

Analisis Komparatif Genetic Algorithm dan Particle Swarm Optimization dalam Optimasi Ensemble Learning untuk Prediksi Coronary Artery Disease

Comparative Analysis of Genetic Algorithm and Particle Swarm Optimization in Ensemble Learning Optimization for Coronary Artery Disease Prediction

Rizal Wahyu Pratama¹, Khulika Malkan^{*2}, Mikhael Setia Budi³, Diah Septiani⁴

^{1,2,3,4} Program Studi Sains Data, Telkom University, Purwokerto, Jawa Tengah, Indonesia

e-mail: ^{*}khulika@student.telkomuniversity.ac.id

Abstrak – Coronary Artery Disease (CAD) merupakan penyebab utama kematian kardiovaskular di seluruh dunia dan proses diagnosisnya masih sangat bergantung pada prosedur invasif yang tidak selalu mudah diakses oleh semua pasien. Penelitian ini membandingkan kinerja Genetic Algorithm (GA) dan Particle Swarm Optimization (PSO) dalam melakukan seleksi fitur sekaligus optimasi hyperparameter secara simultan pada empat model ensemble learning untuk prediksi CAD menggunakan dataset Z-Alizadeh Sani yang memuat 303 data pasien dengan 56 atribut klinis. Kedua algoritma dikonfigurasi secara setara menggunakan representasi solusi berdimensi 43 dengan fungsi fitness berbasis F1-score dan mekanisme early stopping. Hasil penelitian menunjukkan bahwa model terbaik dicapai oleh kombinasi GA dengan Random Forest yang meraih F1-score sebesar 0,9438, akurasi 0,9180, dan ROC-AUC 0,9018. PSO Random Forest dan PSO LightGBM berhasil mencapai recall sempurna 1,0000 sehingga tidak ada satu pun pasien CAD yang terlewat dalam proses klasifikasi. Analisis interpretabilitas menggunakan SHAP mengungkap bahwa Nyeri Dada Tipikal, Usia, dan Fraksi Ejeksi adalah faktor klinis yang paling berpengaruh terhadap prediksi model, konsisten dengan konsensus kardiologi internasional.

Kata kunci – Coronary Artery Disease; Genetic Algorithm; Particle Swarm Optimization; Ensemble Learning; SHAP

Abstract - Coronary Artery Disease (CAD) remains one of the leading causes of cardiovascular mortality worldwide, and its diagnosis still heavily relies on invasive procedures that are not always accessible to all patients. This study compares the performance of Genetic Algorithm (GA) and Particle Swarm Optimization (PSO) in simultaneously performing feature selection and hyperparameter optimization on four ensemble learning models for CAD prediction using the Z-Alizadeh Sani dataset containing 303 patient records with 56 clinical attributes. Both algorithms were configured equivalently using a 43-dimensional solution representation with an F1-score-based fitness function and early stopping mechanism. Results show that the best performance is achieved by the GA-optimized Random Forest model with an F1-score of 0.9438, accuracy of 0.9180, and ROC-AUC of 0.9018. PSO Random Forest and PSO LightGBM achieved perfect recall of 1.0000, successfully identifying all CAD patients without any false negatives. SHAP interpretability analysis reveals that Typical Chest Pain, Age, and Ejection Fraction are the most influential clinical factors, consistent with established cardiology consensus.

Keywords – Coronary Artery Disease; Genetic Algorithm; Particle Swarm Optimization; Ensemble Learning; SHAP

I. PENDAHULUAN

Coronary Artery Disease (CAD) masih menjadi salah satu penyebab kematian tertinggi di dunia hingga hari ini, dengan jutaan kasus baru yang terdeteksi setiap tahunnya di berbagai belahan dunia [1]. Kondisi ini terjadi ketika arteri koroner mengalami penyempitan akibat penumpukan plak aterosklerotik, sehingga pasokan darah dan oksigen ke otot jantung terganggu secara progresif [2]. Negara-negara berkembang termasuk Indonesia menghadapi tekanan ganda karena prevalensi faktor risiko seperti hipertensi, diabetes mellitus, dan gaya hidup tidak sehat terus meningkat di tengah masyarakat yang belum sepenuhnya memiliki akses terhadap layanan diagnostik kardiovaskular yang memadai.

Proses diagnosis CAD konvensional masih mengandalkan prosedur invasif seperti angiografi koroner, yang selain mahal juga memerlukan fasilitas khusus yang tidak tersedia secara merata [3]. Kemajuan teknologi informasi kesehatan membuka peluang besar untuk memanfaatkan data klinis pasien sebagai basis prediksi non-invasif yang lebih efisien dan terjangkau [4]. Data rekam medis, hasil laboratorium, serta temuan elektrokardiografi dan ekokardiografi mengandung pola tersembunyi yang dapat diekstrak oleh algoritma *machine learning* untuk mengidentifikasi risiko CAD sebelum kondisi memburuk secara klinis. Dataset Z-Alizadeh Sani yang memuat 303 rekam pasien dengan 56 atribut klinis menjadi salah satu *benchmark* paling banyak digunakan dalam riset prediksi CAD karena kekayaan informasinya yang mencakup berbagai dimensi pemeriksaan medis [5].

Metode ensemble learning telah terbukti menjadi pendekatan paling efektif dalam klasifikasi data medis berkat kemampuannya menggabungkan kekuatan beberapa model secara sinergis. Random Forest, XGBoost, LightGBM, dan CatBoost merepresentasikan dua paradigma utama dalam *ensemble learning*, yakni *bagging* dan *boosting*, yang masing-masing memiliki karakteristik unik dalam menangani kompleksitas data klinis [6]. Performa keempat model ini sangat bergantung pada konfigurasi *hyperparameter* yang digunakan, sementara pendekatan *tuning* konvensional seperti *grid search* dan *random search* memiliki keterbatasan efisiensi yang semakin terasa ketika ruang pencarian semakin luas dan kompleks [7]. Kebutuhan akan metode optimasi yang lebih adaptif mendorong eksplorasi algoritma metaheuristik seperti Genetic Algorithm dan Particle Swarm Optimization sebagai alternatif yang lebih cerdas dan efisien [8].

Genetic Algorithm terinspirasi dari proses seleksi alam Darwin, di mana populasi kandidat solusi berevolusi dari generasi ke generasi melalui operator seleksi, crossover, dan mutasi hingga menemukan konfigurasi yang semakin optimal [9]. Particle Swarm Optimization mengadopsi perilaku kawanan burung atau ikan dalam mencari sumber daya, di mana setiap partikel bergerak menuju posisi terbaik individu sekaligus posisi terbaik kawanannya secara kolektif [10]. Keduanya mampu melakukan seleksi fitur dan optimasi hyperparameter secara simultan hanya dengan merepresentasikan kombinasi keduanya dalam satu kromosom atau vektor partikel, menjadikan pendekatan ini sangat relevan untuk data berdimensi tinggi seperti dataset Z-Alizadeh Sani [11]. Perbandingan kedua algoritma dalam kondisi yang benar-benar setara masih sangat terbatas dalam literatur yang ada sehingga penelitian ini hadir untuk mengisi kesenjangan tersebut [12].

Aspek interpretabilitas menjadi semakin krusial dalam aplikasi klinis, di mana setiap keputusan prediksi harus dapat dipertanggungjawabkan secara medis kepada praktisi kesehatan [13]. Metode SHAP yang berlandaskan teori permainan Shapley mampu

menjelaskan kontribusi setiap fitur terhadap prediksi model secara lokal maupun global, memungkinkan verifikasi apakah faktor-faktor yang dianggap penting oleh model memang konsisten dengan pengetahuan klinis yang telah mapan [14]. Penelitian ini mengintegrasikan analisis SHAP pada model terbaik sebagai komponen interpretabilitas yang memperkuat nilai klinis dari hasil optimasi. Tujuan penelitian ini mencakup tiga hal utama yaitu membandingkan pengaruh GA dan PSO terhadap peningkatan performa ensemble learning, mengidentifikasi fitur klinis yang secara konsisten dipilih oleh kedua algoritma, dan menganalisis faktor klinis paling berpengaruh berdasarkan interpretabilitas SHAP pada dataset Z-Alizadeh Sani [15].

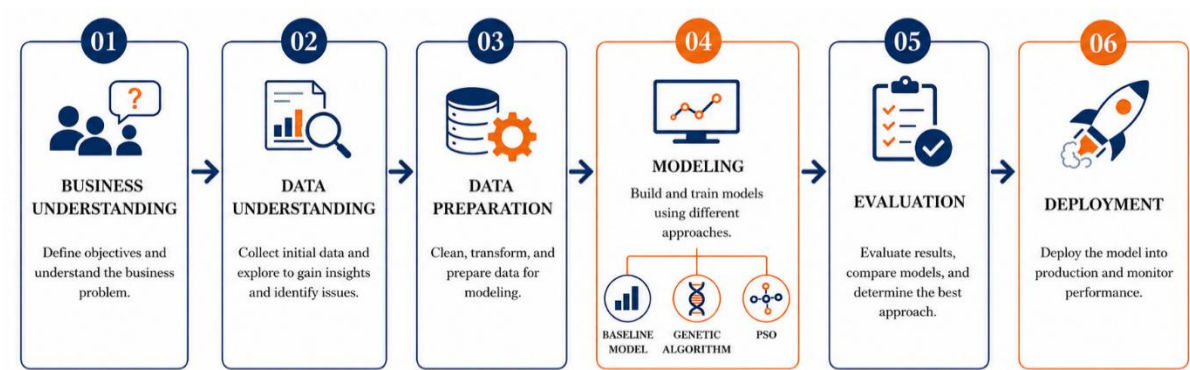
II. TINJAUAN TEORI

Prediksi penyakit jantung koroner berbasis machine learning telah berkembang pesat selama beberapa tahun terakhir dengan berbagai pendekatan yang terus disempurnakan. Kigka dkk. (2022) membangun model berbasis gradient boosting dengan eliminasi fitur rekursif pada 187 pasien dari dua uji klinis besar dan berhasil mencapai akurasi 0,81, menekankan bahwa tuning hyperparameter melalui *randomized search* berkontribusi signifikan terhadap hasil akhir [16]. Studi yang dilakukan Raji dkk. (2022) memformulasikan Genetic Algorithm berbasis seleksi deterministik pada dataset *benchmark* dan membuktikan bahwa GA mampu menemukan kombinasi *hyperparameter* optimal dengan efisiensi komputasi yang jauh lebih baik dibandingkan grid search konvensional [7]. Temuan ini memperkuat posisi GA sebagai alternatif yang layak untuk tuning pada dataset medis berskala kecil hingga menengah.

Abdo dkk. (2024) mencapai akurasi hingga 95,96 persen melalui kombinasi seleksi fitur berbasis Chi-Square dan Particle Swarm Optimization pada dataset *benchmark*, membuktikan bahwa PSO tidak hanya efektif untuk optimasi hyperparameter tetapi juga bekerja sebagai mesin seleksi fitur yang andal [17]. Mehday dkk. (2024) mendemonstrasikan lompatan performa yang signifikan dengan menggabungkan GA dan XGBoost pada data smart grid, di mana akurasi melonjak dari 0,82 menjadi 0,978 setelah optimasi GA diterapkan pada data tidak seimbang [9]. Çakmak dkk. (2026) mengembangkan perspektif yang lebih holistik dengan memasukkan faktor psikososial ke dalam model diagnostik CAD dari 300 pasien, menunjukkan bahwa kompleksitas faktor risiko penyakit ini melampaui sekadar parameter biokimia dan demografis [18].

Sultan dkk. (2025) membangun kerangka *stacking ensemble* dengan validasi silang K-lipatan yang dilengkapi SHAP, memberikan perhatian seimbang antara akurasi prediksi dan kemampuan interpretasi klinis yang dapat dipahami oleh tenaga medis [14]. Penelitian yang paling dekat relevansinya dilakukan oleh Tasmurzayev dkk. (2025) yang menerapkan XGBoost dengan interpretabilitas SHAP pada dataset Z-Alizadeh Sani yang sama, mencapai akurasi 0,9011 dan AUC 0,92, namun tanpa mengeksplorasi optimasi berbasis algoritma evolusioner maupun perbandingan antara GA dan PSO [12]. Alfath dkk. (2025) secara konsisten menunjukkan bahwa kombinasi CatBoost, XGBoost, dan LightGBM sebagai base learner melampaui performa model tunggal manapun dalam prediksi risiko hipertensi, memperkuat pilihan keempat model ensemble yang digunakan dalam penelitian ini [19].

III. METODE



Gambar 3. Alur Penelitian Berbasis Kerangka CRISP-DM

3.1 Eksplorasi dan Persiapan Data

Eksplorasi data mencakup analisis statistik deskriptif, pemeriksaan distribusi fitur, deteksi pencilan menggunakan metode Interquartile Range, analisis korelasi Pearson antar fitur numerik, dan pengukuran relevansi fitur terhadap target menggunakan Mutual Information Score. Hasil eksplorasi mengungkapkan bahwa Nyeri Dada Tipikal memiliki skor Mutual Information tertinggi sebesar 0,1737, jauh melampaui fitur lainnya, sementara 22 fitur mendapat skor nol yang mengindikasikan tidak adanya hubungan statistik terukur dengan target klasifikasi. Dua pasang fitur juga teridentifikasi memiliki korelasi tinggi yaitu Berat Badan dengan Indeks Massa Tubuh ($r=0,73$) dan Limfosit dengan Neutrofil ($r=-0,92$), yang menjadi dasar keputusan reduksi fitur pada tahap preprocessing.

Tabel 2. Keputusan Preprocessing dan Justifikasinya

Langkah	Fitur Terdampak	Justifikasi
Drop zero variance	Exertional CP	Hanya satu nilai unik pada seluruh sampel
Drop MI = 0	22 fitur (CHF, CVA, FH, dll.)	Tidak ada hubungan statistik dengan target
Drop multikolinear	Berat Badan, Limfosit	Korelasi tinggi dengan BMI (0,73) dan Neutrofil (-0,92)
Encoding ordinal	VHD, BBB, Jenis Kelamin	Fitur kategorik dengan hierarki klinis bermakna
Encoding biner	18 fitur Y/N	Konversi langsung Y=1, N=0
RobustScaler	10 fitur numerik kontinu	Tahan terhadap pencilan pada EDA

Tabel 2 merangkum seluruh keputusan *preprocessing* yang diambil berdasarkan bukti analitik dari tahap eksplorasi. Hasilnya adalah reduksi dari 55 fitur awal menjadi 32 fitur final yang siap digunakan untuk pemodelan, dengan pembagian data 80 persen untuk training (242 sampel) dan 20 persen untuk testing (61 sampel). Ketidakseimbangan kelas ditangani melalui penerapan *class weight balanced* pada seluruh model, sebuah pendekatan yang dipilih karena tidak menghasilkan data sintesis yang tidak memiliki dasar klinis nyata sehingga integritas data medis tetap terjaga sepenuhnya [20].

3.2 Pemodelan dan Optimasi

Penelitian ini membangun tiga pipeline yang berjalan secara paralel. Pipeline pertama adalah Baseline yang melatih keempat model ensemble dengan hyperparameter bawaan dan semua 32 fitur sebagai acuan performa awal. Pipeline kedua menggunakan Genetic Algorithm dengan kromosom berdimensi 43 yang mengkodekan keputusan seleksi fitur pada 32 gen pertama dan hyperparameter keempat model pada 11 gen berikutnya, di mana nilai gen fitur di atas 0,5 berarti fitur digunakan dan nilai hyperparameter didekode secara linear ke rentang yang telah ditetapkan. Pipeline ketiga menggunakan Particle Swarm Optimization dengan dimensi partikel yang identik dengan kromosom GA, sehingga satu-satunya variabel yang berbeda antara kedua pipeline adalah mekanisme pencariannya dan perbandingan yang dihasilkan bersifat adil [10].

Tabel 3. Konfigurasi GA dan PSO

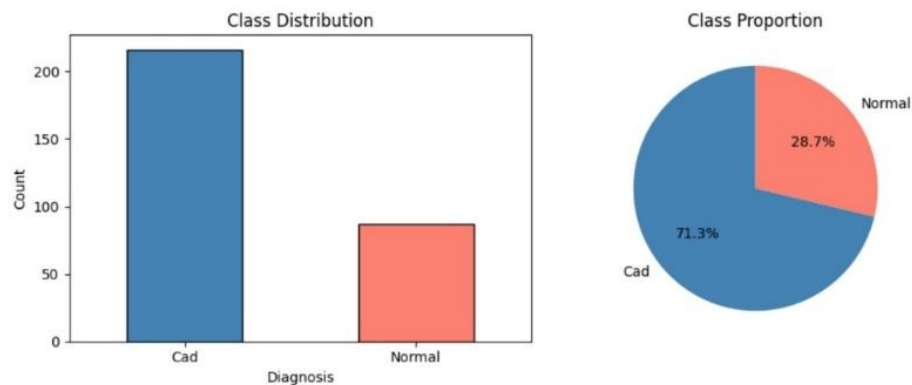
Parameter	Nilai GA	Nilai PSO
Jumlah individu/partikel	100	100
Iterasi/generasi maksimum	50	50
Dimensi solusi	43	43
Fungsi fitness	Rata-rata F1 (3-fold CV)	Rata-rata F1 (3-fold CV)
Early stopping (patience)	5 generasi	5 iterasi
Operator khusus	Crossover blend (0,7), Mutasi Gaussian (0,2)	Bobot inersia 0,9 ke 0,4, $c1=0,5$, $c2=0,3$

Tabel 3 menampilkan konfigurasi lengkap kedua algoritma yang dirancang setara. Fungsi fitness mengukur kualitas setiap solusi dengan melatih keempat model menggunakan kombinasi fitur dan hyperparameter yang dikodekan, kemudian menghitung rata-rata F1-score dari validasi silang 3-lipatan. Mekanisme *early stopping* dengan 5 generasi atau iterasi yang tidak meningkat diterapkan untuk menghentikan pencarian lebih awal apabila nilai fitness terbaik tidak meningkat secara signifikan dengan ambang batas 0,0001, sehingga waktu komputasi lebih efisien tanpa mengorbankan kualitas solusi yang ditemukan [21].

3.3 Evaluasi Model

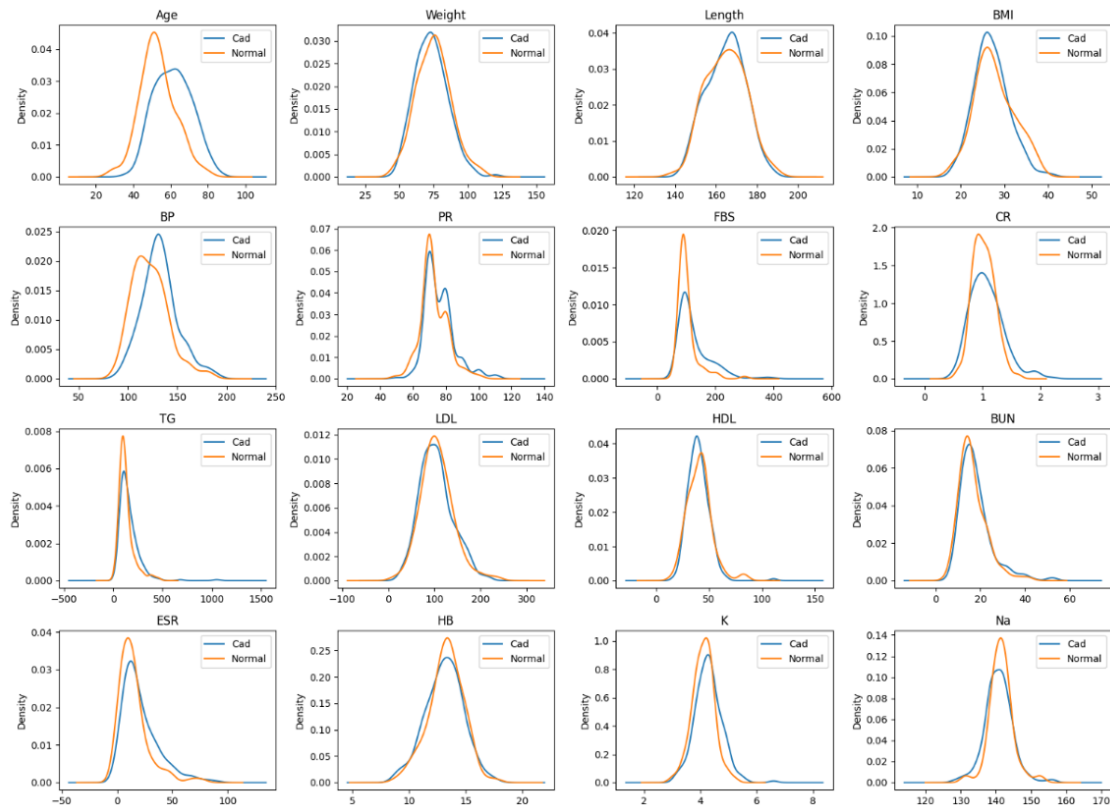
Evaluasi performa dilakukan pada data uji yang sepenuhnya terpisah dari proses pelatihan dan optimasi menggunakan lima metrik komplementer yaitu akurasi, precision, recall, F1-score, dan ROC-AUC. Interpretabilitas model terbaik dianalisis menggunakan SHAP, mencakup visualisasi mean absolute SHAP value dan *beeswarm plot* untuk memahami arah serta distribusi kontribusi setiap fitur terhadap prediksi model.

IV. HASIL DAN PEMBAHASAN



Gambar 4. Distribusi Kelas Target Dataset Z-Alizadeh Sani

Distribusi kelas target yang ditampilkan pada **Gambar 4** menunjukkan ketidakseimbangan yang cukup signifikan dengan rasio hampir 2,5:1 antara pasien CAD dan Normal. Kondisi ini umum dijumpai pada dataset klinis mengingat populasi pasien yang menjalani pemeriksaan kardiovaskular memang cenderung didominasi oleh kasus positif, dan tanpa penanganan yang tepat model akan cenderung memprediksi seluruh sampel sebagai kelas mayoritas sehingga menghasilkan akurasi yang tampak tinggi namun tidak bermakna secara klinis. Penerapan bobot kelas seimbang pada seluruh model menjadi respons yang tepat terhadap kondisi ini karena mendorong model untuk belajar mengenali pola kelas minoritas secara proporsional tanpa mengorbankan integritas data medis.



Gambar 5. Distribusi Fitur Numerik per Kelas (CAD vs Normal)

Gambar 5 menampilkan kurva distribusi kernel untuk 21 fitur numerik kontinu yang memperlihatkan perbedaan distribusi antar kelas secara visual. Usia pasien CAD cenderung terdistribusi ke arah nilai yang lebih tinggi, mengindikasikan bahwa risiko CAD meningkat seiring bertambahnya usia, sementara Fraksi Ejeksi menunjukkan pola terbalik di mana pasien Normal memiliki puncak distribusi lebih tinggi dan lebih sempit sekitar nilai 55 sedangkan distribusi pasien CAD lebih tersebar ke arah nilai yang lebih rendah. Sebaliknya, beberapa fitur seperti Kadar Natrium, Trombosit, dan Hemoglobin menampilkan distribusi yang hampir identik antar kelas, mengindikasikan potensi diskriminatif yang rendah dan menjustifikasi penghapusannya pada tahap preprocessing berdasarkan skor Mutual Information yang mendekati nol.

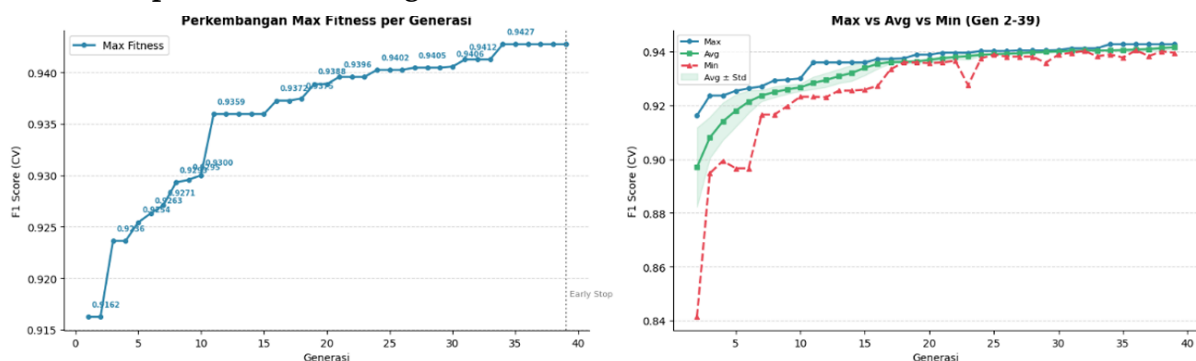
4.2 Hasil Baseline Model

Tabel 4. Hasil Evaluasi Keempat Baseline Model pada Data Uji

Model	Akurasi	Precision	Recall	F1-score	ROC-AUC
Baseline CatBoost	0,8689	0,8723	0,9535	0,9111	0,8734
Baseline RF	0,8525	0,8400	0,9767	0,9032	0,8766
Baseline XGBoost	0,8361	0,8511	0,9302	0,8889	0,8579
Baseline LightGBM	0,8197	0,8478	0,9070	0,8764	0,8579

Tabel 4 menunjukkan bahwa keempat model baseline menghasilkan performa yang cukup kompetitif bahkan tanpa proses optimasi apapun, mencerminkan kualitas fitur yang dihasilkan dari tahap preprocessing yang ketat. CatBoost mencatat F1-score tertinggi sebesar 0,9111 dengan akurasi 0,8689, hasil yang sejalan dengan ekspektasi mengingat dataset Z-Alizadeh Sani memiliki proporsi fitur kategorik yang tinggi dan CatBoost memang dirancang khusus untuk menangani tipe data tersebut. Random Forest menempati posisi kedua dari sisi F1-score namun unggul dalam ROC-AUC (0,8766) dan recall (0,9767), menunjukkan kemampuan deteksi kasus CAD yang lebih sensitif yang sangat relevan dalam konteks klinis.

4.3 Hasil Optimasi Genetic Algorithm



Gambar 6. Analisis Konvergensi Genetic Algorithm selama 39 Generasi

Proses evolusi GA berlangsung selama 39 generasi sebelum early stopping aktif karena nilai fitness terbaik tidak meningkat secara signifikan selama 5 generasi berturut-turut, dengan

total waktu komputasi 7.174 detik. Dinamika konvergensi yang ditampilkan pada **Gambar 6** memperlihatkan peningkatan fitness yang konsisten dari 0,9162 pada generasi pertama menuju 0,9427 pada generasi ke-34, dengan deviasi standar populasi yang turun drastis dari 0,306 pada awal evolusi menuju mendekati nol pada generasi-generasi akhir, mengindikasikan bahwa populasi telah berkonvergensi di sekitar area solusi berkualitas tinggi. Kromosom terbaik hasil evolusi memilih 20 fitur dari 32 fitur yang tersedia, semuanya memiliki relevansi klinis yang dapat dipertanggungjawabkan secara kardiologis, mulai dari faktor risiko klasik seperti Diabetes Mellitus dan Hipertensi hingga temuan ekokardiografi seperti Fraksi Ejeksi dan Kelainan Gerak Dinding Regional.

Tabel 5. Hasil Evaluasi Model GA dan Perbandingan dengan Baseline

Model	Akurasi	Precision	Recall	F1-score	ROC-AUC	Delta F1
GA RF	0,9180	0,9130	0,9767	0,9438	0,9018	+0,0406
GA XGBoost	0,8852	0,8913	0,9535	0,9213	0,8786	+0,0324
GA CatBoost	0,8852	0,8913	0,9535	0,9213	0,8902	+0,0102
GA LightGBM	0,8525	0,8542	0,9535	0,9011	0,8695	+0,0247

Tabel 5 menunjukkan bahwa seluruh model GA tanpa terkecuali mengalami peningkatan F1-score dibandingkan baseline-nya masing-masing, mengkonfirmasi bahwa optimasi GA memberikan nilai tambah yang nyata pada setiap model. GA Random Forest tampil sebagai model terbaik dengan F1-score 0,9438 dan ROC-AUC 0,9018, melampaui seluruh model lainnya di kedua metrik tersebut. Matriks konfusi GA RF memperlihatkan bahwa model berhasil mengklasifikasikan 42 dari 43 pasien CAD dengan benar dan hanya melewatkan satu kasus, sebuah pola kesalahan yang sangat diinginkan karena false negative jauh lebih berbahaya dalam konteks klinis dibandingkan false positive.

4.4 Hasil Optimasi Particle Swarm Optimization

Tabel 6. Hasil Evaluasi Model PSO dan Perbandingan dengan Baseline

Model	Akurasi	Precision	Recall	F1-score	ROC-AUC	Delta F1
PSO RF	0,8852	0,8600	1,0000	0,9247	0,8734	+0,0215
PSO LightGBM	0,8852	0,8600	1,0000	0,9247	0,8566	+0,0483
PSO XGBoost	0,8689	0,8723	0,9535	0,9111	0,8656	+0,0222
PSO CatBoost	0,8689	0,8723	0,9535	0,9111	0,8682	+0,0000

Berdasarkan **Tabel 6** temuan paling menonjol dari hasil PSO yaitu pencapaian recall sempurna 1,0000 secara bersamaan oleh PSO Random Forest dan PSO LightGBM, yang berarti kedua model berhasil mendeteksi seluruh 43 pasien CAD dalam data uji tanpa satu pun yang terlewat. Konsekuensinya terdapat tujuh pasien Normal yang keliru diprediksi sebagai CAD, namun dalam perspektif risiko klinis kesalahan jenis ini jauh lebih dapat diterima dibandingkan melewatkan pasien yang sesungguhnya menderita CAD. Keunggulan PSO

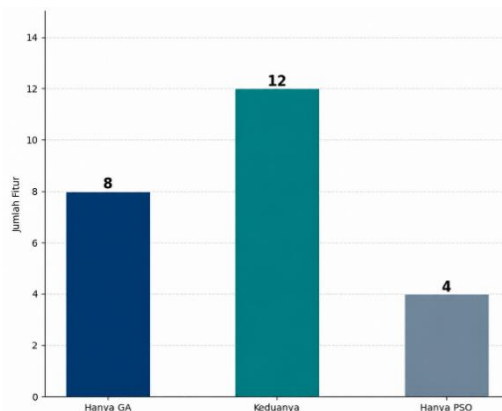
dalam recall ini menjadikannya kandidat yang lebih direkomendasikan untuk aplikasi skrining massal di mana sensitivitas deteksi adalah prioritas absolut.

4.5 Perbandingan Komprehensif dan Analisis Fitur

Tabel 7. Perbandingan Komprehensif Seluruh Model

Pipeline	Model	Akurasi	F1-score	ROC-AUC	Recall
GA	RF	0,9180	0,9438	0,9018	0,9767
PSO	RF	0,8852	0,9247	0,8734	1,0000
PSO	LightGBM	0,8852	0,9247	0,8566	1,0000
GA	XGBoost	0,8852	0,9213	0,8786	0,9535
GA	CatBoost	0,8852	0,9213	0,8902	0,9535
Baseline	CatBoost	0,8689	0,9111	0,8734	0,9535
Baseline	RF	0,8525	0,9032	0,8766	0,9767
Baseline	LightGBM	0,8197	0,8764	0,8579	0,9070

Tabel 7 merangkum perbandingan komprehensif yang menunjukkan beberapa pola penting. Seluruh model yang dioptimasi menghasilkan F1-score lebih tinggi atau setara dibandingkan baseline-nya masing-masing, mengkonfirmasi bahwa proses optimasi memberikan nilai tambah yang konsisten. GA mendominasi posisi teratas berdasarkan F1-score dan ROC-AUC, sementara PSO mendominasi berdasarkan recall. CatBoost merupakan model baseline terkuat namun justru mendapat peningkatan terkecil dari GA, kemungkinan karena nilai bawaannya yang sudah cukup optimal untuk dataset dengan fitur kategorik yang tinggi.

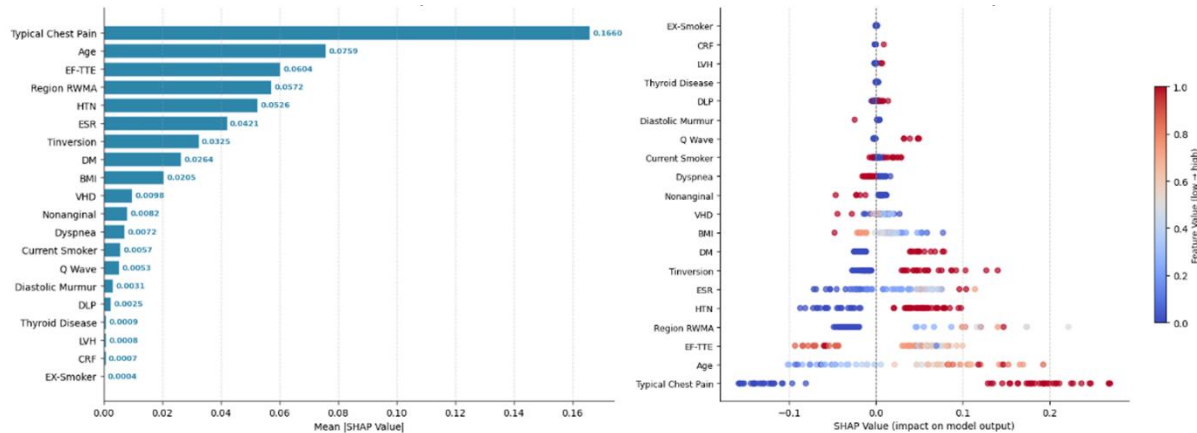


Gambar 9. Perbandingan Fitur Terpilih GA vs PSO

Gambar 9 memperlihatkan bahwa 12 fitur dipilih secara konsisten oleh kedua algoritma, yang dapat dianggap sebagai prediktor paling robust untuk prediksi CAD karena teridentifikasi secara independen oleh dua mekanisme optimasi yang berbeda. Fitur-fitur tersebut mencakup Usia, Indeks Massa Tubuh, Diabetes Mellitus, Hipertensi, Perokok Aktif, Nyeri Dada Tipikal, Inversi Gelombang T, Hipertrofi Ventrikel Kiri, Laju Endap Darah, Fraksi Ejeksi, Kelainan Gerak Dinding Regional, dan Penyakit Katup Jantung. Kedelapan fitur yang hanya dipilih GA mencerminkan pendekatan eksplorasi yang lebih luas melalui mekanisme crossover,

sementara empat fitur unik PSO seperti Kolesterol HDL dan Denyut Nadi menunjukkan bahwa PSO lebih konservatif namun tetap mampu mengidentifikasi prediktor yang memiliki justifikasi klinis valid meski skor Mutual Information-nya tidak tergolong tinggi.

4.6 Interpretabilitas SHAP



Gambar 10. Mean Absolute SHAP Value dan Beeswarm Plot GA Random Forest

Gambar 10 menampilkan hasil analisis SHAP pada model GA Random Forest yang menunjukkan Nyeri Dada Tipikal sebagai fitur paling berpengaruh dengan nilai 0,1660, jauh melampaui fitur lainnya dan menjadikannya faktor tunggal paling menentukan dalam prediksi model ini. Analisis *beeswarm plot* mengungkap pola yang sangat konsisten secara klinis yaitu keberadaan gejala nyeri dada tipikal secara kuat meningkatkan probabilitas prediksi CAD, sementara Fraksi Ejeksi yang rendah juga berkontribusi positif terhadap prediksi CAD sesuai dengan pemahaman bahwa penurunan fungsi pompa jantung merupakan konsekuensi iskemia miokard kronik. Usia yang lebih tua, Hipertensi, dan temuan ekokardiografi abnormal seperti Kelainan Gerak Dinding Regional semuanya berkontribusi positif terhadap prediksi CAD sesuai arah yang diharapkan secara klinis, mengkonfirmasi bahwa model tidak hanya akurat secara statistik tetapi juga telah mempelajari pola yang bermakna secara medis [13].

V. KESIMPULAN DAN SARAN

Penelitian ini berhasil menunjukkan bahwa penerapan Genetic Algorithm dan Particle Swarm Optimization secara simultan untuk seleksi fitur dan optimasi hyperparameter memberikan peningkatan performa yang konsisten pada keempat model *ensemble learning* dibandingkan baseline tanpa optimasi. Model terbaik secara keseluruhan adalah GA Random Forest dengan F1-score 0,9438, akurasi 0,9180, dan ROC-AUC 0,9018, sementara PSO Random Forest dan PSO LightGBM menonjol dengan pencapaian recall sempurna 1,0000 yang menjadikan keduanya tidak melewatkan satu pun pasien CAD dalam data uji. Dua belas fitur yang dipilih secara konsisten oleh kedua algoritma merepresentasikan konsensus prediktor paling robust untuk CAD, dan analisis SHAP mengkonfirmasi bahwa model mempelajari pola yang

bermakna secara medis dengan Nyeri Dada Tipikal, Usia, dan Fraksi Ejeksi sebagai faktor klinis paling berpengaruh sesuai konsensus kardiologi internasional. Penelitian selanjutnya disarankan untuk menggunakan dataset yang lebih besar agar signifikansi statistik dapat diuji dengan lebih memadai.

VI. DAFTAR PUSTAKA

- [1] A. A. Huang dan S. Y. Huang, "Use of machine learning to identify risk factors for coronary artery disease," *PLoS One*, vol. 18, no. 4, hlm. e0284103, Apr 2023, doi: 10.1371/journal.pone.0284103.
- [2] M. Zhang, H. Wang, dan J. Zhao, "Use machine learning models to identify and assess risk factors for coronary artery disease," *PLoS One*, vol. 19, no. 9, hlm. e0307952, Sep 2024, doi: 10.1371/journal.pone.0307952.
- [3] M. Trigka dan E. Dritsas, "Long-Term Coronary Artery Disease Risk Prediction with Machine Learning Models," *Sensors*, vol. 23, no. 3, hlm. 1193, Jan 2023, doi: 10.3390/s23031193.
- [4] B. P. Kaur, H. Singh, R. Hans, S. K. Sharma, C. Sharma, dan Md. M. Hassan, "A Genetic algorithm aided hyper parameter optimization based ensemble model for respiratory disease prediction with Explainable AI," *PLoS One*, vol. 19, no. 12, hlm. e0308015, Des 2024, doi: 10.1371/journal.pone.0308015.
- [5] C. Yu dan H. Pei, "Dynamic Weighting Translation Transfer Learning for Imbalanced Medical Image Classification," *Entropy*, vol. 26, no. 5, hlm. 400, Mei 2024, doi: 10.3390/e26050400.
- [6] T. Chen dan C. Guestrin, "XGBoost: A Scalable Tree Boosting System," dalam *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, New York, NY, USA: ACM, Agu 2016, hlm. 785–794. doi: 10.1145/2939672.2939785.
- [7] I. D. Raji, H. Bello-Salau, I. J. Umoh, A. J. Onumanyi, M. A. Adegboye, dan A. T. Salawudeen, "Simple Deterministic Selection-Based Genetic Algorithm for Hyperparameter Tuning of Machine Learning Models," *Applied Sciences*, vol. 12, no. 3, hlm. 1186, Jan 2022, doi: 10.3390/app12031186.
- [8] H. L. H. S. Warnars, A. S. Sunge, Suzanna, B. Bevyadi, dan M. K. Mueyba, "Prediction of Violence Against Women Using Ensemble Learning Models: A Comparative Study of LightGBM, XGBoost, and Others," *International Journal of Safety and Security Engineering*, vol. 15, no. 4, Apr 2025, doi: 10.18280/ijss.150409.
- [9] A. Mehdiyari, A. Chehri, A. Jakimi, dan R. Saadane, "Hyperparameter Optimization with Genetic Algorithms and XGBoost: A Step Forward in Smart Grid Fraud Detection," *Sensors*, vol. 24, no. 4, Feb 2024, doi: 10.3390/s24041230.
- [10] J. Kennedy dan R. Eberhart, "Particle swarm optimization," dalam *Proceedings of ICNN'95 - International Conference on Neural Networks*, IEEE, hlm. 1942–1948. doi: 10.1109/ICNN.1995.488968.
- [11] F. Yang, Z. Xu, H. Wang, L. Sun, M. Zhai, dan J. Zhang, "A hybrid feature selection algorithm combining information gain and grouping particle swarm optimization for cancer diagnosis," *PLoS One*, vol. 19, no. 3, hlm. e0290332, Mar 2024, doi: 10.1371/journal.pone.0290332.
- [12] N. Tasmurzaev dkk., "Interpretable Machine Learning for Coronary Artery Disease Risk Stratification: A SHAP-Based Analysis," *Algorithms*, vol. 18, no. 11, hlm. 697, Nov 2025, doi: 10.3390/a18110697.
- [13] S. Bin Akter dkk., "Optimizing stability of heart disease prediction across imbalanced learning with interpretable Grow Network," *Comput. Methods Programs Biomed.*, vol. 265, hlm. 108702, Jun 2025, doi: 10.1016/j.cmpb.2025.108702.
- [14] S. Q. Sultan, N. Javaid, N. Alrajeh, dan M. Aslam, "Machine Learning-Based Stacking Ensemble Model for Prediction of Heart Disease with Explainable AI and K-Fold Cross-Validation: A

- Symmetric Approach,” *Symmetry (Basel)*., vol. 17, no. 2, hlm. 185, Jan 2025, doi: 10.3390/sym17020185.
- [15] H. Xie, L. Zhang, C. P. Lim, Y. Yu, dan H. Liu, “Feature Selection Using Enhanced Particle Swarm Optimisation for Classification Models,” *Sensors*, vol. 21, no. 5, hlm. 1816, Mar 2021, doi: 10.3390/s21051816.
- [16] V. I. Kigka dkk., “Machine Learning Coronary Artery Disease Prediction Based on Imaging and Non-Imaging Data,” *Diagnostics*, vol. 12, no. 6, hlm. 1466, Jun 2022, doi: 10.3390/diagnostics12061466.
- [17] A. Abdo, R. Mostafa, dan L. Abdel-Hamid, “An Optimized Hybrid Approach for Feature Selection Based on Chi-Square and Particle Swarm Optimization Algorithms,” *Data (Basel)*., vol. 9, no. 2, hlm. 20, Jan 2024, doi: 10.3390/data9020020.
- [18] A. Çakmak, G. Akyilmaz, A. G. Köse, G. Keskin, dan L. Uğur, “Machine Learning-Based Prediction of Coronary Artery Disease Using Clinical and Behavioral Data: A Comparative Study,” *Diagnostics*, vol. 16, no. 2, hlm. 318, Jan 2026, doi: 10.3390/diagnostics16020318.
- [19] A. S. Alfath, A. K. Wardhana, dan R. Rumini, “Hypertension Risk Prediction Using Stacking Ensemble of CatBoost, XGBoost, and LightGBM: A Machine Learning Approach,” *Journal of Applied Informatics and Computing*, vol. 9, no. 6, hlm. 3146–3156, Des 2025, doi: 10.30871/jaic.v9i6.10370.
- [20] I. Araf, A. Idri, dan I. Chairi, “Cost-sensitive learning for imbalanced medical data: a review,” *Artif. Intell. Rev.*, vol. 57, no. 4, hlm. 80, Mar 2024, doi: 10.1007/s10462-023-10652-8.
- [21] M. Imani, A. Beikmohammadi, dan H. R. Arabnia, “Comprehensive Analysis of Random Forest and XGBoost Performance with SMOTE, ADASYN, and GNUS Under Varying Imbalance Levels,” *Technologies (Basel)*., vol. 13, no. 3, hlm. 88, Feb 2025, doi: 10.3390/technologies13030088.