

**ANALISIS KINERJA ALGORITMA SUPPORT VECTOR MACHINE
DAN GRADIENT BOOSTING MACHINE UNTUK PREDIKSI
PENYAKIT JANTUNG**

SKRIPSI

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



disusun oleh

TEGAR WIRAWAN

21.11.3838

Kepada

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

**ANALISIS KINERJA ALGORITMA SUPPORT VECTOR MACHINE
DAN GRADIENT BOOSTING MACHINE UNTUK PREDIKSI
PENYAKIT JANTUNG**

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



disusun oleh
TEGAR WIRAWAN
21.11.3838

Kepada

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

HALAMAN PERSETUJUAN

SKRIPSI

**Analisis Kinerja Algoritma Support Vector Machine dan
Gradient Boosting Machine untuk Prediksi Penyakit Jantung**

yang disusun dan diajukan oleh

TEGAR WIRAWAN

21.11.3838

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 23 Januari 2025

Dosen Pembimbing,



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

HALAMAN PENGESAHAN

SKRIPSI

**Analisis Kinerja Algoritma Support Vector Machine dan
Gradient Boosting Machine untuk Prediksi Penyakit Jantung**

yang disusun dan diajukan oleh

TEGAR WIRAWAN

21.11.3838

Telah dipertahankan di depan Dewan Penguji
pada tanggal 23 Januari 2025

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Yudi Sutanto, M.Kom
NIK. 190302039

Rizqi Sukma Kharisma, S.Kom., M.Kom.
NIK. 190302215

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



Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 23 Januari 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 : TEGAR WIRAWAN

NIM : 21.11.3838

Menyatakan bahwa Skripsi dengan judul berikut:

Analisis Kinerja Algoritma Support Vector Machine dan Gradient Boosting Machine untuk Prediksi Penyakit Jantung

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, 23 Januari 2025

Yang Menyatakan,



Stamp: AMIKOM YOGYAKARTA, KEPALA TEMPIL, DETALX410759154

Tegar Wirawan

HALAMAN PERSEMBAHAN

Skripsi ini saya persembahkan dengan rasa syukur dan penuh rasa hormat kepada:

1. Kedua Orang Tua Tercinta,
Yang senantiasa selalu mendoakan, memberi dukungan semangat, dan pengorbanan yang tanpa batas, baik moral ataupun material.
2. Dosen Pembimbing.
Bapak Kusnawi, S.Kom., M.Eng. atas bimbinganya, ilmunya, serta arahan selama proses penyusunan skripsi ini.
3. Teman-teman Seperjuangan
Yang telah memberi inspirasi, semangat, motifasi, kebersamaan, serta semangat dalam menyelesaikan skripsi ini.
4. Universitas Amikom Yogyakarta
Tempat saya menimba ilmu dan menempa diri untuk masa depan.

KATA PENGANTAR

Puja dan Puji Syukur penulis panjatkan ke hadirat Allah SWT, yang telah memberikan hidayah, rahmat, serta inayah-Nya, sehingga penulis mampu menyelesaikan skripsi ini dengan judul 'Analisis Kinerja Algoritma Support Vector Machine dan Gradient Boosting Machine untuk Prediksi Penyakit Jantung'.

Skripsi ini disusun untuk dapat memenuhi salah satu syarat untuk memperoleh gelar Sarjana pada Program Studi Informatika, Fakultas Ilmu Komputer, Universitas Amikom Yogyakarta

Dalam skripsi ini, penulis menyadari bahwa keberhasilan ini tidak lepas dari bantuan, dukungan, dan doa dari berbagai pihak. Oleh karena itu, penulis menyampaikan ucapan terimakasih yang sebesar-besarnya kepada:

1. Hanif Al Fatta, S.Kom., M.Kom., Ph.D., selaku dekan Fakultas Ilmu Komputer Universitas Amikom Yogyakarta.
2. Kusnawi, S.Kom., M.Eng., selaku dosen pembimbing yang telah banyak memberikan arahan, bimbingan, serta masukan yang sangat berharga selama proses penyusunan skripsi ini.
3. Tim Dosen Penguji, yang telah memberikan kritik dan saran untuk memperbaiki hasil dari penelitian ini.
4. Kedua Orang Tua Tercinta, yang selalu memberikan doa, dukungan serta semangat tanpa henti.
5. Rekan-rekan seperjuangan, yang telah menjadi teman diskusi, berbagi pengalaman, dan saling mendukung selama proses belajar.
6. Semua pihak yang tidak dapat penulis sebutkan satu per satu yang telah memberikan dukungan dan semangat dalam menyelesaikan skripsi ini.

Penulis menyadari bahwa skripsi ini memiliki kekurangan. Oleh karena itu, kritik dan saran yang membangun sangat penulis harapkan untuk perbaikan di masa mendatang. Akhir kata, semoga semoga skripsi ini dapat memberi manfaat bagi para pembaca serta bisa menjadi inspirasi untuk penelitian selanjutnya.

Yogyakarta, 23 Januari 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 TABEL.....	ix
DAFTAR GAMBAR.....	x
INTISARI	xi
ABSTRACT.....	xii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	2
1.3 Batasan Masalah	3
1.4 Tujuan Penelitian	3
1.5 Manfaat Penelitian	3
1.6 Sistematika Penulisan	4
BAB II TINJAUAN PUSTAKA	6
2.1 Studi Literatur	6
2.2 Dasar Teori.....	16
2.2.1 Data Mining	16
2.2.2 Machine Learning	16
2.2.2.1 Tahapan Machine Learning	16
2.2.3 Support Vector Machine	17
2.2.4 Gradient Boosting Machine	18
2.2.5 Python	19

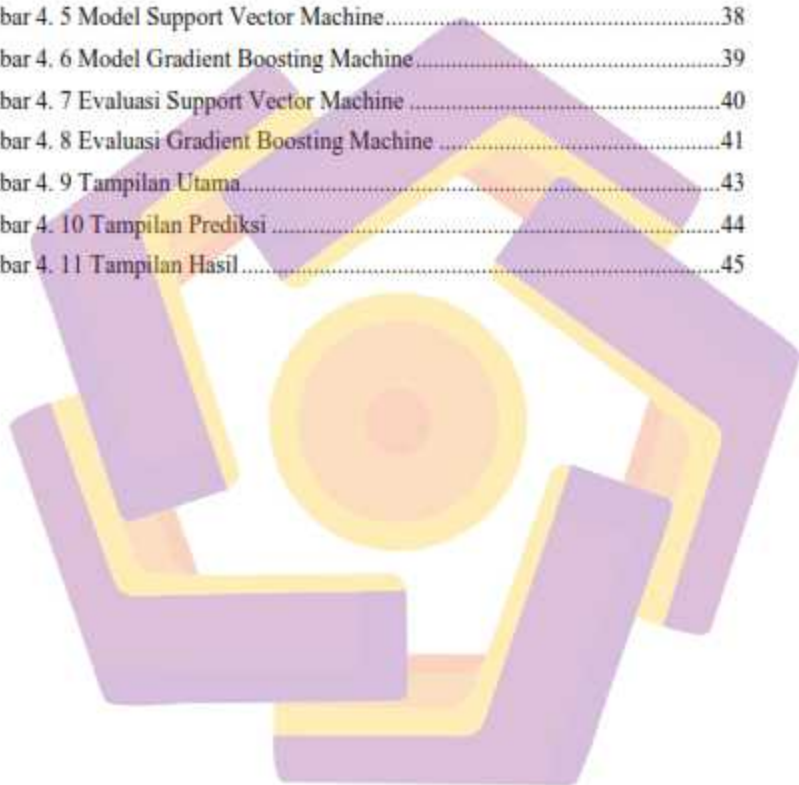
2.2.6 Streamlit.....	19
2.2.7 Preprocessing.....	19
2.2.8 Confusion Matrix.....	20
BAB III METODE PENELITIAN	22
3.1 Objek Penelitian.....	22
3.2 Alur Penelitian.....	22
3.3 Alat dan Bahan.....	29
3.3.2 Alat/instrumen.....	30
4.1 Pengumpulan Data.....	31
4.2 Preprocessing.....	32
4.3 Model.....	37
4.4 Evaluasi.....	39
4.5 Deploymen.....	43
BAB V PENUTUP	46
5.1 Kesimpulan.....	46
5.2 Saran.....	46
REFERENSI	47
LAMPIRAN	50

DAFTAR TABEL

Tabel 2. 1 Keaslian PenelitianTabel	9
Tabel 2. 2 Confussion Matrix	20
Tabel 3. 1 Dataset	24
Tabel 3. 2 Informasi Dataset	25
Tabel 3. 3 Describe Dataset	26
Tabel 3. 4 Missing Value Dataset	27
Tabel 3. 5 Korelasi Dataset	27
Tabel 4. 1 Pngumpulan Data	31
Tabel 4. 2 Label Encoding	32
Tabel 4. 3 Correlation Matrix	33
Tabel 4. 4 Correlation Thresholding	35
Tabel 4. 5 Hasil Correlation Thresholding	35
Tabel 4. 6 Smote	36
Tabel 4. 7 Splite Data	37
Tabel 4. 8 Model Support Vector Machine	38
Tabel 4. 9 Model Gradient Boosting Machine	39
Tabel 4. 10 Evaluasi Support Vector Machine	39
Tabel 4. 11 Evaluasi Gradient Boosting Machine	41
Tabel 4. 12 Perbandingan Model	42

DAFTAR GAMBAR

Gambar 3. 1 Alur penelitian.....	22
Gambar 3. 2 Distribusi Variabel Target.....	28
Gambar 4. 1 Import Data	31
Gambar 4. 2 Label Encoding	33
Gambar 4. 3 Correlation Matrix.....	34
Gambar 4. 4 SMOTE	37
Gambar 4. 5 Model Support Vector Machine.....	38
Gambar 4. 6 Model Gradient Boosting Machine.....	39
Gambar 4. 7 Evaluasi Support Vector Machine	40
Gambar 4. 8 Evaluasi Gradient Boosting Machine	41
Gambar 4. 9 Tampilan Utama.....	43
Gambar 4. 10 Tampilan Prediksi	44
Gambar 4. 11 Tampilan Hasil.....	45



INTISARI

Penyakit jantung merupakan salah satu penyebab utama kematian di dunia, dengan angka kematian yang terus menerus meningkat disetiap tahunnya. Identifikasi terhadap risiko penyakit jantung menjadi penting untuk meningkatkan efektivitas pencegahan maupun pengobatan. Penelitian ini memiliki tujuan untuk membandingkan kinerja dua algoritma yaitu, *Support Vector Machine* dan *Gradient Boosting Machine* untuk prediksi penyakit jantung berdasarkan dataset yang bersumber dari *Kaggle*

Penelitian ini dimulai dari pengumpulan data, dilanjutkan dengan eskplorasi data, *preprocessing* dengan *label encoding*, *SMOTE* untuk menangani ketidak seimbangan antar kelas, serta normalisasi data dengan *StandardScaler*. Seleksi fitur dilakukan dengan menggunakan metode *korelasi thresholding*. Data kemudian dibagi menjadi dua data latih dan data uji untuk membangun model prediktif. Hasil penelitian membandingkan berdasarkan metrix evaluasi seperti *akurasi*, *presisi*, *recall*, dan *f-1 score*.

Hasil dari penelitian ini menunjukkan bahwa model *Gradient Boosting Machine* lebih unggul dengan akurasi yang didapat 98% sedangkan *Support Vector Machine* yang hanya mendapatkan akurasi 93%. Penelitian ini diharapkan dapat berkontribusi dalam praktik Kesehatan untuk mendeteksi risiko penyakit jantung lebih dini. Penelitian selanjutnya untuk menguji dengan algoritma lain ataupun menggunakan dataset yang lebih beragam untuk mendapatkan hasil yang lebih bagus.

Kata kunci: penyakit jantung, *Support Vector Machine*, *Gradient Boosting Machine*, Prediksi, *Machine Learning*.

ABSTRACT

Heart disease is one of the leading causes of death worldwide, with mortality rates increasing annually. Identifying the risk of heart disease is essential to improving the effectiveness of prevention and treatment efforts. This study aims to compare the performance of two machine learning algorithms, Support Vector Machine (SVM) and Gradient Boosting Machine (GBM), for predicting heart disease using a dataset sourced from Kaggle.

The study begins with data collection, followed by data exploration, preprocessing using label encoding, SMOTE to address class imbalance, and data normalization with StandardScaler. Feature selection was conducted using Correlation Thresholding. The data was then split into training and testing sets to build predictive models. The results were compared based on evaluation metrics such as accuracy, precision, recall, and F1-Score.

The findings indicate that the Gradient Boosting Machine outperformed the Support Vector Machine, achieving an accuracy of 98% compared to SVM's accuracy of 93%. This research is expected to contribute to healthcare practices by enabling early detection of heart disease risks. Future research is recommended to explore other algorithms or employ more diverse datasets to achieve better results.

Keyword: heart disease, Support Vector Machine, Gradient Boosting Machine, Prediction, machine learning.