

# Analisis Perbandingan Kinerja Algoritma Klasifikasi dalam Mendeteksi Penyakit Jantung

Abram Setyo Prabowo<sup>1</sup>, Felix Indra Kurniadi<sup>1</sup>

<sup>1</sup>Computer Science Department, School of Computer Science, Bina Nusantara University  
[abram.setyo@binus.ac.id](mailto:abram.setyo@binus.ac.id), [felix.indra@binus.ac.id](mailto:felix.indra@binus.ac.id)

Diterima : 01 September 2023

Disetujui : 28 September 2023

**Abstract**— Deteksi penyakit jantung secara dini dan akurat memiliki dampak signifikan terhadap prognosis pasien serta mengurangi beban penyakit secara keseluruhan. Dalam upaya meningkatkan efektivitas deteksi penyakit jantung, teknik pembelajaran mesin dan algoritma klasifikasi telah muncul sebagai alat yang berpotensi ampuh dalam mendiagnosis kondisi ini dengan tingkat akurasi yang tinggi. Penelitian ini bertujuan untuk memprediksi penyakit jantung dengan menggunakan perbandingan Support Vector Machine (SVM), Random Forest, Logistic Regression, dan AdaBoost. Pada penelitian ini algoritma Random Forest mempunyai model base score untuk training test dengan nilai 1, nilai tersebut merupakan nilai terbaik dibandingkan dengan 3 algoritma yang diusulkan pada penelitian ini. Selama pengujian, hasil yang diperoleh adalah algoritma random forest, SVM, dan AdaBoost merupakan algoritma yang mempunyai nilai terbaik dan nilai yang sama pada hasil pengujian. Untuk nilai akurasi 0.985366, presisi 0.985714, recall 0.985437, dan f1-score 0.985364..

**Keywords** — *Heart Disease, Machine Learning, SVM, AdaBoost, Random Forest, Linear Regression*

## I. PENDAHULUAN

Jantung merupakan salah satu organ vital pada tubuh manusia dan merupakan pusat sistem peredaran darah yang melibatkan pembuluh darah sebagai salurannya. Jantung berperan penting dalam menyediakan oksigen ke seluruh tubuh dan bertugas membersihkan tubuh dari produk metabolisme. Namun, seiring dengan perannya yang penting, jantung juga menjadi salah satu organ tubuh yang paling rentan terserang penyakit.

Menurut World Health Organization (WHO), penyakit jantung merujuk pada sekelompok kondisi di mana terjadi kelainan dalam fungsi jantung atau pembuluh darah. Rentang kondisi ini meliputi berbagai permasalahan seperti penyakit arteri koroner, gagal jantung, gangguan irama jantung, serta kelainan pada katup jantung.

Selama tahun 2022, Indonesia menghadapi tingginya kasus penyakit-penyakit serius, di

antaranya adalah penyakit jantung dengan jumlah kasus mencapai 15,5 juta, kanker sebanyak 3,2 juta kasus, stroke sekitar 2,5 juta kasus, dan gagal ginjal dengan jumlah kasus mencapai 1,3 juta. Pada tahun yang sama, biaya pengobatan untuk penyakit-penyakit serius tersebut menghabiskan sekitar 21% dari total biaya yang ditanggung oleh BPJS dalam program jaminan kesehatan [1], namun penyakit ini sebagian besar disebabkan oleh kebiasaan gaya hidup yang buruk, seperti pola makan yang tidak sehat, kurang berolahraga secara teratur, merokok, mengonsumsi alkohol atau obat-obatan, dan stres yang tinggi. Dalam sebuah penelitian menjelaskan bahwa penyakit jantung dapat dikenali dengan menggunakan pendekatan machine learning [2], penyakit jantung dapat dikenali dari kondisi tubuh pasien. berdasarkan jenis kelamin, tekanan darah, usia, perokok atau tidak serta beberapa indikator yang menjadi ciri penyakit jantung seseorang.

Tingginya angka kematian akibat penyakit jantung mendorong banyak penelitian untuk menyelesaikan masalah ini, berbagai penelitian berbasis komputer dengan menggunakan pendekatan machine learning dilakukan untuk meningkatkan tingkat akurasi dan efektivitas dalam memprediksi penyakit jantung [3]. Pendekatan menggunakan machine learning di bidang kesehatan memberikan kemudahan dalam mendiagnosa penyakit jantung tanpa harus membutuhkan waktu yang lama. Selain itu, machine learning memiliki kemampuan untuk belajar secara mandiri tanpa perlu pengawasan konstan.

Deteksi dini dan akurat penyakit jantung memiliki dampak yang signifikan pada prognosis pasien serta pengurangan beban penyakit secara keseluruhan. Dalam upaya meningkatkan efektivitas deteksi penyakit jantung, teknik machine learning dan algoritma klasifikasi telah muncul sebagai alat yang berpotensi kuat dalam mendiagnosis kondisi ini dengan tingkat akurasi yang tinggi. Algoritma-algoritma ini tergolong dalam kategori reinforcement learning, unsupervised, dan supervised, masing-masing terdiri dari beragam jenis algoritma. Penelitian sebelumnya membandingkan algoritma dalam melakukan data testin untuk mendeteksi penyakit jantung menggunakan C5.0 dan logistic regression [4], hasilnya menunjukkan C5.0 lebih unggul dalam tingkat akurasi yaitu 93,02% dibandingkan dengan logistic regression yang hasilnya 85,25%. Penelitian sebelumnya menggunakan metode random forest untuk melakukan deteksi pasien gagal jantung dan menghasilkan tingkat akurasi hingga 91,45% [5]. Penelitian selanjutnya dengan menggunakan data repositori University of California Irvine untuk digunakan untuk mendeteksi penyakit kardiovaskular [6], menunjukkan bahwa model MLP mencapai tingkat akurasi sebesar 82,47% dibandingkan dengan model K-NN yang menghasilkan 73,77%. Meskipun banyak penelitian yang telah mengusulkan berbagai jenis algoritma, namun diperlukan kebutuhan untuk melakukan perbandingan kinerja secara menyeluruh di antara algoritma klasifikasi. Pertimbangan faktor seperti akurasi, presisi,

recall, dan F1-score perlu dianalisis secara mendalam untuk memahami kelebihan dan kekurangan yang dimiliki oleh masing-masing algoritma. Hal ini penting untuk mengidentifikasi metode machine learning yang paling efektif dalam memprediksi penyakit jantung,

## II. METODE PENELITIAN

### A. Dataset

Dataset yang digunakan pada penelitian ini bersumber dari *Kaggle* [7], terdiri dari 1025 data yang memiliki 14 atribut yang dapat dilihat Tabel.1.

Tabel 1. Deskripsi Atribut Data

Variabel	Deskripsi
Age	Umur
Sex	Jenis kelamin
cp	Nyeri Dada / Chest Pain
trtbps	Hasil tekanan darah selama dirawat di rumah sakit
chol	Kolesterol
fbs	Gula darah puasa
restecg	Hasil elektrokardiografi saat istirahat
thalachh	Detak jantung maksimum
exng	Latihan induksi angina
oldpeak	ST depression
slp	ST segment slope
caa	Jumlah pembuluh darah besar
thall	Hasil tegangan thallium
target/output	1= pasien sakit jantung, 0=pasien sehat

### B. Distribusi Data

Pada tahap ini, dataset dipisahkan menjadi 2 bagian, yaitu data latih (*training data*) dan data uji (*testing data*) dengan perbandingan 80:20. Tidak ada ketentuan yang mengatur jumlah data yang harus digunakan, namun hal penting yang perlu diperhatikan adalah data latih harus lebih besar dibandingkan data uji. Terdapat banyak strategi dalam pembagian data latih dan uji yang dapat digunakan dalam pembuatan model machine learning [8].

### C. Algoritma Klasifikasi

Tahapan ini adalah menentukan algoritma-algoritma klasifikasi yang akan digunakan dan dievaluasi kinerjanya. Model algoritma yang digunakan pada penelitian ini adalah sebagai berikut :

- *Support Vector Machine* (SVM), merupakan suatu algoritma dalam bidang machine learning yang digunakan untuk dua tujuan utama, yakni klasifikasi dan regresi [9]. Dalam konteks klasifikasi, SVM bertujuan untuk menemukan hiperplane terbaik, yang dapat berupa garis atau permukaan, untuk memisahkan dua kelas data dengan margin maksimum. Proses ini dilakukan dengan memilih hiperplane sedemikian rupa sehingga jarak antara hiperplane dan contoh data dari setiap kelas (yang disebut vektor pendukung atau support vectors) menjadi yang paling besar. Selain digunakan untuk klasifikasi, SVM juga dapat diterapkan dalam regresi untuk meramalkan nilai berkelanjutan. Pada kasus regresi, SVM membangun hiperplane dengan margin terlebar
- *Random Forest* merupakan sebuah algoritma di bidang machine learning yang memiliki aplikasi dalam klasifikasi, regresi, dan pemrosesan data [10]. Algoritma ini bekerja berdasarkan konsep ensemble, yang menggabungkan hasil prediksi dari beberapa model pembelajaran yang disebut pohon keputusan, dengan tujuan untuk meningkatkan akurasi dan performa. Setiap pohon dalam Random Forest dibangun dengan menggunakan subset acak dari data pelatihan serta fitur yang ada. Prediksi akhir dihasilkan dengan menggabungkan hasil prediksi dari seluruh pohon dalam ensemble tersebut.
- *Logistic Regression* adalah sebuah metode statistik dalam analisis regresi yang digunakan untuk memodelkan hubungan antara variabel independen (fitur) dengan variabel dependen (klasifikasi biner). Metode ini sering digunakan untuk tugas klasifikasi. *Logistic regression*

menggunakan fungsi logistik (sigmoid) untuk menghasilkan nilai probabilitas prediksi, yang kemudian digunakan untuk mengklasifikasikan data ke dalam kelas yang sesuai [11].

- *AdaBoost* adalah algoritma ensemble dalam machine learning yang digunakan untuk melakukan klasifikasi dan regresi. *Adaboost* sendiri berfungsi untuk meningkatkan kinerja model dengan menggabungkan beberapa model lemah (weak learners) menjadi model yang kuat. Algoritma ini memberikan lebih banyak perhatian pada data yang sebelumnya salah diprediksi dan memungkinkan model berikutnya untuk fokus pada kasus yang lebih sulit. Dengan menggabungkan hasil prediksi dari model-model lemah dengan bobot yang sesuai, Adaboost menghasilkan prediksi akhir yang lebih akurat [12].

## III. HASIL DAN DISKUSI

### A. Model Training

Model training yang dilakukan pada tahapan ini adalah menggunakan model klasifikasi *SVM*, *Random Forest*, *Logistic Regression*, dan *AdaBoost*. Proses pelatihan ini menggunakan prinsip *k-fold cross-validation*. *Cross-validation* memiliki peran penting dalam menetapkan model klasifikasi terbaik. Teknik ini membantu mengatasi *overfitting*, memberikan estimasi kinerja yang baik, dan mampu memabndingkan model klasifikasi dengan lebih obyektif [13].

Teknik ini melibatkan pembagian dataset ke dalam beberapa bagian, atau "lipatan", di mana model dilatih dan diuji pada berbagai kombinasi lipatan. Dari skor evaluasi yang dihasilkan, kita dapat memilih model terbaik yang menunjukkan kinerja yang paling kuat secara keseluruhan. Penting untuk tidak hanya memperhatikan performa pada satu lipatan tertentu, tetapi juga mengamati kinerja yang konsisten di semua lipatan.

Tabel 2 menunjukkan hasil akurasi dari model klasifikasi yang digunakan berdasarkan uji *cross-validation*, dimana *random forest* menunjukkan tingkat akurasi yang tinggi yaitu 0,980488 sedangkan model klasifikasi yang

memiliki tingkat akurasi rendah adalah *logistic regression* dengan nilai 0,862195.

Tabel 2. Hasil Uji Cross-Validation.

Model Klasifikasi	Cross-Validation Accuracy
Logistic Regression	0.862195
Random Forest	<b>0.980488</b>
SVC	0.914634
AdaBoost	0.904878

Selain melakukan *cross-validation*, pada tahapan ini dilakukan juga uji *base score* model untuk data latih. Berdasarkan data pada Tabel 3, terlihat bahwa terlihat bahwa model *Random forest* menunjukkan *base score train* tertinggi dengan nilai 1 sedangkan *base score train* terendah adalah *logistic regression* dengan nilai 0.768293.

Tabel 3. Perbandingan Base Score Model Klasifikasi.

Model Klasifikasi	Base-Train Accuracy
Logistic Regression	0.768293
Random Forest	<b>1.000000</b>
SVC	0.878049
AdaBoost	0.868293

### B. Model Testing

Tahapan ini merujuk pada proses mengukur kinerja model yang telah dilatih pada data yang belum pernah dilihat sebelumnya, hal ini bertujuan untuk memahami sejauh mana model dapat memiliki kinerja yang baik. Ukuran kinerja yang digunakan untuk menguji model klasifikasi adalah presisi, recall, akurasi, dan *f1-score*.

Secara umum, performa model klasifikasi yang baik ditandai oleh presisi, *recall*, dan *f1-score* yang tinggi, serta akurasi yang tinggi. Namun, cara mengartikan metrik-metrik ini tergantung pada konteks masalah tertentu dan pertimbangan antara positif palsu dan negatif palsu. Sebagai contoh, dalam kasus diagnostik medis, kesalahan memprediksi negatif palsu (yaitu, mendiagnosis pasien yang sebenarnya

sakit sebagai sehat) mungkin memiliki dampak lebih serius daripada positif palsu (yaitu, mendiagnosis pasien sehat sebagai sakit). Oleh karena itu, tingkat *recall* yang tinggi (yaitu, mendeteksi sebanyak mungkin kasus positif dengan benar) mungkin lebih diutamakan daripada tingkat presisi yang tinggi (yaitu, menghindari membuat prediksi positif palsu yang banyak) [14].

### C. Hasil Uji

Uji hasil menggunakan matrik performansi untuk mendapatkan akurasi, *recall*, presisi, dan *f1-score*. Hasil dapat dilihat pada Tabel 4 – Tabel 7.

Tabel 4..Perbandingan Hasil Untuk Akurasi

Model Klasifikasi	Akurasi
Logistic Regression	0.809756
Random Forest	<b>0.985366</b>
SVC	<b>0.985366</b>
AdaBoost	<b>0.985366</b>

Tabel 4 menunjukkan hasil akurasi pada hasil uji data. *Random forest*, *SVC*, dan *AdaBoost* menghasilkan nilai yang sama **0.985366**.

Tabel 5. Perbandingan Hasil Untuk Presisi

Model Klasifikasi	Presisi
Logistic Regression	0.814897
Random Forest	<b>0.985714</b>
SVC	<b>0.985714</b>
AdaBoost	<b>0.985714</b>

Tabel 5 menunjukkan hasil presisi pada hasil uji data. *Random forest*, *SVC*, dan *AdaBoost* menghasilkan nilai yang sama **0.985714**.

Tabel 6..Perbandingan Hasil Untuk Recall

Model Klasifikasi	Recall
Logistic Regression	0.809442
Random Forest	<b>0.985437</b>
SVC	<b>0.985437</b>
AdaBoost	<b>0.985437</b>

Tabel 6 menunjukkan hasil *recall* pada hasil uji data. *Random forest*, *SVC*, dan *AdaBoost* menghasilkan nilai yang sama **0.985437**.

Tabel 7..Perbandingan Hasil Untuk *f1-score*

Model Klasifikasi	F1-score
Logistic Regression	0.808865
Random Forest	<b>0.985364</b>
SVC	<b>0.985364</b>
AdaBoost	<b>0.985364</b>

Tabel 7 menunjukkan hasil *f1-score* pada hasil uji data. *Random forest*, *SVC*, dan *AdaBoost* menghasilkan nilai yang sama **0.985364**.

Berdasarkan data yang tersaji dalam Tabel 4-7, terobservasi bahwa nilai precision, recall, accuracy, dan F1-score dari model Random Forest, Support Vector Classifier (SVC), dan AdaBoost tidak menunjukkan perbedaan yang signifikan. Beberapa inferensi dapat diekstraksi untuk menjelaskan fenomena ini.

Pertama, dimensi data yang digunakan dalam studi ini relatif kecil. Masalah ini sangat berdampak pada kapasitas model dalam mempelajari dan generalisasi data dengan efisien, sehingga menghasilkan metrik evaluasi yang serupa.

Kedua, karakteristik fitur dalam dataset tampaknya stabil dan konsisten. Tidak ada fluktuasi atau variabilitas ekstrem dalam fitur yang bisa mempengaruhi kinerja model secara drastis. Ini mengindikasikan bahwa fitur-fitur tersebut telah disaring dan diproses dengan cermat sebelum diintegrasikan ke dalam model, seperti yang telah dibahas pada Bab II.

Terakhir, kompleksitas pola dalam data mungkin tidak mengharuskan adopsi model yang sangat kompleks. Dengan kata lain, dataset kemungkinan besar mengandung pola yang dapat dengan mudah dipelajari dan diinterpretasikan oleh beragam model. Ini tercermin dari kinerja model yang komparatif, sebagaimana terungkap dalam metrik evaluasi yang diperoleh.

#### IV. SIMPULAN

Pada penelitian ini, algoritma *Random Forest* memiliki *base score* model untuk uji latih dengan nilai 1, nilai ini menjadi nilai terbaik dibandingkan 3 algoritma yang diajukan pada penelitian ini. Sedangkan saat melakukan uji, hasil yang diperoleh adalah algoritma *random forest*, *SVM*, dan *AdaBoost* menjadi algoritma yang memiliki nilai terbaik dan bernilai sama dalam hasil ujinya. Untuk nilai akurasi **0.985366**, presisi **0.985714**, *recall* **0.985437**, dan *f1-score* **0.985364**. Sehingga ketiga algoritma tersebut dapat dijadikan rekomendasi untuk membuat model prediksi penyakit jantung. Model ini dapat dikembangkan dengan menerapkan deep learning dan menggunakan dataset yang sama atau dapat menambahkan dataset pasien yang lebih banyak.

#### DAFTAR PUSTAKA

- [1] A. Ahdiat, "Kasus Penyakit Katastropik di Indonesia Meningkat pada 2022." <https://databoks.katadata.co.id/datapublish/2023/07/03/kasus-penyakit-katastropik-di-indonesia-meningkat-pada-2022> (accessed Aug. 26, 2023).
- [2] H. M. Nawawi, J. J. Purnama, and A. B. Hikmah, "Komparasi Algoritma Neural Network dan Naïve Bayes untuk Memprediksi Penyakit Jantung," *J. Pilar Nusa Mandiri*, vol. 15, no. 2, pp. 189–194, Sep. 2019, doi: 10.33480/PILAR.V15I2.669.
- [3] F. D. Telaumbanua, P. Hulu, T. Z. Nadeak, R. R. Lumbantong, and A. Dharma, "Penggunaan Machine Learning Di Bidang Kesehatan," *J. Teknol. DAN ILMU Komput. PRIMA*, vol. 2, no. 2 SE-, pp. 391–399, Jan. 2020, doi: 10.34012/jutikomp.v2i2.657.
- [4] J. J. Pangaribuan, H. Tanjaya, and Kenichi, "Mendeteksi Penyakit Jantung Menggunakan Mechine Learning Dengan Algoritma Logistic Regression," *Mach. Learn.*, vol. 45, no. 13, pp. 40–48, 2021.
- [5] Y. Yuliani, "Algoritma Random Forest Untuk Prediksi Kelangsungan Hidup Pasien Gagal Jantung Menggunakan Seleksi Fitur Bestfirst," *Infotek J. Inform. dan Teknol.*, vol. 5, no. 2, pp. 298–306,

- 2022, doi: 10.29408/jit.v5i2.5896.
- [6] M. Pal, S. Parija, G. Panda, K. Dhama, and R. K. Mohapatra, "Risk prediction of cardiovascular disease using machine learning classifiers," *Open Med.*, vol. 17, no. 1, pp. 1100–1113, 2022, doi: 10.1515/med-2022-0508.
- [7] D. Lapp, "Heart Disease Dataset | Kaggle," 2019. <https://www.kaggle.com/datasets/johnsmith88/heart-disease-dataset> (accessed Aug. 28, 2023).
- [8] F. Handayani, "Komparasi Support Vector Machine, Logistic Regression Dan Artificial Neural Network Dalam Prediksi Penyakit Jantung," *J. Edukasi dan Penelit. Inform.*, vol. 7, no. 3, p. 329, 2021, doi: 10.26418/jp.v7i3.48053.
- [9] A. Damayunita, R. S. Fuadi, and C. Juliane, "Comparative Analysis of Naive Bayes, K-Nearest Neighbors (KNN), and Support Vector Machine (SVM) Algorithms for Classification of Heart Disease Patients," *J. Online Inform.*, vol. 7, no. 2, pp. 219–225, 2022, doi: 10.15575/join.v7i2.919.
- [10] V. Y. Kullarni and P. K. Sinha, "Random Forest Classifier: A Survey and Future Research Directions," *Int. J. Adv. Comput.*, vol. 36, no. 1, pp. 1144–1156, 2013.
- [11] S. N. Ahmed, R. Jhaj, B. Sadasivam, and R. Joshi, "Reversal of hypertensive heart disease: a multiple linear regression model," *Discoveries*, vol. 9, no. 4, p. e138, 2021, doi: 10.15190/d.2021.17.
- [12] T. R. Mahesh *et al.*, "AdaBoost Ensemble Methods Using K-Fold Cross Validation for Survivability with the Early Detection of Heart Disease," *Comput. Intell. Neurosci.*, vol. 2022, p. 9005278, 2022, doi: 10.1155/2022/9005278.
- [13] I. Tougui, A. Jilbab, and J. El Mhamdi, "Impact of the choice of cross-validation techniques on the results of machine learning-based diagnostic applications," *Healthc. Inform. Res.*, vol. 27, no. 3, pp. 189–199, 2021, doi: 10.4258/HIR.2021.27.3.189.
- [14] S. A. Hicks *et al.*, "On evaluation metrics for medical applications of artificial intelligence," *medRxiv*, p. 2021.04.07.21254975, Jan. 2021, doi: 10.1101/2021.04.07.21254975.