

# Pengoptimalan Seleksi Fitur Berbasis Particle Swarm Optimization pada Prediksi Gagal Jantung dengan Random Tree

Ridwansyah <sup>a,1,\*</sup>, Sri Rahayu <sup>a,2</sup>, Jajang Jaya Purnama <sup>a,3</sup>, Verry Riyanto <sup>a,4</sup>, Abdul Hamid <sup>a,5</sup>

<sup>a</sup>Universitas Bina Sarana Informatika, Jln. Kramat Raya No. 98, Jakarta, Indonesia 10450

<sup>1</sup> ridwansyah.rid@bsi.ac.id \*; <sup>2</sup> sri.rry@bsi.ac.id; <sup>3</sup> jajang.jjp@bsi.ac.id; <sup>4</sup> verri.vry@bsi.ac.id; <sup>5</sup> hamid.adh@bsi.ac.id

\* Penulis Koresponden

## INFO ARTIKEL

### Histori Artikel

Pengajuan 2026-01-09

Diperbaiki 2026-06-17

Diterima 2026-06-18

### Kata Kunci

Feature Selection,  
Heart Failure,  
Particle Swarm,  
Optimization (PSO),  
Random Tree

## ABSTRAK

Gagal jantung merupakan salah satu penyebab utama rawat inap dan kematian, terutama pada populasi lanjut usia. Deteksi dini sangat penting untuk mendukung pengambilan keputusan klinis yang lebih efektif. Penelitian ini bertujuan mengembangkan model prediksi gagal jantung menggunakan algoritma Random Tree yang dioptimalkan dengan Particle Swarm Optimization (PSO) untuk seleksi fitur. Random Tree dipilih karena sederhana dan mudah diinterpretasikan, sedangkan PSO digunakan untuk mengidentifikasi fitur yang paling relevan serta mengurangi fitur yang tidak penting. Dataset yang digunakan berasal dari UCI Machine Learning Repository dan terdiri dari 299 data pasien dengan 12 atribut klinis. Kinerja model dievaluasi menggunakan akurasi, presisi, recall, dan Area Under the Curve (AUC). Hasil menunjukkan bahwa model Random Tree tanpa optimasi memperoleh akurasi 75,58% dan AUC 0,632. Setelah penerapan PSO, akurasi meningkat menjadi 82,27% dan AUC menjadi 0,740. Temuan ini menunjukkan bahwa integrasi PSO dan Random Tree efektif meningkatkan performa prediksi gagal jantung serta berpotensi mendukung keputusan klinis.

## ABSTRACT

### Keyword

Feature Selection,  
Heart Failure,  
Particle Swarm,  
Optimization (PSO),  
Random Tree

*Heart failure is one of the leading causes of hospitalization and mortality, particularly among older adults. Early detection is essential to support effective clinical decision-making. This study aims to develop a heart failure prediction model using the Random Tree classification algorithm optimized with Particle Swarm Optimization (PSO) for feature selection. Random Tree was chosen for its simplicity and interpretability, while PSO was employed to identify the most relevant features and remove less important ones. The dataset was obtained from the UCI Machine Learning Repository and consists of 299 patient records with 12 clinical attributes. Model performance was evaluated using accuracy, precision, recall, and Area Under the Curve (AUC). The baseline Random Tree model achieved an accuracy of 75.58% and an AUC of 0.632. After applying PSO-based feature selection, the optimized model achieved an accuracy of 82.27% and an AUC of 0.740. These findings indicate that integrating PSO with Random Tree effectively improves heart failure prediction performance and has potential as a clinical decision-support tool.*

Ini adalah artikel akses terbuka di bawah lisensi [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/).



## 1. Pendahuluan

Dalam era transformasi digital, pengelolaan data kesehatan menjadi komponen krusial dalam upaya peningkatan kualitas layanan medis terutama dalam penanganan penyakit kronis seperti gagal jantung [1]. Gagal jantung merupakan salah satu penyebab utama rawat inap pada kelompok usia lanjut serta memberikan dampak ekonomi dan sosial yang signifikan [2]. Saat ini, data klinis tersedia dalam jumlah besar dari berbagai sumber seperti perangkat monitoring, rekam medis elektronik, dan sistem pencatatan kesehatan lainnya [3]. Namun, tantangan utama bukan lagi pada aspek pengumpulan data, melainkan bagaimana data tersebut dapat diolah menjadi informasi yang bermakna dan bermanfaat dalam pengambilan keputusan klinis yang cepat, akurat, dan berbasis bukti [4].

Dalam konteks prediksi gagal jantung permasalahan utama yang dihadapi tidak hanya terbatas pada kelangkaan atau kualitas data, tetapi juga pada kemampuan model prediktif dalam mengekstraksi dan memanfaatkan fitur-fitur yang paling relevan dan berkontribusi terhadap hasil prediksi [5]. Tingginya kompleksitas data klinis menuntut penerapan metode seleksi fitur yang cerdas [6]. Hal ini menjadi penting karena data klinis umumnya memiliki karakteristik yang kompleks, multidimensi, serta mengandung noise atau informasi yang tidak relevan [7]. Oleh karena itu, diperlukan metode seleksi fitur yang efektif untuk menghasilkan prediksi yang akurat, efisien, dan mudah diinterpretasikan oleh tenaga medis [8]. Seleksi fitur yang tepat sangat penting agar model prediktif yang dikembangkan tidak hanya memiliki performa tinggi, tetapi juga dapat digunakan sebagai alat bantu dalam pengambilan keputusan klinis [9].

Salah satu metode klasifikasi yang sederhana, cepat, dan efisien adalah Random Tree yang merupakan turunan dari metode Decision Tree [10]. Metode ini cocok untuk data berdimensi besar, namun performa metode ini sangat dipengaruhi oleh kualitas fitur yang digunakan dalam pelatihan model [11]. Oleh karena itu, pemilihan fitur yang tepat menjadi aspek penting dalam pengembangan model prediksi dengan memastikan model bekerja secara optimal [12]. Salah satu pendekatan yang menjanjikan dalam proses seleksi fitur ini adalah Particle Swarm Optimization (PSO) yang dikenal efektif dalam menemukan kombinasi fitur optimal secara efisien, bahkan pada ruang pencarian yang kompleks menjadikannya sangat cocok untuk masalah optimasi yang kompleks seperti seleksi fitur dalam data klinis [13].

Penelitian tentang prediksi gagal jantung semakin berkembang seiring dengan meningkatnya ketersediaan data kesehatan dan kemajuan teknologi pembelajaran mesin. Salah satu aspek penting dalam pengembangan model prediktif adalah pemilihan fitur yang optimal. Banyak studi membuktikan bahwa kualitas fitur sangat berpengaruh terhadap kinerja model, terutama pada data klinis yang kompleks [14].

Beberapa penelitian sebelumnya telah mengadopsi teknik optimasi seperti PSO dalam konteks prediksi gagal jantung dan membuktikan peningkatan kinerja secara signifikan [15]. Meski demikian, integrasi antara PSO dan algoritma Random Tree masih jarang dieksplorasi secara mendalam dalam literatur. Padahal, kombinasi ini memiliki potensi besar untuk menghasilkan model prediktif yang tidak hanya akurat tetapi juga efisien.

Oleh karena itu, penelitian ini bertujuan untuk mengembangkan model prediksi gagal jantung berbasis Random Tree dengan optimalisasi fitur menggunakan PSO. Penelitian ini berupaya mengevaluasi apakah integrasi PSO dengan algoritma Random Tree dapat meningkatkan akurasi dan efisiensi prediksi gagal jantung dibandingkan pendekatan konvensional. Model yang dihasilkan diharapkan dapat berfungsi sebagai sistem pendukung keputusan klinis yang andal, praktis, dan aplikatif dalam praktik medis sehari-hari.

Sejumlah studi sebelumnya telah menguji kombinasi algoritma pembelajaran mesin dengan teknik optimasi untuk meningkatkan akurasi prediksi misalnya, mengoptimalkan model Neural Network dengan menggabungkan PSO dan berhasil meningkatkan akurasi dari 77,26% menjadi 85,61% [15]. Sementara itu menggabungkan pendekatan Random Forest yang dioptimasi dengan PSO dan memperoleh peningkatan akurasi dari 84,55% menjadi 87,33% [16]. Pendekatan serupa juga diterapkan menggunakan Decision Tree, yang hasilnya

meningkat dari 79,92% menjadi 85,29% setelah optimasi [17]. Penelitian lain membuktikan efektivitas PSO dalam meningkatkan performa model Naive Bayes dari 74,24% menjadi 81,97% [18]. Di sisi lain, mengevaluasi Support Vector Machine dengan kernel dot dan menerapkan Forward Selection untuk mencapai akurasi 84,30% [19]. Terakhir, Chodidjah menggunakan Random Tree yang dioptimalkan dengan FS, dari akurasi awal 75,56% meningkat menjadi 84,28% [20].

Jika diklasifikasikan, metode optimasi seperti PSO terbukti efektif untuk meningkatkan kinerja berbagai algoritma pembelajaran mesin, terutama pada model kompleks seperti Neural Network dan Random Forest. Namun, metode yang lebih sederhana dan interpretatif seperti Random Tree masih jarang dieksplorasi secara mendalam dengan pendekatan metaheuristik seperti PSO. Hal ini menjadi celah penelitian yang penting, mengingat Random Tree menawarkan keunggulan dalam kecepatan dan kemudahan interpretasi faktor penting dalam penerapan klinis.

Dengan demikian, penelitian ini berangkat dari kesenjangan dalam literatur mengenai kombinasi metode Random Tree dan PSO. Diharapkan bahwa integrasi ini tidak hanya dapat meningkatkan akurasi prediksi gagal jantung, tetapi juga mempertahankan efisiensi komputasi dan interpretabilitas, sehingga dapat menjadi solusi praktis dalam sistem pendukung keputusan klinis.

## 2. Metode penelitian

Penelitian ini mengusulkan pengembangan model prediksi gagal jantung berbasis algoritma Random Tree yang dioptimalkan melalui seleksi fitur menggunakan metode PSO. Random Tree digunakan sebagai metode klasifikasi dalam memprediksi gagal jantung dengan varian dari algoritma decision tree yang membentuk pohon keputusan berdasarkan subset acak dari fitur, sehingga dapat menangani data dengan dimensi tinggi dan memberikan hasil klasifikasi yang baik. Setelah proses klasifikasi dengan Random Tree dilakukan, algoritma PSO digunakan untuk pengoptimalan. PSO merupakan algoritma optimasi dengan tingkah laku kawanan burung atau kawanan ikan dalam mencari sumber makanan, maka dengan teknik seperti itu PSO digunakan untuk mencari kombinasi fitur terbaik yang dapat meningkatkan akurasi prediksi model [21]. Setiap partikel dalam PSO merepresentasikan subset fitur, dan partikel-partikel ini bergerak di ruang pencarian berdasarkan pengalaman terbaik individu dan kelompok tersebut. Dari pengalaman terbaik tersebut dapat mengoptimalkan pemilihan fitur, menemukan subset fitur yang paling relevan serta mengurangi fitur yang tidak signifikan, sehingga dapat mempercepat proses pelatihan dan meningkatkan kinerja model prediktif.

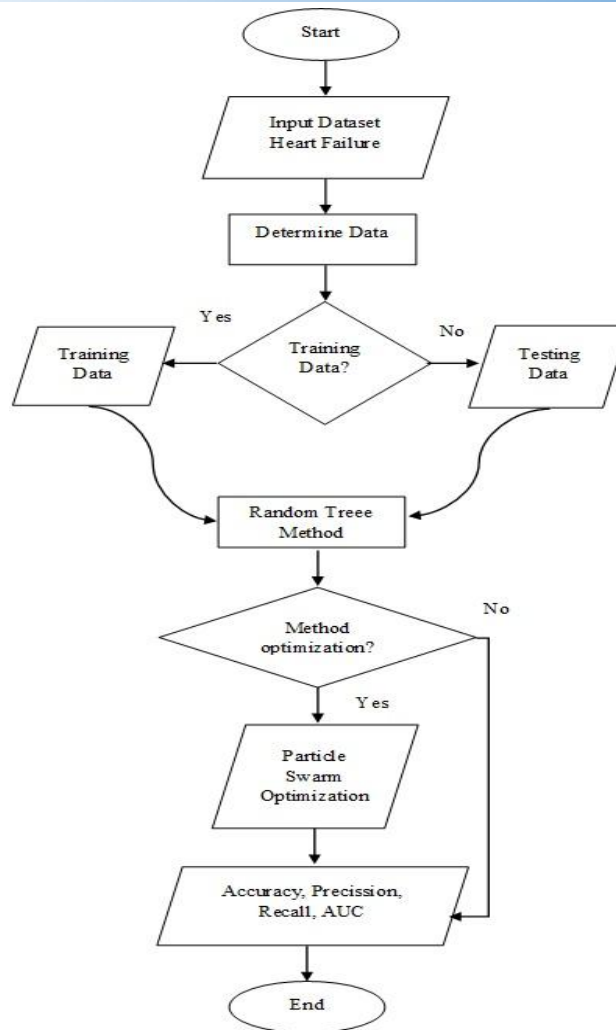
Penelitian ini memerlukan sejumlah perangkat dan bahan pendukung agar proses dan hasil yang diharapkan dapat tercapai. Data yang digunakan dalam penelitian diperoleh dari UCI Machine Learning Repository, dengan judul Heart Failure Clinical Records. Dataset tersebut selanjutnya akan diolah oleh sistem untuk menghasilkan output yang sesuai dengan tujuan penelitian. Salah satu bentuk dataset yang digunakan tersedia dalam format .xls. Dataset ini dapat dilihat pada Gambar. 1

DEATH_EV...	age	anaemia	creatinine_...	diabetes	ejection_fra...	high_blood...	platelets	serum_cre...	serum_sod...	sex	smoking	time
1	75	0	582	0	20	1	265000	1.900	130	1	0	4
1	55	0	7861	0	38	0	263358.030	1.100	136	1	0	6
1	65	0	146	0	20	0	162000	1.300	129	1	1	7
1	50	1	111	0	20	0	210000	1.900	137	1	0	7
1	65	1	160	1	20	0	327000	2.700	116	0	0	8
1	90	1	47	0	40	1	204000	2.100	132	1	1	8
1	75	1	246	0	15	0	127000	1.200	137	1	0	10
1	60	1	315	1	60	0	454000	1.100	131	1	1	10
1	65	0	157	0	65	0	263358.030	1.500	138	0	0	10
1	80	1	123	0	35	1	388000	9.400	133	1	1	10
1	75	1	81	0	38	1	368000	4	131	1	1	10
1	62	0	231	0	25	1	253000	0.900	140	1	1	10
1	45	1	981	0	30	0	136000	1.100	137	1	0	11
1	50	1	168	0	38	1	276000	1.100	137	1	0	11
0	49	1	80	0	30	1	427000	1	138	0	0	12
1	82	1	379	0	50	0	47000	1.300	136	1	0	13
1	87	1	149	0	38	0	262000	0.900	140	1	0	14
1	45	0	582	0	14	0	166000	0.800	127	1	0	14
1	70	1	125	0	25	1	237000	1	140	0	0	15
1	48	1	582	1	55	0	87000	1.900	121	0	0	15
0	65	1	52	0	25	1	276000	1.300	137	0	0	16
1	65	1	128	1	30	1	297000	1.600	136	0	0	20
1	68	1	220	0	35	1	289000	0.900	140	1	1	20

Gambar 1. Contoh Dataset Heart Failure

Pada Gambar 1, dataset ini memiliki karakteristik multivariat dan berasal dari bidang kesehatan. Dataset digunakan untuk berbagai tugas data mining seperti klasifikasi, regresi, dan klustering. Penelitian ini menggunakan tugas datamining klasifikasi dengan terdapat 299 instance (baris data) dan 12 fitur (atribut), dengan tipe data berupa bilangan bulat (*integer*) dan bilangan riil (*real*). Dataset yang digunakan berasal dari data publik terkait pasien gagal jantung yang berisi berbagai atribut klinis seperti usia, tekanan darah, perokok, jenis kelamin, dan lainnya.

Penelitian ini menggunakan pendekatan kuantitatif eksperimental, di mana model prediksi dibangun berdasarkan dataset gagal jantung dan diuji menggunakan pengukuran performa klasifikasi. Penelitian mengenai penyakit gagal jantung menggunakan algoritma Random Tree melibatkan serangkaian tahapan, mulai dari pelatihan model hingga pengujian data. Untuk memperoleh hasil prediksi yang optimal. Tahap awal dalam sistem ini dimulai dengan memasukkan dataset, yaitu proses pengambilan data dari UCI Machine Learning Repository. Setelah data berhasil diperoleh, proses dilanjutkan dengan penentuan sampel yang akan digunakan dalam tahap pelatihan dan pengujian model. Skema alur proses penelitian dapat ditunjukkan pada Gambar 2.



Gambar 2. *Flowchart* Penelitian

Dalam penelitian yang terlihat pada Gambar 2, proses dimulai dari tahap awal yaitu start, yang dilanjutkan dengan memasukkan dataset penyakit gagal jantung yang diambil dari UCI Machine Learning Repository. Dataset ini berisi berbagai fitur klinis seperti usia pasien, tekanan darah, kadar kreatinin, dan lainnya yang relevan dalam memprediksi risiko gagal jantung. Tahap selanjutnya adalah proses pembagian data, di mana sistem menentukan apakah data yang dimasukkan akan digunakan sebagai data pelatihan atau data pengujian. Pemisahan ini dilakukan secara otomatis menggunakan RapidMiner dengan rasio 80:20.

Setelah data terbagi, data pelatihan akan diproses menggunakan metode klasifikasi *Random Tree*. Pada tahap ini, dilakukan pengujian awal terhadap performa model tanpa optimasi untuk mendapatkan model dasar. Kemudian, sistem mengevaluasi apakah diperlukan proses optimasi. Jika ya, maka algoritma PSO diterapkan untuk mengoptimalkan parameter atau pemilihan fitur dari metode *Random Tree*, dengan tujuan meningkatkan akurasi dan efisiensi model.

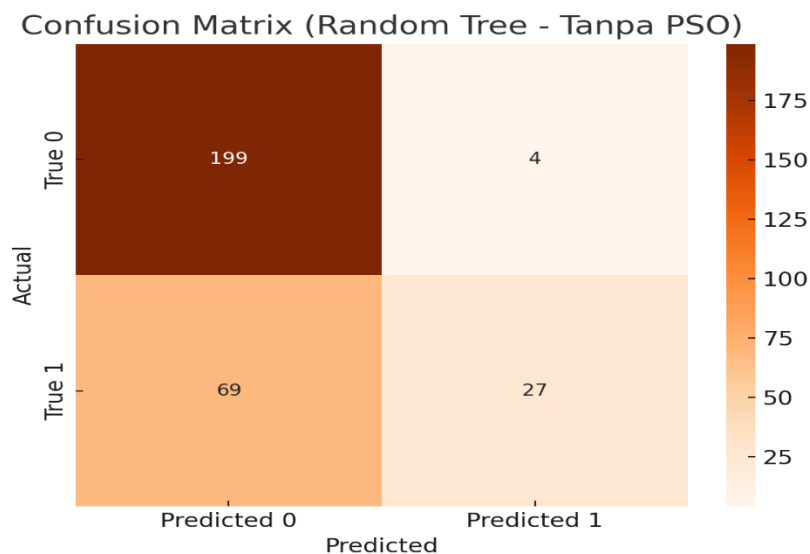
Setelah proses pelatihan dan optimasi selesai, model kemudian dievaluasi menggunakan beberapa metrik performa yaitu akurasi, presisi, *recall*, dan *Area Under the Curve* (AUC). Metrik ini memberikan gambaran komprehensif mengenai kualitas prediksi model baik dari sisi klasifikasi positif maupun negatif. rumus akurasi dihitung berdasarkan proporsi prediksi yang benar terhadap seluruh data, dengan formula sebagai berikut.

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

Hasil dari model Random Tree tanpa optimasi kemudian dibandingkan dengan hasil model yang telah dioptimasi menggunakan PSO. Perbandingan ini bertujuan untuk mengetahui seberapa besar peningkatan performa yang diperoleh setelah diterapkannya optimasi, sehingga dapat ditentukan metode mana yang paling efektif dan efisien dalam mendeteksi potensi gagal jantung pada pasien.

### 3. Hasil dan Analisis

Hasil pengujian awal dilakukan menggunakan algoritma Random Tree tanpa menerapkan metode optimasi apapun, sebagai baseline model. Evaluasi dilakukan dengan menggunakan confusion matrix seperti ditampilkan pada Gambar 3.



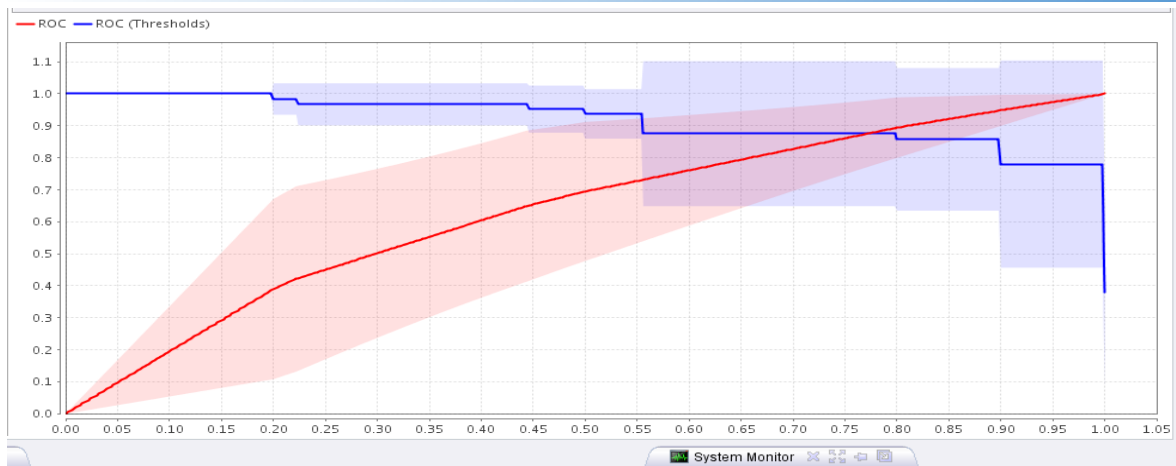
Gambar 3. Confusion Matrix Random Tree (Tanpa PSO)

Berdasarkan Gambar 3, dapat dijelaskan bahwa: Sebanyak 199 data dengan label aktual True 0 berhasil diklasifikasikan dengan benar (True Negative), artinya bahwa prediksi tidak gagal jantung dijawab dengan benar. Sebanyak 4 data dengan label True 0 salah diklasifikasikan sebagai 1 (False Positive), artinya bahwa prediksi gagal jantung dijawab dengan salah yang seharusnya bukan gagal jantung. Sebanyak 27 data dengan label aktual True 1 berhasil diklasifikasikan dengan benar (True Positive), artinya bahwa prediksi gagal jantung dijawab dengan benar. Sebanyak 69 data dengan label True 1 salah diklasifikasikan sebagai 0 (False Negative), artinya tidak gagal jantung dijawab dengan salah yang seharusnya gagal jantung.

$$Accuracy = \frac{226}{229} = 0,755853 \dots \dots \dots (2)$$

Hasil pengujian awal menggunakan algoritma Random Tree tanpa optimasi menghasilkan performa akurasi sebesar 75,58%. Nilai ini menunjukkan bahwa dari seluruh data uji, sekitar tiga per empat data berhasil diklasifikasikan dengan benar. Akurasi yang cukup tinggi ini menunjukkan bahwa model mampu mengenali pola secara umum dalam dataset gagal jantung, meskipun belum tentu seimbang antara kelas positif dan negatif.

Untuk mengevaluasi performa model lebih lanjut, dilakukan pengujian terhadap kurva Receiver Operating Characteristic (ROC) yang menggambarkan kemampuan model dalam membedakan antara dua kelas yaitu pasien gagal jantung dan bukan gagal jantung. Kurva ROC untuk model *Random Tree* ditampilkan pada Gambar 4 berikut:



Gambar 4. Kurva ROC Random Tree (Tanpa PSO)

Berdasarkan Gambar 4, diperoleh nilai Area Under the Curve (AUC) sebesar 0,632. Nilai AUC ini mengindikasikan bahwa kemampuan model dalam membedakan antara pasien gagal jantung dan tidak gagal jantung masih tergolong rendah. Secara interpretatif, AUC mendekati 0,5 menunjukkan performa yang setara dengan tebakan acak, sedangkan AUC mendekati 1,0 menunjukkan klasifikasi yang sempurna. Dengan AUC sebesar 0,632, maka model masih memiliki kemampuan diskriminatif yang lemah dan belum cukup handal untuk diterapkan sebagai alat bantu klinis. Nilai ini juga menegaskan bahwa meskipun model memiliki akurasi yang cukup tinggi kemampuan dalam mendeteksi perbedaan antar kelas masih terbatas, sehingga berisiko menghasilkan prediksi yang bias atau tidak konsisten dalam kasus nyata. Oleh karena itu, diperlukan proses optimasi lanjutan seperti penerapan metode PSO, untuk meningkatkan kualitas seleksi fitur dan akurasi klasifikasi secara keseluruhan.

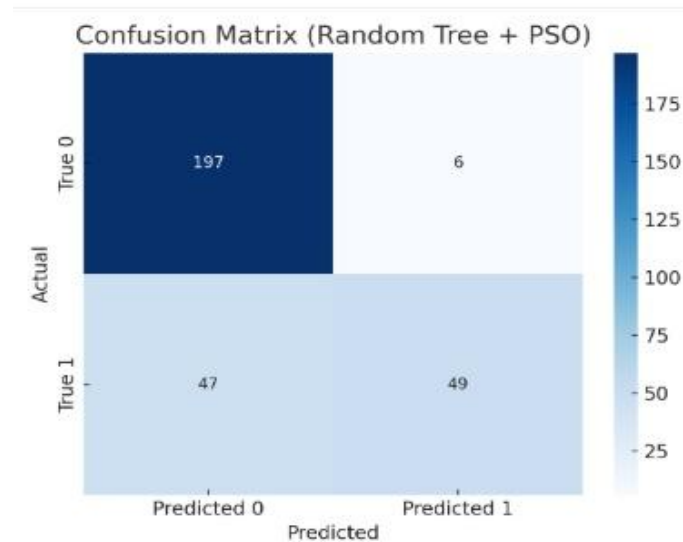
Untuk meningkatkan performa model, dilakukan proses seleksi fitur menggunakan algoritma PSO. PSO bertujuan untuk menemukan subset fitur yang paling relevan dalam membedakan antara pasien gagal jantung dan non-gagal jantung, sekaligus mengurangi fitur yang kurang berpengaruh agar model menjadi lebih efisien dan tidak overfit. Dalam proses ini, setiap partikel pada algoritma PSO merepresentasikan kombinasi fitur tertentu, yang dievaluasi berdasarkan performa model (misalnya akurasi atau AUC). Setelah proses iteratif, diperoleh bobot (weight) kontribusi dari masing-masing fitur terhadap performa klasifikasi, sebagaimana ditunjukkan pada Gambar 5.

attribute	weight
age	1
anaemia	1
creatinine_phosphokinase	0
diabetes	0
ejection_fraction	1
high_blood_pressure	0.056
platelets	1
serum_creatinine	0.320
serum_sodium	1
sex	1
smoking	0
time	0.563

Gambar 5. Bobot Fitur Hasil Seleksi PSO

Berdasarkan Gambar 5, dapat dilihat bahwa fitur-fitur seperti *age*, *anaemia*, *ejection\_fraction*, *platelets*, *serum\_sodium*, dan *sex* memiliki bobot penuh (1.0), yang

menandakan bahwa fitur-fitur ini dianggap sangat penting oleh algoritma PSO dalam meningkatkan prediksi gagal jantung. Sebaliknya, fitur seperti *creatinine\_phosphokinase*, *diabetes*, dan *smoking* mendapatkan bobot nol, yang berarti kontribusinya terhadap akurasi model dianggap sangat rendah atau tidak signifikan. Fitur *serum\_creatinine* dan *time* memiliki bobot menengah (masing-masing 0.320 dan 0.563), yang menunjukkan pengaruhnya cukup relevan tetapi tidak sebesar fitur-fitur utama. Nilai *high\_blood\_pressure* juga memiliki bobot rendah (0.056), menunjukkan kontribusi minor dalam prediksi. Dengan mengeliminasi fitur-fitur yang tidak relevan, model yang dihasilkan setelah optimasi diharapkan memiliki performa yang lebih baik, lebih cepat dalam pelatihan, dan lebih mudah diinterpretasikan secara klinis. Setelah diterapkan metode seleksi fitur menggunakan PSO, dilakukan pelatihan ulang model Random Tree. Hasil evaluasi ditunjukkan pada Gambar 6.

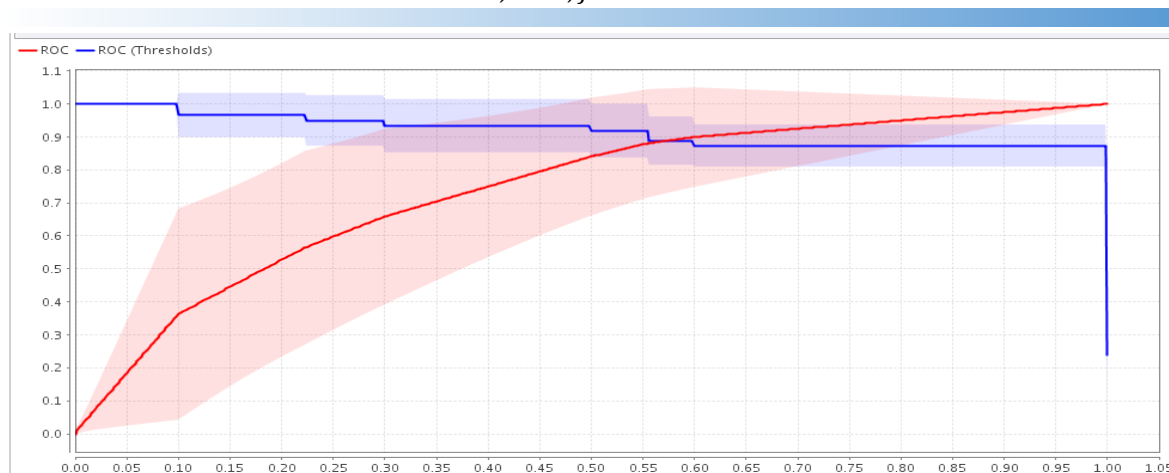


Gambar 6. Confusion Matrix Random Tree dengan PSO

Berdasarkan hasil klasifikasi model *Random Tree* Gambar 6 setelah proses seleksi fitur dengan PSO, diperoleh rincian prediksi sebagai berikut: terdapat 197 kasus negatif (pasien tidak mengalami gagal jantung) yang berhasil dikenali dengan benar oleh model, sementara 6 kasus negatif lainnya salah diklasifikasikan sebagai positif. Untuk kelas positif (pasien dengan gagal jantung), model berhasil mengidentifikasi 49 kasus secara akurat, namun masih terdapat 47 kasus positif yang gagal dikenali, sehingga masuk dalam kategori kesalahan klasifikasi tipe false negative.

$$Accuracy = \frac{246}{299} = 0,822742 \dots \dots \dots (3)$$

Berdasarkan hasil klasifikasi, diperoleh nilai akurasi sebesar 82,27% pada model *Random Tree* yang telah dioptimasi menggunakan PSO. Nilai ini diperoleh dari perbandingan jumlah prediksi yang benar (TP + TN), yaitu 246 data yang diklasifikasikan dengan benar, terhadap total keseluruhan data sebanyak 299. Untuk mengevaluasi performa model lebih lanjut, dilakukan pengujian kembali terhadap kurva ROC yang menggambarkan kemampuan model dalam membedakan antara dua kelas yaitu pasien gagal jantung dan bukan gagal jantung. Kurva ROC untuk model *Random Tree* dengan optimasi PSO ditampilkan pada Gambar 7 berikut:



Gambar 7. Kurva ROC Random Tree dengan PSO

Dalam kurva pada Gambar 7, nilai AUC yang diperoleh sebesar 0.740 yang menunjukkan bahwa model memiliki kemampuan diskriminatif yang cukup baik. Nilai AUC sebesar 0.740 mengindikasikan bahwa model memiliki kemampuan diskriminatif yang cukup baik dalam membedakan pasien gagal jantung dan bukan gagal jantung. Artinya, terdapat peluang sebesar 74% bahwa model dapat secara benar mengidentifikasi perbedaan antara kedua kelas tersebut. Dibandingkan dengan nilai AUC sebelum optimasi (0.632), terjadi peningkatan yang signifikan dalam kemampuan model, yang menegaskan bahwa seleksi fitur menggunakan PSO berhasil meningkatkan kinerja klasifikasi secara keseluruhan.

Peningkatan akurasi ini memperkuat bukti bahwa proses seleksi fitur menggunakan PSO memberikan dampak positif dalam meningkatkan efisiensi dan ketepatan klasifikasi. Dengan jumlah fitur yang lebih sedikit namun lebih relevan, model dapat belajar lebih efektif dari data yang ada, sekaligus mengurangi noise atau informasi yang tidak berguna. Akurasi model meningkat secara signifikan, dan recall meningkat hampir dua kali lipat dibanding model awal. Hal ini menunjukkan adanya perbaikan kinerja model setelah diterapkannya seleksi fitur berbasis PSO.

Hasil penelitian ini menunjukkan bahwa penerapan metode seleksi fitur menggunakan PSO pada algoritma *Random Tree* memberikan dampak positif terhadap performa model dalam memprediksi gagal jantung. Untuk melihat dampaknya secara menyeluruh, dilakukan evaluasi terhadap empat metrik utama: akurasi, presisi, *recall*, dan AUC. Perbandingan kinerja model sebelum dan sesudah seleksi fitur disajikan pada Tabel 1 berikut:

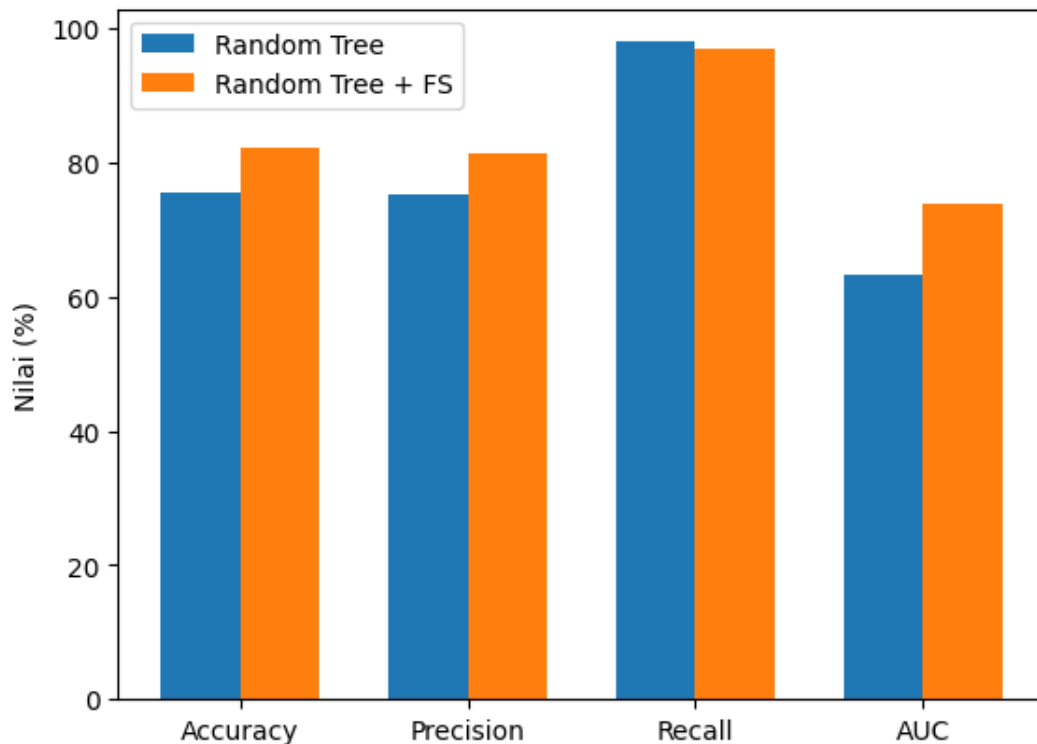
**Tabel 1.** Perbandingan Performa Model Random Tree Sebelum dan Sesudah Optimasi PSO

Algoritma	Random Tree	Random Tree + FS
Accuracy	75.58%	82.27%
Precision	75.38%	81.49%
Recall	98.05%	97.05%
AUC	0.632%	0.740%

Pada tabel 1 Peningkatan akurasi dari 75,58% menjadi 82,27% menunjukkan bahwa model hasil optimasi mampu menghasilkan klasifikasi yang lebih tepat. Begitu juga nilai AUC yang meningkat dari 0,632 menjadi 0,740, mengindikasikan peningkatan kemampuan model dalam membedakan pasien gagal jantung dan non-gagal jantung secara lebih andal. Secara lebih mendalam, peningkatan performa ini dapat dikaitkan dengan kemampuan PSO dalam memilih fitur-fitur yang relevan dan mengeliminasi informasi yang tidak signifikan. Hasil seleksi fitur menunjukkan bahwa fitur seperti *age*, *ejection\_fraction*, *serum\_sodium*, dan *platelets* merupakan faktor penting dalam mendeteksi gagal jantung. Di sisi lain, fitur seperti *smoking*, *diabetes*, dan *creatinine\_phosphokinase* dianggap kurang berpengaruh karena tidak berkontribusi signifikan terhadap akurasi prediksi.

Temuan ini menegaskan bahwa meskipun Random Tree dikenal sebagai algoritma yang cepat dan mudah diinterpretasikan, kualitas fitur input tetap menjadi penentu utama keberhasilan model. Dibandingkan dengan beberapa studi sebelumnya yang menggunakan Neural Network atau Random Forest, hasil penelitian ini menunjukkan bahwa kombinasi metode sederhana dengan teknik optimasi seperti PSO juga mampu menghasilkan performa yang kompetitif. Dengan demikian, hasil penelitian ini menyelesaikan permasalahan klasifikasi klinis dengan pendekatan yang efisien, dan membuka peluang pengembangan sistem prediksi yang praktis, cepat, dan layak digunakan dalam pengambilan keputusan medis.

Gambar 8 berikut memberikan visualisasi yang memperkuat hasil evaluasi antara performa algoritma Random Tree dan Random Tree + PSO.



Gambar 8. Perbandingan Kinerja Random Tree dan Random Tree +PSO

#### 4. Conclusion

Penelitian ini berhasil membuktikan bahwa penerapan metode seleksi fitur menggunakan PSO pada algoritma Random Tree dapat meningkatkan performa model dalam memprediksi gagal jantung. Dengan peningkatan akurasi dari 75,58% menjadi 82,27% dan AUC dari 0,632 menjadi 0,740, model menunjukkan kemampuan klasifikasi yang lebih baik serta potensi aplikasi yang tinggi di lingkungan klinis.

Manfaat dari pendekatan ini adalah terciptanya sistem klasifikasi yang tidak hanya akurat, tetapi juga mudah diinterpretasikan dan efisien secara komputasi. Hal ini penting untuk penerapan nyata dalam sistem pendukung keputusan medis, terutama pada kasus prediksi penyakit kronis seperti gagal jantung yang membutuhkan deteksi dini dan intervensi cepat. Namun demikian, penelitian ini masih memiliki beberapa keterbatasan seperti ukuran dataset yang relatif kecil dan tidak beragam. Untuk studi selanjutnya, disarankan untuk menggunakan dataset dengan ukuran lebih besar, melakukan uji coba dengan data real-time dari rumah sakit, serta membandingkan performa dengan teknik optimasi lain seperti Genetic Algorithm atau Ant Colony Optimization.

## Pengakuan dan Penghargaan

Penelitian ini didukung oleh Fakultas Teknik & Informatika Universitas Bina Sarana Informatika. Dan ucapan terima kasih kepada teman-teman dosen dari Universitas Bina Sarana Informatika yang telah mendukung baik materiil maupun doa.

## References

- [1] R. K. Sachdeva, K. D. Singh, S. Sharma, P. Bathla, and V. Solanki, "An Organized Method for Heart Failure Classification," 2023, doi: 10.1109/ESCI56872.2023.10099809.
- [2] A. Ishaq *et al.*, "Improving the Prediction of Heart Failure Patients' Survival Using SMOTE and Effective Data Mining Techniques," *IEEE Access*, vol. 9, pp. 39707–39716, 2021, doi: 10.1109/ACCESS.2021.3064084.
- [3] D. Chicco and G. Jurman, "The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation," *BMC Genomics*, vol. 21, no. 1, pp. 1–13, 2020, doi: 10.1186/s12864-019-6413-7.
- [4] M. Mamun, A. Farjana, M. Al Mamun, M. S. Ahammed, and M. M. Rahman, "Heart failure survival prediction using machine learning algorithm: am I safe from heart failure?," 2022, doi: 10.1109/IIIoT54504.2022.9817303.
- [5] D. Chicco and L. Oneto, "An Enhanced Random Forests Approach to Predict Heart Failure From Small Imbalanced Gene Expression Data," *IEEE/ACM Trans. Comput. Biol. Bioinforma.*, vol. 18, no. 6, 2021, doi: 10.1109/TCBB.2020.3041527.
- [6] M. Pal and S. Parija, "Prediction of Heart Diseases using Random Forest," *J. Phys. Conf. Ser.*, vol. 1817, no. 1, 2021, doi: 10.1088/1742-6596/1817/1/012009.
- [7] R. A. M. Elghalid, A. Alwirshiffani, A. A. I. Mohamed, F. H. A. Aldeeb, and A. Andiasha, "Comparison of Some Machine Learning Algorithms for Predicting Heart Failure," 2022, doi: 10.1109/ICEMIS56295.2022.9914325.
- [8] Ridwansyah, R. L. Andharsaputri, Yudhistira, Irmawati Carolina, and Suharjanti, "K-Nearest Neighbors Optimization using Particle Swarm Optimization in Selection Digital Payments," *J. Teknol. Inf. dan Terap.*, vol. 12, no. 1, pp. 1–8, 2025, doi: <https://doi.org/10.25047/jtit.v12i1.5911>.
- [9] H. Nurdin, I. Carolina, R. L. Andharsaputri, A. Wuryanto, and Ridwansyah, "Forward Selection as a Feature Selection Method in the SVM Kernel for Student Graduation Data," *Sink. J. dan Penelit. Tek. Inform.*, vol. 8, no. October, pp. 2531–2537, 2024, doi: 10.33395/sinkron.v8i4.14172.
- [10] S. Sartini, S. Sumarna, A. Hamid, A. H. Kahf, and Nicodias Palasar, "Revolusi Diagnosis : Optimasi Random Tree-Pso Untuk Penyakit Ginjal Kronis," *JIRE (Jurnal Inform. Rekayasa Elektron.*, vol. 8, no. 1, pp. 149–158, 2025, doi: <https://doi.org/10.36595/jire.v8i1.1542>.
- [11] I. Ariyati, S. Rosyida, K. Ramanda, V. Riyanto, S. Faizah, and Ridwansyah, "Optimization of the Decision Tree Algorithm Used Particle Swarm Optimization in the Selection of Digital Payments," in *Journal of Physics: Conference Series*, 2020, vol. 1641, no. 1, doi: 10.1088/1742-6596/1641/1/012090.
- [12] W. Widiati, N. Iriadi, I. Ariyati, I. Nawawi, and Sugiono, "Pendekatan Hibrida Decision Tree-Particle Swarm Optimization untuk Deteksi Dini Penyakit Ginjal Kronis," *JASIEK (Jurnal Apl. Sains, Informasi, Elektron. dan Komputer)*, vol. 6, no. 1, pp. 11–21, 2024, doi: 10.26905/jasiek.v6i1.13006.
- [13] H. Nurdin, Suhardjono, A. Wuryanto, D. Yuliandari, and H. Sugiarto, "Naive Bayes and

- Particle Swarm Optimization in Early Detection Of Chronic Kidney Disease," *J. Tek. Inform.*, vol. 5, no. 3, pp. 703-708, 2024, doi: <https://doi.org/10.52436/1.jutif.2024.5.3.1750>.
- [14] T. Prihatin, Sartini, W. E. Pangesti, Yudhistira, and R. Suryadithia, "EFISIENSI PSO DALAM PENGOPTIMALAN NEURAL NETWORK UNTUK CKD," *J. Inform. Manaj. dan Komput.*, vol. 16, no. 2, 2024, doi: <http://dx.doi.org/10.36723/juri.v16i2.707>.
- [15] D. Yuliandari, A. Wuryanto, F. A. Sariasih, Sidik, and F. A. Sariasih, "Improving the Accuracy of Heart Failure Prediction Using the Particle Swarm Optimization Method," *Sink. J. dan Penelit. Tek. Inform.*, vol. 9, no. 1, pp. 210-220, 2024, doi: <https://doi.org/10.33395/sinkron.v9i1.13017> e-ISSN.
- [16] I. Nawawi, "Optimisasi Pemilihan Fitur Untuk Prediksi Gagal Jantung: Fusion Random Forest Dan Particle Swarm Optimization," *INTI NUSA MANDIRI*, vol. 18, no. 2, pp. 122-128, 2024, doi: DOI: <https://doi.org/10.33480/inti.v18i2.5031>.
- [17] Sumarna, Sartini, W. E. Pangesti, R. Suryadithia, and V. Riyanto, "Decision Tree Optimization in Heart Failure Diagnostics: a Particle Swarm Optimization Approach," *J. Tek. Inform.*, vol. 5, no. 3, pp. 739-746, 2024, doi: <https://doi.org/10.52436/1.jutif.2024.5.3.1815>.
- [18] A. Hamid and Ridwansyah, "Optimizing Heart Failure Detection : A Comparison between Naive Bayes and Particle Swarm Optimization," *Paradigma*, vol. 26, no. 1, pp. 30-36, 2024, doi: <https://doi.org/10.31294/p.v26i1.3284>.
- [19] V. Riyanto, H. Destiana, T. Prihatin, Sugiono, and G. Wijaya, "Mengoptimalkan Prediksi Gagal Jantung Dengan Kombinasi Svm Dan Forward Selection," *JIRE (Jurnal Inform. Rekayasa Elektron.*, vol. 8, no. 1, pp. 103-111, 2025, doi: <https://doi.org/10.36595/jire.v8i1.1541>.
- [20] C. Chodidjah, S. Rukiastindari, F. Mutia, L. Rohimah, and A. Aprillia, "Forward Feature Selection for Heart Failure Classification Using the Random Tree Algorithm," *Voteteknika (Vocational Tek. Elektron. dan Inform.*, vol. 13, no. 2, 2025.
- [21] H. Nurdin, I. Nawawi, A. Wuryanto, D. Yuliandari, and H. Sugiarto, "Prediksi Keterlambatan Pembayaran Mahasiswa untuk Mitigasi Risiko Cuti Menggunakan SVM Optimasi PSO," *J. Apl. Sains, Inf., Elektron. dan Komput.*, vol. 7, no. 1, pp. 1-11, 2025, doi: <https://doi.org/10.26905/jasiek.v7i1.15483>.



Ridwansyah, M.Kom lahir di Kota Jakarta, pada 07 Juli 1987. Memperoleh gelar Sarjana Komputer dari Universitas Nusa Mandiri, Indonesia pada 2011, Magister Komputer dari Universitas Nusa Mandiri, Indonesia pada 2014. Saat ini, bekerja di Informatika, Universitas Bina Sarana Informatika sebagai dosen, peneliti dan programming. Minat penelitiannya adalah Datamining, Programming, Sistem Pakar dan Sistem Penunjang Keputusan.

Alamat Email: [ridwansyah.rid@bsi.ac.id](mailto:ridwansyah.rid@bsi.ac.id)



Jajang Jaya Purnama, lahir di Kota Ciamis, Jawa Barat pada hari minggu tanggal 24 Desember 1995, memperoleh gelar Sarjana Komputer dari Universitas BSI Bandung tahun 2018, Magister Ilmu Komputer dari Universitas Nusa Mandiri pada tahun 2020, saat ini bekerja sebagai dosen di Universitas Bina Sarana Informatika dengan bidang ilmu adalah ilmu komputer. Minat penelitiannya adalah Datamining.

Alamat Email: [jajang.jjp@bsi.ac.id](mailto:jajang.jjp@bsi.ac.id)



Sri Rahayu lahir di Kota Tasikmalaya, Jawa Barat pada tanggal 11 Mei 1995, memperoleh gelar Sarjana Komputer dari Universitas BSI Bandung tahun 2018, Magister Ilmu Komputer dari STMIK Nusa Mandiri pada tahun 2020, saat ini bekerja sebagai dosen di Universitas Bina Sarana Informatika. Minat penelitiannya adalah Data Mining.

Alamat Email: [sri.rry@bsi.ac.id](mailto:sri.rry@bsi.ac.id)



Verry Riyanto, M.Kom lahir di Kota Wonogiri, 19 Juli 1988. memperoleh gelar Sarjana dari Universitas Nusa Mandiri, Indonesia pada 2010, Magister dari Universitas Nusa Mandiri, Indonesia pada 2014, dan sedang meneruskan Program Doktorat di IPB. Saat ini, bekerja di Sistem Informasi, Universitas Bina Sarana Informatika sebagai Dosen dan Peneliti. Minat penelitiannya adalah Data Mining, IOT dan Sistem Pakar .

Alamat Email: [verry.vry@bsi.ac.id](mailto:verry.vry@bsi.ac.id)



Abdul Hamid, M.Kom lahir di Kota Cirebon, Jawa Barat, 24 Agustus 1981. memperoleh gelar Sarjana dari Sekolah Tinggi Manajemen Informatika dan Komputer Nusa Mandiri, Indonesia pada 2005, Magister dari Sekolah Tinggi Manajemen Informatika dan Komputer Nusa Mandiri, Indonesia pada 2010, Saat ini, ia bekerja di Sistem Informasi Akuntansi Kampus Kota Bogor, Universitas Bina Sarana Informatika sebagai Dosen dan Peneliti. Minat penelitiannya adalah Data Mining.

Alamat Email: [hamid.adh@bsi.ac.id](mailto:hamid.adh@bsi.ac.id)