Institut Riset dan Publikasi Indonesia (IRPI)
MALCOM: Indonesian Journal of Machine Learning and Computer Science
Journal Homepage: https://journal.irpi.or.id/index.php/malcom
Vol. 5 Iss. 4 October 2025, pp: 1181-1189
ISSN(P): 2797-2313 | ISSN(E): 2775-8575

Implementation of Hyperparameter Tuning for Classification Models in Heart Disease Risk Prediction

Penerapan *Hyperparameter Tuning* pada Model Klasifikasi untuk Prediksi Risiko Penyakit Jantung

Siti Kania Nur Alya Putri[1*], Indah Jumiatin[2], Indri Sulistia[3], Novan Alkaf Bahraini Saputra[4] , Nuruddin Wiranda[5]

[1,2,3,4,5]Program Studi Pendidikan Komputer, Universitas Lambung Mangkurat, Indonesia

E-Mail: [1]stkaniaaa18@gmail.com, [2]indahjumiatin7@gmail.com, [3]indrisulistia.kps@gmail.com, [4]novan.saputra@ulm.ac.id, [5]nuruddin.wd@ulm.ac.id

Received Jun 26th 2025; Revised Aug 22th 2025; Accepted Seo 02nd 2025; Available Online Oct 30th 2025
Corresponding Author: Siti Kania Nur Alya Putri

Abstract

Heart disease is the leading cause of death worldwide, making early detection crucial to reducing mortality rates. The main challenge in this study is improving the effectiveness of classification models in detecting heart disease. The objective of this study is to compare the performance of several classification algorithms and evaluate the impact of hyperparameter tuning on prediction accuracy. The methods used include applying Logistic Regression, Decision Tree, Support Vector Machine (SVM), and K-Nearest Neighbor (K-NN) algorithms to the Cleveland Clinic Heart Disease dataset obtained from Kaggle. Hyperparameter tuning was performed using gridsearchCV and randomizedsearchCV combined with cross-validation. The results showed that after tuning, logistic regression, K-NN, and SVM achieved the highest accuracy of 84%, while the decision tree model recorded the lowest accuracy of 80%. Additionally, precision, recall, and F1-score also improved, particularly for logistic regression and K-NN, which produced the most balanced results. These findings demonstrate that hyperparameter tuning is highly effective in enhancing model performance and support the use of machine learning in the early detection of heart disease.

Keywords: Classification Algorithms, Early Detection, Heart Disease, Hyperparameter Tuning, Machine Learning

Abstrak

Penyakit jantung adalah penyebab utama kematian di seluruh dunia, sehingga penting untuk melakukan deteksi dini secara tepat untuk menurunkan angka kematian. Tantangan utama dalam penelitian ini adalah bagaimana cara meningkatkan efektivitas model klasifikasi dalam mendeteksi penyakit jantung. Tujuan dari penelitian ini adalah untuk membandingkan kinerja beberapa algoritma klasifikasi dan menilai dampak *hyperparameter tuning* terhadap peningkatan akurasi prediksi. Metode yang digunakan mencakup penerapan algoritma *Logistic Regression*, *Decision Tree*, *Support Vector Machine* (SVM), dan *K-Nearest Neighbor* (K-NN) pada *dataset Cleveland Clinic Heart Disease* yang diambil dari Kaggle. Proses *hyperparameter tuning* dilaksanakan dengan menggunakan *gridsearchCV* dan *randomizedsearchCV* bersama dengan cross-validation. Temuan penelitian menunjukkan bahwa setelah dilakukan *tuning*, *logistic regression*, K-NN, dan SVM mencapai akurasi tertinggi yang sama, yaitu 84%. *Decision tree* berada di posisi terendah dengan akurasi 80%. Selain itu, nilai *precision*, *recall*, dan *F1-score* juga meningkat, terutama pada *logistic regression* dan K-NN yang menunjukkan hasil paling seimbang. Hasil ini membuktikan bahwa *hyperparameter tuning* sangat membantu dalam meningkatkan kinerja model klasifikasi dan mendukung penggunaan machine learning untuk deteksi dini penyakit jantung secara lebih efektif.

Kata Kunci: Algoritma Klasifikasi, Deteksi Dini, *Hyperparameter Tuning*, Machine Learning, Penyakit Jantung

1. PENDAHULUAN

Penyakit jantung merupakan penyebab utama kematian di seluruh dunia. Menurut data dari *World Health Organization* (WHO), pada tahun 2019 diperkirakan sebanyak 17,9 juta orang meninggal akibat penyakit ini, yang setara dengan 32% dari seluruh kematian global [17]. Dari jumlah tersebut, sekitar 85% kematian disebabkan oleh serangan jantung dan stroke. Menariknya, lebih dari tiga perempat kematian akibat

penyakit jantung terjadi di negara-negara dengan pendapatan rendah dan menengah. Selain itu, dari 17 juta kematian dini yang terjadi pada orang di bawah usia 70 tahun akibat penyakit tidak menular pada tahun yang sama, 38% di antaranya disebabkan oleh penyakit jantung [2], [15], [17]. Di Indonesia, penyakit jantung juga merupakan ancaman serius. Berdasarkan Riset Kesehatan Dasar (Riskesdas) tahun 2018, prevalensi penyakit jantung koroner mencapai 1,5% dari populasi nasional [18]. Fakta-fakta tersebut menunjukkan pentingnya pengembangan sistem deteksi dini yang akurat, baik pada tingkat global maupun nasional, untuk menekan angka kematian akibat penyakit jantung.

Di sisi lain, tantangan besar dalam diagnosis penyakit jantung secara konvensional adalah ketergantungan pada pengalaman subjektif klinisi dan potensi keterlambatan identifikasi risiko pada tahap awal. Oleh karena itu, diperlukan pendekatan berbasis data yang mampu mengidentifikasi pola klinis secara sistematis dan akurat. Pembelajaran mesin (machine learning) menawarkan potensi besar dalam hal ini karena mampu menangani kompleksitas data medis dan mendeteksi hubungan non-linier antar variabel yang sulit ditangkap secara manual. Implementasi machine learning tidak hanya mempercepat proses prediksi, tetapi juga meningkatkan akurasi dan efisiensi pengambilan keputusan klinis [5], [7].

Dalam beberapa tahun terakhir, algoritma pembelajaran mesin telah diterapkan secara luas dalam pengembangan sistem klasifikasi untuk mendukung diagnosis medis. Model seperti *logistic regression*, *decision tree*, dan *support vector machine* (SVM) telah terbukti efektif dalam mengklasifikasikan kondisi kesehatan pasien berdasarkan data klinis [5], [7]. Penelitian oleh Manikandan et al. [4] menunjukkan bahwa penggabungan teknik seleksi fitur boruta dengan beberapa algoritma klasifikasi mampu menghasilkan akurasi prediksi penyakit jantung lebih dari 85%. Sementara itu, Budholiya et al. [16] menggunakan model XGBoost yang dioptimasi dan berhasil mencapai akurasi hingga 91,82%. Hasil-hasil ini menunjukkan bahwa efektivitas model klasifikasi dapat ditingkatkan secara signifikan melalui pemilihan algoritma dan proses optimasi parameter yang tepat.

Salah satu komponen penting dalam proses pengembangan model klasifikasi adalah *hyperparameter tuning*. Setiap algoritma memiliki parameter-parameter spesifik yang memengaruhi performa model secara langsung. Sebagai contoh, pada *k-nearest neighbor* (K-NN), jumlah tetangga (*k*) menentukan seberapa lokal keputusan klasifikasi dibuat. Pada *Decision Tree*, kedalaman pohon (*max depth*) dan kriteria pemisahan sangat memengaruhi kompleksitas dan akurasi model. Sedangkan pada SVM, nilai *C* dan jenis kernel yang digunakan dapat menentukan keseimbangan antara margin maksimal dan kesalahan klasifikasi. Oleh karena itu, teknik *tuning* seperti *gridsearchCV* dan *randomizedsearchCV* digunakan untuk mengeksplorasi kombinasi *hyperparameter* terbaik secara sistematis melalui validasi silang [6], [13]. Almomany et al. [3] bahkan menunjukkan bahwa penerapan *hyperparameter tuning* pada model K-NN dalam kasus medis mampu meningkatkan akurasi klasifikasi hingga lebih dari 87%.

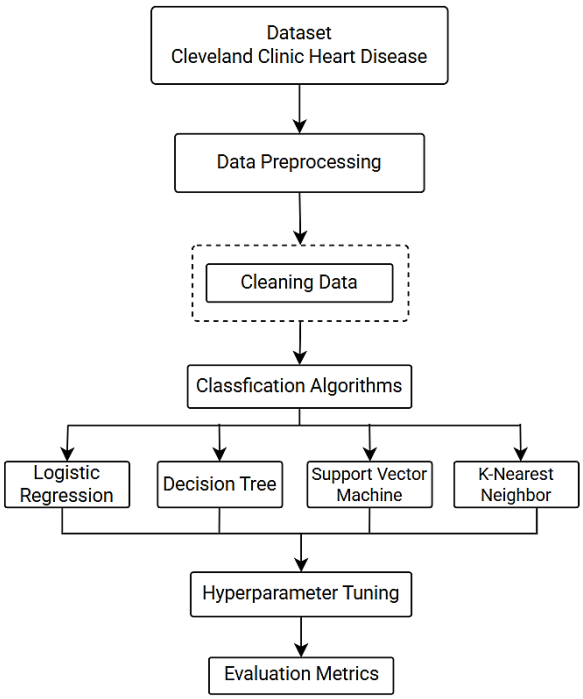
Namun, meskipun beberapa penelitian telah membuktikan efektivitas tuning hyperparameter dalam meningkatkan kinerja model, sebagian besar penelitian sebelumnya lebih menekankan pada pemilihan algoritma terbaik atau penerapan metode ensemble untuk mencapai akurasi tinggi. Hal ini menyebabkan masih terbatasnya analisis mendalam mengenai bagaimana penyesuaian *hyperparameter* memengaruhi performa algoritma tradisional seperti *logistic regression*, decision tree, SVM, dan K-NN. Berdasarkan latar belakang tersebut, penelitian ini bertujuan untuk mengevaluasi dampak *hyperparameter tuning* terhadap performa empat algoritma klasifikasi, yaitu *logistic regression*, *decision tree*, SVM, dan K-NN dalam memprediksi risiko penyakit jantung. Penelitian ini menekankan pentingnya pemilihan strategi *tuning* yang tepat untuk meningkatkan akurasi model secara konsisten. Diharapkan hasil penelitian ini dapat berkontribusi pada pengembangan sistem klasifikasi medis yang lebih presisi dan adaptif, khususnya dalam upaya deteksi dini penyakit jantung secara berbasis data.

2. METODE PENELITIAN

Penelitian ini bersifat eksperimental dengan pendekatan kuantitatif, berfokus pada penerapan dan optimasi algoritma klasifikasi menggunakan teknik *hyperparameter tuning*. Tujuan utamanya adalah membandingkan performa model-model klasifikasi dalam memprediksi risiko penyakit jantung berdasarkan *dataset* terbuka dari *Cleveland Clinic Heart Disease*. Seluruh proses dilakukan menggunakan lingkungan pemrograman Python dengan dukungan pustaka *scikit-learn*. Adapun metodologi dari penelitian dapat dilihat pada Gambar 1.

2.1 *Dataset*

Dataset adalah kumpulan data terstruktur yang berisi informasi klinis dari pasien, disusun dalam format tabel dan siap untuk dianalisis. Dalam penelitian ini, *dataset* yang digunakan adalah *Cleveland Clinic Heart Disease* [19], yang terdiri dari 303 data pasien dan mencakup 14 fitur klinis relevan. Fitur-fitur ini meliputi data demografis, parameter fisiologis, serta hasil laboratorium seperti usia, jenis kelamin, tekanan darah istirahat, kadar kolesterol, detak jantung maksimal, serta data hasil pemeriksaan elektrokardiogram. Keseluruhan informasi ini telah diakui sebagai faktor risiko utama dalam berbagai penelitian medis dan sangat berperan dalam membangun model prediksi penyakit jantung.



Gambar 1. Metodologi Penelitian

2.2 *Data Preprocessing*

Proses *data preprocessing* merupakan tahapan awal yang dilakukan sebelum pelatihan model dimulai. Tujuan utama dari proses ini adalah untuk memastikan bahwa data bersih, konsisten, dan berada dalam format yang sesuai untuk analisis lebih lanjut. Langkah awal yang dilakukan adalah menangani data tidak valid dan data kosong. Selanjutnya, target variabel dikonversi ke format biner (0 = tidak menderita penyakit jantung, 1 = menderita penyakit jantung) agar sesuai untuk tugas klasifikasi biner.

2.3. *Data Cleaning*

Pada tahap *cleaning data*, dilakukan identifikasi terhadap nilai-nilai kosong atau tidak valid pada setiap fitur. Data yang tidak lengkap atau menyimpang dari struktur standar dikecualikan dari proses analisis untuk menjaga integritas *dataset*. Langkah ini penting karena kehadiran nilai hilang atau tidak relevan dapat menurunkan akurasi model. Setiap baris data diperiksa untuk memastikan semua fitur memiliki panjang nilai yang sesuai dan konsisten, sehingga tidak mengganggu proses pelatihan model selanjutnya [1]. Setelah proses pembersihan, dua baris data yang mengandung nilai hilang atau tidak valid dihapus, sehingga jumlah data yang digunakan dalam penelitian ini menjadi 301 baris.

2.4. Algoritma Klasifikasi

Penelitian ini menggunakan empat algoritma klasifikasi populer dalam pembelajaran mesin yang telah banyak digunakan pada penelitian diagnosis penyakit jantung. Keempat model tersebut adalah *logistic regression*, *decision tree*, SVM, dan K-NN. Pemilihan algoritma ini didasarkan pada performanya dalam penelitian-penelitian sebelumnya yang menunjukkan bahwa pendekatan klasifikasi berbasis data dapat memberikan hasil yang akurat dan andal dalam prediksi kondisi klinis [4], [7], [16]. Masing-masing algoritma memiliki karakteristik yang berbeda dalam menangani distribusi data, hubungan antar fitur, dan kompleksitas model. Oleh karena itu, proses *hyperparameter tuning* menjadi penting untuk mengoptimalkan kinerja masing-masing model secara adil.

Logistic regression adalah algoritma yang digunakan untuk menggambarkan hubungan antara variabel bebas dan variabel terikat yang bersifat biner, sehingga cocok untuk diagnosis penyakit jantung yang memiliki kategori “sakit” dan “tidak sakit.” Model ini memanfaatkan fungsi logit guna menghitung kemungkinan keluaran, sehingga setiap nilai input dapat diterjemahkan menjadi angka antara 0 dan 1. Dalam penelitian ini, regresi logistik dipilih karena kesederhanaan dan kemampuannya untuk menciptakan model yang dapat dengan mudah dipahami serta memberikan wawasan mengenai kontribusi tiap fitur terhadap prediksi hasil [4], [5]. Proses penyesuaian dilakukan pada parameter C (nilai regularisasi), *penalty* (L1 atau L2), dan *solver* (liblinear, saga), dengan rentang C antara 0,01 hingga 10. Rentang ini ditentukan untuk menemukan keseimbangan antara bias dan varians, nilai C yang terlalu kecil bisa menyebabkan underfitting, sementara nilai yang besar dapat mengakibatkan overfitting. Dengan demikian, penerapan regresi logistik

dalam penelitian ini lebih dari sekadar pengujian dasar, tetapi juga mencakup optimisasi untuk memastikan model berfungsi dengan baik pada data diagnosis penyakit jantung.

Decision tree adalah metode yang menggunakan struktur pohon untuk membagi *dataset* menjadi cabang-cabang berdasarkan kondisi atribut tertentu. Algoritma ini menerapkan prinsip rekursif dalam memilih atribut terbaik untuk setiap percabangan sehingga dapat menghasilkan aturan klasifikasi yang mudah dipahami. Kelebihan dari *decision tree* adalah kemampuannya menangkap hubungan non-linier di antara variabel dan aspek transparansi dalam pemahaman model [7], [16]. Namun, model ini memiliki kecenderungan untuk mengalami overfitting, terutama jika jumlah sampel dalam *dataset* terbatas seperti dalam penelitian ini. Untuk mengatasi masalah tersebut, penyesuaian dilakukan pada beberapa hyperparameter penting, yaitu max_depth (kedalaman pohon), min_samples_split (jumlah minimum sampel untuk melakukan pembagian), min_samples_leaf (jumlah minimum sampel di bagian ujung pohon), criterion (gini atau entropi), serta max_features (jumlah maksimum fitur yang terpakai). Variasi pada max_depth dan min_samples_split digunakan untuk mengendalikan kompleksitas model dan meningkatkan kemampuan generalisasi. Dengan pendekatan ini, diharapkan *decision tree* dapat menghasilkan model yang tidak hanya tepat tetapi juga stabil untuk digunakan dalam memprediksi penyakit jantung.

SVM beroperasi dengan prinsip menemukan *hyperplane* paling optimal yang memisahkan dua kategori dalam ruang fitur dengan margin terbesar. Algoritma ini dikenal efektif untuk data dengan banyak dimensi dan sering diterapkan dalam berbagai penelitian medis, termasuk untuk diagnosis penyakit jantung, karena kemampuannya menangani masalah non-linear melalui penggunaan kernel [5], [14]. Dalam penelitian ini, SVM dievaluasi dengan kombinasi parameter C (regularisasi), kernel (linear, polynomial, radial basis function), dan gamma (parameter kernel). Parameter C mengatur keseimbangan antara margin yang luas dan kesalahan klasifikasi, sementara gamma mengindikasikan seberapa besar pengaruh suatu data tunggal terhadap pemisahan hyperplane. Berbagai kernel diuji untuk menentukan apakah data lebih baik dipisahkan dengan cara linier atau non-linier. Penyesuaian dilakukan untuk mencari konfigurasi parameter yang paling efektif agar SVM dapat memberikan hasil generalisasi yang lebih baik. Pendekatan ini menjadi penting mengingat karakteristik data medis sering kali rumit dan melibatkan interaksi antar fitur yang tidak linier.

K-NN adalah sebuah algoritma non-parametrik yang beroperasi dengan cara mencari k tetangga paling dekat dari data baru menggunakan ukuran jarak, dan kemudian mengklasifikasikan data itu berdasarkan mayoritas kelas dari tetangga yang ditemukan. K-NN dipilih untuk penelitian ini karena kesederhanaannya serta kemampuannya dalam mendeteksi pola lokal pada *dataset* yang memiliki ukuran kecil hingga menengah, contohnya data tentang penyakit jantung [11], [12]. Namun, kinerja K-NN sangat dipengaruhi oleh pemilihan nilai k (jumlah tetangga), jenis bobot (uniform atau distance), serta jenis ukuran jarak (Euclidean atau Manhattan). Karena itu, penelitian ini melakukan penyesuaian dengan mencoba nilai k = 3, 5, 7, dan 9, serta membandingkan jenis bobot uniform dan distance. Pemilihan ukuran jarak juga dianalisis untuk menentukan apakah distribusi data lebih tepat diukur menggunakan jarak Euclidean atau Manhattan. Dengan demikian, pemanfaatan K-NN dalam penelitian ini tidak hanya memberikan dasar yang sederhana, tetapi juga dieksplorasi secara menyeluruh sehingga hasilnya dapat bersaing dengan algoritma lainnya dalam mendeteksi penyakit jantung.

2.5. *Hyperparameter tuning*

Hyperparameter tuning merupakan proses krusial dalam pembelajaran mesin untuk mengoptimalkan kinerja model. Tidak seperti parameter internal yang dipelajari selama proses pelatihan, *hyperparameter* ditentukan sebelum pelatihan dan berperan langsung dalam mengontrol kompleksitas, bias, dan kapasitas generalisasi dari model yang digunakan. Pemilihan nilai *hyperparameter* yang tidak tepat dapat menyebabkan model mengalami overfitting atau underfitting, yang berdampak pada rendahnya akurasi terhadap data baru atau data yang belum pernah dilihat [6], [8], [9]. Dalam penelitian ini, digunakan dua metode *tuning hyperparameter* yang umum dan banyak diterapkan, yaitu *gridsearchCV* dan *randomizedsearchCV*, yang keduanya tersedia dalam pustaka *scikit-learn*.

1. *ridsearchCV* bekerja dengan cara mengevaluasi seluruh kombinasi parameter dalam ruang pencarian yang telah ditentukan secara eksplisit. Meskipun metode ini menjamin pencarian solusi terbaik dalam ruang parameter tersebut, pendekatan ini memiliki keterbatasan pada waktu komputasi yang tinggi, terutama jika ruang parameter sangat besar [6], [14].
2. *randomizedsearchCV*, di sisi lain, mengevaluasi kombinasi parameter yang dipilih secara acak dari ruang pencarian. Pendekatan ini terbukti lebih efisien secara waktu dan sering kali menghasilkan konfigurasi parameter yang kompetitif, bahkan dalam kasus ruang parameter yang luas dan tidak seragam [6], [14].

Tujuan utama dari *hyperparameter tuning* adalah untuk meningkatkan performa model, baik dari segi akurasi, presisi, *recall*, *F1-score*, maupun daya generalisasi terhadap data baru. Seperti yang ditunjukkan dalam penelitian Putri et al. [6], penerapan *randomizedsearchCV* secara efektif meningkatkan sensitivitas model dalam prediksi kasus medis. Selain itu, penelitian oleh Purnomo et al. [8] dan Firgiawan et al. [9] juga

menegaskan bahwa proses *tuning* yang optimal mampu menghasilkan peningkatan performa model klasifikasi secara signifikan, khususnya dalam konteks medis yang sensitif terhadap kesalahan klasifikasi.

Dalam penelitian ini, *tuning* dilakukan pada masing-masing algoritma klasifikasi dengan ruang pencarian yang dirancang berdasarkan karakteristik struktural tiap algoritma. Tabel berikut merangkum *hyperparameter* yang di tuning untuk setiap model:

Tabel 1. Model dan *Hyperparameter*

Model	*Hyperparameter* yang Digunakan
Logistic Regression	C, penalty, solver
Decision Tree	max_depth, min_samples_split, min_samples_leaf, criterion, max_features
SVM	C, kernel, gamma
K-NN	n_neighbors, weights, metric

2.6. Evaluasi Model

Kinerja model klasifikasi dievaluasi menggunakan beberapa metrik utama, yaitu *accuracy*, *precision*, *recall*, dan *F1-score*. Metrik-metrik ini dihitung berdasarkan data uji dan memberikan gambaran menyeluruh mengenai kemampuan model dalam memprediksi secara akurat. Tujuan dari evaluasi ini adalah untuk mengetahui seberapa baik model dapat memprediksi risiko penyakit jantung.

1. *Accuracy* adalah proporsi prediksi yang benar dibandingkan dengan total prediksi.
2. *Precision* mengukur proporsi hasil positif yang benar dari seluruh hasil positif yang diprediksi.
3. *Recall* (sensitivitas) mengukur proporsi hasil positif yang benar dari semua kasus positif yang sebenarnya.
4. *F1-Score* adalah rata-rata harmonik dari *Precision* dan *Recall*, memberikan keseimbangan antara keduanya.

Rumus perhitungan metrik evaluasi adalah sebagai berikut:

$$Accuracy = \frac{TP + TN}{TP+TN+FP+FN} \quad (1)$$

$$Recall = \frac{TP}{TP+FN} \quad (2)$$

$$Precission = \frac{TP+TN}{TP+FN} \quad (3)$$

$$F1\,Score = 2 \times \frac{recall \times precision}{recall + precision} \quad (4)$$

True Positives (TP) merupakan Jumlah prediksi benar untuk kelas tertentu, *True Negatives* (TN) merupakan Jumlah prediksi benar bukan untuk kelas tertentu (tidak termasuk TP kelas itu), *False Positives* (FP)merupakan Jumlah prediksi salah yang menunjukkan anggota kelas tertentu (tidak termasuk TP kelas itu),dan *False Negatives* (FN) merupakan Jumlah prediksi salah yang menunjukkan bukan anggota kelas tertentu (tidak termasuk baris dan kolom dari kelas tersebut

3. HASIL DAN PEMBAHASAN

3.1 Pengumpulan Data

Pada tahap ini, *dataset* diperoleh dari situs web Kaggle, yang berisi data pasien yang berkaitan dengan diagnosis penyakit jantung. *Dataset* ini awalnya terdiri dari 303 baris data, namun setelah proses pembersihan data dilakukan untuk menghapus nilai yang tidak valid, jumlah data akhir yang dianalisis adalah 301 baris dan mencakup 14 variabel yang tercatat dalam satu tabel. Variabel-variabel tersebut meliputi informasi demografis dan medis pasien, antara lain: *age* (usia), *sex* (jenis kelamin), cp (jenis nyeri dada), *trestbps* (tekanan darah istirahat), *cho*l (kadar kolesterol), *fbs* (gula darah puasa), *restecg* (hasil elektrokardiogram istirahat), *thalach* (detak jantung maksimum), *exang* (angina karena olahraga), *oldpeak* (depresi ST), *slope* (kemiringan segmen ST), *ca* (jumlah pembuluh darah utama), *thal* (kondisi thalassemia), dan target (indikasi penyakit jantung, dengan 1 menunjukkan positif penyakit jantung, dan 0 menunjukkan negatif). *Dataset* prediksi penyakit jantung dapat dilihat pada Tabel 2.

3.2 Akurasi Model Tanpa Optimasi

Evaluasi awal terhadap model klasifikasi tanpa penerapan *tuning hyperparameter* menunjukkan bahwa *logistic regression*, SVM, dan K-NN memiliki performa yang relatif setara, dengan nilai akurasi tertinggi sebesar 0,82. *Logistic regression* mencatat nilai *F1-score* tertinggi yaitu 0,83, yang mencerminkan keseimbangan yang baik antara *precision* (0,82) dan *recall* (0,79). K-NN juga menunjukkan performa stabil

dengan *precision* 0,81 dan *recall* 0,80, menghasilkan *F1-score* 0,80. Sementara itu, meskipun akurasi SVM setara (0,82), model ini mencatat *recall* terendah di antara ketiganya (0,77), dengan *F1-score* 0,80. Di sisi lain, *decision tree* memiliki akurasi terendah (0,74) dan *precision* paling rendah (0,70), meskipun *recall*-nya relatif tinggi (0,77), yang menunjukkan kecenderungan model untuk menghasilkan lebih banyak prediksi positif yang salah. Secara keseluruhan, hasil ini mengindikasikan bahwa *logistic regression* memberikan performa awal yang paling seimbang dan kuat, sedangkan *decision tree* cenderung kurang stabil tanpa proses optimasi parameter. Akurasi model tanpa optimasi tersebut dapat dilihat pada Tabel 3.

Tabel 2. *Dataset* Prediksi Penyakit Jantung

Variabel											
No	1	2	3	4	5	...	297	298	299	300	301
Age	63	67	67	37	41	...	57	45	68	57	57
Sex	1	1	1	1	0	...	0	1	1	1	0
Cp	1	4	4	3	2	...	4	1	4	4	2
Trestbps	145	160	120	130	130	...	140	110	144	130	130
Chol	233	286	229	250	204	...	241	264	193	131	236
Fbs	1	0	0	0	0	...	0	0	1	0	0
Restecg	2	2	2	0	2	...	0	0	0	0	2
Thalach	150	108	129	187	172	...	123	132	141	115	174
Exang	0	1	1	0	0	...	1	0	0	1	0
Oldpeak	2.3	1.5	2.6	3.5	1.4	...	0.2	1.2	3.4	1.2	0.0
Slope	3	2	2	3	1	…	2	2	2	2	2
Ca	0.0	3.0	2.0	0.0	0.0	…	0.0	0.0	2.0	1.0	1.0
Thal	6.0	3.0	7.0	3.0	3.0	…	7.0	7.0	7.0	7.0	3.0
Target	0	1	1	0	0	…	1	1	1	1	1

Tabel 3. Akurasi Model Tanpa Optimasi

Model	*Accuracy*	*Precision*	*Recall*	*F1-Score*	*Support*
Logistic Regression	0.82	0.82	0.79	0.83	137
Decision Tree	0.74	0.70	0.77	0.73	137
SVM	0.82	0.82	0.77	0.80	137
K-NN	0.82	0.81	0.80	0.80	137

3.3 Performa Model Setelah *Hyperparameter tuning*

Penerapan *hyperparameter tuning* menggunakan pendekatan *grid search* menunjukkan adanya peningkatan performa pada semua algoritma klasifikasi. SVM dan K-NN sama-sama mencatat akurasi tertinggi sebesar 0,84. K-NN menunjukkan *precision* tertinggi (0,89), meskipun *recall*-nya relatif lebih rendah (0,75), menghasilkan *F1-score* 0,81. SVM memiliki performa yang lebih seimbang dengan *precision* 0,85 dan *recall* 0,77. Sementara itu, *logistic regression* mempertahankan performa yang stabil dengan akurasi 0,83, *precision* 0,84, dan *F1-score* 0,81. Di sisi lain, meskipun terjadi peningkatan pada *decision tree* dibandingkan sebelum *tuning*, performanya masih paling rendah di antara model lain dengan akurasi 0,78 dan *F1-score* 0,76. Secara keseluruhan, *tuning* dengan *grid search* memberikan dampak positif terhadap semua model, khususnya dalam meningkatkan presisi dan akurasi klasifikasi awal.

Hyperparameter tuning menggunakan *randomized search* memberikan hasil yang sedikit berbeda. *Logistic regression* mencatat peningkatan paling signifikan dengan akurasi mencapai 0,84, *precision* 0,86, dan *recall* 0,80, menghasilkan *F1-score* sebesar 0,82. Model ini menunjukkan keseimbangan kinerja terbaik di antara semua model. K-NN tetap konsisten dengan akurasi 0,84, *precision* 0,85, dan *recall* 0,79, yang menunjukkan bahwa metode ini cukup stabil terhadap variasi parameter. SVM, meskipun mengalami sedikit penurunan akurasi menjadi 0,82 dibandingkan hasil dari *grid search*, masih menunjukkan performa yang kompetitif dengan *F1-score* 0,80. Sebaliknya, *decision tree* mencatat kenaikan akurasi menjadi 0,80, namun metrik lainnya masih tertinggal dibandingkan model lain. Dengan demikian, *tuning* menggunakan *randomized search* secara umum menghasilkan peningkatan metrik yang lebih konsisten pada *logistic regression* dan K-NN, sementara SVM dan *decision tree* menunjukkan hasil yang cenderung naik turun dan kurang stabil. Tabel 4 merupakan hasil Akurasi model dengan *hyperparameter tuning*, *gridsearchCV* dan *randomizedsearchCV*.

Tabel 4. Akurasi Model dengan *Hyperparameter Tuning*, *Gridsearchcv* dan *Randomizedsearchcv*

Model	*GridSearchCV*					*RandomizedSearchCV*				
	Accuracy	*Precision*	*Recall*	*F1-Score*	*Support*	*Accuracy*	*Precision*	*Recall*	*F1-Score*	*Support*
Logistic Regression	0.83	0.84	0.79	0.81	137	0.84	0.86	0.80	0.82	137

Model	*GridSearchCV*					*RandomizedSearchCV*				
	Accuracy	*Precision*	*Recall*	*F1-Score*	*Support*	*Accuracy*	*Precision*	*Recall*	*F1-Score*	*Support*
Decision Tree	0.78	0.77	0.76	0.76	137	0.80	0.82	0.74	0.78	137
SVM	0.84	0.85	0.77	0.81	137	0.82	0.83	0.77	0.80	137
K-NN	0.84	0.89	0.75	0.81	137	0.84	0.85	0.79	0.82	137

3.4 Pembahasan

Penerapan *tuning hyperparameter* memberikan dampak positif terhadap peningkatan kinerja model klasifikasi dalam mengidentifikasi risiko penyakit jantung. Sebelum tuning dilakukan, model *logistic regression*, SVM, dan K-NN menunjukkan akurasi awal yang serupa, yaitu 0,82. Logistic regression mencapai nilai *F1-score* tertinggi di angka 0,83 dengan *precision* sebesar 0,82 dan *recall* 0,79, yang menunjukkan kinerja yang seimbang dalam mengklasifikasikan pasien yang terdeteksi memiliki atau tidak memiliki penyakit jantung. Temuan ini mengindikasikan bahwa meski sederhana, model ini cukup andal dalam memberikan prediksi awal yang konsisten. K-NN dan SVM juga memperoleh *F1-score* sebesar 0,80, dengan perbedaan kecil dalam *recall*, yang menunjukkan bahwa keduanya cukup stabil meskipun masih belum optimal dalam mendeteksi semua kasus positif. Di sisi lain, *decision tree* menunjukkan kinerja yang paling rendah dengan akurasi 0,74 dan *precision* hanya 0,70, walaupun *recall*-nya terbilang baik (0,77). Kondisi ini menunjukkan bahwa *decision tree* cenderung menghasilkan prediksi positif yang berlebihan dan kurang ideal untuk mendukung diagnosis klinis secara langsung.

Setelah tuning dilakukan menggunakan grid search, terlihat peningkatan kinerja di semua model. SVM, K-NN, dan *logistic regression* mencapai akurasi yang sama, yaitu 0,84, namun dengan karakteristik kinerja yang bervariasi. K-NN memperoleh *precision* tertinggi (0,89), yang menunjukkan kemampuannya yang sangat baik dalam mengurangi *false positive*, meskipun dengan iyang lebih rendah (0,75), sehingga masih ada pasien yang memiliki penyakit jantung yang terlewat. Sebaliknya, SVM menunjukkan keseimbangan yang lebih baik antara *precision* dan *recall* (0,85 dan 0,77), menjadikannya andal dalam mendeteksi kedua kategori tersebut. *Logistic regression* mempertahankan performanya dengan *precision* 0,84 dan *F1-score* 0,81, menunjukkan bahwa model ini stabil dan mudah beradaptasi terhadap proses tuning. Di sisi lain, *decision tree* mengalami peningkatan akurasi menjadi 0,78, naik dari 0,74, tetapi tetap tertinggal dibandingkan model lainnya. Temuan ini menegaskan bahwa meskipun mudah dipahami, *decision tree* belum cukup stabil untuk digunakan dalam diagnosis penyakit jantung. Dengan demikian, *grid search* terbukti memberikan kontribusi yang signifikan dalam meningkatkan akurasi dan kestabilan model, terutama pada SVM dan K-NN, yang semakin relevan untuk membantu dalam proses deteksi dini penyakit jantung.

Sementara itu, tuning yang dilakukan menggunakan *randomized search* menunjukkan peningkatan yang paling konsisten pada *logistic regression*, yang mencatat akurasi 0,84 dengan *precision* 0,86 dan *recall* 0,80, menghasilkan *F1-score* 0,82. Model ini menunjukkan keseimbangan terbaik di antara semua model, sehingga dinilai paling cocok untuk tujuan penelitian, yaitu mendeteksi risiko penyakit jantung dengan akurat dan seimbang. K-NN juga mempertahankan kinerjanya yang stabil dengan akurasi 0,84, *precision* 0,85, dan *recall* 0,79, semakin memperkuat posisinya sebagai model sederhana namun efektif dalam klasifikasi medis. Sementara itu, SVM, meskipun mengalami sedikit penurunan dari hasil *grid search* (akurasi 0,82), masih menunjukkan kinerja yang kompetitif dengan *F1-score* 0,80. Sebaliknya, *decision tree* hanya sedikit mengalami peningkatan akurasi menjadi 0,80 dibandingkan hasil dari *grid search*, tetapi *F1-score*-nya tetap lebih rendah dibandingkan model lain, yang menunjukkan kurangnya konsistensi dari algoritma ini dalam konteks penelitian tersebut.

Secara keseluruhan, hasil dari penilaian menunjukkan bahwa penyesuaian hyperparameter dapat memperbaiki mutu klasifikasi dari segi *accuracy*, *precision*, dan *recall*. *Logistic regression* dan K-NN muncul sebagai model yang paling konsisten setelah penyesuaian, dengan performa yang bersaing pada kedua metode tersebut. SVM juga mempertahankan performa yang baik, terutama setelah pencarian grid, meskipun ada sedikit penurunan pada pencarian acak. Di sisi lain, *decision tree* menunjukkan performa yang fluktuatif tergantung pada metode penyesuaian yang diterapkan. Temuan ini menjawab pertanyaan penelitian mengenai perlunya pendekatan penyesuaian hyperparameter yang tepat untuk meningkatkan kinerja model dalam identifikasi dini penyakit jantung. Dengan demikian, algoritma yang paling memenuhi tujuan penelitian adalah *logistic regression* dan K-NN karena keduanya menawarkan keseimbangan terbaik antara *accuracy*, *precision*, dan *recall*, sehingga lebih cocok untuk diterapkan dalam konteks medis. Visualisasi perbandingan dari akurasi model dapat dilihat pada Gambar 2.

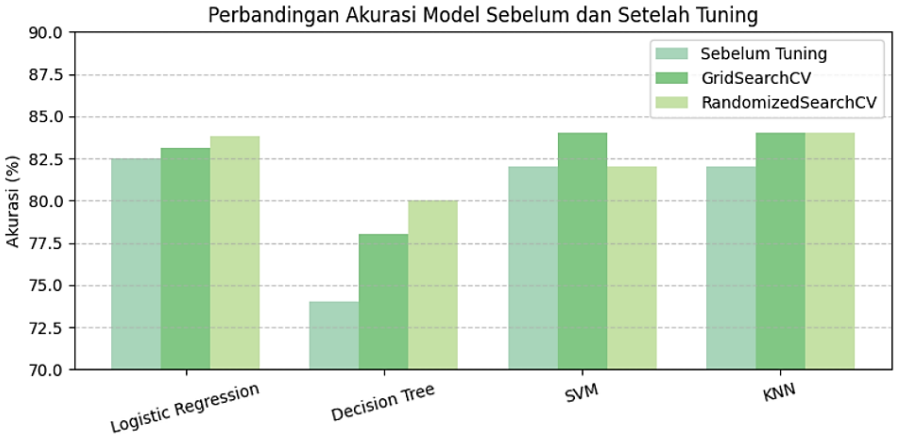
4. KESIMPULAN

Penelitian ini bertujuan untuk mengevaluasi dampak hyperparameter tuning terhadap performa empat algoritma klasifikasi, yaitu logistic regression, decision tree, SVM, dan K-NN, dalam memprediksi risiko

penyakit jantung. Hasil eksperimen menunjukkan bahwa tuning hyperparameter berperan penting dalam meningkatkan akurasi dan konsistensi performa model.

SVM menonjol sebagai algoritma dengan performa terbaik, mencapai akurasi hingga 84% dengan keseimbangan precision dan recall yang tinggi, sehingga sangat potensial untuk menangani klasifikasi data medis yang kompleks. Logistic regression mengalami peningkatan berarti pada aspek recall, yang krusial untuk deteksi dini kasus positif penyakit jantung. K-NN memperlihatkan performa yang relatif stabil pada berbagai skema tuning, menjadikannya opsi yang sesuai untuk *dataset* berukuran kecil. Sementara itu, decision tree tetap kuat dalam mengidentifikasi kasus positif, namun hasilnya cenderung fluktuatif dan sangat dipengaruhi oleh pemilihan parameter yang tepat.

Secara keseluruhan, penelitian ini menegaskan bahwa pemilihan strategi hyperparameter tuning yang tepat dapat meningkatkan akurasi dan stabilitas model secara konsisten, sesuai dengan tujuan penelitian. Temuan ini berkontribusi pada pengembangan sistem klasifikasi medis yang lebih presisi dan adaptif, khususnya dalam mendukung deteksi dini penyakit jantung berbasis data. Untuk penelitian selanjutnya, disarankan eksplorasi pendekatan ensemble, pemanfaatan *dataset* yang lebih besar dan bervariasi, serta integrasi data real-time untuk memperkuat kapabilitas prediksi dalam konteks dunia nyata.



Gambar 2. Perbandingan Akurasi Model

REFERENSI

[1] N. Nurdiansyah, F. S. Febriyan, Z. G. D. Amanta, D. A. Saputra, and W. M. Baihaqi, "Analisis Kesehatan Mental untuk Mencegah Gangguan Mental pada Mahasiswa Menggunakan Algoritma K-Nearest Neighbor (K-NN) dan Random Forest," *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 5, no. 1, pp. 1–9, Nov. 2024, doi: 10.57152/malcom.v5i1.1537.

[2] R. B. Yudhistira, M. Y. Yudhistira, and R. T. Supraptomo, "Perioperative Management in Parturient with Severe Preeclampsia, Obesity, and COVID-19," *Solo Journal of Anesthesi, Pain and Critical Care (SOJA)*, vol. 1, no. 2, p. 67, Oct. 2021, doi: 10.20961/soja.v1i2.54984.

[3] A. Almomany, W. R. Ayyad, and A. Jarrah, "Optimized implementation of an improved KNN classification algorithm using Intel FPGA platform: Covid-19 case study," *Journal of King Saud University - Computer and Information Sciences*, vol. 34, no. 6, pp. 3815–3827, Jun. 2022, doi: 10.1016/j.jksuci.2022.04.006.

[4] G. Manikandan, B. Pragadeesh, V. Manojkumar, A. L. Karthikeyan, R. Roselin, and A. H. Gandomi, "Classification models combined with Boruta feature selection for heart disease prediction," *Inform Med Unlocked*, vol. 44, Jan. 2024, doi: 10.1016/j.imu.2023.101442.

[5] R. Roscher, B. Bohn, M. F. Duarte, and J. Garcke, "Explainable Machine Learning for Scientific Insights and Discoveries," *IEEE Access*, vol. 8, pp. 42200–42216, 2020, doi: 10.1109/ACCESS.2020.2976199.

[6] T. A. E. Putri, T. Widiharih, and R. Santoso, "Penerapan Tuning Hyperparameter Randomsearchcv Pada Adaptive Boosting Untuk Prediksi Kelangsungan Hidup Pasien Gagal Jantung," *Jurnal Gaussian*, vol. 11, no. 3, pp. 397–406, Jan. 2023, doi: 10.14710/j.gauss.11.3.397-406.

[7] C. Janiesch, P. Zschech, and K. Heinrich, "Machine learning and deep learning," *Electronic Markets*, 2021, doi: 10.1007/s12525-021-00475-2.

[8] H. Purnomo, Tad Gonsalves, Evangs Mailoa, F. J. Santoso, and M. R. Pribadi, "Metaheuristics Approach for Hyperparameter Tuning of Convolutional Neural Network," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 8, no. 3, pp. 340–345, Jun. 2024, doi: 10.29207/resti.v8i3.5730.

[9] W. Firgiawan, D. Yustianisa, N. Afiah Nur, and F. Teknik, "Hyperparameter Tuning for Optimizing Stunting Classification with KNN, SVM, and Naïve Bayes Algorithms," *Jurnal TEKNO KOMPAK*, vol. 19, no. 1, 2025, doi: https://doi.org/10.33365/jtk.v19i1.4574.

[10] Hirmayant and E. Utami, "Enhanced Heart Disease Diagnosis Using Machine Learning Algorithms: A Comparison of Feature Selection," *Jurnal RESTI*, vol. 9, no. 2, pp. 385–392, Apr. 2025, doi: 10.29207/resti.v9i2.6175.

[11] P. A. Jusia, A. Rahim, H. Yani, and J. Jasmir, "Improving Performance of KNN and C4.5 using Particle Swarm Optimization in Classification of Heart Diseases," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 8, no. 3, pp. 333–339, Jun. 2024, doi: 10.29207/resti.v8i3.5710.

[12] S. A. Bahanshal, R. S. Baraka, B. Kim, and V. Verdhan, "An Optimized Hybrid Fuzzy Weighted k-Nearest Neighbor with the Presence of Data Imbalance." [Online]. Available: www.ijacsa.thesai.org

[13] A. S. Abdulameer, S. Tiun, N. S. Sani, M. Ayob, and A. Y. Taha, "Enhanced clustering models with wiki-based k-nearest neighbors-based representation for web search result clustering," *Journal of King Saud University - Computer and Information Sciences*, vol. 34, no. 3, pp. 840–850, Mar. 2022, doi: 10.1016/j.jksuci.2020.02.003.

[14] J. Haldy and L. Meily Kurniawidjaja, "Coronary Heart Disease Risk and Associated Risk Factor Among Workers at PT.X in 2023," *Media Publikasi Promosi Kesehatan Indonesia (MPPKI)*, vol. 7, no. 6, pp. 1636–1641, Jun. 2024, doi: 10.56338/mppki.v7i6.5318.

[15] D. Kurniadhi, G. Limbong Masiku, and M. P. Sari, "Description of Risk Factors for Coronary Heart Disease at the Family Medical Center Hospital Heart Clinic In 2021," 2024. [Online]. Available: http://ijsoc.goacademica.com

[16] K. Budholiya, S. K. Shrivastava, and V. Sharma, "An optimized XGBoost based diagnostic system for effective prediction of heart disease," *Journal of King Saud University - Computer and Information Sciences*, vol. 34, no. 7, pp. 4514–4523, Jul. 2022, doi: 10.1016/j.jksuci.2020.10.013.

[17] World Health Organization, "Cardiovascular diseases (CVDs)," WHO Fact Sheet, 2019. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds

[18] Kementerian Kesehatan Republik Indonesia, "Hasil Utama Riskesdas 2018," Badan Penelitian dan Pengembangan Kesehatan, 2018. [Online]. Available: https://dinkesjatengprov.go.id/v2018/storage/2019/09/Hasil-Riskesdas-2018.pdf

[19] A. Avigan, "Cleveland Clinic Heart Disease Dataset," Kaggle, 2020. [Online]. Available: https://www.kaggle.com/datasets/aavigan/cleveland-clinic-heart-disease-dataset