



DETEKSI STROKE MENGGUNAKAN ALGORITMA *GRADIENT BOOSTING* DENGAN OPTIMASI *HYPERPARAMETER BAYESIAN* UNTUK PENANGANAN DATA TIDAK SEIMBANG

SKRIPSI

Untuk memenuhi persyaratan dalam menyelesaikan program sarjana Strata-1 Fisika

Oleh :

ALVARO PUTRA PRASETYA

2111014310008

**PROGRAM STUDI S-1 FISIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMBUNG MANGKURAT
BANJARBARU**

FEBRUARI 2026

LEMBAR PENGESAHAN

SKRIPSI

DETEKSI STROKE MENGGUNAKAN ALGORITMA *GRADIENT BOOSTING* DENGAN OPTIMASI *HYPERPARAMETER BAYESIAN* UNTUK PENANGANAN DATA TIDAK SEIMBANG

Oleh:

ALVARO PUTRA PRASETYA

NIM. 2111014310008


Telah dipertahankan di depan Dosen penguji pada tanggal:

Pembimbing I



Dr. Arfan Eko Fahrudin, S.Si., M.Eng
NIP. 19790904 200501 1 003


Dosen Penguji:

1. Dr. Nurma Sari, S. Si, M. Si. 

Pembimbing II



Prof. Dr. Ichsan Ridwan, S.Si.,
M.Kom
NIP. 19740707 200212 1 003

2. Rahmad Oktafiansyah, S.Si, M.Si. 

Mengetahui,
Ketua Jurusan Fisika



Dr. Nurlina, S.Si., M.Sc.
NIP. 19760414 200312 2 001

LEMBAR PENGESAHAN

SKRIPSI

DETEKSI STROKE MENGGUNAKAN ALGORITMA *GRADIENT BOOSTING* DENGAN OPTIMASI *HYPERPARAMETER BAYESIAN* UNTUK PENANGANAN DATA TIDAK SEIMBANG

Oleh:

ALVARO PUTRA PRASETYA

NIM. 2111014310008

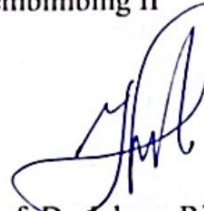
disetujui oleh Dosen Pembimbing untuk disajikan dalam sidang TA Skripsi, pada tanggal:

Pembimbing I



Dr. Arfan Eko Fahrudin, S.Si., M.Eng
NIP. 19790904 200501 1 003

Pembimbing II



Prof. Dr. Ichsan Ridwan, S.Si.,
M.Kom
NIP. 19740707 200212 1 003

Mengetahui,
Ketua Jurusan Fisika



Dr. Nurlina, S.Si., M.Sc.
NIP. 19760414 200312 2 001

LEMBAR PERSEMBAHAN

Penulis persembahkan karya sederhana ini untuk semua pihak yang telah memberikan dukungan, doa, dan motivasi selama proses penyusunan skripsi ini. Semoga karya ini dapat memberikan manfaat dan salah satu sumber referensi kita semua untuk terus berjuang demi kontribusi ilmu pengetahuan di masa depan.

PERNYATAAN

Dengan ini saya menyatakan bahwa dalam skripsi ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar kesarjanaan di suatu Perguruan Tinggi, dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Banjarbaru, 01 Januari 2026



Alvaro Putra Prasetya
NIM. 2111014310008

ABSTRAK

DETEKSI STROKE MENGGUNAKAN ALGORITMA *GRADIENT BOOSTING* DENGAN OPTIMASI *HYPERPARAMETER BAYESIAN* UNTUK PENANGANAN DATA TIDAK SEIMBANG

(Oleh: Alvaro Putra Prasetya | Pembimbing: Dr. Arfan Eko Fahrudin, S.Si., M.Eng., Prof. Dr. Ichsan Ridwan, S.Si., M.Kom., 2025)

Stroke merupakan penyakit kardiovaskular dengan tingkat kematian tinggi yang menjadi beban kesehatan global. Penelitian ini bertujuan mengembangkan model deteksi stroke menggunakan algoritma *Gradient Boosting* (LightGBM dan XGBoost) yang dioptimasi dengan *Bayesian Optimization* untuk menangani ketidakseimbangan data ekstrem (5% kasus positif). Dataset yang digunakan terdiri dari 5.110 data pasien dengan fitur klinis seperti usia, kadar glukosa rata-rata, BMI, *hypertension*, dan riwayat penyakit jantung. *Synthetic Minority Oversampling Technique* (SMOTE) diterapkan untuk menyeimbangkan distribusi kelas dari 249:4.861 menjadi 4.861:4.861. Hasil penelitian menunjukkan bahwa sebelum dilakukan optimasi *Bayesian*, model LightGBM memperoleh *Accuracy* 0,95, *Precision* 0,92, *Recall* 0,97, dan *F1-Score* 0,95, sedangkan XGBoost memperoleh *Accuracy* 0,95, *Precision* 0,94, *Recall* 0,97, dan *F1-Score* 0,96. Setelah optimasi *Bayesian*, model LightGBM mencapai *Accuracy* 0,97, *Precision* 0,96, *Recall* 0,97, dan *F1-Score* 0,97, sedangkan XGBoost memperoleh *Accuracy* 0,97, *Precision* 0,95, *Recall* 0,98, dan *F1-Score* 0,97. Model XGBoost terbaik diimplementasikan dalam aplikasi web berbasis Streamlit untuk memberikan prediksi risiko stroke secara real-time. Penelitian ini diharapkan dapat dikembangkan ke dalam layanan sistem prediksi kesehatan terintegrasi berbasis API agar dapat dihubungkan dengan aplikasi rumah sakit atau platform kesehatan masyarakat.

Kata Kunci : *Stroke, LightGBM, XGBoost, Bayesian Optimization, Data Tidak Seimbang*

ABSTRACT

STROKE DETECTION USING GRADIENT BOOSTING ALGORITHM WITH BAYESIAN HYPERPARAMETER OPTIMIZATION FOR IMBALANCED DATA HANDLING

(By: Alvaro Putra Prasetya | Supervisors: Dr. Arfan Eko Fahrudin, S.Si., M.Eng., Prof. Prof. Dr. Ichsan Ridwan, S.Si., M.Kom., 2025)

Stroke is a cardiovascular disease with a high mortality rate that poses a global health burden. This study aims to develop a stroke detection model using Gradient Boosting algorithms (LightGBM and XGBoost) optimized with Bayesian Optimization to address extreme data imbalance (5% positive cases). The dataset used consisted of 5,110 patient data with clinical features such as age, average glucose levels, BMI, hypertension, and a history of heart disease. The Synthetic Minority Oversampling Technique (SMOTE) was applied to balance the class distribution from 249:4,861 to 4,861:4,861. The results showed that before Bayesian optimization, the LightGBM model achieved an accuracy of 0.95, precision of 0.92, recall of 0.97, and F1-score of 0.95, while XGBoost achieved an accuracy of 0.95, precision of 0.94, recall of 0.97, and F1-score of 0.96. After Bayesian optimization, the LightGBM model achieved an accuracy of 0.97, precision of 0.96, recall of 0.97, and F1-score of 0.97, while XGBoost obtained an accuracy of 0.97, precision of 0.95, recall of 0.98, and F1-score of 0.97. The best XGBoost model was implemented in a Streamlit-based web application to provide real-time stroke risk predictions. This research is expected to be developed into an integrated API-based health prediction system service so that it can be connected to hospital applications or public health platforms.

Keywords: *Stroke, LightGBM, XGBoost, Bayesian Optimization, Imbalanced Data*

PRAKATA

Puji dan syukur penulis panjatkan ke hadirat Allah SWT atas segala rahmat dan karunia-Nya, sehingga penulisan laporan skripsi yang berjudul “Deteksi Stroke Menggunakan Algoritma *Gradient Boosting* dengan Optimasi *Hyperparameter Bayesian* untuk Penanganan Data Tidak Seimbang” ini dapat diselesaikan dengan baik. Penulisan laporan skripsi ini merupakan bagian tugas akademik di Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lambung Mangkurat, sebagai persyaratan untuk menyelesaikan program pendidikan Sarjana (S1) Program Studi Fisika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lambung Mangkurat.

Penulis menyadari bahwa penyelesaian laporan ini tidak lepas dari bantuan berbagai pihak, oleh karena itu pada kesempatan ini penulis mengucapkan terima kasih yang sebesar-besarnya kepada :

1. Bapak Prof. Drs. Abdul Gafur, M.Si., M.Sc., Ph.D. selaku Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lambung Mangkurat.
2. Ibu Dr. Nurlina, S.Si., M.Sc. selaku Ketua Jurusan Fisika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lambung Mangkurat.
3. Bapak Dr. Eka Suarso, S.Si, M.Si selaku dosen Pembimbing Akademik yang selalu memberikan motivasi dan semangat dalam belajar.
4. Bapak Dr. Arfan Eko Fahrudin, S.Si., M.Eng. dan Bapak Prof. Dr. Ichsan Ridwan, S.Si., M.Kom. selaku dosen pembimbing skripsi yang telah meluangkan waktu memberikan bimbingan, kritik dan saran serta dukungan dalam penyusunan Skripsi.
5. Ibu Dr. Nurma Sari, S.Si, M.Si. dan Bapak Rahmad Oktafiansyah, S.Si, M.Si selaku dosen penguji yang telah memberikan kritik dan masukan yang sifatnya membangun, sehingga penelitian ini menjadi lebih baik.
6. Semua dosen FMIPA ULM di Banjarbaru, khususnya dosen Fisika yang telah banyak memberikan ilmu dan pengalamannya selama kuliah.
7. Semua pihak yang selalu menjadi sumber semangat dan inspirasi dalam setiap langkah, baik dalam masa sulit maupun bahagia.

Penulis juga menyadari bahwa laporan ini masih jauh dari kesempurnaan. Oleh karena itu, penulis mengharapkan adanya kritik dan saran yang sifatnya membangun dari semua pihak. Agar diperoleh kesempurnaan dalam pembuatan laporan yang akan datang. Semoga laporan ini dapat berguna bagi penulis pada khususnya dan bagi pembaca pada umumnya.

Banjarbaru, 01 Januari 2026

Alvaro Putra Prasetya

DAFTAR ISI

	Halaman
SKRIPSI.....	i
LEMBAR PENGESAHAN	i
LEMBAR PENGESAHAN	ii
LEMBAR PERSEMBAHAN	iii
PERNYATAAN.....	iv
ABSTRAK	v
ABSTRACT	vi
PRAKATA	vii
DAFTAR ISI.....	ix
DAFTAR GAMBAR	xi
DAFTAR TABEL.....	xii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Perumusan Masalah	3
1.3 Tujuan Penelitian	3
1.4 Manfaat Penelitian	4
1.5 Batasan Masalah.....	4
BAB II TINJAUAN PUSTAKA.....	5
2.1 Stroke	5
2.2 <i>Machine Learning</i> (ML)	7
2.3 ML Dalam Sistem Deteksi dan Optimasi.....	8
2.3.1 LightGBM (<i>Light Gradient Boosting Machine</i>)	8
2.3.2 XGBoost (<i>Extreme Gradient Boosting</i>).....	9
2.4 Penanganan Data Tidak Seimbang dengan SMOTE	10
2.5 Optimasi <i>Hyperparameter</i> dengan <i>Bayesian Optimization</i>	12
2.6 Evaluasi Model.....	13
BAB III METODE PENELITIAN.....	15
3.1 Waktu Dan Tempat Penelitian	15
3.2 Alat Penelitian.....	15
3.3 Prosedur Penelitian.....	15
3.3.1 Tahapan Penelitian	15

3.3.2	Persiapan Alat dan Bahan	16
3.3.3	Pemrosesan Data	16
3.3.4	SMOTE	17
3.3.5	Pembuatan Model.....	19
3.3.5.1	Optimasi <i>Hyperparameter Bayesian</i>	21
3.3.6	Evaluasi Model.....	22
3.3.7	Implementasi Model pada Aplikasi Web.....	23
BAB IV HASIL DAN PEMBAHASAN		24
4.1	Analisis Eksplorasi Data	24
4.2	Penanganan Kelas Tidak Seimbang.....	28
4.3	Pelatihan dan Evaluasi Model ML.....	30
4.4	Optimasi Model Menggunakan <i>Bayesian Optimization</i>	33
4.5	Implementasi Model Untuk Deteksi Stroke.....	35
BAB V PENUTUP.....		37
5.1	Kesimpulan	37
5.2	Saran.....	37
DAFTAR PUSTAKA		38
LAMPIRAN		41
BIODATA PENULIS		62

DAFTAR GAMBAR

	Halaman
Gambar 3.1 Bagan Alur Penelitian Pembuatan Model	16
Gambar 3. 2 Flowchart SMOTE	18
Gambar 3.3 Grafik Kelas Sebelum dan Sesudah SMOTE.....	19
Gambar 3.4 Flowchart Data <i>Training</i> Model ML.....	20
Gambar 3.5 Kode Program Untuk Pelatihan Model LightGBM	21
Gambar 3.6 Kode Program Untuk Pelatihan Model XGBoost	21
Gambar 3.7 Flowchart Data <i>Testing</i> Model ML	21
Gambar 3.8 Kode Program Untuk Optimasi Model LightGBM.....	22
Gambar 3.9 Kode Program Untuk Optimasi Model XGBoost	22
Gambar 3.10 Kode Program Untuk Evaluasi Model	23
Gambar 3.11 Kode Program Untuk Implementasi Model pada Aplikasi Web.....	23
Gambar 4.1 Gambar plot Distribusi Semua Fitur Numerik	27
Gambar 4.2 Gambar Matriks Korelasi	28
Gambar 4.3 Distribusi Kelas Sebelum Proses SMOTE	29
Gambar 4.4 Distribusi Kelas Setelah Proses SMOTE	29
Gambar 4.5 Performa Model LightGBM.....	30
Gambar 4.6 Receiver Operating Characteristic LightGBM.....	30
Gambar 4.7 Performa Model XGBoost	31
Gambar 4.8 <i>Receiver Operating Characteristic</i> XGBoost	31
Gambar 4.9 Perbandingan Performa Model LightGBM dan XGBoost.....	32
Gambar 4.10 Performa Model LightGBM Setelah Optimasi	33
Gambar 4.11 <i>Receiver Operating Characteristic</i> LightGBM Setelah Optimasi ..	33
Gambar 4.12 Performa Model XGBoost Setelah Optimasi	34
Gambar 4.13 <i>Receiver Operating Characteristic</i> XGBoost Setelah Optimasi.....	34
Gambar 4.14 Perbandingan Performa Model Setelah Optimasi	35
Gambar 4.15 Hasil Prediction Menggunakan Streamlit	36

DAFTAR TABEL

	Halaman
Tabel 4.1 Spesifikasi Data.....	24
Tabel 4.2 <i>Missing Value</i> Sebelum <i>Cleaning Data</i>	25
Tabel 4.3 Hasil Analisis Statistik dengan Statistik Deskriptif.....	26