

**ANALISIS PERBANDINGAN MACHINE LEARNING
DENGAN HYPERPARAMETER TUNING TERHADAP
PENYAKIT SERANGAN JANTUNG
SKRIPSI**

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



disusun oleh
ADAM KURNIANTORO
21.11.4170

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2025**

**ANALISIS PERBANDINGAN MACHINE LEARNING
DENGAN HYPERPARAMETER TUNING TERHADAP
PENYAKIT SERANGAN JANTUNG
SKRIPSI**

untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



disusun oleh
ADAM KURNIANTORO
21.11.4170

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2025**

HALAMAN PERSETUJUAN

SKRIPSI

**ANALISIS PERBANDINGAN MACHINE LEARNING DENGAN
HYPERPARAMETER TUNING TERHADAP PENYAKIT SERANGAN
JANTUNG**

yang disusun dan diajukan oleh

Adam Kurniantoro

21.11.4170

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 3 Februari 2025

Dosen Pembimbing,



Kusnawi, S.Kom., M.Eng.

NIK. 190302112

HALAMAN PENGESAHAN

SKRIPSI

ANALISIS PERBANDINGAN MACHINE LEARNING DENGAN
HYPERPARAMETER TUNING TERHADAP PENYAKIT SERANGAN
JANTUNG

yang disusun dan diajukan oleh

Adam Kurniantoro

21.11.4170

Telah dipertahankan di depan Dewan Penguji
pada tanggal 14 Februari 2025

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Dina Maulina, S.Kom., M.Kom.
NIK. 190302250

Ali Mustopa, S.Kom., M.Kom.
NIK. 190302192

Kusnawi, S.Kom., M.Eng.
NIK. 190302112



Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 14 Februari 2025

DEKAN FAKULTAS ILMU KOMPUTER



Hanif Al Fatta, S.Kom., M.Kom., Ph.D.
NIK. 190302096

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Adam Kurniantoro
NIM : 21.11.4170

Menyatakan bahwa Skripsi dengan judul berikut:

Tuliskan Judul Skripsi

Dosen Pembimbing : Kusnawi, S.Kom., M.Eng

1. Karya tulis ini adalah benar-benar **ASLI** dan **BELUM PERNAH** diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan **gagasan, rumusan dan penelitian SAYA** sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab **SAYA**, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini **SAYA** buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka **SAYA** bersedia menerima **SANKSI AKADEMIK** dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 14 Februari 2025

Yang Menyatakan,

A blue ink signature is written over a yellow and red 10,000 Rupiah Indonesian postage stamp. The stamp features the Garuda Pancasila emblem and the text '10000', 'METERAI TEMPEL', and a barcode at the bottom.

Adam Kurniantoro

HALAMAN PERSEMBAHAN

Dengan penuh rasa syukur ke hadirat Allah SWT, atas limpahan rahmat, taufik, dan hidayah-Nya, sehingga skripsi ini dapat terselesaikan dengan baik. Segala puji hanya milik Allah, Dzat yang Maha Pengasih dan Maha Penyayang, yang senantiasa memberikan kekuatan, kesabaran, dan kemudahan dalam setiap langkah kehidupan. Shalawat serta salam senantiasa tercurah kepada Nabi Muhammad SAW, suri teladan bagi seluruh umat manusia, yang membawa cahaya ilmu dan hikmah ke dunia. Dengan segala kerendahan hati, skripsi ini saya persembahkan kepada:

1. Kedua orang tua, Bapak Wiji dan Ibu Dra. Supeniwati dan keluarga tercinta, yang selalu memberikan doa, dukungan, dan kasih sayang tanpa batas.
2. Bapak Kusnawi, S.Kom., M.Eng., selaku dosen pembimbing yang telah memberikan arahan, motivasi, dan ilmu yang sangat berharga.
3. Pasangan tercinta, Arimbi, yang selalu memberikan motivasi, kesabaran, dan dukungan tanpa henti dalam setiap langkah perjalanan akademik ini.
4. Teman-teman seperjuangan, Kresna, Dafi, Novanda, Satria, Tama, Tegar, dan Arya yang selalu mendukung dan berbagi pengalaman dalam Penulisan skripsi ini.
5. Almamater Universitas Amikom Yogyakarta, sebagai tempat menimba ilmu dan mengembangkan diri.

Semoga karya ini dapat memberikan manfaat dan menjadi langkah awal dalam perjalanan akademik dan profesional saya.

KATA PENGANTAR

Puji syukur kehadiran Allah SWT atas rahmat dan hidayah-Nya sehingga penulis dapat menyelesaikan skripsi ini dengan baik. Skripsi ini disusun sebagai salah satu syarat untuk menyelesaikan studi pada Program Sarjana (S1) Informatika di Universitas Amikom Yogyakarta.

Dalam penyusunan skripsi ini, penulis mendapatkan banyak dukungan, bimbingan, serta bantuan dari berbagai pihak. Oleh karena itu, dengan segala kerendahan hati, penulis ingin mengucapkan terima kasih yang sebesar-besarnya kepada:

1. Kedua orang tua, Bapak Wiji dan Ibu Dra. Supeniwati dan keluarga tercinta yang selalu memberikan doa, dukungan, serta semangat yang tiada henti.
2. Prof. Dr. M.Suyanto, MM., Selaku Rektor Universitas Amikom Yogyakarta.
3. Bapak Kusnawi, S.Kom., M.Eng., selaku dosen pembimbing yang telah memberikan arahan, motivasi, dan ilmu yang sangat berharga dalam penyelesaian skripsi ini.
4. Tim Dosen Penguji ibu Dina Maulina, S.Kom., M.Kom. dan Bapak Ali Mustopa, S.Kom., M.Kom. yang telah memberikan masukan serta kritik membangun untuk perbaikan skripsi ini.
5. Bapak Hanif Al Fatta, S.Kom., M.Kom., Ph.D., Selaku Dekan Fakultas Ilmu Komputer Universitas Amikom Yogyakarta.
6. Seluruh dosen dan staf akademik di Program Studi Informatika Universitas Amikom Yogyakarta yang telah memberikan ilmu dan fasilitas selama masa studi.
7. Rekan-rekan mahasiswa yang telah memberikan dukungan dan berbagi ilmu selama proses penyusunan skripsi ini.

Yogyakarta, 14 Februari 2025

Penulis

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERSETUJUAN	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERNYATAAN KEASLIAN SKRIPSI	iv
HALAMAN PERSEMBAHAN	v
KATA PENGANTAR	vi
DAFTAR ISI	vii
DAFTAR TABEL	ix
DAFTAR GAMBAR	x
INTISARI	xi
ABSTRACT	xii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Batasan Masalah	3
1.4 Tujuan Penelitian	4
1.5 Manfaat Penelitian	4
1.6 Sistematika Penulisan	5
BAB II TINJAUAN PUSTAKA	7
2.1 Studi Literatur	7
2.2 Dasar Teori	18
2.2.1 Data Mining	18
2.2.2 Machine Learning	18
2.2.3 Logistic Regression	19
2.2.4 Random Forest	20
2.2.5 Support Vector Machine (SVM)	21
2.2.6 Gradient Boosted Trees (GBT)	23

2.2.7	Pre-Processing.....	23
2.2.8	Evaluasi.....	25
2.2.9	Hyperparameter Tuning.....	26
BAB III METODE PENELITIAN		28
3.1	Objek Penelitian	28
3.2	Alur Penelitian.....	28
3.3	Alat dan Bahan.....	33
3.3.1	Data Penelitian	33
3.3.2	Alat/instrumen	33
BAB IV HASIL DAN PEMBAHASAN		35
4.1	Preprocessing	35
4.1.1	Deteksi Outlier.....	35
4.1.2	Encoding.....	39
4.2	Pemilihan Fitur.....	41
4.3	Normalisasi Data dan Undersampling.....	44
4.4	Pembagian Data.....	47
4.5	Pemodelan.....	49
4.6	Evaluasi Model.....	52
4.7	Hyperparameter Tuning.....	60
BAB V PENUTUP		72
5.1	Kesimpulan.....	72
5.2	Saran.....	72
REFERENSI		74
LAMPIRAN.....		80

DAFTAR TABEL

Tabel 2.1 Keaslian Penelitian.....	9
Tabel 2.2 Confusion Matrix.....	25
Tabel 4.1 Deteksi Outlier.....	35
Tabel 4.2 Drop Outlier.....	38
Tabel 4.3 Encoding.....	39
Tabel 4.4 Visualisasi Korelasi.....	41
Tabel 4.5 Penggunaan ANOVA.....	43
Tabel 4.6 Hasil Seleksi Fitur.....	43
Tabel 4.7 Normalisasi Data.....	44
Tabel 4.8 Under Sampling.....	45
Tabel 4.9 Visualisasi Under Sampling.....	46
Tabel 4.10 Pembagian Data.....	47
Tabel 4.11 Logistic Regression.....	49
Tabel 4.12 Random Forest.....	50
Tabel 4.13 SVM.....	50
Tabel 4.14 Gradient Boosted Tree.....	51
Tabel 4.15 Evaluasi Model Logistic Regression.....	52
Tabel 4.16 Evaluasi Model Random Forest.....	54
Tabel 4.17 Evaluasi Model SVM.....	56
Tabel 4.18 Evaluasi Model GBT.....	58
Tabel 4.19 Evaluasi Model Logistic Regression.....	60
Tabel 4.20 Evaluasi Model Random Forest.....	62
Tabel 4.21 Evaluasi Model SVM.....	65
Tabel 4.22 Evaluasi Model Gradient Boosted Tree.....	67
Tabel 4.23 Perbandingan Evaluasi Model Sebelum Tuning.....	70
Tabel 4.24 Perbandingan Evaluasi Model Setelah Tuning.....	70

DAFTAR GAMBAR

Gambar 2.1 Kurva Logistic Regression	20
Gambar 2.2 Ilustrasi Random Forest[5].....	21
Gambar 2.3 SVM menemukan hyperplane terbaik antara kelas -1 dan +1[24].....	22
Gambar 2.4 Arsitektur Gradient Boosted Trees[27].....	23
Gambar 3.1 Alur Penelitian	28
Gambar 3.2 Informasi Dataset	30
Gambar 4.1 Outlier Numerik Data.....	37
Gambar 4.2 Outlier Pada WeightInKilograms.....	37
Gambar 4.3 Persebaran Data Normal	39
Gambar 4.4 Data Ter-Encoding.....	41
Gambar 4.5 Korelasi Fitur	42
Gambar 4.6 Persebaran Data Target Sebelum Balancing	45
Gambar 4.7 Hasil Undersampling.....	46
Gambar 4.8 Persebaran Data Setelah Undersampling	47
Gambar 4.9 Hasil Pembagian Data.....	49
Gambar 4.10 Model Logistic Regression	50
Gambar 4.11 Model Random Forest.....	50
Gambar 4.12 Model SVM.....	51
Gambar 4.13 Model Gradient Boosted Tree	51
Gambar 4.14 Confusion Matrix Logistic Regression	53
Gambar 4.15 Confusion Matrix Random Forest.....	55
Gambar 4.16 Confusion Matrix SVM	57
Gambar 4.17 Confusion Matrix GBT	59
Gambar 4.18 Parameter LR Terbaik.....	61
Gambar 4.19 Confusion Matrix LR Tuning	62
Gambar 4.20 Parameter RF Terbaik	64
Gambar 4.21 Confusion Matrix RF Tuning.....	64
Gambar 4.22 Parameter SVM Terbaik	66
Gambar 4.23 Confusion Matrix SVM Tuning	67
Gambar 4.24 Parameter GBT Terbaik	69
Gambar 4.25 Confusion Matrix GBT Tuning.....	69

INTISARI

Penyakit serangan jantung termasuk dalam penyakit yang mematikan, di dunia terjadi jutaan kasus dan setiap tahunnya semakin bertambah, sehingga diperlukan analisis yang akurat untuk mendeteksi faktor risiko dan meningkatkan prediksi diagnosis. Penelitian ini membandingkan performa beberapa model machine learning, yaitu Gradient Boosted Tree, Logistic Regression, Random Forest, dan Support Vector Machine (SVM), dalam menganalisis data penyakit serangan jantung dengan dan tanpa hyperparameter tuning. Hasil penelitian menunjukkan bahwa tiga fitur utama yang paling berpengaruh dalam analisis adalah HadAngina (44%), HadStroke (17,7%), dan AgeCategory (17,2%). Dari segi performa, model Gradient Boosted Tree memiliki hasil terbaik dengan akurasi, presisi, F1-score, dan recall sebesar 81%, baik sebelum maupun setelah tuning. Sementara itu, model Logistic Regression, Random Forest, dan SVM memiliki performa sebesar 80% setelah tuning. Hyperparameter tuning tidak memberikan peningkatan signifikan terhadap performa model secara keseluruhan, dengan peningkatan maksimal sebesar 2% pada model Random Forest. Temuan ini menunjukkan bahwa pemilihan model yang tepat lebih berpengaruh dibandingkan tuning parameter dalam analisis penyakit serangan jantung.

Kata Kunci: Penyakit Serangan Jantung, Machine Learning, Hyperparameter Tuning, Gradient Boosted Tree

ABSTRACT

Heart disease is one of the leading causes of death worldwide, with millions of cases occurring each year and increasing annually. Therefore, accurate analysis is required to detect risk factors and improve diagnostic prediction. This study compares the performance of several machine learning models, including Gradient Boosted Tree, Logistic Regression, Random Forest, and Support Vector Machine (SVM), in analyzing heart disease data with and without hyperparameter tuning. The results show that the three most influential features in the analysis are HadAngina (44%), HadStroke (17.7%), and AgeCategory (17.2%). In terms of performance, the Gradient Boosted Tree model achieved the best results with an accuracy, precision, F1-score, and recall of 81%, both before and after tuning. Meanwhile, the Logistic Regression, Random Forest, and SVM models each achieved 80% performance after tuning. Hyperparameter tuning did not significantly improve the overall model performance, with a maximum increase of only 2% in the Random Forest model. These findings indicate that selecting the right model has a greater impact than hyperparameter tuning in heart disease analysis.

Keyword: Heart Disease, Machine Learning, Hyperparameter Tuning, Gradient Boosted Tree