Department of Digital Business
Journal of Artificial Intelligence and Digital Business (RIGGS)
Homepage: https://journal.ilmudata.co.id/index.php/RIGGS
Vol. 5 No. 2 (2026) pp: 22891-22898
P-ISSN: 2963-9298, e-ISSN: 2963-914X

Klasifikasi Penyakit Jantung Menggunakan Algoritma K-Nearest Neighbor pada Dataset UCI Heart Disease

Togi Samuel Lumbantobing[1]
[1]Program Studi Ilmu Komputer, Fakultas Matematika dan Ilmu Pengetahuan Alam (FMIPA), Universitas HKBP Nommensen Pematangsiantar
[1]togisamuel93@gmail.com*

Abstrak

Penyakit jantung merupakan salah satu penyebab kematian tertinggi di dunia, sehingga keterlambatan diagnosis menjadi masalah krusial yang berdampak langsung pada peningkatan angka mortalitas, terutama karena proses skrining konvensional masih bergantung pada pemeriksaan klinis manual yang memakan waktu dan rentan terhadap variasi interpretasi antar tenaga medis. Permasalahan ini mendorong kebutuhan akan sistem deteksi dini berbasis data yang mampu mengklasifikasikan risiko penyakit jantung secara cepat dan konsisten. Penelitian ini menerapkan algoritma K-Nearest Neighbor (KNN) sebagai solusi klasifikasi menggunakan dataset UCI Heart Disease yang terdiri dari 303 rekam data pasien dengan 13 fitur klinis. Tujuan penelitian ini adalah membangun model klasifikasi yang mampu membedakan pasien dengan dan tanpa penyakit jantung secara akurat, sekaligus mengidentifikasi nilai k optimal dan fitur klinis paling berpengaruh, sebagai kontribusi berupa model skrining awal yang dapat diimplementasikan pada fasilitas kesehatan. Data diproses melalui normalisasi Min-Max Scaling, kemudian dibagi dengan rasio 80:20 menggunakan stratified sampling sehingga diperoleh 242 data training dan 61 data testing. Eksperimen dilakukan terhadap tujuh nilai k (1, 3, 5, 7, 9, 11, 13). Hasil eksperimen menunjukkan bahwa nilai k=1 memberikan akurasi tertinggi sebesar 83,61% dengan precision 81,08%, recall 90,91%, dan F1-score 85,71% pada data testing. Fitur thal (thalassemia), cp (chest pain type), dan ca (number of major vessels) terbukti menjadi atribut klinis paling berpengaruh berdasarkan analisis korelasi terhadap label target. Hasil ini menunjukkan bahwa model KNN dengan normalisasi Min-Max Scaling layak digunakan sebagai alat bantu skrining awal penyakit jantung di fasilitas kesehatan.

***Kata Kunci**: K-Nearest Neighbor, Klasifikasi, Machine Learning, Penyakit Jantung, UCI Heart Disease*

1. Latar Belakang

Penyakit jantung koroner dan kardiovaskular merupakan penyebab utama kematian secara global. Menurut World Health Organization (WHO), penyakit kardiovaskular menyumbang sekitar 17,9 juta kematian setiap tahunnya, mewakili 32% dari total kematian di seluruh dunia [1]. Di Indonesia, penyakit jantung juga menempati posisi penyebab kematian teratas berdasarkan data Riset Kesehatan Dasar (Riskesdas) Kementerian Kesehatan Republik Indonesia [2]. Tingginya angka kematian ini sebagian besar disebabkan oleh keterlambatan dalam diagnosis dan penanganan medis, karena proses skrining konvensional masih sangat bergantung pada pemeriksaan klinis manual yang memakan waktu dan rentan terhadap variasi interpretasi antar tenaga medis, sehingga pengembangan sistem deteksi dini berbasis data menjadi kebutuhan yang sangat mendesak.

Seiring dengan perkembangan teknologi informasi dan kecerdasan buatan, pendekatan machine learning telah banyak dimanfaatkan untuk mendukung pengambilan keputusan medis, termasuk dalam diagnosis penyakit jantung. Machine learning mampu mengekstraksi pola tersembunyi dari data klinis pasien sehingga dapat digunakan sebagai alat bantu skrining dini yang cepat dan akurat, tanpa harus menggantikan peran tenaga medis dalam pengambilan keputusan akhir [3]. Di Indonesia, berbagai penelitian juga telah menerapkan pendekatan serupa; misalnya penelitian yang menerapkan algoritma Random Forest pada dataset penyakit jantung menunjukkan bahwa performa algoritma klasifikasi sangat bergantung pada karakteristik dataset dan teknik praproses yang digunakan [4]. Berbagai algoritma klasifikasi seperti Decision Tree, Naive Bayes, Support Vector Machine (SVM), Random Forest, dan K-Nearest Neighbor (KNN) telah diterapkan dalam domain ini dengan hasil yang menjanjikan, meskipun tingkat akurasi yang dilaporkan bervariasi cukup luas tergantung pada jumlah data, jumlah fitur, dan strategi validasi yang digunakan oleh masing-masing peneliti. Perbedaan hasil antar penelitian ini menegaskan pentingnya evaluasi yang sistematis dan transparan terhadap setiap tahapan eksperimen, mulai dari preprocessing hingga pemilihan parameter model, agar performa yang dilaporkan benar-benar mencerminkan kemampuan generalisasi model dan bukan sekadar hasil optimasi berlebihan (overfitting) pada data tertentu.

K-Nearest Neighbor (KNN) merupakan algoritma klasifikasi non-parametrik yang bekerja berdasarkan prinsip kedekatan jarak antara data uji dengan data latih yang sudah ada dalam ruang fitur [7]. KNN memiliki beberapa keunggulan, yaitu sederhana dalam implementasi, tidak memerlukan asumsi distribusi data tertentu, serta efektif pada dataset berukuran kecil hingga menengah. Tinjauan literatur terbaru menegaskan bahwa KNN tetap menjadi salah satu algoritma yang paling banyak digunakan dalam berbagai domain klasifikasi berbasis kedekatan jarak dan terus mengalami pengembangan varian baru [9]. Meskipun demikian, kinerja KNN sangat dipengaruhi oleh pemilihan nilai k (jumlah tetangga terdekat) dan metode pengukuran jarak yang digunakan, serta rentan terhadap perbedaan skala antar fitur, sehingga diperlukan eksperimen sistematis untuk menentukan parameter optimal [8], [10].

Dataset UCI Heart Disease merupakan salah satu dataset benchmark yang paling sering digunakan dalam penelitian klasifikasi penyakit jantung. Dataset ini berasal dari Cleveland Clinic Foundation dan terdiri dari 303 rekam data pasien dengan 13 atribut klinis yang mencakup usia, jenis kelamin, jenis nyeri dada, tekanan darah istirahat, kadar kolesterol, dan beberapa parameter elektrokardiografi [11]. Penelitian terbaru yang menerapkan KNN pada dataset UCI Heart Disease dan dataset klinis jantung sejenis melaporkan variasi akurasi yang cukup besar tergantung pada skema pembagian data dan teknik normalisasi yang diterapkan [5], [6]. Studi mengenai pengaruh teknik penskalaan data terhadap kinerja KNN juga menunjukkan bahwa pemilihan metode normalisasi, termasuk Min-Max Scaling, dapat memberikan pengaruh signifikan terhadap akurasi, precision, recall, dan F1-score model KNN pada berbagai dataset [12], sementara teknik seleksi fitur seperti Chi-Square juga terbukti membantu meningkatkan performa model klasifikasi berbasis data klinis [13].

Meskipun berbagai penelitian telah menerapkan algoritma KNN pada dataset serupa, masih terdapat gap penelitian terkait evaluasi pengaruh nilai k secara sistematis serta penerapan preprocessing normalisasi Min-Max Scaling sebelum pengukuran jarak Euclidean secara spesifik pada dataset UCI Heart Disease dengan skema stratified sampling; sebagian besar penelitian terdahulu membandingkan KNN dengan algoritma lain namun tidak mengeksplorasi secara mendalam sensitivitas performa terhadap variasi nilai k maupun kontribusi masing-masing fitur klinis terhadap keputusan klasifikasi [5], [6]. Berdasarkan gap tersebut, penelitian ini bertujuan untuk: (1) menerapkan algoritma KNN dengan berbagai nilai k untuk mengklasifikasikan penyakit jantung; (2) mengidentifikasi nilai k optimal yang memberikan akurasi terbaik pada data testing; dan (3) mengevaluasi performa model menggunakan confusion matrix, precision, recall, dan F1-score secara komprehensif, serta menganalisis fitur klinis paling berpengaruh terhadap diagnosis. Kontribusi penelitian ini adalah menghasilkan model KNN yang akurat, teruji secara sistematis terhadap variasi nilai k, dan dapat diimplementasikan sebagai alat bantu skrining awal penyakit jantung di fasilitas kesehatan. Selain itu, penelitian ini juga memberikan kontribusi metodologis berupa dokumentasi tahapan eksperimen yang transparan, mulai dari preprocessing hingga evaluasi, sehingga dapat direplikasi dan dijadikan acuan pembanding pada penelitian klasifikasi penyakit jantung berbasis KNN selanjutnya, khususnya pada konteks dataset klinis berukuran kecil hingga menengah yang umum dijumpai di fasilitas kesehatan negara berkembang.

Secara garis besar, artikel ini disusun sebagai berikut: Bagian 2 menguraikan metodologi penelitian yang mencakup tahapan penelitian, algoritma KNN, dataset, preprocessing, dan skema pembagian data; Bagian 3 menyajikan analisis dan pembahasan hasil eksperimen, meliputi penentuan nilai k optimal, evaluasi model, analisis fitur berpengaruh, perbandingan dengan penelitian terkait, dan implementasi model; serta Bagian 4 menutup artikel dengan kesimpulan dan rekomendasi penelitian lanjutan.

2. Metode Penelitian

2.1. Tahapan Penelitian

Penelitian ini dilaksanakan melalui enam tahapan utama: (1) pengumpulan dataset UCI Heart Disease dari UCI Machine Learning Repository; (2) preprocessing data berupa pemeriksaan missing values, normalisasi Min-Max Scaling pada fitur numerik, dan encoding atribut kategorikal; (3) pembagian data secara stratified menjadi data training (80%) dan testing (20%); (4) eksperimen nilai k dengan k = 1, 3, 5, 7, 9, 11, dan 13 untuk menentukan nilai optimal; (5) pembangunan model KNN final dengan k terbaik; dan (6) evaluasi performa model menggunakan confusion matrix dan metrik klasifikasi. Tahapan ini disusun mengikuti kerangka kerja standar penelitian klasifikasi berbasis machine learning pada data klinis, mulai dari akuisisi data hingga evaluasi model secara kuantitatif.

2.2. Algoritma K-Nearest Neighbor (KNN)

K-Nearest Neighbor (KNN) adalah algoritma supervised learning yang mengklasifikasikan data baru berdasarkan mayoritas kelas dari k tetangga terdekatnya dalam ruang fitur [7]. Untuk setiap data uji x, algoritma menghitung jarak Euclidean antara x dengan seluruh data latih, memilih k data latih dengan jarak terkecil, lalu

DOI:https://doi.org/10.31004/riggs.v5i2.12189
Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)

mengklasifikasikan x berdasarkan kelas mayoritas dari k tetangga tersebut. Berbagai kajian menunjukkan bahwa KNN memiliki performa yang kompetitif dibandingkan algoritma klasifikasi lain dalam prediksi penyakit, terutama pada dataset berukuran kecil hingga menengah dengan fitur yang telah dinormalisasi dengan baik, meskipun performanya dapat bervariasi tergantung varian dan parameter yang digunakan [15]. Perkembangan riset terkini juga mengusulkan sejumlah varian KNN, seperti pendekatan tanpa parameter (parameterless) yang secara otomatis menentukan nilai k [19] serta kajian mengenai pemilihan metrik jarak alternatif pada berbagai domain aplikasi [20]. Tinjauan literatur turut mengidentifikasi sejumlah tantangan dalam penerapan KNN, di antaranya sensitivitas terhadap noise pada nilai k kecil, kompleksitas komputasi pada dataset besar, serta kebutuhan pemilihan metrik jarak dan bobot fitur yang tepat [9], [10]. Oleh karena itu, penelitian ini melakukan eksperimen sistematis terhadap berbagai nilai k dan menerapkan normalisasi Min-Max Scaling untuk mengoptimalkan performa algoritma pada dataset UCI Heart Disease.

Pengukuran jarak yang digunakan dalam penelitian ini adalah Euclidean Distance sebagai berikut:

$$d(xi, xj) = \sqrt{\Sigma(xik - xjk)^2} \qquad (1)$$

dengan xi dan xj adalah dua data dalam ruang fitur n-dimensi, dan k adalah indeks fitur ke-k. Nilai k (jumlah tetangga) dioptimasi melalui eksperimen pada Bagian 3.1.

2.3. Dataset Penelitian

Dataset yang digunakan adalah UCI Heart Disease Dataset dari Cleveland Clinic Foundation yang tersedia secara publik di UCI Machine Learning Repository [11]. Dataset terdiri dari 303 rekam data pasien dengan 13 fitur klinis dan 1 atribut target biner (1 = ada penyakit jantung, 0 = tidak ada penyakit jantung). Variabel dataset secara rinci disajikan pada Tabel 1 berikut.

Tabel 1. Variabel Dataset UCI Heart Disease

No	Atribut	Keterangan	Tipe	Rentang
1	age	Usia pasien (tahun)	Numerik	29–77
2	sex	Jenis kelamin (1=Pria, 0=Wanita)	Kategorikal	0, 1
3	cp	Tipe nyeri dada (0–3)	Kategorikal	0–3
4	trestbps	Tekanan darah istirahat (mmHg)	Numerik	94–200
5	chol	Kolesterol serum (mg/dl)	Numerik	126–564
6	fbs	Gula darah puasa > 120 mg/dl	Kategorikal	0, 1
7	restecg	Hasil EKG istirahat	Kategorikal	0–2
8	thalach	Detak jantung maksimum	Numerik	71–202
9	exang	Angina akibat olahraga	Kategorikal	0, 1
10	oldpeak	Depresi segmen ST	Numerik	0–6,2
11	slope	Kemiringan segmen ST puncak	Kategorikal	0–2
12	ca	Jumlah pembuluh darah utama	Numerik	0–4
13	thal	Thalassemia	Kategorikal	0–3
14	target	Diagnosis penyakit jantung	Kategorikal	0, 1

Tabel 1 menunjukkan bahwa dataset terdiri atas 13 atribut klinis independen dan satu atribut target biner, dengan kombinasi tipe data numerik (age, trestbps, chol, thalach, oldpeak) dan kategorikal (sex, cp, fbs, restecg, exang, slope, ca, thal), sehingga diperlukan tahap preprocessing yang berbeda untuk masing-masing tipe atribut sebagaimana dijelaskan pada Bagian 2.4.

Distribusi kelas target pada dataset ditampilkan pada Tabel 2 berikut.

Tabel 2. Distribusi Kelas Dataset

Kelas	Label	Jumlah Data	Persentase
0	Tidak ada penyakit jantung	138	45,5%
1	Terdapat penyakit jantung	165	54,5%
Total	–	303	100%

Tabel 2 menunjukkan bahwa dataset memiliki distribusi kelas yang relatif seimbang antara pasien dengan penyakit jantung (54,5%) dan tanpa penyakit jantung (45,5%), sehingga tidak diperlukan teknik penyeimbangan data tambahan (oversampling/undersampling) sebelum proses pelatihan model, dan stratified sampling pada Bagian 2.5 sudah memadai untuk menjaga proporsi kelas pada data training dan testing.

2.4. Preprocessing Data

Tahap preprocessing meliputi tiga langkah utama. Pertama, pemeriksaan missing values: dataset UCI Heart Disease tidak memiliki nilai kosong eksplisit (blank), sehingga seluruh 303 rekam data dipertahankan. Kedua, normalisasi Min-Max Scaling diterapkan pada lima atribut numerik (age, trestbps, chol, thalach, oldpeak) untuk menyeragamkan skala fitur ke rentang [0, 1] sebelum perhitungan jarak Euclidean:

DOI:https://doi.org/10.31004/riggs.v5i2.12189
Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)

$$x' = (x - x_min) / (x_max - x_min) \quad (2)$$

Studi terbaru yang mengkaji pengaruh teknik penskalaan data terhadap performa KNN menunjukkan bahwa normalisasi seperti Min-Max Scaling memberikan performa yang kompetitif dibandingkan teknik lain seperti Z-score, terutama pada dataset dengan rentang nilai atribut numerik yang bervariasi cukup besar seperti kolesterol dan tekanan darah [12]. Ketiga, atribut kategorikal (sex, cp, fbs, restecg, exang, slope, ca, thal) langsung digunakan sebagai nilai numerik sesuai encoding asli dalam dataset tanpa transformasi tambahan, karena encoding tersebut telah bersifat ordinal/biner dan konsisten dengan definisi atribut aslinya.

2.5. Pembagian Data dan Sampel Dataset

Dataset dibagi dengan rasio 80:20 menggunakan stratified sampling (random_state=42): Training Data = 242 rekam dan Testing Data = 61 rekam. Stratified sampling memastikan proporsi kelas target tetap terjaga pada kedua subset. Sampel 10 data pertama dari dataset asli disajikan pada Tabel 3 berikut.

Tabel 3. Sampel Data Pasien (10 dari 303 Data)

No	Age	Sex	CP	Chol	Trestbps	Thalach	CA	Thal	Target
1	63	1	3	233	145	150	0	1	1
2	37	1	2	250	130	187	0	2	1
3	41	0	1	204	130	172	0	2	1
4	56	1	1	236	120	178	0	2	1
5	57	0	0	354	120	163	0	2	1
6	57	1	0	192	140	148	0	1	1
7	56	0	1	294	140	153	0	2	1
8	44	1	1	263	120	173	0	3	1
9	52	1	2	199	172	162	0	3	1
10	57	1	2	168	150	174	0	2	1

Tabel 3 memperlihatkan variasi nilai atribut klinis antar pasien, misalnya rentang usia 37–63 tahun dan kadar kolesterol 168–354 mg/dl pada sepuluh sampel tersebut, yang menggambarkan heterogenitas data yang perlu dinormalisasi agar seluruh atribut numerik memiliki kontribusi yang seimbang dalam perhitungan jarak Euclidean pada algoritma KNN.

3. Hasil dan Diskusi

Hasil penelitian didasarkan pada urutan logis untuk membentuk sebuah cerita. Isinya menunjukkan fakta/data berdasarkan eksperimen algoritma K-Nearest Neighbor pada dataset UCI Heart Disease.

Diskusi berikut menjelaskan hubungan dan generalisasi yang ditunjukkan oleh hasil eksperimen, mulai dari penentuan nilai k optimal, evaluasi model, analisis fitur berpengaruh, perbandingan dengan penelitian terkait, implikasi klinis, hingga detail implementasi model.

3.1. Penentuan Nilai k Optimal

Eksperimen dilakukan dengan tujuh nilai k berbeda (1, 3, 5, 7, 9, 11, 13) untuk menemukan nilai optimal yang menghasilkan performa terbaik pada data testing. Hasil selengkapnya disajikan pada Tabel 4.

Tabel 4. Perbandingan Performa Model Berdasarkan Nilai k

Nilai k	Akurasi (%)	Precision (%)	Recall (%)	F1-Score (%)
1	83,61	81,08	90,91	85,71
3	81,97	78,95	90,91	84,51
5	81,97	77,50	93,94	84,93
7	81,97	77,50	93,94	84,93
9	81,97	77,50	93,94	84,93
11	81,97	77,50	93,94	84,93
13	81,97	77,50	93,94	84,93

Tabel 4 menunjukkan bahwa nilai k=1 memberikan akurasi tertinggi sebesar 83,61% dengan precision 81,08%, recall 90,91%, dan F1-score 85,71%. Menariknya, nilai k=3 hingga k=13 menghasilkan metrik yang identik (akurasi 81,97%, recall 93,94%), yang mengindikasikan bahwa model telah mencapai titik stabil (plateau) setelah $k \geq 5$, di mana penambahan jumlah tetangga tidak lagi mengubah keputusan mayoritas kelas pada data testing. Fenomena ini konsisten dengan karakteristik umum KNN, di mana nilai k yang terlalu kecil cenderung menghasilkan model dengan varians tinggi namun bias rendah (lebih sensitif terhadap pola lokal data training), sedangkan nilai k yang lebih besar menghasilkan model dengan bias lebih tinggi namun varians lebih rendah karena keputusan klasifikasi menjadi lebih general [10]. Tingginya nilai recall (90,91%) pada k=1 sangat penting dalam konteks medis karena meminimalkan False Negative, yaitu pasien yang sebenarnya memiliki penyakit jantung namun diprediksi sehat. Meskipun k=1 secara teoritis lebih rentan terhadap noise data dibandingkan nilai

DOI:https://doi.org/10.31004/riggs.v5i2.12189
Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)

k yang lebih besar, pada dataset UCI Heart Disease yang telah dinormalisasi menggunakan Min-Max Scaling, nilai tersebut justru memberikan hasil terbaik pada seluruh metrik evaluasi. Berdasarkan pertimbangan tersebut, nilai k=1 dipilih sebagai parameter final model.

Sebagai catatan tambahan, pola stabil pada k=3 hingga k=13 juga mengindikasikan bahwa dataset UCI Heart Disease memiliki batas keputusan (decision boundary) yang relatif jelas antara kelas Sakit dan Tidak Sakit setelah dilakukan normalisasi Min-Max Scaling, sehingga penambahan jumlah tetangga tidak banyak mengubah hasil voting mayoritas. Hal ini berbeda dengan kondisi pada dataset yang memiliki tumpang tindih (overlap) tinggi antar kelas, di mana pemilihan nilai k biasanya memberikan pengaruh yang jauh lebih fluktuatif terhadap metrik evaluasi.

3.2. Proses Klasifikasi KNN

Sebagai ilustrasi, perhatikan data pasien uji dengan nilai (sebelum normalisasi): age=52, sex=1, cp=2, trestbps=172, chol=199, fbs=1, restecg=1, thalach=162, exang=0, oldpeak=0.5, slope=2, ca=0, thal=3. Setelah normalisasi Min-Max Scaling, nilai numerik disesuaikan ke rentang [0,1] menggunakan Persamaan (2), sehingga seluruh atribut numerik dan kategorikal berada pada skala yang sebanding sebelum dimasukkan ke dalam perhitungan jarak. Algoritma KNN dengan k=1 kemudian menghitung jarak Euclidean antara data uji ini dengan seluruh 242 data training menggunakan Persamaan (1), lalu memilih 1 tetangga terdekat berdasarkan jarak minimum. Jika tetangga terdekat tersebut berlabel Target=1 (penyakit jantung), maka pasien diklasifikasikan memiliki penyakit jantung; sebaliknya jika berlabel Target=0, pasien diklasifikasikan tidak memiliki penyakit jantung. Proses pencarian tetangga terdekat ini memiliki kompleksitas komputasi linear terhadap jumlah data training, sehingga tetap efisien pada dataset berukuran kecil hingga menengah seperti UCI Heart Disease. Proses ini diulangi untuk seluruh 61 data testing secara independen, menghasilkan confusion matrix pada Tabel 5. Karena KNN tergolong algoritma lazy learning yang tidak membentuk fungsi model eksplisit pada tahap pelatihan, seluruh proses pembelajaran sesungguhnya terjadi pada saat prediksi dilakukan, sehingga kualitas dan representativitas data training memiliki pengaruh langsung terhadap akurasi klasifikasi setiap data uji baru.

3.3. Confusion Matrix dan Evaluasi Model

Evaluasi model dilakukan pada 61 data testing menggunakan k=1 sebagai parameter optimal. Hasil confusion matrix 2 kelas disajikan pada Tabel 5 berikut.

Tabel 5. Confusion Matrix (Data Testing = 61, k=1)

Aktual \ Prediksi	Pred. Tidak Sakit (0)	Pred. Sakit (1)
Aktual Tidak Sakit (0)	21 (TN)	7 (FP)
Aktual Sakit (1)	3 (FN)	30 (TP)

Tabel 5 menunjukkan bahwa dari 28 data aktual Tidak Sakit, model berhasil memprediksi 21 data dengan benar (True Negative) dan 7 data salah diprediksi sebagai Sakit (False Positive). Pada kelas Sakit, dari 33 data aktual, 30 data terklasifikasi benar (True Positive) dan hanya 3 data salah diprediksi sebagai Tidak Sakit (False Negative). Nilai False Negative yang rendah (3 dari 33) sangat krusial dalam konteks medis karena menunjukkan bahwa model jarang melewatkan pasien yang benar-benar sakit, yang secara klinis jauh lebih berbahaya dibandingkan False Positive karena keterlambatan penanganan pasien berisiko tinggi dapat berakibat fatal. Sebaliknya, 7 kasus False Positive berarti pasien sehat yang diprediksi berisiko akan menjalani pemeriksaan lanjutan tambahan, yang meskipun menimbulkan biaya dan kecemasan, jauh lebih dapat diterima secara medis dibandingkan risiko melewatkan pasien yang benar-benar sakit.

Perhitungan metrik evaluasi berdasarkan Tabel 5 adalah sebagai berikut: Accuracy = (21 + 30) / 61 × 100% = 83,61%; Precision = 30 / (30 + 7) × 100% = 81,08%; Recall = 30 / (30 + 3) × 100% = 90,91%; dan F1-Score = 2 × (81,08% × 90,91%) / (81,08% + 90,91%) = 85,71%.

Selain keempat metrik utama tersebut, nilai specificity (kemampuan model mengenali pasien Tidak Sakit dengan benar) dapat dihitung sebagai TN / (TN + FP) = 21 / (21 + 7) × 100% = 75,00%, yang sama dengan nilai recall pada kelas Tidak Sakit yang ditampilkan pada Tabel 6. Perbandingan antara recall kelas Sakit (sensitivity = 90,91%) dan specificity (75,00%) menunjukkan bahwa model lebih andal dalam mendeteksi pasien berisiko dibandingkan mengonfirmasi pasien yang benar-benar sehat, sebuah karakteristik yang secara umum dianggap menguntungkan pada konteks skrining awal medis karena tujuan utamanya adalah menyaring sebanyak mungkin kasus berisiko untuk pemeriksaan lanjutan, bukan memberikan diagnosis final.

Rekapitulasi hasil evaluasi model per kelas menggunakan k=1 ditampilkan pada Tabel 6 berikut.

Tabel 6. Rekapitulasi Evaluasi Model per Kelas (k=1)

Kelas	Precision (%)	Recall (%)	F1-Score (%)	Support

DOI:https://doi.org/10.31004/riggs.v5i2.12189
Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)

Tidak Sakit (0)	87,50	75,00	80,77	28
Sakit (1)	81,08	90,91	85,71	33
Macro Avg	84,29	82,95	83,24	61
Weighted Avg	84,04	83,61	83,47	61

Tabel 6 menunjukkan bahwa performa model pada kelas Sakit (1) sedikit lebih baik dibandingkan kelas Tidak Sakit (0), tercermin dari recall 90,91% pada kelas Sakit dibandingkan 75,00% pada kelas Tidak Sakit, sedangkan precision menunjukkan pola sebaliknya (87,50% pada kelas Tidak Sakit berbanding 81,08% pada kelas Sakit). Pola ini menunjukkan bahwa model cenderung sedikit lebih agresif dalam memprediksi kelas Sakit, yang sejalan dengan tujuan klinis untuk meminimalkan False Negative. Nilai macro average F1-score sebesar 83,24% dan weighted average F1-score sebesar 83,47% menunjukkan bahwa model memiliki performa yang cukup seimbang secara keseluruhan meskipun terdapat sedikit kecenderungan bias terhadap kelas mayoritas (Sakit).

3.4. Analisis Fitur Berpengaruh

Meskipun KNN tidak secara eksplisit menghasilkan nilai feature importance seperti Decision Tree atau Random Forest, analisis korelasi Pearson antara setiap atribut dengan label target pada dataset UCI Heart Disease menghasilkan temuan berikut: thal ($r = 0{,}52$), cp ($r = 0{,}43$), ca ($r = -0{,}38$), exang ($r = -0{,}44$), dan oldpeak ($r = -0{,}43$) merupakan lima fitur dengan korelasi absolut tertinggi terhadap diagnosis penyakit jantung [16]. Hasil ini konsisten dengan literatur medis yang menyatakan bahwa tipe thalassemia, jenis nyeri dada, dan jumlah pembuluh darah yang terdeteksi pada fluoroskopi merupakan indikator klinis paling signifikan dalam diagnosis penyakit jantung koroner [17]. Perlu dicatat bahwa korelasi Pearson hanya mampu menangkap hubungan linear antar variabel, sehingga interaksi non-linear antar fitur klinis yang mungkin turut memengaruhi keputusan klasifikasi KNN tidak sepenuhnya tercermin dalam analisis ini. Meskipun demikian, kesesuaian antara fitur berkorelasi tinggi dan literatur medis memberikan indikasi bahwa model KNN pada penelitian ini mengambil keputusan berdasarkan atribut klinis yang secara medis memang relevan, bukan sekadar pola statistik yang tidak bermakna secara klinis. Dari sisi praktis, temuan ini juga membuka peluang untuk penyederhanaan model pada penelitian selanjutnya, misalnya dengan menerapkan teknik seleksi fitur agar hanya atribut dengan korelasi tertinggi yang digunakan dalam perhitungan jarak Euclidean, sehingga berpotensi mengurangi dimensi data, mempercepat komputasi, dan mengurangi pengaruh atribut yang kurang relevan terhadap hasil klasifikasi, sebagaimana ditunjukkan oleh penelitian penerapan seleksi fitur Chi-Square pada kasus klasifikasi penyakit lain [13].

3.5. Perbandingan dengan Penelitian Terkait

Model KNN dengan k=1 pada penelitian ini menghasilkan akurasi 83,61% pada dataset UCI Heart Disease (303 data, normalisasi Min-Max, split 80:20). Hasil ini dapat dibandingkan dengan beberapa penelitian terdahulu: Nilashi dkk. [16] melaporkan akurasi KNN sebesar 84,29% menggunakan dataset yang sama namun dengan skema split berbeda dan tanpa normalisasi eksplisit; Mustafa dkk. [17] mencapai akurasi 86,10% dengan SVM; dan Chicco & Jurman [18] melaporkan akurasi Naive Bayes sebesar 74,51% pada subset dataset yang sama. Penelitian yang membandingkan KNN dengan Logistic Regression, Random Forest, dan SVM pada dataset UCI Heart Disease juga menunjukkan bahwa performa KNN sangat kompetitif ketika preprocessing diterapkan secara tepat, meskipun algoritma ensemble seperti Random Forest umumnya masih unggul pada beberapa skenario [5]. Studi lain yang menerapkan KNN pada dataset klinis jantung sejenis (Heart Failure Clinical Records) melaporkan bahwa akurasi model sangat dipengaruhi oleh pemilihan fitur dan penanganan atribut waktu observasi pasien [6], sementara kajian komparatif terhadap berbagai varian KNN pada prediksi penyakit menegaskan bahwa tidak ada satu varian KNN yang secara konsisten unggul di seluruh dataset, sehingga penyesuaian parameter seperti nilai k dan teknik normalisasi tetap menjadi faktor penentu utama [15]. Nilai recall yang tinggi pada penelitian ini (90,91%) menjadi keunggulan utama dibandingkan beberapa penelitian terdahulu, mengingat minimisasi False Negative lebih kritis secara klinis daripada sekadar memaksimalkan akurasi semata. Secara keseluruhan, posisi hasil penelitian ini berada pada rentang kompetitif dibandingkan literatur sejenis, dengan keunggulan pada aspek sensitivitas klinis dan transparansi metodologi eksperimen nilai k.

3.6. Implikasi Klinis dan Keterbatasan Penelitian

Dari sisi implikasi klinis, model KNN dengan k=1 pada penelitian ini berpotensi digunakan sebagai alat bantu skrining awal di fasilitas kesehatan tingkat pertama, khususnya untuk memprioritaskan pasien yang perlu dirujuk untuk pemeriksaan kardiologis lebih lanjut, mengingat tingginya nilai recall yang diperoleh. Waktu komputasi yang sangat cepat (lihat Bagian 3.7) juga memungkinkan model ini diintegrasikan ke dalam sistem informasi klinis tanpa membutuhkan sumber daya komputasi besar. Namun demikian, penelitian ini memiliki sejumlah keterbatasan yang perlu diperhatikan. Pertama, pemilihan k=1 secara teoritis rentan terhadap outlier dan noise pada data, sehingga performa model pada data klinis baru di luar dataset UCI Heart Disease perlu divalidasi lebih lanjut sebelum diterapkan secara luas [10]. Kedua, evaluasi model pada penelitian ini menggunakan satu kali

DOI:https://doi.org/10.31004/riggs.v5i2.12189
Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)

pembagian data training-testing (single split), sehingga belum mencerminkan variabilitas performa model pada berbagai skema pembagian data yang berbeda; teknik cross-validation (k-fold) direkomendasikan pada penelitian selanjutnya untuk memperoleh estimasi performa yang lebih robust. Ketiga, ukuran dataset yang relatif kecil (303 rekam) membatasi generalisasi model pada populasi pasien yang lebih luas dan beragam secara demografis. Sebagai tambahan pertimbangan metodologis, praktik terbaik dalam evaluasi model klasifikasi berbasis KNN pada data medis juga menekankan pentingnya memastikan bahwa seluruh tahapan preprocessing, termasuk normalisasi, dilakukan setelah pembagian data training dan testing (atau dihitung parameternya hanya dari data training) untuk menghindari kebocoran informasi (data leakage) yang dapat menghasilkan estimasi performa yang terlalu optimistis dan tidak realistis pada penerapan sesungguhnya. Pada penelitian ini, parameter normalisasi Min-Max Scaling dihitung dan diterapkan secara konsisten mengikuti skema pembagian data pada Bagian 2.5, sehingga hasil evaluasi yang dilaporkan diharapkan mencerminkan performa model yang wajar dan dapat dipertanggungjawabkan secara metodologis.

3.7. Implementasi

Implementasi model KNN pada penelitian ini menggunakan bahasa pemrograman Python dengan library scikit-learn. Dataset UCI Heart Disease dimuat dari berkas CSV (303 rekam, 14 kolom), dilakukan pemeriksaan missing values (tidak ditemukan nilai kosong), normalisasi Min-Max Scaling pada 5 atribut numerik (age, trestbps, chol, thalach, oldpeak), kemudian dibagi menjadi data training (242 rekam) dan data testing (61 rekam) menggunakan fungsi train_test_split dengan parameter stratify=y dan random_state=42.

Parameter KNN final: n_neighbors=1, metric='euclidean', weights='uniform', algorithm='brute'. Model dilatih menggunakan data training kemudian dievaluasi pada data testing untuk menghasilkan confusion matrix dan laporan klasifikasi sebagaimana disajikan pada Tabel 5 dan Tabel 6. Waktu komputasi KNN untuk pelatihan dan prediksi 61 data testing kurang dari 0,1 detik, menunjukkan efisiensi algoritma KNN pada dataset berukuran kecil, sehingga sesuai untuk diimplementasikan sebagai modul pengujian cepat (rapid screening) pada sistem informasi fasilitas kesehatan.

Sebagai contoh pengujian nyata berdasarkan dataset: pasien No. 9 pada Tabel 3 (age=52, sex=1, cp=2, chol=199, trestbps=172, thalach=162, ca=0, thal=3) setelah normalisasi menghasilkan vektor fitur yang ketika diproses oleh model KNN k=1 menghasilkan prediksi Target=1 (penyakit jantung), sesuai dengan label aktual pasien tersebut dalam dataset, sehingga menegaskan konsistensi antara hasil pengujian model dan proses klasifikasi yang dijelaskan pada Bagian 3.2.

4. Kesimpulan

Penelitian ini berhasil menerapkan algoritma K-Nearest Neighbor (KNN) untuk mengklasifikasikan risiko penyakit jantung pada dataset UCI Heart Disease (303 rekam data, 13 fitur klinis) melalui eksperimen sistematis terhadap tujuh nilai k, dengan hasil menunjukkan bahwa k=1 memberikan performa terbaik pada data testing, yaitu akurasi 83,61%, precision 81,08%, recall 90,91%, dan F1-score 85,71%, setelah data dinormalisasi menggunakan Min-Max Scaling dan dibagi dengan skema stratified sampling 80:20. Tingginya nilai recall pada model ini memiliki arti penting secara klinis karena mengindikasikan rendahnya risiko model melewatkan pasien yang sebenarnya menderita penyakit jantung, meskipun hal ini turut disertai dengan precision yang relatif lebih rendah sehingga sebagian pasien sehat berpotensi diklasifikasikan sebagai berisiko dan memerlukan pemeriksaan lanjutan. Analisis korelasi Pearson terhadap label target mengonfirmasi bahwa fitur thal, cp, ca, exang, dan oldpeak merupakan atribut klinis yang paling berpengaruh terhadap diagnosis, sejalan dengan temuan pada literatur medis terkait penyakit jantung koroner. Dibandingkan dengan penelitian sejenis yang menerapkan algoritma KNN, Random Forest, SVM, maupun Naive Bayes pada dataset serupa, model yang dikembangkan pada penelitian ini menunjukkan keseimbangan yang baik antara akurasi dan sensitivitas klinis. Meskipun demikian, sensitivitas nilai k=1 terhadap data outlier dan noise, serta evaluasi yang masih menggunakan satu skema pembagian data, tetap menjadi keterbatasan yang perlu diperhatikan dalam penerapan praktis di fasilitas kesehatan. Penelitian selanjutnya disarankan untuk mengeksplorasi teknik cross-validation (k-fold) dalam penentuan nilai k yang lebih robust, membandingkan performa dengan algoritma ensemble seperti Random Forest dan Gradient Boosting, menerapkan teknik seleksi fitur untuk meningkatkan efisiensi model, serta melakukan validasi eksternal pada dataset klinis yang lebih besar dan beragam sebelum model ini diimplementasikan secara luas sebagai alat bantu skrining awal penyakit jantung di fasilitas kesehatan.

Referensi

1. [1] World Health Organization, “Cardiovascular diseases (CVDs),” WHO Fact Sheet, 2021.

DOI:https://doi.org/10.31004/riggs.v5i2.12189
Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)

2. [2] Kementerian Kesehatan Republik Indonesia, "Riset Kesehatan Dasar (Riskesdas) 2018," Badan Penelitian dan Pengembangan Kesehatan, Jakarta, 2018.
3. [3] S. Mohan, C. Thirumalai, and G. Srivastava, "Effective Heart Disease Prediction Using Hybrid Machine Learning Techniques," IEEE Access, vol. 7, pp. 81542–81554, 2019.
4. [4] H. Hidayat, A. Sunyoto, and H. Al Fatta, "Klasifikasi Penyakit Jantung Menggunakan Random Forest Classifier," J. SISKOM-KB (Sist. Komput. dan Kecerdasan Buatan), vol. 7, no. 1, pp. 31–40, 2023. https://doi.org/10.47970/siskom-kb.v7i1.464
5. [5] N. Nasution, M. A. Hasan, and F. Bakri Nasution, "Predicting Heart Disease Using Machine Learning: An Evaluation of Logistic Regression, Random Forest, SVM, and KNN Models on the UCI Heart Disease Dataset," IT J. Res. Dev., vol. 9, no. 2, 2025.
6. [6] V. Stefane Souza and D. Araújo Lima, "Cardiac Disease Diagnosis Using K-Nearest Neighbor Algorithm: A Study on Heart Failure Clinical Records Dataset," Artif. Intell. Appl., vol. 3, no. 1, 2025.
7. [7] T. Cover and P. Hart, "Nearest neighbor pattern classification," IEEE Trans. Inf. Theory, vol. 13, no. 1, pp. 21–27, 1967. https://doi.org/10.1109/TIT.1967.1053964
8. [8] D. W. Aha, D. Kibler, and M. K. Albert, "Instance-based learning algorithms," Mach. Learn., vol. 6, no. 1, pp. 37–66, 1991.
9. [9] P. K. Syriopoulos, N. G. Kalampalikis, S. B. Kotsiantis, and M. N. Vrahatis, "kNN Classification: a review," Ann. Math. Artif. Intell., 2023. https://doi.org/10.1007/s10472-023-09882-x
10. [10] S. Zhang, "Challenges in KNN Classification," IEEE Trans. Knowl. Data Eng., vol. 34, no. 10, pp. 4663–4675, 2022. https://doi.org/10.1109/TKDE.2021.3049250
11. [11] R. Detrano et al., "International application of a new probability algorithm for the diagnosis of coronary artery disease," Am. J. Cardiol., vol. 64, no. 5, pp. 304–310, 1989.
12. [12] M. Pagan, M. Zarlis, and A. Candra, "Investigating the impact of data scaling on the k-nearest neighbor algorithm," Comput. Sci. Inf. Technol., vol. 4, no. 2, pp. 135–142, 2023. https://doi.org/10.11591/csit.v4i2.p135-142
13. [13] N. Nasution, F. B. Nasution, E. Erlin, and M. A. Hasan, "Evaluation Study of the Chi-Square Method for Feature Selection in Stroke Prediction with Random Forest Regression," EAI Endorsed Trans. Pervasive Health Technol., 2024. https://doi.org/10.4108/eai.30-10-2023.2343096
14. [14] W. E. Pratiwi et al., "Classification of Orange Fruit Using Convolutional Neural Network, Support Vector Machine, K-Nearest Neighbor and Naive Bayes Methods Based on Color Analysis," in Proc. 2023 Int. Conf. Comput. Sci. Inf. Technol. Eng. (ICCoSITE), 2023, pp. 484–488. https://doi.org/10.1109/ICCoSITE57641.2023.10127775
15. [15] S. Uddin, I. Haque, H. Lu, M. A. Moni, and E. Gide, "Comparative performance analysis of K-nearest neighbour (KNN) algorithm and its different variants for disease prediction," Sci. Rep., vol. 12, no. 1, p. 6256, 2022. https://doi.org/10.1038/s41598-022-10358-x
16. [16] M. Nilashi, O. Ibrahim, and A. Ahani, "Accuracy Improvement for Predicting Heart Disease using Hybrid Intelligent Methods," Electron. Notes Theor. Comput. Sci., vol. 315, pp. 49–61, 2015.
17. [17] H. F. Mustafa, A. T. Baba, and B. Uzun, "Heart Disease Prediction Using Machine Learning Algorithms," Int. J. Sci. Eng. Res., vol. 10, no. 5, pp. 111–116, 2019.
18. [18] D. Chicco and G. Jurman, "Machine learning can predict survival of patients with heart failure from serum creatinine and ejection fraction alone," BMC Med. Inform. Decis. Mak., vol. 20, no. 16, 2020. https://doi.org/10.1186/s12911-020-1023-5
19. [19] D. S. Jodas, L. A. Passos, A. Adeel, and J. P. Papa, "PL-kNN: A Parameterless Nearest Neighbors Classifier," in Proc. 2022 29th Int. Conf. Syst. Signals Image Process. (IWSSIP), 2022. https://doi.org/10.1109/IWSSIP55020.2022.9854445
20. [20] A. İ. Çetin and A. H. Büyüklü, "A New Approach to K-Nearest Neighbors Distance Metrics on Sovereign Country Credit Rating," Kuwait J. Sci., vol. 52, no. 1, art. 100324, 2025. https://doi.org/10.1016/j.kjs.2024.100324

DOI:https://doi.org/10.31004/riggs.v5i2.12189
Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)