e- ISSN: 2828–299X

Analisis Perbandingan Model Machine Learning untuk Prediksi Penyakit Jantung

Juli Sulaksono[1], Danar Putra Pamungkas[2], Danang Wahyu Widodo[3]
[1,2]Teknik Informatika, Fakultas Teknik dan Ilmu Komputer, Universitas Nusantara PGRI Kediri
[3]Teknik Mesin, Fakultas Teknik, Universitas Nusantara PGRI Kediri
E-mail: *[1] *jsulaksono@unpkediri.ac.id,* [2]danar@unpkediri.ac.id,
[3]*danangwahyuwidodo@unpkediri.ac.id*

Abstrak – *Penyakit jantung merupakan salah satu penyebab utama kematian di dunia yang memerlukan deteksi dini secara akurat. Penelitian ini bertujuan untuk menganalisis dan memprediksi risiko penyakit jantung menggunakan dataset kesehatan publik melalui pendekatan machine learning. Metode yang digunakan melibatkan perbandingan dua algoritma klasifikasi, yaitu Random Forest dan Logistic Regression. Data diproses melalui tahapan pembersihan, pembagian dataset (80:20), serta normalisasi khusus untuk regresi logistik. Hasil evaluasi menunjukkan bahwa algoritma Random Forest memberikan performa sempurna dengan akurasi 100%, sementara Logistic Regression menghasilkan akurasi sebesar 87,80%. Penelitian ini menyimpulkan bahwa Random Forest lebih unggul dalam menangkap pola medis yang kompleks, sementara Logistic Regression unggul dalam aspek interpretabilitas koefisien fitur*.

Kata Kunci *— Heart disease, logistic regression, machine learning, prediction model, random forest*

1. PENDAHULUAN

Penyakit jantung (*heart disease*) atau kardiovaskular merupakan kondisi medis kritis akibat penyempitan atau penyumbatan pembuluh darah yang dapat memicu serangan jantung fatal. Penyakit ini tetap menjadi penyebab kematian tertinggi di dunia dan Indonesia tanpa memandang latar belakang usia maupun gaya hidup [1]. Mengingat kompleksitas faktor risikonya, seperti hipertensi, kadar kolesterol, dan detak jantung maksimum, diperlukan sistem deteksi dini yang akurat untuk mendukung tindakan preventif medis.

Penelitian mengenai prediksi penyakit jantung menggunakan *machine learning* telah berkembang pesat dalam lima tahun terakhir. Anjeli & Rizki (2025) melakukan perbandingan algoritma dan menemukan bahwa *AdaBoost* memberikan akurasi 86,96% dalam deteksi dini [2]. Sementara itu, Jasman dkk. (2022**)** membandingkan *Logistic Regression*, *Random Forest*, dan *Perceptron*, di mana *Random Forest* dengan teknik SMOTE mencapai akurasi tertinggi sebesar 94% [3]. Nasution dkk. (2025) mengevaluasi dataset UCI dan menunjukkan bahwa tanpa seleksi fitur, *Random Forest* tetap konsisten dengan akurasi 89,7%, mengungguli *Logistic Regression* (84,2%) [4].

Penelitian lain oleh Putri dkk. (2025) menyoroti pentingnya *hyperparameter tuning* untuk meningkatkan efektivitas model klasifikasi penyakit jantung [5]. Di sisi lain, Yuliasari & Rahmatulloh (2025) mengonfirmasi bahwa *Random Forest* unggul dengan akurasi 90,16% dibandingkan algoritma lain seperti KNN [6]. Terkait penggunaan teknik penyeimbangan data, penelitian Alexander (2025**)** membuktikan bahwa optimasi melalui *Feature Importance* pada *Random Forest* meningkatkan efisiensi model secara signifikan [7]. Wahidin dkk. (2024) juga melakukan komparasi enam metode dan menemukan bahwa *Naive Bayes* dan *Logistic Regression* memberikan performa kompetitif di kisaran 84% [8].

Selanjutnya, Simanjuntak dkk. (2025) membandingkan model tradisional dengan *Deep Learning*, mencatat bahwa meski *Deep Neural Network* mencapai akurasi 92,1%, model *machine learning* tradisional seperti SVM dan *Random Forest* tetap sangat kompetitif untuk data tabular medis [9]. Penelitian Putri (2021) di Indonesia menggunakan KNN menunjukkan akurasi yang optimal melalui *cross-validation* pada data kardiovaskular [10]. Terakhir, studi oleh Farida & Bahri (2024) menekankan pentingnya tahap *preprocessing* dan deteksi *outlier* untuk meningkatkan akurasi klasifikasi penyakit jantung pada populasi di Asia [11].

Meskipun banyak algoritma tersedia, perbandingan antara *Random Forest* yang bersifat non-linear dengan *Logistic Regression* yang bersifat linear tetap menjadi topik yang relevan untuk dikaji guna memahami sejauh mana kompleksitas model mempengaruhi akurasi pada data klinis. Oleh karena itu, penelitian ini bertujuan untuk melakukan analisis perbandingan performa antara model *Random Forest* dan *Logistic Regression* dalam memprediksi risiko penyakit jantung. Fokus utama penelitian ini adalah mengevaluasi metrik akurasi, presisi, *recall*, dan *f1-score* guna menentukan model yang paling andal dalam mengidentifikasi pola risiko kesehatan pasien secara tepat.

e- ISSN: 2828–299X

2. METODE PENELITIAN

Metodologi yang digunakan dalam penelitian ini mengikuti alur kerja standar eksperimen *machine learning*, mulai dari pengumpulan data hingga evaluasi model.

2.1 Alur Penelitian

Tahapan penelitian ini dimulai dari identifikasi masalah, pengumpulan dataset, pra-pemrosesan data (*preprocessing*), pelatihan model menggunakan dua algoritma berbeda, dan diakhiri dengan evaluasi performa model menggunakan *confusion matrix*.

2.2 Dataset

Dataset yang digunakan dalam penelitian ini adalah *Heart Disease Dataset* yang diperoleh dari sumber publik. Dataset terdiri dari 1026 sampel dengan 14 atribut, di mana 13 atribut merupakan fitur klinis (seperti *age, sex, cp, trestbps, chol, fbs, restecg, thalach, exang, oldpeak, slope, ca, thal*) dan 1 atribut sebagai label target (0 untuk tidak berisiko, 1 untuk berisiko).

2.3 Preprocessing

Tahapan *preprocessing* dilakukan untuk memastikan kualitas data sebelum masuk ke tahap pelatihan:

1. Pembersihan Data: Memeriksa adanya *missing values* untuk menjaga konsistensi data.
2. Pemisahan Fitur dan Target: Memisahkan variabel independen (X) dari variabel dependen (y).
3. Normalisasi Fitur: Khusus untuk algoritma *Logistic Regression*, dilakukan standarisasi menggunakan StandardScaler agar seluruh fitur memiliki skala yang sama, sehingga mempercepat konvergensi model.
4. Pembagian Data: Dataset dibagi menjadi dua bagian, yaitu *training set* sebesar 80% (820 data) untuk melatih model dan *testing set* sebesar 20% (206 data) untuk menguji performa prediksi

2.4 Algoritma Klasifikasi

Penelitian ini membandingkan dua pendekatan algoritma yang berbeda secara fundamental: satu berbasis *ensemble* pohon keputusan dan satu berbasis fungsi probabilitas linear.

Random Forest

Random Forest adalah algoritma *ensemble learning* yang bekerja dengan membangun sekumpulan pohon keputusan (*decision trees*) selama fase pelatihan. Algoritma ini menggunakan teknik *Bootstrap Aggregating* (Bagging) dan seleksi fitur acak untuk mengurangi varians dan mencegah *overfitting*. Tahapan algoritmanya adalah sebagai berikut:

1. Bootstrapping: Mengambil sampel data secara acak dengan pengembalian dari dataset pelatihan.
2. Membangun Tree: Untuk setiap sampel, dibangun sebuah *decision tree*. Pada setiap *node*, pemisahan (*split*) dipilih dari subset fitur yang diambil secara acak.
3. Kriteria Pemisahan: Penentuan split terbaik didasarkan pada Gini Impurity atau Information Gain (Entropy). Persamaan Gini Impurity (G) adalah:

$$G = 1 - \sum_{i=1}^{e}(p_i) \quad ^2 \quad \text{.....(1)}$$

Di mana P_i adalah probabilitas kelas i pada node tersebut.

4. Voting (Agregasi): Setelah semua pohon terbentuk, prediksi akhir ditentukan berdasarkan suara terbanyak (*majority voting*) dari seluruh pohon.

Logistic Regression

Logistic Regression digunakan untuk memprediksi probabilitas variabel target biner. Meskipun namanya regresi, algoritma ini digunakan untuk klasifikasi dengan memetakan input linear ke dalam fungsi logistik (Sigmoid). Tahapannya meliputi:

1. Linear Combination: Menghitung kombinasi linear dari fitur input (x) dan bobot (w):

$$z = w_0 + w_1x_1 + w_2x_2 + ... + w_nx_n \text{ .....(2)}$$

2. Fungsi Sigmoid: Memetakan nilai z ke dalam rentang probabilitas 0 hingga 1 menggunakan fungsi:

$$\sigma(z) = \frac{1}{1+e^{-z}} \text{ .....(3)}$$

3. **Thresholding**: Jika $\sigma(z) \geq 0.5$, maka diklasifikasikan sebagai kelas 1 (berisiko), jika kurang maka kelas 0.

4. **Optimasi**: Model diperbarui menggunakan *Gradient Descent* untuk meminimalkan fungsi *Log Loss* atau *Cross-Entropy Error*.

Metrik Evaluasi

Akurasi (Accuracy)

Akurasi merupakan ukuran dasar yang dipakai untuk menilai kinerja model kalsifikasi. Nilai akurasi dihitung dengan membandingkan jumlah klasifikasi yang tepat dengan total jumlah kasus yang dianalisis (*Evaluation Metrics in Machine Learning - GeeksforGeeks*, n.d.). Berikut adalah rumus untuk menghitung akurasi :

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \text{.....(4)}$$

Di mana :
TP = True Positive, TN = True Negative, FP = False Positive, FN = False Negative.

Presisi (Precision)

Presisi adalah salah satu indikator dalam menilai model klasifikasi yang menghitung seberapa akurat prediksi untuk kategori kelas positif. Presisi menggambarkan seberapa banyak dari total data yang diidentifikasi sebagai kategori positif benar-benar merupakan bagian dari kategori positif (*Evaluation Metrics in Machine Learning - GeeksforGeeks*, n.d.). Berikut adalah rumus untuk menghitung Presisi :

$$Precision = \frac{TP}{TP+FP} \text{.....(5)}$$

Di mana :
TP = True Positive, FP = False Positive.

Recall (Sensitivity)

Sensitivity, atau yang sering disebut sebagai recall, adalah salah satu indikator penilaian dalam model klasifikasi yang menilai seberapa baik model dapat mendeteksi kasus positif dengan akurat. Nilai sensitivity yang mencapai 100% menunjukkan bahwa semua data yang benar-benar berada dalam kategori positif telah berhasil diidentifikasi dengan tepat oleh model, yang berarti tidak ada kasus positif yang terlewat atau salah diklasifikasikan sebagai negatif (*Evaluation Metrics in Machine Learning - GeeksforGeeks*, n.d.). Berikut adalah rumus untuk menghitung recall :

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

Di mana :
TP = True Positive, FN = False Negative.

F1-Score

F1-Score merupakan suatu metrik evaluasi komposis yang berperan sebagai satu metrik untuk menilai performa model klasifikasi dengan lebih mendalam. Perhitungan metrik ini dilakukan berdasarkan rata-rata harmonik antara nilai presisi dan recall. Ketika F1-Score mencapai angka 100%, itu menandakan situasi ideal di mana model telah berhasil mencapai keseimbangan sempurna antara presisi dan recall. Ini menunjukkan bahwa model tersebut tidak hanya sangat tepat dalam meremalkan instance positif (presisi yang tinggi), tetapi juga sangat efisien dalam mengidentifikasi seluruh instance positif yang ada (recall yang tinggi) (*Evaluation Metrics in Machine Learning - GeeksforGeeks*, n.d.). Berikut adalah rumus untuk menghitung F1-score :

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

e- ISSN: 2828–299X

3. HASIL DAN PEMBAHASAN

Bagian ini menyajikan hasil evaluasi kinerja model Random Forest dan Logistic Regression secara kuantitatif dan kualitatif. Penyajian hasil tidak hanya bersifat deskriptif, tetapi juga didukung oleh data evaluasi yang dapat dipertanggungjawabkan berupa metrik klasifikasi dan confusion matrix

3.1 Hasil Evaluasi Kinerja Model

Evaluasi dilakukan menggunakan data uji sebanyak 205 data (20% dari total dataset). Metrik evaluasi yang digunakan meliputi akurasi, precision, recall, dan F1-score. Ringkasan hasil evaluasi kedua model disajikan pada Tabel 1

Tabel 1. Perbandingan Kinerja Model Random Forest dan Logistic Regression

Model	Akurasi	Precision	Recall	F1-Score
Random Forest	1.00	1.00	1.00	1.00
Logistic Regression	0.878	0.878	0.878	0.878

Tabel 1 menyajikan perbandingan kinerja kedua model berdasarkan empat metrik evaluasi utama, yaitu akurasi, precision, recall, dan F1-score. Nilai akurasi menunjukkan proporsi prediksi yang benar terhadap seluruh data uji. Berdasarkan tabel tersebut, model Random Forest mencapai nilai akurasi sebesar 1,00, yang berarti seluruh data uji berhasil diklasifikasikan dengan benar. Sementara itu, model Logistic Regression memperoleh akurasi sebesar 0,878, yang menunjukkan bahwa sekitar 87,8% data uji dapat diprediksi secara tepat.

Nilai precision menggambarkan tingkat ketepatan model dalam memprediksi kelas positif, sedangkan recall menunjukkan kemampuan model dalam mengenali seluruh data yang benar-benar termasuk dalam kelas positif. Pada Tabel 1 terlihat bahwa Random Forest menghasilkan nilai precision dan recall yang sempurna, sedangkan Logistic Regression memiliki nilai precision sebesar 0,88 dan recall sebesar 0,87. Perbedaan ini mengindikasikan bahwa Random Forest lebih konsisten dalam meminimalkan kesalahan prediksi pada kedua kelas.

F1-score merupakan rata-rata harmonis antara precision dan recall. Nilai F1-score pada Random Forest sebesar 1,00 menunjukkan keseimbangan yang sangat baik antara ketepatan dan kelengkapan prediksi. Adapun Logistic Regression memiliki nilai F1-score sebesar 0,88, yang masih tergolong baik namun menunjukkan adanya penurunan performa dibandingkan Random Forest.

3.2 Classification Report dan Confusion Matrix Random Forest

Hasil classification report untuk model Random Forest menunjukkan bahwa model mampu mengklasifikasikan seluruh data uji dengan sangat baik. Nilai precision, recall, dan F1-score pada kelas 0 (tidak memiliki penyakit jantung) dan kelas 1 (memiliki penyakit jantung) masing-masing bernilai 1,00.

Tabel 2. Classification Report Model Random Forest

Kelas	Precision	Recall	F1-Score	Support
0	1.00	1.00	1.00	98
1	1.00	1.00	1.00	107
		Akurasi	1.00	205

Tabel 2 menampilkan classification report dari model Random Forest secara rinci untuk masing-masing kelas. Kolom *support* menunjukkan jumlah data uji pada setiap kelas, yaitu 98 data untuk kelas 0 (tidak memiliki penyakit jantung) dan 107 data untuk kelas 1 (memiliki penyakit jantung). Kondisi ini menunjukkan bahwa distribusi data uji relatif seimbang, sehingga evaluasi kinerja model dapat dilakukan secara objektif.

Nilai precision, recall, dan F1-score pada kedua kelas masing-masing bernilai 1,00. Hal ini menunjukkan bahwa model Random Forest mampu memprediksi seluruh data pada setiap kelas dengan benar tanpa menghasilkan kesalahan klasifikasi. Dengan kata lain, tidak terdapat data kelas 0 yang salah diprediksi sebagai kelas 1 maupun sebaliknya.

Nilai akurasi keseluruhan sebesar 1,00 yang ditunjukkan pada baris terakhir tabel mengonfirmasi bahwa seluruh data uji berhasil diklasifikasikan dengan tepat. Hasil ini mengindikasikan bahwa Random Forest memiliki kemampuan yang sangat kuat dalam mengenali pola pada dataset yang digunakan, meskipun diperlukan pengujian lanjutan untuk memastikan kemampuan generalisasi model pada data yang berbeda.

e- ISSN: 2828–299X

3.3 Classification Report dan Confusion Matrix Logistic Regression
Model Logistic Regression menghasilkan performa yang cukup baik, meskipun masih terdapat beberapa kesalahan klasifikasi. Hasil classification report Logistic Regression disajikan pada Tabel 3

Tabel 3. Classification Report Model Logistic Regression

Kelas	Precision	Recall	F1-Score	Support
0	0.88	0.83	0.86	90
1	0.88	0.91	0.89	115
		Akurasi	0.878	205

Tabel 3 menyajikan classification report dari model Logistic Regression untuk masing-masing kelas. Pada kelas 0, model menghasilkan nilai precision sebesar 0,88 dan recall sebesar 0,83. Nilai tersebut menunjukkan bahwa sebagian besar data yang tidak memiliki penyakit jantung dapat diprediksi dengan benar, meskipun masih terdapat beberapa data yang salah diklasifikasikan sebagai kelas berisiko. Pada kelas 1, Logistic Regression menghasilkan nilai recall yang lebih tinggi, yaitu sebesar 0,91, dengan precision sebesar 0,88. Nilai recall yang tinggi ini menunjukkan bahwa model cukup efektif dalam mengenali pasien yang benar-benar memiliki risiko penyakit jantung, walaupun masih terdapat sejumlah kecil kesalahan prediksi.

Nilai F1-score pada kedua kelas berada pada kisaran 0,86 hingga 0,89, yang menandakan keseimbangan yang cukup baik antara precision dan recall. Akurasi keseluruhan sebesar 0,878 menunjukkan bahwa Logistic Regression mampu mengklasifikasikan sebagian besar data uji dengan tepat, namun performanya masih berada di bawah Random Forest. Secara kuantitatif, Random Forest unggul pada seluruh metrik evaluasi dibandingkan Logistic Regression. Keunggulan ini menunjukkan bahwa Random Forest lebih mampu menangkap hubungan nonlinier dan interaksi kompleks antar fitur medis. Namun, performa yang sangat tinggi pada Random Forest juga perlu dianalisis secara kritis, karena kemungkinan dipengaruhi oleh karakteristik dataset dan potensi overfitting pada skenario pengujian tertentu.

Logistic Regression, meskipun memiliki performa yang lebih rendah, tetap menunjukkan kestabilan yang baik dan keunggulan dari sisi interpretabilitas. Koefisien fitur pada Logistic Regression dapat digunakan untuk menjelaskan kontribusi masing-masing variabel terhadap risiko penyakit jantung, yang menjadi aspek penting dalam konteks klinis. Dengan demikian, hasil penelitian ini menunjukkan bahwa Random Forest lebih unggul dari sisi akurasi prediksi, sedangkan Logistic Regression lebih unggul dari sisi transparansi dan kemudahan interpretasi. Penyajian hasil secara kuantitatif ini memperkuat kesimpulan bahwa pemilihan model harus disesuaikan dengan kebutuhan aplikasi, khususnya dalam sistem pendukung keputusan medis.

4. SIMPULAN

Berdasarkan hasil analisis perbandingan yang telah dilakukan, dapat disimpulkan bahwa penerapan *machine learning* untuk memprediksi risiko penyakit jantung memberikan hasil yang sangat signifikan. Dataset *Heart Disease* yang digunakan memiliki karakteristik yang sesuai untuk klasifikasi biner, dengan kombinasi fitur medis yang mampu merepresentasikan kondisi kesehatan pasien secara menyeluruh.

Hasil evaluasi menunjukkan bahwa algoritma **Random Forest** merupakan model terbaik dalam penelitian ini dengan tingkat akurasi mencapai **100%**, serta nilai *precision, recall,* dan *F1-score* yang sempurna (1,00). Hal ini membuktikan bahwa *Random Forest* sangat handal dalam menangkap pola hubungan non-linear antar fitur klinis. Di sisi lain, algoritma **Logistic Regression** juga menunjukkan performa yang stabil dengan akurasi sebesar **87,80%**. Meskipun lebih rendah dari *Random Forest*, model ini memiliki keunggulan dari sisi interpretabilitas koefisien fitur yang lebih mudah dipahami oleh tenaga medis. Secara keseluruhan, penggunaan model klasifikasi ini sangat layak dipertimbangkan sebagai alat bantu pendukung keputusan (*decision support system*) di bidang kesehatan

5. SARAN

Berdasarkan hasil penelitian dan evaluasi yang telah dilakukan, terdapat beberapa saran yang dapat diajukan untuk pengembangan penelitian selanjutnya:

1. Penambahan Dataset: Menggunakan dataset yang lebih bervariasi dan mencakup populasi yang lebih luas (multi-etnis) untuk memastikan generalisasi model tetap terjaga pada data medis yang berbeda secara demografis.
2. Optimasi Model: Melakukan eksperimen dengan teknik *Hyperparameter Tuning* seperti *Grid Search* atau *Random Search* pada algoritma *Random Forest* untuk memastikan model tetap efisien tanpa mengalami *overfitting* yang berlebihan.
3. Eksplorasi Algoritma: Menguji algoritma lain yang lebih kompleks seperti *XGBoost*, *LightGBM*, atau *Deep Learning* (Neural Networks) sebagai pembanding tambahan guna melihat konsistensi performa pada data yang lebih besar.

DAFTAR PUSTAKA

[1] WHO, "Cardiovascular diseases (CVDs) Fact Sheets," 2024.

[2] Anjeli, L., & Rizki, F. 2025. Analisis Prediksi Penyakit Jantung Menggunakan Perbandingan Algoritma Machine Learning. *JTINFO: Jurnal Teknik Informatika*, 4(2), 262-269.

[3] Jasman, T. Z., dkk. 2022. Perbandingan Logistic Regression, Random Forest, dan Perceptron pada Klasifikasi Pasien Gagal Jantung. *CSRID Journal*, 14(3), 271-286.

[4] Nasution, N., dkk. 2025. An Evaluation of Logistic Regression, Random Forest, SVM, and KNN Models on the UCI Heart Disease Dataset. *IT Journal Research and Development*.

[5] Putri, S. K. N. A., dkk. 2025. Penerapan Hyperparameter Tuning pada Model Klasifikasi untuk Prediksi Risiko Penyakit Jantung. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, 5(4).

[6] Yuliasari, S., & Rahmatulloh, A. 2025. Performance Analysis and Accuracy of Machine Learning Algorithms for Heart Disease Prediction. *Telematika*, 22(3).

[7] Alexander, M. 2025. Penerapan Algoritma Random Forest untuk Deteksi Penyakit Jantung dan Optimisasi Menggunakan Feature Importance. *Bachelor Thesis, Universitas Multimedia Nusantara*.

[8] Wahidin, A. J., dkk. 2024. Machine Learning Untuk Klasifikasi Penyakit Jantung. *Aisyah Journal of Informatics and Electrical Engineering*, 6(1).

[9] Simanjuntak, M. S., dkk. 2025. A Comparative Study of Machine Learning and Deep Learning Models for Heart Disease Classification. *Journal of Applied Informatics and Computing*, 9(6).

[10] Putri, I. P. 2021. Analisis Performa Metode K-Nearest Neighbor (KNN) dan Crossvalidation pada Data Penyakit Cardiovascular. *Indonesian Journal of Data Science*, 2(1), 21-28.

[11] Farida, L. N., & Bahri, S. 2024. Klasifikasi Gagal Jantung Menggunakan Metode SVM (Support Vector Machine). *Komputika: Jurnal Sistem Komputer*, 13(2).