



UNIVERSITAS  
MERCU BUANA

**PERBANDINGAN PERFORMA ALGORITMA XGBOOST, CATBOOST  
DAN GBM DALAM PREDIKSI PENYAKIT KARDIOVASKULAR**

LAPORAN TUGAS AKHIR



PANWASTO SAMOSIR P  
41521010178

UNIVERSITAS  
MERCU BUANA

**PROGRAM STUDI TEKNIK INFORMATIKA  
FAKULTAS ILMU KOMPUTER  
UNIVERSITAS MERCU BUANA  
JAKARTA  
2024**



**PERBANDINGAN PERFORMA ALGORITMA XGBOOST, CATBOOST  
DAN GBM DALAM PREDIKSI PENYAKIT KARDIOVASKULAR**

LAPORAN TUGAS AKHIR

PANWASTO SAMOSIR P  
41521010178

UNIVERSITAS  
MERCU BUANA

**PROGRAM STUDI TEKNIK INFORMATIKA  
FAKULTAS ILMU KOMPUTER  
UNIVERSITAS MERCU BUANA  
JAKARTA  
2024**

## HALAMAN PERNYATAAN KARYA SENDIRI

Saya yang bertanda tangan di bawah ini:

Nama Mahasiswa : Panwasto Samosir P  
NIM : 41521010178  
Program Studi : Teknik Informatika  
Judul Laporan Skripsi : Perbandingan Performa Algoritma XGBoost, CatBoost Dan GBM Dalam Prediksi Penyakit Kardiovaskular

Menyatakan bahwa Laporan Skripsi ini adalah hasil karya saya sendiri dan bukan plagiat, serta semua sumber baik yang dikutip maupun dirujuk telah saya nyatakan dengan benar. Apabila ternyata ditemukan di dalam Laporan Skripsi saya terdapat unsur plagiat, maka saya siap mendapatkan sanksi akademis yang berlaku di Universitas Mercu Buana.



Jakarta, 17 Desember 2024



Panwasto Samosir P

UNIVERSITAS  
MERCU BUANA

## HALAMAN PENGESAHAN




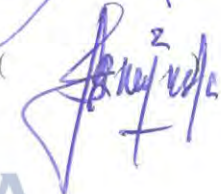
Laporan Skripsi ini diajukan oleh:

Nama : Panwasto Samosri P  
NIM : 41521010178  
Program Studi : Teknik Informatika  
Judul Laporan Skripsi : Perbandingan Performa Algoritma XGBoost, CatBoost dan GBM dalam Prediksi Penyakit Kardiovaskular

Telah berhasil dipertahankan pada sidang di hadapan Dewan Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana Strata 1 pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer Universitas Mercu Buana.

Disahkan oleh:

Pembimbing : Umniy Salamah, S.T., MMSI  
NIDN : 0306098104  
Ketua Penguji : Wawan Gunawan, S.Kom., M.T  
NIDN : 0424108104  
Penguji 1 : Mohamad Yusuf, S.Kom., MCS  
NIDN : 0307097606  
Penguji 2 : Muhammad Rifqi, S.Kom.,  
M.Kom  
NIDN : 0301067101

()  
()  
()  
()

UNIVERSITAS  
MERCU BUANA

Jakarta, 14 Januari 2025

Mengetahui,

Dekan



Dr. Bambang Jokonowo, S.Si., MTI  
NIDN : 0320037002

Ketua Program Studi



Dr. Hadi Santoso, S.Kom., M.Kom  
NIDN : 0225067701

## KATA PENGANTAR

Puji syukur kehadiran Tuhan yang Maha Esa, atas segala rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan tugas akhir yang merupakan salah satu persyaratan kelulusan Program Studi Strata Satu (S1) pada jurusan Teknik Informatika, Universitas Mercu Buana.

Penulis menyadari bahwa tugas akhir ini masih jauh dari sempurna, karena kesempurnaan sejatinya hanya milik Tuhan yang Maha Esa. Oleh karena itu, saran dan masukan yang membangun senantiasa penulis terima dengan senang hati. Serta berkat dukungan, motivasi, bantuan, bimbingan, dan doa dari banyak pihak, penulis mengucapkan terima kasih kepada:

1. Bapak Prof. Dr. Andi Adriansyah, M. Eng selaku Rektor Universitas Mercu Buana
2. Bapak Dr. Bambang Jokonowo, S.Si., M.T.I selaku Dekan Fakultas Ilmu Komputer.
3. Bapak Dr. Hadi Santoso, S.Kom, M.Kom selaku Ketua Program Studi Teknik Informatika Universitas Mercu Buana.
4. Ibu Umniy Salamah, S.Kom, MMSI selaku Dosen Pembimbing tugas akhir yang telah memberikan pengarahan, motivasi, menyediakan waktu, tenaga, dan pikiran sehingga selama pembuatan tugas akhir ini terjadwal dengan baik.
5. Kedua Orang Tua dan saudara-saudara saya yang selalu mensupport dan mendukung saya selama menjalani masa studi sebagai mahasiswa Universitas Mercubuana.
6. Semua teman kuliah yang selalu berbagi informasi dan memberikan dukungan dalam bentuk yang berbeda-beda.

Akhir kata, penulis berharap Tuhan Yang Maha Esa berkenan membalas segala kebaikan semua pihak yang telah membantu. Semoga Tugas Akhir ini membawa manfaat bagi pengembangan ilmu.

Jakarta, 10 Desember 2024



Panwasto Samosir P

**HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS  
AKHIR UNTUK KEPENTINGAN AKADEMIS**

Sebagai sivitas akademik Universitas Mercu Buana, saya yang bertanda tangan di bawah ini:

Nama : Panwasto Samosir P  
NIM : 41521010178  
Program Studi : Teknik Informatika  
Judul Laporan Skripsi : Perbandingan Performa Algoritma XGBoost, CatBoost dan GBM dalam Prediksi Penyakit Kardiovaskular

Demi pengembangan ilmu pengetahuan, dengan ini memberikan izin dan menyetujui untuk memberikan kepada Universitas Mercu Buana **Hak Bebas Royalti Non-Eksklusif (*Non-exclusive Royalti-Free Right*)** atas karya ilmiah saya yang berjudul di atas beserta perangkat yang ada (jika diperlukan).

Dengan Hak Bebas Royalti Non-Eksklusif ini Universitas Mercu Buana berhak menyimpan, mengalihmedia/format-kan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan mempublikasikan Laporan Skripsi saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

UNIVERSITAS  
MERCU BUANA

Jakarta, 10 Desember 2024

Yang menyatakan,



Panwasto Samosir P

## ABSTRAK

Nama : Panwasto Samosir P  
NIM : 41521010178  
Program Studi : Teknik Informatika  
Judul Laporan Penelitian : Perbandingan Performa Algoritma XGBoost  
CatBoost dan GBM dalam Prediksi Penyakit  
Kardiovaskular  
Pembimbing : Umniy Salamah, S.Kom, MMSI

Penyakit kardiovaskular merupakan penyebab utama kematian di seluruh dunia, melibatkan gangguan pada jantung dan pembuluh darah, seperti penyakit jantung koroner dan hipertensi. Faktor risiko penyakit kardiovaskular mencakup gaya hidup tidak sehat serta faktor non-modifikasi seperti usia dan riwayat keluarga. Untuk mengatasi tantangan dalam deteksi dini dan prediksi penyakit kardiovaskular, pendekatan machine learning, khususnya algoritma boosting, telah menunjukkan potensi yang signifikan. Penelitian ini bertujuan untuk membandingkan kinerja tiga algoritma boosting utama, yaitu XGBoost, CatBoost, dan Gradient Boosting, dalam memprediksi risiko penyakit kardiovaskular menggunakan dataset yang tersedia secara online. Hasil penelitian menunjukkan bahwa CatBoost memiliki performa terbaik dengan akurasi sebesar 75%, Precision 0.83, dan ROC AUC 0.81, yang mengindikasikan kemampuannya dalam menghasilkan prediksi yang lebih akurat. Gradient Boosting memiliki akurasi 70% dan menunjukkan keseimbangan yang baik antara Recall dan Precision, sementara XGBoost memiliki kinerja terendah dengan akurasi 63.3% di semua metrik evaluasi. Berdasarkan hasil ini, CatBoost adalah model yang paling efektif untuk memprediksi risiko penyakit kardiovaskular.

**Kata Kunci:** Kardiovaskular, Boosting Algorithms, XGBoost, CatBoost, Gradient Boosting Machine (GBM)

## ABSTRACT

Name : Panwasto Samosir P  
NIM : 41521010178  
Study Program : Teknik Informatika  
Thesis Title : Comparison of XGBoost, CatBoost, and GBM  
Algorithm Performance in Predicting  
Cardiovascular Disease  
Counsellor : Umniy Salamah, S.Kom, MMSI

*Cardiovascular disease is the leading cause of death worldwide, involving disorders of the heart and blood vessels, such as coronary artery disease and hypertension. Risk factors for cardiovascular disease include unhealthy lifestyle choices as well as non-modifiable factors such as age and family history. To address the challenges in early detection and prediction of cardiovascular disease, machine learning approaches, particularly boosting algorithms, have shown significant potential. This study aims to compare the performance of three main boosting algorithms: XGBoost, CatBoost, and Gradient Boosting, in predicting the risk of cardiovascular disease using publicly available datasets. The results indicate that CatBoost outperforms the other models with an accuracy of 75%, Precision of 0.83, and ROC AUC of 0.81, demonstrating its superior ability to generate accurate predictions. Gradient Boosting achieves an accuracy of 70% and demonstrates a good balance between Recall and Precision, while XGBoost has the lowest performance with an accuracy of 63.3% across all evaluation metrics. Based on these findings, CatBoost is the most effective model for predicting the risk of cardiovascular disease.*

**Keywords:** Cardiovascular Diseases, Boosting Algorithms, XGBoost, CatBoost, Gradient Boosting Machine (GBM)



## DAFTAR ISI

<b>HALAMAN JUDUL</b> .....	i
<b>HALAMAN PERNYATAAN KARYA SENDIRI</b> .....	ii
<b>HALAMAN PENGESAHAN</b> .....	iii
<b>KATA PENGANTAR</b> .....	iv
<b>HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI</b> .....	v
<b>ABSTRAK</b> .....	vi
<b>ABSTRACT</b> .....	vii
<b>DAFTAR ISI</b> .....	viii
<b>DAFTAR TABEL</b> .....	x
<b>DAFTAR GAMBAR</b> .....	xi
<b>DAFTAR LAMPIRAN</b> .....	xii
<b>BAB I PENDAHULUAN</b> .....	1
1.1 Latar Belakang .....	1
1.2 Perumusan Masalah .....	3
1.3 Tujuan Penelitian .....	3
1.4 Manfaat Penelitian .....	3
1.5 Batasan Masalah .....	4
<b>BAB II TINJAUAN PUSTAKA</b> .....	5
2.1 Penelitian Terdahulu .....	5
2.2 Teori Pendukung .....	13
2.2.1 Boosting .....	13
2.2.2 Algoritma XGBoost .....	14
2.2.3 Algoritma CatBoost .....	14
2.2.4 Algoritma Gradien Boosting Machine .....	15
2.2.5 Decision Trees .....	16
2.2.6 Penyakit Kardiovaskular .....	17
2.2.7 Python .....	17
2.2.8 Confusion Matrix .....	18
<b>BAB III METODE PENELITIAN</b> .....	19
3.1 Jenis Penelitian .....	19
3.2 Tahapan Penelitian .....	20
<b>BAB IV HASIL DAN PEMBAHASAN</b> .....	25

4.1	Dataset.....	25
4.2	Eksplorasi Data .....	26
4.3	Preprocessing Data.....	28
4.4	Split Data .....	36
4.5	Pelatihan Model .....	37
4.6	Evaluation Model.....	39
<b>BAB V KESIMPULAN DAN SARAN .....</b>		<b>42</b>
5.1	Kesimpulan .....	42
5.2	Saran .....	43
<b>DAFTAR PUSTAKA .....</b>		<b>44</b>
<b>LAMPIRAN.....</b>		<b>47</b>



UNIVERSITAS  
MERCU BUANA

## DAFTAR TABEL

Tabel 2. 1 Tinjauan Penelitian .....	5
Tabel 2. 2 Confusion Matrix .....	18



## DAFTAR GAMBAR

Gambar 2. 1 Pohon Keputusan XGBoost .....	14
Gambar 2. 2 Konsep Model GBM .....	16
Gambar 3. 1 Tahapan Penelitian .....	20
Gambar 3. 2 Prosedur Penelitian.....	24
Gambar 4. 1 Hasil Dataframe .....	27
Gambar 4. 2 Informasi Atribut.....	27
Gambar 4. 3 Statistik Data .....	27
Gambar 4. 4 Missing Value .....	28
Gambar 4. 5 Data Frame .....	29
Gambar 4. 6 Variabel Distribusi Kematian Pasien .....	30
Gambar 4. 7 Correlation Plot .....	31
Gambar 4. 8 Distribusi Fitur Numerik .....	32
Gambar 4. 9 Distribusi Fitur Biner .....	33
Gambar 4. 10 Numerical Correlation Plot .....	35
Gambar 4. 11 Binary Correlation Plot .....	36
Gambar 4. 12 Hasil Akurasi Algoritma .....	38
Gambar 4. 13 Confusion Matrix .....	40
Gambar 4. 14 Feature Importances Catboost.....	41

## DAFTAR LAMPIRAN

Lampiran 1 Kartu Asistensi .....	47
Lampiran 2 Halaman Persetujuan .....	48
Lampiran 3 Submit Jurnal .....	49
Lampiran 4 Curriculum Vitae .....	50
Lampiran 5 Surat Pernyataan HAKI .....	51
Lampiran 6 Sertifikat BNSP .....	53
Lampiran 7 Form Revisi Dosen Penguji .....	54
Lampiran 8 Hasil Cek Turnitin .....	56

