STORAGE – Jurnal Ilmiah Teknik dan Ilmu Komputer, Vol. 5 No. 3, Agustus 2026, 302 – 311
e-ISSN: 2828-5344 | DOI: 10.55123

ANALISIS PERBANDINGAN PERFORMA ALGORITMA RANDOM FOREST, XGBOOST, DAN K-NEAREST NEIGHBOR DALAM KLASIFIKASI PENYAKIT JANTUNG

[1)]Risma Silitonga, [2)]Deni Risdiansyah, [3)]Erni
[1,2,3)]Informatika, Fakultas Teknik Informatika, Universitas Bina Sarana Informatika Pontianak
[1)] risma.sltnga@gmail.com, [2)]deni.drx@bsi.ac.id, [3)]erx@bsi.ac.id

INFO ARTIKEL

Riwayat Artikel :
Diterima : 10 Juni 2026
Disetujui : 22 Juni 2026

Kata Kunci :
Penyakit Jantung, Random Forest, XGBoost, K-Nearest Neighbor, Optuna.

ABSTRAK

Penyakit jantung merupakan salah satu penyebab kematian tertinggi di dunia sehingga diperlukan metode klasifikasi berbasis data yang mampu mendukung proses *skrining* awal secara efektif. Berbagai algoritma *machine learning* telah digunakan dalam klasifikasi penyakit jantung, namun masih terbatas penelitian yang membandingkan performa Random Forest, XGBoost, dan K-Nearest Neighbor (K-NN) dengan penanganan ketidakseimbangan data serta optimasi *hyperparameter* secara sistematis melalui *ablation study*. Penelitian ini bertujuan untuk menganalisis dan membandingkan kinerja ketiga algoritma tersebut dalam klasifikasi penyakit jantung menggunakan Heart Failure Prediction Dataset dari Kaggle yang terdiri dari 918 sampel dan 12 atribut. Tahapan penelitian meliputi *preprocessing* data (imputasi, encoding, deteksi *outlier* menggunakan Isolation Forest), feature engineering, pembagian data dengan rasio 80:20, normalisasi dengan MinMaxScaler, seleksi fitur menggunakan Information Gain, penyeimbangan data menggunakan *Random Oversampling*, serta optimasi *hyperparameter* menggunakan Optuna berbasis *Tree-structured Parzen Estimator* (TPE). Evaluasi model dilakukan menggunakan metrik Accuracy, Precision, Recall, F1-Score, dan Area Under Curve (AUC) melalui empat skenario *ablation study:* model dasar (S1), model dengan Random Oversampling (S2), model dengan Optuna (S3), dan model dengan kombinasi keduanya (S4). Hasil penelitian menunjukkan bahwa optimasi Optuna menghasilkan peningkatan yang lebih konsisten dibandingkan kombinasi Optuna dengan *Random Oversampling*. Random Forest pada Skenario 3 memperoleh kinerja terbaik dengan akurasi 89,14%, F1-Score 90,73%, dan AUC 94,95%. Hasil ini menunjukkan bahwa Optuna efektif meningkatkan kinerja klasifikasi, dengan Random Forest sebagai algoritma terbaik untuk klasifikasi penyakit jantung pada dataset yang digunakan sebagai sistem pendukung skrining awal.

ARTICLE INFO

Article History :
Received : Jun 10, 2026
Accepted : Jun 22, 2026

Keywords:
Heart Disease, Random Forest,XGBoost, K-Nearest Neighbor, Optuna.

ABSTRACT

Heart disease is one of the highest causes of death in the world, so a data-based classification method is needed that is able to support the initial screening process effectively. Various machine learning algorithms have been used in heart disease classification, but there is still limited research comparing the performance of Random Forest, XGBoost, and K-Nearest Neighbor (K-NN) with handling data imbalances and systematic hyperparameter optimization through ablation studies. This

study aims to analyze and compare the performance of the three algorithms in classifying heart disease using the Heart Failure Prediction Dataset from Kaggle which consists of 918 samples and 12 attributes. The research stages include data preprocessing (imputation, encoding, outlier detection using Isolation Forest), feature engineering, data division with a ratio of 80:20, normalization with MinMaxScaler, feature selection using Information Gain, data balancing using Random Oversampling, and hyperparameter optimization using Optuna based on Tree-structured Parzen Estimator (TPE). Model evaluation was carried out using Accuracy, Precision, Recall, F1-Score, and Area Under Curve (AUC) metrics through four ablation study scenarios: basic model (S1), model with Random Oversampling (S2), model with Optuna (S3), and model with a combination of both (S4). The research results show that Optuna optimization produces more consistent improvements compared to the combination of Optuna with Random Oversampling. Random Forest in Scenario 3 obtained the best performance with an accuracy of 89.14%, F1-Score 90.73%, and AUC 94.95%. These results show that Optuna is effective in improving classification performance, with Random Forest as the best algorithm for classifying heart disease on the dataset used as an initial screening support system.

1. PENDAHULUAN

Penyakit jantung mencakup serangkaian kondisi klinis yang meliputi berbagai gangguan fungsi jantung termasuk penyakit kardiovaskular, penyakit jantung koroner, dan serangan jantung (Hoki, Yuliana, and Robet 2026) Penyakit kardiovaskular adalah penyebab utama kematian di dunia, dengan 19,8 juta kematian pada tahun 2022 atau sekitar 32% dari total kematian di dunia 85% di antaranya disebabkan oleh serangan jantung dan stroke. Lebih dari tiga perempat kasus ini terjadi di negara-negara berpenghasilan rendah hingga menengah (WHO 2025). Di Indonesia, peningkatan penyakit jantung tidak lagi terbatas pada kelompok lanjut usia, melainkan juga menyerang usia produktif, dipicu oleh pola makan tidak sehat, kurangnya aktivitas fisik, tekanan psikologis, obesitas, serta kebiasaan merokok (Aini, Fitriani, and Awwalin 2026). Kondisi ini semakin mengkhawatirkan karena banyak kasus baru terdeteksi ketika penyakit telah memasuki fase komplikasi serius karena terbatasnya akses ke layanan kesehatan yang berfokus pada pencegahan dan deteksi dini (Febrianti, Hermanto, and Sunandar 2025).

Identifikasi dini penyakit jantung memegang peran vital dalam menekan angka mortalitas sekaligus meringankan beban sistem layanan kesehatan. Pendekatan diagnostik tradisional menghadapi berbagai kendala, antara lain durasi pemeriksaan yang panjang, biaya yang tidak sedikit, serta ketergantungan pada interpretasi subjektif tenaga medis (Jumardin and Noprisson 2024). Tantangan mendasar dalam penanganan penyakit jantung terletak pada sulitnya melakukan *skrining* awal secara cepat dan akurat. Mayoritas individu tidak mengetahui kondisi kesehatan jantung mereka karena penyakit ini kerap tidak menunjukkan gejala yang jelas pada fase awal sampai komplikasi serius terjadi (Setiawan et al. 2025).Keadaan ini mendorong kebutuhan untuk menggunakan cara yang lebih efektif, salah satunya dengan menggunakan teknik pembelajaran mesin.

Oleh karena itu, teknik *Machine Learning* menawarkan potensi besar sebagai sistem pendukung *skrining.* Teknik pembelajaran mesin dapat secara otomatis menganalisis data klinis yang kompleks dan membuat prediksi berdasarkan pola yang ada dalam rekam medis (Baghdadi et al. 2023). Melalui pendekatan ini, profesional medis dapat mengalokasikan sumber daya secara lebih tepat sasaran kepada pasien yang memerlukan evaluasi klinis lanjutan. Ketidakseimbangan kelas terjadi saat proporsi antar kelas dalam dataset tidak

proporsional, sehingga model cenderung memihak kelas mayoritas dan menghasilkan akurasi global yang tampak tinggi namun lemah dalam mendeteksi kelas minoritas yang justru kritis secara klinis(Agyemang et al. 2025).

Random Oversampling (ROS) merupakan metode yang digunakan untuk menangani ketidakseimbangan distribusi kelas dengan cara memperbanyak sampel dari kelas minoritas secara acak hingga proporsi antar kelas menjadi setara(Wongvorachan, He, and Bulut 2023).

Selain itu, kinerja model *machine learning* turut ditentukan oleh konfigurasi *hyperparameter,* yaitu sekumpulan parameter yang harus ditentukan sebelum proses pelatihan dimulai dan tidak dapat diperoleh secara langsung dari data latih (Bischl et al. 2023). Optuna adalah kerangka kerja optimasi *hyperparameter* otomatis yang berbasis pada *Tree structured Parzen Estimator* (TPE) yang bekerja dengan pendekatan optimasi Bayesian untuk menjelajahi ruang *hyperparameter* secara lebih efisien (Bischl et al. 2023). Sejumlah penelitian telah mengkaji klasifikasi penyakit jantung menggunakan beragam algoritma *Machine Learning* yang dikombinasikan dengan teknik penanganan ketidakseimbangan data dan strategi optimasi *hyperparameter*.

Namun demikian, variasi dalam pendekatan metodologis, pilihan dataset, teknik penyeimbangan, dan metode optimasi masih menjadi pembeda utama di antara studi-studi yang ada. Oleh karena itu, tinjauan sistematis terhadap penelitian sebelumnya diperlukan untuk mengidentifikasi kesenjangan yang belum teratasi.Pemetaan sistematis penelitian sebelumnya disajikan pada tabel 1 sebagai berikut:

Tabel 1. State of the Art

Penulis (Tahun)	Dataset & Algoritma	Imbalance / Optimasi	Hasil Utama	Limitations
(Amritha et al. 2026)	Kaggle 918 sampel RF, SVM, DT	Tidak Ada / Grid Search	RF: AUC=0,9517 SVM: AUC=0,9130 DT: AUC=0,8513	Tanpa KNN, tanpa penyeimbangan kelas
(Beny et al. 2026)	Kaggle 918 sampel RF, AdaBoost, XGBoost	Tidak Ada / Tidak Ada	RF: Acc=86,8%, AUC=92,2% XGB: Acc=84,8%	Tanpa penyeimbangan Kelas,tanpa optimasi
(Caroline and Rachmat 2025)	UCI 303 spl XGBoost, LightGBM	SMOTE / RandomizedCV	LGB: Acc=82,85% AUC=90,37% XGB: Prec=85,46%	Hanya 2 algoritma; dataset kecil
(Hutagalung and Andrianingsih 2026)	UCI 303 sampel RF, XGBoost +PSO	Tidak Ada / PSO, Bayesian	RF+PSO: AUC=90,89% F1=81,88% XGB: AUC=90,84%	Tanpa penyeimbangan kelas; tanpa KNN
(Andi et al. 2026)	UCI 303 spl LR, SVM, KNN, GB	Tidak Ada / Grid/Random/ Bayes	LR&SVM: Acc=88,33% KNN: Acc=86,7%	Tanpa penyeimbangan kelas; hanya akurasi

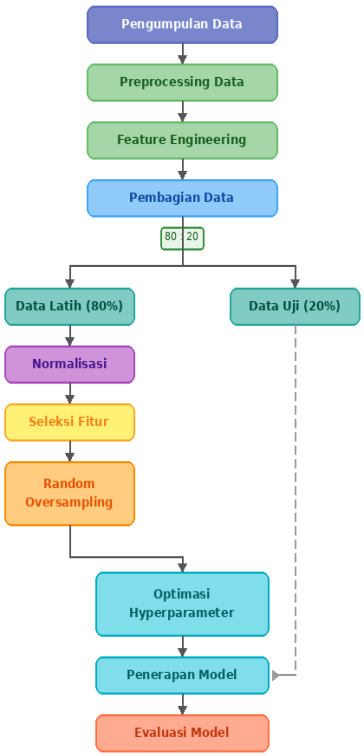
Kajian Tabel 1 menunjukkan bahwa tiga kesenjangan penelitian (*research gap*) belum dibahas secara menyeluruh dalam literatur yang ada. Pertama, banyak penelitian belum melihat pengaruh random oversampling secara terpisah sebagai skenario eksperimen mandiri. Akibatnya, kontribusi teknik penyeimbangan kelas yang terpisah terhadap masing-masing algoritma belum dapat diukur secara akurat. Kedua, belum ada penelitian yang memeriksa penggunaan Optuna sebagai teknik optimasi hyperparameter secara terpisah dari metode lain. Akibatnya, belum ada cara yang jelas untuk mengetahui seberapa besar kontribusi optimasi ini terhadap peningkatan kinerja model. Ketiga, belum ada penelitian yang membandingkan Random Forest, XGBoost, dan K-Nearest Neighbor secara bersamaan dalam empat skenario eksperimen terstruktur *(ablation study*) pada dataset Heart Failure Prediction 918 sampel.

Kebaruan penelitian ini adalah evaluasi terstruktur pengaruh *Random Oversampling* dan Optuna terhadap tiga algoritma melalui empat skenario *ablation study* dengan prosedur ketat pencegahan *data leakage* menggunakan dataset 918 sampel yang lebih besar dibandingkan studi terdahulu.

2. METODE

Penelitian ini menggunakan pendekatan kuantitatif dengan menggunakan metode analisis perbandingan untuk mengevaluasi kinerja tiga algoritma klasifikasi yaitu Random Forest, XGBoost, dan K-Nearest Neighbor dalam

memprediksi penyakit jantung.Tahapan dalam penelitian ini dapat dilihat pada gambar 1.



Gambar 1. Kerangka Tahapan Penelitian

1. Pengumpulan Data

Dataset yang digunakan dalam penelitian ini adalah Heart Failure Prediction Dataset yang dipublikasikan oleh fedesoriano di Kaggle. Dataset ini terdiri dari 918 sampel dengan 12 atribut. Dengan 1 atribut sebagai label target, pada dataset yang digunakan terdapat 410 sampel merupakan pasien tidak sakit jantung (kelas 0), dan 508 sampel merupakan pasien sakit jantung (kelas 1). Dataset ini merupakan integrasi dari lima dataset kardiovaskular internasional, yaitu Cleveland, Hungaria, VA Long Beach, Swiss, dan Stalog.

2. *Preprocessing* Data

Tahap *preprocessing* bertujuan meningkatkan kualitas data sebelum tahap pelatihan model dilaksanakan. Seluruh proses *preprocessing* dilakukan sebelum pembagian dataset guna memastikan data yang dibagi telah bersih dari *outlier*. Tahapan preprocessing yang digunakan dalam penelitian ini meliputi:

Imputasi (Penanganan *Missing Value*) Dalam studi ini, penanganan *missing value* dilakukan. Terdapat dua atribut dengan nilai yang tidak lengkap pada dataset yang digunakan, yaitu Cholesterol dan RestingBP, yang ditangani menggunakan imputasi nilai median secara keseluruhan. Adapun jumlah pada dataset yang *missing value* dapat dilihat pada tabel 1 sebagai berikut:

Tabel 2. Penanganan *Missing Value*

Atribut	Missing Value	Metode Imputasi
Cholesterol	172	Median
RestingBP	1	Median

Konversi data kategorikal ke format numerik dilakukan agar dapat diproses oleh algoritma machine learning. Tiga skema encoding diterapkan: binary encoding untuk atribut biner (Sex, ExerciseAngina), ordinal encoding untuk atribut berurutan (ST_Slope, RestingECG), dan one-hot encoding untuk atribut nominal (ChestPainType). Pemilihan skema ini bertujuan mempertahankan struktur informasi kategori agar model dapat mengekstrak pola secara lebih efektif.

Penghapusan *Outlier(Isolation Forest)* penghapusan outlier dilakukan dengan metode *Isolation Forest* guna mengidentifikasi data yang menyimpang secara signifikan dari distribusi mayoritas. Metode ini dipilih karena bersifat *model-based* dan mampu mendeteksi anomali pada data berdimensi tinggi secara otomatis tanpa asumsi distribusi tertentu. Hasil deteksi mengungkapkan terdapat 46 *outlier* yang kemudian dihilangkan sehingga menghasilkan data yang lebih bersih.

3. *Feature Engineering* (Rekayasa Fitur)

Rekayasa fitur dilakukan untuk menghasilkan fitur baru berdasarkan kombinasi atribut yang sudah ada. Pada penelitian ini ditambah lima fitur yang bertujuan untuk memperkaya informasi yang digunakan dalam proses klasifikasi sehingga model dapat lebih baik mengenali pola penyakit jantung. Fitur tambahan beserta penjelasannya dapat dilihat pada tabel berikut:

Tabel 3. Penambahan Fitur

No	Nama Fitur	Formula	Keterangan
1	Age_MaxHR_ratio	Age/(MaxHR+1)	Rasio usia vs detak jantung maks
2	ST_Angina_risk	Oldpeak x ExerciseAngina	Interaksi depresi ST dan angina
3	Age_DM_risk	Age x FastingBS	Risiko diabetes berbasis usia

4	Cardio_Stress	(Age/(MaxHR+1)) x(Oldpeak+1)	Indeks stres kardiovaskular
5	BP_Age_norm	RestingBP / Age	Tekanan darah relatif usia

4. Pembagian Data

Setelah data melalui tahap *preprocessing* (imputasi, encoding, deteksi outlier) dan feature engineering, sebanyak 46 sampel teridentifikasi sebagai outlier dihapus, sehingga total data bersih berjumlah 872 sampel. Dataset kemudian dibagi menjadi data pelatihan sebanyak 697 sampel (80%) dan data uji sebanyak 175 sampel (20%) menggunakan *stratified split*. Rasio 80:20 dipilih untuk memaksimalkan jumlah data pelatihan mengingat ukuran dataset yang terbatas pasca penghapusan outlier.

5. Normalisasi

Normalisasi dilakukan menggunakan metode *MinMaxScaler* untuk mentransformasi nilai atribut numerik ke dalam rentang [0, 1]. Scaler di-fit hanya pada data latih kemudian diterapkan pada data uji untuk mencegah kebocoran data.

6. Seleksi Fitur

Seleksi fitur dilakukan pada data latih untuk mencegah kebocoran data. Metode yang digunakan adalah Information Gain dengan nilai ambang batas 0,01. Atribut dengan skor Information Gain di bawah ambang tersebut dieliminasi karena dinilai tidak memberikan kontribusi informatif yang signifikan bagi proses klasifikasi. Hasil seleksi menghasilkan 18 fitur terpilih dari total 19 fitur setelah feature engineering, dengan mengeliminasi satu fitur yaitu Age_DM_risk yang memiliki nilai Information Gain di bawah ambang batas.

7. Random Oversampling

Random oversampling diterapkan pada data latih guna menangani ketimpangan distribusi kelas sakit dan sehat. Dalam penelitian ini, Random Oversampling hanya diterapkan pada Skenario 2 (Model Oversampling) dan Skenario 4 (Model Optuna dengan *Oversampling*).

Oversampling bekerja dengan menduplikasi sampel secara acak dari kelas minoritas hingga jumlahnya setara dengan kelas mayoritas (Hadiansyah, Rozikin, and Fatchan 2024). Setelah proses oversampling, jumlah data latih meningkat dari 697 sampel (kelas 0: 308, kelas 1: 389) menjadi 778 sampel dengan distribusi kelas yang seimbang (kelas 0: 389, kelas 1: 389).

8. Optimasi *Hyperparameter*

Optimasi *hyperparameter* menggunakan metode Optuna. Optimasi hyperparameter diterapkan pada Skenario 3 (Model Optuna) dan Skenario 4 (Model Optuna dengan *Oversampling*).

Optuna merupakan pustaka optimasi *hyperparameter* yang menggunakan pendekatan *define by run,* sehingga ruang pencarian hyperparameter dapat dikonstruksi secara dinamis dan adaptif selama proses optimisasi berlangsung(Lai et al. 2024).

Tabel 4
Ruang Pencarian Dan Nilai Terbaik Random Forest (400 Trial, Seed=42)

Hyperparameter	Ruang Pencarian	Nilai Terbaik
n_estimators	[300–2500, step 100]	2.200
max_depth	[5–40]	36
max_features	[sqrt, log2, 0.3–0.7]	log2
class_weight	[balanced, balanced_subsample, None]	balanced_subsample
criterion	[gini, entropy]	gini
min_samples_split	[2–8]	2
min_samples_leaf	[1–3]	1
max_samples	[0.7–1.0]	0,9946

Random Forest menggunakan 400 trial (seed=42) karena memiliki ruang pencarian terluas (8 hyperparameter). Jumlah trial yang lebih banyak diperlukan agar Optuna dapat mengeksplorasi ruang pencarian secara menyeluruh.

Tabel 5
Ruang Pencarian Dan Nilai Terbaik XGBoost (300 Trial, Seed=42)

Hyperparameter	Ruang Pencarian	Nilai Terbaik
n_estimators	[300–2000, step 100]	1.000
max_depth	[3–8]	8
learning_rate	[0,005–0,10]	0,0438
subsample	[0,65–1,0]	0,8449
gamma	[0,0–0,8]	0,1213
reg_alpha	[0,0–0,5]	0,1867
min_child_weight	[1–5]	2
colsample_bytree	[0,65–1,0]	0,6900
colsample_bylevel	[0,65–1,0]	0,8606
reg_lambda	[0,5–3,0]	0,6159
max_delta_step	[0–5]	2

XGBoost menggunakan 300 trial (seed=42) dengan regularisasi berlapis (gamma, reg_alpha, reg_lambda) untuk mencegah overfitting. Learning rate kecil (0,0438) dikompensasi dengan jumlah pohon yang besar (1.000).

Tabel 6
Ruang Pencarian Dan Nilai Terbaik K-NN (150 Trial, Seed=42)

Hyperparameter	Ruang Pencarian	Nilai Terbaik
n_neighbors	[1–30]	30

Hyperparameter	Ruang Pencarian	Nilai Terbaik
metric	[euclidean, manhattan, chebyshev, minkowski]	manhattan
weights	[uniform, distance]	distance
algorithm	[auto, ball_tree, kd_tree, brute]	ball_tree
leaf_size	[10–60]	44
p	[1–4]	1

9. Penerapan Model

Penelitian ini mengaplikasikan algoritma Random Forest, XGBoost, dan K-Nearest Neighbor (K-NN) untuk mengklasifikasikan penyakit jantung. Adapun pengertian dari masing-masing algoritma sebagai berikut:

Random Forest adalah algoritma pembelajaran mesin berbasis *ensemble* yang membangun banyak pohon keputusan secara acak dan kemudian menggunakan suara mayoritas dari setiap pohon untuk mendapatkan hasil prediksi akhir (Sumwiza et al. 2023).

XGBoost (Extreme Gradient Boosting) merupakan algoritma pembelajaran mesin yang membangun pohon keputusan secara bertahap, dimana setiap pohon baru bertugas memperbaiki kesalahan prediksi pohon sebelumnya hingga menghasilkan model yang akurat (Budholiya, Shrivastava, and Sharma 2022).

K-Nearest Neighbor (K-NN) merupakan algoritma pembelajaran mesin non-parametrik yang mengklasifikasikan data baru berdasarkan kemiripannya dengan sejumlah k data terdekat dalam ruang fitur, yang kelasnya ditentukan melalui suara terbanyak dari tetangga terdekat(Khan et al. 2024).

10. Evaluasi Model

Kinerja model dievaluasi menggunakan lima metrik yaitu Akurasi, Presisi, Recall, F1-Score, dan Area Under Curve (AUC). Berikut rumus dari masing-masing metrik evaluasi:

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

$$\text{Precision} = \frac{TP}{TP+FP} \quad (2)$$

$$\text{Recall} = \frac{TP}{TP+FN} \quad (3)$$

$$\text{F1-Score} = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (4)$$

Akurasi digunakan untuk mengukur seberapa tepat suatu model mengklasifikasikan data (Aryanti, Misriati, and Hidayat 2023).

Presisi untuk menilai keakuratan prediksi yang positif.

Recall mengukur seberapa akurat model mengidentifikasi semua kasus positif (Hakim et al. 2025).

Sedangkan F1-Score digunakan untuk menilai keseimbangan antara Precision dan Recall.

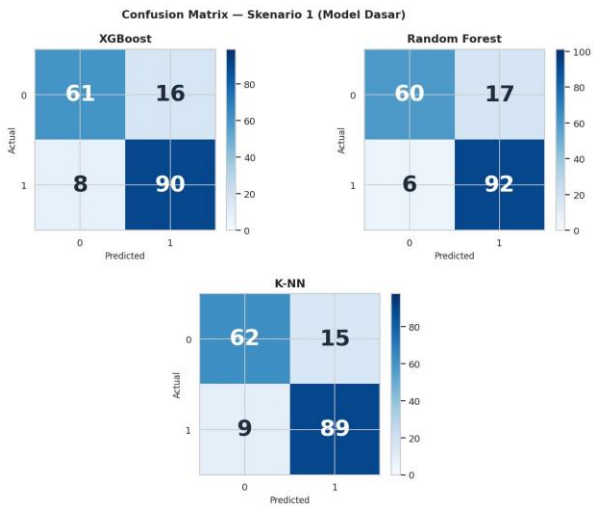
3. HASIL DAN PEMBAHASAN

Berdasarkan eksperimen yang menerapkan empat skenario ablasi, studi ini mengevaluasi kinerja Random Forest, XGBoost, dan K-Nearest Neighbor melalui empat tahap: model dasar (Skenario 1), model dengan Random Oversampling (Skenario 2), model dengan optimasi Optuna (Skenario 3), dan kombinasi keduanya (Skenario 4). Semua evaluasi dilakukan menggunakan metrik Akurasi, Presisi, Recall, F1-Score, dan AUC. Berikut adalah hasil dari keempat skenario tersebut:

3.1. Skenario 1 Model Dasar (*Default*)

Tabel 1
Hasil Skenario 1 Model Dasar

Model	Acc.	Prec.	Recall	F1	AUC
Random Forest	86,86 %	84,40 %	93,88 %	88,89 %	94,76 %
XGBoost	86,29 %	84,91 %	91,84 %	88,24 %	94,35 %
K-NN	86,29 %	85,58 %	90,82 %	88,12 %	93,09 %



Gambar 1. Confusion Matrix Skenario 1 (Model Dasar)

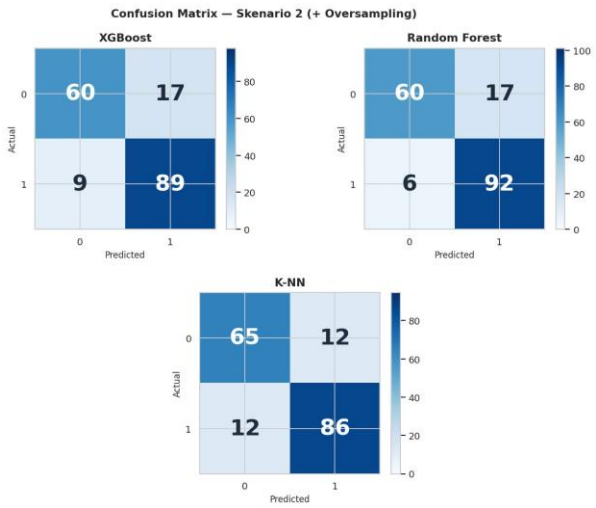
Pada skenario model dasar, Random Forest berhasil meraih akurasi dan recall tertinggi. Hal ini disebabkan oleh mekanisme *bagging* yang menstabilkan prediksi melalui agregasi output dari ratusan pohon keputusan, sehingga varians model dapat ditekan secara signifikan. XGBoost menunjukkan performa

yang kompetitif melalui mekanisme gradient boosting yang memperbaiki kesalahan prediksi secara bertahap. K-NN memperoleh AUC terendah karena sensitif terhadap ketidakseimbangan distribusi kelas dalam ruang fitur. Secara keseluruhan, ketiga model menunjukkan kemampuan dasar yang baik namun masih terdapat ruang peningkatan melalui optimasi lebih lanjut.

3.2. Skenario 2 Model dengan *Oversampling*

Tabel 2.
Hasil Skenario 2 (Model *Oversampling)*

Model	Acc.	Prec.	Rec.	F1	AUC
Random Forest	86,86 %	84,40 %	93,88 %	88,89 %	93,86 %
XGBoost	85,14 %	83,96 %	90,82 %	87,25 %	93,20 %
K-NN	86,29 %	87,76 %	87,76 %	87,76 %	92,17 %



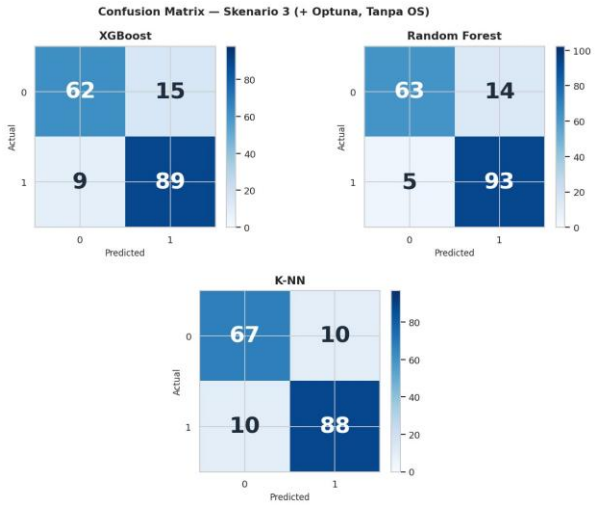
Gambar 2. Confusion Matrix Skenario 2
(Model dengan *Oversampling*)

Penerapan random *oversampling* tidak menghasilkan peningkatan yang merata pada semua algoritma. Hal ini disebabkan karena teknik ini hanya menduplikasi sampel yang telah ada tanpa menambah keragaman data. Random Forest mempertahankan akurasi dan F1-Score, tetapi mengalami penurunan AUC dibandingkan dengan skenario dasar karena duplikasi sampel yang mengurangi variasi data pelatihan, sementara XGBoost mengalami penurunan pada beberapa metrik akibat pola data yang berulang dan tidak variatif. Sementara itu, K-NN mengalami peningkatan presisi karena distribusi kelas yang lebih seimbang, namun nilai recallnya menurun.

3.3. Skenario 3 Model dengan Optuna

Tabel 3. Hasil Skenario 3 (Model Optuna)

Model	Acc.	Prec.	Rec.	F1	AUC
Random Forest	89,14 %	86,92 %	94,90 %	90,73 %	94,95 %
XGBoost	86,29 %	85,58 %	90,82 %	88,12 %	94,14 %
K-NN	88,57 %	89,80 %	89,80 %	89,80 %	95,52 %



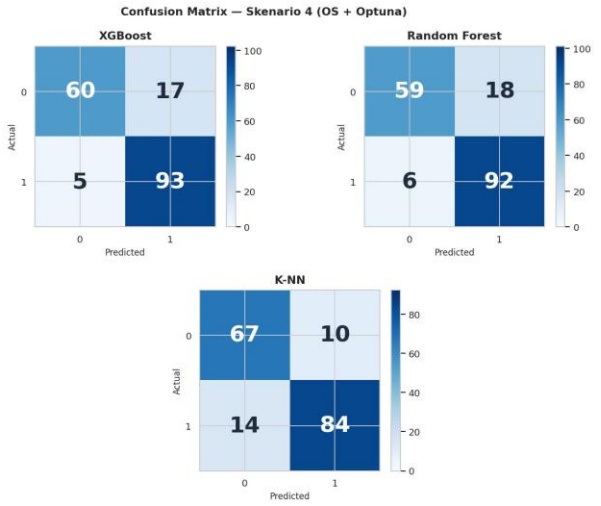
Gambar 3. Confusion Matrix Skenario 3
(Model dengan Optuna)

Optimasi menggunakan Optuna memberikan peningkatan kinerja yang lebih stabil dan konsisten dibandingkan *oversampling* karena pendekatan berbasis *Tree Structured Parzen Estimator* secara adaptif menyempitkan ruang pencarian berdasarkan distribusi probabilitas dari percobaan sebelumnya, sehingga konfigurasi optimal dapat ditemukan dengan lebih efisien.

3.4. Skenario 4 Model Optuna dengan Oversampling

Tabel 4. Skenario 4 Model Optuna dengan Oversampling

Model	Acc.	Prec.	Rec.	F1	AUC
Random Forest	86,29 %	83,64 %	93,88 %	88,46 %	94,40 %
XGBoost	87,43 %	84,55 %	94,90 %	89,42 %	93,04 %
K-NN	86,29 %	89,36 %	85,71 %	87,50 %	94,66 %



Gambar 4. Confusion Matrix Skenario 4 (Oversampling dengan Optuna)

Pada skenario 4, mekanisme gradient boosting XGBoost bekerja dengan baik dengan kombinasi *oversampling* dan Optuna dalam mendeteksi kelas positif, yang menghasilkan recall dan akurasi tertinggi dari ketiga model.

Dibandingkan dengan Skenario 3, kinerja Random Forest pada Skenario 4 menurun. Hal ini disebabkan oleh hasil *oversampling* duplikat yang mengurangi keragaman antar pohon dalam mekanisme bagging sehingga mencegah optimasi *hyperparameter* Optuna untuk bekerja secara optimal.

Performa K-NN pada Skenario 4 juga menurun dibandingkan dengan Skenario 3. Sebagai algoritma berbasis jarak, keberadaan sampel duplikat akibat *oversampling* menyebabkan batas keputusan kurang representatif terhadap distribusi data asli.

4. PENUTUP

4.1. Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan, optimasi hyperparameter menggunakan Optuna menghasilkan peningkatan kinerja yang lebih konsisten dibandingkan Random *Oversampling*.

Random Forest pada Skenario 3 (Optuna tanpa *Oversampling*) menghasilkan kinerja terbaik dengan akurasi 89,14%, F1-Score 90,73%, dan AUC 94,95%. Sementara itu, penambahan Random Oversampling tidak selalu meningkatkan performa karena pada beberapa skenario justru menurunkan metrik tertentu akibat pola data duplikat yang berulang memperkenalkan bias distribusi.

Hasil ini mengindikasikan bahwa pemilihan strategi penyeimbangan data dan optimasi hyperparameter perlu disesuaikan dengan karakteristik masing-masing algoritma. Model ini diposisikan sebagai alat bantu *skrining* awal penyakit jantung, sehingga tidak dapat menggantikan penilaian klinis dokter dan masih memerlukan validasi lebih lanjut sebelum diterapkan dalam praktik medis.

4.2. Saran

Penelitian selanjutnya disarankan menggunakan *stratified k-fold cross-validation* untuk estimasi performa yang lebih stabil, serta melakukan validasi pada dataset klinis dari sumber yang berbeda. Perbandingan dengan teknik penyeimbangan kelas lain seperti SMOTE atau ADASYN, penerapan interpretabilitas model menggunakan SHAP, dan eksplorasi algoritma alternatif seperti LightGBM atau *deep learning* juga dapat menjadi arah pengembangan yang bernilai.

4.3.Keterbatasan

Penelitian ini menggunakan *single train-test split* 80:20 tanpa *stratified k-fold cross-validation*, sehingga estimasi performa berpotensi bervariasi pada partisi data yang berbeda. Dataset yang digunakan berasal dari satu sumber publik, sehingga generalisasi model terhadap data klinis nyata masih memerlukan validasi lebih lanjut.

5. DAFTAR PUSTAKA

Agyemang, Edmund Fosu, Joseph Agyapong Mensah, Eric Nyarko, Dennis Arku, Benedict Mbeah-Baiden, Enock Opoku, and Ezekiel Nii Noye Nortey. 2025. "Addressing Class Imbalance Problem in Health Data Classification: Practical Application From an Oversampling Viewpoint." *Applied Computational Intelligence and Soft Computing* 2025(1). doi: 10.1155/acis/1013769.

Aini, Siti Azizah, Diah Fitriani, and Mohammad Alif Awwalin. 2026. "Klasifikasi Penyakit Jantung Berdasarkan Faktor Klinis Menggunakan Algoritma XG Boost, Random Forest, Support Vector Machine (SVM) Dan K-Nearest Neighbor (KNN)." 10(2):3449–54.

Amritha, Yadhurani Dewi, Ni Luh, Putu Ika, and Wira Putra Dananjaya. 2026. "Model Machine Learning Yang Dioptimalkan Untuk Prediksi Penyakit Jantung Menggunakan R Shiny." 8(01):1–10.

Andi, Tri, Amelia Ritahani, Andri Pranolo, Candra Juni, and Cahyo Kusuma. 2026. "Classification of Heart Disease with Machine Learning : A Comparison of Grid Search , Random Search , and Bayesian Optimization." 10(January):262–70.

Aryanti, Riska, Titik Misriati, and Rahmat Hidayat. 2023. "Klasifikasi Risiko Kesehatan Ibu Hamil Menggunakan Random Oversampling Untuk Mengatasi Ketidakseimbangan Data." *KLIK: Kajian Ilmiah Informatika Dan Komputer* 3(5):409–16.

Baghdadi, Nadiah A., Sally Mohammed Farghaly Abdelaliem, Amer Malki, Ibrahim Gad, Ashraf Ewis, and Elsayed Atlam. 2023. "Advanced Machine Learning Techniques for Cardiovascular Disease Early Detection and Diagnosis." *Journal of Big Data* 10(1):1–29. doi: 10.1186/s40537-023-00817-1.

Beny, Beny, Herti Yani, Gangga Ramadhan, and Putra Yupu. 2026. "Telematika Performance Analysis of Ensemble Learning Models in Heart Failure Prediction : Random Forest , AdaBoost , and XGBoost." 19(1):1–10.

Bischl, Bernd, Martin Binder, Michel Lang, Tobias Pielok, Jakob Richter, Stefan Coors, Janek Thomas, Theresa Ullmann, Marc Becker, Anne Laure Boulesteix, Difan Deng, and Marius Lindauer. 2023. "Hyperparameter Optimization: Foundations, Algorithms, Best Practices, and Open Challenges." *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 13(2):1–43. doi: 10.1002/widm.1484.

Budholiya, Kartik, Shailendra Kumar Shrivastava, and Vivek Sharma. 2022. "An Optimized XGBoost Based Diagnostic System for Effective Prediction of Heart Disease." *Journal of King Saud University - Computer and Information Sciences* 34(7):4514–23. doi: 10.1016/j.jksuci.2020.10.013.

Caroline, Fionna, and Nur Rachmat. 2025. "Comparison of XGBoost and LightGBM Algorithms in Predicting Heart Disease." 5(2):1232–39.

Febrianti, Nisa, Teguh Iman Hermanto, and Muhamad Agus Sunandar. 2025. "Diagnosa Penyakit Jantung Berdasarkan Kondisi Tubuh Dengan Metode Artificial Neural Network." 15(2):408–14.

Hadiansyah, Zikri, Zaenur Rozikin, and Muhamad Fatchan. 2024. "Implementasi Algoritma K-Nearest Neighbor Dalam Klasifikasi Penyakit Kanker Paru Paru." *Journal of Computer System and Informatics (JoSYC)* 6(1):96–106. doi: 10.47065/josyc.v6i1.6195.

Hakim, Lukman, Ahmad Sobri, Lukman Sunardi, and Deni Nurdiansyah. 2025. "Prediksi Penyakit Jantung Berbasis Mesin Learning Dengan Menggunakan Metode K-Nn." *Jurnal Digital Teknologi Informasi* 7(2):14. doi: 10.32502/digital.v7i2.9429.

Hoki, Leony, Yuliana, and Robet. 2026. "Comparative Analysis of XGBoost , KNN , and SVM Algorithms for Heart Disease Prediction Using SMOTE-Tomek Balancing." 10(1):305–16.

Hutagalung, Pancar Hizkia, and Andrianingsih Andrianingsih. 2026. "Heart Disease Classification Using Optimised XGBoost and Random Forest with SHAP Explanations." *Sinkron* 10(1):330–42. doi: 10.33395/sinkron.v10i1.15544.

Jumardin, and Handrie Noprisson. 2024. "Perancangan Prototipe Aplikasi Prediksi Kematian Akibat Gagal Jantung Menggunakan Metode Machine Learning Berdasarkan Data Heart Failure Clinical Records." 7(3):724–29.

Khan, Muhammad Amir, Tehseen Mazhar, Muhammad Mateen Yaqoob, Muhammad Badruddin Khan, Abdul Khader Jilani Saudagar, Yazeed Yasin Ghadi, Umar Farooq Khattak, and Mohammad Shahid. 2024. "Optimal Feature Selection for Heart Disease Prediction Using Modified Artificial Bee Colony (M-ABC) and K-Nearest Neighbors (KNN)." *Scientific Reports* 14(1):1–15. doi: 10.1038/s41598-024-78021-1.

Lai, Li Hsing, Ying Lei Lin, Yu Hui Liu, Jung

Pin Lai, Wen Chieh Yang, Hung Pin Hou, and Ping Feng Pai. 2024. "The Use of Machine Learning Models with Optuna in Disease Prediction." *Electronics (Switzerland)* 13(23):1–20. doi: 10.3390/electronics13234775.

Setiawan, Dita, Ali Muhammad, Siti Herawati, and Fransiska Dewi. 2025. "Penerapan Algoritma Klasifikasi Untuk Deteksi Dini Penyakit Jantung Koroner Berdasarkan Gejala Klinis Koroner Berdasarkan Data Klinis Pasien Menggunakan Algoritma Pembelajaran Mesin [7]. Komputasional Berbasis Algoritma Pembelajaran Mesin Untuk Klasifik." 5:18–26.

Sumwiza, Kellen, Celestin Twizere, Gerard Rushingabigwi, Pierre Bakunzibake, and Peace Bamurigire. 2023. "Enhanced Cardiovascular Disease Prediction Model Using Random Forest Algorithm." *Informatics in Medicine Unlocked* 41(July):101316. doi: 10.1016/j.imu.2023.101316.

WHO. 2025. "Cardiovascular Diseases (CVDs)." Retrieved April 15, 2026 (https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)?).

Wongvorachan, Tarid, Surina He, and Okan Bulut. 2023. "A Comparison of Undersampling, Oversampling, and SMOTE Methods for Dealing with Imbalanced Classification in Educational Data Mining." *Information (Switzerland)* 14(1). doi: 10.3390/info14010054.