

Optimasi Seleksi Fitur Adaptive Particle Swarm Optimization Untuk Klasifikasi Penyakit Jantung Dengan Ensemble Learning

Bagas Adi Nata^{1*}, Solikhun²

^{1*}Program Studi Sistem Informasi, STIKOM Tunas Bangsa, Pematangsiantar, Indonesia

²Program Studi Teknik Informatika, STIKOM Tunas Bangsa, Pematangsiantar, Indonesia

Email: ^{1*}adinatabagas364@gmail.com, ²solikhun@amiktunasbangsa.ac.id

Email Penulis Korespondensi: adinatabagas364@gmail.com

Abstrak—Klasifikasi penyakit jantung berbasis machine learning memerlukan fitur yang relevan dan model yang mampu menggeneralisasi pola klinis secara konsisten. Penelitian terdahulu pada dataset Heart Failure Prediction menunjukkan bahwa K-Nearest Neighbor (KNN) yang dioptimasi Particle Swarm Optimization (PSO) mencapai Akurasi 89,09% dan area under the curve (AUC) 0,935. Namun, penggunaan bobot inerti tetap dan ketergantungan pada single learner masih berpotensi membatasi keseimbangan eksplorasi-eksploitasi serta ketahanan model. Penelitian ini mengusulkan seleksi fitur berbasis Adaptive Particle Swarm Optimization (APSO) yang menurunkan bobot inerti secara bertahap dari 0,90 menjadi sekitar 0,42 selama 30 iterasi, kemudian memanfaatkan subset fitur terbaik pada ensemble learning berbasis soft voting. Dataset terdiri atas 918 rekaman, 11 fitur prediktor, dan satu kelas target HeartDisease. Berdasarkan keluaran eksperimen, model APSO dan ensemble menghasilkan Akurasi 89,71%, F1-score 0,8986, dan AUC 0,9466. Confusion matrix menunjukkan 90 true negative, 12 false positive, 9 false negative, dan 93 true positive pada 204 data uji. Dibandingkan baseline KNN dan PSO, metode yang diusulkan meningkatkan Akurasi sebesar 0,62 poin persentase dan AUC sebesar 0,0116, sekaligus mempertahankan recall kelas penyakit sebesar 91,18%. Hasil ini menunjukkan bahwa adaptasi dinamika pencarian dan penggabungan model heterogen dapat meningkatkan daya diskriminasi klasifikasi penyakit jantung, meskipun validasi dengan pembagian data identik dan pengujian eksternal masih diperlukan.

Kata Kunci: Adaptive Particle Swarm Optimization; Ensemble Learning; Klasifikasi; Penyakit Jantung; Seleksi Fitur

Abstract—Heart disease classification using machine learning requires relevant features and predictive models capable of consistently generalizing clinical patterns. Previous studies on the Heart Failure Prediction dataset demonstrated that K-Nearest Neighbor (KNN) optimized with Particle Swarm Optimization (PSO) achieved an accuracy of 89.09% and an Area Under the Curve (AUC) of 0.935. However, the use of a fixed inertia weight and reliance on a single learner may limit the balance between exploration and exploitation, thereby reducing model robustness. This study proposes a feature selection approach based on Adaptive Particle Swarm Optimization (APSO), in which the inertia weight is gradually decreased from 0.90 to approximately 0.42 over 30 iterations. The optimal feature subset is subsequently utilized in a soft voting ensemble learning model. The dataset consists of 918 records, 11 predictive features, and one target class (HeartDisease). Experimental results indicate that the proposed APSO-based ensemble model achieved an accuracy of 89.71%, an F1-score of 0.8986, and an AUC of 0.9466. The confusion matrix yielded 90 true negatives, 12 false positives, 9 false negatives, and 93 true positives on 204 testing instances. Compared with the baseline KNN-PSO model, the proposed method improved classification accuracy by 0.62 percentage points and increased the AUC by 0.0116, while maintaining a disease-class recall of 91.18%. These findings demonstrate that combining adaptive search dynamics with heterogeneous ensemble learning enhances the discriminative capability of heart disease classification, although further validation using identical data partitioning strategies and external datasets is still required.

Keywords: Adaptive Particle Swarm Optimization; Ensemble Learning; Classification; Heart Disease; Feature Selection

1. PENDAHULUAN

Penyakit jantung merupakan salah satu masalah kesehatan yang memerlukan deteksi dan pengambilan keputusan secara cepat [1], [2], [3]. Pertumbuhan data klinis memungkinkan penerapan machine learning untuk membantu proses klasifikasi pasien berdasarkan kombinasi karakteristik demografis, hasil pemeriksaan, dan respons terhadap aktivitas fisik. Meskipun demikian, performa model tidak hanya dipengaruhi oleh pemilihan algoritma klasifikasi, tetapi juga oleh kualitas representasi fitur yang digunakan.

Dalam beberapa tahun terakhir, perkembangan machine learning telah memberikan kontribusi yang signifikan terhadap bidang kesehatan, khususnya dalam membantu proses diagnosis dan prediksi berbagai penyakit. Algoritma machine learning mampu mempelajari pola yang tersembunyi dari data klinis sehingga dapat menghasilkan prediksi yang lebih cepat dan konsisten dibandingkan analisis manual. Pada kasus penyakit jantung, kemampuan model dalam mengidentifikasi pasien yang berisiko sejak dini menjadi sangat penting karena dapat mendukung tenaga medis dalam menentukan langkah penanganan yang lebih tepat. Oleh karena itu, peningkatan performa model klasifikasi menjadi salah satu fokus utama dalam penelitian berbasis kecerdasan buatan di bidang kesehatan.

Dataset penyakit jantung umumnya memuat fitur numerik dan kategorikal dengan skala serta kontribusi yang berbeda. Fitur yang tidak relevan, redundan, atau mengandung nilai yang tidak konsisten dapat meningkatkan kompleksitas model dan menurunkan kemampuan generalisasi [4]. Selain meningkatkan waktu komputasi, keberadaan fitur yang kurang informatif juga berpotensi menyebabkan model mengalami overfitting sehingga performanya menurun ketika dihadapkan pada data baru. Oleh sebab itu, proses seleksi fitur menjadi tahap yang penting karena mampu menyederhanakan ruang pencarian, mengurangi noise, meningkatkan efisiensi komputasi, serta membantu model memusatkan pembelajaran pada atribut yang benar-benar berkontribusi terhadap proses klasifikasi [5].

Berbagai metode optimasi telah digunakan untuk melakukan seleksi fitur, salah satunya adalah Particle Swarm Optimization (PSO). Algoritma ini dikenal memiliki kemampuan pencarian global yang baik dan relatif sederhana untuk

diimplementasikan pada berbagai permasalahan optimasi. Dalam bidang klasifikasi penyakit, PSO sering dimanfaatkan untuk memperoleh kombinasi fitur yang mampu meningkatkan performa model sekaligus mengurangi jumlah atribut yang digunakan. Meskipun demikian, efektivitas PSO masih dipengaruhi oleh pengaturan parameter yang digunakan selama proses optimasi.

Penelitian sebelumnya oleh Jusia dkk. (2024) menguji K-Nearest Neighbor (KNN) dan C4.5 dengan pendekatan optimasi PSO pada Heart Failure Prediction Dataset dari Kaggle. Penelitian tersebut menggunakan 918 rekaman, 11 atribut prediktor, pembagian data 70:30, dan RapidMiner sebagai alat eksperimen. Hasil terbaik diperoleh oleh KNN berbasis PSO dengan akurasi 89,09%, precision 89,61%, recall 90,79%, dan AUC 0,935. Pada konfigurasi tersebut, PSO menggunakan population size 5, maksimum 30 generasi, inertia weight 1,0, local best weight 1,0, dan global best weight 1,0.

Temuan tersebut membuktikan bahwa PSO mampu meningkatkan performa klasifikasi [6], tetapi masih terdapat beberapa keterbatasan. Penggunaan inertia weight yang bernilai tetap belum mampu menyesuaikan kebutuhan eksplorasi pada tahap awal dan eksploitasi pada tahap akhir proses pencarian. Kondisi tersebut berpotensi menyebabkan partikel terjebak pada solusi lokal sehingga proses pencarian subset fitur belum optimal. Selain itu, penggunaan single learner membuat performa akhir model sangat bergantung pada karakteristik satu algoritma klasifikasi. Sebagai contoh, KNN sangat dipengaruhi oleh skala data dan pemilihan tetangga terdekat, sedangkan algoritma berbasis pohon memiliki karakteristik kesalahan yang berbeda.

Salah satu pendekatan yang dapat digunakan untuk mengatasi keterbatasan tersebut adalah Adaptive Particle Swarm Optimization (APSO). Berbeda dengan PSO konvensional, APSO menerapkan mekanisme penyesuaian inertia weight secara bertahap selama proses iterasi sehingga keseimbangan antara eksplorasi dan eksploitasi dapat dipertahankan dengan lebih baik. Pendekatan ini diharapkan mampu menghasilkan subset fitur yang lebih representatif sehingga kualitas model klasifikasi dapat meningkat.

Selain optimasi seleksi fitur, penggunaan ensemble learning juga menjadi strategi yang banyak diterapkan untuk meningkatkan kemampuan generalisasi model. Ensemble learning mengombinasikan beberapa algoritma klasifikasi sehingga keputusan akhir tidak hanya bergantung pada satu model. Melalui mekanisme soft voting, probabilitas prediksi dari beberapa classifier digabungkan untuk menghasilkan keputusan yang lebih stabil. Pendekatan ini diharapkan mampu mengurangi kelemahan masing-masing algoritma sekaligus meningkatkan ketahanan model terhadap variasi data.

Berdasarkan celah penelitian tersebut, penelitian ini mengusulkan seleksi fitur menggunakan Adaptive Particle Swarm Optimization (APSO) dan memanfaatkan subset fitur terbaik sebagai masukan bagi model ensemble learning. Kontribusi utama penelitian meliputi penggunaan inertia weight adaptif untuk menyeimbangkan eksplorasi dan eksploitasi, integrasi seleksi fitur biner dengan soft voting ensemble yang menggabungkan beberapa base learner heterogen, serta evaluasi performa menggunakan accuracy, precision, recall, F1-score, confusion matrix, dan Area Under the Curve (AUC). Pendekatan tersebut diharapkan tidak hanya meningkatkan akurasi klasifikasi, tetapi juga menghasilkan model yang memiliki kemampuan diskriminasi dan generalisasi yang lebih baik dibandingkan metode sebelumnya.

Berdasarkan uraian tersebut, penelitian ini mengusulkan penerapan Adaptive Particle Swarm Optimization (APSO) sebagai metode seleksi fitur yang diintegrasikan dengan ensemble learning berbasis soft voting untuk meningkatkan performa klasifikasi penyakit jantung. Kontribusi utama penelitian ini meliputi: (1) penerapan mekanisme adaptive inertia weight pada APSO untuk memperoleh subset fitur yang lebih representatif dibandingkan PSO konvensional; (2) integrasi hasil seleksi fitur dengan model ensemble learning yang mengombinasikan beberapa base learner melalui mekanisme soft voting guna meningkatkan kemampuan generalisasi model; dan (3) evaluasi kinerja menggunakan accuracy, precision, recall, F1-score, AUC, serta confusion matrix untuk menunjukkan efektivitas pendekatan yang diusulkan. Melalui kontribusi tersebut, penelitian ini diharapkan mampu menghasilkan model klasifikasi penyakit jantung yang lebih akurat, stabil, dan memiliki kemampuan diskriminasi yang lebih baik dibandingkan metode sebelumnya.

2. METODOLOGI PENELITIAN

2.1 Desain Penelitian

Penelitian ini menggunakan pendekatan eksperimental[7]. Alur dimulai dari pengambilan dataset, pra-pemrosesan, seleksi fitur menggunakan APSO, pelatihan ensemble pada subset fitur terbaik, dan evaluasi pada data uji yang terpisah. Kerangka penelitian ditunjukkan pada Gambar 1.



Gambar 1. Kerangka Penelitian APSO dan Ensemble Learning

2.2 Dataset dan Pra-pemrosesan

Dataset yang digunakan adalah Heart Failure Prediction Dataset yang juga dipakai pada penelitian sebelumnya. Dataset terdiri atas 918 rekaman dengan 11 fitur prediktor dan satu kelas target HeartDisease. Label 0 menunjukkan kondisi normal, sedangkan label 1 menunjukkan penyakit jantung. Ringkasan dataset disajikan pada Tabel 1.

Tabel 1. Ringkasan karakteristik dataset

Komponen	Keterangan
Jumlah Rekaman	981 data pasien
Fitur Prediktor	11 atribut
Kelas target	HeartDisease (0 dan 1)
Distribusi Kelas	410 Normal dan 580 Penyakit jantung
Fitur Numerik	Age, RestingBp, Cholestrol, MaxHr, Old peak
Fitur	Sex, ChestPainType, FastingBps

Pra-pemrosesan dilakukan dengan memeriksa duplikasi, nilai kosong, dan nilai nol yang tidak fisiologis pada atribut seperti RestingBP dan Cholesterol[8], [9], [10]. Imputasi, encoding fitur kategorikal, serta standardisasi fitur numerik harus dipelajari hanya dari data latih untuk mencegah kebocoran informasi. Pada keluaran eksperimen yang dianalisis, set uji berisi 204 data dengan distribusi seimbang, yaitu 102 data untuk masing-masing kelas.

2.3 Adaptive Particle Swarm Optimization

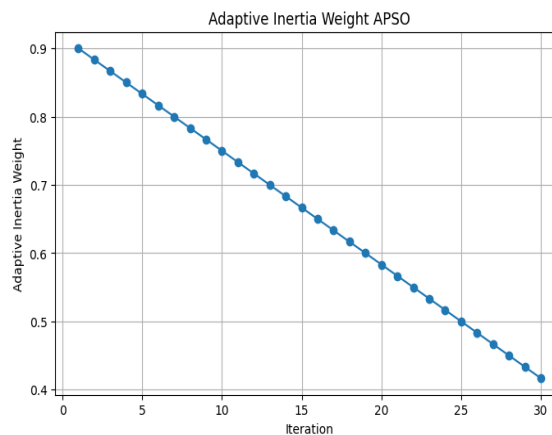
Adaptive Particle Swarm Optimization (APSO) merupakan pengembangan dari metode Particle Swarm Optimization (PSO) yang dirancang untuk meningkatkan kemampuan pencarian solusi optimal secara lebih adaptif[11], [12], [13]. Metode ini terinspirasi dari perilaku kawanan burung atau ikan dalam menemukan posisi terbaik melalui proses interaksi antar partikel di dalam ruang pencarian[14], [15]. APSO merepresentasikan setiap partikel sebagai vektor biner sepanjang jumlah fitur. Nilai 1 menandakan fitur dipilih, sedangkan nilai 0 menandakan fitur tidak digunakan. Posisi partikel merupakan kandidat subset fitur, dan kualitas setiap kandidat dinilai berdasarkan performa model pada data latih/validasi serta jumlah fitur yang digunakan.

$$v_i(t+1) = w(t)v_i(t) + c_1r_1(pbest_i - x_i(t)) + c_2r_2(gbest - x_i(t)) \quad (1)$$

$$x_i(t+1) = \begin{cases} 1, & \text{if } \text{sigmoid}(v_i(t+1)) > rand \\ 0, & \text{otherwise} \end{cases} \quad (2)$$

Berbeda dari PSO konvensional pada penelitian sebelumnya yang menggunakan inertia weight tetap 1,0, penelitian ini memakai inertia weight adaptif. Bobot yang lebih besar pada fase awal mendorong eksplorasi subset fitur, sedangkan bobot yang lebih kecil pada fase akhir memperkuat eksploitasi di sekitar solusi terbaik.

Dengan $w_{\max} = 0,90$, $w_{\min} = 0,40$, dan $T = 30$ iterasi, nilai inertia menurun hingga sekitar 0,42 pada iterasi terakhir. Perubahan bobot tersebut ditunjukkan pada Gambar 2.



Gambar 2. Penerapan adaptive inertia weight selama 30 menit

Fungsi fitness dirancang untuk meminimalkan kesalahan klasifikasi dan proporsi fitur yang dipilih. Bentuk umum fungsi fitness dituliskan sebagai berikut.

$$Fitness(S) = \alpha(1 - Acc_{cv}(S)) + (1 - \alpha) \left(\frac{|S|}{d} \right) \quad (3)$$

Pada persamaan tersebut, S adalah subset fitur, Acc_cv adalah Akurasi validasi, d adalah jumlah seluruh fitur, dan alpha merupakan bobot yang memprioritaskan performa prediksi. Dengan formulasi ini, APSO diarahkan untuk menemukan subset yang ringkas tanpa mengorbankan kemampuan klasifikasi.

2.4 Ensemble Learning

Ensemble learning merupakan pendekatan machine learning yang mengombinasikan beberapa model klasifikasi untuk menghasilkan keputusan prediksi yang lebih stabil dan akurat dibandingkan penggunaan single classifier [16]. Pendekatan ini bekerja dengan memanfaatkan keberagaman karakteristik pembelajaran dari setiap model sehingga kelemahan pada satu algoritma dapat dikurangi oleh algoritma lainnya. Pada kasus klasifikasi medis, ensemble learning banyak digunakan karena mampu meningkatkan kemampuan generalisasi model dan mengurangi risiko overfitting terhadap data latih.

Subset fitur terbaik dari APSO digunakan sebagai masukan bagi ensemble learning berbasis soft voting. Ensemble menggabungkan base learner heterogen, seperti KNN, Random Forest, dan Support Vector Machine, sehingga kesalahan dari satu model dapat dikompensasi oleh model lain [17]. Keputusan akhir ditentukan dari rata-rata probabilitas prediksi seluruh anggota ensemble.

$$P(y = 1 | x) = \frac{1}{M} \sum_{m=1}^M P_m(y = 1 | x) \quad (4)$$

Sampel diklasifikasikan sebagai penyakit jantung apabila probabilitas gabungan lebih besar atau sama dengan ambang 0,5. Penggunaan soft voting dipilih karena memanfaatkan tingkat keyakinan model, bukan hanya label kelas akhir.

2.4 Ensemble Learning

Kinerja model dievaluasi menggunakan Akurasi, precision, recall, F1-score, specificity, confusion matrix, dan AUC. Akurasi mengukur proporsi prediksi benar, precision mengukur ketepatan prediksi kelas penyakit [18], [19], [20], recall mengukur kemampuan mendeteksi pasien penyakit jantung [21], [22], sedangkan F1-score menyeimbangkan precision dan recall. AUC digunakan untuk menilai kemampuan diskriminasi model pada berbagai ambang keputusan [23], [24], [25].

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (5)$$

$$Recall = \frac{TP}{TP+FN} \quad (6)$$

$$F1-Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (7)$$

3. HASIL DAN PEMBAHASAN

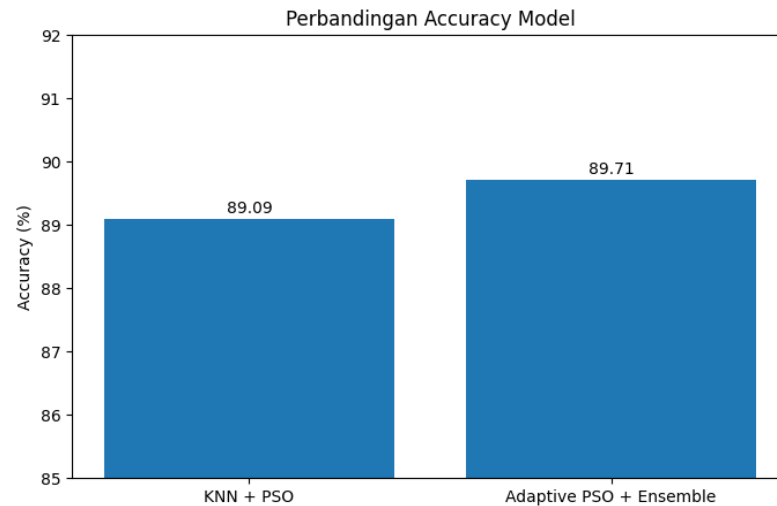
3.1 Hasil Kinerja Model

Model APSO dan ensemble menghasilkan Akurasi 0,8971 atau 89,71%. Nilai F1-score kelas penyakit sebesar 0,8986 dan AUC sebesar 0,9466 menunjukkan bahwa model tidak hanya menghasilkan prediksi benar pada sebagian besar data uji, tetapi juga memiliki kemampuan diskriminasi yang sangat baik antara kelas normal dan penyakit jantung.

Tabel 2. Ringkasan karakteristik dataset

Metrik	Nilai
Akurasi	89,71%
Precision kelas 1	88,57%
Recall Kelas 1	91,18%
F1-score kelas 1	0,8986
Specificity	88,42%
AUC	0,9466

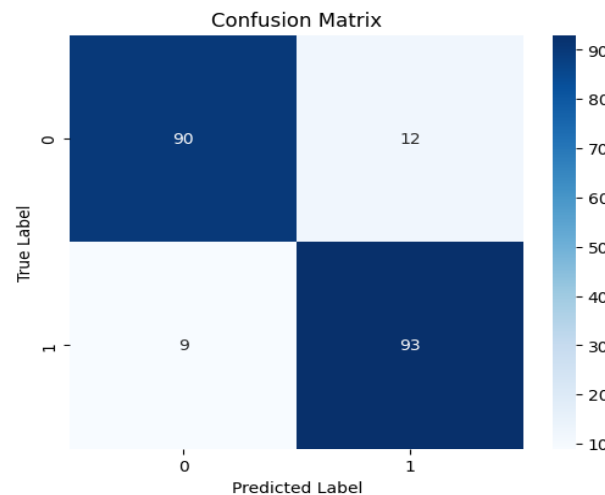
Gambar 3 memperlihatkan bahwa Akurasi APSO dan ensemble sebesar 89,71% berada di atas hasil KNN dan PSO pada penelitian sebelumnya sebesar 89,09%. Selisih 0,62 poin persentase memang tidak besar, tetapi menunjukkan bahwa adaptasi dinamika pencarian dan agregasi model dapat memberi peningkatan tambahan pada dataset yang sama.



Gambar 3. Perbandingan akurasi dengan hasil terbaik penelitian sebelumnya

3.2 Analisis Confusion Matrix

Confusion matrix pada Gambar 4 menunjukkan 90 data normal diprediksi dengan benar dan 93 data penyakit jantung diprediksi dengan benar. Kesalahan terdiri atas 12 false positive dan 9 false negative. Jumlah false negative yang lebih rendah penting pada konteks skrining karena kesalahan tersebut merepresentasikan pasien penyakit jantung yang diprediksi sebagai normal.



Gambar 4. Confusion matrix model APSO dan ensemble

Selanjutnya pada Tabel 3 yaitu Ringkasan karakteristik dataset.

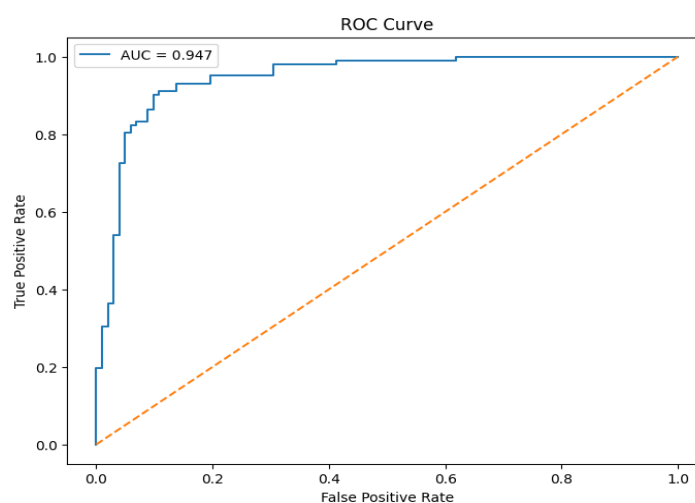
Tabel 3. Ringkasan karakteristik dataset

Kelas	Precision	Recall	F1
0	0,9091	0,8824	0,8955
1	0,8857	0,9118	0,8986
Macro avg	0,8974	0,8971	0,8970
Weighted avg	0,8974	0,8971	0,8970
Kelas	Precision	Recall	F1

Kinerja kedua kelas relatif seimbang. Precision kelas normal lebih tinggi daripada kelas penyakit, sedangkan recall kelas penyakit lebih tinggi daripada kelas normal. Pola ini menunjukkan bahwa model cenderung menjaga sensitivitas deteksi penyakit, dengan konsekuensi sejumlah data normal diprediksi sebagai penyakit.

3.3 Analisis Kurva ROC

Kurva ROC pada Gambar 5 berada jauh di atas garis diagonal dengan AUC 0,9466. Nilai tersebut menunjukkan probabilitas yang tinggi bahwa model memberikan skor lebih besar kepada pasien penyakit jantung dibandingkan pasien normal. AUC juga penting karena tidak bergantung pada satu ambang klasifikasi tertentu.



Gambar 5. Kurva ROC mode APSO dan ensemble learning dalam AUC 0,9466

3.4 Perbandingan Dengan Penelitian Sebelumnya

Perbandingan terhadap penelitian sebelumnya disajikan pada Tabel 4. Penelitian sebelumnya menggunakan PSO konvensional dengan inertia weight tetap dan menilai KNN serta C4.5 sebagai single learner. Penelitian ini mengganti mekanisme pencarian menjadi adaptif dan menggabungkan beberapa model melalui ensemble learning.

Tabel 4. Ringkasan karakteristik dataset

Aspek	Jusia dkk. (2024)	Penelitian ini
Optimasi	PSO konvensional	APSO
Inertia weight	Tetap 1,0	0,90 ke sekitar 0,42
Klasifikator	KNN/C4.5 tunggal	Soft voting ensemble
Accuracy	89,09%	89,71%
Precision	89,61%	88,57%
Recall	90,79%	91,18%
AUC	0,935	0,9466

Dibandingkan hasil terbaik penelitian sebelumnya, Akurasi meningkat sebesar 0,62 poin persentase dan AUC meningkat sebesar 0,0116. Recall kelas penyakit juga meningkat sekitar 0,39 poin persentase, sedangkan precision menurun sekitar 1,04 poin persentase. Dengan demikian, manfaat utama metode yang diusulkan terletak pada peningkatan daya diskriminasi dan sensitivitas deteksi penyakit, bukan pada peningkatan seluruh metrik secara seragam.

Novelty penelitian ini bukan sekadar mengganti algoritma klasifikasi, tetapi mengubah dua komponen yang saling terkait. APSO mengatur intensitas pencarian subset fitur secara dinamis, sementara ensemble mengurangi ketergantungan pada satu bias model. Kombinasi tersebut relevan untuk data klinis yang mengandung fitur dengan tipe dan pola hubungan yang beragam.

Walaupun demikian, selisih hasil harus ditafsirkan secara hati-hati. Perbedaan ukuran set uji, proses pra-pemrosesan, alat implementasi, dan seed acak dapat memengaruhi nilai akhir. Oleh karena itu, perbandingan pada

penelitian ini menunjukkan potensi peningkatan, tetapi belum dapat dianggap sebagai bukti statistik bahwa APSO dan ensemble selalu lebih unggul pada seluruh skenario.

3.5 Perbandingan Dengan Penelitian Sebelumnya

Penelitian ini memiliki beberapa keterbatasan. Pertama, evaluasi masih menggunakan satu dataset publik sehingga generalisasi ke populasi klinis lain belum dapat dipastikan. Kedua, nilai nol pada beberapa atribut memerlukan penanganan yang konsisten karena dapat merepresentasikan data yang hilang. Ketiga, perbandingan dengan penelitian sebelumnya belum dilakukan pada pembagian data dan perangkat eksperimen yang identik. Keempat, ensemble meningkatkan kompleksitas komputasi dan mengurangi kemudahan interpretasi dibandingkan single learner.

Untuk memastikan replikasi, implementasi final perlu menyimpan seed acak, subset fitur gbest, anggota dan bobot ensemble, parameter APSO, serta waktu komputasi. Pengujian menggunakan repeated stratified cross-validation, nested validation, dan dataset eksternal juga diperlukan untuk mengukur stabilitas hasil.

3.6 Pembahasan

Hasil penelitian menunjukkan bahwa kombinasi Adaptive Particle Swarm Optimization (APSO) dan ensemble learning mampu meningkatkan performa klasifikasi penyakit jantung dibandingkan pendekatan PSO konvensional pada penelitian rujukan. Model yang diusulkan menghasilkan accuracy sebesar 89,71% dengan nilai AUC mencapai 0,9466. Hasil tersebut menunjukkan bahwa metode yang diterapkan tidak hanya mampu meningkatkan ketepatan klasifikasi, tetapi juga meningkatkan kemampuan model dalam membedakan kelas normal dan penyakit jantung secara lebih baik.

Peningkatan performa dipengaruhi oleh penggunaan adaptive inertia weight pada proses optimasi fitur. Pada penelitian ini, inertia weight diturunkan secara bertahap dari 0,90 hingga sekitar 0,42 selama 30 iterasi. Nilai inertia weight yang tinggi pada awal iterasi membantu partikel melakukan eksplorasi ruang solusi secara lebih luas sehingga peluang menemukan kombinasi fitur terbaik menjadi lebih besar. Selanjutnya, penurunan inertia weight pada iterasi akhir membantu proses eksploitasi agar pencarian lebih fokus pada solusi terbaik yang telah ditemukan sebelumnya. Pendekatan tersebut membuat proses seleksi fitur menjadi lebih adaptif dibandingkan PSO konvensional yang menggunakan inertia weight tetap.

Grafik adaptive inertia weight menunjukkan pola penurunan linear yang stabil selama proses optimasi. Kondisi tersebut mengindikasikan bahwa APSO mampu menjaga keseimbangan antara eksplorasi dan eksploitasi secara bertahap. Pada tahap awal, partikel memiliki pergerakan yang lebih luas untuk menghindari konvergensi prematur, sedangkan pada tahap akhir pergerakan partikel menjadi lebih stabil sehingga proses pencarian lebih terarah menuju subset fitur optimal. Mekanisme ini menjadi salah satu faktor yang mendukung peningkatan performa klasifikasi dibandingkan pendekatan PSO standar pada penelitian sebelumnya.

Selain proses optimasi fitur, penggunaan ensemble learning berbasis soft voting juga memberikan kontribusi terhadap peningkatan performa model. Ensemble learning menggabungkan beberapa base learner heterogen sehingga keputusan klasifikasi tidak bergantung pada satu model saja. Pada penelitian ini, kombinasi model digunakan untuk memanfaatkan karakteristik pembelajaran yang berbeda dari setiap algoritma klasifikasi. KNN memiliki kemampuan mengenali pola berdasarkan kedekatan antar-data, Random Forest mampu menangani hubungan nonlinier dan mengurangi overfitting, sedangkan Support Vector Machine efektif dalam membentuk batas pemisah antar-kelas. Penggabungan probabilitas prediksi dari beberapa model membantu meningkatkan stabilitas keputusan klasifikasi.

Hasil confusion matrix menunjukkan bahwa model berhasil mengklasifikasikan 90 data normal dan 93 data penyakit jantung secara benar. Kesalahan klasifikasi terdiri atas 12 false positive dan 9 false negative. Jumlah false negative yang lebih rendah menjadi aspek penting dalam sistem deteksi penyakit karena kesalahan tersebut dapat menyebabkan pasien yang sebenarnya memiliki penyakit jantung diprediksi sebagai normal. Oleh sebab itu, recall kelas penyakit sebesar 91,18% menunjukkan bahwa model memiliki sensitivitas yang cukup baik dalam mendeteksi pasien berisiko.

Kinerja model pada masing-masing kelas juga menunjukkan distribusi performa yang relatif seimbang. Precision kelas normal sedikit lebih tinggi dibandingkan kelas penyakit, sedangkan recall kelas penyakit lebih tinggi dibandingkan recall kelas normal. Kondisi tersebut menunjukkan bahwa model cenderung lebih sensitif dalam mendeteksi penyakit jantung walaupun masih terdapat beberapa data normal yang diprediksi sebagai penyakit. Dalam konteks klasifikasi medis, kondisi tersebut masih dapat diterima karena kesalahan false negative umumnya lebih berisiko dibandingkan false positive.

Nilai AUC sebesar 0,9466 memperlihatkan bahwa model memiliki kemampuan diskriminasi yang sangat baik. Kurva ROC yang berada jauh di atas garis diagonal menunjukkan bahwa model mampu memberikan probabilitas prediksi yang lebih tinggi pada pasien penyakit jantung dibandingkan pasien normal. AUC menjadi indikator penting karena evaluasi tidak hanya bergantung pada satu threshold klasifikasi tertentu. Dengan demikian, peningkatan performa model tidak hanya terlihat pada accuracy, tetapi juga pada kualitas pemisahan antar-kelas.

Dibandingkan penelitian sebelumnya yang menggunakan KNN berbasis PSO konvensional, penelitian ini menunjukkan peningkatan accuracy sebesar 0,62 poin persentase dan peningkatan AUC sebesar 0,0116. Walaupun peningkatan accuracy relatif kecil, peningkatan tersebut tetap menunjukkan adanya kontribusi dari mekanisme adaptive inertia weight dan ensemble learning terhadap kualitas klasifikasi. Selain itu, recall kelas penyakit juga meningkat dibandingkan

penelitian sebelumnya, yang menunjukkan bahwa metode yang diusulkan lebih sensitif dalam mendeteksi pasien penyakit jantung.

4. KESIMPULAN

Penelitian ini mengusulkan optimasi seleksi fitur berbasis Adaptive Particle Swarm Optimization (APSO) untuk klasifikasi penyakit jantung menggunakan ensemble learning berbasis soft voting. APSO menerapkan mekanisme adaptive inertia weight yang menurunkan nilai inertia weight dari 0,90 menjadi sekitar 0,42 selama 30 iterasi untuk menyeimbangkan eksplorasi dan eksploitasi, sehingga diperoleh subset fitur terbaik yang digunakan sebagai masukan model ensemble. Berdasarkan hasil eksperimen, model yang diusulkan memperoleh akurasi sebesar 89,71%, F1-score sebesar 0,8986, recall kelas penyakit sebesar 91,18%, dan AUC sebesar 0,9466. Dibandingkan metode KNN yang dioptimasi PSO pada penelitian sebelumnya, pendekatan yang diusulkan meningkatkan akurasi sebesar 0,62 poin persentase dan AUC sebesar 0,0116, serta menghasilkan confusion matrix dengan hanya 9 false negative dari 102 data penyakit jantung pada data uji. Hasil tersebut menunjukkan bahwa kombinasi seleksi fitur adaptif dan ensemble learning berpotensi meningkatkan kemampuan diskriminasi model dalam klasifikasi penyakit jantung. Meskipun demikian, penelitian selanjutnya perlu menerapkan protokol pembagian data yang identik, validasi silang berulang, pengujian pada dataset eksternal, serta metode interpretabilitas agar keandalan dan transparansi model dapat dievaluasi secara lebih komprehensif.

REFERENCES

- [1] R. Pasaribu, E. Simbolon, T. Supriadi Sianturi, R. Sains Laia, and F. Matematika dan Ilmu Pengetahuan Alam, “Prediksi Penyakit Jantung Menggunakan Metode Decision Tree Berdasarkan Faktor Klinis Pasien,” *Jurnal Matematika dan Aplikasi*, vol. 01, no. 2, pp. 61–65, 2025, [Online]. Available: <https://ejournal.samudrailmu.com/index.php/jma>
- [2] R. Y. Firdaus and H. Mulyani, “Penerapan Data Mining untuk Prediksi Penyakit Jantung Menggunakan Metode Decision Tree,” *JRIT*, vol. 3, no. 1, pp. 65–72, 2026.
- [3] K. Chandra, D. Juan, and S. Prasetyo, “Prosiding SENAM 2024: Prediksi Penyakit Jantung Koroner Menggunakan Metode K-NN dan Regresi Logistik Berdasarkan Kerangka Kerja CRISP-DM,” *Sistem Informasi & Informatika*, vol. 4, pp. 241–248, 2024.
- [4] A. Wahid and A. Muliawan, “Strategi Retensi Pelanggan Berbasis Historis: Optimalisasi Model Prediksi Churn Menggunakan Machine Learning,” *SemanTIK : Teknik Informasi*, vol. 11, no. 2, Dec. 2025, doi: 10.55679/semantik.v11i2.237.
- [5] I. Gusti, A. Ngurah, R. Semadi, M. Samsudin, and K. Dharmendra, “2,3,4 Sistem Informasi; Institut Teknologi dan Bisnis STIKOM Bali; Jalan Raya Puputan No. 86 Renon, Denpasar, (0361) 244445; Cara sitasi: Edwar,” *Journal of informatics*, vol. 8, no. 1, pp. 11–18.
- [6] M. I. Adrian and E. A. Laksana, “Optimalisasi Parameter Support Vector Machine dengan Algoritma PSO untuk Tugas Klasifikasi Sentimen Ulasan IMDb,” *Jurnal Algoritma*, vol. 22, no. 1, pp. 288–299, Jun. 2025, doi: 10.33364/algoritma/v.22-1.2306.
- [7] J. Jusnita and D. Denur, “Analisis Perbandingan Suhu Rem Tromol dengan Penggunaan Fluida Sebagai Pendingin: Pendekatan Eksperimental,” *Jurnal Surya Teknik*, vol. 10, no. 2, pp. 908–911, Dec. 2023, doi: 10.37859/jst.v10i2.6440.
- [8] A. H. Daeli and D. Anubhakti, “OPTIMASI K-MEDOIDS DENGAN PCA UNTUK KLASTERISASI INDIKATOR KESEHATAN IBU HAMIL DI PUSKESMAS LAHOMI,” 2026. [Online]. Available: <http://jom.fti.budiluhur.ac.id/index.php/IDEALIS/indexAnisahHasratniwatiDaeli><http://jom.fti.budiluhur.ac.id/index.php/IDEALIS/index>
- [9] D. Yulvida, S. Quinevera, R. Mardianto, and S. Joses, “Klasifikasi Pemohon Pinjaman dengan Hyperparameter Tuning dan Teknik Penyeimbangan Data,” *Journal of Applied Computer Science and Technology*, vol. 6, no. 2, pp. 92–100, Dec. 2025, doi: 10.52158/kjtrh05.
- [10] Hesti Sabrila Aulia, Muhammad Arifin, and Diana Laily Fithri, “PENERAPAN ALGORITMA MACHINE LEARNING UNTUK PENGELOMPOKAN SISWA BERDASARKAN ASPEK AKADEMIK DAN NON-AKADEMIK,” *Rabit : Jurnal Teknologi dan Sistem Informasi Univrab*, vol. 11, no. 1, pp. 1200–1210, Jan. 2026, doi: 10.36341/rabit.v11i1.7249.
- [11] K. Dwi, N. Cahyo, and M. F. Dwidasaputra, “Optimasi Penempatan Node pada WSN dengan Adaptive PSO untuk Efisiensi Energi dan Cakupan Optimal,” *Karapan Network S Journal*, vol. I, No.I, 2025, doi: 10.20473/KNJ.X.X.134-145.
- [12] A. M. Safira, T. Trimono, and K. M. Hindrayani, “PREDIKSI HARGA SAHAM DI INDONESIA DENGAN EXTREME GRADIENT BOOSTING YANG DIOPTIMALKAN OLEH ADAPTIVE PARTICLE SWARM OPTIMIZATION,” *Jurnal Sistem Informasi dan Informatika (Simika)*, vol. 9, no. 1, pp. 1–12, Feb. 2026, doi: 10.47080/5r67ag12.
- [13] A. Latifa, N. HIKMAH, H. Kurniawan, K. Rohmat Hidayat, N. Larasati, and Rumini, “Implementasi Deep Learning Algoritma Convolutional Neural Network untuk Klasifikasi Kesegaran Buah dan Sayur,” *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 13, no. 2, pp. 319–328, Apr. 2026, doi: 10.25126/jtiik.2026131.
- [14] G. Urva, K. Azmi, and W. Desriyati, “Model Optimalisasi Pemilihan Ekstrakurikuler Menggunakan Algoritma Particle Swarm Optimization (PSO),” *Jurnal Algoritma*, vol. 22, no. 2, pp. 268–276, Nov. 2025, doi: 10.33364/algoritma/v.22-2.2488.
- [15] A. N. Aisyi Maulidhia, M. Ardiana, M. Syai’in, Y. Andika, and R. S. F. J. Putra, “Perbandingan Metode Particle Swarm Optimization dan Firefly Algorithm untuk Optimasi Virtual Inertia Control Berbasis Capacitor Energy Storage,” *Jurnal Riset Rekayasa Elektro*, vol. 7, no. 2, pp. 237–250, Dec. 2025, doi: 10.30595/jrre.v7i2.28429.
- [16] L. M. Cendani and A. Wibowo, “Perbandingan Metode Ensemble Learning pada Klasifikasi Penyakit Diabetes,” *Jurnal Masyarakat Informatika*, vol. 13, no. 1, pp. 33–44, May 2022, doi: 10.14710/jmasif.13.1.42912.
- [17] V. I. Yani, A. Aradea, and H. Mubarak, “Optimasi Prakiraan Cuaca Menggunakan Metode Ensemble pada Naïve Bayes dan C4.5,” *Jurnal Teknik Informatika dan Sistem Informasi*, vol. 8, no. 3, Dec. 2022, doi: 10.28932/jutisi.v8i3.5455.

- [18] Ihya Bahrul Alam, Hasbi Firmansyah, and Wahyu Asriyani, “Evaluasi Klasifikasi Akurasi dan Weighted Mean Precision pada Gradient Boosted Trees untuk Risiko Diabetes Awal,” *Jurnal Dinamika Informatika*, vol. 15, no. 1, pp. 68–85, Mar. 2026, doi: 10.31316/jdi.v15i1.423.
- [19] E. Safitri, R. Heppy Ria Sibarani, Y. SM Sidabutar, and D. Kiswanto, “KLASIFIKASI PENYAKIT DAUN ANGGUR BERBASIS CITRA MENGGUNAKAN METODE K-NEAREST NEIGHBORS (KNN),” *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 8, no. 6, pp. 12633–12642, Nov. 2024, doi: 10.36040/jati.v8i6.12004.
- [20] N. Khoirun Nisa, D. Ahkam Sani, and M. Udin, “PREDIKSI DINI RESIKO PENYAKIT DIABETES MENGGUNAKAN JARINGAN SYARAF TIRUAN BACKPROPAGATION,” *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 9, no. 5, pp. 9096–9102, Jul. 2025, doi: 10.36040/jati.v9i5.15134.
- [21] M. A. Saputra and T. Sugihartono, “Evaluasi Kinerja Model LSTM Untuk Prediksi Risiko Penyakit Jantung Menggunakan Dataset,” *Jurnal Pendidikan dan Teknologi Indonesia*, vol. 5, no. 7, pp. 1823–1833, Jul. 2025, doi: 10.52436/1.jpti.821.
- [22] D. Fabiyanto and Z. Pratama Putra, “Validasi Efektivitas Logistic Regression untuk Diagnosa Penyakit Jantung melalui Pendekatan Machine Learning,” *Jurnal Ilmiah FIFO*, vol. 16, no. 2, p. 158, Nov. 2024, doi: 10.22441/fifo.2024.v16i2.006.
- [23] J. Homepage, E. W. Solang, F. X. Adu, A. Dharma, and N. Gunantara, “MALCOM: Indonesian Journal of Machine Learning and Computer Science Machine Learning Evaluation for Hotel Cancellation Prediction with Threshold Adjustment and Cost-Based Evaluation Evaluasi Machine Learning untuk Prediksi Pembatalan Hotel dengan Threshold Adjustment dan Cost-Based Evaluation,” vol. 6, no. 1, pp. 193–204, 2026, doi: 10.57152/malcom.v6i1.2466.
- [24] W. D. Gumilang, R. Safitri, and L. Riyandari, “Komparasi Algoritma Machine Learning dengan SMOTE untuk Prediksi Retensi Donor Darah,” *Blend Sains Jurnal Teknik*, vol. 4, no. 4, pp. 827–839, Apr. 2026, doi: 10.56211/blendsains.v4i4.1658.
- [25] Adil Setiawan, Andri Armaginda Siregar, N. Setiawan, Jalaluddin Nasution, Naufal Dhiya Putra Dalimunthe, and Farhan Sardy Abdillah, “Optimasi Performa Model SVM dan Random Forest untuk Klasifikasi Kanker Payudara Menggunakan Penyetelan Hyperparameter,” *Jurnal Komputer Teknologi Informasi Sistem Informasi (JUKTISI)*, vol. 4, no. 3, pp. 2141–2149, Jan. 2026, doi: 10.62712/juktisi.v4i3.789.