



SKRIPSI

PENERAPAN SMOTE-ENN DAN XGBOOST DENGAN OPTIMASI BAYESIAN UNTUK MENGATASI *IMBALANCE CLASS* DALAM DETEKSI PENYAKIT GAGAL JANTUNG

BINTANG TIARA PRAMESTI

NPM 21081010333

DOSEN PEMBIMBING

Fetty Tri Anggraeny, S.Kom., M.Kom

Wahyu Syaifullah J. S., S.Kom., M.Kom

**KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, SAINS, DAN TEKNOLOGI
UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAWA TIMUR
FAKULTAS ILMU KOMPUTER
PROGRAM STUDI INFORMATIKA
SURABAYA
2025**



SKRIPSI

**PENERAPAN SMOTE-ENN DAN XGBOOST
DENGAN OPTIMASI BAYESIAN UNTUK
MENGATASI *IMBALANCE CLASS* DALAM
DETEKSI PENYAKIT GAGAL JANTUNG**

BINTANG TIARA PRAMESTI

NPM 21081010333

DOSEN PEMBIMBING

Fetty Tri Anggraeny, S.Kom., M. Kom.

Wahyu Syaifullah J.S., S.Kom., M.Kom.

**KEMENTERIAN PENDIDIKAN TINGGI, SAINS, DAN TEKNOLOGI
UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAWA TIMUR
FAKULTAS ILMU KOMPUTER
PROGRAM STUDI INFORMATIKA
SURABAYA
2025**

Halaman ini sengaja dikosongkan

LEMBAR PENGESAHAN

PENERAPAN SMOTE-ENN DAN XGBOOST DENGAN OPTOMASI BAYESIAN UNTUK MENGATASI *IMBALANCE CLASS* DALAM DETEKSI PENYAKIT GAGAL JANTUNG

Oleh :

BINTANG TIARA PRAMESTI
NPM.21081010333

Telah dipertahankan dihadapan dan diterima oleh Tim Penguji Skripsi Prodi Informatika Fakultas Ilmu Komputer Universitas Pembangunan Nasional Veteran Jawa Timur Pada tanggal 5 Juni 2025.

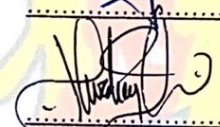
Menyetujui,

Fetty Tri Anggraeny, S.Kom., M.Kom
NIP. 19820211 2021212 005



(Pembimbing I)

Wahyu Syaifullah J. S., S.Kom., M.Kom
NIP. 19860825 202121 1 003



(Pembimbing II)

Chrystia Aji Putra, S.Kom, M.T.
NIP. 19861008 2021211 001



(Ketua Penguji)

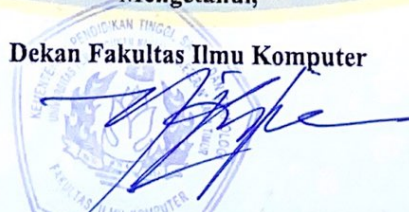
Eka Prakarsa Mandyartha, S.T., M.Kom
NIP. 19880525 2018031 001



(Anggota Penguji)

Mengetahui,

Dekan Fakultas Ilmu Komputer



Prof. Dr. Ir. Novirina Hendrasarie, M.T.
NIP. 19681126 199403 2 001

Halaman ini sengaja dikosongkan

LEMBAR PERSETUJUAN

**PENERAPAN SMOTE-ENN DAN XGBOOST DENGAN OPTIMASI
BAYESIAN UNTUK MENGATASI *IMBALANCE CLASS* DALAM
DETEKSI PENYAKIT GAGAL JANTUNG**

Oleh:

BINTANG TIARA PRAMESTI

NPM. 21081010333



Menyetujui,

Koordinator Program Studi Informatika

Fakultas Ilmu Komputer

Fetty Tri Anggraeny, S.Kom., M.Kom

NIP. 19820211 2021212 005

Halaman ini sengaja dikosongkan

SURAT PERNYATAAN BEBAS PLAGIASI

Yang bertandatangan di bawah ini:

Nama Mahasiswa : Bintang Tiara Pramesti
NPM : 21081010333
Program : Sarjana (S1)
Program Studi : Informatika
Fakultas : Ilmu Komputer

Menyatakan bahwa dalam dokumen ilmiah Skripsi ini tidak terdapat bagian dari karya ilmiah lain yang telah diajukan untuk memperoleh gelar akademik di suatu Lembaga Pendidikan Tinggi, dan juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang/lembaga lain, kecuali yang secara tertulis disitasi dalam dokumen ini dan disebutkan secara lengkap dalam daftar pustaka.

Dan saya menyatakan bahwa dokumen ilmiah ini bebas dari unsur-unsur plagiasi. Apabila dikemudian hari ditemukan indikasi plagiat pada Skripsi ini, saya bersedia menerima sanksi sesuai dengan peraturan perundang-undangan yang berlaku.

Demikian surat pernyataan ini saya buat dengan sesungguhnya tanpa ada paksaan dari siapapun juga dan untuk dipergunakan sebagaimana mestinya.



Surabaya, 1 September 2025

Yang Membuat Pernyataan,



Bintang Tiara Pramesti
NPM. 21081010333

Halaman ini sengaja dikosongkan

ABSTRAK

Nama Mahasiswa / NPM : Bintang Tiara Pramesti / 21081010333

Judul Skripsi : Penerapan Smote-Enn Dan Xgboost Dengan Optimasi Bayesian Untuk Mengatasi Imbalance Class Dalam Deteksi Penyakit Gagal Jantung

Dosen Pembimbing : 1. Fetty Tri Anggraeny, S.Kom., M.Kom.
2. Wahyu Syaifullah J.S., S.Kom., M.Kom.

Ketidakseimbangan kelas (*class imbalance*) merupakan permasalahan umum dalam penerapan machine learning pada data medis, khususnya dalam mendeteksi penyakit gagal jantung. Kondisi ini menyebabkan model lebih cenderung mengenali kelas mayoritas (sehat) dibandingkan kelas minoritas (gagal jantung). Penelitian ini bertujuan mengatasi masalah tersebut dengan menerapkan SMOTE-ENN serta membandingkan performa XGBoost pada tiga skenario: tanpa *balancing*, dengan SMOTE-ENN Default, dan dengan SMOTE-ENN yang dioptimasi menggunakan *Bayesian Optimization* (BO). Data diproses melalui tahap *preprocessing*, penanganan *outlier*, pembagian dataset (70:15:15), serta evaluasi menggunakan metrik *Accuracy*, *Precision*, *Recall*, *F1-score*, dan *Stratified K-Fold Cross-Validation*.

Hasil penelitian menunjukkan bahwa XGBoost tanpa SMOTE menghasilkan akurasi tinggi (90,82%) tetapi sangat bias dengan nilai *Recall* hanya 09,17%. Penerapan SMOTE-ENN *Default* meningkatkan performa dengan *Recall* 94,08% dan *F1-score* 92,91%, sedangkan SMOTE-ENN Fine-Tuned BO memberikan *Recall* tertinggi (94,24%) dengan *F1-score* 94,71%. Meskipun akurasi sedikit menurun dibandingkan SMOTE-ENN *Default*, model *fine-tuned* terbukti lebih optimal dalam mendeteksi kelas minoritas. Dengan demikian, kombinasi SMOTE-ENN dan optimasi *Bayesian* dapat meningkatkan sensitivitas model XGBoost serta menjadi strategi efektif untuk mengatasi ketidakseimbangan data medis.

Kata kunci : Gagal Jantung, *Imbalance Class*, SMOTE-ENN, XGBoost, *Bayesian Optimization*, Hyperparameter Tuning

Halaman ini sengaja dikosongkan

ABSTRACT

Student Name / NPM : Bintang Tiara Pramesti / 21081010333
Thesis Title : Implementation of SMOTE-ENN and XGBoost
with Bayesian Optimization to Address Class
Imbalance in Heart Failure Detection
Advisor : 1. Fetty Tri Anggraeny, S.Kom., M.Kom.
2. Wahyu Syaifullah J.S., S.Kom., M.Kom

Class imbalance is a common issue in applying machine learning to medical data, especially in detecting heart failure. This condition makes models biased toward the majority class (healthy) while ignoring the minority (heart failure). This study addresses the problem by applying SMOTE-ENN and comparing the performance of XGBoost under three scenarios: without balancing, with Default SMOTE-ENN, and with Fine-Tuned SMOTE-ENN using Bayesian Optimization (BO). The dataset was preprocessed, outliers handled, split into 70:15:15, and evaluated using Accuracy, Precision, Recall, F1-score, and Stratified K-Fold Cross-Validation.

The results show that XGBoost without SMOTE achieved high Accuracy (94.27%) but had poor Recall (09.17%). Applying Default SMOTE-ENN significantly improved performance with Recall of 94.08% and F1-score of 92.91%, while Fine-Tuned SMOTE-ENN BO achieved the highest Recall (94.24%) with an F1-score of 94.71%. Although Accuracy was slightly lower than the default, the fine-tuned model proved more effective in detecting minority cases. In conclusion, combining SMOTE-ENN with Bayesian Optimization enhances the sensitivity of XGBoost and provides an effective strategy to address class imbalance in medical datasets.

Keywords: Heart Failure, Imbalance Class, SMOTE-ENN, XGBoost, Bayesian Optimization, Hyperparameter Tuning.

Halaman ini sengaja dikosongkan

KATA PENGANTAR

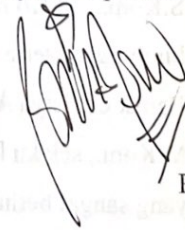
Puji syukur kehadiran Allah SWT atas segala rahmat, hidayah dan karunia-Nya kepada penulis sehingga skripsi dengan judul **“PENERAPAN SMOTE-ENN DAN XGBOOST DENGAN OPTIMASI BAYESIAN UNTUK MENGATASI IMBALANCE CLASS DALAM DETEKSI PENYAKIT GAGAL JANTUNG”** dapat terselesaikan dengan baik.

Skripsi ini disusun sebagai salah satu syarat untuk memperoleh gelar Sarjana Komputer pada Program Studi Informatika, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional “Veteran” Jawa Timur. Penyusunan skripsi ini tidak lepas dari bantuan, bimbingan, dan dukungan dari berbagai pihak. Oleh karena itu, penulis mengucapkan terima kasih yang sebesar-besarnya kepada:

1. Ibu Prof. Dr. Ir. Novirina Hendrasarie, MT, selaku Dekan Fakultas Ilmu Komputer Universitas Pembangunan Nasional “Veteran” Jawa Timur.
2. Ibu Fetty Try Anggraeny, S.Kom., M.Kom., selaku Ketua Program Studi Informatika Fakultas Ilmu Komputer Universitas Pembangunan Nasional “Veteran” Jawa Timur.
3. Ibu Fetty Try Anggraeny, S.Kom., M.Kom. dan Bapak Wahyu Syaifullah J.S., S.Kom., M.Kom., selaku Dosen Pembimbing yang telah memberikan arahan, bimbingan, serta motivasi kepada penulis selama proses penyusunan skripsi ini.
4. Bapak Christya Aji Putra, S.Kom., M.T. dan Bapak Eka Prakarsa Mandyartha, S.T., M.Kom., selaku Dosen Penguji yang telah memberikan masukan, kritik, serta saran yang sangat berharga demi kesempurnaan skripsi ini.
5. Seluruh dosen Program Studi Informatika yang telah memberikan ilmu, bimbingan, dan inspirasi selama proses perkuliahan yang menjadi bekal penting dalam penyusunan skripsi ini.
6. Kepada Mama Tercinta, Rina Trisyawati, yang telah menjadi sumber motivasi dan kekuatan serta dukungan moral dan material yang tak ternilai harganya.
7. Kepada Nenek dan Adik, yang turut serta mendoakan dan memberikan dukungan bagi penulis.

8. Yang teristimewa, Ahmad Mubarak Al bira, yang selalu setia mendampingi, meluangkan waktu serta memberikan dukungan tanpa henti, sehingga penulis dapat menyelesaikan skripsi ini.
 9. Azila Lailannafisa dan Athaya Aqilah, selaku sahabat yang selalu menemani penulis dari awal perkuliahan hingga akhir perkuliahan.
 10. Sahabat terdekat yang lainnya, Arab, Farras, Aqilah, Hilmi, Mahen dan Samuel, yang turut memberikan motivasi dan mewarnai hari penulis selama di perkuliahan.
 11. Kepada teman-teman terdekat penulis lainnya yang senantiasa memberikan semangat dikala penulis dalam keadaan sedih serta selalu mengajak penulis untuk bermain agar tidak terlalu stress memikirkan masalah yang dihadapi
 12. Semua pihak yang tidak bisa disebutkan satu-persatu yang telah banyak membantu memberikan pemikiran demi kelancaran dan keberhasilan penulisan skripsi ini.
- Penulis menyadari bahwa di dalam penyusunan skripsi ini banyak terdapat kekurangan. Untuk itu kritik dan saran yang membangun dari semua pihak sangat diharapkan demi kesempurnaan penulisan skripsi ini. Akhirnya, dengan segala keterbatasan yang penulis miliki semoga laporan ini dapat bermanfaat bagi semua pihak umumnya dan penulis pada khususnya.

Surabaya, 1 September 2025



Penulis

DAFTAR ISI

LEMBAR PENGESAHAN	iv
LEMBAR PERSETUJUAN	vi
SURAT PERNYATAAN BEBAS PLAGIASI.....	viii
ABSTRAK	x
ABSTRACT	xii
KATA PENGANTAR.....	xiv
DAFTAR ISI.....	xvi
DAFTAR GAMBAR.....	xviii
DAFTAR TABEL	xix
BAB I.....	21
1.1 Latar Belakang	21
1.2 Rumusan Masalah	22
1.3 Tujuan Penelitian.....	23
1.4 Manfaat Penelitian.....	23
1.5 Batasan Masalah.....	24
BAB II	25
2.1 Penelitian Terdahulu.....	25
2.2 Gagal Jantung	27
2.3 <i>Imbalanced Class</i>	28
2.4 SMOTE (<i>Synthetic Minority Oversampling Technique</i>)	29
2.5 ENN (<i>Edited Nearest Neighbors</i>).....	30
2.6 Teknik <i>Ensemble</i>	30
2.7 <i>Boosting</i>	31
2.8 XGBoost (<i>Extreme Gradient Boosting</i>)	32
2.9 <i>Hyperparameter Tuning</i>	36
2.10 <i>Confusion Matrix</i>	39
BAB III.....	42
3.1 Tahapan Penelitian	42
3.2 Studi Literatur.....	42
3.3 Pengumpulan Data	42

3.4 Preprocessing.....	44
3.5 Pembuatan Model.....	46
3.5.1 Teknik Resampling SMOTE-ENN.....	47
3.5.2 Implementasi XGBoost.....	49
3.5.3 Penyetelan <i>Hyperparameter Tuning</i>	53
3.6 Evaluasi Metrik	55
3.7 Skenario Pengujian.....	56
BAB IV	58
4.1 <i>Import Library</i>	58
4.2 Persiapan Data.....	61
4.2.1 Pemuatan Dataset.....	61
4.2.2 <i>Outlier Handling</i>	62
4.2.3 Transformasi <i>Skewness</i>	63
4.2.4 <i>Split</i> Dataset	63
4.2.5 Normalisasi	65
4.3 Arsitektur Model	66
4.3.1 Fungsi SMOTE-ENN	67
4.3.2 Fungsi XGBoost	71
4.4 Pelatihan Model.....	73
4.5 Skenario Pengujian.....	76
4.5.1 Skenario Perbandingan Dataset	77
4.5.2 Skenario Pengujian Hyperparameter Tuning.....	78
4.5.3 Skenario Pengujian Evaluasi Model	80
4.5.4 Skenario Pengujian Evaluasi Model dengan Perbandingan <i>Cross Validation</i>	85
4.6 Testing	94
BAB V.....	96
5.1 Kesimpulan.....	96
5.2 Saran.....	97
DAFTAR PUSTAKA.....	98
LAMPIRAN.....	100

DAFTAR GAMBAR

Gambar 2. 1 Ilustrasi SMOTE	29
Gambar 2. 2 Ilustrasi Proses Ensemble.....	31
Gambar 2. 3 Ilustrasi Proses Boosting	32
Gambar 2. 4 Proses Ilustrasi XGBoost	32
Gambar 3. 1 Flowchart Tahapan Penelitian.....	42
Gambar 3. 2 Flowchart Preprocessing	44
Gambar 3. 3 Flowchart Alur Pembuatan Model	47
Gambar 3. 4 Flowchart Teknik Resampling SMOTE-ENN	47
Gambar 3. 5 Flowchart XGBoost	49
Gambar 3. 6 Flowchart Hyperparameter Tuning	53
Gambar 4. 1 Hasil Pencarian Hyperparameter Tuning	79
Gambar 4. 2 Classification Report XGBoost Tanpa SMOTE	80
Gambar 4. 3 Confusion Matrix XGBoost Tanpa SMOTE.....	81
Gambar 4. 4 Classification Report XGBoost SMOTE-ENN (Default).....	82
Gambar 4. 5 Confusion Matrix XGBoost SMOTE-ENN (Default)	82
Gambar 4. 6 Classification Report XGBoost SMOTE-ENN dengan BO (Fine-Tuned)	83
Gambar 4. 7 Confusion Matrix XGBoost SMOTE-ENN dengan BO (Fine-Tuned)	83
Gambar 4. 8 Fold 1 Tanpa SMOTE	86
Gambar 4. 9 Fold 2 Tanpa SMOTE	86
Gambar 4. 10 Fold 3 Tanpa SMOTE	87
Gambar 4. 11 Fold 4 Tanpa SMOTE	87
Gambar 4. 12 Fold 5 Tanpa SMOTE	88
Gambar 4. 13 Fold 1 XGBoost SMOTE-ENN Default	88
Gambar 4. 14 Fold 2 XGBoost SMOTE-ENN Default	89
Gambar 4. 15 Fold 3 XGBoost SMOTE-ENN Default	89
Gambar 4. 16 Fold 4 XGBoost SMOTE-ENN Default	90
Gambar 4. 17 Fold 5 XGBoost SMOTE-ENN Default	90
Gambar 4. 18 Fold 1 XGBoost SMOTE-ENN BO.....	91
Gambar 4. 19 Fold 2 XGBoost SMOTE-ENN BO.....	91
Gambar 4. 20 Fold 3 XGBoost SMOTE-ENN BO.....	92
Gambar 4. 21 Fold 4 XGBoost SMOTE-ENN BO.....	92
Gambar 4. 22 Fold 5 XGBoost SMOTE-ENN BO.....	93

DAFTAR TABEL

Tabel 2. 1 Tabel Confusion Matrix	39
Tabel 3. 1 Rincian Fitur Dataset	43
Tabel 3. 2 Contoh Implementasi Data.....	48
Tabel 3. 3 Data Sesudah Penambahan Data Sintetis.....	48
Tabel 3. 4 Data Sesudah Penambahan Data Sintesis	49
Tabel 3. 5 Hasil Gradien dan Hessian	51
Tabel 3. 6 Contoh Implementasi Bayesian Optimization	53
Tabel 3. 7 Skenario Pengujian	56
Tabel 4. 1 Hasil Perbandingan Dataset	77
Tabel 4. 2 Hasil Pengujian Parameter	78
Tabel 4. 3 Hasil Testing	95

Halaman ini sengaja dikosongkan