Building of Informatics, Technology and Science (BITS)
Volume 7, No 4, March 2026 Page: 2586−2596
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v7i4.9493

Analisis Performa K-Nearest Neighbor dengan Optimasi F1-Score dan Teknik SMOTE dalam Klasifikasi Risiko Serangan Jantung

Fikri Luqman Pratama*, Muhamad Akrom

Fakultas Ilmu Komputer, Program Studi Teknik Informatika, Universitas Dian Nuswantoro, Semarang, Indonesia
Email: [1,*]111202214814@mhs.dinus.ac.id, [2]m.akrom@dsn.dinus.ac.id
Email Penulis Korespondensi: 111202214814@mhs.dinus.ac.id
Submitted: **07/03/2026**; Accepted: **19/03/2026**; Published**: 19/03/2026**

Abstrak−Serangan jantung merupakan salah satu penyebab utama kematian di dunia sehingga diperlukan metode prediksi yang mampu mendeteksi risiko secara lebih akurat sejak dini. Namun, dalam banyak dataset medis terdapat permasalahan ketidakseimbangan distribusi kelas (class imbalance), di mana jumlah data pasien dengan risiko tinggi jauh lebih sedikit dibandingkan dengan pasien normal. Kondisi ini dapat menyebabkan model machine learning cenderung bias terhadap kelas mayoritas dan menurunkan kemampuan model dalam mendeteksi kasus berisiko tinggi. Penelitian ini bertujuan untuk menganalisis kinerja algoritma K-Nearest Neighbor (KNN) yang dioptimalkan menggunakan F1-score dan dikombinasikan dengan teknik Synthetic Minority Over-sampling Technique (SMOTE) dalam klasifikasi risiko serangan jantung. Dataset yang digunakan merupakan Heart Attack Dataset yang terdiri atas fitur numerik dan kategorikal. Metode penelitian menggunakan pendekatan eksperimen dengan membangun pipeline machine learning yang mencakup pra-pemrosesan data, penanganan missing value, standardisasi fitur, oversampling menggunakan SMOTE, serta optimasi hiperparameter melalui GridSearchCV dengan F1-score sebagai metrik utama. Evaluasi model dilakukan menggunakan Stratified 5-Fold Cross-Validation dengan metrik accuracy, precision, recall, F1-score, dan ROC-AUC. Hasil penelitian menunjukkan bahwa model KNN baseline menghasilkan accuracy sebesar 98,50%, precision 95,27%, recall 81,47%, dan ROC-AUC 0,9278. Sementara itu, model KNN dengan SMOTE memperoleh recall sebesar 87,27% dan ROC-AUC 0,9484, yang menunjukkan peningkatan kemampuan deteksi kasus serangan jantung serta penurunan false negative hingga 31%, meskipun precision menurun menjadi 72,15%. Berdasarkan hasil tersebut, integrasi SMOTE dan optimasi hiperparameter terbukti efektif dalam meningkatkan sensitivitas model, sehingga lebih sesuai diterapkan dalam konteks medis yang memprioritaskan keselamatan pasien.

Kata Kunci: K-Nearest Neighbor; SMOTE; Klasifikasi Penyakit Jantung; Machine Learning Medis; Class Imbalance

Abstract−Heart attack is one of the leading causes of death worldwide, making early risk prediction essential for improving patient outcomes. However, many medical datasets suffer from class imbalance, where the number of high-risk cases is significantly smaller than normal cases. This condition may cause machine learning models to be biased toward the majority class and reduce their ability to detect high-risk patients. This study aims to analyze the performance of the K-Nearest Neighbor (KNN) algorithm optimized using F1-score and combined with the Synthetic Minority Over-sampling Technique (SMOTE) for heart attack risk classification. The dataset used is the Heart Attack Dataset, which consists of numerical and categorical features. The research applies an experimental approach by developing a machine learning pipeline that includes data preprocessing, missing value handling, feature standardization, oversampling using SMOTE, and hyperparameter optimization through GridSearchCV with F1-score as the main evaluation metric. Model evaluation is conducted using Stratified 5-Fold Cross-Validation with accuracy, precision, recall, F1-score, and ROC-AUC metrics. The results show that the baseline KNN model achieves an accuracy of 98.50%, precision 95.27%, recall 81.47%, and ROC-AUC 0.9278. Meanwhile, the KNN model integrated with SMOTE attains a recall of 87.27% and ROC-AUC of 0.9484, indicating improved detection of heart attack cases and a reduction in false negatives by 31%, although precision decreases to 72.15%. These findings demonstrate that the integration of SMOTE and hyperparameter optimization effectively improves model sensitivity, making it more suitable for medical applications that prioritize patient safety.

Keywords: K-Nearest Neighbor; SMOTE; Heart Disease Classification; Medical Machine Learning; Class Imbalance

1. PENDAHULUAN

Perkembangan teknologi machine learning dalam bidang kesehatan telah menunjukkan potensi yang signifikan dalam mendukung proses diagnosis, prediksi penyakit, serta pengambilan keputusan klinis berbasis data. Penyakit kardiovaskular khususnya serangan jantung, merupakan salah satu penyebab utama kematian di dunia dan menimbulkan beban signifikan pada sistem kesehatan masyarakat. Deteksi dini terhadap individu yang memiliki risiko tinggi menjadi krusial untuk menurunkan angka kematian serta mengurangi biaya perawatan. Seiring dengan meningkatnya digitalisasi layanan kesehatan, data medis yang tersedia menjadi semakin besar dan kompleks, sehingga menuntut penerapan teknologi machine learning untuk analisis yang cepat, akurat, dan berbasis bukti [1], [2], [3].

Berbagai algoritma machine learning telah digunakan untuk prediksi risiko penyakit jantung, termasuk K-Nearest Neighbor (KNN), Random Forest, dan XGBoost [4], [5]. KNN menjadi salah satu algoritma populer karena sederhana dan efektif dalam menangani permasalahan klasifikasi berbasis jarak antar data. Namun, performa algoritma ini dapat menurun apabila dataset mengalami ketidakseimbangan kelas, di mana jumlah pasien berisiko tinggi lebih sedikit dibanding kelas mayoritas [6], [7]. Kondisi ini mengakibatkan model cenderung bias terhadap kelas mayoritas sehingga kurang sensitif dalam mendeteksi kasus minoritas.

Salah satu metode yang banyak digunakan untuk mengatasi ketidakseimbangan kelas adalah Synthetic Minority Over-sampling Technique (SMOTE), yang bekerja dengan membuat data sintetis pada kelas minoritas sehingga distribusi kelas menjadi lebih seimbang [8], [9]. Penelitian sebelumnya menunjukkan bahwa penerapan SMOTE pada dataset penyakit jantung mampu meningkatkan akurasi model serta memperbaiki kemampuan deteksi

Building of Informatics, Technology and Science (BITS)
Volume 7, No 4, March 2026 Page: 2586−2596
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v7i4.9493

kasus minoritas[10], [11]. Namun, sebagian besar penelitian masih berfokus pada algoritma ensemble atau boosting, sementara analisis empiris pengaruh SMOTE terhadap performa KNN secara spesifik masih terbatas [12], [13].

Selain itu, evaluasi performa model prediksi penyakit jantung pada kondisi class imbalance memerlukan metrik yang representatif. Metrik umum seperti akurasi sering menyesatkan karena lebih dipengaruhi kelas mayoritas. Oleh karena itu, F1-score menjadi metrik yang lebih relevan untuk menilai keseimbangan antara presisi dan recall pada kasus minoritas [14], [15]. Meski demikian, penggunaan F1-score dalam evaluasi KNN yang dipadukan dengan SMOTE masih jarang dilakukan secara sistematis, sehingga terdapat gap penelitian dalam hal ini [16].

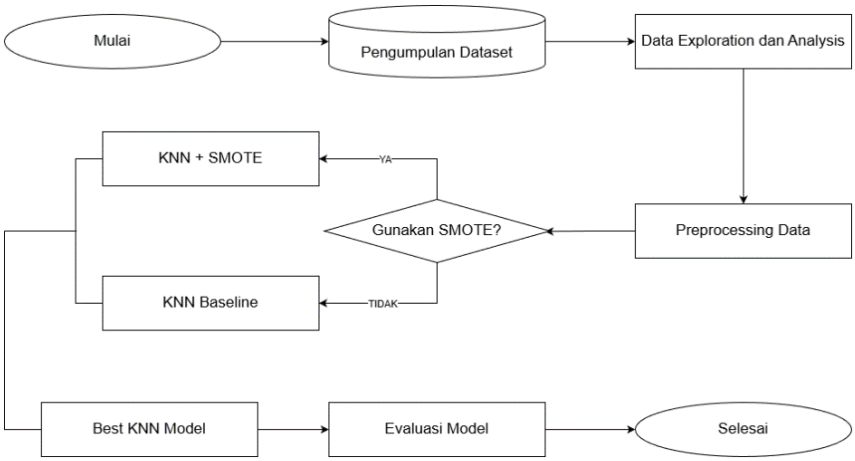
Berdasarkan studi sebelumnya, terlihat bahwa pertama, penelitian empiris mengenai pengaruh SMOTE terhadap performa KNN secara khusus masih minim. Kedua, perbandingan sistematis antara KNN baseline dan KNN dengan SMOTE jarang dilakukan. Ketiga, optimasi parameter KNN untuk dataset tidak seimbang belum dilakukan secara menyeluruh. Keempat, evaluasi menggunakan metrik F1-score masih terbatas pada penelitian sebelumnya [17], [18]. Kekurangan tersebut memperlihatkan perlunya penelitian yang menitikberatkan pada kombinasi KNN dengan SMOTE serta evaluasi performa menggunakan metrik yang tepat.

Penelitian ini fokus pada analisis performa algoritma K-Nearest Neighbor (KNN) dalam klasifikasi risiko penyakit jantung dengan penerapan SMOTE untuk menangani ketidakseimbangan kelas. Selain itu, penelitian ini juga melakukan optimasi parameter KNN untuk memaksimalkan F1-score, sehingga memberikan gambaran akurat tentang kemampuan model dalam mendeteksi kasus minoritas. Selain itu, penelitian ini membandingkan performa KNN baseline dengan KNN+SMOTE untuk menilai pengaruh teknik resampling terhadap peningkatan klasifikasi [19].

Penelitian ini mengenai pengaruh SMOTE terhadap performa KNN, penyajian perbandingan sistematis antara KNN baseline dan KNN dengan SMOTE, serta optimasi parameter KNN untuk meningkatkan F1-score pada dataset yang tidak seimbang. Selain itu, penelitian ini menjadi referensi bagi pengembangan sistem prediksi penyakit jantung berbasis machine learning yang lebih robust dan sensitif terhadap kasus minoritas [20]. Dengan pendekatan ini, diharapkan penelitian dapat memberikan kontribusi nyata dalam pengembangan sistem prediksi risiko penyakit jantung yang tidak hanya akurat, tetapi juga sensitif terhadap kasus minoritas, sehingga dapat mendukung deteksi dini dan pengambilan keputusan klinis yang lebih efektif.

2. METODOLOGI PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif dengan metode eksperimen untuk mengevaluasi kinerja algoritma machine learning dalam mengklasifikasikan risiko serangan jantung berdasarkan data kesehatan. Pendekatan ini dipilih untuk memperoleh hasil yang objektif dan terukur melalui pengujian model secara sistematis. Algoritma K-Nearest Neighbor (KNN) digunakan sebagai metode utama dalam penelitian, dengan penerapan dua skenario pemodelan, yaitu tanpa penanganan ketidakseimbangan data dan dengan penerapan teknik Synthetic Minority Over-sampling Technique (SMOTE). Setiap model dievaluasi menggunakan beberapa metrik kinerja guna menganalisis efektivitas metode yang digunakan serta dampaknya terhadap kemampuan klasifikasi, khususnya pada kelas minoritas. Untuk rangkaian penelitian tersebut dapat dilihat pada Gambar 1 berikut:



Gambar 1. Diagram Alur Tahapan Penilitian

2.1 Pengumpulan data

Dataset yang digunakan dalam penelitian ini adalah Heart Attack Dataset yang memuat kombinasi fitur numerik dan kategorikal. Dataset tersebut mencerminkan karakteristik demografis serta kondisi medis pasien, meliputi variabel seperti gender, age, body_mass_index, systolic_blood_pressure, dan beberapa penyakit penyerta seperti diabetes, chronic_kidney_disease, serta atrial_fibrillation. Variabel heart_attack digunakan sebagai target klasifikasi. Secara keseluruhan, dataset ini terdiri atas 14 variabel prediktor dan 1 variabel target, dan diperoleh dari platform Kaggle yang disediakan oleh Ryan Whittaker.

Building of Informatics, Technology and Science (BITS)
Volume 7, No 4, March 2026 Page: 2586−2596
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v7i4.9493

Dataset ini terdiri dari N sampel pasien dengan 14 variabel prediktor dan 1 variabel target yaitu *heart_attack*. Variabel target bersifat biner yang menunjukkan apakah pasien mengalami serangan jantung atau tidak. Berdasarkan analisis awal, distribusi kelas pada dataset menunjukkan kondisi ketidakseimbangan kelas, di mana jumlah sampel pada kelas non-serangan jantung lebih dominan dibandingkan kelas serangan jantung. Kondisi ini berpotensi menurunkan performa model klasifikasi sehingga diperlukan teknik penyeimbangan data seperti SMOTE untuk meningkatkan kemampuan model dalam mengenali kelas minoritas.

2.1.1 Struktur dan Komposisi Dataset

Struktur dan komposisi dataset digunakan untuk memberikan gambaran mengenai jenis fitur yang terlibat dalam penelitian ini serta asal-usul masing-masing fitur. Dataset yang digunakan dalam penelitian ini adalah Heart Attack CSV Dataset yang diperoleh dari platform penyedia dataset publik Kaggle, yang diunggah oleh Ryan Whittaker. Dataset ini berisi informasi medis pasien yang berkaitan dengan faktor risiko dan kejadian serangan jantung, meliputi variabel demografis, kondisi fisiologis, riwayat kesehatan, serta gaya hidup pasien. Dataset terdiri dari kombinasi fitur numerik dan kategorikal yang mencerminkan kondisi fisiologis, perilaku harian, dan interaksi gaya hidup responden. Selain fitur asli yang berasal langsung dari data mentah, terdapat pula beberapa fitur hasil rekayasa (*feature engineering*) yang ditambahkan untuk memperkuat representasi pola non-linear yang dibutuhkan oleh model KNN. Tabel 1 menyajikan daftar lengkap fitur yang digunakan, termasuk tipe datanya, deskripsi singkat, dan apakah fitur tersebut merupakan data asli atau hasil rekayasa. Untuk memberikan gambaran mengenai karakteristik data, ringkasan fitur yang digunakan pada Tabel 1 berikut:

Tabel 1. Struktur Dan Komposisi Dataset

Nama Fitur	Tipe Data	Deskripsi
gender	Kategorikal	Jenis kelamin pasien
age	Numerik	Usia pasien dalam tahun
body_mass_index	Numerik	Indeks massa tubuh
smoker	Biner	Status perokok
systolic_blood_pressure	Numerik	Tekanan darah sistolik
hypertension_treated	Biner	Status pengobatan hipertensi
family_history_of_cardiovascular_disease	Biner	Riwayat penyakit kardiovaskular
atrial_fibrillation	Biner	Fibrilasi atrium
chronic_kidney_disease	Biner	Penyakit ginjal kronis
rheumatoid_arthritis	Biner	Rheumatoid arthritis
diabetes	Biner	Status diabetes
chronic_obstructive_pulmonary_disorder	Biner	Penyakit paru obstruktif kronis
forced_expiratory_volume_1	Numerik	Volume ekspirasi dalam 1 detik
time_to_event_or_censoring	Numerik	Waktu hingga kejadian
patient_id	Identifier	ID unik pasien
heart_attack	Target	Variabel target klasifikasi

2.2 Pra-pemrosesan data

Tahap pra-pemrosesan data dilakukan untuk meningkatkan kualitas data sebelum digunakan dalam proses pelatihan model. Proses ini diimplementasikan menggunakan pipeline agar seluruh transformasi data dapat dilakukan secara konsisten pada tahap pelatihan dan pengujian model.

Penanganan nilai hilang dilakukan menggunakan SimpleImputer, dengan strategi *median* untuk fitur numerik dan *most frequent* untuk fitur kategorikal. Pendekatan ini dipilih untuk mempertahankan distribusi data sekaligus mengatasi permasalahan data yang tidak lengkap.

Selanjutnya, fitur kategorikal dikonversi menggunakan One-Hot Encoding dengan parameter handle_unknown="ignore" untuk menghindari kesalahan yang mungkin terjadi ketika muncul kategori baru pada data pengujian. Fitur numerik kemudian distandardisasi menggunakan StandardScaler agar setiap fitur memiliki skala yang seragam dan konsisten. Proses standardisasi ini sangat penting karena algoritma K-Nearest Neighbor bergantung sepenuhnya pada perhitungan jarak antar data, yang secara langsung memengaruhi hasil prediksi. Penerapan standardisasi menggunakan StandardScaler memastikan setiap variabel memberikan kontribusi yang seimbang, sehingga model tidak bias terhadap fitur dengan rentang nilai yang lebih besar. Standardisasi dilakukan dengan pendekatan Z-score, yang memanfaatkan nilai rata-rata (mean) dan simpangan baku (standard deviation), sehingga setiap fitur memiliki peran proporsional dalam proses pengukuran jarak antar titik data. Selain itu, beberapa fitur tambahan dihitung secara cermat untuk memperkaya representasi data masukan, termasuk fitur non-linear yang terbukti memiliki pengaruh signifikan terhadap performa model, berdasarkan analisis permutation importance yang dilakukan sebelumnya. Tahapan pra-pemrosesan ini, sehingga memberikan gambaran menyeluruh mengenai transformasi data sebelum diterapkan pada model KNN, termasuk langkah-langkah pembersihan data, deteksi outlier, dan normalisasi yang mendukung stabilitas hasil prediksi dijabarkan secara rinci pada Tabel 2.

Building of Informatics, Technology and Science (BITS)
Volume 7, No 4, March 2026 Page: 2586−2596
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v7i4.9493

Tabel 2. Tahapan Pra-Pemrosesan dan Tools yang Digunakan

Deskripsi	Komponen / Tools
Analisis awal dataset	Pandas, Matplotlib
Identifikasi tipe fitur (numerik/kategorikal)	select_dtypes
Pemeriksaan missing values	isnull().sum()
Pra-pemrosesan manual (imputasi mean/modus, encoding)	SimpleImputer, pd.get_dummies
Oversampling manual untuk eksplorasi	SMOTE
Pipeline preprocessing (imputasi median, standardisasi, encoding)	SimpleImputer, StandardScaler, OneHotEncoder
Setup cross-validation (5-fold stratified)	StratifiedKFold
Hyperparameter tuning KNN tanpa SMOTE	GridSearchCV + KNeighborsClassifier
Hyperparameter tuning KNN dengan SMOTE	GridSearchCV + SMOTE + ImbPipeline
Evaluasi performa model	accuracy_score, precision_score, recall_score, f1_score, roc_auc_score
Visualisasi hasil	ConfusionMatrixDisplay, roc_curve

2.3 Penanganan Ketidakseimbangan Data dengan SMOTE

Untuk menangani permasalahan ketidakseimbangan kelas pada variabel target, penelitian ini menerapkan Synthetic Minority Oversampling Technique (SMOTE). Metode ini bekerja dengan membangkitkan sampel sintetis pada kelas minoritas melalui proses interpolasi antar data yang saling berdekatan dalam ruang fitur, sehingga jumlah data pada setiap kelas menjadi lebih seimbang.

Pada tahap awal, SMOTE diterapkan secara terpisah untuk memungkinkan pengamatan yang lebih mendalam terhadap perubahan distribusi data. Selanjutnya, teknik ini diintegrasikan ke dalam pipeline menggunakan Imbalanced Pipeline guna memastikan konsistensi penerapan SMOTE selama proses cross-validation dan mencegah terjadinya data leakage.

Hasil analisis sebelum dan sesudah penerapan SMOTE menunjukkan adanya perbaikan pada distribusi kelas target menjadi lebih seimbang. Evaluasi lanjutan terhadap distribusi fitur numerik serta matriks korelasi setelah proses oversampling dilakukan untuk memastikan bahwa pembangkitan data sintetis tidak mengubah karakteristik utama dataset. Secara umum, SMOTE bekerja dengan menentukan sejumlah k-nearest neighbors untuk setiap sampel pada kelas minoritas, kemudian memilih salah satu tetangga terdekat secara acak dan membentuk sampel sintetis melalui proses interpolasi antara kedua titik data tersebut.

2.3.1 Pembangunan *Pipeline* Model *Machine Learning*

Pada tahap ini dibangun enam pipeline model KNN untuk mengevaluasi pengaruh pra-pemrosesan data, tuning hiperparameter, dan penyeimbangan kelas secara sistematis. Pipeline pertama menggunakan KNN dasar sebagai model baseline tanpa modifikasi tambahan. Pipeline kedua mengombinasikan KNN dengan SMOTE untuk mengevaluasi dampak penyeimbangan kelas terhadap performa model.

Pipeline ketiga dan keempat menerapkan tuning hiperparameter menggunakan GridSearchCV, masing-masing tanpa dan dengan SMOTE. Pendekatan ini bertujuan untuk memperoleh konfigurasi KNN terbaik berdasarkan kombinasi nilai n_neighbors, weights, dan distance metric. Selanjutnya, pipeline kelima dan keenam mengintegrasikan pra-pemrosesan terstandarisasi, SMOTE, dan tuning hiperparameter ke dalam satu kesatuan pipeline berbasis scikit-learn dan imbalanced-learn. Struktur pipeline ini memastikan seluruh tahapan pemrosesan data dilakukan secara konsisten pada setiap iterasi pelatihan dan evaluasi model. *Hiperparameter* yang dituning pada setiap model dijelaskan dalam Tabel 3 berikut:

Tabel 3. Ruang Pencarian Hiperparameter KNN

Hiperparameter	Notasi	Nilai yang Diuji
Jumlah tetangga	k	3, 5, 7, 9, 11, 13, 15
Bobot tetangga	w	uniform, distance
Metrik jarak	d	euclidean, Manhattan
Skema validasi	–	Stratified K-Fold (5-fold)
Fungsi evaluasi	–	F1-score

2.4 Optimasi Hyperparameter K-Nearest Neighbors

Proses optimasi hiperparameter pada algoritma K-Nearest Neighbors (KNN) dilaksanakan dengan memanfaatkan GridSearchCV, dengan ruang pencarian yang meliputi parameter n_neighbors, weights, dan metric. F1-score dipilih sebagai metrik evaluasi utama karena dinilai lebih representatif dalam menangani dataset yang memiliki ketidakseimbangan kelas, serta mampu mencerminkan keseimbangan antara nilai precision dan recall, khususnya pada konteks diagnosis medis.

Building of Informatics, Technology and Science (BITS)
Volume 7, No 4, March 2026 Page: 2586−2596
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v7i4.9493

Tahapan evaluasi dilakukan menggunakan metode Stratified 5-Fold Cross-Validation guna menjaga konsistensi proporsi kelas pada setiap fold. Pengujian dilakukan pada dua skenario, yaitu model KNN tanpa penerapan SMOTE dan model KNN yang dikombinasikan dengan SMOTE. Untuk menghindari terjadinya data leakage antara data pelatihan dan validasi, digunakan pendekatan pipeline, sementara proses komputasi dipercepat melalui penerapan pemrosesan paralel.

2.5 Evaluasi dan Perbandingan Performa Model

Evaluasi performa model dilakukan menggunakan beberapa metrik klasifikasi untuk memberikan gambaran yang komprehensif terhadap kualitas prediksi. Metrik pertama yang digunakan adalah akurasi, yang mengukur proporsi prediksi yang benar terhadap seluruh data dan dirumuskan sebagai berikut:

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \tag{1}$$

Selain itu digunakan *F1-score*, yaitu metrik harmonisasi antara *precision* dan *recall* yang sangat berguna ketika terdapat ketidakseimbangan kelas. Rumus *F1-score* adalah:

$$F1 = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision}+\text{Recall}} \tag{2}$$

Selain akurasi dan *F1-score*, penelitian ini juga menggunakan *precision*, *recall*, dan AUC untuk mengevaluasi performa model. Setiap metrik dihitung menggunakan rumus berikut.
Presisi mengukur proporsi prediksi positif yang benar. Dirumuskan sebagai:

$$\text{Precision} = \frac{TP}{TP+FP} \tag{3}$$

Recall menilai sejauh mana model mampu mendeteksi seluruh sampel positif.

$$\text{Recall} = \frac{TP}{TP+FN} \tag{4}$$

Dirumuskan sebagai: AUC merepresentasikan kemampuan model dalam membedakan kelas positif dan negatif pada berbagai threshold keputusan. Secara matematis dituliskan sebagai:

$$\text{AUC} = \int_0^1 \quad TPR(FPR)\, d(FPR) \tag{5}$$

dengan:

$$TPR = \frac{TP}{TP+FN}, FPR = \frac{FP}{FP+TN} \tag{6}$$

ROC–AUC *(Receiver Operating Characteristic – Area Under the Curve)* yang mengukur kemampuan model dalam membedakan kelas positif dan negatif pada berbagai ambang keputusan. AUC menggambarkan luas area di bawah kurva ROC, di mana nilai mendekati 1 menunjukkan bahwa model mampu membedakan kedua kelas dengan sangat baik. ROC–AUC juga lebih informatif dibanding akurasi ketika dataset memiliki distribusi kelas tidak seimbang, karena memperhatikan *trade-off* antara *True Positive Rate* dan *False Positive Rate*.

Kombinasi metrik-metrik tersebut memberikan evaluasi yang lebih robust terhadap performa model, baik dari sisi ketepatan prediksi, kemampuan mengenali kelas minoritas, maupun stabilitas model dalam memisahkan kedua kelas.

3. HASIL DAN PEMBAHASAN

Bagian ini menyajikan rangkaian hasil evaluasi terhadap dua konfigurasi model K-Nearest Neighbor (KNN) yang dikembangkan dalam penelitian ini. Kedua model diuji menggunakan 5-fold Stratified Cross-Validation, dengan stratifikasi yang menjaga proporsi kelas pada setiap fold agar distribusi kelas tetap konsisten. Pendekatan ini menjamin evaluasi yang objektif, robust, dan relevan untuk dataset yang tidak seimbang, seperti kasus risiko serangan jantung.

Secara keseluruhan, kedua model menunjukkan performa yang baik, namun memiliki karakteristik dan keunggulan masing-masing. KNN Tuned (F1-score) menunjukkan peningkatan performa setelah optimasi hyperparameter menggunakan Grid Search, sedangkan penerapan SMOTE pada pipeline menghasilkan peningkatan signifikan dalam mengenali kelas minoritas (pasien berisiko tinggi). Kombinasi Grid Search + SMOTE menunjukkan performa terbaik dalam keseimbangan antara precision dan recall, serta kemampuan diskriminatif model (ROC-AUC). Temuan ini menegaskan pentingnya hyperparameter tuning dan penyeimbangan kelas dalam pengembangan model KNN untuk prediksi risiko serangan jantung.

3.1 Kinerja Model KNN Baseline (KNN Tuned tanpa SMOTE)

Model K-Nearest Neighbors (KNN) baseline dikembangkan melalui proses tuning hiperparameter menggunakan GridSearchCV dengan F1-score sebagai metrik evaluasi utama. Ruang pencarian parameter meliputi n_neighbors dengan nilai [3, 5, 7, 9, 11, 13, 15], weights yang terdiri dari uniform dan distance, serta distance metric berupa

Building of Informatics, Technology and Science (BITS)
Volume 7, No 4, March 2026 Page: 2586−2596
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v7i4.9493

euclidean dan manhattan. Evaluasi model dilakukan menggunakan Stratified 5-Fold Cross-Validation untuk memastikan distribusi kelas target tetap proporsional pada setiap fold, sehingga hasil evaluasi yang diperoleh bersifat stabil dan tidak bias.

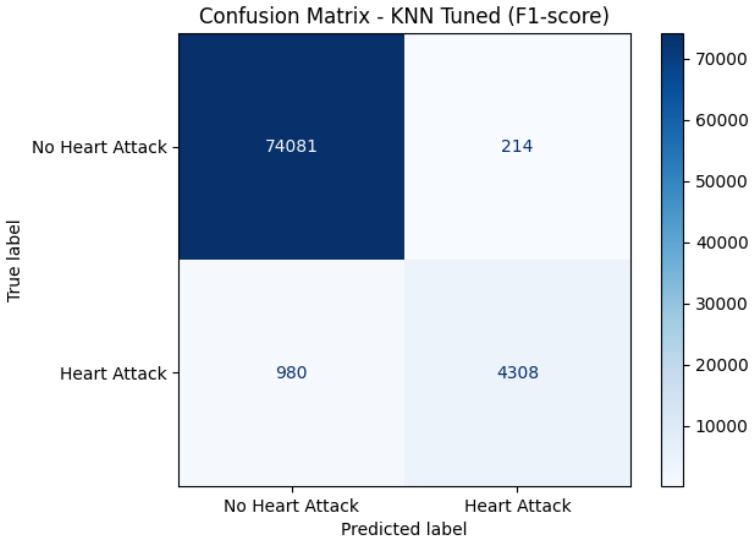
Berdasarkan hasil evaluasi, model KNN baseline menunjukkan performa klasifikasi yang sangat baik secara keseluruhan. Nilai Accuracy sebesar 0,9850 dan ROC-AUC sebesar 0,9278 mengindikasikan bahwa model memiliki kemampuan yang tinggi dalam membedakan kelas heart attack dan no heart attack. Selain itu, nilai Precision sebesar 0,9527 menunjukkan bahwa sebagian besar prediksi positif yang dihasilkan model merupakan prediksi yang benar, sehingga tingkat kesalahan false positive dapat ditekan, sebagaimana dapat dilihat pada Tabel 4.

Namun demikian, nilai Recall sebesar 0,8147 menunjukkan bahwa masih terdapat sejumlah kasus serangan jantung yang tidak teridentifikasi oleh model. Kondisi ini mengindikasikan keterbatasan model KNN baseline dalam mendeteksi kelas minoritas akibat ketidakseimbangan distribusi kelas pada dataset. Nilai F1-score sebesar 0,8783 mencerminkan keseimbangan yang cukup baik antara precision dan recall, meskipun peningkatan sensitivitas terhadap kasus berisiko tinggi masih diperlukan, sebagaimana ditunjukkan pada Tabel 4 berikut :

Tabel 4. Hasil Evaluasi KNN Tuned (F1-score)

Metrik	Nilai
Accuracy	0.9850
Precision	0.9527
Recall	0.8147
F1-Score	0.8783
ROC-AUC	0.9278

Kinerja model KNN baseline yang telah dioptimasi melalui tuning hiperparameter dengan F1-score sebagai metrik evaluasi utama. Hasil evaluasi menunjukkan bahwa model memiliki tingkat akurasi dan kemampuan diskriminasi kelas yang sangat baik, tercermin dari nilai accuracy dan ROC-AUC yang tinggi. Nilai precision yang besar mengindikasikan bahwa prediksi positif yang dihasilkan model cenderung tepat, sehingga tingkat kesalahan false positive relatif rendah. Namun, capaian nilai recall yang masih terbatas menunjukkan bahwa sebagian kasus serangan jantung belum berhasil teridentifikasi oleh model. Kondisi ini menegaskan bahwa meskipun performa model secara umum sudah kuat, kemampuan deteksi terhadap kelas minoritas masih belum optimal, sehingga diperlukan pendekatan lanjutan untuk meningkatkan sensitivitas model dapat dilihat pada Gambar 2.



Gambar 2. Confusion matrix KNN Tuned (F1-score)

Confusion matrix dari model KNN baseline yang telah dituning. Model berhasil mengklasifikasikan 74.081 data No Heart Attack secara benar sebagai true negative, dengan tingkat kesalahan yang relatif kecil pada kelas mayoritas. Pada kelas Heart Attack, model mampu mengidentifikasi 4.308 data sebagai true positive, namun masih terdapat 980 kasus false negative. Kondisi ini menunjukkan bahwa model lebih optimal dalam mengenali kelas mayoritas dibandingkan kelas minoritas. Sehingga penanganan ketidakseimbangan kelas menjadi langkah penting pada tahap pengembangan berikutnya.

3.2 Kinerja Model KNN + SMOTE

Model K-Nearest Neighbors (KNN) yang diintegrasikan dengan Synthetic Minority Over-sampling Technique (SMOTE) dikembangkan menggunakan ImbPipeline sebagai pendekatan yang lebih komprehensif dalam menangani permasalahan ketidakseimbangan kelas. SMOTE diterapkan pada tahap prapemrosesan untuk menghasilkan data

Building of Informatics, Technology and Science (BITS)
Volume 7, No 4, March 2026 Page: 2586−2596
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v7i4.9493

sintetis pada kelas minoritas, sehingga representasi kelas heart attack menjadi lebih seimbang. Integrasi SMOTE ke dalam pipeline memastikan bahwa proses oversampling hanya dilakukan pada data pelatihan di setiap fold selama cross-validation, sehingga potensi terjadinya data leakage yang dapat menyebabkan hasil evaluasi terlalu optimistis dapat dihindari.

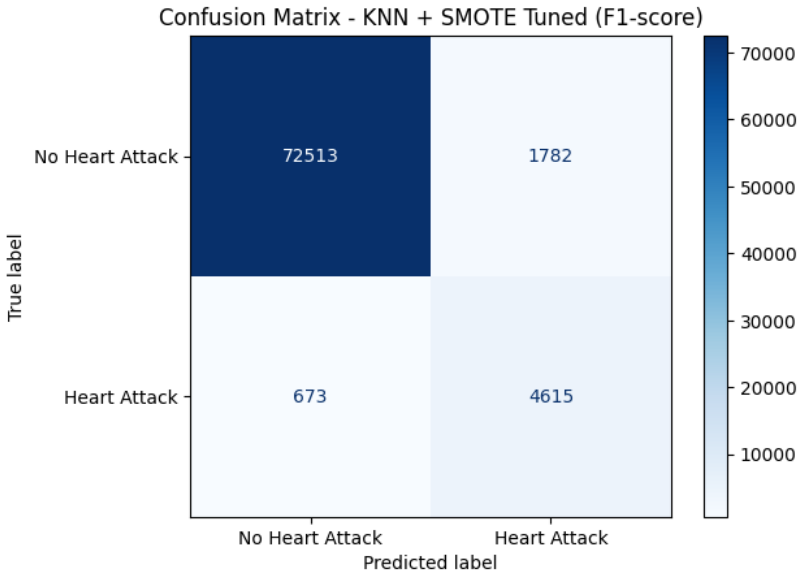
Proses tuning hiperparameter pada model KNN + SMOTE dilakukan menggunakan ruang pencarian parameter yang sama dengan model KNN baseline, guna menjamin perbandingan performa yang adil (fair comparison). Pemanfaatan ImbPipeline memungkinkan integrasi yang mulus antara tahap prapemrosesan, SMOTE, dan algoritma KNN dalam satu alur kerja terpadu, yang selanjutnya dapat dioptimalkan secara sistematis menggunakan GridSearchCV dengan F1-score sebagai metrik evaluasi utama.

Hasil evaluasi performa model KNN + SMOTE Tuned disajikan pada Tabel 5, yang mencakup metrik Accuracy, Precision, Recall, F1-score, dan ROC-AUC. Secara umum, penerapan SMOTE diharapkan dapat meningkatkan nilai recall dan F1-score dibandingkan model baseline, karena model memperoleh kemampuan yang lebih baik dalam mengenali kelas minoritas, sebagaimana dapat dilihat pada Tabel 5 berikut:

Tabel 5. Hasil Evaluasi KNN Tuned (F1-score)

Metrik	Nilai
Accuracy	0.9692
Precision	0.7215
Recall	0.8727
F1-Score	0.7899
ROC-AUC	0.9484

Hasil evaluasi performa model KNN setelah diterapkan teknik SMOTE dan dilakukan tuning hiperparameter menggunakan F1-score sebagai metrik utama. Berbeda dengan model baseline, integrasi SMOTE menghasilkan peningkatan yang nyata pada kemampuan model dalam mengenali kasus serangan jantung, yang tercermin dari nilai *recall* yang lebih tinggi. Peningkatan sensitivitas ini menunjukkan bahwa model menjadi lebih efektif dalam menangkap pola pada kelas minoritas. Di sisi lain, terjadi penurunan nilai *precision* yang menandakan bertambahnya prediksi positif yang tidak sepenuhnya akurat, sebagai konsekuensi dari proses oversampling. Meskipun demikian, keseimbangan antara *precision* dan *recall* yang ditunjukkan oleh nilai *F1-score* mengindikasikan bahwa penerapan SMOTE memberikan kontribusi positif terhadap kinerja model secara keseluruhan, khususnya dalam konteks deteksi risiko serangan jantung.



Gambar 3. Confusion Matrix KNN + SMOTE Tuned (F1-score)

Gambar 3 menjelaskan confusion matrix dari model KNN + SMOTE yang telah melalui proses tuning hiperparameter. Dibandingkan dengan model KNN baseline, jumlah false negative pada kelas Heart Attack mengalami penurunan, yang menunjukkan peningkatan kemampuan model dalam mendeteksi pasien berisiko tinggi. Distribusi prediksi yang lebih seimbang antara kelas mayoritas dan minoritas mencerminkan efektivitas penerapan SMOTE dalam mengatasi ketidakseimbangan kelas.

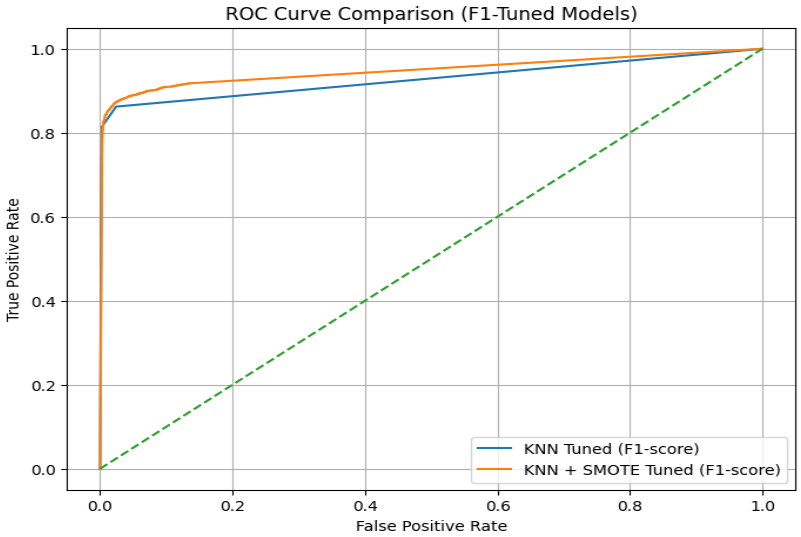
3.3 Analisis ROC Curve dan Diskriminasi Model

Analisis ROC Curve digunakan sebagai pendekatan visual yang efektif untuk membandingkan kemampuan diskriminasi kedua model pada berbagai nilai ambang (threshold) probabilitas. Kurva ROC menggambarkan

Building of Informatics, Technology and Science (BITS)
Volume 7, No 4, March 2026 Page: 2586−2596
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v7i4.9493

hubungan antara True Positive Rate (sensitivitas) dan False Positive Rate (1–spesifisitas), sehingga memungkinkan evaluasi kinerja model pada berbagai skenario klasifikasi. Dalam konteks diagnosis medis, analisis ini memiliki peran penting karena penentuan threshold yang optimal dapat disesuaikan dengan kebutuhan klinis, misalnya untuk memprioritaskan deteksi dini kasus berisiko tinggi atau menekan tingkat kesalahan prediksi positif.

Selain evaluasi visual, nilai Area Under the Curve (AUC) digunakan sebagai metrik tunggal yang merangkum kemampuan diskriminatif model secara keseluruhan. Perbandingan nilai AUC antara model KNN baseline dan model KNN + SMOTE memberikan gambaran mengenai kemampuan relatif masing-masing model dalam membedakan pasien dengan risiko tinggi dan risiko rendah terhadap serangan jantung pada seluruh kemungkinan threshold klasifikasi.



Gambar 4. ROC Curve Comparison (F1-Tuned Models)

Pada Gambar 4 perbandingan kurva ROC antara model KNN Tuned (F1-score) dan KNN + SMOTE Tuned (F1-score). Kedua kurva berada jauh di atas garis diagonal acak, yang menunjukkan bahwa kedua model memiliki kemampuan diskriminasi yang sangat baik dalam membedakan kelas *heart attack* dan *no heart attack*. Kurva ROC model KNN + SMOTE tampak berada sedikit di atas kurva KNN baseline pada sebagian besar rentang *False Positive Rate*, yang mengindikasikan peningkatan *True Positive Rate* pada berbagai threshold klasifikasi.

Perbedaan posisi kurva tersebut mencerminkan bahwa integrasi SMOTE memberikan kontribusi positif terhadap sensitivitas model, terutama dalam mendeteksi kasus serangan jantung pada kelas minoritas. Hal ini sejalan dengan peningkatan nilai *recall* dan *ROC-AUC* pada model KNN + SMOTE dibandingkan model tanpa SMOTE. Dengan demikian, hasil visualisasi ROC ini menegaskan bahwa penerapan SMOTE mampu meningkatkan kemampuan diskriminatif model secara keseluruhan tanpa mengorbankan stabilitas performa secara signifikan.

3.4. Interpretasi dan Pembahasan Komparatif

Penerapan Synthetic Minority Over-sampling Technique (SMOTE) memberikan dampak yang signifikan terhadap peningkatan performa model, khususnya pada metrik recall, yang meningkat dari 0,8147 menjadi 0,8727 atau sebesar 7,12%. Metrik recall ini menjadi indikator yang paling krusial dalam konteks prediksi risiko serangan jantung, karena secara langsung berkaitan dengan kemampuan model dalam mengidentifikasi pasien yang benar-benar berada pada kondisi berisiko tinggi. Peningkatan nilai recall tersebut menunjukkan bahwa model K-Nearest Neighbor (KNN) yang diintegrasikan dengan SMOTE mampu mendeteksi kasus heart attack secara lebih akurat, sehingga secara signifikan menurunkan jumlah false negative, yang dalam praktik klinis dapat menimbulkan konsekuensi fatal bagi pasien. Proses pembangkitan data sintetis melalui SMOTE memungkinkan model untuk mempelajari batas keputusan (decision boundary) yang lebih representatif bagi kelas minoritas, sekaligus memperkuat generalisasi model terhadap data baru. Selain itu, penerapan SMOTE membantu mengurangi bias terhadap kelas mayoritas dan meningkatkan keseimbangan distribusi kelas, sehingga keseluruhan performa model menjadi lebih stabil dan andal, sebagaimana ditunjukkan pada Tabel 6.

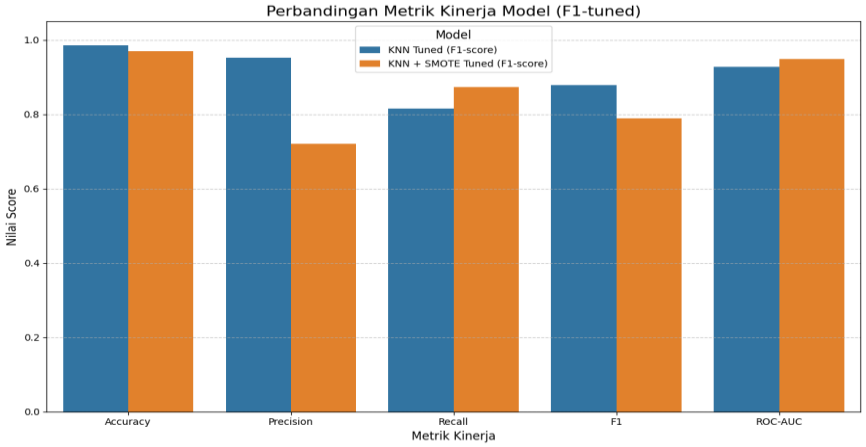
Tabel 6. Perbandingan Performa Model KNN (F1-Tuned)

Metrik	KNN Tuned	KNN + SMOTE	Selisih	Perubahan (%)	Interpretasi Klinis
Accuracy	0.9850	0.9692	-0.0158	-1.60%	Slight overall decrease
Precision	0.9527	0.7215	-0.2312	-24.27%	Increase in false positives
Recall	0.8147	0.8727	+0.0580	+7.12%	Improved heart attack detection
F1-Score	0.8783	0.7899	-0.0884	-10.06%	Balanced metric decline

Building of Informatics, Technology and Science (BITS)
Volume 7, No 4, March 2026 Page: 2586−2596
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v7i4.9493

Metrik	KNN Tuned	KNN + SMOTE	Selisih	Perubahan (%)	Interpretasi Klinis
ROC-AUC	0.9278	0.9484	+0.0206	+2.22%	Superior discrimination ability

Perbandingan performa antara model KNN baseline dan model KNN + SMOTE menunjukkan adanya perubahan karakteristik kinerja yang mencerminkan trade-off antar metrik evaluasi. Visualisasi pada Gambar 5 memperlihatkan bahwa penerapan SMOTE meningkatkan nilai recall dan ROC-AUC, namun diikuti dengan penurunan pada metrik precision dan F1-score. Hal ini menunjukkan adanya pergeseran fokus model dari akurasi keseluruhan menuju peningkatan sensitivitas dalam mendeteksi kasus serangan jantung



Gambar 5. Perbandingan Metrik Kinerja Model KNN Tuned dan KNN + SMOTE

Dalam gambar 5 menunjukan meskipun performa antara model KNN baseline dan model KNN + SMOTE menunjukkan adanya perubahan karakteristik kinerja yang mencerminkan trade-off antar metrik evaluasi. Meskipun nilai *accuracy* dan *precision* mengalami penurunan masing-masing sebesar 1,60% dan 24,27%, peningkatan nilai *recall* sebesar 7,12% dan *ROC-AUC* sebesar 2,22% menegaskan bahwa model dengan SMOTE memiliki kemampuan yang lebih baik dalam mendeteksi dan membedakan pasien berisiko tinggi dan rendah terhadap serangan jantung. Penurunan *precision* mengindikasikan meningkatnya jumlah *false positive*, namun kondisi ini masih dapat diterima dalam konteks medis karena konsekuensi dari *false negative* jauh lebih serius dibandingkan dengan *false positive*, yang umumnya hanya berdampak pada kebutuhan pemeriksaan lanjutan, sebagaimana dapat dilihat pada Tabel 6 dan Tabel 7.

Tabel 7. Analisis Performance Trade-off

Aspek	KNN Tuned	KNN + SMOTE	Dampak Klinis
Sensitivity (Recall)	81.47%	87.27%	5.8% lebih banyak kasus terdeteksi
Specificity	Sangat tinggi	Moderat	Lebih banyak alarm, masih dapat diterima
False Negative Rate	18.53%	12.73%	Penurunan 31% kasus terlewat
Clinical Safety	Baik	Sangat Baik	Perlindungan pasien lebih optimal
Aspek	KNN Tuned	KNN + SMOTE	Dampak Klinis

Analisis lebih lanjut menunjukkan bahwa peningkatan sensitivitas model berdampak langsung pada penurunan *false negative rate* dari 18,53% menjadi 12,73%, atau berkurang sekitar 31%. Penurunan ini mencerminkan peningkatan tingkat keselamatan pasien karena lebih sedikit kasus serangan jantung yang terlewat dalam proses klasifikasi. Meskipun tingkat spesifisitas menurun secara moderat akibat peningkatan *false positive*, kondisi tersebut masih berada dalam batas yang dapat diterima secara klinis, mengingat tujuan utama sistem prediksi ini adalah meminimalkan risiko kegagalan deteksi pada kondisi yang mengancam jiwa, sebagaimana ditunjukkan pada Tabel 7.

Dari sisi kemampuan diskriminasi, nilai *Area Under the Curve* (AUC) yang lebih tinggi pada model KNN + SMOTE, yaitu sebesar 0,9484 dibandingkan 0,9278 pada model baseline, menunjukkan keunggulan model dalam membedakan pasien dengan risiko tinggi dan rendah pada berbagai nilai ambang probabilitas. Peningkatan AUC sebesar 2,22% memberikan fleksibilitas yang lebih besar bagi praktisi medis dalam menentukan nilai *cut-off* klasifikasi sesuai dengan konteks klinis dan tingkat toleransi risiko yang diinginkan. Model dengan nilai AUC yang tinggi juga menunjukkan *robustness* yang lebih baik dalam melakukan pemeringkatan pasien berdasarkan tingkat risiko serangan jantung, sebagaimana dirangkum pada Tabel 8.

Tabel 8. Interpretasi ROC-AUC dan Signifikansi Klinis

Model	AUC	Kategori Performa	Interpretasi Klinis
KNN Tuned	0.9278	Excellent	Discriminasi sangat baik
KNN + SMOTE	0.9484	Excellent+	Discriminasi superior
Improvement	+0.0206	–	Peningkatan stratifikasi risiko

Building of Informatics, Technology and Science (BITS)
Volume 7, No 4, March 2026 Page: 2586−2596
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v7i4.9493

Meskipun nilai *F1-score* pada model KNN + SMOTE mengalami penurunan sebesar 10,06% dibandingkan model baseline, penurunan ini merupakan konsekuensi langsung dari trade-off antara *precision* dan *recall* yang mengutamakan peningkatan sensitivitas. Dalam aplikasi *medical screening*, prioritas terhadap *recall* dinilai lebih penting dibandingkan keseimbangan metrik secara matematis, terutama untuk penyakit dengan tingkat fatalitas tinggi seperti serangan jantung. Oleh karena itu, penurunan *F1-score* tidak serta-merta menunjukkan penurunan kualitas model secara klinis.

Berdasarkan analisis risiko dan manfaat, model KNN + SMOTE Tuned memberikan nilai klinis yang lebih unggul dibandingkan model KNN baseline. Model ini mampu menurunkan jumlah kasus serangan jantung yang terlewat secara signifikan, meningkatkan keselamatan pasien, serta mengurangi potensi risiko hukum akibat kegagalan diagnosis. Meskipun terjadi peningkatan jumlah *false positive*, dampak klinis dan biaya yang ditimbulkan relatif lebih rendah dibandingkan risiko fatal akibat *false negative*. Oleh karena itu, model KNN + SMOTE Tuned direkomendasikan untuk implementasi klinis, dengan strategi penyesuaian nilai threshold guna mengoptimalkan keseimbangan antara *precision* dan *recall* sesuai dengan kebutuhan klinis dan ketersediaan sumber daya, sebagaimana dirangkum pada Tabel 9.

Penelitian ini juga sejalan dengan beberapa penelitian terdahulu yang menunjukkan bahwa penerapan teknik penyeimbangan kelas dapat meningkatkan kemampuan model dalam mendeteksi kelas minoritas pada dataset medis. Penelitian sebelumnya melaporkan bahwa penggunaan SMOTE pada dataset penyakit jantung mampu meningkatkan nilai recall dan sensitivitas model, meskipun sering kali diikuti dengan penurunan nilai precision akibat bertambahnya prediksi positif. Temuan serupa juga dilaporkan pada penelitian yang menerapkan algoritma K-Nearest Neighbor pada data medis tidak seimbang, di mana integrasi teknik oversampling seperti SMOTE membantu model dalam mempelajari pola pada kelas minoritas secara lebih representatif.

Dibandingkan dengan penelitian-penelitian tersebut, hasil penelitian ini menunjukkan kecenderungan yang konsisten, yaitu adanya peningkatan kemampuan deteksi kasus serangan jantung setelah penerapan SMOTE. Peningkatan nilai recall sebesar 7,12% serta peningkatan ROC-AUC sebesar 2,22% menunjukkan bahwa metode yang digunakan mampu meningkatkan sensitivitas model dalam mengenali pasien berisiko tinggi. Hal ini memperkuat temuan sebelumnya bahwa penanganan ketidakseimbangan kelas merupakan faktor penting dalam pengembangan model prediksi pada bidang kesehatan, terutama untuk kasus penyakit dengan risiko fatal seperti serangan jantung.

Dengan demikian, penelitian ini tidak hanya mengonfirmasi efektivitas teknik SMOTE dalam meningkatkan performa model pada dataset medis yang tidak seimbang, tetapi juga memberikan kontribusi tambahan melalui analisis komparatif antara model KNN baseline dan model KNN yang dioptimalkan menggunakan SMOTE serta tuning hiperparameter berbasis F1-score.

Tabel 9. Clinical Risk Assessment dan Cost–Benefit Analysis

Jenis Kesalahan	Dampak KNN Tuned	Dampak KNN + SMOTE	Biaya Klinis
False Negative	~18.5% kasus terlewat	~12.7% kasus terlewat	Sangat tinggi (potensi fatal)
False Positive	Sangat rendah	Meningkat moderat	Rendah (tes tambahan)
Patient Safety	Baik	Lebih unggul	–
Legal Risk	Moderat	Lebih rendah	Risiko gugatan tinggi
Jenis Kesalahan	Dampak KNN Tuned	Dampak KNN + SMOTE	Biaya Klinis

4. KESIMPULAN

Berdasarkan evaluasi menyeluruh terhadap kinerja algoritma K-Nearest Neighbor yang dioptimalkan menggunakan F1-score dan dikombinasikan dengan teknik SMOTE dalam klasifikasi risiko serangan jantung, hasil penelitian ini menunjukkan bahwa penerapan SMOTE memberikan pengaruh yang signifikan terhadap peningkatan kemampuan deteksi kasus *heart attack*. Model KNN + SMOTE Tuned berhasil mencapai nilai *recall* sebesar 87,27%, lebih tinggi dibandingkan model KNN baseline yang memperoleh 81,47%, sehingga menghasilkan peningkatan deteksi kasus positif sebesar 7,12% serta penurunan jumlah diagnosis yang terlewat (*false negative*) hingga 31%. Meskipun terjadi konsekuensi berupa penurunan nilai *precision* dari 95,27% menjadi 72,15%, peningkatan nilai *ROC-AUC* dari 0,9278 menjadi 0,9484 (kenaikan 2,22%) mengindikasikan kemampuan diskriminatif model yang lebih unggul. Dalam konteks medis, model KNN yang terintegrasi dengan SMOTE direkomendasikan karena mengutamakan tingkat sensitivitas yang tinggi, yang lebih krusial dalam mencegah kesalahan diagnosis fatal, sementara proses pembangkitan data sintetis terbukti efektif dalam mengatasi permasalahan ketidakseimbangan kelas pada dataset serangan jantung.

REFERENCES

[1] S. Wan, F. Wan, X. D.-A. of C. Diseases, and undefined 2025, "Machine learning approaches for cardiovascular disease prediction: A review," *ElsevierS Wan, F Wan, X DaiArchives of Cardiovascular Diseases, 2025•Elsevier*, 2025, doi: 10.1016/j.acvd.2025.04.055.

[2] I. Akbar, F. Supriadi, and D. I. Junaedi, "Pemanfaatan Machine Learning Di Bidang Kesehatan," *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 9, no. 1, pp. 1744–1749, Jan. 2025, doi: 10.36040/jati.v9i1.12663.

Building of Informatics, Technology and Science (BITS)
Volume 7, No 4, March 2026 Page: 2586−2596
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v7i4.9493

[3] J. Yang and J. Guan, "A Heart Disease Prediction Model Based on Feature Optimization and Smote-Xgboost Algorithm," *Information 2022, Vol. 13,* vol. 13, no. 10, Oct. 2022, doi: 10.3390/info13100475.
[4] M. S. Simanjuntak, R. Robet, and L. Hoki, "A Comparative Study of Machine Learning and Deep Learning Models for Heart Disease Classification," *Journal of Applied Informatics and Computing*, vol. 9, no. 6, pp. 3405–3409, Dec. 2025, doi: 10.30871/jaic.v9i6.11546.
[5] N. Tyagi and P. Jain, "A Review of Machine Learning Algorithms for Predicting Heart Disease," *2024 2nd International Conference on Disruptive Technologies, ICDT 2024*, pp. 961–965, 2024, doi: 10.1109/ICDT61202.2024.10488917.
[6] A. K. Yadav, R. Shukla, and T. R. Singh, "Machine learning in expert systems for disease diagnostics in human healthcare," *Machine Learning, Big Data, and IoT for Medical Informatics*, pp. 179–200, Jan. 2021, doi: 10.1016/B978-0-12-821777-1.00022-7.
[7] C. Boukhatem, H. Y. Youssef, and A. B. Nassif, "Heart Disease Prediction Using Machine Learning," *2022 Advances in Science and Engineering Technology International Conferences, ASET 2022*, 2022, doi: 10.1109/ASET53988.2022.9734880.
[8] A. Ishaq *et al.*, "Improving the Prediction of Heart Failure Patients' Survival Using SMOTE and Effective Data Mining Techniques," *IEEE Access*, vol. 9, pp. 39707–39716, 2021, doi: 10.1109/ACCESS.2021.3064084.
[9] N. Nasution, F. B. Nasution, M. A. Hasan, and L. Kuning, "Predicting Heart Disease Using Machine Learning: An Evaluation of Logistic Regression, Random Forest, SVM, and KNN Models on the UCI Heart Disease Dataset," *IT Journal Research and Development*, vol. 9, no. 2, pp. 140–150, Apr. 2025, doi: 10.25299/itjrd.2025.17941.
[10] E. Richardson, R. Trevizani, J. A. Greenbaum, H. Carter, M. Nielsen, and B. Peters, "The receiver operating characteristic curve accurately assesses imbalanced datasets," *Patterns*, vol. 5, no. 6, p. 100994, Jun. 2024, doi: 10.1016/j.patter.2024.100994.
[11] D. Chicco and G. Jurman, "The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation," *BMC Genomics*, vol. 21, no. 1, Jan. 2020, doi: 10.1186/s12864-019-6413-7.
[12] S. S. Yadav* and G. P. Bhole, "Learning from Imbalanced Data in Classification," *International Journal of Recent Technology and Engineering (IJRTE)*, vol. 8, no. 5, pp. 1907–1016, Jan. 2020, doi: 10.35940/ijrte.e6286.018520.
[13] A. H. Shaker, I. A. Ibrahim, and S. K. Gharghan, "Machine learning techniques for cardiovascular disease detection through heart sound analysis: A review," *AIP Conf. Proc.*, vol. 3232, no. 1, Oct. 2024, doi: 10.1063/5.0236263.
[14] S. P. Aulia, B. Rahmat, and A. Junaidi, "Enhancing Heart Disease Prediction through SMOTE-ENN Balancing and RFECV Feature Selection," *Journal of Artificial Intelligence and Engineering Applications (JAIEA)*, vol. 4, no. 3, pp. 1968–1973, Jun. 2025, doi: 10.59934/jaiea.v4i3.1057.
[15] M. Kavitha, G. Gnaneswar, R. Dinesh, Y. R. Sai, and R. S. Suraj, "Heart Disease Prediction using Hybrid machine Learning Model," *Proceedings of the 6th International Conference on Inventive Computation Technologies, ICICT 2021*, pp. 1329–1333, Jan. 2021, doi: 10.1109/ICICT50816.2021.9358597.
[16] M. Rahardi, B. P. Asaddulloh, A. Aminuddin, F. F. Abdulloh, I. Saifudin, and F. P. Kusumawijaya, "Optimizing Machine Learning Models for Class Imbalance in Heart Disease Prediction," *Engineering, Technology & Applied Science Research*, vol. 15, no. 3, pp. 23599–23604, Jun. 2025, doi: 10.48084/etasr.10407.
[17] J. J. Wibowo, D. A. Kristiyanti, and J. Wiratama, "Enhancing Heart Disease Classification: A Comparative Analysis of SMOTE and Naïve Bayes on Imbalanced Data," *JOIV : International Journal on Informatics Visualization*, vol. 9, no. 5, pp. 2072–2079, Sep. 2025, doi: 10.62527/joiv.9.5.3248.
[18] S. Akinola, R. Leelakrishna, and V. Varadarajan, "Enhancing cardiovascular disease prediction: A hybrid machine learning approach integrating oversampling and adaptive boosting techniques," *AIMS Med. Sci.*, vol. 11, no. 2, pp. 58–71, 2024, doi: 10.3934/medsci.2024005.
[19] S. Gupta, A. Tripathi, and C. Srivastava, "Machine Learning Models for Heart Disease Prediction: Balancing Accuracy and Transparency," *2nd IEEE International Conference on IoT, Communication and Automation Technology, ICICAT 2024*, pp. 1605–1611, 2024, doi: 10.1109/ICICAT62666.2024.10922965.
[20] D. Purwanto, S. C. Hidayati, D. I. Ricoida, and K. A. Putri, "Optimized Machine Learning Models for Heart Disease Prediction: A Performance Analysis," *Proceedings - 2024 International of Seminar on Application for Technology of Information and Communication: Smart And Emerging Technology for a Better Life, iSemantic 2024*, pp. 559–562, 2024, doi: 10.1109/iSemantic63362.2024.10762500.