E-ISSN: 3025-5481
DINASTI INFORMATION AND TECHNOLOGY
DINASTI RESEARCH
https://dinastires.org/DIT dinasti.info@gmail.com +62 811 7404 455
DOI: https://doi.org/10.38035/dit.v4i1

Analisis Perbandingan Imputasi Mean, Median, dan KNN pada Decision Tree Berbasis Entropi untuk Penanganan Nilai Hilang dalam Prediksi Komplikasi Infark Miokard

Muhammad Rizal Arifin[1], Zainal Abidin[2]
[1]Program Studi Informatika, Sekolah Tinggi Teknik Pati, Pati, Indonesia, ra6478428@gmail.com
[2]Program Studi Informatika, Sekolah Tinggi Teknik Pati, Pati, Indonesia

Corresponding Author: ra6478428@gmail.com[1]

Abstract: *Clinical datasets often contain missing values that can affect the reliability of prediction models. This study compares mean, median, and KNN imputation in handling missing values in the Myocardial Infarction Complications dataset. The initial dataset contained 1,700 patient records, 111 input features, and LET_IS as the prediction target. After converting missing value symbols, cleaning columns with high missing values, and removing rows that could not be used for the research target, the data processed in the train-test split became 1,350 records, consisting of 1,080 training records and 270 test records. Each imputation scenario was evaluated using an Entropy-based Decision Tree as a practical approach to the C4.5 splitting principle. The latest Google Colab output in the single-split scenario showed that median imputation produced the highest accuracy of 89.63%, followed by KNN at 88.52% and mean at 88.15%. The main recommendation is median because it produced the highest accuracy in the main testing scenario.*

Keyword: *Entropy-Based Decision Tree, Data Imputation, Myocardial Infarction, KNN Imputer, Missing Values*

Abstrak: Dataset klinis sering memuat nilai hilang yang dapat memengaruhi keandalan model prediksi. Penelitian ini membandingkan imputasi mean, median, dan KNN dalam menangani nilai hilang pada dataset Myocardial Infarction Complications. Data awal terdiri atas 1.700 rekam pasien, 111 fitur masukan, dan LET_IS sebagai target prediksi. Setelah simbol nilai hilang dikonversi, kolom dengan nilai hilang tinggi dibersihkan, dan baris yang tidak dapat digunakan pada target penelitian dihapus, data yang diproses pada tahap train-test split menjadi 1.350 rekam, yaitu 1.080 data latih dan 270 data uji. Setiap skenario imputasi dievaluasi menggunakan Decision Tree berbasis Entropi sebagai pendekatan praktis terhadap prinsip pemisahan C4.5. Hasil Google Colab terbaru pada skenario single-split menunjukkan bahwa imputasi median menghasilkan akurasi tertinggi sebesar 89,63%, diikuti KNN sebesar 88,52% dan mean sebesar 88,15%. Rekomendasi utama penelitian ini adalah median karena memberikan akurasi tertinggi pada pengujian utama.

Kata Kunci: Data Imputasi,Decision Tree Berbasis Entropi, Infark Miokard, KNN Imputer, Missing Values

PENDAHULUAN

Infark miokard merupakan kejadian kardiovaskular serius karena dapat menimbulkan komplikasi akut dan kematian dini apabila kondisi pasien tidak diantisipasi dengan tepat. Pengambilan keputusan klinis pada konteks ini bergantung pada banyak variabel, seperti data demografis, riwayat klinis, tanda vital, hasil laboratorium, indikator elektrokardiografi, dan komorbiditas. Berbagai variabel tersebut membuka peluang pemanfaatan data mining dan machine learning untuk mendukung penilaian risiko awal. Namun, dataset klinis jarang berada dalam kondisi lengkap. Nilai hilang sering terjadi karena tidak semua pemeriksaan dilakukan, sebagian observasi tertunda, atau rekam medis tidak tersedia pada waktu yang sama (Liu & others, 2023). Kondisi ini menjadikan prapemrosesan data sebagai tahap penting sebelum model prediksi dibangun (Golovenkin & others, 2020).

Dataset *Myocardial Infarction Complications* dipilih karena merepresentasikan masalah klasifikasi klinis nyata dengan banyak atribut masukan, beberapa luaran komplikasi, dan nilai hilang. Dataset ini memiliki 1.700 rekam pasien, 111 fitur masukan, serta 12 variabel keluaran yang berkaitan dengan komplikasi setelah infark miokard. Repositori dataset menjelaskan bahwa kolom 2 sampai 112 digunakan sebagai masukan, sedangkan kolom 113 sampai 124 merupakan keluaran atau kemungkinan komplikasi (Golovenkin & others, 2020). Struktur tersebut membuat dataset ini sesuai untuk menguji pengaruh metode penanganan nilai hilang terhadap performa klasifikasi.

Penanganan nilai hilang bukan sekadar tahap teknis yang netral. Strategi imputasi yang berbeda dapat mengubah distribusi variabel masukan dan selanjutnya memengaruhi batas keputusan yang dipelajari oleh model klasifikasi. Imputasi sederhana, seperti *mean* dan *median*, mudah diterapkan serta efisien secara komputasi, tetapi keduanya merangkum setiap fitur hanya dengan satu nilai pusat. Imputasi *KNN* menggunakan nilai dari sampel terdekat dan dapat mempertahankan pola lokal ketika struktur ketetanggaan data bermakna. Akan tetapi, *KNN* juga dapat dipengaruhi oleh skala variabel, perhitungan jarak, dimensi data yang tinggi, dan atribut klinis yang berisik. Oleh karena itu, perbandingan ketiga metode tersebut dalam pengaturan klasifikasi yang sama diperlukan untuk memahami apakah metode yang lebih kompleks selalu menghasilkan performa prediksi yang lebih baik.

Model *Decision Tree* banyak digunakan dalam klasifikasi karena strukturnya lebih mudah diinterpretasikan dibandingkan banyak model black-box. Algoritma C4.5 memperkenalkan information gain ratio untuk memilih atribut pemisah dan menjadi salah satu metode pohon keputusan yang berpengaruh (Breiman et al., 1984). Dalam implementasi Python, pengklasifikasi C4.5 murni tidak tersedia pada modul standar scikit-learn. Oleh karena itu, penelitian ini menggunakan *DecisionTreeClassifier* dengan kriteria entropy sebagai *Decision Tree* berbasis Entropi yang mengikuti prinsip information gain pada pemisahan bergaya C4.5. Penjelasan ini penting agar penelitian tidak mengklaim implementasi sebagai C4.5/J48 murni, tetapi tetap konsisten dengan konsep pohon keputusan C4.5 (Powers, 2011).

Permasalahan lain pada dataset adalah ketidakseimbangan kelas target LET_IS. Kelas mayoritas mendominasi dataset, sedangkan beberapa kelas minoritas memiliki jumlah rekam yang sangat sedikit. Dalam kondisi ini, akurasi dapat terlihat tinggi ketika model mampu mengenali kelas mayoritas, tetapi belum tentu mewakili performa pada semua kelas. Oleh sebab itu, akurasi dalam penelitian ini ditempatkan sebagai metrik utama sesuai keluaran *Google Colab*, sedangkan metrik tambahan seperti precision, recall, dan F1-score direkomendasikan untuk pengembangan evaluasi berikutnya. Penjelasan ini diperlukan agar interpretasi hasil tetap proporsional dan tidak berlebihan(Ghafari & others, 2023; Saito & Rehmsmeier, 2015).

Berdasarkan pertimbangan tersebut, penelitian ini bertujuan membandingkan imputasi *mean*, *median*, dan *KNN* dalam menangani nilai hilang untuk memprediksi komplikasi infark miokard menggunakan *Decision Tree* berbasis Entropi. Kebaruan penelitian ini bukan pada pengusulan algoritma baru, melainkan pada penyajian perbandingan terkontrol terhadap tiga skenario imputasi dengan dataset klinis yang sama, target yang sama, pembagian data yang sama, pengklasifikasi yang sama, dan metrik evaluasi yang sama. Hasil penelitian diharapkan membantu peneliti menentukan apakah strategi imputasi sederhana atau strategi berbasis ketetanggaan lebih sesuai ketika membangun model *Decision Tree* berbasis Entropi pada data klinis dengan nilai hilang dan distribusi kelas yang tidak seimbang (Li & others, 2024; Satty et al., 2024).

METODE

Penelitian ini menggunakan desain eksperimen kuantitatif untuk membandingkan performa tiga skenario prapemrosesan data, yaitu imputasi *mean*, imputasi *median*, dan imputasi *KNN*. Ketiga metode tersebut dipilih karena mewakili pendekatan imputasi sederhana berbasis ukuran pemusatan dan pendekatan berbasis kedekatan antarobservasi. Setiap skenario diproses menggunakan prosedur klasifikasi yang sama, mulai dari pembagian data, pelatihan model, hingga evaluasi hasil. Dengan desain ini, perbedaan performa dapat dikaitkan secara lebih langsung dengan metode imputasi yang digunakan, bukan dengan perbedaan model, target, atau skema pengujian. Variabel keluaran yang digunakan dalam penelitian ini adalah LET_IS, yaitu kategori luaran fatal dalam dataset. Satu target dipilih agar eksperimen tetap fokus dan perbandingan antarmetode imputasi tidak bercampur dengan karakteristik luaran komplikasi lain yang mungkin memiliki distribusi dan tingkat kesulitan klasifikasi berbeda (Thygesen & others, 2018).

Dataset yang digunakan dalam penelitian ini diperoleh dari UCI Machine Learning Repository dan sumber data University of Leicester. Dataset tersebut dipilih karena memuat data klinis nyata tentang komplikasi infark miokard dan memiliki karakteristik nilai hilang yang relevan untuk pengujian metode imputasi. File mentah dibaca sebagai berkas comma-separated tanpa header, sehingga penentuan kolom masukan dan keluaran mengikuti struktur asli dataset. Simbol tanda tanya pada data ditafsirkan sebagai nilai hilang dan dikonversi ke format NaN agar dapat diproses menggunakan pustaka Python. Kolom pertama diperlakukan sebagai identitas pasien dan dihapus dari fitur karena tidak merepresentasikan prediktor klinis yang bermakna. Pada tahap pembersihan, kolom dengan nilai hilang lebih dari 50% dihapus dan baris yang tidak dapat digunakan untuk target penelitian dikeluarkan dari proses pemodelan.

Dengan demikian, data awal yang terdiri atas 1.700 rekam disesuaikan menjadi 1.350 rekam yang diproses pada tahap *train-test split*. Proporsi 20% dari data tersebut menghasilkan 270 sampel uji, sedangkan 80% lainnya menjadi 1.080 sampel latih. Penjelasan ini menjawab perbedaan antara jumlah awal 1.700 rekam dan total support 270 yang terlihat pada classification report dari *Google Colab*. Setelah seleksi kolom, matriks masukan tetap mengikuti deskripsi dataset yang mengidentifikasi kolom 2 sampai 112 sebagai variabel masukan dan kolom 113 sampai 124 sebagai keluaran komplikasi (Afkanpour et al., 2024; Golovenkin & others, 2020). Oleh karena itu, tahapan prapemrosesan dalam penelitian ini disusun agar mengikuti struktur dataset asli dan dapat direplikasi pada eksperimen serupa (Luque et al., 2019).

Sumber: Hasil penelitian (2025)

Gambar 1. Alur kerja penelitian untuk membandingkan metode imputasi

Prapemrosesan dilakukan secara hati-hati untuk menghindari kebocoran data dan menjaga validitas evaluasi model. Data terlebih dahulu dibersihkan dengan menghapus kolom yang memiliki nilai hilang tinggi dan baris yang tidak dapat digunakan untuk target penelitian. Setelah proses ini, data yang diproses terdiri atas 1.350 sampel. Data kemudian dibagi menjadi 80% data latih atau 1.080 sampel dan 20% data uji atau 270 sampel sebelum imputasi dilakukan. Data latih digunakan untuk melatih imputer, kemudian statistik imputasi atau informasi ketetanggaan yang dipelajari dari data latih diterapkan pada data uji. Urutan ini penting karena informasi dari data uji tidak boleh memengaruhi parameter prapemrosesan yang dipelajari dari data latih. Pembagian stratified digunakan agar proporsi kelas LET_IS pada data uji tetap konsisten dengan distribusi kelas asli, terutama karena target penelitian memiliki karakteristik kelas yang tidak seimbang (Little & Rubin, 2019; Luque et al., 2019; Thygesen & others, 2018).

Imputasi *mean* mengganti setiap nilai hilang dengan rata-rata aritmetika dari fitur terkait yang dihitung dari data latih. Metode ini sederhana dan sering digunakan sebagai baseline untuk fitur klinis numerik. Imputasi *median* mengganti nilai hilang dengan *median* dari fitur. Median lebih tahan terhadap nilai ekstrem dan dapat berguna ketika variabel klinis memiliki distribusi miring. Imputasi *KNN* memperkirakan nilai hilang menggunakan nilai rata-rata dari tetangga terdekat berdasarkan ukuran jarak yang dapat menangani ketidakhadiran nilai. Dalam eksperimen ini, jumlah tetangga ditetapkan lima sesuai implementasi umum pada scikit-learn. Ketiga metode tersebut relevan karena strategi imputasi yang berbeda dapat memengaruhi distribusi fitur dan performa model prediksi pada dataset klinis yang tidak lengkap (Breiman et al., 1984; Luque et al., 2019; van Buuren, 2018)

Pengklasifikasi yang digunakan pada semua skenario adalah *Decision Tree* berbasis Entropi. Model diimplementasikan menggunakan kelas *DecisionTreeClassifier* dengan kriteria entropy dan *random_state* tetap untuk meningkatkan reprodusibilitas. Entropy dipilih karena mengukur kualitas pemisahan berdasarkan information gain Shannon. Mekanisme ini secara konseptual dekat dengan pemisahan berbasis informasi pada C4.5, meskipun implementasinya tidak identik dengan algoritma C4.5 asli. Oleh karena itu, model dalam artikel ini disebut *Decision Tree* berbasis Entropi, bukan C4.5 murni (Breiman et al., 1984; Quinlan, 1993).

Model dievaluasi menggunakan akurasi sebagai metrik utama sesuai keluaran eksperimen terbaru dari *Google Colab*oratory. Akurasi mengukur proporsi prediksi benar dari seluruh data uji, sehingga dapat digunakan untuk membandingkan performa umum antar skenario imputasi. Selain itu, precision, recall, dan F1-score dilaporkan sebagai metrik pendukung berdasarkan classification report agar pembaca dapat melihat pengaruh ketidakseimbangan kelas secara lebih proporsional. Metrik pendukung ini penting karena pada dataset yang tidak seimbang, akurasi dapat dipengaruhi oleh dominasi kelas mayoritas dan belum tentu sepenuhnya menggambarkan kemampuan model pada kelas minoritas (Ghafari & others, 2023; Little & Rubin, 2019; Powers, 2011). Untuk menjaga konsistensi laporan, tabel hasil, pembahasan, dan kesimpulan menggunakan nilai terbaru dari keluaran *Google Colab*.

Eksperimen komputasional dilakukan menggunakan Python melalui *Google Colab*oratory. Paket utama yang digunakan adalah pandas untuk manipulasi data, scikit-learn untuk prapemrosesan, pemodelan *Decision Tree* berbasis Entropi dan perhitungan metrik, serta NumPy untuk pemrosesan numerik. Penggunaan ekosistem Python sumber terbuka mendukung reprodusibilitas karena prosedur dapat diulang menggunakan dataset, target, pembagian data, pengaturan imputasi, dan *random_state* yang sama (Harris & others, 2020; Pedregosa & others, 2011).

Algoritma 1. Prosedur eksperimen untuk membandingkan metode imputasi

Input: Dataset Myocardial Infarction Complications / Output: Akurasi untuk imputasi mean, median, dan KNN

1. Memuat MI.data dan mengonversi simbol ? menjadi nilai hilang.
2. Menghapus kolom ID dan memilih kolom 2 sampai 112 sebagai fitur masukan.
3. Memilih LET_IS sebagai target keluaran.
4. Membersihkan kolom dengan nilai hilang lebih dari 50% dan baris yang tidak dapat digunakan untuk target penelitian, kemudian membagi data menjadi 80% data latih atau 1.080 sampel dan 20% data uji atau 270 sampel menggunakan stratifikasi.
5. Untuk setiap metode imputasi dalam {mean, median, KNN}:
   a. Melatih imputer hanya menggunakan fitur data latih.
   b. Mentransformasi fitur data latih dan data uji menggunakan imputer.
   c. Melatih pengklasifikasi Decision Tree berbasis Entropi.
   d. Memprediksi label pada data uji.
   e. Menghitung akurasi dengan membandingkan label aktual dan label prediksi.
6. Membandingkan nilai akurasi dan menentukan metode imputasi terbaik berdasarkan keluaran Google Colab terbaru.

Ringkasan prosedur pada Algoritma 1 menjadi dasar penyusunan deskripsi dataset dan skenario eksperimen. Informasi utama tentang dataset, target, hasil pembersihan data, dan pembagian data ditunjukkan pada Tabel 1.

Tabel 1. Deskripsi dataset yang digunakan dalam penelitian

Deskripsi	Nilai
Sumber dataset	UCI Machine Learning Repository
Nama dataset	Myocardial Infarction Complications
Data awal	1.700
Kolom mentah	124
Kolom setelah pembersihan	117
Fitur masukan	111
Keluaran tersedia	12
Target penelitian	LET_IS
Nilai hilang awal	15.974
Persentase nilai hilang awal	7,58%
Nilai hilang setelah penambahan 20%	54.973
Data yang diproses	1.350
Data latih	1.080 sampel atau 80%
Data uji	270 sampel atau 20%
Pembagian train-test	80% : 20% stratified

Berdasarkan deskripsi dataset pada Tabel 1, eksperimen kemudian disusun ke dalam tiga skenario imputasi. Ketiga skenario menggunakan target, pembagian data, dan pengklasifikasi yang sama sehingga perbedaan hasil dapat ditelusuri pada metode imputasi yang digunakan. Rincian skenario eksperimen dan parameter implementasi disajikan pada Tabel 2.

Tabel 2. Skenario eksperimen dan parameter implementasi

Skenario	Metode imputasi	Model dan parameter
S1	Imputasi mean	Decision Tree berbasis Entropi, DecisionTreeClassifier, criterion = entropy, random_state = 42
S2	Imputasi median	Decision Tree berbasis Entropi, DecisionTreeClassifier, criterion = entropy, random_state = 42
S3	Imputasi KNN	KNNImputer, n_neighbors = 5; Decision Tree berbasis Entropi, DecisionTreeClassifier, criterion = entropy, random_state = 42

HASIL DAN PEMBAHASAN

Pemeriksaan awal terhadap dataset menunjukkan bahwa terdapat 15.974 nilai hilang dari total 210.800 sel data. Jumlah ini setara dengan sekitar 7,58% missingness. Secara umum, persentase tersebut dapat dianggap tidak terlalu besar apabila dilihat dari keseluruhan dataset. Namun, nilai ini tetap perlu diperhatikan karena nilai hilang tidak terdistribusi merata pada semua variabel. Beberapa atribut hanya memiliki sedikit nilai hilang, sedangkan atribut lain memiliki jumlah nilai hilang yang lebih tinggi. Kondisi ini menunjukkan bahwa masalah data hilang tidak dapat dipahami hanya dari persentase total, tetapi juga perlu dilihat dari pola distribusinya pada setiap fitur. Dalam dataset klinis, pola seperti ini cukup umum karena tidak semua pemeriksaan medis dilakukan kepada seluruh pasien, baik karena kondisi pasien, keterbatasan rekam medis, perbedaan prosedur pemeriksaan, maupun ketersediaan hasil laboratorium saat pencatatan (Luque et al., 2019; Mazdadi et al., 2024; Thygesen & others, 2018).

Keberadaan nilai hilang menjadi alasan penting untuk menerapkan strategi imputasi sebelum proses klasifikasi. Jika data yang tidak lengkap langsung dihapus menggunakan pendekatan listwise deletion, banyak rekam pasien yang masih memuat informasi klinis penting juga dapat ikut terhapus. Penghapusan langsung juga dapat mengurangi ukuran sampel, menurunkan representativitas data, dan mengubah distribusi karakteristik pasien dalam dataset. Oleh karena itu, penelitian ini tidak menggunakan penghapusan data sebagai pendekatan utama, tetapi membandingkan beberapa metode imputasi, yaitu *mean*, *median*, dan *KNN*. Perbandingan ini dilakukan untuk mengidentifikasi metode imputasi yang paling sesuai dalam mempertahankan informasi data sekaligus menghasilkan performa klasifikasi terbaik pada pengaturan eksperimen yang digunakan (Breiman et al., 1984; van Buuren, 2018).

Selain masalah nilai hilang, distribusi target LET_IS juga menunjukkan ketidakseimbangan kelas yang kuat. Kelas 0 merupakan kelas mayoritas dengan 1.429 rekam atau 84,06% dari seluruh dataset. Sebaliknya, kelas lain memiliki jumlah rekam yang jauh lebih kecil, dan kelas dengan frekuensi terendah hanya memiliki 12 rekam. Kondisi ini menunjukkan bahwa model klasifikasi berhadapan dengan distribusi target yang tidak seimbang, sehingga model cenderung lebih mudah mempelajari pola kelas mayoritas dibandingkan pola kelas minoritas. Ketika dataset dibagi menjadi 80% data latih dan 20% data uji, beberapa kelas minoritas dapat memiliki support yang sangat terbatas pada data uji. Akibatnya, evaluasi terhadap kelas minoritas menjadi lebih sulit karena jumlah contoh yang tersedia tidak sebanding dengan kelas mayoritas (Little & Rubin, 2019).

Ketidakseimbangan kelas merupakan konteks penting dalam menafsirkan nilai akurasi. Pada dataset yang didominasi oleh satu kelas, akurasi dapat tampak tinggi ketika model memprediksi kelas mayoritas dengan baik, meskipun performa pada kelas minoritas belum optimal. Kondisi ini perlu diperhatikan karena akurasi hanya menghitung proporsi prediksi benar secara keseluruhan, sehingga tidak menunjukkan secara rinci apakah model benar-benar mengenali pola setiap kelas target. Dalam kasus target LET_IS yang tidak seimbang, dominasi kelas 0 dapat membuat model cenderung memprediksi kelas mayoritas, sedangkan kelas dengan ukuran sampel kecil kurang terwakili selama proses pembelajaran.

Oleh karena itu, akurasi dalam penelitian ini digunakan sebagai ukuran umum performa model sesuai keluaran *Google Colab* terbaru, tetapi tidak diinterpretasikan sebagai satu-satunya

indikator yang sepenuhnya mewakili kemampuan model pada semua kelas. Hasil perlu ditafsirkan secara proporsional dengan mempertimbangkan dominasi kelas 0 dan rendahnya jumlah rekam pada beberapa kelas lain. Jika performa hanya dilihat dari akurasi, model dapat tampak kuat meskipun kemampuannya dalam mengidentifikasi kelas minoritas masih terbatas. Karena itu, penelitian berikutnya disarankan menambahkan metrik seperti precision, recall, F1-score, dan confusion matrix agar evaluasi kelas minoritas dapat dijelaskan lebih rinci. Penggunaan metrik tambahan tersebut penting untuk memberikan gambaran yang lebih lengkap tentang kelebihan dan kelemahan model, terutama dalam membedakan kelas mayoritas dan minoritas pada dataset klinis yang tidak seimbang (Ghafari & others, 2023; Saito & Rehmsmeier, 2015)

Tabel 3. Distribusi kelas target LET_IS

Kelas LET_IS	Jumlah rekam	Persentase
0	1.429	84,06%
1	110	6,47%
2	18	1,06%
3	54	3,18%
4	23	1,35%
5	12	0,71%
6	27	1,59%
7	27	1,59%
Total	1.700	100,00%

Hasil evaluasi terbaru dari *Google Colab* pada 270 data uji menunjukkan bahwa metode imputasi memengaruhi performa *Decision Tree* berbasis Entropi. Imputasi *median* menghasilkan akurasi tertinggi sebesar 89,63%. Imputasi *KNN* berada pada posisi kedua dengan akurasi 88,52%, sedangkan imputasi *mean* memperoleh akurasi 88,15%. Berdasarkan hasil tersebut, *median* menjadi metode imputasi terbaik dalam eksperimen *single-split* ini ketika penilaian didasarkan pada akurasi. Temuan ini menunjukkan bahwa strategi imputasi dapat memengaruhi distribusi fitur yang diterima model, kemudian berdampak pada performa klasifikasi yang dihasilkan (Luque et al., 2019; Thygesen & others, 2018; van Buuren, 2018). Perubahan nilai ini menggantikan hasil pada versi artikel sebelumnya sehingga isi naskah, tabel, dan keluaran kode menjadi konsisten.

Perbedaan akurasi antar metode relatif kecil, tetapi tetap penting untuk dilaporkan secara konsisten. Selisih antara *median* dan *mean* adalah 1,48 percentage point, sedangkan selisih antara *median* dan *KNN* adalah 1,11 percentage point. Kondisi ini menunjukkan bahwa semua metode imputasi masih menghasilkan performa yang cukup berdekatan, tetapi *median* memberikan hasil tertinggi. Dengan demikian, pembahasan tidak lagi menyatakan bahwa *mean* adalah metode terbaik berdasarkan macro F1-score karena keluaran Colab terbaru yang digunakan dalam naskah ini berfokus pada akurasi sebagai metrik utama. Namun, akurasi tetap harus ditafsirkan secara hati-hati karena target LET_IS memiliki distribusi kelas yang tidak seimbang, sehingga metrik tambahan seperti precision, recall, dan F1-score tetap penting untuk menjelaskan performa model pada kelas minoritas (Ghafari & others, 2023; Little & Rubin, 2019; Saito & Rehmsmeier, 2015).

Tabel 4. Perbandingan akurasi model berdasarkan metode imputasi

Metode	Akurasi	Akurasi (%)	Peringkat
Mean	0,8815	88,15%	3
Median	0,8963	89,63%	1
KNN	0,8852	88,52%	2

Imputasi *KNN* bukan metode terbaik, tetapi performanya masih lebih tinggi daripada imputasi *mean* berdasarkan akurasi. Hasil ini menunjukkan bahwa pendekatan berbasis tetangga tetap dapat mempertahankan sebagian pola lokal dalam data klinis yang memiliki nilai

hilang. Namun, *KNN* belum mampu mengungguli *median* karena perhitungan jarak pada data berdimensi tinggi dapat dipengaruhi oleh skala fitur, jumlah atribut, dan relevansi klinis yang berbeda antarvariabel. Pada dataset dengan banyak atribut, sampel yang tampak dekat secara numerik belum tentu setara secara klinis. Oleh karena itu, performa *KNN* dapat ditingkatkan pada penelitian berikutnya melalui feature scaling, seleksi fitur, reduksi dimensi, atau penggunaan metrik jarak yang lebih sesuai agar struktur ketetanggaan lebih merepresentasikan kemiripan data (Luque et al., 2019; Pedregosa & others, 2011; Thygesen & others, 2018).

Nilai utama pada Tabel 4 telah diselaraskan dengan keluaran *Google Colab* terbaru. Penyelarasan ini penting karena abstrak, tabel hasil, pembahasan, dan kesimpulan merupakan bagian yang paling mudah diperiksa oleh dosen atau reviewer. Dengan menggunakan nilai *mean* 88,15%, *median* 89,63%, dan *KNN* 88,52% secara konsisten, naskah menjadi lebih aman untuk diverifikasi kembali. Konsistensi antara kode eksperimen, tabel hasil, dan narasi pembahasan juga mendukung prinsip reprodusibilitas dalam pelaporan penelitian berbasis machine learning (Harris & others, 2020; Pedregosa & others, 2011).

Tabel 5. Selisih akurasi setiap metode imputasi terhadap median

Metode imputasi	Akurasi (%)	Selisih dari median	Interpretasi
Median	89,63%	0,00 pp	Akurasi tertinggi
KNN	88,52%	-1,11 pp	Peringkat kedua
Mean	88,15%	-1,48 pp	Akurasi terendah

Temuan ini sejalan dengan literatur machine learning yang menyatakan bahwa keputusan prapemrosesan dapat memengaruhi performa klasifikasi, terutama pada dataset klinis yang tidak lengkap. Dalam penelitian ini, metode imputasi yang berbeda menghasilkan nilai akurasi yang berbeda meskipun pengklasifikasi, target, pembagian data, dan parameter utama dibuat sama. Hal tersebut menunjukkan bahwa imputasi tidak boleh dipandang sebagai tahap teknis rutin. Imputasi menentukan bentuk matriks fitur yang diterima model, sehingga nilai yang dihasilkan dapat memengaruhi pola pemisahan pada *Decision Tree* berbasis Entropi. Oleh karena itu, pemilihan metode imputasi perlu dijelaskan secara jelas dalam artikel agar proses eksperimen dapat dipahami dan direplikasi dengan lebih mudah (Breiman et al., 1984; Thygesen & others, 2018; van Buuren, 2018).

Dari sudut pandang praktis, imputasi *median* merupakan pilihan paling sesuai dalam eksperimen *single-split* ini karena menghasilkan akurasi tertinggi. Median relatif tahan terhadap nilai ekstrem sehingga dapat memberikan representasi pusat yang lebih stabil untuk variabel klinis yang distribusinya tidak selalu normal. Keunggulan ini dapat membantu *Decision Tree* berbasis Entropi membentuk ambang pemisahan yang lebih menguntungkan pada data uji. Namun, selisih akurasi yang kecil menunjukkan bahwa *mean* dan *KNN* tetap dapat menjadi alternatif. Hasil validasi silang tetap perlu dibaca sebagai uji stabilitas tambahan, bukan sebagai pengganti hasil utama *single-split*.

Hasil ini juga menunjukkan keterbatasan penggunaan satu model *Decision Tree* berbasis Entropi. Meskipun *Decision Tree* mudah diinterpretasikan, model ini dapat sensitif terhadap perubahan kecil pada data. Perubahan nilai hasil imputasi dapat membuat satu sampel masuk ke cabang pohon yang berbeda. Oleh karena itu, performa model dapat berubah ketika metode imputasi diganti. Untuk meningkatkan ketahanan model, penelitian selanjutnya dapat membandingkan hasil ini dengan random forest, gradient boosting, XGBoost, atau *Decision Tree* berbasis Entropi dengan pengaturan pruning. Perbandingan model lanjutan penting karena studi sebelumnya menunjukkan bahwa metode machine learning yang berbeda dapat menghasilkan performa berbeda dalam memprediksi komplikasi infark miokard (Khamis et al., 2024; Satty et al., 2024).

Keterbatasan lain adalah penggunaan LET_IS sebagai target tunggal. Dataset *Myocardial Infarction Complications* memiliki 12 variabel keluaran, dan setiap keluaran dapat memiliki distribusi serta interpretasi klinis yang berbeda. Oleh karena itu, hasil akurasi pada target

LET_IS tidak dapat langsung digeneralisasi ke semua keluaran komplikasi. Pengujian tambahan pada 12 keluaran dapat memberikan gambaran yang lebih komprehensif. Namun, penggunaan satu target tetap bermanfaat karena membuat perbandingan antar metode imputasi lebih terkontrol dan lebih mudah ditelusuri (Golovenkin & others, 2020).

Implikasi metodologis dari hasil ini adalah bahwa kesesuaian antara kode dan artikel harus menjadi perhatian utama. Dalam penelitian berbasis machine learning, perubahan kecil pada kode, target, *random_state*, atau perhitungan metrik dapat menghasilkan nilai yang berbeda. Jika naskah tidak diperbarui sesuai keluaran terbaru, hasil yang dilaporkan dapat dianggap tidak konsisten. Oleh karena itu, setiap tabel hasil sebaiknya disusun berdasarkan keluaran final yang benar-benar digunakan oleh peneliti. Prinsip ini penting untuk menjaga reprodusibilitas dan kredibilitas artikel, terutama pada studi yang menggunakan perangkat lunak dan pustaka sumber terbuka (Harris & others, 2020; Pedregosa & others, 2011).

Imputasi *mean* dan *median* merupakan metode sederhana, tetapi keduanya tetap relevan sebagai baseline. Dalam eksperimen ini, *mean* menghasilkan akurasi 88,15%, sedangkan *median* meningkat menjadi 89,63%. Perbedaan ini menunjukkan bahwa *median* lebih sesuai untuk pengaturan data ini, kemungkinan karena lebih tahan terhadap nilai ekstrem. Meskipun demikian, *mean* tetap menunjukkan performa cukup tinggi dan tidak tertinggal jauh dari *KNN*. Hasil ini menunjukkan bahwa metode imputasi sederhana masih layak dipertimbangkan pada eksperimen awal, terutama ketika peneliti membutuhkan metode prapemrosesan yang mudah dijelaskan dan direplikasi (Breiman et al., 1984; van Buuren, 2018).

Secara teoretis, imputasi *KNN* memberikan estimasi yang lebih kontekstual karena menggunakan informasi dari sampel terdekat. Pada keluaran *Google Colab* terbaru, *KNN* menghasilkan akurasi 88,52% dan menempati peringkat kedua. Hasil ini menunjukkan bahwa *KNN* dapat memberikan performa kompetitif, tetapi belum melampaui *median*. Salah satu kemungkinan penyebabnya adalah kompleksitas ruang fitur klinis yang memuat banyak variabel dengan skala dan makna berbeda. Tanpa scaling atau seleksi fitur, jarak antarsampel mungkin tidak merepresentasikan kemiripan klinis yang sebenarnya. Oleh karena itu, penerapan feature scaling, seleksi fitur, atau reduksi dimensi dapat dipertimbangkan dalam penelitian selanjutnya (Luque et al., 2019).

Meskipun akurasi menjadi metrik utama pada keluaran terbaru, ketidakseimbangan kelas tetap perlu dicatat sebagai konteks interpretasi. Target LET_IS didominasi oleh kelas 0, sehingga akurasi dapat dipengaruhi oleh keberhasilan model mengenali kelas mayoritas. Oleh sebab itu, precision, recall, dan F1-score dari classification report dirangkum sebagai metrik pendukung pada Tabel 6. Ringkasan tersebut menunjukkan bahwa nilai *weighted average* relatif tinggi karena dipengaruhi oleh dominasi kelas mayoritas, sedangkan nilai *macro average* lebih rendah karena performa pada beberapa kelas minoritas masih terbatas (Ghafari & others, 2023; Little & Rubin, 2019; Saito & Rehmsmeier, 2015).

Untuk menjaga transparansi pelaporan, classification report tidak hanya dibaca dari akurasi, tetapi juga dari *macro average* dan *weighted average*. Macro average memperlakukan setiap kelas secara setara, sehingga lebih sensitif terhadap performa kelas minoritas, sedangkan *weighted average* mempertimbangkan support setiap kelas. Dengan demikian, perbedaan antara *macro average* dan *weighted average* membantu menjelaskan bahwa performa agregat model masih sangat dipengaruhi oleh dominasi kelas mayoritas. Ringkasan metrik pendukung berdasarkan keluaran *Google Colab* terbaru disajikan pada Tabel 6.

Tabel 6. Ringkasan classification report berdasarkan keluaran Google Colab terbaru

Metode	Akurasi	Macro P/R/F1	Weighted P/R/F1
Mean	88,15%	0,29 / 0,29 / 0,29	0,87 / 0,88 / 0,87
Median	89,63%	0,31 / 0,30 / 0,29	0,87 / 0,90 / 0,88
KNN	88,52%	0,36 / 0,28 / 0,29	0,87 / 0,89 / 0,87

Setelah ringkasan metrik pendukung dipastikan selaras dengan keluaran Colab, interpretasi singkat setiap metode imputasi disajikan pada Tabel 7 untuk memperjelas posisi *median*, *KNN*, dan *mean* dalam hasil eksperimen.

Tabel 7. Interpretasi hasil akurasi untuk setiap metode imputasi

Metode	Akurasi	Peringkat	Kekuatan	Catatan
Median	89,63%	1	Tertinggi	Metode utama
KNN	88,52%	2	Kompetitif	Perlu scaling/seleksi fitur
Mean	88,15%	3	Baseline sederhana	Tetap layak dibandingkan

Berdasarkan interpretasi tersebut, beberapa pengembangan eksperimen lanjutan dirangkum pada Tabel 8 agar evaluasi model dapat diperluas dalam penelitian berikutnya.

Tabel 8. Rekomendasi pengembangan eksperimen lanjutan

Aspek	Rekomendasi	Tujuan	Penjelasan	Prioritas
Validasi	Repeated stratified CV	Menguji stabilitas semua metode	Melengkapi 5-fold CV	Tinggi
Metrik	Laporkan precision/recall/F1	Menilai kelas minoritas	Dirangkum pada Tabel 6	Tinggi
KNN	Feature scaling	Memperbaiki perhitungan jarak	Sesuai untuk data banyak fitur	Sedang
Model	Random forest/XGBoost	Membandingkan pengklasifikasi	Eksperimen lanjutan	Sedang

Keluaran terbaru memperjelas hubungan antara metode imputasi dan akurasi model. Imputasi *median* menghasilkan akurasi tertinggi sehingga menjadi metode utama yang direkomendasikan dalam eksperimen *single-split* ini. *KNN* berada pada posisi kedua dan menunjukkan bahwa pendekatan berbasis tetangga masih cukup kompetitif. Imputasi *mean* menjadi metode dengan akurasi terendah pada *single-split*, tetapi selisihnya dari *KNN* dan *median* tidak terlalu besar. Oleh karena itu, pembahasan menekankan bahwa *median* lebih unggul berdasarkan keluaran *Google Colab* terbaru pada skenario pengujian utama, bukan berdasarkan tabel hasil lama. Temuan ini mendukung pandangan bahwa strategi prapemrosesan, terutama imputasi, dapat memengaruhi bentuk data yang dipelajari model dan berdampak pada performa klasifikasi (Afkanpour et al., 2024; Li & others, 2024; Little & Rubin, 2019; van Buuren, 2018)

Hasil validasi silang 5-fold pada keluaran *Google Colab* menunjukkan bahwa *mean* memiliki akurasi CV sebesar 87,82% dengan simpangan baku ±0,76%. Hasil ini diposisikan sebagai uji stabilitas model karena menunjukkan bahwa performa *mean* relatif konsisten pada pembagian data yang berbeda. Namun, hasil CV tersebut tidak mengubah rekomendasi utama penelitian karena skenario perbandingan utama dalam artikel menggunakan *single-split* dengan 270 sampel uji, dan pada skenario tersebut *median* menghasilkan akurasi tertinggi sebesar 89,63%. Dengan demikian, *mean* dicatat memiliki stabilitas CV yang baik, sedangkan *median* tetap direkomendasikan sebagai metode imputasi utama berdasarkan performa tertinggi pada pengujian utama.

Eksperimen ini juga menekankan pentingnya pemisahan data latih dan data uji. Imputer harus dilatih hanya pada data latih, kemudian diterapkan pada data uji. Jika imputer dilatih pada seluruh dataset sebelum pembagian data, informasi dari data uji dapat masuk ke tahap prapemrosesan. Jenis *data leakage* ini dapat membuat akurasi tampak lebih tinggi daripada kondisi sebenarnya. Oleh karena itu, penjelasan tentang urutan *train-test split*, pembersihan data, dan imputasi perlu dipertahankan pada bagian metode agar proses evaluasi tetap valid dan dapat direplikasi (Luque et al., 2019; Pedregosa & others, 2011; Thygesen & others, 2018).

Penggunaan istilah C4.5 juga perlu dijelaskan secara hati-hati. *DecisionTreeClassifier* dengan kriteria entropy bukan implementasi C4.5 murni karena algoritma C4.5 asli memiliki

gain ratio, pruning, dan mekanisme penanganan atribut tertentu. Namun, penggunaan entropy masih dapat dijelaskan sebagai pendekatan terhadap prinsip pemisahan berbasis informasi. Oleh karena itu, istilah yang digunakan secara konsisten dalam naskah adalah *Decision Tree* berbasis Entropi. Penjelasan ini membuat naskah lebih aman secara ilmiah karena reviewer tidak akan menganggap artikel mengklaim implementasi C4.5 murni yang tidak sesuai dengan kode (Breiman et al., 1984; Quinlan, 1993).

Dataset juga memiliki interpretasi temporal yang perlu diperhatikan. Deskripsi dataset menjelaskan bahwa beberapa fitur dapat berasal dari waktu yang berbeda selama perawatan pasien. Artikel ini menggunakan seluruh 111 fitur sebagai pengaturan eksperimen umum untuk membandingkan metode imputasi. Oleh karena itu, hasil penelitian lebih tepat diposisikan sebagai perbandingan metodologis, bukan sebagai model klinis siap pakai untuk prediksi saat pasien masuk. Jika tujuan penelitian berikutnya adalah prediksi dini, fitur dari periode lanjutan perlu ditinjau agar model benar-benar sesuai dengan konteks penggunaan klinis yang dimaksud (Afkanpour et al., 2024; Golovenkin & others, 2020).

Dalam penyusunan naskah, bagian utama yang harus sesuai dengan keluaran Colab adalah abstrak, Tabel 4, Tabel 6, pembahasan, dan kesimpulan. Bagian-bagian tersebut menjadi fokus pembaca ketika menilai apakah hasil penelitian konsisten. Jika nilai pada bagian tersebut berbeda dari kode, naskah mudah dipertanyakan saat proses review. Oleh karena itu, semua nilai utama telah disesuaikan menjadi *mean* 88,15%, *median* 89,63%, *KNN* 88,52%, dan akurasi CV *mean* 87,82% ± 0,76%. Penyesuaian ini membuat artikel lebih siap diperiksa oleh dosen dan reviewer serta mendukung prinsip reprodusibilitas dalam pelaporan eksperimen berbasis machine learning (Harris & others, 2020; Pedregosa & others, 2011).

Kontribusi penelitian ini tetap jelas meskipun hasilnya telah diperbarui. Penelitian ini tidak mengusulkan algoritma imputasi baru atau pengklasifikasi baru. Kontribusinya terletak pada perbandingan empiris tiga metode imputasi pada dataset klinis terbuka. Hasil menunjukkan bahwa pilihan prapemrosesan dapat mengubah performa akurasi model. Kontribusi ini relevan bagi jurnal informatika terapan karena menghubungkan dataset terbuka, prapemrosesan nilai hilang, dan model klasifikasi yang mudah direplikasi.

Secara keseluruhan, pembahasan menghasilkan dua rekomendasi praktis. Pertama, metode terbaik harus ditentukan berdasarkan keluaran eksperimen yang benar-benar digunakan dalam kode. Pada keluaran *single-split* terbaru, *median* merupakan metode terbaik berdasarkan akurasi. Kedua, hasil akurasi perlu dilengkapi dengan konteks ketidakseimbangan kelas agar interpretasi model tidak berlebihan. Dengan demikian, *median* dapat direkomendasikan pada skenario utama, sementara evaluasi lanjutan tetap perlu memperluas metrik, target, dan skema validasi agar hasil penelitian semakin kuat.

Konsistensi antara kode, tabel hasil, abstrak, pembahasan, dan kesimpulan membuat artikel lebih kuat serta mengurangi risiko koreksi selama proses review.

Tabel 9. Ancaman validitas dan strategi mitigasi

Aspek	Potensi ancaman	Mitigasi dalam penelitian atau saran penelitian berikutnya
Validitas internal	Data leakage selama prapemrosesan	Imputer dilatih hanya pada data latih dan kemudian diterapkan pada data uji.
Validitas konstruk	Akurasi dapat dipengaruhi kelas mayoritas pada target tidak seimbang	Akurasi dilaporkan sesuai keluaran Colab; metrik tambahan direkomendasikan untuk penelitian berikutnya.
Validitas eksternal	Hasil hanya berasal dari satu dataset dan satu target	Penelitian berikutnya dapat mengevaluasi seluruh 12 keluaran komplikasi dan dataset klinis lain.
Validitas algoritma	Entropy-based Decision Tree bukan C4.5 murni	Istilah yang digunakan disesuaikan dengan implementasi scikit-learn dan dijelaskan dalam metode.
Reprodusibilitas	Perubahan random_state atau pembagian data dapat mengubah hasil	Pengaturan eksperimen, target, pembagian data, dan parameter utama dilaporkan secara jelas.

Aspek	Potensi ancaman	Mitigasi dalam penelitian atau saran penelitian berikutnya
Interpretasi klinis	Fitur dapat berasal dari waktu perawatan berbeda	Hasil diposisikan sebagai perbandingan metodologis, bukan model klinis siap pakai.

Ancaman validitas pada Tabel 9 perlu dipertimbangkan ketika menafsirkan temuan. Validitas internal terutama berkaitan dengan cara nilai hilang ditangani. Jika proses imputasi menggunakan informasi dari data uji, hasil model dapat menjadi bias. Dalam penelitian ini, imputer dilatih hanya pada data latih untuk mengurangi risiko tersebut. Validitas konstruk berkaitan dengan kesesuaian metrik evaluasi. Akurasi mudah dipahami dan konsisten dengan keluaran *Google Colab* terbaru, tetapi pada target yang tidak seimbang, akurasi perlu didukung oleh precision, recall, dan F1-score.

Validitas eksternal terbatas karena eksperimen menggunakan satu dataset dan satu target. Dataset *Myocardial Infarction Complications* bernilai karena bersifat terbuka dan memiliki missingness realistis, tetapi hasil pada LET_IS tidak otomatis berlaku pada keluaran komplikasi lain. Validitas algoritma juga penting. Algoritma C4.5 asli merupakan metode khusus, sedangkan implementasi yang digunakan adalah *Decision Tree* berbasis Entropi pada scikit-learn. Karena reviewer dapat mempertanyakan perbedaan ini, naskah menjelaskan istilah model secara hati-hati.

Reprodusibilitas bergantung pada pelaporan pengaturan eksperimen yang jelas. Artikel ini melaporkan sumber dataset, jumlah rekam, seleksi fitur, target keluaran, konversi nilai hilang, strategi imputasi, pembagian data, model, parameter, dan metrik evaluasi. Pelaporan tersebut membantu pembaca mengulangi prosedur yang sama. Selain itu, interpretasi klinis juga perlu dibatasi karena penelitian ini berfokus pada perbandingan metodologis, bukan pada pengembangan sistem klinis yang langsung digunakan di rumah sakit.

KESIMPULAN

Penelitian ini membandingkan imputasi *mean*, *median*, dan *KNN* dalam menangani nilai hilang untuk prediksi komplikasi infark miokard menggunakan *Decision Tree* berbasis Entropi. Berdasarkan keluaran *Google Colab* terbaru pada skenario *single-split*, imputasi *median* menghasilkan akurasi tertinggi sebesar 89,63%, diikuti *KNN* sebesar 88,52% dan *mean* sebesar 88,15%. Dengan demikian, metode imputasi yang direkomendasikan pada pengaturan eksperimen utama adalah *median* karena memberikan akurasi terbaik pada 270 data uji. Hasil validasi silang 5-fold menunjukkan bahwa *mean* memiliki stabilitas dengan akurasi 87,82% ± 0,76%, tetapi hasil tersebut diposisikan sebagai evaluasi tambahan dan tidak mengubah rekomendasi utama berbasis *single-split*. Temuan ini menunjukkan bahwa strategi imputasi perlu dipilih dan dilaporkan secara cermat karena dapat memengaruhi distribusi fitur, pola pemisahan pada *Decision Tree*, serta performa klasifikasi.

Penelitian ini menegaskan bahwa penanganan nilai hilang perlu dilaporkan secara jelas karena dapat mengubah performa dan interpretasi model. Penelitian selanjutnya dapat diperluas dengan mengevaluasi seluruh 12 keluaran komplikasi, menggunakan repeated stratified cross-validation, menambahkan precision, recall, F1-score, confusion matrix, serta membandingkan model lain seperti random forest, gradient boosting, dan XGBoost. Pengembangan tersebut diperlukan agar hasil tidak hanya menunjukkan akurasi tertinggi pada satu skenario, tetapi juga memberikan gambaran yang lebih stabil dan proporsional terhadap performa model pada kelas mayoritas maupun minoritas.

Ucapan Terima Kasih

Penulis mengucapkan terima kasih kepada UCI Machine Learning Repository dan repositori data University of Leicester yang telah menyediakan akses terbuka terhadap dataset *Myocardial Infarction Complications* yang digunakan dalam penelitian ini. Penulis juga

berterima kasih atas dukungan pemanfaatan perangkat lunak sumber terbuka yang memungkinkan eksperimen ini direplikasi.

REFERENSI

Afkanpour, M., Hosseinzadeh, E., & Tabesh, H. (2024). Identify the Most Appropriate Imputation Method for Handling Missing Values in Clinical Structured Datasets: A Systematic Review. *BMC Medical Research Methodology*, *24*(1), 188. https://doi.org/10.1186/s12874-024-02310-6

Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). *Classification and Regression Trees*. Wadsworth.

Ghafari, R., & others. (2023). Prediction of the Fatal Acute Complications of Myocardial Infarction via Machine Learning Algorithms. *Journal of Tehran University Heart Center*, *18*(4), 278–287. https://doi.org/10.18502/jthc.v18i4.14827

Golovenkin, S. E., & others. (2020). Trajectories, Bifurcations, and Pseudo-Time in Large Clinical Datasets: Applications to Myocardial Infarction and Diabetes Data. *GigaScience*, *9*(11), giaa128. https://doi.org/10.1093/gigascience/giaa128

Harris, C. R., & others. (2020). Array Programming with NumPy. *Nature*, *585*, 357–362. https://doi.org/10.1038/s41586-020-2649-2

Khamis, G. S. M., Mohammed, Z. M. S., Alanazi, S. M., Mahmoud, A. F. A., Abdalla, F. A., & Bkheet, S. A. (2024). Prediction of Myocardial Infarction Complications Using Gradient Boosting. *Engineering, Technology & Applied Science Research*, *14*(5), 17486–17492.

Li, J., & others. (2024). Comparison of the Effects of Imputation Methods for Missing Data in Predictive Modelling of Cohort Study Datasets. *BMC Medical Research Methodology*, *24*, 41. https://doi.org/10.1186/s12874-024-02173-x

Little, R. J. A., & Rubin, D. B. (2019). *Statistical Analysis with Missing Data* (3rd ed.). Wiley.

Liu, M., & others. (2023). Handling Missing Values in Healthcare Data: A Systematic Review of Deep Learning-Based Imputation Techniques. *Artificial Intelligence in Medicine*, *142*, 102587. https://doi.org/10.1016/j.artmed.2023.102587

Luque, A., Carrasco, A., Martin, A., & de Las Heras, A. (2019). The Impact of Class Imbalance in Classification Performance Metrics Based on the Binary Confusion Matrix. *Pattern Recognition*, *91*, 216–231. https://doi.org/10.1016/j.patcog.2019.02.023

Mazdadi, M. I., Saragih, T. H., Budiman, I., Farmadi, A., & Tajali, A. (2024). The Effectiveness of Data Imputations on Myocardial Infarction Complication Classification Using Machine Learning Approach with Hyperparameter Tuning. *Jurnal Ilmiah Teknik Elektro Komputer Dan Informatika*, *10*(3), 520–533. https://doi.org/10.26555/jiteki.v10i3.29479

Pedregosa, F., & others. (2011). Scikit-Learn: Machine Learning in Python. *Journal of Machine Learning Research*, *12*, 2825–2830.

Powers, D. M. W. (2011). Evaluation: From Precision, Recall and F-Measure to ROC, Informedness, Markedness and Correlation. *Journal of Machine Learning Technologies*, *2*(1), 37–63.

Quinlan, J. R. (1993). *C4.5: Programs for Machine Learning*. Morgan Kaufmann.

Saito, T., & Rehmsmeier, M. (2015). The Precision-Recall Plot is More Informative Than the ROC Plot When Evaluating Binary Classifiers on Imbalanced Datasets. *PLoS ONE*, *10*(3), e0118432. https://doi.org/10.1371/journal.pone.0118432

Satty, A. A., Salih, M. M. Y., Hassaballa, A. A., Gumma, E. A. E., Abdallah, A., & Khamis, G. S. M. (2024). Comparative Analysis of Machine Learning Algorithms for Investigating Myocardial Infarction Complications. *Engineering, Technology & Applied Science Research*, *14*(1), 12775–12779. https://doi.org/10.48084/etasr.6691

Thygesen, K., & others. (2018). Fourth Universal Definition of Myocardial Infarction (2018). *Circulation*, *138*(20), e618–e651. https://doi.org/10.1161/CIR.0000000000000617

van Buuren, S. (2018). *Flexible Imputation of Missing Data* (2nd ed.). CRC Press.