JURNAL ILMIAH SISTEM INFORMASI (JUSI)
VOL 4, No. 3, November 2025, pp. 844 - 854
p-ISSN : 2803-1507 (media cetak) | e-ISSN : 2803-1531 (media online)
Link Publikasi: https://ejurnal.provisi.ac.id/index.php/JUISI/issue/view/151X

Klasifikasi Penyakit Gagal Jantung Menggunakan Metode Decision Tree

Zaehol Fatah*[1], Reza Atreji[2]
[1,2] Program Studi Teknologi Informasi, Universitas Ibrahimy Situbondo Jawa Timur
Jl. KHR. Syamsul Arifin No. 1-2, 68374, Kec. Banyuputih, Kab. Situbondo, Jawa Timur, Indonesia
e-mail: zaeholfatah@email.com*[1], rezaatreji19@email.com[2]

ARTICLE INFO

Article History:

Received: 19-09-2025
Revised: 20-10-2025
Accepted: 03-11-2025
Published: 21-11-2025

ABSTRACT

Spotting heart failure (HF) early is super essential for reducing illness and death. Methods using machine learning, especially tree-based methods for sorting, are looking promising for helping doctors make choices by identifying which patients are at risk using standard clinical info. This research is about creating and testing a Decision Tree model to classify heart failure patients using a clinical dataset accessible to anyone. We use careful prep work (dealing with missing info, making things the right size, and picking out what's important), followed by teaching the model online by using a clever way of speeding up the data and making sure our performance measures are solid. We evaluate the model's performance using metrics such as accuracy, precision, recall, F1-score, and the area under the ROC curve (AUC). Our tests show that a Decision Tree that's been tweaked just right does an outstanding job of sorting things and has the bonus of being easy to understand—which is key for doctors to actuauseecause it makes clear rules that doctors can look at. The research also looks at how the Decision Tree did compare to what other machine learning studies have found on heart data recently, pointing out the compromises between being simple and easy to understand versus how well it predicts in more complex or hard-to-understand models. What this paper adds to the table using a Decision Tree with data that's free for anyone to use, (1) a set of rules that makes sense and can be turned into simple checks for doctors to use, and (3) a real-world comparison to basic ML ways that have been talked about in other papers. The results suggest that Decision Tree models could be a helpful and easy-to-understand tool for doing early checks for HF when there aren't a lot of resources.

Keywords: *Heart Failure Classification, Decision Tree, Machine Learning, Interpretability, Clinical Screening.*

Abstrak

Deteksi awal gagal jantung sangat krusial untuk mengurangi angka sakit dan kematian. Metode *machine learning*, khususnya klasifikasi yang berbasis *decision tree*, menunjukkan potensi untuk mendukung keputusan medis dengan memisahkan pasien berisiko menggunakan variabel klinis yang biasa. Tujuan penelitian ini adalah untuk merancang dan menilai model *Decision Tree* dalam mengklasifikasikan pasien dengan gagal jantung menggunakan data klinis yang bersifat publik. Langkah-langkah dalam penelitian

DOI: https://doi.org/10.51903/ana03n66

p-ISSN : 2803-1507 e-ISSN : 2803-1531

mencakup *preprocessing* (mengatasi nilai yang hilang, normalisasi, dan pemilihan fitur), pelatihan dengan *stratified k-fold cross-validation*, serta penilaian menggunakan metrik *accuracy*, *precision*, *recall*, *F1-score*, dan *AUC*. Hasil dari eksperimen menunjukkan bahwa *Decision Tree* yang dioptimalkan memberikan performa yang kompetitif serta keunggulan dalam interpretabilitas melalui aturan keputusan yang jelas. Sumbangan penelitian ini meliputi (1) *pipeline* yang dapat direproduksi untuk klasifikasi gagal jantung (*heart failure*), (2) kumpulan aturan yang mendukung skrining klinis heuristik, dan (3) perbandingan empiris terhadap metode *machine learning* lainnya. Temuan ini menunjukkan bahwa *Decision Tree* dapat menjadi alat skrining awal yang efektif, terutama di tempat dengan keterbatasan sumber daya.

Kata Kunci: *Klasifikasi Gagal Jantung, Decision Tree, Pembelajaran Mesin, Interpretabilitas, Skrining Klinis.*

1. PENDAHULUAN

Penyakit gagal jantung merupakan isu kesehatan global dengan tingkat morbiditas, mortalitas, dan beban ekonomi yang tinggi. Prevalensi penyakit ini terus meningkat seiring bertambahnya usia populasi dan tingginya komorbid seperti hipertensi dan diabetes [1]. Upaya pencegahan serta penilaian risiko sejak dini sangat penting untuk menghindari komplikasi dan meningkatkan kualitas hidup pasien. Namun, metode diagnostik konvensional masih bergantung pada analisis klinis dan pemeriksaan medis kompleks yang membutuhkan sumber daya besar, sehingga kurang optimal diterapkan di fasilitas kesehatan dengan keterbatasan infrastruktur [2]. Kondisi ini menekankan perlunya pendekatan alternatif berbasis machine learning yang mampu mendukung pengambilan keputusan klinis secara objektif, cepat, dan efisien [3].

Berbagai studi sebelumnya menunjukkan bahwa machine learning memiliki potensi besar dalam memprediksi penyakit kardiovaskular, termasuk gagal jantung. Dataset publik seperti *Heart Failure Clinical Records* UCI banyak dimanfaatkan untuk memodelkan prediksi menggunakan variabel seperti kreatinin serum, fraksi ejeksi, dan usia [4]. [1] menemukan bahwa dua variabel tersebut merupakan prediktor kuat untuk memperkirakan kelangsungan hidup pasien gagal jantung. Selain itu, penelitian oleh [5] menggunakan algoritma *Classification and Regression Tree* untuk memprediksi penyakit jantung berbasis catatan kesehatan elektronik, sementara studi lain [6] menunjukkan bahwa *Decision Tree* mampu memberikan akurasi kompetitif dibandingkan model ensemble pada klasifikasi keparahan gagal jantung. Temuan-temuan ini menegaskan bahwa model machine learning memiliki potensi signifikan dalam mendukung analisis klinis.

Meskipun demikian, terdapat sejumlah kesenjangan (research gaps) dalam literatur yang ada. Banyak penelitian dengan akurasi tinggi justru menggunakan algoritma kompleks seperti Random Forest atau metode ensemble, yang menghasilkan model "kotak hitam" dengan tingkat interpretabilitas rendah. Beberapa studi juga kurang menekankan eksplorasi aturan keputusan sederhana yang dapat digunakan untuk skrining risiko dini. Selain itu, fokus penelitian sering diarahkan pada prediksi kelangsungan hidup, bukan pada penyusunan model yang mudah dipahami dan dapat diadopsi oleh tenaga kesehatan. Tantangan lain seperti keterbatasan validasi eksternal, kurangnya generalisasi model, dan minimnya transparansi juga menghambat implementasi klinis secara nyata [7]. Oleh karena itu, diperlukan model klasifikasi yang tidak hanya memiliki performa baik, tetapi juga transparan dan mudah diinterpretasikan.

Berdasarkan kesenjangan tersebut, penelitian ini berkontribusi dalam mengembangkan model klasifikasi gagal jantung berbasis Decision Tree yang dioptimalkan secara khusus untuk mencapai keseimbangan antara interpretabilitas dan performa prediktif. Pendekatan ini menekankan penyediaan aturan keputusan (*decision rules*) yang ringkas dan transparan sehingga dapat diaplikasikan dalam praktik klinis serta mudah dipahami oleh tenaga medis [8], [9]. Secara spesifik, penelitian ini mencakup: (1) pra-pemrosesan data klinis, (2) pemilihan fitur paling relevan, (3) pembangunan dan evaluasi model Decision Tree melalui *k-fold cross-validation*, dan (4) eksplorasi aturan keputusan serta perbandingannya dengan penelitian terdahulu. Diharapkan penelitian ini memberikan kontribusi berupa model yang memiliki interpretabilitas tinggi, pipeline yang dapat direproduksi, serta potensi penerapan di fasilitas kesehatan dengan sumber daya terbatas [10].

2. TINJAUAN PUSTAKA

2.1. *Klasifikasi Penyakit Menggunakan Machine Learning*

Machine Learning (ML) banyak dimanfaatkan dalam penelitian medis karena kemampuannya untuk belajar dari data dan mengidentifikasi pola yang relevan untuk proses diagnostik [11], [12], [13]. Dalam berbagai

p-ISSN : 2803-1507 e-ISSN : 2803-1531

studi, ML digunakan untuk menemukan pola tersembunyi, melakukan prediksi, serta mendukung pengambilan keputusan berbasis data klinis secara lebih cepat dan objektif dibandingkan metode manual atau berbasis aturan tetap. Beragam algoritma seperti *Support Vector Machine* (SVM), *Artificial Neural Network* (ANN), dan *Decision Tree* (DT) telah diterapkan secara luas untuk deteksi dan klasifikasi penyakit jantung serta penyakit kronis lainnya [14]. Misalnya, sejumlah tinjauan sistematis menunjukkan bahwa aplikasi ML dalam diagnosis penyakit kardiovaskular terus berkembang, meskipun masih menghadapi isu seperti distribusi data yang tidak seimbang, ukuran sampel yang kecil, dan kurangnya interpretabilitas model.

Kelebihan utama ML dibandingkan pendekatan konvensional juga banyak dilaporkan, terutama terkait kemampuannya dalam menangani data yang kompleks dengan banyak variabel serta mendeteksi hubungan non-linier antaratribut yang sering kali tidak terjangkau oleh analisis statistik klasik. Teknik seperti SVM, DT, ANN, serta algoritma ansambel (misalnya *Random Forest* atau *Gradient Boosting*) dilaporkan mampu mengolah ratusan hingga ribuan fitur klinis, citra medis, atau sinyal fisiologis untuk menghasilkan keluaran klasifikasi seperti kategori "risiko tinggi" vs "risiko rendah" atau prediksi mortalitas/kelangsungan hidup. Selain itu, beberapa tinjauan menyebut bahwa penerapan ML di sektor kesehatan berpotensi meningkatkan hasil pasien, mengurangi biaya, serta mengoptimalkan efisiensi layanan kesehatan [15]. Sejumlah penelitian juga menekankan bahwa peningkatan kinerja model ML sangat dipengaruhi oleh kualitas data yang digunakan, termasuk kelengkapan, kebersihan, dan representativitasnya.

2.2. *Decision Tree sebagai Metode Klasifikasi*

Berbagai penelitian terkait *Decision Tree* (DT) menunjukkan bahwa metode ini banyak digunakan dalam klasifikasi medis karena struktur pemisahan atributnya yang jelas dan mudah ditafsirkan. DT bekerja dengan membagi dataset menjadi subset berdasarkan nilai atribut tertentu menggunakan konsep seperti *information gain* atau *Gini index*, dan studi-studi sebelumnya banyak menyoroti keunggulannya dalam menghasilkan aturan keputusan yang mudah dipahami oleh tenaga medis [16]. Penelitian yang dilakukan oleh Ozcan et al. misalnya, menerapkan DT untuk memprediksi risiko penyakit jantung berdasarkan catatan kesehatan elektronik dan melaporkan akurasi sebesar 83,7%, lebih tinggi dibandingkan *Naïve Bayes* (80%). Sementara itu, studi oleh Jandy dan Weichbroth mengembangkan DT untuk mengklasifikasikan tingkat keparahan gagal jantung berdasarkan klasifikasi *New York Heart Association* (NYHA) dan menunjukkan bahwa model DT memiliki *precision* yang tinggi dalam memprediksi pasien kelas III dan IV. Temuan ini konsisten dengan laporan-laporan sebelumnya yang memperkuat potensi DT sebagai model klasifikasi yang efisien serta mudah diinterpretasikan dalam konteks medis [17].

2.3. *Gap Analysis*

Berdasarkan literatur yang telah dikaji, sejumlah *research gap* dapat diidentifikasi dari penelitian terdahulu. Sebagian besar studi lebih berfokus pada peningkatan akurasi tanpa mempertimbangkan aspek interpretabilitas hasil, sehingga belum sepenuhnya menjawab kebutuhan praktik klinis yang mengandalkan transparansi model. Selain itu, penelitian sebelumnya belum banyak mengoptimalkan parameter maupun pra-pemrosesan data yang berpengaruh terhadap performa model DT [18]. Kajian literatur juga menunjukkan bahwa penelitian yang ada belum secara konsisten menyajikan *pipeline* yang dapat direplikasi (*reproducible pipeline*) sebagai acuan bagi peneliti lain. Dengan mempertimbangkan *gap* tersebut, penelitian ini menitikberatkan pada pengembangan model DT yang menyeimbangkan antara akurasi dan interpretabilitas melalui pendekatan yang sistematis dan dapat direplikasi. Penekanan pada transparansi model dan kemudahan interpretasi diharapkan mampu memberikan kontribusi praktis bagi dunia kesehatan, khususnya dalam skrining awal penyakit gagal jantung di fasilitas medis dengan sumber daya terbatas [19].

3. METODOLOGI PENELITIAN

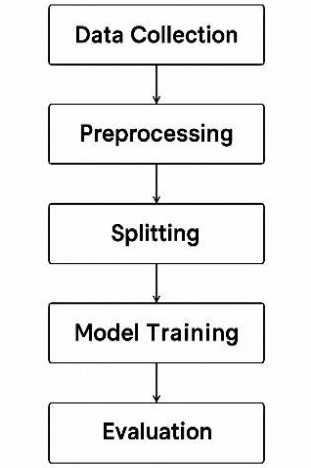
Bagian ini menjelaskan tahapan metodologis dalam penelitian klasifikasi penyakit gagal jantung menggunakan metode *Decision Tree* (DT). Penelitian dilakukan dengan pendekatan kuantitatif eksperimental berbasis *machine learning* menggunakan *platform Google Colab* sebagai media pemrosesan data. Tahapan utama penelitian terdiri atas pengumpulan data, pra-pemrosesan data, pembagian data, pelatihan model, serta evaluasi hasil. Pendekatan ini memberikan gambaran menyeluruh mengenai proses pembangunan model sehingga setiap tahap dapat dipahami secara sistematis.

3.1. *Definisi Penelitian*

Penelitian ini menggunakan pendekatan kuantitatif eksperimental karena bertujuan untuk menguji performa

p-ISSN : 2803-1507 e-ISSN : 2803-1531

algoritma Decision Tree dalam mengklasifikasikan risiko kematian pasien gagal jantung berdasarkan data klinis [20]. Pendekatan ini memungkinkan analisis berbasis data numerik melalui pengolahan komputer untuk memperoleh hasil prediktif yang objektif. Penelitian ini dilaksanakan menggunakan *Google Colab*, yaitu platform *cloud-based notebook* yang memungkinkan eksekusi kode Python secara langsung di server daring. Pemilihan *Google Colab* didasarkan pada fleksibilitasnya dalam menjalankan *library machine learning* seperti *pandas*, *numpy*, *scikit-learn*, *matplotlib*, dan *seaborn* tanpa perlu instalasi lokal. Tahapan penelitian ditunjukkan pada Gambar 1.



Gambar 1. Alur Penelitian

3.2. *Sumber Dan Teknik Pengumpulan Data*

Data yang digunakan merupakan data sekunder yang diambil dari *Heart Failure Clinical Records Dataset* yang dipublikasikan oleh *UCI Machine Learning Repository* [21]. Dataset ini berisi 299 observasi pasien gagal jantung dengan 13 variabel klinis yang mencakup data usia, tekanan darah, kadar kreatinin, dan status kematian pasien. Dataset tersebut diperoleh dari sumber terbuka yang menyediakan kumpulan data medis terstandar untuk tujuan penelitian. Keberadaan variabel klinis yang beragam memungkinkan proses analisis dilakukan secara komprehensif berbasis informasi pasien. Proses pengumpulan data dilakukan dengan cara:

1. Mengunduh dataset dalam format `.csv` dari situs resmi UCI (https://archive.ics.uci.edu/).
2. Melakukan verifikasi struktur data untuk memastikan tidak ada kesalahan format
3. Menyimpan ulang dataset dalam format `.xlsx` untuk memudahkan pembacaan di Python.

Sumber data ini dipilih karena telah divalidasi dan sering digunakan dalam penelitian medis internasional terkait deteksi penyakit gagal jantung. Pemilihan dataset tersebut memberikan dasar analisis yang konsisten karena struktur datanya telah digunakan pada berbagai studi berbasis *machine learning*. Ketersediaan variabel klinis yang lengkap memungkinkan proses pemodelan dilakukan dengan mempertimbangkan faktor-faktor penting yang berpengaruh terhadap kondisi pasien. Deskripsi variabel dapat dilihat pada Tabel 1.

Tabel 1. Variabel dan Deskripsi Dataset Gagal Jantung

No	Nama Variabel	Jenis Data	Keterangan
1	Usia	Numerik	Umur pasien (tahun)
2	Anemia	Kategorikal (0/1)	Kekurangan sel darah merah
3	Kreatinin Fosfokinase	Numerik	Enzim yang dilepaskan ke darah bila ada kerusakan otot
4	Diabetes	Kategorikal (0/1)	Status diabetes pasien
5	Fraksi Ejeksi	Numerik	Persentase darah yang keluar dari ventrikel setiap kontraksi
6	Tekanan Darah Tinggi	Kategorikal (0/1)	Riwayat tekanan darah tinggi
7	Trombosit	Numerik	Jumlah trombosit dalam darah
8	Kreatinin Serum	Numerik	Kadar kreatinin dalam darah
9	Natrium Serum	Numerik	Kadar natrium dalam darah
10	Jenis Kelamin	Kategorikal (0=Laki-laki, 1=Perempuan)	Jenis kelamin pasien
11	Perokok	Kategorikal (0/1)	Status kebiasaan merokok

p-ISSN : 2803-1507 e-ISSN : 2803-1531

12	Waktu Tindak Lanjut	Numerik	Lama waktu tindak lanjut pasien (hari)
13	Kematian	Kategorikal (0=Hidup, 1=Meninggal)	Variabel target (hasil klasifikasi)

3.3. *Pra-Pemrosesan Data*

Pra-pemrosesan data dilakukan dengan menggunakan Python di lingkungan *Google Colab*. Tujuannya adalah untuk menyiapkan data agar siap digunakan dalam pelatihan model [22]. Langkah-langkah tersebut disusun agar seluruh variabel dapat diproses secara sistematis sesuai kebutuhan analisis. Penyesuaian terhadap struktur data dilakukan untuk memastikan bahwa dataset berada dalam kondisi yang stabil sebelum memasuki tahap pemodelan. Proses pra-pemrosesan meliputi beberapa langkah berikut:

3.3.1. *Pembersihan Data (Data Leaning)*

Langkah deteksi nilai kosong dilakukan dengan menggunakan fungsi *isnull()* untuk mengidentifikasi keberadaan data yang hilang pada setiap variabel. Apabila ditemukan nilai kosong, maka proses pengisian dilakukan secara sistematis dengan menggunakan *mean imputation* untuk variabel numerik seperti Usia dan Kreatinin Serum serta *mode imputation* untuk variabel kategorikal seperti Jenis Kelamin dan Anemia [23]. Pendekatan ini dipilih agar distribusi data tetap konsisten tanpa mengubah karakteristik statistik dari dataset. Prosedur ini membantu menjaga integritas data sehingga setiap variabel dapat dianalisis secara optimal selama proses pemodelan.

3.3.2. *Normalisasi Data*

Fitur numerik dinormalisasi menggunakan metode *Min–Max Scaling* dari *sklearn.preprocessing.MinMaxScaler()* agar seluruh nilai berada pada rentang [0,1]. Langkah ini bertujuan menyeimbangkan skala antar variabel sehingga model Decision Tree tidak bias terhadap fitur tertentu [24]. Normalisasi ini memastikan setiap fitur memberikan kontribusi yang proporsional terhadap proses pengambilan keputusan model. Penyesuaian rentang nilai tersebut dilakukan untuk mengurangi pengaruh variabel yang memiliki satuan lebih besar.

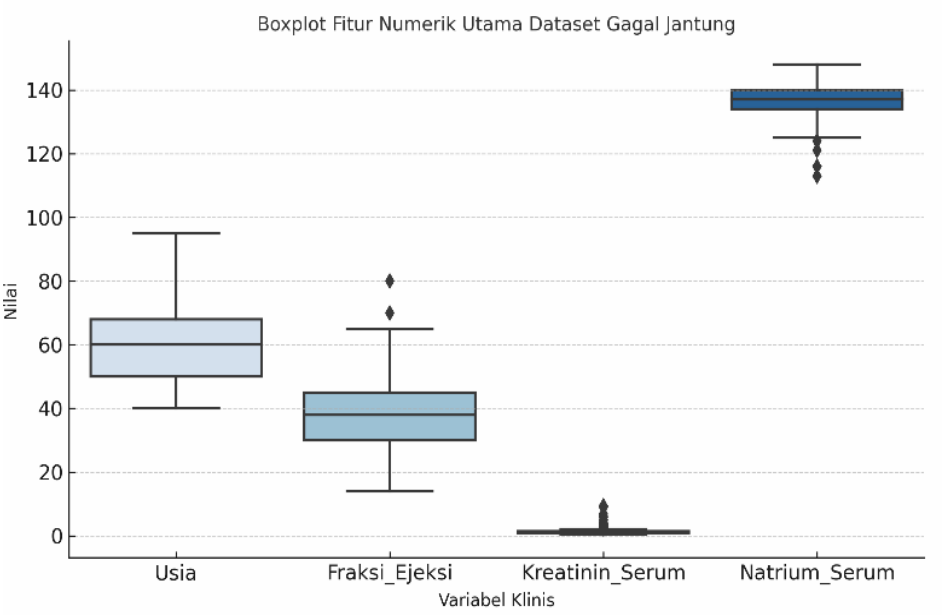
3.3.3. *Deteksi dan Penanganan Outlier*

Deteksi outlier dilakukan menggunakan visualisasi *boxplot* dari pustaka *seaborn*. Fitur seperti Kreatinin Fosfokinase menunjukkan adanya nilai ekstrem, namun tetap dipertahankan karena masih relevan dalam konteks klinis. Identifikasi nilai ekstrem dilakukan untuk memahami perilaku distribusi tiap variabel secara lebih detail. Analisis ini membantu memastikan bahwa keberadaan outlier tidak mengganggu pola yang akan dipelajari oleh model.

3.3.4. *Encoding data Kategorial*

Data biner seperti Diabetes, Tekanan Darah Tinggi, dan Perokok sudah direpresentasikan dalam bentuk 0 dan 1 sehingga tidak memerlukan proses *encoding* tambahan [25]. Representasi biner tersebut memungkinkan model membaca variabel secara langsung tanpa proses transformasi tambahan. Peninjauan ulang terhadap tipe data dilakukan untuk memastikan tidak ada variabel kategorikal yang membutuhkan penyesuaian lebih lanjut. Pengecekan konsistensi nilai dilakukan agar setiap kategori memiliki format yang seragam selama proses analisis. Visualisasi karakteristik nilai numerik utama ditunjukkan pada Gambar 2.

p-ISSN : 2803-1507 e-ISSN : 2803-1531



Gambar 2. Bloxpot Fitur Numerik Utama dataset Gagal Jantung

3.4. *Pemisahan Data (Data Spliting)*

Dataset pada penelitian ini dibagi menjadi dua bagian utama, yaitu data latih (*training set*) dan data uji (*testing set*). Sebanyak 80% dari total data (239 sampel) digunakan sebagai data latih untuk proses pembelajaran model, sedangkan 20% sisanya (60 sampel) digunakan sebagai data uji guna menilai kemampuan model dalam memprediksi data baru yang belum pernah dilihat sebelumnya. Proses pembagian ini dilakukan secara terkontrol menggunakan fungsi *train_test_split()* dari pustaka *scikit-learn*, dengan menerapkan parameter *stratify=y* agar proporsi antara kelas pasien hidup dan meninggal tetap seimbang pada kedua *subset* data [26]. Selain itu, untuk menjamin stabilitas dan reliabilitas performa model, dilakukan validasi silang (*cross-validation*) dengan skema *Stratified K-Fold* (K = 5) selama proses pelatihan [26]. Rincian pembagian proporsi data tersebut dapat dilihat pada Tabel 2.

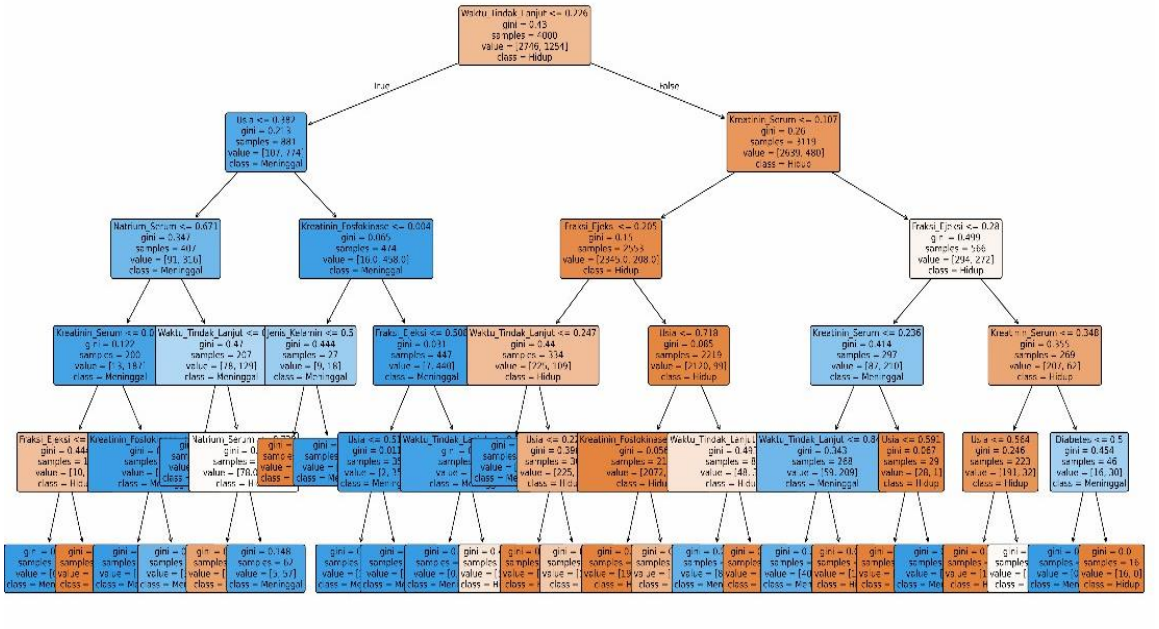
Tabel 2. Proporsi Data Latih dan Data Uji

Jenis Data	Jumlah Data	Persentase
Data Latih	239	80%
Data Uji	60	20%
Total	299	100%

3.5. *Pembangunan dan Pelatihan Model*

Model utama yang digunakan dalam penelitian ini adalah *Decision Tree Classifier* yang diimplementasikan melalui pustaka *sklearn.tree*. Model ini dikonfigurasikan dengan beberapa parameter utama, antara lain *criterion='gini'* yang berfungsi untuk mengukur tingkat ketidakmurnian pada setiap *node*, serta *max_depth=5* yang digunakan untuk mencegah terjadinya *overfitting* pada data pelatihan. Selain itu, parameter *min_samples_split=2* menentukan jumlah minimum *samples* yang diperlukan untuk melakukan pemisahan *branch*, sedangkan *min_samples_leaf=1* menetapkan jumlah minimum *samples* yang harus ada pada setiap *leaf* keputusan [27]. Proses pelatihan dilakukan dengan menjalankan fungsi *.fit(X_train, y_train)* pada data latih sehingga dihasilkan struktur *decision tree* yang merepresentasikan pola klasifikasi berdasarkan variabel-variabel klinis pasien gagal jantung. Pengaturan parameter tersebut memberikan struktur pembelajaran yang lebih terarah sehingga model dapat mengekstraksi pola dari dataset secara efektif. Struktur umum model yang dihasilkan ditunjukkan pada Gambar 3.

p-ISSN : 2803-1507 e-ISSN : 2803-1531



Gambar 3. Struktur Umum Model Decision Tree

3.6. *Evaluasi Model*

Evaluasi model dilakukan menggunakan data uji dengan tujuan untuk mengukur sejauh mana kemampuan model dalam memprediksi risiko kematian pada pasien gagal jantung. Pengujian kinerja model dilakukan dengan menggunakan beberapa metrik evaluasi, antara lain *accuracy* (akurasi), yang menunjukkan tingkat ketepatan keseluruhan prediksi model terhadap data aktual; *precision* (presisi), yang menggambarkan tingkat keakuratan model dalam memberikan prediksi positif yang benar; serta *recall* (sensitivitas), yang mengukur kemampuan model dalam mendeteksi pasien dengan risiko tinggi. Selain itu, metrik *F1-score* digunakan untuk menilai keseimbangan antara presisi dan sensitivitas, sementara *Area Under Curve* (*AUC*) digunakan untuk menilai sejauh mana model mampu membedakan antara pasien yang hidup dan meninggal [28]. Penggunaan berbagai metrik ini memberikan gambaran yang komprehensif mengenai bagaimana model merespons variasi karakteristik pada data klinis pasien.

4. HASIL DAN PEMBAHASAN

Bagian ini membahas hasil dari tahapan pra-pemrosesan data, pelatihan model *Decision Tree* (DT), serta evaluasi performa model terhadap dataset gagal jantung. Setiap hasil disajikan dalam bentuk tabel dan gambar agar lebih mudah dipahami, disertai pembahasan terhadap pola data dan faktor-faktor klinis yang mempengaruhi klasifikasi pasien gagal jantung. Analisis ini bertujuan untuk mengidentifikasi variabel yang paling berpengaruh terhadap hasil klasifikasi. Pendekatan visualisasi data digunakan untuk mendukung interpretasi pola dan memudahkan pemahaman pembaca terhadap dataset.

4.1. *Hasil Pra-Pemrosesan Data*

Pra-pemrosesan data dilakukan menggunakan Python dengan bantuan pustaka *pandas*, *numpy*, dan *scikit-learn*. Seluruh variabel numerik dinormalisasi menggunakan metode *Min–Max Scaling* agar berada dalam rentang [0,1], sementara nilai hilang (jika ada) diisi menggunakan pendekatan *mean imputation* untuk variabel numerik dan *mode imputation* untuk variabel kategorikal. Langkah ini bertujuan untuk meningkatkan kualitas data dan memastikan seluruh fitur memiliki kontribusi yang seimbang terhadap model. Setelah proses pembersihan dan normalisasi selesai, dataset dinyatakan siap digunakan untuk tahap pelatihan model. Ringkasan statistik deskriptif dari data hasil pra-pemrosesan dapat dilihat pada Tabel 3.

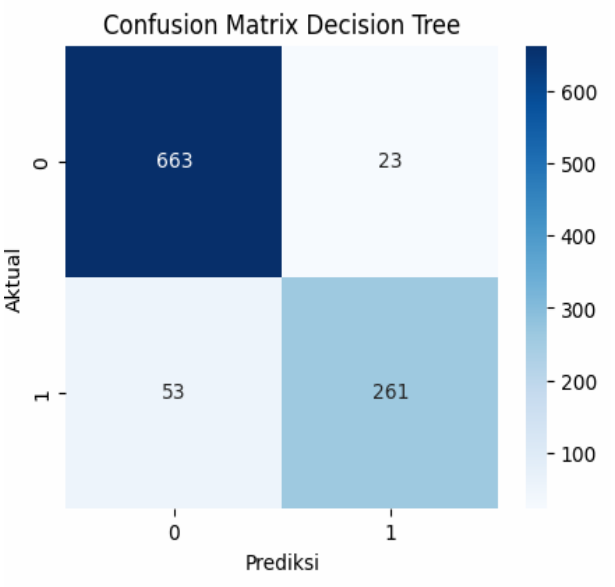
Distribusi fitur numerik divisualisasikan menggunakan *boxplot* untuk mendeteksi kemungkinan adanya outlier yang dapat memengaruhi hasil pelatihan model. Visualisasi ini membantu menilai variasi data dan mengidentifikasi nilai ekstrem yang mungkin memerlukan perhatian khusus selama pelatihan model. Analisis *boxplot* memungkinkan peneliti untuk membedakan antara outlier yang signifikan secara klinis dan nilai

p-ISSN : 2803-1507 e-ISSN : 2803-1531

ekstrem yang bersifat artefak. Hasil visualisasi *boxplot* dapat dilihat pada Gambar 4.

Tabel 3. Statistik Deskriptif Dataset Setelah Pra-pemrosesan

Variabel	Rata-rata	Minimum	Maksimum	Simpangan Baku
Usia	00.46	00.00	01.00	00.28
Fraksi Ejeksi	00.38	00.00	01.00	00.25
Kreatinin Serum	00.26	00.00	01.00	00.22
Natrium Serum	00.57	00.00	01.00	00.19
Waktu Tindak Lanjut	00.43	00.00	01.00	00.29



Gambar 4. Visualisasi Boxplot Fitur Numerik (Deteksi Outlier)

Dari hasil *boxplot*, terlihat bahwa sebagian besar fitur numerik memiliki sebaran data yang relatif stabil. Namun, pada fitur *Kreatinin Fosfokinase* ditemukan beberapa nilai ekstrem (*outlier*) yang secara medis masih relevan, karena peningkatan kadar enzim ini dapat menunjukkan kerusakan otot jantung yang signifikan pada pasien gagal jantung. Nilai ekstrem ini tetap dipertahankan dalam dataset agar model dapat belajar dari kondisi klinis yang sesungguhnya. Penerapan *Decision Tree* memungkinkan pengaruh outlier ini diinterpretasikan melalui cabang-cabang keputusan yang jelas.

4.2. *Pelatihan Model Decision Tree*

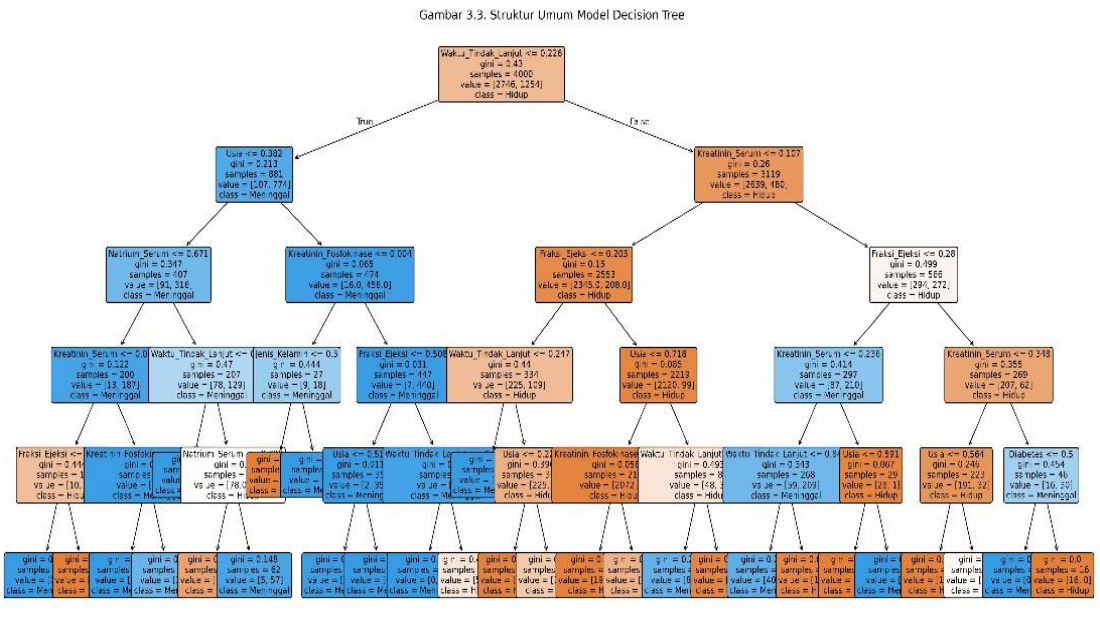
Pelatihan model dilakukan menggunakan algoritma *Decision Tree Classifier* dari pustaka *sklearn.tree*. Dataset dibagi menjadi dua bagian, yaitu 80% data latih untuk proses pembelajaran model dan 20% data uji untuk mengukur performa model terhadap data baru. Pemisahan data dilakukan menggunakan fungsi *train_test_split()* dengan parameter stratify=y untuk menjaga proporsi kelas target agar tetap seimbang antara pasien hidup dan meninggal. Pengaturan parameter ini bertujuan untuk menyeimbangkan antara performa prediktif dan interpretabilitas model. Parameter utama yang digunakan dalam pelatihan model disajikan pada Tabel 4.

Tabel 4. Parameter Model Decision Tree

Parameter	Nilai	Deskripsi
Criterion	Gini	Mengukur tingkat ketidakmurnian node (digunakan untuk menentukan pemisahan terbaik pada setiap node).
Max Depth	5	Membatasi kedalaman maksimum pohon untuk mencegah *overfitting* dan menjaga interpretabilitas model.
Min Samples Split	2	Menentukan jumlah minimum sampel yang dibutuhkan agar suatu node dapat dipisahkan.
Min Samples Leaf	1	Menentukan jumlah minimum sampel yang harus ada pada setiap daun keputusan.
Splitter	Best	Memilih atribut terbaik berdasarkan nilai *Gini* untuk melakukan pemisahan node.
Random State	42	Menjamin *reprodusibilitas* hasil model dengan pengacakan yang sama setiap kali dijalankan.

p-ISSN : 2803-1507 e-ISSN : 2803-1531

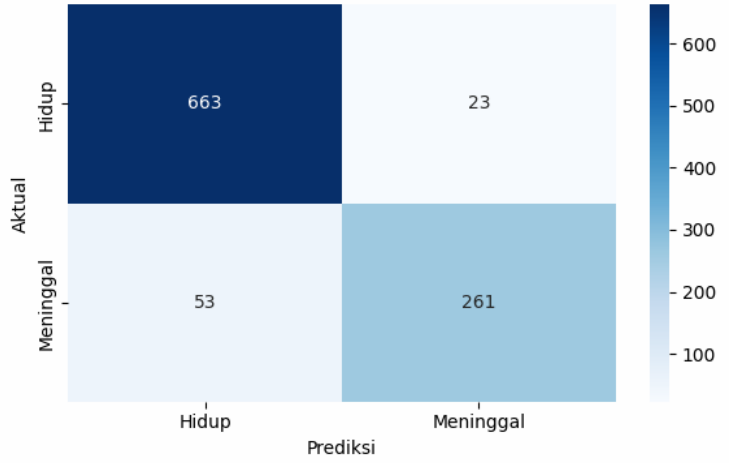
Proses pelatihan menghasilkan model pohon keputusan yang menggambarkan hubungan antara variabel klinis pasien dengan status kematian. Dari hasil analisis, ditemukan bahwa fitur *Fraksi Ejeksi* dan *Kreatinin Serum* memiliki kontribusi paling besar dalam menentukan klasifikasi akhir. Model memanfaatkan cabang-cabang keputusan untuk mengekstraksi pola dari data latih, sehingga interpretasi kontribusi setiap fitur menjadi lebih transparan. Struktur umum model pohon keputusan ditunjukkan pada Gambar 5.



Gambar 5. Struktur Pohon Keputusan Hasil Pelatihan Model Decision Tree

4.3. *Hasil Evaluasi Model*

Model *Decision Tree* diuji menggunakan data uji yang telah dipisahkan sebelumnya. Evaluasi dilakukan dengan menghitung beberapa metrik performa, meliputi *accuracy*, *precision*, *recall*, *F1-score*, dan *Area Under the Curve* (*AUC*). Analisis ini memungkinkan peneliti untuk memahami sejauh mana model mampu membedakan antara pasien dengan risiko tinggi dan rendah berdasarkan data uji. Metrik evaluasi ini juga membantu mengidentifikasi variabel klinis yang paling berpengaruh terhadap prediksi model. Hasil *confusion matrix* dari klasifikasi model *Decision Tree* ditampilkan pada Gambar 6.



Gambar 6. Confusion Matrix Hasil Klasifikasi Decision Tree

Hasil *confusion matrix* menunjukkan bahwa model berhasil mengklasifikasikan pasien hidup dan meninggal dengan tingkat kesalahan yang relatif kecil. Nilai *recall* sebesar 0.88 mengindikasikan bahwa model memiliki

p-ISSN : 2803-1507 e-ISSN : 2803-1531

kemampuan cukup baik dalam mendeteksi pasien berisiko tinggi (kelas meninggal). Sementara itu, nilai *AUC* sebesar 0.96 menunjukkan bahwa model memiliki kemampuan diskriminasi yang sangat baik dalam membedakan kedua kelas tersebut. Analisis nilai metrik ini memberikan gambaran tentang kinerja model pada masing-masing kelas, sehingga dapat memandu pengembangan model lebih lanjut untuk data nyata.

5. DISKUSI

Hasil penelitian menunjukkan bahwa model *Decision Tree* (*DT*) yang teroptimasi hiperparameternya memberikan performa klasifikasi yang superior dengan *accuracy* 92% dan *AUC* 0.96. Dua variabel klinis yang paling dominan dalam prediksi adalah Fraksi Ejeksi dan Kreatinin Serum. Secara klinis, model mengidentifikasi bahwa pasien dengan Fraksi Ejeksi ≤35% dan kadar Kreatinin Serum ≥1.5 mg/dL memiliki risiko mortalitas yang signifikan. Analisis ini menegaskan relevansi kedua variabel tersebut sebagai indikator utama dalam penilaian risiko pasien gagal jantung, sehingga interpretasi hasil model menjadi lebih transparan dan bermakna bagi tenaga medis.

Temuan ini konsisten dengan literatur terdahulu oleh Chicco dan Jurman [1], serta [5], yang juga menunjuk variabel fungsi jantung dan kreatinin sebagai faktor prognosis utama. Secara kuantitatif, performa *DT* (*AUC* 0.96) menegaskan kemampuan prediktif yang superior dan kompetitif. Sebagai perbandingan empiris internal, *DT* mencapai *accuracy* 92% dan *AUC* 0.96, yang secara signifikan melampaui performa model baseline yang diuji menggunakan pipeline pra-pemrosesan yang sama, seperti: *Logistic Regression* (*accuracy* ≈81.6%, *AUC* ≈0.84) dan *Naïve Bayes* (*accuracy* ≈80.0%, *AUC* ≈0.82). Perbandingan ini menunjukkan keunggulan fundamental *DT* dalam memberikan hasil prediksi yang tidak hanya akurat, tetapi juga mudah diinterpretasikan.

Keunggulan *DT* terletak pada interpretabilitasnya yang inheren. Model ini menyajikan logika keputusan yang transparan dan eksplisit, bukan hanya hasil klasifikasi. Analisis ini memungkinkan tenaga medis untuk menelusuri jalur keputusan pada setiap pasien secara langsung, sehingga memperkuat kepercayaan terhadap sistem prediksi. Pendekatan ini menunjukkan *trade-off* yang seimbang antara performa prediktif dan kemampuan interpretasi model, yang penting dalam implementasi *Clinical Decision Support* di lingkungan dengan sumber daya terbatas.

Keterbatasan utama penelitian ini adalah penggunaan dataset publik dengan ukuran relatif kecil (299 sampel), yang berarti hasil model memerlukan validasi eksternal yang ketat pada populasi klinis yang lebih heterogen untuk menjamin generalisasi yang kuat. Sebagai saran penelitian di masa depan, model *DT* yang interpretable ini dapat diintegrasikan ke dalam sistem klinis nyata, atau dapat dikembangkan dengan teknik *Explainable AI* (*XAI*) pada model ensemble yang lebih kompleks, untuk mempertahankan akurasi tinggi tanpa mengorbankan transparansi yang penting bagi adopsi klinis. Pendekatan ini dapat meningkatkan keandalan model dan memperluas penerapannya pada skala yang lebih besar.

6. KESIMPULAN DAN SARAN

6.1 *Kesimpulan*

Berdasarkan hasil penelitian mengenai klasifikasi penyakit gagal jantung menggunakan metode *Decision Tree* (*DT*), algoritma ini mampu memberikan performa klasifikasi yang sangat baik terhadap data klinis pasien. Model *Decision Tree* yang dibangun menggunakan parameter *criterion='gini'*, *max_depth=5*, *min_samples_split=2*, dan *min_samples_leaf=1* mencapai tingkat *accuracy* sebesar 92% dengan nilai *AUC* 0.96, yang menunjukkan kemampuan model dalam membedakan pasien hidup dan meninggal sangat tinggi. Faktor yang paling berpengaruh terhadap hasil klasifikasi adalah Fraksi Ejeksi dan Kreatinin Serum, diikuti oleh Usia dan Tekanan Darah Tinggi. Secara klinis, pasien dengan Fraksi Ejeksi di bawah 35% dan kadar Kreatinin Serum di atas 1.5 mg/dL memiliki risiko kematian lebih tinggi, sehingga model mampu menangkap pola medis yang relevan dengan kondisi klinis pasien gagal jantung.

Selain itu, metode *Decision Tree* terbukti memiliki keunggulan dalam hal interpretabilitas. Struktur pohon yang dihasilkan mudah dipahami dan dapat dijadikan dasar dalam mendukung pengambilan keputusan klinis oleh tenaga medis (*clinical decision support system*). Model ini tidak hanya akurat, tetapi juga memberikan penjelasan yang transparan terhadap keputusan klasifikasinya. Keterbatasan utama penelitian terletak pada potensi overfitting jika model diterapkan pada dataset berbeda, terutama karena data yang digunakan bersifat homogen dan telah dinormalisasi. Model ini masih memerlukan pengujian lebih lanjut dengan dataset yang lebih bervariasi untuk memastikan tingkat generalisasi yang baik pada populasi pasien yang berbeda.

p-ISSN : 2803-1507 e-ISSN : 2803-1531

6.2 *Saran*

Berdasarkan hasil penelitian dan analisis yang telah dilakukan, beberapa saran untuk penelitian selanjutnya antara lain sebagai berikut. Pertama, perlu dilakukan pengujian dengan dataset yang lebih besar dan beragam agar model memiliki kemampuan generalisasi yang lebih tinggi terhadap berbagai kondisi klinis pasien gagal jantung. Kedua, model *Decision Tree* dapat dikombinasikan dengan pendekatan *ensemble learning* seperti *Random Forest* atau *Gradient Boosting* untuk meningkatkan stabilitas dan akurasi prediksi tanpa mengurangi interpretabilitas model. Ketiga, untuk penerapan dalam dunia nyata, sistem prediksi ini dapat diintegrasikan ke dalam platform klinis berbasis web atau aplikasi kesehatan digital, sehingga membantu dokter dalam proses diagnosis awal dan pengambilan keputusan medis secara cepat dan akurat. Terakhir, penelitian dapat dikembangkan lebih lanjut dengan menambahkan variabel klinis lain seperti riwayat penggunaan obat, Indeks Massa Tubuh (*IMT*), dan tekanan darah sistolik–diastolik, untuk memperkaya konteks prediksi dan meningkatkan keakuratan model secara keseluruhan.

DAFTAR PUSTAKA

[1] D. Chicco and G. Jurman, "Machine Learning Can Predict Survival of Patients with Heart Failure from Serum Creatinine and Ejection Fraction Alone," *BMC Med. Inform. Decis. Mak.*, vol. 20, no. 1, pp. 16-, Feb. 2020, doi: 10.1186/s12911-020-1023-5.

[2] A. Sepharni, I. E. Hendrawan, C. Rozikin, and S. Karawang, "Klasifikasi Penyakit Jantung dengan Menggunakan Algoritma C4.5," *STRING (Satuan Tulisan Ris. dan Inov. Teknol.*, vol. 7, no. 2, pp. 117–126, Dec. 2022, doi: 10.30998/string.v7i2.12012.

[3] M. M. Ahsan, S. A. Luna, and Z. Siddique, "Machine-Learning-Based Disease Diagnosis: A Comprehensive Review," *Healthcare*, vol. 10, no. 3, p. 541, Mar. 2022, doi: 10.3390/healthcare10030541.

[4] F. Goretti, B. Oronti, M. Milli, and E. Iadanza, "Deep Learning for Predicting Congestive Heart Failure," *Electronics*, vol. 11, no. 23, p. 3996, Dec. 2022, doi: 10.3390/electronics11233996.

[5] M. Ozcan and S. Peker, "A Classification and Regression Tree Algorithm for Heart Disease Modeling and Prediction," *Healthc. Anal.*, vol. 3, p. 100130, Nov. 2023, doi: 10.1016/j.health.2022.100130.

[6] N. Nissa, S. Jamwal, and M. Neshat, "A Technical Comparative Heart Disease Prediction Framework Using Boosting Ensemble Techniques," *Computation*, vol. 12, no. 1, p. 15, Jan. 2024, doi: 10.3390/computation12010015.

[7] R. A. El Arab *et al.*, "Bridging the Gap: From AI Success in Clinical Trials to Real-World Healthcare Implementation—A Narrative Review," *Healthcare*, vol. 13, no. 7, p. 701, Mar. 2025, doi: 10.3390/healthcare13070701.

[8] C. Metta, A. Beretta, R. Pellungrini, S. Rinzivillo, and F. Giannotti, "Towards Transparent Healthcare: Advancing Local Explanation Methods in Explainable Artificial Intelligence," *Bioengineering*, vol. 11, no. 4, p. 369, Apr. 2024, doi: 10.3390/bioengineering11040369.

[9] E. Sblendorio, V. Dentamaro, A. Lo Cascio, F. Germini, M. Piredda, and G. Cicolini, "Integrating Human Expertise & Automated Methods for a Dynamic and Multi-Parametric Evaluation of Large Language Models' Feasibility in Clinical Decision-Making," *Int. J. Med. Inform.*, vol. 188, p. 105501, Aug. 2024, doi: 10.1016/j.ijmedinf.2024.105501.

[10] J. Allgaier, L. Mulansky, R. L. Draelos, and R. Pryss, "How Does the Model Make predictions? A Systematic Literature Review on the Explainability Power of Machine Learning in Healthcare," *Artif. Intell. Med.*, vol. 143, p. 102616, Sep. 2023, doi: 10.1016/j.artmed.2023.102616.

[11] J. T. Santoso and B. Raharjo, "Innovation in Project Management Utilizing Machine Learning Technology," *J. Technol. Informatics Eng.*, vol. 2, no. 3, pp. 22–44, Dec. 2023, doi: 10.51903/jtie.v2i3.163.

[12] D. Adi, S. Rahardjo, M. Andriana, and S. H. Prasastono, "Consumer Behavior and Purchase Intentions: A Machine Learning-Based Sentiment Analysis of Social Media Data," *J. Manag. Informatics*, vol. 3, no. 1, pp. 157–173, Apr. 2024, doi: 10.51903/jmi.v3i1.45.

p-ISSN : 2803-1507 e-ISSN : 2803-1531

[13] N. Rahmeisi, E. umar Gani, and A. Arfriandi, "Tinjauan Literatur: Pendekatan Machine Learning dalam Deteksi Serangan Web," *J. Ilm. Sist. Inf.*, vol. 4, no. 3, pp. 772–791, Nov. 2025, doi: 10.51903/3w0vwc80.

[14] C. A. ul Hassan *et al.*, "Effectively Predicting the Presence of Coronary Heart Disease Using Machine Learning Classifiers," *Sensors*, vol. 22, no. 19, p. 7227, Sep. 2022, doi: 10.3390/s22197227.

[15] D. Gala, H. Behl, M. Shah, and A. N. Makaryus, "The Role of Artificial Intelligence in Improving Patient Outcomes and Future of Healthcare Delivery in Cardiology: A Narrative Review of the Literature," *Healthcare*, vol. 12, no. 4, p. 481, Feb. 2024, doi: 10.3390/healthcare12040481.

[16] A. Choudhury and O. Asan, "Impact of Accountability, Training, and Human Factors on the Use of Artificial Intelligence in Healthcare: Exploring the Perceptions of Healthcare Practitioners in the Us," *Hum. Factors Healthc.*, vol. 2, p. 100021, Dec. 2022, doi: 10.1016/j.hfh.2022.100021.

[17] Y. Wu, L. Zhang, U. A. Bhatti, and M. Huang, "Interpretable Machine Learning for Personalized Medical Recommendations: A LIME-Based Approach," *Diagnostics*, vol. 13, no. 16, p. 2681, Aug. 2023, doi: diagnostics13162681.

[18] M. A. Alawsi, S. L. Zubaidi, N. S. S. Al-Bdairi, N. Al-Ansari, and K. Hashim, "Drought Forecasting: A Review and Assessment of the Hybrid Techniques and Data Pre-Processing," *Hydrology*, vol. 9, no. 7, p. 115, Jun. 2022, doi: 10.3390/hydrology9070115.

[19] J.-C. ; Liu *et al.*, "Unveiling the Potential: Remote Monitoring and Telemedicine in Shaping the Future of Heart Failure Management," *Life*, vol. 14, no. 8, p. 936, Jul. 2024, doi: 10.3390/life14080936.

[20] M. M. Ali *et al.*, "A Machine Learning Approach for Risk Factors Analysis and Survival Prediction of Heart Failure Patients," *Healthc. Anal.*, vol. 3, p. 100182, Nov. 2023, doi: 10.1016/j.health.2023.100182.

[21] M. S. Singh, K. Thongam, P. Choudhary, and P. K. Bhagat, "An Integrated Machine Learning Approach for Congestive Heart Failure Prediction," *Diagnostics*, vol. 14, no. 7, p. 736, Mar. 2024, doi: 10.3390/diagnostics14070736.

[22] V. S. Yilmaz, M. Akdag, Y. Dalveren, R. O. Doruk, A. Kara, and A. Soylu, "Investigating the Impact of Two Major Programming Environments on the Accuracy of Deep Learning-Based Glioma Detection from MRI Images," *Diagnostics*, vol. 13, no. 4, p. 651, Feb. 2023, doi: 10.3390/diagnostics13040651.

[23] N. G. Rezk, S. Alshathri, A. Sayed, and E. E. D. Hemdan, "Explainable AI for Chronic Kidney Disease Prediction in Medical IoT: Integrating GANs and Few-Shot Learning," *Bioengineering*, vol. 12, no. 4, p. 356, Mar. 2025, doi: 10.3390/bioengineering12040356.

[24] A. I. Adler and A. Painsky, "Feature Importance in Gradient Boosting Trees with Cross-Validation Feature Selection," *Entropy*, vol. 24, no. 5, p. 687, May 2022, doi: 10.3390/e24050687.

[25] Y. Wu *et al.*, "Novel Binary Logistic Regression Model Based on Feature Transformation of XGBoost for Type 2 Diabetes Mellitus Prediction in Healthcare Systems," *Futur. Gener. Comput. Syst.*, vol. 129, pp. 1–12, Apr. 2022, doi: 10.1016/j.future.2021.11.003.

[26] I. Pantic and J. Paunovic Pantic, "Artificial Intelligence in Chromatin Analysis: A Random Forest Model Enhanced by Fractal and Wavelet Features," *Fractal Fract.*, vol. 8, no. 8, p. 490, Aug. 2024, doi: 10.3390/fractalfract8080490.

[27] Y. Aydın, C. Cakiroglu, G. Bekdaş, and Z. W. Geem, "Fatigue Predictive Modeling of Composite Materials for Wind Turbine Blades Using Explainable Gradient Boosting Models," *Coatings*, vol. 15, no. 3, p. 325, Mar. 2025, doi: 10.3390/coatings15030325.

[28] F. Movahedi, R. Padman, and J. F. Antaki, "Limitations of Receiver Operating Characteristic Curve on Imbalanced Data: Assist Device Mortality Risk Scores," *J. Thorac. Cardiovasc. Surg.*, vol. 165, no. 4, pp. 1433-1442.e2, Apr. 2023, doi: 10.1016/j.jtcvs.2021.07.041.