

KLASIFIKASI FAKTOR RISIKO PENYAKIT JANTUNG MENGGUNAKAN *MACHINE LEARNING*

Mailiana Anita¹, Imelda Grecea Dwi Yulianti², Swarno Varestama Pasaribu³

Program Studi Teknik Informatika, Universitas Widya Dharma Pontianak

Jl. H.O.S. Cokroaminoto, Pontianak – Kalimantan Barat, Indonesia

Email: ¹anitamailiana@gmail.com, ²imeldagdy@gmail.com, ³swarnovarestama@gmail.com

ABSTRAK

Penyakit jantung merupakan salah satu penyebab utama kematian di seluruh dunia, termasuk di Indonesia. Deteksi dini terhadap risiko penyakit jantung sangat penting untuk mencegah komplikasi serius namun tantangan utama yang dihadapi adalah deteksi dini penyakit ini sering terlambat karena proses klasifikasinya belum cukup akurat. Penelitian ini bertujuan membangun model klasifikasi risiko penyakit jantung dengan pendekatan *machine learning*. Dataset yang digunakan berasal dari Kaggle, terdiri dari 920 data pasien dengan 13 fitur medis yang relevan. Proses penelitian meliputi pra-pemrosesan data, seleksi fitur menggunakan metode *Recursive Feature Elimination (RFE)*, pelatihan model dengan enam algoritma (*Random Forest*, *Support Vector Machine*, *K-Nearest Neighbors*, *Naive Bayes*, *Decision Tree*, dan *Logistic Regression*), serta evaluasi kinerja model menggunakan metrik akurasi, *precision*, *recall*, *f1-score*, ROC-AUC, dan analisis tingkat kepentingan fitur. Hasil penelitian menunjukkan bahwa algoritma *Support Vector Machine (SVM)* memberikan performa terbaik dengan akurasi sebesar 83,91% dan *AUC* sebesar 0,92, diikuti oleh *Random Forest* dengan akurasi 82,61%. Fitur yang paling berkontribusi terhadap hasil klasifikasi adalah tipe nyeri dada (cp), jumlah pembuluh darah utama (ca), dan jenis talasemia (thal). Temuan ini menunjukkan bahwa penerapan *machine learning* dengan seleksi fitur yang tepat dapat meningkatkan akurasi klasifikasi risiko penyakit jantung dan berpotensi digunakan sebagai sistem pendukung keputusan dalam dunia medis.

Kata kunci: penyakit jantung, klasifikasi, *machine learning*, seleksi fitur, *Support Vector Machine*

ABSTRACT

Heart disease is one of the leading causes of death worldwide, including in Indonesia. Early detection of heart disease risk is very important to prevent serious complications but the main challenge faced is that early detection of this disease is often late because the classification process is not accurate enough. This study aims to build a heart disease risk classification model with a machine learning approach. The dataset used comes from Kaggle, consisting of 920 patient data with 13 relevant medical features. The research process includes data pre-processing, feature selection using the Recursive Feature Elimination (RFE) method, model training with six algorithms (Random Forest, Support Vector Machine, K-Nearest Neighbors, Naive Bayes, Decision Tree, and Logistic Regression), and model performance evaluation using accuracy, precision, recall, f1-score, ROC-AUC metrics, and feature importance analysis. The results showed that the Support Vector Machine (SVM) algorithm provided the best performance with an accuracy of 83.91% and AUC of 0.92, followed by Random Forest with an accuracy of 82.61%. The features that contributed most to the classification results were chest pain type (cp), number of major blood vessels (ca), and thalassemia type (thal). These findings suggest that the application of machine learning with appropriate feature selection can improve the accuracy of heart disease risk classification and has the potential to be used as a decision support system in the medical world.

Keywords: heart disease, classification, machine learning, feature selection, *Support Vector Machine*

1. PENDAHULUAN

Penyakit jantung merupakan penyebab kematian paling umum di seluruh dunia. Menurut laporan terbaru WHO, penyakit jantung merupakan penyebab kematian utama di seluruh dunia, yang merenggut sekitar 17,9 juta jiwa setiap tahunnya [1]. Di Indonesia sendiri, menurut menteri kesehatan RI, Budi Gunadi Sadikin, stroke dan

Copyright (c) HOAQ (High Education of Organization Archive Quality) : Jurnal Teknologi Informasi



penyakit jantung merupakan penyebab kematian utama di Indonesia, dengan perkiraan 1,5 juta kematian setiap tahunnya [2]. Deteksi dini penyakit jantung menjadi tantangan besar dalam dunia medis karena sering dikaitkan dengan rendahnya akurasi dalam klasifikasi penyakit jantung [3]. Banyak orang yang baru menyadari dirinya mengidap penyakit jantung ketika sudah mencapai stadium yang sangat parah, hal ini dapat menyebabkan penanganan menjadi tertunda dan membahayakan nyawa. Rendahnya akurasi dalam klasifikasi penyakit jantung menimbulkan kesulitan dimana model atau algoritma klasifikasi sulit membedakan berbagai jenis penyakit jantung. Penelitian oleh Dedik dan Zico (2025) juga menekankan pentingnya penerapan algoritma klasifikasi dalam deteksi dini penyakit jantung koroner berdasarkan gejala klinis, sehingga intervensi medis dapat dilakukan lebih cepat [4]. Oleh karena itu, penting untuk mengembangkan suatu sistem prediksi yang mampu mendeteksi risiko sejak dini.

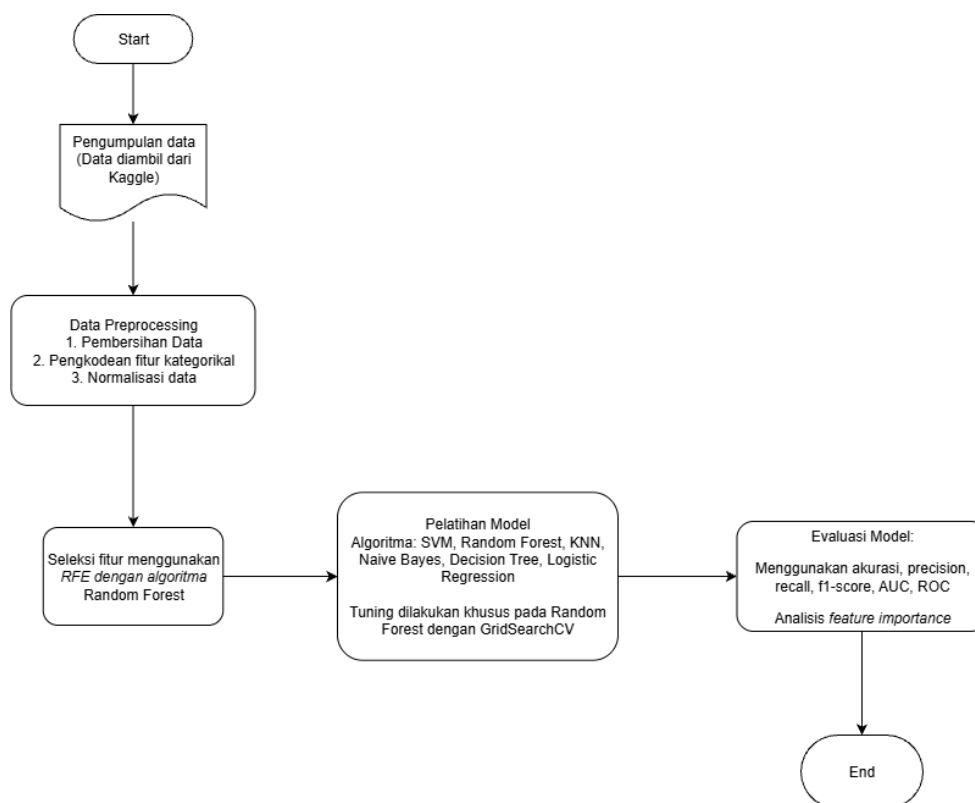
Dengan kemajuan teknologi yang luar biasa, telah mendorong sejumlah penelitian yang bertujuan memfasilitasi diagnosis penyakit jantung melalui penerapan *machine learning* [5]. Dalam beberapa tahun terakhir, *machine learning* telah menjadi pendekatan populer untuk menganalisis data medis karena kemampuannya mengidentifikasi pola dari kumpulan data besar. Penelitian ini bertujuan untuk mengembangkan model klasifikasi penyakit jantung menggunakan pendekatan *feature selection* dan mengevaluasi beberapa model *machine learning*. Evaluasi eksperimental akan dilakukan untuk menganalisis hasil percobaan, dengan harapan dapat meningkatkan akurasi identifikasi pasien yang berisiko terkena penyakit jantung. Tujuan utama penelitian ini adalah untuk berkontribusi dalam meningkatkan diagnosis dini dan penanganan penyakit jantung.

Dalam penelitian ini menggunakan metode *Random Forest*, *Decision Tree*, *Naïve Bayes*, *KNN*, *SVM* dan *Logistic Regression* untuk meningkatkan akurasi dan stabilitas prediksi dengan membandingkan beberapa algoritma tersebut. Penggunaan keenam algoritma ini bertujuan untuk membandingkan performa model dan memilih algoritma terbaik berdasarkan metrik seperti *akurasi*, *precision*, *recall*, *F1-score*, dan *AUC*. Kemudian untuk menggali karakteristik masing-masing model, seperti interpretabilitas (*Decision Tree*, *Logistic Regression*), keandalan klasifikasi (*SVM*, *Random Forest*), dan efisiensi komputasi (*Naive Bayes*). Serta menjamin generalisasi dan stabilitas hasil, dengan uji silang (*cross-validation*) dan tuning parameter. Penggunaan berbagai algoritma ini bertujuan untuk menemukan model terbaik berdasarkan akurasi, *AUC*, dan kemudahan interpretasi dalam mendukung diagnosis penyakit jantung.

Penelitian serupa telah dilakukan menggunakan algoritma *K-Nearest Neighbors(KNN)* dan *Support Vector Machine(SVM)* dalam kualifikasi penyakit serangan jantung, yang menunjukkan bahwa kedua metode ini efektif dalam mengidentifikasi pasien berisiko tinggi [6]. Selain itu penelitian yang sama juga membahas pemanfaatan *machine learning* dalam klasifikasi penyakit jantung dan mendukung pentingnya pendekatan algoritmik dalam diagnosis medis [7]. Penelitian sebelumnya juga membandingkan 3 algoritma *machine learning* yaitu *KNN*, *Naive Bayes*, dan *Logistic Regression* yang dimana *KNN* menghasilkan akurasi tertinggi sebesar 91% [8]. Penelitian yang serupa juga membandingkan tujuh metode *machine learning* (*Naive Bayes*, *K-Nearest Neighbors*, *Random Forest*, *Logistic Regression*, *Support Vector Machine*, *Decision Tree* dan *AdaBoost*) yang menggunakan dataset Cleveland dari UCI dengan hasil menunjukkan bahwa *Naive Bayes* menghasilkan akurasi tertinggi sebesar 84,67%, diikuti oleh *Logistic Regression* 84,30% dan *Random Forest* 81,70% [9]. Namun, sebagian besar penelitian tersebut hanya membandingkan beberapa algoritma tanpa seleksi fitur yang sistematis, tanpa *tuning hyperparameter*, dan tidak menampilkan kontribusi tiap fitur terhadap hasil prediksi. Berbeda dengan penelitian sebelumnya yang memfokuskan pada perbandingan performa algoritma tanpa pendekatan seleksi fitur atau optimasi hyperparameter, penelitian ini menggunakan perbandingan enam algoritma yang dikombinasikan dengan metode seleksi fitur *Recursive Feature Elimination (RFE)* serta evaluasi yang lebih komprehensif (*precision*, *recall*, *F1-score*, dan *ROC-AUC*). Dengan demikian, pendekatan ini menawarkan hasil yang lebih mendalam dibanding studi sebelumnya, meskipun masih perlu dikembangkan dengan penerapan teknik penanganan data tidak seimbang serta analisis interpretabilitas lanjutan seperti *SHAP* atau *LIME* untuk meningkatkan kepercayaan pengguna medis terhadap sistem.

2. METODE PENELITIAN

Berikut langkah-langkah penelitian:



Gambar 1. Langkah Penelitian

Pengumpulan Data

Pengumpulan data menggunakan dataset publik yang dimana dataset yang digunakan berasal dari *platform Kaggle*, berisi 920 data pasien dengan 13 fitur merepresentasikan berbagai atribut medis dan demografis pasien, yaitu usia (age), jenis kelamin (sex), tipe nyeri dada (cp), tekanan darah istirahat (trestbps), kolesterol serum (chol), gula darah puasa (fbs), hasil elektrokardiogram istirahat (restecg), detak jantung maksimum yang dicapai (thalch), angina yang diinduksi oleh latihan (exang), depresi segmen ST yang diinduksi oleh latihan relatif terhadap istirahat (oldpeak), kemiringan segmen ST (slope), jumlah pembuluh darah utama yang diwarnai oleh fluoroskopi (ca), dan tipe talasemia (thal)[10]. Target klasifikasi adalah status penyakit jantung pasien yang diberi label num, dikodekan sebagai 0 (tidak sakit) dan 1-4 (sakit). Kumpulan data ini telah melalui proses pembersihan dan validasi sehingga dapat digunakan untuk pengembangan model *machine learning*. Kumpulan data ini telah mendapat pengakuan di komunitas *Kaggle* dan diberi label peringkat kualitas gold, sehingga mencakup kumpulan data mana yang layak dipercaya dan populer.

Tabel 1. 13 fitur atribut dalam dataset

Fitur	Keterangan
age	Umur Pasien
sex	Jenis Kelamin Pasien 0 = perempuan 1 = laki-laki
Cp	Jenis Nyeri Dada: 0 = typical angina 1 = atypical angina 2 = non-anginal 3 = asymptomatic
Trestbps	Tekanan Darah (mm Hg)
Chol	Kolestrol (mg/dl)
Fbs	Gula darah puasa 0 = false 1 = true

Restecg	Resting Relectrocardiographic” 0 = normal 1 = ST-T abnormality 2 = lv hypertrophy
Thalch	Denyut jantung maksimum
Exang	Olahraga: 0 = True 1 = False
Oldpeak	Depresi
Slope	The slope of the peak exercise segment: 0 = upsloping 1 = flat 2 = downsloping
Ca	Flurosopi
Thal	Jenis Talasemia 0 = normal 1 = fixed defect 2 = reversable defect
Num	Diagnosa 0 = sehat 1 – 4 = sakit jantung

Data Preprocessing

Proses data preprocessing merupakan tahapan krusial dalam pengolahan data sebelum dilakukan pelatihan model pembelajaran mesin. Pada penelitian ini, tahapan preprocessing dilakukan melalui beberapa langkah sistematis sebagai berikut:

1. Pembersihan Data
Dilakukan pemeriksaan terhadap data numerik untuk mengidentifikasi serta memperbaiki anomali atau ketidaksesuaian format yang dapat mengganggu proses analisis lebih lanjut.
2. Pengkodean Fitur Kategorikal
Seluruh fitur kategorikal dikonversi ke dalam format numerik menggunakan metode *Label Encoding* agar dapat dikenali dan diproses secara optimal oleh algoritma pembelajaran mesin.
3. Normalisasi Data
Data numerik kemudian dinormalisasi menggunakan teknik *StandardScaler* guna memastikan bahwa semua fitur memiliki skala distribusi yang seragam (mean = 0, standard deviation = 1), sehingga menghindari dominasi fitur tertentu dalam proses pelatihan model.

Seleksi fitur

Proses seleksi fitur dalam penelitian ini menggunakan *metode Recursive Feature Elimination (RFE)* dengan algoritma *Random Forest* sebagai estimator. *Random Forest* dipilih karena memiliki kemampuan bawaan dalam mengukur tingkat kepentingan fitur secara akurat, mampu menangani hubungan non-linear serta interaksi antar fitur, dan relatif stabil terhadap overfitting. Selain itu, *Random Forest* tidak sensitif terhadap skala data, sehingga cocok digunakan dalam proses seleksi fitur berulang tanpa memerlukan transformasi tambahan. Oleh karena itu, algoritma ini dinilai lebih efektif dalam mengoptimalkan pemilihan fitur yang relevan dibandingkan algoritma yang lain.

Pendekatan ini merupakan metode seleksi berbasis wrapper (wrapper-based) yang mengevaluasi pentingnya fitur berdasarkan performa model. Karena tidak melibatkan kombinasi dengan teknik seleksi lainnya seperti metode filter atau embedded, maka *RFE* yang diterapkan bersifat non-hybrid.

RFE bekerja dengan membangun model secara rekursif dan mengeliminasi fitur-fitur yang memiliki kontribusi paling kecil terhadap performa model. Dari total 13 fitur yang tersedia, dipilih 7 fitur terbaik berdasarkan kontribusinya dalam meningkatkan akurasi model. Ketujuh fitur tersebut adalah: age, cp (chest pain type), trestbps (resting blood pressure), chol (serum cholesterol), thalach (maximum heart rate achieved), ca (number of major vessels colored by fluoroscopy), dan thal (thalassemia).

Pelatihan Model

Dalam penelitian ini, tuning hyperparameter dilakukan hanya pada algoritma *Random Forest* menggunakan metode *GridSearchCV* dengan *5-fold cross-validation*. Parameter yang dituning mencakup *n_estimators* (jumlah pohon), *max_depth* (kedalaman maksimum), dan *criterion* (fungsi pemisahan: *gini* atau *entropy*). Sementara itu, algoritma lain seperti *SVM*, *KNN*, *Logistic Regression*, *Naive Bayes*, dan *Decision Tree* menggunakan konfigurasi default dari pustaka *scikit_learn*, tanpa proses tuning tambahan. Hal ini dilakukan untuk menjaga fokus pada baseline performa serta menghindari bias tuning tidak merata antar model. Adapun konfigurasi hyperparameter yang digunakan pada masing-masing algoritma dirangkum dalam Tabel 2.

Tabel 2. Konfigurasi Hyperparameter Model

Algoritma	Hyperparameter Utama	Nilai Yang Digunakan
<i>Random Forest</i>	<i>n_estimators, max_depth, criterion</i>	Dituning via <i>GridSearchCV</i>
<i>SVM</i>	<i>C = 1.0, kernel = 'rbf', gamma = 'scale'</i>	Default <i>scikit-learn</i>
<i>K-Nearest Neighbors</i>	<i>n_neighbors = 5, metric = 'minkowski'</i>	Default <i>scikit-learn</i>
<i>Logistic Regression</i>	<i>solver = 'lbfgs', penalty = 'l2', C = 1</i>	Default <i>scikit-learn</i>
<i>Naive Bayes</i>	<i>GaussianNB</i> (asumsi distribusi normal)	Default <i>scikit-learn</i>
<i>Decision Tree</i>	<i>criterion = 'gini', max_depth = None</i>	Default <i>scikit-learn</i>

Penggunaan nilai default pada sebagian besar algoritma bertujuan untuk membandingkan baseline performa model tanpa intervensi tuning, sementara tuning pada *Random Forest* dilakukan untuk mengoptimalkan performa klasifikasi.

Evaluasi Model

Evaluasi model dalam penelitian ini dilakukan untuk menilai kestabilan dan kemampuan generalisasi dari algoritma yang digunakan dalam klasifikasi risiko penyakit jantung. Tahapan evaluasi dilakukan dengan pendekatan metode validasi silang (*cross-validation*) dan pengukuran performa model menggunakan berbagai metrik evaluasi.

1. Metode Evaluasi

Untuk memastikan bahwa model tidak hanya memiliki performa yang baik pada data pelatihan tetapi juga mampu bekerja secara konsisten pada data yang tidak terlihat sebelumnya, penelitian ini menggunakan *5-Fold Cross-Validation*. Metode ini membagi data menjadi lima bagian yang seimbang, empat bagian digunakan untuk pelatihan model dan satu bagian untuk pengujian, yang kemudian diputar sebanyak lima kali agar seluruh data digunakan baik secara data latih maupun data uji. Pendekatan ini bertujuan untuk mengurangi bias dan varian pada proses evaluasi serta memberikan gambaran yang lebih representatif terhadap performa model secara umum.

Selain itu, untuk algoritma *Random Forest*, dilakukan proses penyetelan hyperparameter menggunakan teknik *GridSearchCV* yang juga mengimplementasikan validasi silang. Tujuan dari penyetelan ini adalah untuk menemukan kombinasi parameter terbaik (seperti jumlah estimator, kedalaman pohon, dan fungsi pemisah) yang dapat mengoptimalkan performa model.

2. Metrik Evaluasi Kinerja Model

Untuk mengukur kinerja klasifikasi dari setiap algoritma, digunakan sejumlah metrik evaluasi yang mencakup aspek presisi, sensitivitas, keseimbangan performa, dan diskriminasi kelas. Metrik-metrik tersebut meliputi :

1. *Confusion Matrix*

Confusion matrix digunakan untuk merepresentasikan jumlah prediksi benar dan salah dari model pada masing-masing kelas. Matriks ini terdiri atas empat komponen utama: True Positif (TP), False Positive (FP), True Negative (TN), dan False Negative (FN).

Pada penelitian ini, *confusion matrix* hanya ditampilkan untuk model *Random Forest* sebagai representasi performa klasifikasi. Pemilihan ini didasarkan pada dua alasan utama: pertama, *Random Forest* merupakan model dengan performa tertinggi kedua setelah *SVM*, dengan akurasi 82,61% dan *AUC* 0,91; kedua, model ini digunakan sebagai estimator dalam proses seleksi fitur *RFE*, sehingga memiliki peran sentral dalam alur penelitian. Selain itu, *Random Forest* dipilih karena memiliki interpretabilitas yang baik dan distribusi prediksinya dapat menjelaskan pola klasifikasi secara lebih jelas. Metrik evaluasi lain seperti akurasi, *precision*, *recall*, *F1-score*, dan *AUC* tetap digunakan untuk semua model untuk menjaga konsistensi analisis. Berikut hasil *confusion matrix* dari *Random Forest* dalam penelitian ini:

[[81 24],

[16 109]]

Copyright (c) HOAQ (High Education of Organization Archive Quality) : Jurnal Teknologi Informasi

Yang mengindikasikan 81 prediksi negatif benar (TN), 109 prediksi positif benar (TP), 24 prediksi positif salah (FP), dan 16 prediksi negatif salah (FN).

2. Akurasi (*Accuracy*)

Mengukur proporsi prediksi yang benar terhadap seluruh jumlah data:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

3. Presisi (*Precision*)

Menggambarkan tingkat ketepatan model dalam memprediksi kelas positif:

$$Precision = \frac{TP}{TP + FP}$$

4. Sensitivitas (*Recall*)

Mengukur kemampuan model dalam mengidentifikasi seluruh kasus positif:

$$Recall = \frac{TP}{TP + FN}$$

5. *F1-score*

Merupakan rata-rata harmonis dari precision dan recall, digunakan saat diperlukan keseimbangan antara keduanya:

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

6. *ROC Curve* dan *AUC* (Area Under the Curve)

Kurva *ROC* menggambarkan hubungan antara True Positive Rate (TPR) dan False Positive Rate (FPR) untuk berbagai ambang batas klasifikasi. Sementara itu, nilai *AUC* digunakan untuk mengukur luas dibawah kurva *ROC*, yang menunjukkan kemampuan model dalam membedakan antara kelas positif dan negatif. Nilai *AUC* yang mendekati 1 mengindikasikan performa klasifikasi yang sangat baik.

Penggunaan kombinasi metrik tersebut memberikan pemahaman yang menyeluruh mengenai kinerja klasifikasi dari masing-masing model, baik dari segi akurasi umum, sensitivitas terhadap kasus positif, hingga keseimbangan antara kesalahan tipe I dan tipe II. Pendekatan ini penting khususnya dalam konteks medis, dimana kesalahan klasifikasi dapat berdampak serius terhadap keputusan klinis.

Rumus masing-masing model

Rumus dari masing-masing algoritma *machine learning* yang digunakan dalam penelitian ini:

1. Support Vector Machine (SVM)

SVM berusaha mencari hyperplane terbaik untuk memisahkan dua kelas.

Rumus umum:

$$\min_{w,b} \frac{1}{2} \|w\|^2 \quad \text{dengan syarat} \quad y_i(w \cdot x_i + b) \geq 1$$

- w : vector bobot
- b : bias
- y_i : label kelas $\in \{-1, +1\}$
- x_i : vector fitur

2. Random Forest

Random Forest adalah Kumpulan pohon keputusan (*decision trees*) dengan teknik voting. Prediksi ditentukan berdasarkan mayoritas suara dari pohon-pohon tersebut.

$$\hat{y} = \text{mode}(h_1(x), h_2(x), \dots, h_k(x))$$

- $h_k(x)$: output dari decision tree ke-k

3. K-Nearest Neighbors (KNN)

Prediksi dilakukan berdasarkan k tetangga terdekat menggunakan jarak Euclidean:

$$d(x, x_i) = \sqrt{\sum_{j=1}^n (x_j - x_{ij})^2}$$

- x : data baru
- x_i : data dalam training set
- n : jumlah fitur

4. Logistic Regression

Model probabilitas klasifikasi biner menggunakan fungsi sigmoid:

$$P(y = 1|x) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n)}}$$

- β : koefisien regresi
- x : vector fitur

5. Naïve Bayes

Berdasarkan Teorema Bayes dengan asumsi independensi antar fitur :

$$P(C_k|x) = \frac{P(x|C_k) \cdot P(C_k)}{P(x)}$$

- C_k : Kelas ke- k yang akan diprediksi.
- x : data baru atau fitur masukan.
- $P(C_k | x)$: probabilitas bahwa data x termasuk dalam kelas C_k (posterior probability).
- $P(x | C_k)$: probabilitas munculnya fitur x jika diketahui berada dalam kelas C_k (likelihood).
- $P(C_k)$: probabilitas awal terjadinya kelas C_k tanpa melihat data (prior).
- $P(x)$: probabilitas munculnya data x di semua kelas (evidence).

6. Decision Tree

Menggunakan pemisahan berdasarkan information gain atau Gini impurity.

$$Gini(D) = 1 - \sum_{i=1}^c P_i^2$$

- P_i : proporsi kelas i dalam dataset D
- C : jumlah kelas

3. HASIL DAN PEMBAHASAN

Penelitian ini mengevaluasi performa enam algoritma *machine learning*, yaitu *Random Forest*, *Support Vector Machine (SVM)*, *K-Nearest Neighbors (KNN)*, *Naive Bayes*, *Decision Tree*, dan *Logistic Regression*, dalam mengklasifikasikan risiko penyakit jantung[11]. Pengujian dilakukan menggunakan bahasa pemrograman *python* dengan bantuan beberapa pustaka populer seperti *scikit-learn* untuk permodelan dan evaluasi, *pandas* untuk manipulasi data yaitu mengganti teks menjadi angka; membersihkan data numerik dari anomali format; dan imputasi nilai kosong dengan rata-rata kolom terkait, *matplotlib* dan *seaborn* untuk visualisasi, serta *GridSearchCV* untuk *tuning hyperparameter*.

Sebelum pelatihan model, dilakukan proses *feature selection* menggunakan *Recursive Feature Elimination (RFE)*, yang berhasil menyeleksi 7 fitur terbaik dari total 13 fitur dalam dataset. 7 fitur tersebut meliputi age, cp, trestbps, chol, thalach, ca, dan thal. Setelahnya dilakukan evaluasi untuk tiap algoritma yang ada, evaluasi dilakukan dengan metrik akurasi, *precision*, *recall*, dan *f1-score*, serta analisis *feature importance* untuk memahami kontribusi masing-masing fitur terhadap prediksi. Kemudian melakukan *feature importance* untuk melihat fitur mana yang paling berpengaruh dalam proses pengambilan keputusan model *machine learning*.

Hasil Analisis Kepentingan Fitur

Tabel 3. *Feature Importance* dari Semua Model

Fitur	<i>Random Forest</i>	<i>Naive Bayes</i>	<i>Decision Tree</i>	<i>Logistic Regression</i>	<i>SVM</i>	<i>KNN</i>
cp	0.24	0.02	0.27	0.80	0.04	0.03
thal	0.21	0.01	0.12	0.40	0.03	0.06
ca	0.18	0.08	0.15	1.20	0.085	0.04
thalach	0.15	0.01	0.17	0.60	0.01	0.005
chol	0.12	0.03	0.11	0.35	0.045	0.035
age	0.08	0.01	0.10	0.20	0.02	0.01
trestbps	0.05	0.005	0.09	0.30	0.01	0.005

Tabel 3 menunjukkan perbandingan skor kepentingan fitur (*feature importance*) dari tujuh variabel utama yang digunakan dalam model prediksi penyakit jantung, berdasarkan enam algoritma klasifikasi yaitu *Random Forest*, *SVM*, *Naive Bayes*, *Decision Tree*, *KNN*, dan *Logistic Regression*. Fitur-fitur seperti cp (jenis nyeri dada), ca (jumlah pembuluh darah utama yang terlihat melalui fluoroskopi), dan thal (kelainan perfusi darah) memiliki

skor yang tinggi dalam sebagian besar model, hal itu menunjukkan bahwa ketiga fitur tersebut sangat memengaruhi hasil klasifikasi. Fitur-fitur seperti chol (kadar kolestrol), age, dan trestbps (tekanan darah istirahat) cenderung memiliki pengaruh yang lebih rendah terhadap output pada semua algoritma dan skor yang lebih rendah dari fitur-fitur ini menunjukkan bahwa walaupun tetap relevan secara klinis mereka tidak memberikan kontribusi sebesar cp, ca, dan thal dalam klasifikasi akhir.

Model *Naive Bayes* dan *Decision Tree* menunjukkan bahwa fitur cp merupakan fitur dengan skor tertinggi, yaitu masing-masing sebesar 0.24 dan 0.27. Sementara itu, *Logistic Regression* memberikan skor tertinggi kepada fitur ca yaitu sebesar 1.20, artinya fitur ini sangat dominan dalam menentukan prediksi. Model *Naive Bayes* cenderung memberikan skor rendah pada semua fitur, tetapi fitur ca tetap yang paling berpengaruh dibandingkan fitur lainnya. Pada *SVM* dan *KNN*, distribusi skor cukup merata dan lebih kecil, menandakan bahwa kedua algoritma ini lebih konservatif dalam menetapkan pentingnya fitur paling besar dalam prediksi risiko penyakit jantung. Hal ini sesuai dengan kajian klinis bahwa gejala nyeri dada (cp) serta hasil tes visualisasi pembuluh darah (ca) dan perfusi darah (thal) merupakan indikator penting dalam diagnosis penyakit jantung.

Temuan ini memberikan implikasi bahwa fitur-fitur tersebut perlu diperhatikan dalam sistem pendukung keputusan berbasis *machine learning*, karena memberikan kontribusi terbesar dalam akurasi prediksi. Sistem ini dapat digunakan untuk membantu dokter dalam melakukan deteksi dini dan memberikan intervensi yang tepat.

Hasil Evaluasi Kinerja Model

Berdasarkan hasil evaluasi, didapatkan performa masing-masing model sebagai berikut:

Tabel 4. Hasil Evaluasi Kinerja dari Semua Model

Model	Akurasi	Precision	Recall	F1-Score	AUC
<i>SVM</i>	83.91%	0.84	0.83	0.83	0.91
<i>Random Forest</i>	82.61%	0.82	0.82	0.82	0.91
<i>KNN</i>	81.74%	0.81	0.81	0.81	0.87
<i>Logistic Regression</i>	80.00%	0.79	0.80	0.79	0.87
<i>Naive Bayes</i>	79.57%	0.79	0.79	0.79	0.89
<i>Decision Tree</i>	78.26%	0.78	0.77	0.77	0.77

Tabel 4 menyajikan hasil evaluasi kinerja dari enam algoritma *machine learning* yang digunakan untuk klasifikasi risiko penyakit jantung, yaitu *Support Vector Machine (SVM)*, *Random Forest*, *K-Nearest Neighbors (KNN)*, *Logistic Regression*, *Naive Bayes*, dan *Decision Tree*. Evaluasi dilakukan berdasarkan metrik akurasi, *precision*, *recall*, *F1-score*, dan *Area Under the Curve (AUC)*, yang secara kolektif memberikan gambaran menyeluruh mengenai performa masing-masing model.

Dalam penelitian ini, algoritma *Random Forest* digunakan dalam dua peran utama yaitu yang pertama sebagai estimator dalam metode seleksi fitur berbasis *wrapper Recursive Feature Elimination (RFE)* dan kedua, sebagai salah satu algoritma klasifikasi yang dievaluasi setelah proses seleksi fitur. Hal ini bertujuan untuk mengoptimalkan pemilihan fitur relevan sekaligus menguji kemampuan prediktif *Random Forest* dibandingkan model lainnya.

Dari berbagai algoritma yang dicoba, *SVM* menunjukkan performa terbaik, dengan akurasi 83,91% dan AUC 0,91. Model memiliki kemampuan yang sangat baik dalam membedakan antara pasien yang berisiko dan tidak berisiko penyakit jantung. *Precision* dan *recall* yang cukup seimbang 0,84, 0,83 menunjukkan bahwa model tidak hanya mampu mengenali sebagian besar kasus positif secara benar, tetapi juga mengurangi jumlah prediksi salah yang bersifat positif palsu.

Dengan akurasi 82.61% dan AUC yang sama tinggi, yaitu 0.91, model *Random Forest* menempati posisi kedua di antara dataset itu. Model ini menunjukkan keseimbangan yang cukup baik antara sensitivitas dan spesifisitas, serta menunjukkan stabilitas performa di seluruh metrik evaluasi. Hasil serupa juga diperoleh dengan penerapan teknik *SMOTEEN* pada algoritma *Random Forest* berhasil meningkatkan akurasi prediksi penyakit jantung, mendukung hasil penelitian ini yang juga menempatkan *Random Forest* sebagai salah satu model dengan performa tinggi [12]. Selain itu, penelitian yang serupa juga membandingkan algoritma *Random Forest* dan *SVM* dalam klasifikasi penyakit jantung, yang memperlihatkan kedua model ini unggul dalam mendeteksi risiko penyakit jantung [13]. Hal ini menjadikan *Random Forest* dan *SVM* sebagai kandidat kuat untuk diterapkan dalam skenario klinis yang membutuhkan model prediksi yang andal dan konsisten.

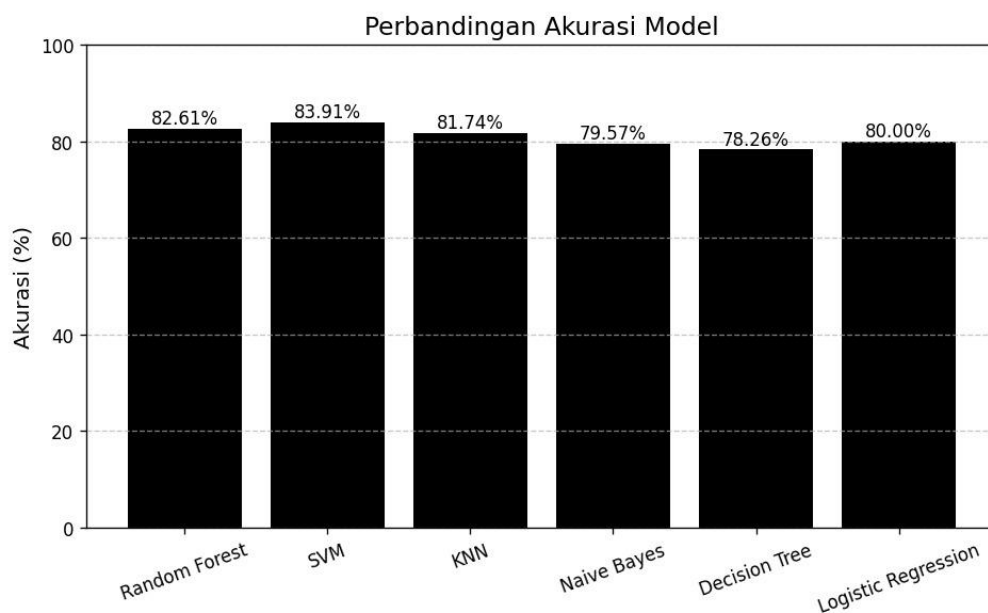
Algoritma *K-Nearest Neighbors (KNN)* mencatatkan akurasi sebesar 81.74% dengan nilai *F1-score* dan *recall* sebesar 0.81. Meskipun performanya berada sedikit di bawah *SVM* dan *Random Forest*, *KNN* tetap dapat menjadi alternatif yang baik untuk aplikasi yang memerlukan interpretasi sederhana serta implementasi yang relatif mudah.

Logistic Regression dan *Naive Bayes* menunjukkan performa yang cukup kompetitif, masing-masing dengan akurasi 80.00% dan 79.57%, serta nilai *AUC* yang juga tinggi, yakni 0.87 dan 0.89. Meskipun kedua algoritma ini tidak sekuat *SVM* dalam hal akurasi, namun keduanya memiliki keunggulan dari sisi efisiensi komputasi dan interpretabilitas, sehingga cocok digunakan sebagai baseline model atau pada sistem dengan sumber daya terbatas.

Sebaliknya, *Decision Tree* mencatatkan kinerja terendah dengan akurasi 78.26% dan *AUC* 0.77. Meskipun *recall* pada kelas positif (penderita penyakit jantung) cukup tinggi, yaitu 0.77, *precision*-nya cenderung rendah. Hal ini menunjukkan bahwa model ini berpotensi menghasilkan banyak prediksi positif palsu, yang dalam konteks medis dapat menyebabkan overdiagnosis dan kecemasan yang tidak perlu pada pasien. Adapun penelitian sebelumnya juga menerapkan algoritma *Decision Tree* dalam prediksi penyakit jantung dan memperoleh akurasi sebesar 80% dengan *F1-score* sebesar 0,82 [14]. Namun, penelitian tersebut hanya berfokus pada satu algoritma tanpa perbandingan dengan model lain atau penerapan teknik seleksi fitur. Hal ini memperkuat keunggulan penelitian ini yang tidak hanya mengevaluasi enam algoritma berbeda, tetapi juga mengintegrasikan metode seleksi fitur *RFE* serta analisis *feature importance* sebagai upaya untuk meningkatkan akurasi dan interpretabilitas model dalam konteks medis.

Secara keseluruhan, temuan ini mengindikasikan bahwa *SVM* dan *Random Forest* merupakan dua algoritma paling optimal dalam mendeteksi risiko penyakit jantung berdasarkan data yang digunakan. Hal ini diperkuat dengan nilai *AUC* yang tinggi, yang menandakan bahwa kedua model memiliki kemampuan generalisasi yang baik dalam membedakan kelas target. Oleh karena itu, model-model ini berpotensi besar untuk diintegrasikan dalam sistem pendukung keputusan klinis berbasis data.

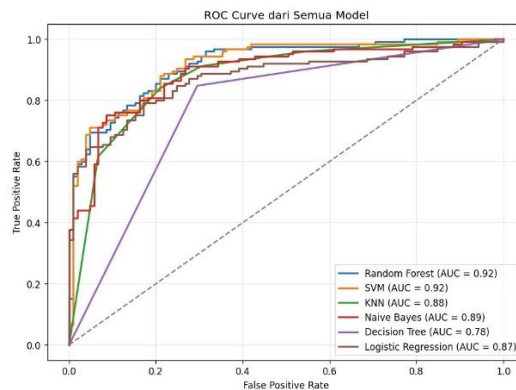
Perbandingan Akurasi Model



Gambar 2. Perbandingan Akurasi Model

Seperti yang terlihat pada Gambar 2, algoritma *Support Vector Machine (SVM)* menghasilkan akurasi tertinggi sebesar 83.91%, kemudian diikuti oleh *Random Forest* sebesar 82.61%, dan *KNN* sebesar 81.74%, di sisi lain, algoritma *Naive Bayes*, *Decision Tree* dan *Logistic Regression* masing-masing mencatatkan akurasi sebesar 79.57%, 78.26%, dan 80.00%. Oleh karena itu, dari tabel ini diketahui bahwa *SVM* dan *Random Forest* lebih mampu menangkap pola kompleks pada data medis dibanding model lainnya.

Analisis ROC Curve dan AUC



Gambar 3. ROC Curve dari Semua model

Gambar 3 memperlihatkan kurva *ROC* dari masing-masing model. Nilai *AUC* (*Area Under Curve*) menunjukkan keseluruhan performa klasifikasi. Baik *Random Forest* maupun *SVM* memiliki *AUC* tertinggi yaitu 0.92, menandakan performa yang sangat baik. Model *Naive Bayes* dan *KNN* mengikuti dengan *AUC* sebesar 0.89 dan 0.88. *Logistic Regression* mencatat 0.87, sedangkan *Decision Tree* berada diposisi terakhir dengan *AUC* 0.78. Ini mengonfirmasi bahwa *Random Forest* dan *SVM* tidak hanya unggul dalam akurasi tetapi juga dalam keseimbangan antara sensitivitas dan spesifisitas.

4. SIMPULAN

Kesimpulannya, hasil penelitian ini menunjukkan bahwa penerapan algoritma *machine learning* dalam klasifikasi penyakit jantung merupakan kinerja yang menjanjikan. Hasil terbaik diperoleh dari algoritma *SVM* yang memberikan akurasi yang baik dan seimbang dalam menebak kelas positif dan negatif. *Random Forest* juga memberikan hasil yang kompetitif. Efektivitas model meningkat secara signifikan setelah pemilihan fitur dalam alur kerja, yang dilakukan dengan menggunakan metode *RFE*. Hal ini menyederhanakan jumlah input dengan menyisakan fitur yang paling relevan.

Selain itu, analisis lebih lanjut tentang pentingnya fitur juga menunjukkan bahwa variabel medis memiliki peran dominan dalam klasifikasi risiko penyakit jantung. Oleh karena itu, temuan ini memperkuat bahwa model berbasis *machine learning* dapat diterapkan pada pengendalian kualitas klasifikasi dengan memberikan wawasan yang signifikan bagi para pengambil keputusan klinis berdasarkan data.

Secara rinci, temuan dari penelitian ini dapat dibagi menjadi:

1. Algoritma *SVM* menunjukkan kinerja tertinggi, dengan akurasi 83.91% dan nilai *AUC* 0.92, yang berarti jika *SVM* adalah model yang efektif dalam mencari pasien yang berisiko kena penyakit jantung.
2. Metode seleksi fitur *RFE* berhasil meningkatkan efisiensi model, dengan hanya mempertahankan tujuh fitur dari tiga belas yang tersedia tanpa menurunkan performa klasifikasi secara signifikan.
3. Elemen-elemen yang paling berpengaruh dalam klasifikasi mencakup *cp* (tipe nyeri dada), *ca* (jumlah arteri utama yang terlihat melalui fluoroskopi), dan *thal* (kategori talasemia), yang secara konsisten mendapatkan nilai tinggi di hampir semua model.
4. *Random forest* menunjukkan hasil yang kompetitif dan konsisten, membuatnya menjadi opsi yang baik, terutama dalam situasi praktik klinis yang memerlukan model yang dapat diandalkan dengan tingkat pemahaman yang baik.
5. Penerapan *machine learning* sebagai sistem pendukung keputusan terbukti memiliki potensi untuk meningkatkan akurasi diagnosis penyakit jantung, serta berkontribusi pada proses pengambilan keputusan medis yang lebih berbasis data dan objektif.

Penelitian selanjutnya disarankan untuk menggunakan teknik augmentasi data seperti *SMOTE* atau *SMOTEEN* untuk mengatasi ketidakseimbangan kelas, serta mempertimbangkan penggunaan model *deep learning* atau teknik ensemble lainnya untuk meningkatkan performa prediksi. Selain itu, pengujian pada dataset yang lebih besar dan heterogen dapat memperkuat generalisasi model pada populasi yang lebih luas.

DAFTAR PUSTAKA

- [1] World Health Organization (WHO), "Cardiovascular diseases," World Health Organization (WHO). Accessed: May 22, 2025. [Online]. Available: https://www.who.int/health-topics/cardiovascular-diseases#tab=tab_1
- [2] M. Sani, "Menkes: Stroke dan Jantung Jadi Penyebab 1,5 Juta Kematian per Tahun di Indonesia," Marwah Kepri.com. Accessed: May 22, 2025. [Online]. Available: <https://marwahkepri.com/2025/01/17/menkes-stroke-dan-jantung-jadi-penyebab-15-juta-kematian-per-tahun-di-indonesia/>
- [3] R. Antika, A. Rifa, F. Dikananda, D. Indriya Efendi, and R. Narasati, "Penerapan Algoritma Decision Tree Berbasis Pohon Keputusan Dalam Klasifikasi Penyakit Jantung," *Jurnal Mahasiswa Teknik Informatika*, vol. 7, no. 6, 2023, doi: <https://doi.org/10.36040/jati.v7i6.8264>.
- [4] D. Fabiyanto and Z. Pratama Putra, "Validasi Efektivitas Logistic Regression untuk Diagnosa Penyakit Jantung melalui Pendekatan Machine Learning," *Jurnal Ilmiah FIFO*, vol. 16, no. 2, p. 158, Nov. 2024, doi: 10.22441/fifo.2024.v16i2.006.
- [5] A. Dewi, A. Safira Surya, and Y. Yamasari, "Penerapan Algoritma Naïve Bayes (NB) untuk Klasifikasi Penyakit Jantung," *Journal of Informatics and Computer Science*, vol. 5, No. 3, pp. 447-445, 2024, doi: 10.26740/jinacs.v5n3.447.
- [6] S. N. Nuraini Arif, A. Mutoi Siregar, S. Faisal, and A. Ratna Juwita, "Klasifikasi Penyakit Serangan Jantung Menggunakan Metode Machine Learning K-Nearest Neighbors (KNN) dan Support Vector Machine (SVM)," *Jurnal Media Informatika Budidarma*, vol. 8, no. 3, pp. 1617–1626, Jul. 2024, doi: 10.30865/mib.v8i3.7844.
- [7] A. Carolina Wibowo, S. Ardi Lestari, and Nurchim, "Analisis Penggunaan Machine Learning Dalam Klasifikasi Penentuan Penyakit Jantung," *Jurnal Sistem Informasi dan Teknik Komputer*, vol. 9, no. 2, p. 395, 2024, doi: <https://doi.org/10.51876/simtek.v9i2.395>.
- [8] M. Napiah and S. Heristian, "Perbandingan Algoritma Machine Learning pada Klasifikasi Penyakit Jantung," *Jurnal Infortech*, vol. 6, no. 1, p. 21888, Jun. 2024, doi: 10.31294/infortech.v6i1.21888.
- [9] Ratnasari, A. Jurnaidi Wahidin, A. Eko Setiawan, and P. Bintoro, "Machine Learning Untuk Klasifikasi Penyakit Jantung," *Aisyah Journal of Informatics and Electrical Engineering*, vol. 6, no. 1, 2024, doi: <https://doi.org/10.30604/jti.v6i1.272>.
- [10] S. Muhammad Abdullah, "Heart Disease Dataset | Cleaned," Kaggle. Accessed: May 22, 2025. [Online]. Available: <https://www.kaggle.com/datasets/abdmental01/heart-disease-dataset>
- [11] M. Anita, I. G. Dwi Yulianti, and S. Varestama Pasaribu, "Projectrti," 2025. Accessed: Jun. 11, 2025. [Online]. Available: <https://github.com/Nicholaslorenzo/projectrti.git>
- [12] A. Rahmada and E. Redy Susanto, "Peningkatan Akurasi Prediksi Penyakit Jantung dengan Teknik SMOTEENN pada Algoritma Random Forest," *Jurnal Pendidikan dan Teknologi Indonesia (JPTI)*, vol. 4, no. 12, pp. 795–803, Dec. 2024, doi: <https://doi.org/10.52436/1.jpti.524>.
- [13] D. Setiawan, A. Muhammad, and S. H. Fransiska Dewi, "Penerapan Algoritma Klasifikasi untuk Deteksi Dini Penyakit Jantung Koroner Berdasarkan Gejala Klinis," *Teknik: Jurnal Ilmu Teknik dan Informatika*, vol. 5, no. 1, pp. 18–26, May 2025, doi: 10.30605/onoma.v9i2.2961.
- [14] I. Sabily Karima, "Penerapan Machine Learning untuk memprediksi Resiko Pengidap Penyakit Jantung menggunakan Algoritma decision tree," *Joglo, Kec. Kembangan*, vol. 14(1), no. 1, pp. 73–81, 2025, doi: <http://dx.doi.org/10.22441/format.2025.v14.i1.007>.

