



Klasifikasi Data Tidak Seimbang Menggunakan SMOTE-ENN dengan Optimasi Bayesian

Bintang Tiara Pramesti¹, Fetty Tri Anggraeny², Wahyu Syaifullah Jauharis Saputra³

^{1,2,3}Universitas Pembangunan Nasional "Veteran" Jawa Timur, Indonesia

E-mail: 21081010333@student.upnjatim.ac.id, fettyanggraeny.if@upnjatim.ac.id,
wahyu.s.j.saputra.if@upnjatim.ac.id

Article Info	Abstract
Article History Received: 2025-08-05 Revised: 2025-09-12 Published: 2025-10-20 Keywords: <i>Heart Failure;</i> <i>Imbalance Class, SMOTE-ENN;</i> <i>XGBoost;</i> <i>Bayesian Optimization;</i> <i>Hyperparameter Tuning.</i>	Early heart failure detection is vital, yet medical data often suffers from imbalance class, biasing predictive models and reducing accuracy for the critical minority class. This study addresses this by integrating Synthetic Minority Oversampling Technique (SMOTE) and Edited Nearest Neighbors (ENN) for data balancing and cleaning, combined with Extreme Gradient Boosting (XGBoost) as the primary classifier. Model performance is further optimized via Bayesian Optimization for efficient hyperparameter tuning. Experimental results demonstrate SMOTE-ENN effectively balances class distribution and reduces false negatives. Bayesian Optimization successfully increased the model's Area Under the Curve (AUC) to 0.84. Substantial improvements in Recall for the positive class, alongside consistent performance across visual analyses, confirm this hybrid approach's effectiveness in building a robust, accurate, and reliable heart failure detection model.
Artikel Info	Abstrak
Sejarah Artikel Diterima: 2025-08-05 Direvisi: 2025-09-12 Dipublikasi: 2025-10-20 Kata kunci: <i>Gagal Jantung;</i> <i>Imbalance Class;</i> <i>SMOTE-ENN;</i> <i>XGBoost;</i> <i>Bayesian Optimization;</i> <i>Hyperparameter Tuning.</i>	Deteksi dini gagal jantung krusial, namun data medis sering menghadapi imbalance class, menyebabkan bias model dan akurasi rendah pada kelas minoritas. Penelitian ini mengusulkan pendekatan komprehensif: mengintegrasikan Synthetic Minority Oversampling Technique (SMOTE) dan Edited Nearest Neighbors (ENN) untuk penyeimbangan dan pembersihan data, serta menggunakan Extreme Gradient Boosting (XGBoost) sebagai model klasifikasi utama. Kinerja model ditingkatkan melalui hyperparameter tuning dengan Bayesian Optimization yang efisien. Hasil eksperimen menunjukkan SMOTE-ENN efektif menyeimbangkan distribusi kelas dan mengurangi false negative. Bayesian Optimization berhasil meningkatkan Area Under the Curve (AUC) model hingga 0.84. Peningkatan signifikan pada Recall kelas positif, didukung konsistensi performa visual, menegaskan efektivitas pendekatan hibrida ini dalam membangun model deteksi gagal jantung yang robust, akurat, dan andal.

I. PENDAHULUAN

Penyakit gagal jantung merupakan kondisi medis yang serius dan memengaruhi jutaan orang di seluruh dunia (Verizarie, 2020). Ditandai dengan ketidakmampuan jantung untuk memompa darah secara efisien, penyakit ini menjadi salah satu penyebab utama kematian global, dengan lebih dari 17 juta kematian setiap tahunnya menurut data Organisasi Kesehatan Dunia (WHO) (Anindita and Margarini, 2021). Di Indonesia, angka kematian akibat penyakit jantung koroner, stroke, dan gagal jantung yang terkait dengan hipertensi menunjukkan betapa pentingnya deteksi dini. Intervensi medis yang tepat waktu dapat mengurangi angka morbiditas dan mortalitas, sehingga upaya untuk meningkatkan metode deteksi menjadi sangat krusial (Anindita and Margarini, 2021).

Dalam konteks analisis data medis, penerapan machine learning untuk deteksi penyakit menghadapi tantangan signifikan, terutama

masalah imbalance class (Fransiskus Sitanggang and Sitompul, 2024). Fenomena ini terjadi ketika jumlah sampel dalam dataset tidak seimbang, di mana jumlah pasien gagal jantung (kelas minoritas) jauh lebih sedikit dibandingkan individu sehat (kelas mayoritas). Akibatnya, model prediktif sering kali bias terhadap kelas mayoritas, menghasilkan akurasi tinggi secara keseluruhan tetapi performa buruk dalam mengidentifikasi kelas minoritas yang justru paling penting (Syukron, Santoso and Widiharih, 2020). Hal ini menunjukkan perlunya pendekatan yang lebih efektif untuk menangani masalah ini dalam pengembangan model prediktif (Rizky Mubarok and Herteno, 2022).

Penelitian sebelumnya oleh Wang dkk. (2021) telah mengeksplorasi kombinasi SMOTE-ENN dan XGBoost untuk klasifikasi data tidak seimbang, menunjukkan hasil yang menjanjikan dalam kecepatan klasifikasi dan pembersihan noise. Namun, studi tersebut masih menunjukkan

keterbatasan dalam parameter evaluasi seperti Precision dan F1-Score, yang sering kali terkait dengan kompleksitas penyetelan hyperparameter dalam algoritma machine learning. Untuk mengatasi keterbatasan ini, Bayesian Optimization muncul sebagai metode penyetelan hyperparameter yang lebih efisien (Wu et al., 2019). Dengan pendekatan probabilistik yang adaptif, metode ini dapat menemukan kombinasi parameter optimal dengan lebih cepat, mengurangi risiko overfitting, dan meningkatkan akurasi model (Rizky Mubarak and Herteno, 2022). Integrasi teknik penanganan imbalance class dengan optimasi parameter yang canggih diharapkan dapat menghasilkan sistem klasifikasi yang lebih robust dan akurat untuk diagnosis gagal jantung (Syukron, Saputro and Widodo, 2023).

II. METODE PENELITIAN

1. Pengumpulan Data

Dataset penelitian ini berasal dari Kaggle <https://www.kaggle.com/datasets/alexteboul/heart-disease-health-indicators-dataset/data> dengan total 253.680 data. Dari jumlah tersebut, sebanyak 229.787 responden tidak memiliki atau tidak pernah mengalami penyakit jantung, sedangkan 23.893 responden tercatat pernah mengalami penyakit jantung. Distribusi data tersebut menunjukkan ketidakseimbangan kelas yang cukup signifikan.

2. Preprocessing Data

- Data yang digunakan bersumber dari [Pemisahan Fitur dan Target: Kolom HeartDiseaseorAttack sebagai target (y), sisanya fitur (X).
- Outlier Handling: Winsorizing pada BMI (13-60) dan clipping pada MentHlth serta PhysHlth (maksimum 30 hari).
- Transformasi Skewness: np.log1p() pada MentHlth dan PhysHlth.
- Split Dataset: Data dibagi menjadi training, validation, dan test set secara stratified dengan dua proporsi: 80:10:10 dan 70:15:15.
- Normalisasi: StandardScaler untuk menstandarisasi fitur.

3. Pembuatan Model

- Teknik Resampling SMOTE-ENN: Menggunakan implementasi SMOTE-GPU dan ENN-GPU untuk menyeimbangkan dan membersihkan data training.

- Implementasi XGBoost: Model XGBoost dibangun dengan komponen seperti InitialLogOddsCalculator, GradientHessianCalculator, XGBDataPreparer, XGBParamConfigurator, XGBTrainer, dan HybridXGBoostGPU untuk memanfaatkan akselerasi GPU.

- Penyetelan Hyperparameter (Bayesian Optimization): Mengoptimalkan hyperparameter XGBoost (n_estimators, max_depth, learning_rate, gamma, reg_lambda) pada validation set untuk memaksimalkan Recall.

4. Evaluasi Metrik

Kinerja model dievaluasi menggunakan Confusion Matrix untuk menghitung Accuracy, Precision, Recall, dan F1-Score. Selain itu, digunakan juga Calibration Curve, ROC-AUC Curve, PR Curve, dan F1-Score vs Threshold untuk analisis visual.

III. HASIL DAN PEMBAHASAN

1. Hasil Perbandingan Split Dataset

Tabel 1. Hasil Perbandingan Split Dataset

Split Dataset	Akurasi	Presisi	Recall	F1-Score
80:10:10	0.935749	0.940935	0.940787	0.940859
70:15:15	0.937473	0.943218	0.941726	0.942471

Skenario pembagian dataset 70:15:15 secara signifikan lebih unggul dibandingkan 80:10:10 dalam semua metrik evaluasi, dengan akurasi, precision, recall, dan F1-Score yang lebih tinggi, sehingga direkomendasikan sebagai konfigurasi optimal untuk deteksi penyakit gagal jantung.

2. Hasil Pengujian Hyperparameter Tuning

Tabel 2. Hasil Pengujian Hyperparameter Tuning

Hyperparameter Tuning	Nilai
n_estimators	107
max_depth	9
learning_rate	0.0377
gamma	1.3386
reg_lambda	1.0444
recall	0.9418

Bayesian Optimization berhasil menemukan kombinasi hyperparameter yang menghasilkan Recall tertinggi (0.9418), menunjukkan efisiensinya dalam optimasi.

3. Hasil Pengujian Perbandingan Model

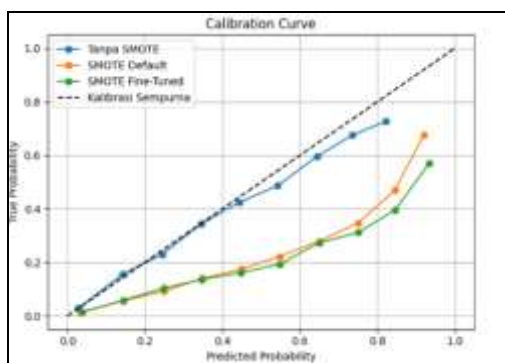
Tabel 3. Hasil Pengujian Model

Model	Accuracy	Precision	Recall	F1-score
SMOTE-ENN BO	0.9375	0.9432	0.9417	0.9425
SMOTE-ENN Default	0.9319	0.9449	0.9288	0.9368
Tanpa SMOTE	0.9079	0.5533	0.1138	0.1888

Hasil pengujian menunjukkan bahwa model SMOTE-ENN dengan optimasi Bayesian mencapai kinerja terbaik, dengan akurasi 0,9375, precision 0,9432, recall 0,9417, dan F1-score 0,9425, menandakan kemampuannya yang superior dalam mendeteksi kasus positif. Sebaliknya, model SMOTE-ENN dengan parameter default menunjukkan performa kompetitif, meskipun sedikit lebih rendah, dengan akurasi 0,9319 dan recall 0,9288, yang menggarisbawahi dampak positif dari optimasi hiperparameter.

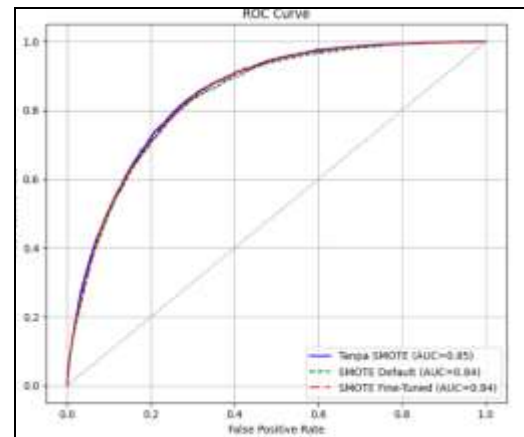
Di sisi lain, model tanpa SMOTE mengalami penurunan signifikan, dengan recall hanya 0,1138 dan F1-score 0,1888, serta precision 0,5533, menunjukkan ketidakmampuan dalam mendeteksi kasus gagal jantung dan meningkatkan risiko false negative. Temuan ini menegaskan pentingnya penerapan teknik penanganan ketidakseimbangan data seperti SMOTE-ENN untuk meminimalkan risiko tidak terdeteksinya pasien dengan penyakit.

4. Evaluasi Model (Kurva)



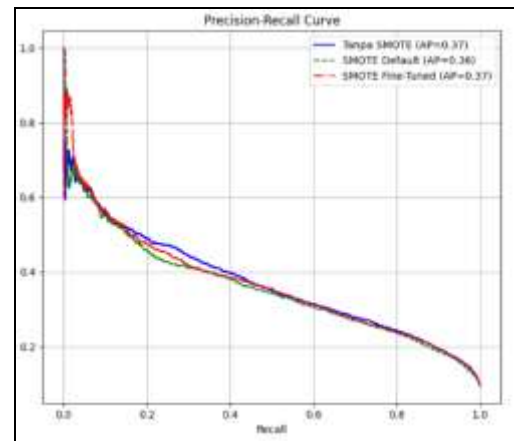
Gambar 1. Calibration Curve

Pada Gambar 1 model tanpa SMOTE lebih terkalibrasi. Model SMOTE cenderung underestimate probabilitas pada rentang menengah hingga tinggi, menunjukkan perlunya kalibrasi tambahan.



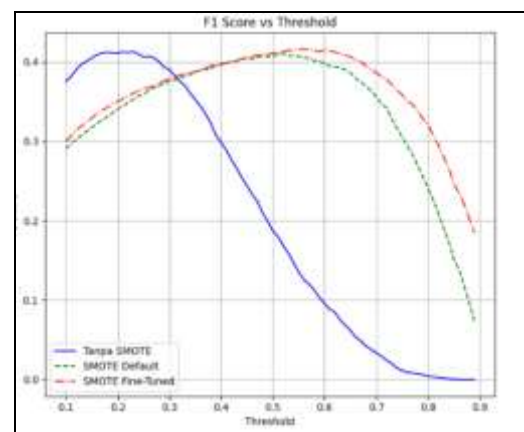
Gambar 2. ROC-AUC Curve

ROC-AUC Curve: Nilai AUC sekitar 0.84-0.85 untuk semua skenario, menunjukkan kemampuan diskriminasi yang baik, namun SMOTE tidak memberikan peningkatan signifikan pada AUC.



Gambar 3. PR Curve

Nilai Average Precision sekitar 0.36-0.37 untuk semua skenario. Penerapan SMOTE tidak memberikan peningkatan signifikan pada trade-off antara Precision dan Recall.



Gambar 4. F1-score vs Threshold

Model tanpa SMOTE optimal pada threshold rendah (0.2-0.25). Model SMOTE (default dan fine-tuned) lebih stabil pada rentang threshold yang lebih luas (0.5-0.6) dengan F1-Score puncak sekitar 0.42-0.43.

5. Testing

Tabel 4. Hasil Testing

Sample	Label Asli	Probabilitas	Prediksi
1	1	0.5311	1
2	1	0.8492	1
3	1	0.3103	0
6	0	0.1499	0
7	0	0.1585	0
8	0	0.0272	0

Hasil Testing Individual Model mampu mengklasifikasikan sebagian besar sampel dengan benar, baik kelas positif maupun negatif, dengan probabilitas prediksi yang meyakinkan. Terjadi misclassification pada kasus borderline (probabilitas mendekati threshold), menunjukkan tantangan dalam data medis dengan karakteristik mirip antar kelas.

IV. SIMPULAN DAN SARAN

A. Simpulan

Penerapan SMOTE-ENN secara efektif menyeimbangkan distribusi kelas, meningkatkan Recall kelas positif dari 0.78 menjadi 0.89 (0.9417 dengan cross-validation), yang berarti sensitivitas model dalam mendeteksi pasien gagal jantung meningkat signifikan.

Tingkat akurasi prediksi model pada pengujian data individu menunjukkan kemampuan model yang konsisten dalam membedakan pasien gagal jantung dan non-gagal jantung. Meskipun ada kesalahan pada kasus borderline, probabilitas prediksi yang dihasilkan mengindikasikan keyakinan tinggi model pada sebagian besar kasus.

B. Saran

Untuk pengembangan lebih lanjut, disarankan untuk melakukan optimasi nilai threshold klasifikasi guna menyeimbangkan Recall dan Precision, khususnya dalam mengurangi risiko false negative yang fatal. menyeimbangkan Recall dan Precision demi mengurangi false negative. Disarankan pula penambahan fitur klinis pada dataset, validasi model dengan data eksternal untuk generalisasi, integrasi ke CDSS, serta

eksplorasi algoritma alternatif seperti LightGBM atau deep learning.

DAFTAR RUJUKAN

- ANINDITA, M. AND MARGARINI, E., 2021. Peringatan Hari Jantung Sedunia 2021: Jaga Jantungmu untuk Hidup Lebih Sehat. [online] Available at: <<https://ayosehat.kemkes.go.id/peringatan-hari-jantung-sedunia-2021-jaga-jantungmu-untuk-hidup-lebih-sehat>>.
- FRANSISKUS SITANGGANG, B. AND SITOMPUL, P., 2024. Menggunakan Machine Learning Metode Random Forest. *INNOVATIVE: Journal Of Social Science Research*, 4, pp.3347-3357.
- RIZKY MUBAROK, M. AND HERTENO, R., 2022. Hyper-Parameter Tuning Pada Xgboost Untuk Prediksi Keberlangsungan Hidup Pasien Gagal Jantung. [online] Available at: <<https://klik.ulm.ac.id/index.php/klik/article/view/484>> [Accessed 18 August 2025].
- SYUKRON, A., SAPUTRO, E. AND WIDODO, P., 2023. Penerapan Metode Smote Untuk Mengatasi Ketidakseimbangan Kelas Pada Prediksi Gagal Jantung. [online] *Jurnal Teknologi Informasi dan Terapan (J-TIT)*, Available at: <<https://doi.org/10/25047/jtit.v10i1.312>>.
- SYUKRON, M., SANTOSO, R. AND WIDIHARIH, T., 2020. Perbandingan Metode Smote Random Forest Dan Smote Xgboost Untuk Klasifikasi Tingkat Penyakit Hepatitis C Pada Imbalance Class Data. [online] Available at: <<https://ejournal3.undip.ac.id/index.php/gaussian/>>.
- VERIZARIE, R., 2020. Gagal Jantung: Jenis, Gejala, Penyebab, Pengobatan, dll - DokterSehat. [online] Available at: <<https://doktersehat.com/penyakit-a-z/gagal-jantung/>> [Accessed 11 February 2025].
- WU, J., CHEN, X.Y., ZHANG, H., XIONG, L.D., LEI, H. AND DENG, S.H., 2019. Hyperparameter optimization for machine learning models based on Bayesian optimization. *Journal of Electronic Science and Technology*, 17(1), pp.26-40. <https://doi.org/10.11989/JEST.1674-862X.80904120>.