Infotek : Jurnal Informatika dan Teknologi
Vol. 9 No. 2, Juli 2026
Hal.432-442

e-ISSN 2614-8773

DOI : 10.29408/jit.v9i2.34799 Link : https://dx.doi.org/10.29408/jit.v9i2.34799

Evaluasi Performa Random Forest Dengan Penyesuaian Threshold Klasifikasi Pada Prediksi Penyakit Jantung

Lidya Shafadhila[1*], Andreas Perdana[2]
[1,2]Program Studi Teknik Informatika, Universitas Dharma Wacana
*lidyashafadhila13@gmail.com

Abstrak

Penyakit jantung merupakan salah satu penyebab utama kematian di dunia sehingga diperlukan sistem prediksi yang akurat untuk mendukung deteksi dini. Penelitian ini bertujuan mengevaluasi performa algoritma Random Forest dalam prediksi penyakit jantung melalui penerapan metode threshold adjustment. Penelitian ini menerapkan threshold adjustment berbasis Youden Index dengan pengujian pada dataset Statlog dan Cleveland untuk menilai konsistensi performa model pada karakteristik data yang berbeda namun memiliki fitur yang serupa. Metode tersebut digunakan untuk menentukan threshold optimal guna memperoleh keseimbangan terbaik antara recall dan specificity. Selain itu, penelitian menganalisis pengaruh perubahan threshold terhadap metrik accuracy, recall, specificity, dan F1-score guna meminimalkan kesalahan false negative serta meningkatkan relevansi klinis hasil prediksi. Pengujian dilakukan menggunakan Cleveland Heart Disease Dataset dan Statlog Heart Disease Dataset dengan teknik stratified sampling. Hasil penelitian menunjukkan bahwa threshold adjustment mampu meningkatkan performa klasifikasi secara signifikan. Dataset Cleveland menghasilkan accuracy sebesar 0.867, recall sebesar 0.893, dan Youden Index sebesar 0.737 pada threshold optimal 4.0, sedangkan dataset Statlog memperoleh accuracy sebesar 0.833, recall sebesar 0.833, dan Youden Index sebesar 0.667 pada threshold 5.0. Secara keseluruhan, kombinasi Random Forest dan threshold adjustment terbukti efektif dalam meningkatkan kualitas prediksi penyakit jantung.

Kata kunci : False negative, Machine Learning, penyakit jantung, Random Forest, threshold adjustment, Youden Index

Abstract

Heart disease remains one of the leading causes of death worldwide, necessitating an accurate prediction system to support early detection. This study aims to evaluate the performance of the Random Forest algorithm in predicting heart disease through the application of a threshold adjustment method. This research implements a Youden Index-based threshold adjustment, tested on the Statlog and Cleveland datasets, to assess the consistency of model performance across different data characteristics that share similar features. This method is utilized to determine the optimal threshold to achieve the best balance between recall and specificity. Additionally, the study analyzes the impact of threshold variations on accuracy, recall, specificity, and F1-score metrics to minimize false-negative errors and enhance the clinical relevance of the prediction results. Testing was conducted using the Cleveland Heart Disease Dataset and the Statlog Heart Disease Dataset with stratified sampling techniques. The results demonstrate that threshold adjustment significantly improves classification performance. The Cleveland dataset yielded an accuracy of 0.867, a recall of 0.893, and a Youden Index of 0.737 at an optimal threshold of 4.0, while the Statlog dataset achieved an accuracy of 0.833, a recall of 0.833, and a Youden Index of 0.667 at a threshold of 5.0. Overall, the combination of Random Forest and threshold adjustment proven to be effective in improving the quality of heart disease predictions.

Keywords : *False negative, Machine Learning, heart disease, Random Forest, threshold adjustment, Youden Index*

Infotek : Jurnal Informatika dan Teknologi
Vol. 9 No. 2, Juli 2026
Hal.432-442

e-ISSN 2614-8773

DOI : 10.29408/jit.v9i2.34799 Link : https://dx.doi.org/10.29408/jit.v9i2.34799

1. Pendahuluan

Penyakit jantung merupakan salah satu penyebab utama kematian di dunia sehingga diperlukan sistem deteksi dini yang akurat. Perkembangan *artificial intelligence* mendorong pemanfaatan *machine learning* dalam diagnosis berbasis data[1], [2]. Salah satu algoritma yang banyak digunakan adalah Random Forest karena mampu menangani data nonlinier, mengurangi *overfitting*, dan menghasilkan performa yang stabil[3], [4].

Penelitian sebelumnya umumnya berfokus pada optimasi parameter dan seleksi fitur[8], [10], [15]. Namun, dalam konteks medis, performa model tidak hanya ditentukan oleh akurasi, tetapi juga kemampuan meminimalkan *false negative*[8]. Selain itu, penggunaan *threshold* statis 0.5 masih menjadi keterbatasan karena belum mempertimbangkan distribusi data[6], [7].

Oleh karena itu, penelitian ini menerapkan *threshold adjustment* berbasis *Youden Index* pada model Random Forest untuk menentukan ambang batas optimal. Penelitian bertujuan mengevaluasi performa model serta menganalisis keseimbangan *recall* dan *specifitiy*[11]. Pengujian dilakukan menggunakan dataset Statlog dan Cleveland untuk menilai konsistensi performa pada karakteristik data yang berbeda dengan fitur yang serupa[12].

Dengan pendekatan ini, penelitian diharapkan dapat meningkatkan akurasi prediksi dan memberikan kontribusi dalam pengembangan sistem pendukung keputusan medis yang lebih adaptif[13], [14], [15].

2. Tinjauan Pustaka

2.1. Penelitian Terkait

Berbagai penelitian mengenai prediksi penyakit jantung telah dilakukan dengan memanfaatkan algoritma *machine learning* untuk meningkatkan akurasi diagnosis. Salah satu algoritma yang banyak digunakan adalah Random Forest karena memiliki kemampuan klasifikasi yang baik pada data medis dengan karakteristik kompleks. Berikut adalah lima penelitian utama yang menjadi acuan dalam penelitian ini :

1. Penelitian oleh S. Swarna dkk.[1] menunjukkan bahwa Random Forest efektif dalam memprediksi penyakit jantung, meskipun evaluasi masih berfokus pada akurasi. S. Rasheed dkk.[3] meningkatkan performa model melalui optimasi *hyperparameter*, tetapi masih menggunakan *threshold* statis. Y. Rimal dkk.[8] juga menunjukkan performa Random Forest yang baik dibandingkan algoritma lain, namun belum menganalisis *false negative* dan perubahan *threshold*.
2. Sementara itu, E. W. Solang dkk.[6] membuktikan bahwa *threshold adjustment* dapat meningkatkan keseimbangan performa model, dan O. A. Abioye serta M. E. Irhebhude[11] menunjukkan bahwa *Youden Index* efektif dalam menentukan *threshold* optimal berdasarkan keseimbangan *recall* dan *specificity*.

Infotek : Jurnal Informatika dan Teknologi
Vol. 9 No. 2, Juli 2026
Hal.432-442

e-ISSN 2614-8773

DOI : 10.29408/jit.v9i2.34799 Link : https://dx.doi.org/10.29408/jit.v9i2.34799

Berdasarkan penelitian terdahulu, sebagian besar studi masih berfokus pada optimasi dan akurasi, sedangkan analisis *threshold adjustment* pada Random Forest dengan lebih dari satu dataset masih terbatas. Oleh karena itu, penelitian ini menerapkan *threshold adjustment* berbasis *Youden Index* pada dataset Statlog dan Cleveland untuk memperoleh *threshold* optimal dan meningkatkan kualitas prediksi penyakit jantung.

2.2. Landasan Teori

1. Penyakit Jantung

Penyakit jantung adalah gangguan pada sistem kardiovaskular yang dapat menghambat aliran darah dan menjadi salah satu penyebab utama kematian karena sulit dideteksi sejak dini[1], [2]. Faktor risikonya meliputi usia, tekanan darah tinggi, kolesterol, diabetes, obesitas, merokok, dan riwayat keluarga[8].

2. Machine Learning

Machine learning adalah bagian dari *artificial intelligence* yang memungkinkan komputer belajar dari data untuk melakukan prediksi atau klasifikasi, termasuk dalam bidang kesehatan[2], [9] Salah satu metode yang umum digunakan adalah *supervised learning*[12].

3. Random Forest

Random Forest adalah algoritma *ensemble learning* yang menggabungkan banyak *decision tree* untuk meningkatkan akurasi dan mengurangi *overfitting*, serta efektif pada data medis[1],[3],[10], [13].

4. Threshold Adjustment

Threshold adjustment adalah penyesuaian ambang batas klasifikasi (tidak selalu 0.5) untuk meningkatkan performa model, terutama dalam meningkatkan *recall* dan mengurangi *false negative*[6], [7].

5. Confusion Matrix

Confusion matrix adalah tabel yang digunakan untuk mengevaluasi hasil klasifikasi model. Tabel ini terdiri atas *True Positive* (TP), *True Negative* (TN), *False Positive* (FP), dan *False Negative* (FN). Dari komponen tersebut dapat dihitung nilai *accuracy*, *precision*, *recall*, *specificity*, dan *F1-score*[8], [12].

6. Receiver Operating Characteristic (ROC) dan AUC

Kurva *Receiver Operating Characteristic* (ROC) digunakan untuk melihat kemampuan model dalam membedakan kelas positif dan negatif[9], [12] Sementara itu, *Area Under Curve* (AUC) menunjukkan kualitas model dengan nilai antara 0 - 1. Semakin tinggi nilai AUC, semakin baik performa model[2], [10] .

7. Youden Index

Youden Index merupakan ukuran statistik yang digunakan untuk menentukan titik potong optimal pada klasifikasi biner. Nilai *Youden Index* dihitung dengan persamaan:

Infotek : Jurnal Informatika dan Teknologi
Vol. 9 No. 2, Juli 2026
Hal.432-442

e-ISSN 2614-8773

DOI : 10.29408/jit.v9i2.34799 Link : https://dx.doi.org/10.29408/jit.v9i2.34799

$$J = Sensitivity + Specificity - 1$$

Nilai J tertinggi menunjukkan keseimbangan terbaik antara *recall* dan *specificity*.

8. Metrik Evaluasi Model

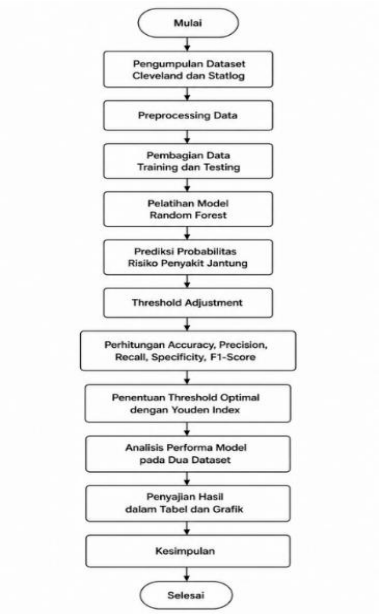
Evaluasi model menggunakan *accuracy*, *precision*, *recall*, *specificity*, dan *F1-score*. *Accuracy* menunjukkan ketepatan prediksi, *recall* untuk mendeteksi kasus positif, *specificity* untuk mengenali kasus negatif, dan *F1-score* untuk melihat keseimbangan *precision* dan *recall*[8], [12], [15].

3. Metode Penelitian

3.1. Tahapan Penelitian

Penelitian ini menggunakan metode kuantitatif dengan pendekatan eksperimen komputasional untuk mengevaluasi model prediksi penyakit jantung. Evaluasi dilakukan menggunakan *accuracy*, *precision*, *recall*, *specificity*, *F1-score*, *ROC-AUC*, dan *confusion matrix*, serta menerapkan *threshold adjustment* berbasis *Youden Index*.

Alur metodologi penelitian secara keseluruhan ditunjukkan pada Gambar 1.



Gambar 1. Tahapan Penelitian

3.2. Sumber dan Jenis Data

Data yang digunakan dalam penelitian ini merupakan data sekunder yang diperoleh dari *UCI Machine Learning Repository*, yaitu *Cleveland Heart Disease Dataset* dan *Statlog Heart Disease Dataset*[1], [3], [8], [12]. Kedua dataset dipilih karena memiliki atribut klinis yang relevan dan sering digunakan dalam penelitian prediksi penyakit jantung. Penggunaan dua dataset bertujuan untuk menguji konsistensi performa model Random Forest pada karakteristik data yang berbeda[2], [9].

Data yang digunakan berupa data kuantitatif dalam format numerik dan kategorikal yang telah dikodekan ke bentuk numerik agar dapat diproses menggunakan algoritma klasifikasi

Jenis Deskripsi masing-masing variabel disajikan pada Tabel 1.

Infotek : Jurnal Informatika dan Teknologi
Vol. 9 No. 2, Juli 2026
Hal.432-442

e-ISSN 2614-8773

DOI : 10.29408/jit.v9i2.34799 Link : https://dx.doi.org/10.29408/jit.v9i2.34799

Table 1. Deskripsi Variabel Dataset Penyakit Jantung

No	Atribut (Fitur)	Nama Kolom	Deskripsi	Tipe Data
1.	Age	age	Usia pasien (tahun).	Numerik
2.	Sex	sex	Jenis kelamin (1 = Laki-laki; 0 = Perempuan).	Kategorik
3.	Chest Pain	cp	Tipe nyeri dada (Skala 1-4).	Kategorik
4.	Resting BP	trestbps	Tekanan darah istirahat (mm Hg).	Numerik
5.	Cholesterol	chol	Kadar kolesterol serum (mg/dl).	Numerik
6.	Fasting BS	fbs	Gula darah puasa > 120 mg/dl (1 = Ya; 0 = Tidak).	Kategorik
7.	Resting ECG	restecg	Hasil elektrokardiografi istirahat (Skala 0-2).	Kategorik
8.	Max HR	thalach	Detak jantung maksimal yang dicapai.	Numerik
9.	Exang	exang	Angina akibat olahraga (1 = Ya; 0 = Tidak).	Kategorik
10.	Oldpeak	oldpeak	Depresi ST akibat olahraga vs istirahat.	Numerik
11.	Slope	slope	Kemiringan puncak segmen ST (Skala 1-3).	Kategorik
12.	CA	ca	Jumlah pembuluh darah utama (0-3).	Numerik
13.	Thal	thal	Status talasemia (Normal, Fixed, Reversable).	Kategorik
-	Target	target	Diagnosis (0 = Sehat; 1 = Berisiko)	Label (Y)

Dataset Cleveland memiliki beberapa *missing values* sehingga dilakukan proses pembersihan data sebelum analisis. Sementara itu, dataset Statlog digunakan sebagai dataset pembanding karena memiliki atribut yang serupa dan data yang lebih terstruktur. Kedua dataset digunakan dalam proses pelatihan, pengujian, dan evaluasi model menggunakan *threshold adjustment* untuk menentukan ambang batas optimal pada prediksi penyakit jantung[6], [11].

3.3. Preprocessing Data

Tahap *preprocessing data* dilakukan untuk menyiapkan dataset agar dapat diproses secara optimal oleh algoritma Random Forest. Penelitian ini menggunakan dua dataset, yaitu *Cleveland Heart Disease Dataset* dan *Statlog Heart Disease Dataset*. Dataset Statlog dibaca menggunakan pemisah spasi, sedangkan dataset Cleveland dibaca dalam format *comma-separated values* (CSV) dengan struktur atribut yang serupa.

Pada dataset Cleveland terdapat data hilang (*missing values*) yang ditandai dengan simbol "?", sehingga nilai tersebut diubah menjadi *null* dan baris yang mengandung data hilang dihapus. Selanjutnya, seluruh data dikonversi ke dalam tipe numerik (*float*). Setelah itu, dilakukan pemisahan variabel input dan variabel target, di mana variabel target pada kedua dataset diseragamkan menjadi klasifikasi biner, yaitu 0 (tidak berisiko) dan 1 (berisiko).

Tahap berikutnya adalah pembagian data menjadi data latih sebesar 80% dan data uji sebesar 20% menggunakan metode *stratified*

Infotek : Jurnal Informatika dan Teknologi
Vol. 9 No. 2, Juli 2026
Hal.432-442

e-ISSN 2614-8773

DOI : 10.29408/jit.v9i2.34799 Link : https://dx.doi.org/10.29408/jit.v9i2.34799

sampling untuk menjaga keseimbangan distribusi kelas. Hasil dari tahap *preprocessing* adalah dataset yang bersih dan siap digunakan untuk pelatihan model serta analisis *threshold adjustment*. Untuk memperjelas perubahan data yang terjadi setelah tahap *preprocessing*, perbandingan kondisi dataset sebelum dan sesudah *preprocessing* disajikan pada Tabel 2 dan 3.

Table 2. Perbandingan Dataset Statlog Sebelum dan Sesudah Preprocessing

Aspek	Sebelum	Sesudah
Jumlah Data	Data lengkap tanpa missing values	Tetap sama
Jumlah Atribut	13 atribut prediktor dan 1 target	Tetap 13 atribut prediktor dan 1 target
Missing Values	Tidak terdapat missing values	Tetap tidak terdapat missing values
Tipe Data	Data numerik	Tetap numerik
Variabel Target	Nilai target terdiri atas 1 dan 2	Disesuaikan menjadi biner: 0 = tidak berisiko, 1 = berisiko
Kesiapan Pemodelan	Hampir siap digunakan	Siap digunakan untuk pelatihan model

Table 3. Perbandingan Dataset Cleveland Sebelum dan Sesudah Preprocessing

Aspek	Sebelum	Sesudah
Jumlah Data	Data asli masih mengandung beberapa baris dengan missing values	Jumlah data berkurang setelah baris yang mengandung data hilang dihapus
Jumlah Atribut	13 atribut prediktor dan 1 target	Tetap 13 atribut prediktor dan 1 target
Missing Values	Terdapat simbol “?” pada beberapa atribut	Tidak terdapat missing values
Tipe Data	Sebagian atribut masih terbaca sebagai string/objek	Seluruh atribut telah dikonversi ke format numerik
Variabel Target	Nilai target terdiri atas 0 sampai 4	Disederhanakan menjadi biner: 0 = tidak berisiko, 1 = berisiko
Kesiapan Pemodelan	Belum siap digunakan langsung	Siap digunakan untuk pelatihan model

Berdasarkan Tabel 2 dan Tabel 3, proses *preprocessing* berhasil menyeragamkan format data, membersihkan *missing values*, serta menyesuaikan variabel target sehingga kedua dataset siap digunakan pada tahap pelatihan dan evaluasi model Random Forest.

3.4. Pembagian Data Training dan Testing

Setelah tahap *preprocessing*, data dibagi menjadi data latih (*training*) dan data uji (*testing*) dengan proporsi 80% : 20%. Data latih digunakan untuk melatih model Random Forest, sedangkan data uji digunakan untuk mengevaluasi kinerja model secara objektif. Pembagian data dilakukan menggunakan metode *stratified sampling* melalui fungsi *train_test_split()* dari pustaka *scikit-learn* dengan parameter *stratify = y* untuk menjaga keseimbangan distribusi kelas. Selain itu, digunakan *random_state = 42* agar hasil pembagian data konsisten dan dapat direproduksi. Hasilnya, dataset Cleveland dan Statlog siap digunakan untuk pelatihan dan evaluasi model dengan *threshold adjustment*.

1. Metode Random Forest

Random Forest adalah algoritma klasifikasi berbasis *ensemble learning* yang

Infotek : Jurnal Informatika dan Teknologi
Vol. 9 No. 2, Juli 2026
Hal.432-442

e-ISSN 2614-8773

DOI : 10.29408/jit.v9i2.34799 Link : https://dx.doi.org/10.29408/jit.v9i2.34799

menggabungkan banyak *decision tree* menggunakan *majority voting* untuk meningkatkan akurasi dan mengurangi *overfitting*. Dalam penelitian ini, model digunakan dengan parameter *n_estimators* = *200* dan *random_state* = *42*, serta dilatih menggunakan data hasil *preprocessing*. Model kemudian menghasilkan prediksi berupa probabilitas melalui fungsi *predict_proba()* yang menunjukkan kemungkinan pasien berisiko atau tidak berisiko penyakit jantung..

2. Prediksi Probabilitas

Model Random Forest menghasilkan probabilitas melalui fungsi *predict_proba()* untuk menunjukkan kemungkinan pasien berisiko. Nilai ini (0–1) menjadi dasar dalam penentuan *threshold* dan dikonversi ke skala 0–10 untuk mempermudah analisis.

3. Threshold Adjustment

Dilakukan penyesuaian ambang batas dari 1.0–10.0 untuk melihat pengaruhnya terhadap performa model. *Threshold* rendah meningkatkan *recall*, sedangkan *threshold* tinggi meningkatkan *specificity*.

4. Penentuan Threshold Optimal (Youden Index)

Threshold terbaik dipilih menggunakan *Youden Index*, yaitu:

$$J = Sensitivity + Specificity - 1$$

Nilai tertinggi menunjukkan keseimbangan terbaik antara *sensitivity* dan *specificity*.

5. Metrik Evaluasi

Kinerja model diukur dengan *accuracy*, *precision*, *recall*, *specificity*, *F1-score*, dan *ROC-AUC* untuk analisis yang menyeluruh.

3.5. Teknik Analisis Data

Analisis dilakukan dengan membandingkan performa model pada berbagai *threshold* dan memilih yang optimal. Fokus utama adalah meminimalkan *false negative* agar deteksi risiko penyakit jantung lebih akurat.

4. Hasil dan Pembahasan

4.1. Hasil Penelitian

Bagian ini menyajikan hasil pengujian dan pembahasan model Random Forest dengan metode *threshold adjustment* pada prediksi penyakit jantung menggunakan dataset Statlog dan Cleveland.

Hasil ditampilkan dalam bentuk tabel dan grafik yang menunjukkan pengaruh variasi *threshold* terhadap *accuracy*, *precision*, *recall*, *specificity*, *F1-score*, dan *Youden Index*. Selanjutnya, ditentukan *threshold* optimal berdasarkan nilai *Youden Index* tertinggi untuk memperoleh keseimbangan terbaik antara *recall* dan *specificity*.

1. Hasil Pengujian Model

Penelitian ini menguji model Random Forest pada dataset Statlog dan Cleveland dengan *threshold adjustment*. Evaluasi menggunakan *accuracy*,

Infotek : Jurnal Informatika dan Teknologi
Vol. 9 No. 2, Juli 2026
Hal.432-442

e-ISSN 2614-8773

DOI : 10.29408/jit.v9i2.34799 Link : https://dx.doi.org/10.29408/jit.v9i2.34799

precision, *recall*, *specificity*, *F1-score*, dan *Youden Index*, serta didukung kurva *ROC* dan *confusion matrix*.

2. Hasil Pengujian pada Dataset Statlog

Hasil pengujian model pada dataset *Statlog* disajikan pada Tabel 4. Pengujian dilakukan dengan menerapkan variasi nilai *threshold* untuk mengamati perubahan performa model dalam mengklasifikasikan data penyakit jantung.

Table 4. Hasil Evaluasi Model pada Dataset Statlog

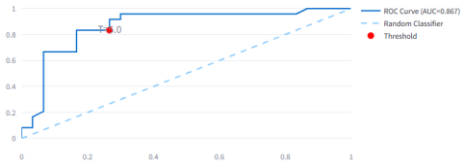
Threshold	*Accuracy*	*Precision*	*Recall*	*F1-Score*	*Youden Index*
1.0	0.519	0.479	0.958	0.639	0.125
1.1	0.519	0.479	0.958	0.639	0.125
1.2	0.556	0.5	0.958	0.657	0.192
1.3	0.574	0.511	0.958	0.667	0.225
1.4	0.63	0.548	0.958	0.697	0.325
1.5	0.648	0.561	0.958	0.708	0.358
1.6	0.667	0.575	0.958	0.719	0.392
1.7	0.685	0.59	0.958	0.73	0.425
1.8	0.704	0.605	0.958	0.742	0.458
1.9	0.759	0.657	0.958	0.78	0.558
2.0	0.778	0.676	0.958	0.793	0.592
....	....	....	....	....	....
10.0	0.556	0.0	0.0	0.0	0.0

Berdasarkan Tabel 4, pada *threshold* rendah (1.0–3.2), model menghasilkan nilai *recall* yang sangat tinggi (0.958), yang menunjukkan kemampuan model dalam mendeteksi hampir seluruh kasus positif. Namun, pada kondisi ini nilai *precision* masih rendah sehingga jumlah *false positive* relatif tinggi.

Seiring meningkatnya nilai *threshold*, *precision* mengalami peningkatan sementara *recall* menurun, yang menunjukkan adanya *trade-off* antara ketepatan prediksi dan kemampuan deteksi kasus positif. Nilai *Youden Index* tertinggi diperoleh pada rentang *threshold* 4.9–5.6, yang menandakan keseimbangan terbaik antara *recall* dan *specificity*.

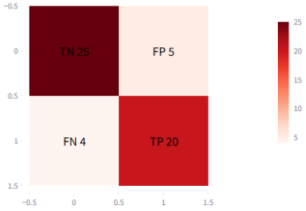
Meskipun beberapa nilai dalam rentang tersebut menghasilkan performa yang serupa, *threshold* 5.0 dipilih sebagai nilai optimal karena lebih mudah diinterpretasikan serta tetap mampu mempertahankan *recall* yang tinggi untuk meminimalkan *false negative* dalam konteks klinis.

Visualisasi hubungan antara *threshold* dan performa model pada dataset *Statlog* ditunjukkan pada Gambar 2 dan Gambar 3.



Gambar 2. Kurva ROC pada Dataset Statlog

Kurva *ROC* pada Gambar 2 digunakan untuk melihat kemampuan model membedakan kelas, sedangkan nilai *AUC* menjadi indikator performa keseluruhan model Random Forest.



Gambar 3. Confusion Matrix Dataset Statlog

Gambar 3 menunjukkan bahwa model mampu mengklasifikasikan sebagian besar data dengan

Infotek : Jurnal Informatika dan Teknologi
Vol. 9 No. 2, Juli 2026
Hal.432-442

e-ISSN 2614-8773

DOI : 10.29408/jit.v9i2.34799 Link : https://dx.doi.org/10.29408/jit.v9i2.34799

benar, dengan tingkat *false negative* yang rendah. Hal ini menandakan model cukup reliabel dalam mendeteksi risiko penyakit jantung, yang penting dalam konteks klinis

3. Hasil Pengujian pada Dataset Cleveland

Hasil pengujian model Random Forest pada dataset *Cleveland* disajikan pada Tabel 5. Evaluasi dilakukan dengan menerapkan variasi nilai *threshold* untuk menganalisis perubahan performa model dalam proses klasifikasi.

Table 5. Hasil Evaluasi Model pada Dataset Cleveland

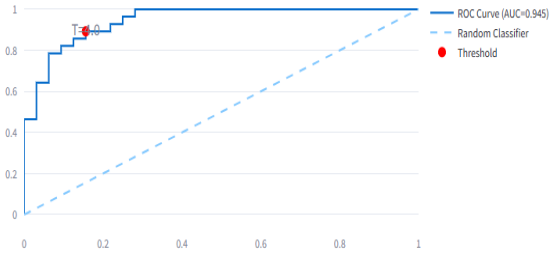
Threshold	***Accuracy***	***Precision***	***Recall***	***F1-Score***	***Youden Index***
1.0	0.683	0.596	1.0	0.747	0.406
1.1	0.717	0.622	1.0	0.767	0.469
1.2	0.767	0.667	1.0	0.8	0.562
1.3	0.783	0.683	1.0	0.812	0.594
1.4	0.783	0.683	1.0	0.812	0.594
1.5	0.783	0.683	1.0	0.812	0.594
1.6	0.8	0.7	1.0	0.824	0.625
1.7	0.817	0.718	1.0	0.836	0.656
1.8	0.817	0.718	1.0	0.836	0.656
1.9	0.817	0.718	1.0	0.836	0.656
2.0	0.817	0.718	1.0	0.836	0.656
....	....	....	....	....	....
10.0	0.533	0.0	0.0	0.0	0.0

Berdasarkan Tabel 5, pada *threshold* rendah (1.0–2.2) model menghasilkan nilai *recall* sebesar 1.000, yang berarti seluruh data positif berhasil dideteksi. Namun, kondisi ini diikuti dengan nilai *precision* yang masih rendah, sehingga jumlah *false positive* relatif tinggi dan model cenderung mengklasifikasikan lebih banyak data sebagai positif.

Seiring meningkatnya *threshold*, nilai *precision* meningkat secara bertahap, sedangkan *recall* menurun. Hal ini menunjukkan adanya *trade-off* antara *recall* dan *precision*, di mana peningkatan *threshold* membuat model lebih selektif dalam menentukan kelas positif, tetapi berpotensi meningkatkan *false negative*.

Nilai *Youden Index* tertinggi diperoleh pada rentang *threshold* 3.9–4.0 sebesar 0.737, yang menunjukkan performa klasifikasi paling optimal. Dari kedua nilai tersebut, *threshold* 4.0 dipilih sebagai nilai optimal karena lebih mudah diinterpretasikan dan memiliki performa yang setara dengan 3.9.

Visualisasi performa model pada *dataset Cleveland* ditunjukkan pada Gambar 4 dan Gambar 5.



Gambar 4. Kurva ROC pada Dataset Cleveland

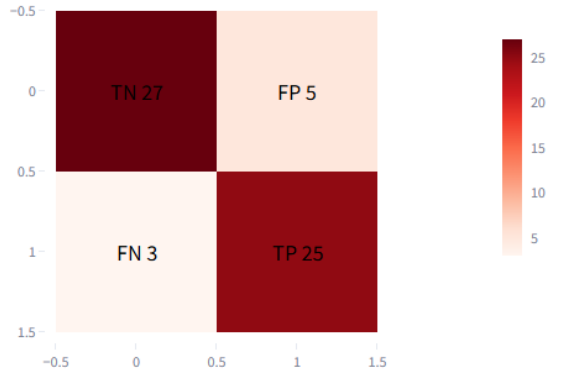
Gambar 4 menunjukkan melalui kurva *ROC* bahwa model memiliki kemampuan pemisahan kelas yang sangat baik. Hal ini didukung oleh nilai

Infotek : Jurnal Informatika dan Teknologi
Vol. 9 No. 2, Juli 2026
Hal.432-442

e-ISSN 2614-8773

DOI : 10.29408/jit.v9i2.34799 Link : https://dx.doi.org/10.29408/jit.v9i2.34799

AUC sebesar 0.945, yang mengindikasikan performa klasifikasi yang tinggi.



Gambar 5. Confusion Matrix Dataset Cleveland

Gambar 5 menunjukkan bahwa sebagian besar data berhasil diklasifikasikan dengan benar, dengan jumlah *false negative* yang relatif rendah. Hal ini menandakan model cukup efektif dalam mendeteksi pasien berisiko, yang penting dalam konteks medis.

Secara keseluruhan, model Random Forest dengan *threshold adjustment* menunjukkan performa yang baik, ditandai dengan nilai *accuracy* dan *F1-score* yang tinggi serta keseimbangan antara *sensitivity* dan *specificity*.

4. Analisis Pengaruh Threshold Adjustment

Berdasarkan hasil pengujian pada kedua dataset, perubahan nilai *threshold* memberikan pengaruh yang signifikan terhadap performa model Random Forest, terutama pada keseimbangan *recall*, *specificity*, *precision*, dan *F1-score*. Pada *threshold* rendah, model menghasilkan *recall* yang tinggi karena sebagian besar data positif berhasil terdeteksi. Namun, kondisi ini menyebabkan *specificity* menurun sehingga *false positive* meningkat. Sebaliknya, pada *threshold* tinggi, *specificity* dan *precision* meningkat karena model lebih selektif, tetapi *recall* menurun sehingga *false negative* bertambah, yang perlu dihindari dalam konteks medis.

Hasil penelitian menunjukkan bahwa *threshold* standar 0.5 tidak selalu optimal. Pada dataset Statlog, *threshold* terbaik berada pada rentang 4.9–5.6 (*Youden Index* = 0.667), sedangkan pada dataset Cleveland berada pada 3.9–4.0 (*Youden Index* = 0.737). Perbedaan ini menunjukkan bahwa karakteristik data memengaruhi nilai *threshold* optimal. Penggunaan *Youden Index* terbukti efektif dalam menentukan *threshold* yang mampu menyeimbangkan *recall* dan *specificity*, sehingga lebih adaptif dibandingkan *threshold* statis. Selain itu, model pada dataset Cleveland menunjukkan performa yang lebih baik dibandingkan Statlog, ditandai dengan nilai evaluasi yang lebih tinggi dan *false negative* yang lebih rendah. Secara keseluruhan, metode *threshold adjustment* mampu meningkatkan kualitas prediksi dan membantu meminimalkan *false negative*, sehingga lebih relevan untuk mendukung pengambilan keputusan dalam bidang medis.

4.2. Pembahasan

Hasil penelitian menunjukkan bahwa penerapan *threshold adjustment* pada algoritma Random Forest mampu meningkatkan keseimbangan

Infotek : Jurnal Informatika dan Teknologi
Vol. 9 No. 2, Juli 2026
Hal.432-442

e-ISSN 2614-8773

DOI : 10.29408/jit.v9i2.34799 Link : https://dx.doi.org/10.29408/jit.v9i2.34799

performa model dalam prediksi penyakit jantung. Penyesuaian nilai *threshold* memberikan fleksibilitas dalam menyesuaikan performa model, terutama untuk meningkatkan deteksi kasus positif dan menekan kesalahan *false negative* yang penting dalam konteks klinis.

Penggunaan *Youden Index* terbukti efektif dalam menentukan *threshold* optimal karena mampu menyeimbangkan *recall* dan *specificity*, serta lebih adaptif dibandingkan *threshold* statis. Hasil pengujian pada dua dataset menunjukkan bahwa nilai *threshold* optimal dipengaruhi oleh karakteristik data, sehingga tidak dapat digeneralisasi. Secara keseluruhan, hasil penelitian ini juga memperkuat temuan sebelumnya bahwa *threshold adjustment* dapat meningkatkan kualitas performa model klasifikasi, khususnya dalam menjaga keseimbangan antara sensitivitas dan spesifisitas. Kontribusi utama penelitian ini terletak pada evaluasi konsistensi performa model Random Forest melalui penerapan *threshold adjustment* berbasis *Youden Index* pada dua dataset berbeda

5. Kesimpulan

Berdasarkan hasil penelitian, penerapan metode *threshold adjustment* pada algoritma Random Forest terbukti mampu meningkatkan keseimbangan performa model dalam prediksi penyakit jantung. Penyesuaian nilai *threshold* berpengaruh signifikan terhadap *recall*, *specificity*, *precision*, dan *F1-score*, sehingga model dapat disesuaikan dengan kebutuhan klinis, terutama untuk menekan *false negative*.

Penentuan *threshold* optimal menggunakan *Youden Index* efektif dalam memperoleh keseimbangan terbaik antara *recall* dan *specificity*. Pada dataset Cleveland diperoleh nilai *Youden Index* sebesar 0,737 dengan *threshold* 4,0, sedangkan pada Statlog sebesar 0,667 dengan *threshold* 5,0, yang menunjukkan pengaruh karakteristik data terhadap hasil optimal. Secara keseluruhan, model Random Forest menunjukkan performa yang baik pada kedua dataset, dengan nilai *accuracy*, *F1-score*, dan *ROC-AUC* yang tinggi. Kombinasi Random Forest dan *threshold adjustment* berbasis *Youden Index* menjadi pendekatan yang efektif dan adaptif dalam meningkatkan akurasi prediksi serta mendukung pengambilan keputusan medis.

6. Daftar Pustaka

[1] S. Swarna, G. Anil, P. Shivamani, and K. Madhu Babu, “Machine Learning Based Heart Disease Prediction Using Random Forest,” *International Journal of Scientific Research in Engineering and Management*, 2024, doi: 10.55041/IJSREM31320.

[2] M. D. Teja and G. M. Rayalu, “Optimizing heart disease diagnosis with advanced machine learning models: a comparison of predictive performance,” *BMC Cardiovasc. Disord.*, vol. 25, no. 1, Dec. 2025, doi: 10.1186/s12872-025-04627-6.

[3] S. Rasheed, G. K. Kumar, D. M. Rani, M. V. V. P. Kantipudi, and M. Anila, “Heart

Infotek : Jurnal Informatika dan Teknologi
Vol. 9 No. 2, Juli 2026
Hal.432-442

e-ISSN 2614-8773

DOI : 10.29408/jit.v9i2.34799 Link : https://dx.doi.org/10.29408/jit.v9i2.34799

Disease Prediction Using GridSearchCV and Random Forest," *EAI Endorsed Trans. Pervasive Health Technol.*, vol. 10, Jan. 2024, doi: 10.4108/eetpht.10.5523.

[4] Z. Noroozi, A. Orooji, and L. Erfannia, "Analyzing the impact of feature selection methods on machine learning algorithms for heart disease prediction," *Sci. Rep.*, vol. 13, no. 1, Dec. 2023, doi: 10.1038/s41598-023-49962-w.

[5] M. Kristanaya, Melinda Putri Azzahra, Trimono, and Mohammad Idhom, "Klasifikasi Status Rujukan Pasien Poliklinik Bandara Berbasis Random Forest dan Interpretabilitas Model Menggunakan SHAP," *Data Sciences Indonesia (DSI)*, vol. 5, no. 1, pp. 60–74, Jul. 2025, doi: 10.47709/dsi.v5i1.6078.

[6] E. W. Solang, F. X. Adu, A. Dharma, and N. Gunantara, "Machine Learning Evaluation for Hotel Cancellation Prediction with Threshold Adjustment and Cost-Based Evaluation," *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 6, no. 1, pp. 193–204, Feb. 2026, doi: 10.57152/malcom.v6i1.2466.

[7] Y. H. Riskandya and A. Tjahyanto, "Integrasi Analytic Hierarchy Process (AHP)–Machine Learning yang Dinamis dalam Prediksi Risiko Kredit," *JURNAL LOCUS: Penelitian & Pengabdian*, vol. 4, pp. 9448–9455, 2025.

[8] Y. Rimal, N. Sharma, S. Paudel, A. Alsadoon, M. P. Koirala, and S. Gill, "Comparative analysis of heart disease prediction using logistic regression, SVM, KNN, and random forest with cross-validation for improved accuracy," *Sci. Rep.*, vol. 15, no. 1, Dec. 2025, doi: 10.1038/s41598-025-93675-1.

[9] A. A. Lamir, S. Razzagzadeh, and Z. Rezaei, "A Comprehensive Machine Learning Framework for Heart Disease Prediction: Performance Evaluation and Future Perspectives," in *3rd National Conference on Computing (QICAR)*, May 2025. Accessed: Apr. 20, 2026. [Online]. Available: https://arxiv.org/abs/2505.09969

[10] A. Naik, G. G. Tejani, and S. J. Mousavirad, "SGO enhanced random forest and extreme gradient boosting framework for heart disease prediction," *Sci. Rep.*, vol. 15, no. 1, Dec. 2025, doi: 10.1038/s41598-025-02525-7.

[11] O. A. Abioye and M. E. Irhebhude, "Big Data-Driven Health Risk Stratification: A Health Index-Based Approach Using Feature Importance and PySpark," *Journal of Computing Theories and Applications*, vol. 2, no. 4, pp. 456–469, Mar. 2025, doi: 10.62411/jcta.12327.

[12] N. Chandrasekhar and S. Peddakrishna, "Enhancing Heart Disease Prediction Accuracy through Machine Learning Techniques and Optimization," *Processes*, vol. 11, no. 4, pp. 1–31, Apr. 2023, doi: 10.3390/pr11041210.

[13] M. G. El-Shafiey, A. Hagag, E. S. A. El-Dahshan, and M. A. Ismail, "A hybrid GA and PSO optimized approach for heart-disease prediction based on random forest," *Multimed. Tools Appl.*, vol. 81, no. 13, pp. 18155–18179, May 2022, doi: 10.1007/s11042-022-12425-x.

[14] B. Kalaivani and A. Ranichitra, "Optimizing Cardiovascular Disease Prediction with a Hybrid Gradient Descent Adaptive Algorithm and Random Forest Classifier," Oct. 2025. [Online]. Available: https://www.jisem-journal.com/

[15] K. Dissanayake and M. G. M. Johar, "Comparative study on heart disease prediction using feature selection techniques on classification algorithms," *Applied Computational Intelligence and Soft Computing*, vol. 2021, 2021, doi: 10.1155/2021/5581806.