

SKRIPSI

**IMPLEMENTASI METODE RANDOM FOREST UNTUK
KLASIFIKASI PASIEN YANG TERKENA
GAGAL JANTUNG**

***IMPLEMENTATION OF THE RANDOM FOREST METHOD ON
THE CLASSIFICATION OF PATIENTS AFFECTED BY
HEART FAILURE***



Disusun oleh

**NABILA ZULFIKA HEMADEWI
19101018**

**PROGRAM STUDI S1 TEKNIK TELEKOMUNIKASI
FAKULTAS TEKNIK TELEKOMUNIKASI DAN ELEKTRO
INSTITUT TEKNOLOGI TELKOM PURWOKERTO**

2023

SKRIPSI

**IMPLEMENTASI METODE RANDOM FOREST UNTUK
KLASIFIKASI PASIEN YANG TERKENA
GAGAL JANTUNG**

***IMPLEMENTATION OF THE RANDOM FOREST METHOD ON
THE CLASSIFICATION OF PATIENTS AFFECTED BY
HEART FAILURE***



Disusun oleh

**NABILA ZULFIKA HEMADEWI
19101018**

**PROGRAM STUDI S1 TEKNIK TELEKOMUNIKASI
FAKULTAS TEKNIK TELEKOMUNIKASI DAN ELEKTRO
INSTITUT TEKNOLOGI TELKOM PURWOKERTO**

2023

**IMPLEMENTASI METODE RANDOM FOREST UNTUK
KLASIFIKASI PASIEN YANG TERKENA
GAGAL JANTUNG**

***IMPLEMENTATION OF THE RANDOM FOREST METHOD ON
THE CLASSIFICATION OF PATIENTS AFFECTED BY
HEART FAILURE***

**Skripsi ini digunakan sebagai salah satu syarat untuk memperoleh
Gelar Sarjana Teknik (S.T.)
Di Institut Teknologi Telkom Purwokerto
2023**

Disusun oleh

**NABILA ZULFIKA HEMADEWI
19101018**

DOSEN PEMBIMBING

**Rahmat Widadi, S.Pd., M.Eng.
Irmayatul Hikmah, S.Si., M.Si.**

**PROGRAM STUDI S1 TEKNIK TELEKOMUNIKASI
FAKULTAS TEKNIK TELEKOMUNIKASI DAN ELEKTRO
INSTITUT TEKNOLOGI TELKOM PURWOKERTO
2023**

HALAMAN PENGESAHAN


IMPLEMENTASI METODE RANDOM FOREST UNTUK KLASIFIKASI PASIEN YANG TERKENA GAGAL JANTUNG


IMPLEMENTATION OF THE RANDOM FOREST METHOD ON THE CLASSIFICATION OF PATIENTS AFFECTED BY HEART FAILURE

Disusun oleh
NABILA ZULFIKA HEMADEWI
19101018


Telah dipertanggungjawabkan di hadapan Tim Penguji pada tanggal 14 Februari
2023

Susunan Tim Penguji

Pembimbing Utama : Rahmat Widadi, S. Pd., M.Eng. ()
NIDN. 0631039201

Pembimbing Pendamping : Irmayatul Hikmah, S.Si., M.Si. ()
NIDN. 0610069301

Penguji 1 : Zein Hanni Pradana, S.T., M.T. ()
NIDN. 0604039001


Penguji 2 : Dr. Wahyu Pamungkas, S.T., M.T. ()
NIDN. 0606037801

Mengetahui,

Ketua Program Studi S1 Teknik Telekomunikasi

Institut Teknologi Telkom Purwokerto

a/n


Prasetyo Yulianto, S.T., M.T.

NIDN. 0620079201

HALAMAN PERNYATAAN ORISINALITAS

Dengan ini saya, **NABILA ZULFIKA HEMADEWI**, menyatakan bahwa skripsi dengan judul " **IMPLEMENTASI METODE RANDOM FOREST UNTUK KLASIFIKASI PASIEN YANG TERKENA GAGAL JANTUNG** " adalah benar-benar karya saya sendiri. Saya tidak melakukan penjiplakan kecuali melalui pengutipan sesuai dengan etika keilmuan yang berlaku. Saya bersedia menanggung risiko ataupun sanksi yang dijatuhkan kepada saya apabila ditemukan pelanggaran terhadap etika keilmuan dalam skripsi saya ini.

Purwokerto, 02 Februari 2023

Yang menyatakan,



(Nabila Zulfika Hemadewi)

PRAKATA

Puji dan syukur penulis panjatkan kehadiran Allah SWT yang telah melimpahkan kasih dan sayang-Nya sehingga penulis dapat menyelesaikan skripsi yang berjudul **"Implementasi Metode Random Forest Untuk Klasifikasi Pasien Yang Terkena Gagal Jantung"**.

Maksud dari penyusunan skripsi ini adalah untuk memenuhi salah satu syarat dalam menempuh ujian sarjana Teknik Telekomunikasi pada Fakultas Teknik Telekomunikasi dan Elektro Institut Teknologi Telkom Purwokerto.

Dalam penyusunan skripsi ini, banyak pihak yang sangat membantu penulis dalam berbagai hal. Oleh karena itu, penulis sampaikan rasa terima kasih yang sedalam-dalamnya kepada:

1. Keluarga tercinta yang telah mendoakan saya dan memberikan perhatian.
2. Bapak Rahmat Widadi, S. Pd., M.Eng. selaku pembimbing I.
3. Ibu Irmayatul Hikmah, S.Si., M.Si. selaku pembimbing II.
4. Bapak Prasetyo Yuliantoro, S.T., M.T. ketua Program Studi S1 Teknik Telekomunikasi.
5. Bapak Dr. Arfianto Fahmi, S.T., M.T. selaku Rektor Institut Teknologi Telkom Purwokerto.
6. Seluruh dosen, staf dan karyawan Program studi S1 Teknik Telekomunikasi Institut Teknologi Telkom Purwokerto.
7. Teman – teman yang telah memberikan semangat dan dukungan.

Purwokerto, 02 Februari 2023



(Nabila Zulfika Hemadewi)

DAFTAR ISI

HALAMAN JUDUL	I
HALAMAN PENGESAHAN.....	IV
HALAMAN PERNYATAAN ORISINALITAS	V
PRAKATA	VI
ABSTRAK	VII
ABSTRACT	VIII
DAFTAR ISI.....	IX
DAFTAR GAMBAR.....	XI
DAFTAR TABEL	XV
BAB 1 PENDAHULUAN	1
1.1 LATAR BELAKANG	1
1.2 RUMUSAN MASALAH.....	3
1.3 BATASAN MASALAH.....	3
1.4 TUJUAN.....	4
1.5 MANFAAT	4
1.6 SISTEMATIKA PENULISAN.....	4
BAB 2 DASAR TEORI.....	6
2.1 KAJIAN PUSTAKA.....	6
2.2 DASAR TEORI	12
2.2.1 <i>ARTIFICIAL INTELLIGENCE</i>	12
2.2.2 <i>MACHINE LEARNING</i>	13
2.2.3 KLASIFIKASI.....	15
2.2.4 <i>RANDOM FOREST</i>	15
2.2.5 <i>FEATURE IMPORTANCES</i>	17
2.2.6 <i>FEATURE SCALING</i>	18
2.2.7 <i>MIN MAX NORMALIZATION</i>	18
2.2.8 <i>IMBALANCE CLASS</i>	19
2.2.9 <i>ADAPTIVE SYNTETHIC (ADASYN)</i>	20
2.2.10 <i>K-FOLD CROSS VALIDATION</i>	22
2.2.11 <i>CONFUSION MATRIX</i>	24
2.2.12 BAHASA PEMROGRAMAN	25
BAB 3 METODE PENELITIAN.....	26
3.1 SUMBER DATA	26

3.2	<i>TOOLS</i>	41
3.3	ALUR PENELITIAN	42
3.4	ALUR PERANCANGAN PROGRAM	44
3.5	PENGUJIAN PROGRAM.....	48
BAB 4 HASIL DAN PEMBAHASAN		49
4.1	PARAMETER PENELITIAN	49
4.2	ANALISIS HASIL PENELITIAN.....	49
4.2.1	ANALISIS <i>PREPROCESSING</i> DATA.....	49
4.2.3	ANALISIS HASIL KLASIFIKASI	56
BAB V PENUTUP		76
5.1	KESIMPULAN	76
5.2	SARAN.....	76
DAFTAR PUSTAKA		78
LAMPIRAN.....		84

DAFTAR GAMBAR

Gambar 2.1 Kerangka <i>Artificial Intelligence</i> dan <i>Machine Learning</i>	13
Gambar 2.2 Arsitektur Algoritma <i>Random Forest</i>	16
Gambar 2.3 Ilustrasi <i>Min Max Normalization</i>	19
Gambar 2.4 Ilustrasi <i>10-Fold Cross Validation</i>	23
Gambar 3.1 Persentase Peristiwa Kematian Pasien Gagal Jantung.....	27
Gambar 3.2 Sebaran Pasien Yang Meninggal/Selamat Dari Gagal Jantung 28	
Gambar 3.3 Persentase Gender Pasien Gagal Jantung.....	29
Gambar 3.4 Sebaran Gender Pasien Gagal Jantung.....	29
Gambar 3.5 Sebaran Umur Pasien Penyakit Gagal Jantung	30
Gambar 3.6 Persentase Pasien Gagal Jantung Yang Memiliki Anemia.....	31
Gambar 3.7 Pasien Yang Memiliki Anemia	31
Gambar 3.8 Persentase Pasien Gagal Jantung Yang Memiliki Diabetes	32
Gambar 3.9 Pasien Yang Memiliki Diabetes.....	33
Gambar 3.10 Persentase Pasien Gagal Jantung Yang Memiliki Tekanan Darah Tinggi.....	34
Gambar 3.11 Pasien Yang Memiliki Tekanan Darah Tinggi	34
Gambar 3.12 Persentase Pasien Gagal Jantung Yang Merokok.....	36
Gambar 3.13 Pasien Yang Merokok	36
Gambar 3.14 Sebaran Fraksi Ejeksi Pasien Penyakit Gagal Jantung.....	37
Gambar 3.15 Sebaran <i>Creatinine phosphokinase</i> Pasien Penyakit Gagal Jantung.....	38
Gambar 3.16 Sebaran <i>Serum creatinine</i> Pasien Penyakit Gagal Jantung	39
Gambar 3.17 Sebaran <i>Serum sodium</i> Pasien Penyakit Gagal Jantung.....	39
Gambar 3.18 Sebaran Trombosit Pasien Penyakit Gagal Jantung.....	40
Gambar 3.19 Sebaran Periode <i>Follow Up</i> Pasien Penyakit Gagal Jantung ..	41
Gambar 3.20 Alur Penelitian	42
Gambar 3.21 Alur Perancangan Program.....	44
Gambar 4.1 Fitur Yang Paling Mempengaruhi Target	52
Gambar 4.2 Perbandingan Tiap Fitur Sebelum Di Scaling dan Sesudah Di Scaling	54

Gambar 4.3 Data Sebelum Di Seimbangkan Dengan ADASYN	54
Gambar 4.4 Jumlah Data Variabel Target Yang Belum Di Seimbangkan Dengan ADASYN.....	55
Gambar 4.5 Data Setelah Di Seimbangkan Dengan ADASYN.....	55
Gambar 1 <i>Feature Importances</i> Dari Skema 12 Fitur	84
Gambar 2 <i>Feature Importances</i> Dari Skema 11 Fitur	84
Gambar 3 <i>Feature Importances</i> Dari Skema 10 Fitur	85
Gambar 4 <i>Feature Importances</i> Dari Skema 9 Fitur	85
Gambar 5 <i>Feature Importances</i> Dari Skema 8 Fitur	86
Gambar 6 <i>Feature Importances</i> Dari Skema 7 Fitur	86
Gambar 7 <i>Feature Importances</i> Dari Skema 6 Fitur	87
Gambar 8 <i>Feature Importances</i> Dari Skema 5 Fitur	87
Gambar 9 <i>Feature Importances</i> Dari Skema 4 Fitur	88
Gambar 10 <i>Feature Importances</i> Dari Skema 3 Fitur	88
Gambar 11 <i>Feature Importances</i> Dari Skema 2 Fitur	89
Gambar 12 <i>Feature Importances</i> Dari Skema 1 Fitur	89
Gambar 13 <i>Feature Scaling</i> Dari Skema 12 Fitur.....	89
Gambar 14 <i>Feature Scaling</i> Dari Skema 11 Fitur.....	90
Gambar 15 <i>Feature Scaling</i> Dari Skema 10 Fitur.....	90
Gambar 16 <i>Feature Scaling</i> Dari Skema 9 Fitur	90
Gambar 17 <i>Feature Scaling</i> Dari Skema 8 Fitur	90
Gambar 18 <i>Feature Scaling</i> Dari Skema 7 Fitur	91
Gambar 19 <i>Feature Scaling</i> Dari Skema 6 Fitur	91
Gambar 20 <i>Feature Scaling</i> Dari Skema 5 Fitur	91
Gambar 21 <i>Feature Scaling</i> Dari Skema 4 Fitur	91
Gambar 22 <i>Feature Scaling</i> Dari Skema 3 Fitur	92
Gambar 23 <i>Feature Scaling</i> Dari Skema 2 Fitur	92
Gambar 24 <i>Feature Scaling</i> Dari Skema 1 Fitur	92
Gambar 25 Data Pada Skema 12 Fitur Yang Belum Di Seimbangkan Dengan Menggunakan ADASYN	93
Gambar 26 Data Pada Skema 11 Fitur Yang Belum Di Seimbangkan Dengan Menggunakan ADASYN	93

Gambar 27 Data Pada Skema 10 Fitur Yang Belum Di Seimbangkan Dengan Menggunakan ADASYN	94
Gambar 28 Data Pada Skema 9 Fitur Yang Belum Di Seimbangkan Dengan Menggunakan ADASYN	94
Gambar 29 Data Pada Skema 8 Fitur Yang Belum Di Seimbangkan Dengan Menggunakan ADASYN	95
Gambar 30 Data Pada Skema 7 Fitur Yang Belum Di Seimbangkan Dengan Menggunakan ADASYN	95
Gambar 31 Data Pada Skema 6 Fitur Yang Belum Di Seimbangkan Dengan Menggunakan ADASYN	96
Gambar 32 Data Pada Skema 5 Fitur Yang Belum Di Seimbangkan Dengan Menggunakan ADASYN	96
Gambar 33 Data Pada Skema 4 Fitur Yang Belum Di Seimbangkan Dengan Menggunakan ADASYN	97
Gambar 34 Data Pada Skema 3 Fitur Yang Belum Di Seimbangkan Dengan Menggunakan ADASYN	97
Gambar 35 Data Pada Skema 2 Fitur Yang Belum Di Seimbangkan Dengan Menggunakan ADASYN	98
Gambar 36 Data Pada Skema 12 Fitur Yang Sudah Di Seimbangkan Dengan Menggunakan ADASYN	98
Gambar 37 Data Pada Skema 11 Fitur Yang Sudah Di Seimbangkan Dengan Menggunakan ADASYN	99
Gambar 38 Data Pada Skema 10 Fitur Yang Sudah Di Seimbangkan Dengan Menggunakan ADASYN	99
Gambar 39 Data Pada Skema 9 Fitur Yang Sudah Di Seimbangkan Dengan Menggunakan ADASYN	100
Gambar 40 Data Pada Skema 8 Fitur Yang Sudah Di Seimbangkan Dengan Menggunakan ADASYN	100
Gambar 41 Data Pada Skema 7 Fitur Yang Sudah Di Seimbangkan Dengan Menggunakan ADASYN	101
Gambar 42 Data Pada Skema 6 Fitur Yang Sudah Di Seimbangkan Dengan Menggunakan ADASYN	101

Gambar 43 Data Pada Skema 5 Fitur Yang Sudah Di Seimbangkan Dengan Menggunakan ADASYN	102
Gambar 44 Data Pada Skema 4 Fitur Yang Sudah Di Seimbangkan Dengan Menggunakan ADASYN	102
Gambar 45 Data Pada Skema 3 Fitur Yang Sudah Di Seimbangkan Dengan Menggunakan ADASYN	103
Gambar 46 Data Pada Skema 2 Fitur Yang Sudah Di Seimbangkan Dengan Menggunakan ADASYN	103

DAFTAR TABEL

Tabel 2.1 Perbandingan Metode.....	9
Tabel 2.2 <i>Confusion Matrix</i>	24
Tabel 3.1 Variabel dan Deskripsi	26
Tabel 3.2 Peristiwa Kematian Pasien Gagal Jantung.....	28
Tabel 3.3 Gender Pasien Gagal Jantung.....	29
Tabel 3.4 Pasien Yang Memiliki Anemia.....	31
Tabel 3.5 Pasien Yang Memiliki Diabetes.....	33
Tabel 3.6 Pasien Yang Memiliki Tekanan Darah Tinggi.....	35
Tabel 3.7 Pasien Yang Merokok.....	36
Tabel 4.1 Data Awal Berdasarkan Kolom 10 Teratas	50
Tabel 4.2 Statistik Deskriptif Data Awal	50
Tabel 4.3 Data <i>Scaling</i> Berdasarkan 10 Kolom Teratas	53
Tabel 4.4 Statistik Deskriptif Data <i>Scaling</i>	53
Tabel 4.5 Jumlah Kelas Mayoritas Dan Minoritas Tiap Fitur Berdasarkan Total Data	55
Tabel 4.6 Sebaran Data <i>Training</i> Dan Data <i>Testing</i> Dari Skema 12 Fitur	56
Tabel 4.7 Sebaran Data <i>Training</i> Dan Data <i>Testing</i> Dari Skema 11 Fitur	57
Tabel 4.8 Sebaran Data <i>Training</i> Dan Data <i>Testing</i> Dari Skema 10 Fitur	58
Tabel 4.9 Sebaran Data <i>Training</i> Dan Data <i>Testing</i> Dari Skema 9 Fitur	59
Tabel 4.10 Sebaran Data <i>Training</i> Dan Data <i>Testing</i> Dari Skema 8 Fitur	60
Tabel 4.11 Sebaran Data <i>Training</i> Dan Data <i>Testing</i> Dari Skema 7 Fitur	61
Tabel 4.12 Sebaran Data <i>Training</i> Dan Data <i>Testing</i> Dari Skema 6 Fitur	61
Tabel 4.13 Sebaran Data <i>Training</i> Dan Data <i>Testing</i> Dari Skema 5 Fitur	62
Tabel 4.14 Sebaran Data <i>Training</i> Dan Data <i>Testing</i> Dari Skema 4 Fitur	63
Tabel 4.15 Sebaran Data <i>Training</i> Dan Data <i>Testing</i> Dari Skema 3 Fitur	64
Tabel 4.16 Sebaran Data <i>Training</i> Dan Data <i>Testing</i> Dari Skema 2 Fitur	65
Tabel 4.17 Sebaran Data <i>Training</i> Dan Data <i>Testing</i> Dari Skema 1 Fitur	65
Tabel 4.18 Rincian Data <i>Training</i> Dan Data <i>Testing</i> Dari Setiap Skema	66
Tabel 4.19 <i>Confusion</i> Tabel Data <i>Testing</i> Semua Skema	67

Tabel 4.20 Nilai Akurasi, <i>Precision</i> , <i>Recall</i> dan <i>F1-score</i> Pada Setiap Skema	70
Tabel 4.21 <i>Confusion</i> Tabel Data <i>Training</i> Semua Skema.....	72
Tabel 4.22 Nilai Akurasi, <i>Precision</i> , <i>Recall</i> dan <i>F1-score</i> Pada Setiap Skema	74
Tabel 1. <i>Confusion Matrix</i> Skema 12 Fitur.....	103
Tabel 2. <i>Confusion Matrix</i> Skema 11 Fitur.....	104
Tabel 3. <i>Confusion Matrix</i> Skema 10 Fitur.....	104
Tabel 4. <i>Confusion Matrix</i> Skema 9 Fitur.....	104
Tabel 5. <i>Confusion Matrix</i> Skema 8 Fitur.....	104
Tabel 6. <i>Confusion Matrix</i> Skema 7 Fitur.....	104
Tabel 7. <i>Confusion Matrix</i> Skema 6 Fitur.....	105
Tabel 8. <i>Confusion Matrix</i> Skema 5 Fitur.....	105
Tabel 9. <i>Confusion Matrix</i> Skema 4 Fitur.....	105
Tabel 10. <i>Confusion Matrix</i> Skema 3 Fitur.....	105
Tabel 11. <i>Confusion Matrix</i> Skema 2 Fitur.....	105
Tabel 12. <i>Confusion Matrix</i> Skema 1 Fitur.....	106
Tabel 13 <i>Confusion Matrix</i> Data <i>Testing Fold</i> 10 Dari Skema 12 Fitur.....	106
Tabel 14 <i>Confusion Matrix</i> Data <i>Testing Fold</i> 10 Dari Skema 11 Fitur.....	106
Tabel 15 <i>Confusion Matrix</i> Data <i>Testing Fold</i> 9 & 10 Dari Skema 10 Fitur	106
Tabel 16 <i>Confusion Matrix</i> Data <i>Testing Fold</i> 2 Dari Skema 9 Fitur.....	106
Tabel 17 <i>Confusion Matrix</i> Data <i>Testing Fold</i> 10 Dari Skema 8 Fitur.....	107
Tabel 18 <i>Confusion Matrix</i> Data <i>Testing Fold</i> 10 Dari Skema 7 Fitur.....	107
Tabel 19 <i>Confusion Matrix</i> Data <i>Testing Fold</i> 10 Dari Skema 6 Fitur.....	107
Tabel 20 <i>Confusion Matrix</i> Data <i>Testing Fold</i> 10 Dari Skema 5 Fitur.....	107
Tabel 21 <i>Confusion Matrix</i> Data <i>Testing Fold</i> 7 & 10 Dari Skema 4 Fitur..	107
Tabel 22 <i>Confusion Matrix</i> Data <i>Testing Fold</i> 4 Dari Skema 3 Fitur.....	108
Tabel 23 <i>Confusion Matrix</i> Data <i>Testing Fold</i> 5 Dari Skema 2 Fitur.....	108
Tabel 24 <i>Confusion Matrix</i> Data <i>Testing Fold</i> 4 Dari Skema 1 Fitur.....	108
Tabel 25 <i>Confusion Matrix</i> Data <i>Training</i> Dari Skema 12 Fitur.....	108
Tabel 26 <i>Confusion Matrix</i> Data <i>Training</i> Dari Skema 11 Fitur.....	108
Tabel 27 <i>Confusion Matrix</i> Data <i>Training</i> Dari Skema 10 Fitur.....	109

Tabel 28 <i>Confusion Matrix</i> Data Training Dari Skema 9 Fitur	109
Tabel 29 <i>Confusion Matrix</i> Data Training Dari Skema 8 Fitur	109
Tabel 30 <i>Confusion Matrix</i> Data Training Dari Skema 7 Fitur	109
Tabel 31 <i>Confusion Matrix</i> Data Training Dari Skema 6 Fitur	109
Tabel 32 <i>Confusion Matrix</i> Data Training Dari Skema 5 Fitur	110
Tabel 33 <i>Confusion Matrix</i> Data Training Dari Skema 4 Fitur	110
Tabel 34 <i>Confusion Matrix</i> Data Training Dari Skema 3 Fitur	110
Tabel 35 <i>Confusion Matrix</i> Data Training Dari Skema 2 Fitur	110
Tabel 36 <i>Confusion Matrix</i> Data Training Dari Skema 1 Fitur	110