

Model Machine Learning yang Dioptimalkan untuk Prediksi Penyakit Jantung Menggunakan R Shiny

Yadburani Dewi Amritha^{1*}, Ni Luh Putu Ika Candrawengi¹, Md Wira Putra Dananjaya¹, Made Ari Riska Dayanti¹

¹*Universitas Pendidikan Nasional, Denpasar, Indonesia*

yadburani@undiknas.ac.id^{*}

Received: 07/12/2025 | Revised: 22/12/2025 | Accepted: 03/01/2026 |

Copyright©2025 by authors. Authors agree that this article remains permanently open access under the terms of the Creative Commons

Abstrak

Penyakit jantung terus menjadi kontributor utama kematian global, yang menyoroti pentingnya deteksi dini dalam meningkatkan luaran pasien. Peningkatan ketersediaan dataset klinis terstruktur telah memungkinkan penerapan sistem cerdas untuk prediksi risiko dan dukungan diagnostik. Dalam makalah ini, efektivitas tiga algoritma pembelajaran terawasi (supervised learning)—Random Forest (RF), Support Vector Machine (SVM), dan Decision Tree (DT)—dievaluasi untuk tugas prediksi penyakit jantung. Investigasi ini didasarkan pada dataset *Heart Failure Prediction* yang bersumber dari platform Kaggle. Proses pelatihan untuk setiap model melibatkan validasi silang 10-lipatan (10-fold cross-validation), dengan hiperparameter yang kemudian dioptimalkan menggunakan pencarian grid (grid search). Efikasi model diukur berdasarkan tolok ukur klasifikasi standar, mencakup akurasi, sensitivitas, spesifisitas, dan area di bawah kurva ROC (AUC). Model Random Forest muncul sebagai yang paling efektif, menunjukkan kinerja unggul dengan AUC sebesar 0,9517, sensitivitas 81,18%, dan spesifisitas 90,44%. Untuk memfasilitasi penggunaan klinis, model ini kemudian diintegrasikan ke dalam alat web ramah pengguna yang dibangun dengan kerangka kerja R Shiny. Antarmuka ini memungkinkan pengguna untuk memasukkan data klinis tingkat pasien dan memperoleh prediksi waktu-nyata (*real-time*), beserta visualisasi kepentingan fitur dan probabilitas risiko. Implementasi ini menjembatani kesenjangan antara pengembangan algoritma dan aplikasi praktis, menawarkan alat pendukung keputusan yang mudah digunakan untuk skrining awal penyakit jantung. Temuan ini menegaskan bahwa model machine learning, ketika dioptimalkan dan divalidasi dengan benar, dapat berfungsi sebagai alat yang efektif dan dapat diinterpretasikan dalam pengambilan keputusan klinis. Karya ini berkontribusi pada kemajuan *e-health* dan integrasi model berbasis AI ke dalam alur kerja medis.

Kata kunci: *machine learning, random forest, biostatistika, R Shiny, e-health*

Abstract

Heart disease continues to be a major contributor to global mortality, highlighting the critical importance of early detection in enhancing patient outcomes. The increasing availability of structured clinical datasets has enabled the application of intelligent systems for risk prediction and diagnostic support. In this paper, the effectiveness of three supervised learning algorithms—Random Forest (RF), Support Vector Machine (SVM), and Decision Tree (DT)—is evaluated for the task of heart disease prediction. This investigation is based on the Heart Failure Prediction dataset sourced from the Kaggle platform. The training process for each model involved a 10-fold cross-validation, with its hyperparameters later being tuned using grid search optimization. Model efficacy was measured against standard classification benchmarks, including accuracy, sensitivity, specificity, and the area under the ROC curve (AUC). The Random Forest model emerged as the most effective, demonstrating superior performance with an AUC of 0.9517, sensitivity of 81.18%, and specificity of 90.44%. To facilitate clinical use, this model was subsequently integrated into a user-friendly web tool built with the R Shiny framework. The interface allows users to input patient-level clinical data and obtain real-time predictions, along with visualizations of feature importance and risk probability. This implementation bridges the gap between algorithm development and practical application, offering a user-friendly decision support tool for early heart disease screening. The findings affirm that machine learning models, when properly tuned and validated, can serve as effective and interpretable tools in clinical decision-making. This work contributes to the advancement of e-health and the integration of AI-driven models into medical workflows.

Keywords: machine learning, random forest, biostatistics, R Shiny, e-health

Pendahuluan

Penyakit jantung, sebagai bagian dari kondisi kardiovaskular, terus menjadi penyumbang utama tingkat kematian global. Menurut World Health Organization (2021), diperkirakan 17,9 juta orang meninggal akibat penyakit kardiovaskular setiap tahunnya, yang mewakili 32% dari seluruh kematian global. Beban penyakit ini tidak hanya berdampak pada angka mortalitas, tetapi juga menciptakan tekanan ekonomi yang signifikan bagi sistem kesehatan di seluruh dunia akibat biaya perawatan jangka panjang dan hilangnya produktivitas. Metode diagnosis konvensional, seperti angiografi koroner, seringkali bersifat invasif, mahal, dan memerlukan keahlian teknis yang tinggi, sehingga tidak selalu dapat diakses dengan cepat di fasilitas kesehatan tingkat pertama (Ansari et al., 2023). Oleh karena itu, deteksi dini dan intervensi tepat waktu sangat penting untuk mengurangi beban penyakit jantung, namun banyak kasus tetap tidak terdiagnosa hingga mencapai tahap kritis.

Kompleksitas diagnosis penyakit jantung terletak pada banyaknya faktor risiko yang terlibat, mulai dari data demografis, tanda-tanda vital, hingga hasil uji laboratorium klinis. Dokter sering kali dihadapkan pada tantangan untuk menganalisis data multivariat ini secara manual, yang rentan terhadap kesalahan manusia atau subjektivitas. Di sinilah teknologi komputasi memegang peranan vital. Evolusi sains data dan teknik machine learning telah memungkinkan pendekatan inovatif untuk deteksi penyakit, khususnya melalui analisis dataset klinis. Teknik

pembelajaran terawasi (supervised learning), termasuk Random Forest (RF), Support Vector Machine (SVM), dan Decision Tree (DT), telah menunjukkan keberhasilan yang cukup besar di bidang prediksi medis, termasuk klasifikasi penyakit jantung (Ansari et al., 2023; Zriqat et al., 2017). Algoritma ini dapat mengidentifikasi pola rumit dalam data pasien dan menghasilkan prediksi probabilistik yang membantu dalam evaluasi risiko dan pengambilan keputusan klinis.

Pemilihan ketiga algoritma tersebut dalam penelitian ini didasarkan pada karakteristik unik masing-masing dalam menangani data medis. Decision Tree menawarkan keunggulan dalam hal interpretabilitas, memungkinkan dokter memahami alur logika keputusan, meskipun sering kali rentan terhadap overfitting. Di sisi lain, Support Vector Machine dikenal efektif dalam menangani data berdimensi tinggi namun memiliki kompleksitas komputasi yang lebih besar. Sementara itu, Random Forest, sebagai metode ensemble, hadir untuk menutupi kelemahan model tunggal dengan meningkatkan akurasi dan stabilitas prediksi melalui agregasi banyak pohon keputusan (Erni & Sa'adah, 2021). Evaluasi komparatif terhadap ketiga metode ini menjadi krusial untuk menentukan model mana yang paling optimal menyeimbangkan antara akurasi prediksi dan efisiensi komputasi dalam konteks data penyakit jantung.

Literatur yang ada menampilkan banyak studi komparatif mengenai kinerja algoritma machine learning konvensional untuk prediksi penyakit jantung. Sebagai contoh, Prasanna Sai Teja & Veeramani (2022) melakukan studi komparatif RF, SVM, dan DT pada dataset UCI Heart Disease dan menemukan bahwa Random Forest menghasilkan akurasi yang lebih tinggi. Demikian pula, studi oleh Erni & Sa'adah (2021) menyoroti bahwa metode ensemble seperti RF dan Extra Trees mengungguli pengklasifikasi individu karena ketahanannya terhadap overfitting. Namun, terdapat kesenjangan dalam literatur di mana evaluasi sering kali terbatas pada penggunaan parameter default tanpa eksplorasi mendalam (Bouthillier et al., 2021). Padahal, studi teoritis dan empiris menunjukkan bahwa kinerja model machine learning sangat bergantung pada hyperparameter tuning yang tepat. Optimasi ini krusial untuk memastikan model tidak hanya menghafal data latih, tetapi juga mampu melakukan generalisasi dengan baik pada data pasien baru yang belum pernah dilihat sebelumnya (Yang & Shami, 2020; Bergstra & Bengio, 2012).

Selain masalah optimasi model, tantangan terbesar dalam adopsi Artificial Intelligence (AI) di dunia medis adalah kesenjangan antara pengembangan algoritma dan implementasi praktis. Banyak model prediksi berkinerja tinggi berakhir hanya sebagai kode eksperimental yang tidak dapat diakses oleh praktisi medis karena ketiadaan antarmuka yang intuitif. Isu "Black Box" dalam AI sering kali membuat dokter enggan menggunakan sistem pendukung keputusan karena kurangnya transparansi dan kesulitan dalam pengoperasian. Sebagian besar studi yang ada terutama berfokus pada kinerja algoritmik dan tidak memperluas model mereka menjadi alat praktis bagi pengguna akhir.

Untuk mengatasi kesenjangan ini, studi ini tidak hanya melakukan evaluasi empiris terhadap tiga model machine learning klasik (RF, SVM, DT) pada dataset Kaggle Heart Failure Prediction dengan proses optimasi grid search yang ketat, tetapi juga mengintegrasikan model dengan kinerja terbaik ke dalam antarmuka interaktif yang ramah pengguna menggunakan R Shiny. Sistem purwarupa ini memungkinkan dokter atau peneliti untuk memasukkan fitur spesifik pasien dan mendapatkan prediksi risiko penyakit secara real-time. Berbeda dengan studi sebelumnya yang berfokus terutama pada pelaporan metrik akurasi, penelitian ini menutup kesenjangan antara desain algoritma dan implementasi klinis dengan menyediakan antarmuka pengguna yang dapat diinterpretasikan dan diterapkan. Dengan memajukan bidang e-health,

penelitian ini meletakkan dasar bagi sistem pendukung keputusan klinis cerdas yang secara efektif mengintegrasikan wawasan berpusat pada data dengan prinsip desain yang berpusat pada pengguna.

Metodologi Penelitian

Studi ini membandingkan kinerja klasifikasi dari tiga algoritma *supervised machine learning* (ML)—*Random Forest* (RF), *Support Vector Machine* (SVM), dan *Decision Tree* (DT)—secara eksperimental untuk tugas deteksi penyakit jantung. Alur kerja metodologis mencakup pra-pemrosesan data, pengembangan model, optimasi hiperparameter menggunakan validasi silang, evaluasi model, dan penyebaran sistem melalui antarmuka grafis yang dibangun dengan R Shiny.

Dataset Analisis

Studi ini didasarkan pada Heart Failure Prediction Dataset dari Kaggle, kumpulan 918 contoh data dan 12 atribut. Atribut ini mencakup campuran variabel, termasuk demografis (misalnya, usia, jenis kelamin), klinis (misalnya, tekanan darah istirahat, kolesterol), dan data elektrokardiografis. Variabel target adalah *HeartDisease*, yang dikodekan sebagai biner (1 = ada penyakit, 0 = tidak ada penyakit).

Pra-pemrosesan Data

Pra-pemrosesan data melibatkan langkah-langkah berikut:

- Menangani variabel kategorikal dengan mengubahnya menjadi tipe faktor (misalnya, Sex, ChestPainType, RestingECG, dll.).
- Pelabelan (label encoding) variabel target menjadi NoDisease dan Disease.
- Kami mempartisi dataset menjadi set pelatihan (70%) dan set pengujian (30%). Teknik pengambilan sampel berstrata (stratified sampling) diterapkan selama pembagian ini untuk memastikan distribusi kelas yang konsisten dipertahankan di kedua subset (Kuhn & Johnson, 2013).
- Tidak ada nilai yang hilang atau pencilan (outliers) yang memerlukan imputasi yang diidentifikasi selama analisis eksplorasi.

Pelatihan dan Penyetelan Model

Setiap model (RF, SVM, DT) dilatih menggunakan kerangka kerja caret di R, yang menawarkan antarmuka yang konsisten dan modular untuk evaluasi model (Lantz, 2019).

- Validasi Silang: Untuk meningkatkan generalisasi model dan mengurangi overfitting, validasi silang 10-lipatan diterapkan menggunakan kode di bawah ini: `trainControl(method = "cv", number = 10)` (8) Metode ini membagi dataset pelatihan menjadi 10 subset yang sama, secara iteratif melatih pada 9 lipatan sambil memvalidasi pada satu lipatan yang tersisa, diulang di seluruh lipatan. Persamaan untuk menghitung rata-rata akurasi validasi silang disediakan di Bagian II-C.
- Penyetelan Hiperparameter: Setiap algoritma menjalani optimasi pencarian grid untuk mengidentifikasi konfigurasi parameter yang paling efektif:
 1. RF: mtry – jumlah fitur yang dipilih secara acak pada setiap pemisahan.
 2. SVM: C – parameter biaya yang mengontrol lebar margin.

3. DT: cp – parameter kompleksitas untuk mengontrol pemangkasan pohon. Parameter yang menghasilkan AUC maksimum di seluruh pemisahan validasi dipilih sebagai optimal.

Metrik Evaluasi

Untuk menilai kinerja model, kami menghitung beberapa metrik klasifikasi standar dari matriks konfusi, yaitu: akurasi, sensitivitas, spesifisitas, presisi, skor F1, dan AUC (Powers, 2011). Rumus rinci dan deskripsi teoretis metrik ini disediakan di Bagian II-B. AUC dihitung menggunakan paket pROC di R (Robin et al., 2011) yang memberikan estimasi kuat tentang kemampuan diskriminatif model.

Pengembangan Antarmuka Pengguna dengan R Shiny

Untuk meningkatkan kegunaan dan relevansi klinis, model dengan kinerja terbaik (RF) disebarluaskan ke dalam antarmuka R Shiny yang interaktif. UI menerima input data pasien secara real-time dan memberikan hasil prediksi langsung serta representasi visual dari kepentingan fitur. Implementasi ini mengikuti prinsip desain yang berpusat pada pengguna (ISO, 2010), yang bertujuan untuk aksesibilitas tinggi, beban kognitif minimal, dan keterjelasan (*explainability*) dalam luaran prediksi (Kushniruk & Patel, 2004).

Hasil dan Pembahasan

Evaluasi Kinerja Model

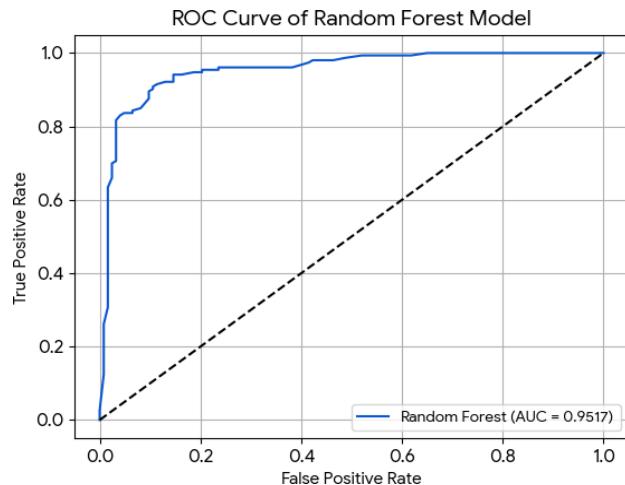
Setelah melatih dan mengoptimalkan ketiga model *machine learning* menggunakan validasi silang 10-lipatan dan penyetelan hiperparameter, kami melakukan evaluasi kinerja akhir. Penilaian ini dilakukan pada set pengujian (*held-out test set*) dari Heart Failure Prediction Dataset, berdasarkan metrik termasuk akurasi, sensitivitas, spesifisitas, dan area di bawah kurva ROC (AUC). Seperti yang ditunjukkan pada Tabel 1, model *Random Forest* (RF) mengungguli SVM dan DT dalam hal ROC-AUC dan spesifisitas. Meskipun sensitivitasnya (81,18%) sedikit lebih rendah daripada SVM, model Random Forest tetap mencapai kinerja yang paling seimbang dan kuat, menunjukkan kemampuan diskriminatif yang kuat. Selanjutnya, nilai AUC dari model Random Forest mencapai 0,9517 (berdasarkan komputasi AUC terpisah), mengonfirmasi kemampuan luar biasanya untuk membedakan antara pasien dengan dan tanpa penyakit jantung.

Tabel 1. Kinerja Model Berdasarkan Hasil Validasi Silang

Model	AUC (ROC)	Sensitivity	Specificity	Tuned Parameter
Random Forest	0.9181	0.8118	0.9044	Mtry=2
SVM	0.9130	0.8046	0.8793	C = 0.5
DT	0.8513	0.7563	0.8312	Cp = 0.01

Analisis Kurva ROC

Gambar 1 mengilustrasikan kurva ROC yang dihasilkan oleh model Random Forest. Posisinya yang signifikan di atas garis dasar diagonal mencerminkan kinerja prediktif yang kuat. Nilai AUC sebesar 0,9517 memperkuat kemampuan diskriminatif model yang kuat dalam membedakan antara kelas target.



Gambar 1. Kurva ROC Model Random Forest

Kepentingan Fitur (*Feature Importance*)

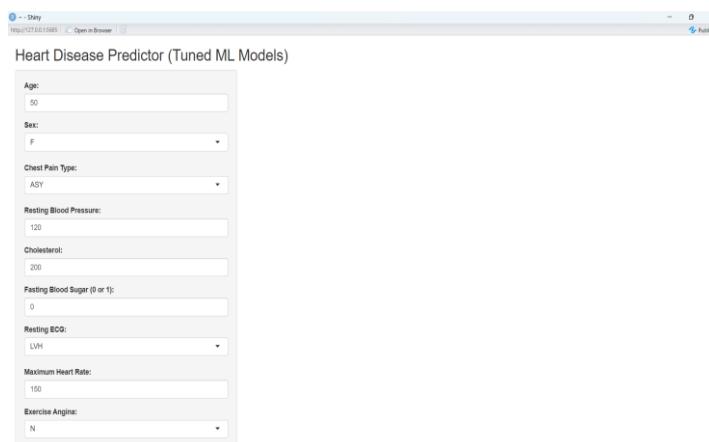
Salah satu kritik utama terhadap algoritma machine learning yang kompleks seperti Random Forest adalah sifatnya yang sering dianggap sebagai "kotak hitam" (black box), di mana mekanisme pengambilan keputusan sulit dipahami oleh manusia. Untuk mengatasi hal ini dan meningkatkan interpretabilitas model dalam konteks klinis, penelitian ini menerapkan analisis kepentingan fitur (feature importance analysis). Dalam algoritma Random Forest, kepentingan fitur diukur untuk menentukan seberapa besar kontribusi setiap variabel prediktor dalam meningkatkan kemurnian (purity) node atau mengurangi ketidakmurnian (impurity) di seluruh pohon keputusan yang dibangun.

Secara spesifik, studi ini menggunakan metrik Mean Decrease Gini atau Mean Decrease Accuracy. Metrik ini menghitung rata-rata penurunan impuritas Gini yang dihasilkan oleh suatu variabel setiap kali variabel tersebut dipilih untuk memisahkan node dalam pohon. Semakin besar nilai penurunan ini, semakin signifikan peran variabel tersebut dalam membedakan antara pasien yang memiliki penyakit jantung dan yang tidak. Implementasi teknis dilakukan menggunakan fungsi varImpPlot() yang tersedia dalam paket caret dan randomForest di lingkungan R. Fungsi ini memvisualisasikan peringkat variabel secara hierarkis, memungkinkan identifikasi cepat terhadap determinan utama penyakit.

Berdasarkan analisis yang dilakukan, visualisasi dari varImpPlot() mengungkapkan bahwa Age (Usia), ChestPainType (Tipe Nyeri Dada), STSlope (Kemiringan segmen ST pada EKG), dan Cholesterol (Kadar Kolesterol) adalah prediktor yang memiliki dampak paling signifikan terhadap performa model. Temuan ini sangat krusial karena memvalidasi bahwa model tidak sekadar mempelajari "kebisingan" (noise) data, melainkan menangkap pola biologis yang nyata.

Integrasi ke dalam Antarmuka R Shiny

Untuk mendemonstrasikan penyebaran praktis, model *Random Forest* dengan kinerja terbaik diintegrasikan ke dalam aplikasi web R Shiny yang interaktif. Seperti yang diilustrasikan pada Gambar 2, antarmuka pengguna memungkinkan input parameter klinis individual, termasuk usia, jenis kelamin, kategori nyeri dada, kadar kolesterol, gula darah puasa, dan atribut pasien relevan lainnya. Setelah dikirimkan, aplikasi mengembalikan prediksi real-time mengenai risiko penyakit jantung, bersama dengan penjelasan visual seperti skor kepentingan fitur. Implementasi ini mengikuti prinsip desain yang berpusat pada pengguna. Dengan menyediakan tata letak berbasis formulir yang bersih, ini memastikan kemudahan penggunaan dan interpretabilitas—keduanya penting untuk adopsi klinis. Arsitektur sistem menjembatani pemodelan berbasis data dengan kegunaan dunia nyata, meningkatkan potensi machine learning dalam dukungan keputusan kesehatan.



Gambar 2. Antarmuka R Shiny untuk Prediksi Penyakit Jantung

Pembahasan

Analisis Komparatif Kinerja Algoritma Hasil eksperimen ini sejalan dengan literatur yang ada di mana model ensemble seperti Random Forest (RF) biasanya mengungguli pengklasifikasi yang lebih sederhana dalam dataset layanan kesehatan (Erni & Sa'adah, 2021). Superioritas Random Forest dalam studi ini dapat dikaitkan dengan mekanisme dasarnya yang memanfaatkan teknik bagging (Bootstrap Aggregating) dan feature randomness. Berbeda dengan Decision Tree tunggal yang membangun satu pohon keputusan dari keseluruhan data, RF membangun hutan yang terdiri dari banyak pohon keputusan yang tidak saling berkorelasi. Dengan mengambil suara mayoritas dari prediksi setiap pohon, RF secara efektif mengurangi varians dan risiko overfitting yang sering menjadi kelemahan utama pada model pohon tunggal. Hal ini sangat krusial dalam data medis yang sering kali memiliki noise atau variasi biologis yang tinggi antar pasien.

Di sisi lain, Support Vector Machine (SVM) juga menunjukkan kinerja yang kuat, khususnya dalam metrik spesifisitas. Keberhasilan SVM terletak pada kemampuannya untuk menemukan hyperplane optimal yang memisahkan kelas data dalam ruang dimensi tinggi. Namun, meskipun akurasinya kompetitif, SVM memiliki kelemahan inheren terkait kompleksitas komputasi saat volume data meningkat dan tantangan dalam interpretabilitas yang lebih rendah. SVM sering dianggap sebagai model "Black Box" karena keputusan klasifikasinya didasarkan

pada transformasi matematis kompleks (kernel trick) yang sulit diterjemahkan ke dalam aturan klinis sederhana yang dapat dipahami oleh tenaga medis awam.

Dilema Akurasi versus Interpretabilitas Model Decision Tree (DT) tetap menjadi yang paling dapat diinterpretasikan karena struktur logisnya yang menyerupai cara berpikir manusia (aturan if-then). Dokter dapat dengan mudah menelusuri alur keputusan dari root hingga leaf untuk memahami mengapa seorang pasien diklasifikasikan berisiko tinggi. Namun, kemudahan interpretasi ini harus dibayar dengan performa prediksi yang inferior, sebagaimana ditunjukkan oleh metrik kinerja keseluruhan yang lebih rendah dibandingkan RF dan SVM. DT sangat sensitif terhadap perubahan kecil pada data latih, yang dapat menyebabkan perubahan drastis pada struktur pohon (instabilitas). Oleh karena itu, dalam konteks diagnosis penyakit kritis seperti penyakit jantung di mana kesalahan prediksi (False Negative) dapat berakibat fatal, mengandalkan DT tunggal mungkin terlalu berisiko meskipun transparansinya tinggi.

Kontribusi Translasional: Dari Model ke Aplikasi Penting untuk dicatat, penelitian ini berkontribusi melampaui metrik akurasi dengan menyebarkan aplikasi praktis. Tantangan terbesar dalam penerapan kecerdasan buatan di dunia kesehatan saat ini bukanlah kurangnya algoritma yang akurat, melainkan kurangnya integrasi ke dalam alur kerja klinis. Sebagian besar model penelitian berhenti pada tahap evaluasi kode dan tidak pernah menyentuh tangan praktisi.

Mengintegrasikan model yang terlatih ke dalam aplikasi R Shiny menjembatani kesenjangan translasional antara evaluasi algoritmik dan implementasi dunia nyata. Aplikasi berbasis web ini mentransformasi model matematika yang abstrak menjadi antarmuka visual yang nyata, memungkinkan simulasi risiko secara instan tanpa memerlukan instalasi perangkat lunak pemrograman yang rumit. Hal ini mendukung konsep Clinical Decision Support System (CDSS) modern, di mana teknologi tidak menggantikan dokter, melainkan bertindak sebagai "opini kedua" yang cerdas dan objektif. Dengan fitur visualisasi probabilitas risiko dan indikator variabel penting, aplikasi ini membantu dokter dalam memprioritaskan pasien dan merencanakan intervensi preventif yang lebih personal, sehingga secara efektif mendukung pengembangan alat pendukung keputusan dalam praktik klinis.

Kesimpulan

Penelitian ini memperkenalkan pendekatan *machine learning* yang kuat untuk prediksi penyakit jantung memanfaatkan dataset klinis terstruktur. Algoritma *Random Forest* (RF) muncul sebagai model unggulan dalam analisis komparatif kami terhadap *Support Vector Machine* (SVM) dan *Decision Tree* (DT). RF mencapai skor ROC AUC tertinggi (0,9517) dan menunjukkan keseimbangan yang kuat antara spesifikasi tinggi (90,44%) dan sensitivitas yang kuat (81,18%).

Untuk meningkatkan ketahanan dan generalisasi model, studi ini menggunakan proses optimasi hiperparameter yang ketat yang menggabungkan validasi silang 10-lipatan dengan pencarian grid. Berbeda dengan banyak studi sebelumnya yang mengandalkan parameter default atau pembagian train-test sederhana, pendekatan ini mendukung stabilitas model di berbagai partisi data.

Di luar kinerja prediktif, penelitian ini menekankan interpretabilitas dan kegunaan model. Alat seperti analisis kurva ROC dan plot kepentingan fitur memungkinkan pengguna untuk

memahami dan mempercayai hasil prediksi—pertimbangan penting dalam pengambilan keputusan klinis di mana transparansi sangat penting.

Yang paling penting, model dengan kinerja terbaik (*Random Forest*) tidak hanya dievaluasi tetapi juga disebarluaskan dalam aplikasi R Shiny yang interaktif, memungkinkan pengguna untuk memasukkan nilai spesifik pasien dan mendapatkan prediksi risiko *real-time*. Antarmuka yang berpusat pada pengguna ini menjembatani kesenjangan antara pengembangan algoritma dan penyebarluasan praktis, meningkatkan dampak potensial dari karya ini dalam lingkungan perawatan kesehatan dunia nyata.

Namun, penelitian ini memiliki keterbatasan tertentu, terutama terkait penggunaan dataset sekunder tunggal yang mungkin belum sepenuhnya menangkap variabilitas demografis pasien secara global atau mencakup fitur klinis dinamis yang lebih kompleks.

Secara ringkas, studi ini menunjukkan bahwa dengan optimasi, validasi, dan desain yang berfokus pada pengguna yang tepat, model machine learning dapat berfungsi sebagai alat yang akurat, dapat diinterpretasikan, dan praktis untuk diagnosis dini penyakit jantung.

Saran untuk Penelitian Selanjutnya:

- Menerapkan model pada dataset yang lebih besar dan lebih beragam untuk generalisasi yang lebih luas dan validasi eksternal.
- Mengintegrasikan fitur klinis tambahan (misalnya, biomarker, hasil pencitraan) untuk meningkatkan cakupan dan akurasi prediktif.
- Melakukan pengujian kegunaan formal dengan profesional kesehatan untuk menilai kegunaan antarmuka dan relevansi klinis.
- Memperluas aplikasi R Shiny menjadi modul pendukung keputusan klinis yang berfungsi penuh yang terintegrasi dengan sistem rekam medis elektronik (EHR).

Daftar Pustaka

- Ansari, F., Sharma, S., & Garg, A. (2023). Performance evaluation of machine learning techniques for heart disease prediction. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 13(1), e1441.
- Bergstra, J., & Bengio, Y. (2012). Random search for hyper-parameter optimization. *Journal of Machine Learning Research*, 13(2)
- Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319–340.
- Erni, A. P., & Sa'adah, U. (2021). Comparison of Decision Tree, Naïve Bayes, and Random Forest Algorithm in Detecting Heart Disease. *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, 5(1), 101–108.
- ISO. (2010). *Ergonomics of human-system interaction — Part 210: Human-centred design for interactive systems* (ISO 9241-210:2010). International Organization for Standardization.
- Kohavi, R. (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection. *IJCAI*, 2, 1137–1145.

- Kuhn, M., & Johnson, K. (2013). *Applied predictive modeling*. Springer.
- Kushniruk, A. W., & Patel, V. L. (2004). Cognitive and usability engineering methods for evaluation of clinical information systems. *Journal of Biomedical Informatics*, 37(1), 56–76.
- Lantz, B. (2019). *Machine Learning with R: Expert techniques for predictive modeling*. Packt Publishing Ltd.
- Ong, W. M., Lee, J., & Tan, T. S. (2020). User-centered design of clinical prediction tools: usability evaluation of a prototype for type 2 diabetes screening. *BMC Medical Informatics and Decision Making*, 20(1), 1–10.
- Powers, D. M. (2011). Evaluation: From precision, recall and F-measure to ROC, informedness, markedness and correlation. *Journal of Machine Learning Technologies*, 2(1), 37–63.
- Prasanna Sai Teja, P., & Veeramani, T. (2022). Comparing the Efficiency of Heart Disease Prediction Using Machine Learning Techniques. *Cardiometry*, (23), 494–500.
- Robin, X., Turck, N., Hainard, A., et al. (2011). pROC: An open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics*, 12, 77.
<https://doi.org/10.1186/1471-2105-12-77>
- World Health Organization. (2021). *Cardiovascular diseases (CVDs)*. Retrieved from [https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-\(cvds\)](https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds))
- Yang, L., & Shami, A. (2020). On hyperparameter optimization of machine learning algorithms: Theory and practice. *Neurocomputing*, 415, 295–316.
- Zhou, T., Jin, X., & Yang, Z. (2020). Heart disease prediction model based on decision tree and grid search. *Journal of Intelligent & Fuzzy Systems*, 38(3), 3229–3239.
- Zriqat, I. A., Al-Dubai, M. M., & Al-Sharabi, T. H. (2017). A comparative study for predicting heart diseases using data mining classification methods. *arXiv preprint arXiv:1706.09969*.