewasa Hingga Lansia Dengan Menggunakan Metode K-Nearest Neighbor

Tugas Akhir ini telah diperiksa dan disidangkan sebagai salah satu persyaratan untuk memperoleh gelar Sarjana pada Program Studi Sistem Informasi, Fakultas Ilmu Komputer, Universitas Mercu Buana.

Jakarta, 01 Agustus 2023

Menyetujui,

Pembimbing Misbahul Fajri, Dr. M.TI

Penguji 1 Dr.Bambang Jokonowo, S.Si , MTI

Penguji 2 Andi Nugroho, ST, M. Kom

Mengetahui,

UNIVERSITAS
MERCU BUANA

(Dr.Bambang Jokonowo,S.Si., M.T.I)
Dekan Fakultas Ilmu Komputer

((Dr. Ruci Meiyanti, M.Kom)
Ka.Prodi Sistem Informasi

KATA PENGANTAR

Puji syukur kami panjatkan kehadirat Allah SWT, yang telah melimpahkan rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan Metodologi Penelitian Teknologi Informasi dengan judul “MODEL DATA MINING UNTUK PREDIKSI PENYAKIT GAGAL JANTUNG PADA USIA DEWASA HINGGA LANSIA DENGAN MENGGUNAKAN METODE K-NEAREST NEIGHBOR”.

Penulis menyadari bahwa laporan Metodologi Penelitian Teknologi Informasi ini tidak akan selesai tanpa bantuan dan bimbingan dari berbagai pihak. Oleh karena itu penulis mengucapkan terimakasih kepada:

1. Bapak Prof. Dr. Andi Adriansyah, M. Eng selaku Rektor Universitas Mercu Buana
2. Bapak Dr. Bambang Jokonowo, S.Si.,M.T.i selaku Dekan Fasilkom dan Dosen Penguji 1
3. Ibu Dr. Ruci Meiyanti, M.Kom selaku Kepala Program Studi Sistem Informasi
4. Bapak Misbahul Fajri, Dr., M.T.I selaku dosen pembimbing yang telah memberikan bimbingan, bantuan, serta saran sehingga penulis dapat menyelesaikan laporan ini.
5. Bapak Andi Nugroho, ST, M.Kom selaku dosen pengampu tugas akhir dan dosen penguji
6. Semua Bapak dan Ibu dosen Program Studi Informasi
7. Kepada kedua orangtua yang selalu memberikan doa, motivasi dan semangat untuk menyelesaikan laporan ini.
8. Sahabat dan semua pihak yang telah membantu namun tidak bisa disebutkan satu persatu

Akhir kata, penulis berharap segala bentuk saran serta masukan bahkan kritik yang membangun dari berbagai pihak. Semoga laporan tugas akhir ini dapat bermanfaat bagi semua pihak dan khususnya dalam bidang teknologi informasi.

Jakarta, 01 Agustus 2023

(Muhamad Rizki)



HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS

Sebagai sivitas akademik Universitas Mercu Buana, saya yang bertanda tangan di bawah ini:

Nama Mahasiswa (1) : Muhamad Rizki
NIM : 41819010005
Nama Mahasiswa (2) : Raka Shakuralim Kusmana
NIM : 41819010056
Nama Mahasiswa (3) : Ahmad Wildan
NIM : 41819010006
Program Studi : Sistem Informasi
Judul Laporan Skripsi : Model Data Mining Untuk Prediksi Penyakit Gagal Jantung Pada Usia Dewasa Hingga Lansia Dengan Menggunakan Metode K-Nearest Neighbor

Demi pengembangan ilmu pengetahuan, dengan ini memberika izin dan menyetujui untuk memberikan kepada Universitas Mercu Buana **Hak Bebas Royalti Non-Eksklusif (Non-exclusive Royalty-Free Right)** atas karya ilmiah saya yang berjudul di atas beserta perangkat yang ada (jika diperlukan).

Dengan Hak Bebas Royalti Non-Eksklusif ini Universitas Mercu Buana berhak menyimpan, mengalihmedia/format-kan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan mempublikasikan Laporan Skripsi saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta. Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, 01 Agustus 2023

Yang menyatakan,


(Muhamad Rizki)

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERNYATAAN KARYA SENDIRI	ii
HALAMAN PENGESAHAN.....	iii
KATA PENGANTAR	iv
HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS	vi
ABSTRAK	vii
<i>ABSTRACT</i>	viii
DAFTAR ISI.....	ix
DAFTAR TABEL.....	xi
DAFTAR GAMBAR	xii
DAFTAR LAMPIRAN	xiii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	4
1.3 Tujuan Penelitian	5
1.4 Batasan Masalah.....	5
1.5 Manfaat Penelitian.....	6
1.6 Sistematika Penulisan.....	6
BAB II LANDASAN TEORI.....	8
2.1 Teori/Konsep Terkait	8
2.1.1 DATA MINING	8
2.1.2 JUPYTERLAB.....	8
2.1.3 DATASET	9
2.1.4 METODE ALGORITMA K-NEAREST NEIGHBOR	9

2.1.5	METODE ALGORITMA LOGISTIC REGRESSION	10
2.1.6	METODE ALGORITMA RANDOM FOREST	10
2.1.7	METODE ALGORITMA NEURAL NETWORK.....	12
2.1.8	SPSS (STATISTICAL PROGRAM FOR SOCIAL SCIENCE)	12
2.1.9	CONFUSION MATRIX	12
2.1.10.	PYTHON	15
2.2	Penelitian Terdahulu	15
2.3	Analisis Literature Review	36
BAB III METODE PENELITIAN.....		37
3.1	Deskripsi Sumber Data.....	37
3.2	Teknik Pengumpulan Data	50
3.3	Diagram Alir Penelitian	51
3.4	<i>Machine Learning Life Cycle</i>	57
BAB IV HASIL DAN PEMBAHASAN		59
4.1.	Hasil Pengujian	59
4.1.1	Hasil	59
4.2.	Tools Yang Digunakan	72
4.3.	Pembahasan.....	73
BAB V KESIMPULAN DAN SARAN		75
5.1.	Kesimpulan	75
5.2.	Saran.....	75
DAFTAR PUSTAKA		76
LAMPIRAN.....		82
	Lampiran 1 Asistensi Kartu Bimbingan	82
	Lampiran 2 <i>Curriculum Vitae</i>	84

DAFTAR TABEL

Tabel 2. 1 Confusion Matrix	14
Tabel 2. 2 Literature Review.....	16
Tabel 3. 1 Statistik Total	37
Tabel 3. 2 <i>Age</i>	38
Tabel 3. 3 <i>Sex</i>	39
Tabel 3. 4 <i>Chest Pain Type</i>	40
Tabel 3. 5 <i>Resting BP</i>	40
Tabel 3. 6 <i>Cholesterol</i>	42
Tabel 3. 7 <i>Fasting BS</i>	45
Tabel 3. 8 <i>Resting ECG</i>	45
Tabel 3. 9 <i>Max HR</i>	45
Tabel 3. 10 <i>Exercise</i>	47
Tabel 3. 11 <i>Oldpeak</i>	48
Tabel 3. 12 <i>ST Slope</i>	50
Tabel 3. 13 <i>Heart Disease</i>	50

DAFTAR GAMBAR

Gambar 1. 1 Anatomi Jantung	1
Gambar 1. 2 Perbandingan Jantung Normal Dan Gagal Jantung.....	3
Gambar 3.4. 1 <i>Machine Learning Life Cycle</i>	57
Gambar 4.1. 1 <i>Import Package</i>	59
Gambar 4.1. 2 <i>Read Data</i>	59
Gambar 4.1. 3 <i>Informasi Data</i>	60
Gambar 4.1. 4 Statistik Deskriptif Data	60
Gambar 4.1. 5 <i>Missing Value Detection</i>	61
Gambar 4.1. 6 <i>Duplicate Value Detection</i>	61
Gambar 4.1. 7 <i>Drop Column</i>	62
Gambar 4.1. 8 <i>Mendefinisikan Tipe Data</i>	62
Gambar 4.1. 9 <i>Memeriksa Outliers</i>	63
Gambar 4.1. 10 <i>Mendefinisikan Outliers dan Removing Outliers</i>	64
Gambar 4.1. 11 <i>Heatmap Korelasi Analisis</i>	65
Gambar 4.1. 12 <i>Melakukan Encoding</i>	66
Gambar 4.1. 13 <i>Mendefinisikan Variabel Feature dan Target</i>	66
Gambar 4.1. 14 <i>SMOTE</i>	66
Gambar 4.1. 15 <i>Train Test Split</i>	67
Gambar 4.1. 16 <i>Confusion Matrix K-Nearest Neighbor</i>	68
Gambar 4.1. 17 <i>Confusion Matrix Neural Network</i>	69
Gambar 4.1. 18 <i>Confusion Matrix Logistic Regression</i>	70
Gambar 4.1. 19 <i>Confusion Matrix Random Forest</i>	71
Gambar 4.1. 20 <i>Model Comparisson</i>	71

DAFTAR LAMPIRAN

Lampiran 1 Asistensi Kartu Bimbingan	82
Lampiran 2 <i>Curriculum Vitae</i>	84

