

**Komparasi Algoritma KNN Dan Random Forest
untuk Mendiagnosa Penyakit Jantung Koroner
(Studi Kasus: RSUD Dr. Soeselo Slawi)**



SKRIPSI

Diajukan sebagai syarat meraih gelar Sarjana Komputer

Oleh:

Ahmad Miftakhuddin

NIM: 21195001

**PROGRAM STUDI TEKNIK INFORMATIKA
SEKOLAH TINGGI MANAJEMEN INFORMATIKA & KOMPUTER
STMIK YMI TEGAL**

2025

HALAMAN PERSETUJUAN

Pembimbing Skripsi memberikan rekomendasi kepada:

Nama : Ahmad Miftakhuddin

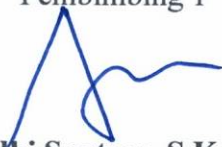
NIM : 21159001

Program Studi : Teknik Informatika

Judul Skripsi : Komparasi Algoritma KNN dan Random Forest untuk Mendiagnosa Penyakit Jantung Koroner (Studi Kasus: RSUD Dr. Soeselo Slawi)

Mahasiswa tersebut telah dinyatakan selesai melaksanakan bimbingan dan dapat mengikuti Ujian Skripsi pada tahun akademik 2024/2025.

Pembimbing 1



Nugroho Adhi Santoso, S.Kom., M.Kom
NIPY.2024.10.008

Tegal, 31 Juli 2025

Pembimbing 2



Bayu Aji Santoso, S.Kom., M.Kom
NIPY.2024.10.019

HALAMAN PENGESAHAN

Nama : Ahmad Miftakhuddin
NIM : 21159001
Program Studi : Teknik Informatika
Judul Skripsi : Komparasi Algoritma KNN Dan Random Forest untuk Mendiagnosa Penyakit Jantung Koroner (Studi Kasus: RSUD Dr. Soeselo Slawi)

Dinyatakan LULUS/~~TIDAK LULUS~~ setelah dipertahankan di hadapan Dewan Penguji Skripsi Program Studi Teknik Informatika/Sistem Informasi STMIK YMI Tegal.

Tegal, 31 Juli 2025

Dewan Penguji :

1. Ketua : Sarif Surejo, S.E., M.Kom

2. Anggota I : Erni Unggul Sedya Utami, S.E., M.Si

3. Anggota II : Zaenul Arif, S.Kom., M.Kom

Tanda Tangan
1. 
2. 
3. 

Mengetahui

Ketua Program Studi Teknik Informatika



Aang Alim Murtopo, S.Kom., M.Kom

NIPY.2024.10.002

LEMBAR PERNYATAAN ORISINALITAS

Saya yang bertanda tangan dibawah ini:

Nama : Ahmad Miftakhuddin
NIM : 21195001
Tempat, Tanggal Lahir : Tegal, 04 April 1990
Alamat : Desa Dukuhwringin Rt.08 Rw.03 Kec. Slawi
Kab.Tegal - Jawa Tengah

Dengan ini menyatakan bahwa skripsi dengan judul **Komparasi Algoritma KNN Dan Random Forest Untuk Mendiagnosa Penyakit Jantung Koroner (Studi Kasus: RSUD Dr. Soeselo Slawi)** adalah hasil pekerjaan saya dan seluruh ide, pendapat atau materi dari sumber lain telah dikutip dengan cara penulisan referensi yang sesuai.

Pernyataan ini saya buat dengan sebenar-benarnya dan jika pernyataan ini tidak sesuai dengan kenyataan, maka saya bersedia menanggung sanksi yang akan dikenakan kepada saya termasuk pencabutan gelar Sarjana Komputer (S.Kom) yang telah saya dapatkan.

Tegal, 31 Juli 2025



Ahmad Miftakhuddin

NIM. 21195001

ABSTRAK

Penyakit jantung koroner (PJK) merupakan salah satu penyebab utama kematian di Indonesia, sehingga deteksi dini menjadi aspek krusial dalam upaya mencegah komplikasi dan meningkatkan efektivitas intervensi medis. Penelitian ini bertujuan untuk melakukan analisis komparatif terhadap dua algoritma *machine learning*, yaitu *K-Nearest Neighbors* (KNN) dan *Random Forest*, dalam klasifikasi diagnosis penyakit jantung koroner berdasarkan data rekam medis pasien RSUD Dr. Soeselo Slawi. Metodologi penelitian meliputi tahapan preprocessing data, transformasi fitur kategorikal, normalisasi, pembagian data dengan rasio 80:20, serta pelatihan dan evaluasi model menggunakan metrik akurasi, *precision*, *recall*, dan *F1-score*. Hasil evaluasi menunjukkan bahwa algoritma *Random Forest* memiliki performa lebih unggul dengan akurasi sebesar 84.00%, *precision* 84.02%, *recall* 84.00%, dan *F1-score* 83.97%. Sementara *K-Nearest Neighbors* (KNN) memperoleh akurasi 82.00% dengan nilai evaluasi yang sedikit lebih rendah. Pengujian tambahan menggunakan *K-Fold Cross Validation* menunjukkan bahwa *Random Forest* lebih stabil terhadap variasi data, dengan akurasi tertinggi 84.80% pada fold ke-4. Penelitian ini juga melakukan pembagian data dengan rasio (60:40, 70:30, dan 90:10). Berdasarkan hasil tersebut, *Random Forest* direkomendasikan sebagai algoritma terbaik untuk implementasi sistem pendukung keputusan medis dalam diagnosis penyakit jantung koroner. Penelitian ini diharapkan dapat berkontribusi dalam pengembangan teknologi kesehatan berbasis kecerdasan buatan yang lebih akurat, efisien, dan dapat diandalkan di lingkungan rumah sakit.

Kata Kunci: Penyakit Jantung Koroner, *Machine Learning*, *K-Nearest Neighbors*, *Random Forest*, Diagnosis Medis.

ABSTRACT

Coronary heart disease (CHD) is a leading cause of death in Indonesia, making early detection crucial for preventing complications and increasing the effectiveness of medical interventions. This study aims to conduct a comparative analysis of two algorithms machine learning, that is K-Nearest Neighbors(KNN) and Random Forest, in the classification of coronary heart disease diagnosis based on patient medical records data from Dr. Soeselo Slawi Regional Hospital. The research methodology includes data preprocessing, categorical feature transformation, normalization, data splitting with an 80:20 ratio, and model training and evaluation using accuracy metrics, precision, recall, And F1-score. The evaluation results show that the Random Forest algorithm has superior performance with an accuracy of 84.00%, precision 84.02%, recall 84.00%, and F1-score 83.97%. Meanwhile K-Nearest Neighbors(KNN) obtained an accuracy of 82.00% with a slightly lower evaluation value. Additional testing using K-Fold Cross Validation show that Random Forest more stable against data variations, with the highest accuracy of 84.80% on the 4th fold. This study also divided the data with ratios (60:40, 70:30, and 90:10). Based on these results, Random Forest is recommended as the best algorithm for implementing a medical decision support system in the diagnosis of coronary heart disease. This research is expected to contribute to the development of more accurate, efficient, and reliable artificial intelligence-based health technology in hospital settings.

Keywords: Coronary Heart Disease, Machine Learning, K-Nearest Neighbors, Random Forest, Medical Diagnosis.

KATA PENGANTAR

Puji dan syukur dipanjatkan ke hadirat Allah SWT yang telah melimpahkan rahmat dan karunia-Nya, akhirnya penulis dapat menyelesaikan skripsi dengan judul “Komparasi Algoritma KNN Dan *Random Forest* Untuk Mendiagnosa Penyakit Jantung Koroner (Studi Kasus : RSUD Dr. Soeselo Slawi)”. Semoga segala usaha dan kerja keras yang telah dilakukan dapat bermanfaat bagi penulis dan juga bagi para pembaca.

Penyusunan skripsi ini tidak lepas dari dukungan berbagai pihak. Oleh karena itu, penulis menyampaikan terima kasih yang sebesar-besarnya kepada:

1. Ketua STMIK YMI Tegal, Bapak Gunawan Adib Ahmadi, S.Pt., M.Pd., yang telah memberikan fasilitas, dukungan, serta suasana akademik yang kondusif sehingga penulis dapat menyelesaikan penyusunan skripsi ini dengan baik.
2. Ketua Program Studi Teknik Informatika, Bapak Aang Alim Murtopo, S.Kom., M.Kom., yang telah memberikan arahan, motivasi, serta kesempatan bagi penulis untuk terus berkembang dalam proses akademik.
3. Dosen Pembimbing I, Bapak Nugroho Adhi Santoso, S.Kom., M.Kom., dan Dosen Pembimbing II, Bapak Bayu Aji Santoso, S.Kom., M.Kom., yang telah dengan sabar membimbing, memberikan masukan yang konstruktif, serta mendampingi penulis selama penyusunan skripsi ini, sehingga dapat terselesaikan dengan baik dan sesuai harapan.
4. Seluruh jajaran RSUD Dr. Soeselo Slawi atas izin dan kerja sama dalam penyediaan data penelitian.
5. Orang tua yang telah memberikan doa, cinta, dan dukungan tanpa henti, serta selalu memberikan semangat dalam setiap langkah yang saya tempuh.
6. Teman-teman seperjuangan yang telah memberikan semangat dan bantuan selama proses ini.

Penulis menyadari bahwa skripsi ini masih jauh dari sempurna. Oleh karena itu, kritik dan saran yang membangun sangat diharapkan demi perbaikan di masa

mendatang. Semoga karya ini dapat memberikan manfaat bagi penulis dan pembaca.

Tegal, 31 Juli 2025

DAFTAR ISI

	Halaman
HALAMAN JUDUL	i
HALAMAN PERSETUJUAN	ii
HALAMAN PENGESAHAN.....	iii
LEMBAR PERNYATAAN ORISINALITAS	iv
ABSTRAK	v
ABSTRACT	vi
KATA PENGANTAR.....	vii
DAFTAR ISI.....	ix
DAFTAR GAMBAR.....	xi
DAFTAR TABEL	xiii
DAFTAR LAMPIRAN	xiv
BAB I PENDAHULUAN.....	1
1.1.Latar Belakang Masalah.....	1
1.2.Rumusan Masalah	3
1.3.Batasan Masalah.....	3
1.4.Tujuan Penelitian	4
1.5.Manfaat Penelitian	4
BAB II LANDASAN TEORI	6
2.1 Landasan Teori.....	6
2.2 Penelitian Terkait	12
BAB III METODOLOGI PENELITIAN	15
3.1 Metode Penelitian.....	15
3.2 Teknik Pengumpulan Data.....	16
3.3 Teknik Analisis Data.....	18
BAB IV HASIL DAN PEMBAHASAN	20
4.1. Hasil Penelitian	20
4.2. Pembahasan Hasil Penelitian	39
BAB V KESIMPULAN DAN SARAN	43

5.1 Kesimpulan	43
5.2 Saran	44
DAFTAR PUSTAKA.....	45

DAFTAR GAMBAR

	Halaman
Gambar 2.1. Ilustrasi <i>Random Forest</i>	8
Gambar 3.1. Tahapan Penelitian.....	15
Gambar 4.1. <i>Source Code Import Library</i>	20
Gambar 4.2. <i>Source Code</i> Pembacaan Dataset.....	21
Gambar 4.3. Hasil Pembacaan Dataset.....	21
Gambar 4.4. <i>Source Code</i> Pembersihan Dataset	22
Gambar 4.5. Hasil Pembersihan Dataset	22
Gambar 4.6. <i>Source Code Cek Missing Values</i>	22
Gambar 4.7. Hasil Cek Missing Values	23
Gambar 4.8. <i>Source Code</i> Transformasi Kategorikal Dataset.....	23
Gambar 4.9. Hasil Transformasi Kategorikal Dataset.....	24
Gambar 4.10. <i>Source Code</i> Kolerasi Antar Fitur	25
Gambar 4.11. Hasil Kolerasi Antar Fitur	25
Gambar 4.12. <i>Source Code</i> Fitur dan Hasil.....	26
Gambar 4.13. <i>Source Code</i> Normalisasi	26
Gambar 4.14. Hasil Normalisasi	27
Gambar 4.15. <i>Source Code</i> Split Data dan Hasil	27
Gambar 4.16. <i>Source Code</i> Model <i>K-Nearest Neighbors</i>	29
Gambar 4.17. <i>Source Code</i> <i>Error Rate</i> Dan Akurasi	29
Gambar 4.18. Hasil <i>Error Rate</i> Dan Akurasi	30
Gambar 4.19. <i>Source Code</i> Evaluasi <i>K-Nearest Neighbors</i>	31
Gambar 4.20. Hasil Evaluasi <i>K-Nearest Neighbors</i>	31
Gambar 4.21. Hasil <i>Confusion Matrix</i> KNN.....	32
Gambar 4.22. <i>Source Code</i> <i>k-fold cross validation</i> KNN	33
Gambar 4.23. Hasil <i>k-fold cross validation</i> KNN	33
Gambar 4.24. <i>Source Code</i> Model <i>Random Forest</i>	34
Gambar 4.25. <i>Source Code</i> Evaluasi <i>Random Forest</i>	35
Gambar 4.26. Hasil Evaluasi Model <i>Random Forest</i>	35

Gambar 4.27. Hasil <i>Confusion Matrix Random Forest</i>	35
Gambar 4.28. <i>Source Code k-fold cross validation Random Forest</i>	36
Gambar 4.29. Hasil <i>K-fold cross validation Random Forest</i>	37
Gambar 4.30. Hasil Perbandingan <i>Random Forest</i> dan KNN(60:40).....	37
Gambar 4.31. Hasil Perbandingan <i>Random Forest</i> dan KNN(70:30).....	37
Gambar 4.32. Hasil Perbandingan <i>Random Forest</i> dan KNN(80:20).....	38
Gambar 4.33. Hasil Perbandingan <i>Random Forest</i> dan KNN(90:10).....	38

DAFTAR TABEL

	Halaman
Tabel 2.1. <i>Confusion Matrix</i>	10
Tabel 3.1. Dataset Pasien	16
Tabel 3.2. Metadata Dataset	18
Tabel 4.1. Perbandingan Hasil Penelitian.....	38
Tabel 4.2. Komparasi Tingkat Akurasi Alogoritma KNN dan <i>Random Forest</i> ...	40

DAFTAR LAMPIRAN

	Halaman
Lampiran 1 Surat Izin Penelitian.....	46