



**IMPLEMENTASI ALGORITMA K-NEAREST NEIGHBOR UNTUK  
PREDIKSI PASIEN GAGAL JANTUNG**

**SKRIPSI**

**RENDY**

**1810512085**

**UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAKARTA**

**FAKULTAS ILMU KOMPUTER**

**PROGRAM STUDI SISTEM INFORMASI**

**JAKARTA**

**2022**



**IMPLEMENTASI ALGORITMA K-NEAREST NEIGHBOR UNTUK  
PREDIKSI PASIEN GAGAL JANTUNG**

**SKRIPSI**

**Diajukan Sebagai Syarat Untuk Memperoleh Gelar  
Sarjana Komputer**

**RENDY**

**1810512085**

**UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAKARTA**

**FAKULTAS ILMU KOMPUTER**

**PROGRAM STUDI SISTEM INFORMASI**

**JAKARTA**

**2022**

## LEMBAR ORISINALITAS

Skripsi ini adalah hasil karya saya sendiri dan semua sumber yang dikutip maupun yang dirujuk telah saya nyatakan dengan benar:

Nama : Rendy  
NIM : 1810512085  
Program Studi : S1 Sistem Informasi  
Judul : Implementasi Algoritma K-Nearest Neighbor Untuk Prediksi Pasien Gagal Jantung

Bilamana dikemudian hari ditemukan ketidaksesuaian dengan pernyataan saya ini, maka saya bersedia dituntut dan diproses dengan ketentuan yang berlaku.

Jakarta, 3 November 2022

Yang membuat pernyataan,



Rendy

**PERNYATAAN PERSETUJUAN PUBLIKASI SKRIPSI UNTUK  
KEPENTINGAN AKADEMIS**

Sebagai civitas akademik Universitas Pembangunan Nasional Veteran Jakarta, saya yang bertanda tangan dibawah ini:

Nama : Rendy  
NIM : 1810512085  
Fakultas : Ilmu Komputer  
Program Studi : S1 Sistem Informasi

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Pembangunan Nasional Veteran Jakarta Hak Bebas Royalti Non eksklusif (*Non-exclusive Royalty Free Right*) atas karya ilmiah saya yang berjudul:

**Implementasi Algoritma K-Nearest Neighbor Untuk Prediksi Pasien Gagal Jantung**

Beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti ini Universitas Pembangunan Nasional Veteran Jakarta berhak menyimpan, mengalih media/formatkan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan mempublikasikan Skripsi saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta. Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, 3 November 2022

Yang membuat pernyataan,



Rendy

## LEMBAR PENGESAHAN

Dengan ini dinyatakan bahwa Tugas Akhir berikut:

Nama : Rendy

NIM : 1810512085

Program Studi : Sistem Informasi

Judul Tugas Akhir : IMPLEMENTASI ALGORITMA K-NEAREST NEIGHBOR  
UNTUK PREDIKSI PASIEN GAGAL JANTUNG

Telah berhasil dipertahankan di hadapan Tim Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana Komputer pada Program Studi Sistem Informasi, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional Veteran Jakarta.



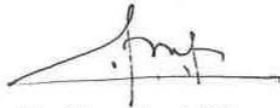
Ati Zaidiah, S.Kom, M.TI.

Penguji I



Sarika M.Kom.

Penguji II



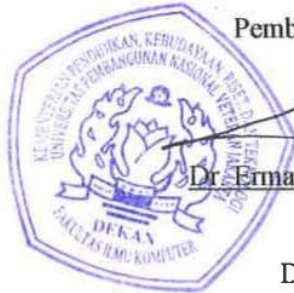
Dr. Ermatita, M.Kom.

Pembimbing I



Ika Nurlaili, S.Kem, M.Sc.

Pembimbing II



Dr. Ermatita, M.Kom.

Dekan



Helena Nurramdhani Irmada, S.Pd.

M.Kom.

Ketua Program Studi

Diterapkan di : Jakarta

Tanggal Ujian : 7 Desember 2022



# IMPLEMENTASI ALGORITMA K-NEAREST NEIGHBOR UNTUK PREDIKSI PASIEN GAGAL JANTUNG

Rendy

## ABSTRAK

Pada tahun 2019, *World Health Organization (WHO)* telah memperkirakan sebesar 17,9 juta orang meninggal diakibatkan oleh penyakit kardiovaskuler atau lebih umumnya gagal jantung. Penelitian ini bertujuan untuk membuat suatu prediksi kematian terjadinya akan akibat gagal jantung. Prediksi ini dilakukan dengan membuat suatu klasifikasi dengan sejumlah kriteria yaitu keadaan tubuh pasien dan penyakit kronis yang pernah dan sedang diderita pasien. Dalam penelitian ini akan digunakan algoritma *K-Nearest Neighbor* dalam *Machine Learning* untuk melakukan suatu klasifikasi. Sampel data yang digunakan didapat dari seorang peneliti yang telah menganalisis data pasien. *Dataset* ini awalnya memiliki sebesar 299 jumlah *record* data, dimana akan dilakukan analisis dengan pembagian sebesar 20% menjadi data uji dan 80% menjadi data latih. Dalam penelitian untuk mempermudah dalam melakukan proses analisis data akan digunakan bantuan bahasa pemrograman *Python* untuk mendapat suatu model prediksi sederhana. Hasil penelitian ini akan menunjukkan tingkat akurasi prediksi, apakah layak untuk digunakan dan kemudian diimplementasi kedalam sebuah sistem sederhana berbasis *website*.

**Kata Kunci:** *Data Mining*, klasifikasi, *K-Nearest Neighbor*, gagal jantung.

# ***K-NEAREST NEIGHBOR ALGORITHM IMPLEMENTATION FOR HEART FAILURE PATIENT PREDICTION***

**Rendy**

## ***ABSTRACT***

*In 2019, the World Health Organization (WHO) has estimated that 17.9 million people died due to cardiovascular disease or more generally heart failure. This study aims to make a prediction of death due to heart failure. This prediction is made by making a classification with a number of criteria, namely the patient's body condition and chronic diseases that the patient has and is currently suffering from. In this research, we will use the K-Nearest Neighbor algorithm in Machine Learning to carry out a classification. The data sample used was obtained from a researcher who had analyzed patient data. This dataset initially has 299 data records, which will be analyzed by dividing 20% into test data and 80% into other data. In research to make it easier to carry out the process of data analysis, the help of the Python programming language will be used to obtain a simple predictive model. The results of this study will show the level of prediction accuracy, whether it is feasible to use and then implement it into a simple website-based system.*

***Keywords:*** *Data Mining, classification, K-Nearest Neighbor, heart failure.*

## **KATA PENGANTAR**

Puji dan syukur penulis panjatkan kepada Tuhan Yesus Kristus karena atas berkat-Nya sehingga penulisan Tugas Akhir ini yang berjudul “Implementasi Algoritma K-Nearest Neighbor Untuk Prediksi Pasien Gagal Jantung” dapat diselesaikan. Selain itu penulis juga ingin mengucapkan rasa terima kasih untuk:

1. Ibu Dr. Ermatita, M.Kom selaku Dekan Fakultas Ilmu Komputer dan juga dosen pembimbing pertama yang membimbing penulis dalam penulisan Tugas Akhir ini dengan baik dan tepat waktu.
2. Ibu Helena Nurramdhani Irmada selaku Ketua Program Studi S1 Sistem Informasi
3. Ibu Ika Nurlaili, S.Kom,M.Sc. selaku dosen pembimbing kedua yang juga telah menuntun dan memberi saran penulisan pada Tugas Akhir ini.
4. Bapak Rio Wirawan, S.Kom., MMSI. selaku dosen pembimbing akademik
5. Ibu dan Bapak Dosen S1 Sistem Informasi UPNVJ atas segala ilmu dan pengajaran yang telah diberikan.
6. Teman-teman Jurusan S1 Sistem Informasi UPNVJ yang telah saling mendukung dan memberikan diskusi.

Jakarta, 3 November 2022

Penulis

Rendy



## DAFTAR ISI

LEMBAR ORISINALITAS .....	i
PERNYATAAN PERSETUJUAN PUBLIKASI SKRIPSI UNTUK KEPENTINGAN AKADEMIS.....	ii
LEMBAR PENGESAHAN .....	iii
ABSTRAK .....	iv
ABSTRACT .....	v
KATA PENGANTAR .....	vi
DAFTAR ISI.....	vii
DAFTAR GAMBAR .....	ix
DAFTAR TABEL.....	x
DAFTAR LAMPIRAN.....	xi
BAB I PENDAHULUAN.....	1
1.1 . Latar Belakang.....	1
1.2 . Rumusan Masalah.....	2
1.3 . Ruang Lingkup Penelitian .....	3
1.4 . Tujuan Penelitian .....	3
1.5 . Manfaat Penelitian .....	3
1.6 . Sistematika Penulisan .....	3
BAB II TINJAUAN PUSTAKA.....	5
2.1 . Gagal Jantung .....	5
2.2 . Data Mining .....	5
2.3 . Metode Klasifikasi .....	10
2.4 . Machine Learning .....	11
2.5 . Algoritma K-Nearest Neighbor .....	12
2.6 . Python .....	13
2.7 . Penelitian Terdahulu .....	13
BAB III METODOLOGI PENELITIAN.....	16
3.1 . Tahapan Penelitian.....	16
3.2 . Pengumpulan Data .....	17
3.3 . Pemahaman Data .....	17

3.4 .	Proses Persiapan Dataset (Praproses Data).....	18
3.5 .	Pengolahan Data .....	20
3.6 .	Hasil dan Evaluasi .....	22
3.7 .	Alat Bantu Penelitian .....	22
3.8 .	Jadwal Penelitian` .....	23
BAB IV HASIL DAN PEMBAHASAN .....		24
4.1 .	Pemahaman Data .....	24
4.2 .	Praproses Data .....	24
4.2.1 .	Pembersihan Data.....	25
4.2.2 .	Pemilihan fitur (variabel) .....	26
4.3 .	Pengolahan Data .....	30
4.4 .	Hasil dan Evaluasi .....	31
4.5 .	Implementasi Algoritma <i>K-Nearest Neighbor</i> dengan <i>Python</i> .....	32
4.5.1 .	Import <i>Library</i> dan Membuka <i>Dataset</i> .....	32
4.5.2 .	Melakukan Pengecekan Data .....	33
4.5.3 .	Pembagian Data .....	34
4.5.4 .	Mencari Nilai K Optimal .....	34
4.5.5 .	Membuat Model Klasifikasi Algoritma <i>K-Nearest Neighbor</i> .....	35
4.5.6 .	Membuat Confusion Matrix.....	35
4.5.7 .	Melakukan Ekspor Model.....	36
4.6 .	Pembuatan Sistem Sederhana .....	36
4.6.1 .	Use Case Diagram.....	36
4.6.2 .	Halaman Utama.....	37
4.6.3 .	Halaman Prediksi .....	39
BAB V PENUTUP.....		40
5.1 .	Kesimpulan .....	40
5.2 .	Saran .....	40
DAFTAR PUSTAKA .....		41
RIWAYAT HIDUP.....		43
LAMPIRAN.....		45

## DAFTAR GAMBAR

Gambar 1 Tahapan Data Mining.....	6
Gambar 2 Kerangka Berpikir Penelitian .....	16
Gambar 3 Missing Value dari Dataset .....	25
Gambar 4 Heatmap untuk korelasi.....	26
Gambar 5 Korelasi .....	27
Gambar 6 Distribusi age berdasarkan DEATH_EVENT.....	27
Gambar 7 Distribusi ejection_fraction berdasarkan DEATH_EVENT .....	28
Gambar 8 Distribusi serum_creatinine berdasarkan DEATH_EVENT.....	28
Gambar 9 Distribusi serum_sodium berdasarkan DEATH_EVENT.....	29
Gambar 10 Distribusi time berdasarkan DEATH_EVENT .....	29
Gambar 11 Import Library .....	32
Gambar 12 Membuka dataset.....	33
Gambar 13 Pengecekan Data .....	33
Gambar 14 Pembagian Data .....	34
Gambar 15 Mencari nilai k .....	34
Gambar 16 Implementasi algoritma k-nearest neighbor.....	35
Gambar 17 Confusion Matrix .....	35
Gambar 18 Ekspor model dengan pickle .....	36
Gambar 19 Use Case Diagram.....	36
Gambar 20 Tampilan Halaman Utama .....	37
Gambar 21 Bagian Prediksi .....	38
Gambar 22 Bagian About Us .....	38
Gambar 23 Tampilan halaman prediksi .....	39

## DAFTAR TABEL

Tabel 1 Confusion Matrix .....	8
Tabel 2 Dataset.....	17
Tabel 3 Variabel .....	18
Tabel 4 Contoh Perhitungan Korelasi .....	19
Tabel 5 Sample Data latih .....	21
Tabel 6 Sampel data uji.....	21
Tabel 7 Hasil perhitungan dan pengurutan .....	22
Tabel 8 Data .....	24
Tabel 9 Jumlah Data.....	30
Tabel 10 5 Nilai k dengan akurasi tertinggi .....	30
Tabel 11 Hasil Perhitungan.....	31
Tabel 12 Confusion Matrix .....	31

## DAFTAR LAMPIRAN

LAMPIRAN 1 Dataset Awal .....	45
LAMPIRAN 2 Source Code App.py .....	53
LAMPIRAN 3 Hasil Turnitin .....	54