

Klasifikasi Penyakit Gagal Jantung Menggunakan Algoritma *Naive Bayes*

Muhammad Az-Zauqy Faradeya^{1*}, Egia Rosi Subhiyakt^{1,2}

¹Faculty of Computer Science, Universitas Dian Nuswantoro, Semarang, 50131, Indonesia

²Research Center for Intelligent Distributed Surveillance and Security (IDSS), Universitas Dian Nuswantoro, Semarang, 50131, Indonesia

*email: 111202113346@mhs.dinus.ac.id

Info Artikel	ABSTRAK
Dikirim: 20 Desember 2024 Diterima: 5 Februari 2025 Diterbitkan: 23 Mei 2025	Penyakit gagal jantung menunjukkan peningkatan prevalensi global yang mengkhawatirkan dengan kompleksitas dampak klinis yang signifikan. Studi ini mengimplementasikan algoritma <i>Naive Bayes</i> untuk memprediksi risiko gagal jantung, menghadirkan solusi yang lebih efisien dalam komputasi dan interpretasi dibandingkan Random Forest yang membutuhkan waktu komputasi tinggi atau SVM dengan akurasi 92%. Pendekatan metodologi mencakup preprocessing data terstruktur, meliputi penanganan <i>missing value</i> , pengembangan fitur, normalisasi skala, dan penyeimbangan dataset. Penerapan K-Fold Cross Validation dengan variasi K (2, 4, 5, 10) mencapai kinerja optimal pada K=4 dengan akurasi 85,1%, yang memungkinkan pengurangan tingkat kesalahan diagnosis hingga 14,9%. Pencapaian presisi 81,1%, recall 86,1%, dan AUC-ROC 0,914 berkontribusi pada penghematan biaya perawatan melalui ketepatan identifikasi dini. Sistem ini dapat diintegrasikan dalam <i>skrining</i> otomatis untuk efisiensi alokasi sumber daya medis, menghasilkan penghematan operasional signifikan melalui prioritasasi pasien berisiko tinggi dan intervensi preventif yang tepat waktu. Stabilitas performa dengan AUC-ROC konsisten (0,91-0,92) menjadikannya landasan handal untuk sistem pendukung keputusan klinis yang meningkatkan <i>outcome</i> pasien secara menyeluruh.
Kata kunci: <i>Classification;</i> <i>Data Mining;</i> Gagal Jantung; <i>Naive Bayes;</i>	

1. PENDAHULUAN

Data dari Organisasi Kesehatan Dunia (WHO) menunjukkan bahwa penyakit jantung menyumbang sepertiga dari kematian di seluruh dunia. Meningkatnya insiden kondisi gagal jantung dengan tingkat kematian yang tinggi memberikan tekanan yang signifikan pada sistem perawatan kesehatan global. Sekitar 17,9 juta orang di seluruh dunia kehilangan nyawa karena penyakit gagal jantung setiap tahunnya, dengan Asia memiliki frekuensi yang lebih besar dari kondisi tersebut [1]. Prevalensi gagal jantung di negara berkembang terus meningkat dan menjadi salah satu penyebab utama morbiditas dan mortalitas. Sekitar setengah dari pasien yang didiagnosis dengan gagal jantung masih memiliki peluang untuk bertahan hidup hingga lima tahun. Namun, lebih dari 250.000 pasien meninggal setiap tahun akibat gagal jantung, baik secara langsung maupun tidak langsung. Angka ini telah meningkat hingga enam kali lipat selama empat dekade terakhir [2]. Negara berpendapatan rendah dan menengah menunjukkan perbedaan signifikan dalam penyebab, karakteristik klinis, dan pengelolaan gagal jantung dibandingkan negara berpendapatan tinggi. Faktor-faktor seperti hipertensi, diabetes melitus, dislipidemia, merokok, dan konsumsi alkohol berlebihan sering tidak terdiagnosis atau tidak terkelola dengan baik di negara-negara tersebut [3] [4]. Selain itu, penyakit jantung rematik, perikarditis tuberkulosis, dan penyakit kardiovaskular yang terkait dengan HIV masih memberikan kontribusi signifikan

terhadap tingginya angka gagal jantung di wilayah ini [4]. Usia pasien yang lebih muda di negara-negara berkembang mencerminkan kelemahan sistem kesehatan dalam mendeteksi dini faktor risiko. Hal ini diperparah dengan akses terbatas terhadap layanan kesehatan dan terapi medis canggih [5]. Secara keseluruhan, gagal jantung menjadi tantangan kesehatan global yang memerlukan pendekatan sistemik untuk pengendalian faktor risiko dan peningkatan deteksi dini. Pasien dengan gagal jantung sering kali merasa sulit untuk melakukan tugas sehari-hari, yang membuat mereka sangat rentan terhadap rasa khawatir, stres, putus asa, dan kesulitan dalam mengelola emosi mereka sendiri. Kualitas hidup pasien dapat terpengaruh secara negatif oleh pemikiran mereka tentang biaya pengobatan, prognosis kondisi mereka, dan waktu yang dibutuhkan untuk pulih [6]. Untuk menghindari gagal jantung pada tahap awal, sangat penting untuk memahami variabel-variabel yang berkontribusi terhadap gagal jantung. Dengan menggunakan teknik data mining dan klasifikasi, salah satu pendekatannya adalah mengidentifikasi gejala gagal jantung berdasarkan riwayat medis pasien.

Data mining adalah teknik yang membantu menganalisis data dan kemudian menghasilkan informasi yang berguna. Teknik ini mengintegrasikan atau menggabungkan dengan disiplin ilmu lain seperti statistik, kecerdasan buatan, dan machine learning. Data mining ialah teknik untuk menemukan, mencari, atau mengeksplorasi informasi atau pengetahuan baru dari sekumpulan data yang sangat besar [7]. Dalam ranah data mining, terdapat berbagai algoritma klasifikasi yang dapat dimanfaatkan untuk menganalisis dan mengategorikan data. Mengingat beragamnya pilihan algoritma tersebut, sangat penting untuk melakukan perbandingan guna menentukan tingkat akurasi dan efektivitas masing-masing metode. Algoritma-algoritma klasifikasi yang sering digunakan, seperti *Random Forest*, *Naïve Bayes*, *Support Vector Machine*, dan *Decision Tree*, menawarkan pendekatan berbeda dalam mengolah dan mengklasifikasikan data, sehingga membutuhkan evaluasi komprehensif untuk memahami kinerja dan keunggulan spesifik setiap algoritma [8]. Istilah Belanda untuk klasifikasi adalah *Classification*, yang berasal dari kata klasifikasi dalam bahasa Prancis. Frasa ini menggambarkan teknik untuk mengatur fakta secara metodis sesuai dengan pedoman atau aturan yang telah ditetapkan sebelumnya. Karena ada kategori variabel tujuan dalam kategorisasi [9]. Klasifikasi di dalam Data Mining merupakan metode analitik yang memungkinkan pengelompokan data ke dalam kategori tertentu dengan mempertimbangkan atribut dan karakteristik spesifik yang dimiliki. Pada hakikatnya, teknik ini bertujuan untuk mengembangkan model atau fungsi prediktif yang dapat menentukan kelas atau kategori dari suatu entitas data baru berdasarkan parameter dan ciri-ciri yang telah diidentifikasi sebelumnya [10]. Pada penelitian klasifikasi penyakit gagal jantung ini akan berfokus pada algoritma *Naive Bayes*. Pendekatan *Naive Bayes*, yang dapat memberikan probabilitas kelas keluaran berdasarkan gejala atau karakteristik yang termasuk dalam kumpulan data, digunakan oleh para peneliti dalam penelitian ini untuk kategorisasi. Karena mudah diinterpretasikan dan diaplikasikan, pendekatan *Naive Bayes* banyak digunakan dalam pengembangan sistem pakar. Metode *Naive Bayes* merupakan pendekatan kategorisasi yang menggunakan fitur atau sifat intrinsik data untuk menempatkannya ke dalam kategori yang telah ditentukan. Dasar dari pendekatan ini adalah teori probabilitas Bayes, yang menyatakan bahwa kemungkinan terjadinya suatu kejadian tergantung pada kemungkinan kejadian-kejadian yang terkait. Pendekatan *Naive Bayes* digunakan dalam sistem pakar untuk klasifikasi [11]. *Gaussian Naive Bayes* (GNB) merupakan salah satu varian metode klasifikasi statistik yang mengaplikasikan teorema Bayes dengan pendekatan distribusi Gaussian. Metode ini memungkinkan pengklasifikasian suatu observasi ke dalam kategori tertentu dengan memanfaatkan informasi dari variabel prediktor, di mana perhitungan probabilitas didasarkan pada karakteristik distribusi normal dari data yang dianalisis. Melalui pendekatan matematis ini, GNB mampu memberikan estimasi yang efektif untuk mengategorikan data berdasarkan probabilitas kemunculan pada setiap kelas yang ada [12].

Dari penelitian sebelum nya pada tahun 2022, penelitian yang dilakukan oleh Quswatun Hasanah, dkk. untuk memprediksi gagal jantung menggunakan metode *Naive Bayes* dengan hasil akurasi 69% [13], pada tahun 2022, penelitian dilakukan oleh Sepharni, dkk untuk melakukan klasifikasi penyakit jantung menggunakan algoritma C4.5 dengan hasil akurasi 79% [14], pada tahun 2022 penelitian yang dilakukan Pratama, dkk untuk melakukan klasifikasi Penyakit Gagal Jantung Menggunakan Algoritma *K-Nearest Neighbor* dengan hasil akurasi 70,65% [15], dan pada tahun 2022 penelitian yang dilakukan Edric, dan Saut Parsaoran Tamba untuk melakukan prediksi Penyakit Gagal Jantung Dengan Menggunakan *Random Forest* dengan hasil akurasi sebesar 82,6087% yang kemudian di optimasi dengan teknik K-Fold dan Grid Search CV menjadi 85,058% [16]. Pada tahun 2023 penelitian yang dilakukan Akhdan Arifuddin, dkk melakukan penelitian dengan judul

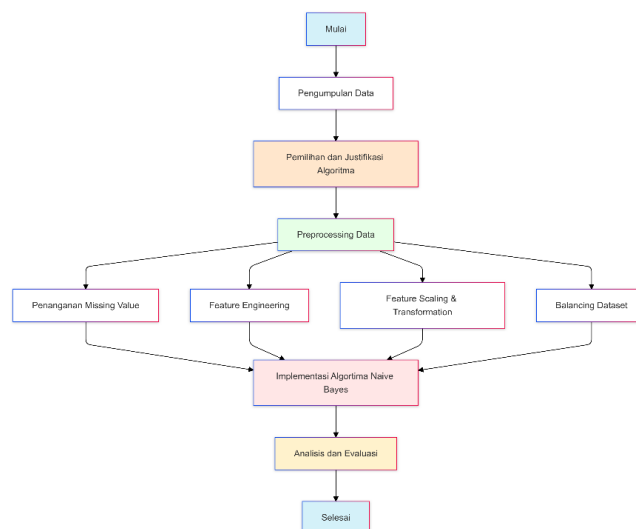
Perbandingan Kinerja Pohon Keputusan dan Vektor Pendukung Algoritma Mesin untuk Prediksi Gagal Jantung (dalam bahasa Indonesia) dan mendapatkan hasil akurasi 92% pada metode *Support Vector Machine* dan 81% *Decision Tree*[17].

Namun, penelitian sebelumnya masih memiliki beberapa keterbatasan. Metode seperti Random Forest membutuhkan waktu komputasi tinggi pada dataset besar, sementara akurasi K-Nearest Neighbor masih rendah (70,65%). Penelitian oleh Quswatun Hasanah dkk. menggunakan Naïve Bayes tanpa melakukan preprocessing pada dataset yang sama. Dalam penelitian ini, dilakukan preprocessing untuk meningkatkan kualitas data, serta evaluasi ulang terhadap kinerja Naïve Bayes pada data yang telah diproses. Tujuannya adalah membuktikan bahwa metode ini tidak hanya memberikan akurasi yang baik, tetapi juga praktis dan mudah diterapkan dalam konteks nyata.

2. METODE PENELITIAN

2.1. Tahapan Penelitian

Implementasi klasifikasi penyakit gagal jantung memerlukan metodologi penelitian yang sistematis dan terstruktur untuk mencapai hasil yang optimal. Untuk memastikan ketercapaian tujuan penelitian, berikut diuraikan diagram alur metodologi yang menggambarkan serangkaian tahapan penelitian secara sekuensial:



Gambar 1. Diagram Alur

2.2. Pengumpulan Data

Penelitian ini menggunakan dataset dari Kaggle (<https://www.kaggle.com/jackleenrasmybareh/heart-failure>) yang berisi 1000 kasus gagal jantung terdokumentasi pada tahun 2020. Dataset ini dipilih karena representatif dan kompleks, dengan distribusi seimbang antara kasus survival (508) dan mortalitas (492), sehingga mengurangi potensi bias kelas. Terdapat 31 variabel klinis yang mencakup aspek demografis, komorbiditas, dan riwayat perawatan, memberikan gambaran menyeluruh terhadap profil pasien dan faktor risiko gagal jantung. Namun, keterbatasan dataset perlu diperhatikan, seperti ukuran sampel yang terbatas, potensi bias geografis karena bersumber dari satu institusi, serta missing values pada variabel quintile (6 data) dan ethnicgroup (43 data). Selain itu, tidak terdapat informasi rinci tentang tingkat keparahan penyakit maupun jenis intervensi medis. Meskipun begitu, dengan preprocessing yang tepat, dataset ini tetap memadai untuk membangun model prediktif yang andal.

2.3. Pemilihan dan Justifikasi Algoritma *Naive Bayes*

Naive Bayes memiliki keunggulan signifikan dibanding algoritma lain, terutama dalam efisiensi, interpretabilitas, dan kemampuannya menangani berbagai jenis dataset, termasuk data medis. Algoritma ini mengandalkan pendekatan probabilistik yang sederhana, memudahkan interpretasi hasil dan cocok untuk data dengan asumsi independensi antar fitur [18]. Penelitian menunjukkan bahwa dibanding *K-Nearest Neighbor* (KNN), *Naive Bayes* lebih efisien dalam perhitungan dan pengolahan data berskala besar [19]. Bahkan pada dataset kecil, seperti dalam klasifikasi keluhan masyarakat, algoritma ini tetap mampu memberikan hasil yang akurat dan cepat [20]. Oleh karena itu, *Naive Bayes* menjadi pilihan ideal untuk analisis yang menuntut efisiensi, akurasi, dan kemudahan implementasi.

2.4. Preprocessing Data

Pra pemrosesan data merupakan prosedur kritis yang melibatkan manipulasi dan persiapan data secara komprehensif sebelum dilakukannya analisis lanjutan [21]. Dalam konteks pembelajaran mesin atau metodologi penelitian ilmiah, tahapan ini memiliki peran fundamental untuk mengoptimalkan kualitas dan validitas data. Penelitian kami mengimplementasikan serangkaian prosedur pra pemrosesan yang dirancang secara sistematis, dengan fokus utama pada penanganan karakteristik data medis yang memiliki kompleksitas tinggi, guna menjamin kehandalan dan akurasi analisis yang akan dilakukan.

2.4.1. Penanganan *Missing Value*

Kehilangan data berdampak besar terhadap integritas metodologis dataset, yang dapat menyebabkan distorsi statistik dan menurunkan kualitas pengambilan keputusan. Data yang tidak lengkap bukan hanya menimbulkan ketidakakuratan dalam klasifikasi dan pengelompokan, tetapi juga melemahkan validitas serta reliabilitas analisis secara keseluruhan [22]. Dalam dataset gagal jantung yang digunakan, terdapat 1000 kasus dengan data hilang terutama pada informasi demografis: etnis (43 data) dan status sosial-ekonomi (6 data). Penanganan missing value menjadi krusial karena setiap atribut pasien dapat memengaruhi hasil diagnosis. Untuk itu, kami menerapkan tiga pendekatan penanganan data hilang yang disesuaikan dengan jenis data.

2.4.1.1. *K-Nearest Neighbor Imputation*

Untuk data numerik seperti usia dan lama rawat inap, kami menggunakan metode KNN Imputer yang mengisi data hilang berdasarkan kemiripan karakteristik antar pasien. Metode ini dipilih karena pada kasus gagal jantung, pasien dengan karakteristik serupa cenderung memiliki pola parameter medis yang mirip, sehingga estimasi yang dihasilkan lebih mencerminkan kondisi klinis yang realistis.

2.4.1.2. *Multiple Imputation by Chained Equations (MICE)*

MICE digunakan untuk memvalidasi dan melengkapi proses imputasi, khususnya untuk menangani keterkaitan kompleks antar faktor risiko gagal jantung. Metode ini mempertimbangkan hubungan antara berbagai kondisi komorbid, parameter vital, dan karakteristik pasien dalam proses imputasi, yang penting untuk meningkatkan akurasi prediksi gagal jantung.

2.4.1.3. *Mode-based Imputation*

Untuk data kategorikal seperti jenis penyakit komorbid, kami menerapkan imputasi berbasis mode dengan mempertimbangkan kelompok umur pasien. Pendekatan ini dipilih karena manifestasi penyakit sering berkorelasi dengan kelompok usia tertentu pada kasus gagal jantung. Strategi ini memastikan nilai yang diimputasi tetap konsisten dengan karakteristik demografis dan pola klinis yang umum ditemui pada kelompok usia tersebut.

Kombinasi ketiga metode ini menghasilkan dataset yang lebih lengkap dan andal untuk pengembangan model prediksi, dengan tetap mempertahankan pola klinis yang relevan untuk diagnosis gagal jantung. Hasil evaluasi

menunjukkan bahwa data yang diimputasi konsisten dengan karakteristik klinis pasien gagal jantung, sehingga mendukung pengembangan model prediksi yang lebih akurat.

2.4.2. Feature Engineering

Feature Engineering merupakan teknik sistematis dalam analisis data yang mencakup serangkaian prosedur untuk meningkatkan kualitas variabel input, dengan tujuan mengoptimalkan performa model prediksi gagal jantung [23], Kami melakukan beberapa pengembangan fitur yang disesuaikan dengan karakteristik data medis:

2.4.2.1. Kategorisasi Variabel Dasar

Untuk meningkatkan interpretabilitas model, kami melakukan kategorisasi pada variabel umur dan lama rawat inap menggunakan metode q-cut dengan 20 kelompok. Pendekatan ini memungkinkan pengelompokan pasien berdasarkan karakteristik yang serupa, yang penting untuk analisis risiko gagal jantung.

2.4.2.2. Sistem Scoring Kondisi Medis

Kami mengembangkan sistem pembobotan untuk kondisi komorbid berdasarkan tingkat keparahan dan dampaknya terhadap gagal jantung. Kanker diberi bobot tertinggi (3.0), diikuti penyakit ginjal (2.7) dan aritmia (2.5), sementara kondisi seperti COPD (2.3), hipertensi (2.2), dan diabetes (2.0) diberi bobot sesuai signifikansi klinisnya dalam perkembangan gagal jantung.

2.4.2.3. Fitur Interaksi

Untuk menangkap kompleksitas hubungan antar variabel medis, kami mengembangkan beberapa fitur interaksi:

- 1) Perhitungan jumlah komorbiditas untuk mengukur kompleksitas kondisi pasien
- 2) Risk density sebagai indikator kumulatif risiko gagal jantung
- 3) Interaksi antara usia dengan faktor risiko dan komorbiditas
- 4) Efisiensi lama rawat inap untuk evaluasi penanganan pasien

Pengembangan fitur ini dirancang khusus untuk meningkatkan kemampuan model dalam memprediksi gagal jantung dengan mempertimbangkan berbagai aspek kondisi medis pasien secara komprehensif. Hasil evaluasi menunjukkan bahwa fitur-fitur yang dikembangkan berkontribusi signifikan dalam meningkatkan akurasi prediksi model *Naive Bayes*.

2.4.3. Feature Scaling dan Transformation

Feature scaling merupakan tahap kritis dalam preprocessing data medis untuk meningkatkan akurasi prediksi gagal jantung [24][25]. Proses transformasi ini mengkonversi variabel-variabel dengan rentang nilai yang berbeda ke dalam skala yang sebanding, seperti usia pasien (30-102 tahun), lama rawat inap (0-89 hari), dan jumlah kunjungan medis. Standardisasi ini penting untuk memastikan setiap variabel memberikan kontribusi proporsional dalam prediksi, mengingat beragamnya parameter medis yang digunakan dalam diagnosis gagal jantung.

2.4.3.1. RobustScaler

RobustScaler diimplementasikan untuk menangani outlier dalam data medis dengan lebih efektif. Teknik ini menggunakan statistik yang kuat (median dan kuartil) untuk menghilangkan pengaruh nilai ekstrem pada parameter medis seperti lama rawat inap yang sangat panjang atau usia yang jauh dari rata-rata, sehingga meningkatkan keandalan prediksi gagal jantung.

2.4.3.2. PowerTransformer

PowerTransformer dengan metode Yeo-Johnson digunakan untuk menormalkan distribusi variabel medis yang tidak seimbang. Metode ini efektif untuk menangani data seperti jumlah komorbiditas atau skor risiko yang cenderung memiliki distribusi miring, membantu meningkatkan akurasi prediksi gagal jantung dengan membuat distribusi data lebih normal.

2.4.3.3. MinMaxScaler

MinMaxScaler mentransformasi semua fitur ke dalam rentang yang seragam (-1,1), memastikan setiap variabel medis memberikan kontribusi seimbang dalam prediksi. Pendekatan ini penting untuk algoritma *Naive Bayes* yang sensitif terhadap skala data, membantu meningkatkan akurasi prediksi dengan menyeimbangkan pengaruh setiap parameter medis.

Kombinasi ketiga metode ini menghasilkan dataset yang lebih seimbang dan terstandardisasi untuk prediksi gagal jantung, dengan evaluasi menunjukkan peningkatan signifikan dalam performa model.

2.4.4. Balancing Dataset

Dataset gagal jantung sering mengalami ketidakseimbangan kelas yang dapat mempengaruhi akurasi prediksi, dengan rasio 50.8% kasus survival dan 49.2% kasus mortalitas. Untuk mengatasi hal ini, kami menerapkan serangkaian teknik penyeimbangan data yang dirancang khusus untuk kasus medis:

2.4.4.1. Pembersihan Data

Kombinasi EditedNearestNeighbours dan TomekLinks diterapkan untuk menghapus noise data yang dapat mengaburkan pola prediksi gagal jantung. Teknik ini dipilih karena kemampuannya mengidentifikasi dan menghapus kasus-kasus yang berpotensi membingungkan model tanpa menghilangkan informasi klinis penting.

2.4.4.2. SMOTE (Synthetic Minority Over-sampling Technique)

SMOTE diimplementasikan dengan parameter $k_neighbors=5$ untuk menyeimbangkan kelas minoritas dalam dataset gagal jantung. Metode ini menciptakan sampel sintetis dengan menginterpolasi antara pasien yang memiliki karakteristik serupa, khususnya dalam parameter medis yang kritis. SMOTE membantu meningkatkan jumlah kasus kematian tanpa menimbulkan overfitting, karena sampel baru dibuat berdasarkan pola nyata dalam data.

2.4.4.3. ADASYN (Adaptive Synthetic)

ADASYN diterapkan dengan $n_neighbors=7$ sebagai metode oversampling sekunder yang berfokus pada kasus-kasus yang sulit diprediksi. Berbeda dengan SMOTE, ADASYN secara adaptif menghasilkan lebih banyak sampel sintetis untuk kasus gagal jantung yang memiliki pola lebih kompleks. Parameter $n_neighbors$ yang lebih besar memastikan sampel sintetis mencerminkan variasi yang lebih luas dalam karakteristik pasien.

2.4.4.4. Pengendalian Outlier

Penanganan outlier menggunakan metode z-score dengan threshold 3.0 diterapkan untuk memastikan keandalan model. Pendekatan ini mempertahankan variasi alami dalam data medis sambil menghilangkan nilai-nilai ekstrem yang dapat mengganggu akurasi prediksi. Kombinasi metode penyeimbangan ini menghasilkan dataset yang lebih representatif untuk pengembangan model prediksi gagal jantung, dengan tetap mempertahankan karakteristik klinis yang penting untuk diagnosis.

2.5. Implementasi Algoritma Naïve Bayes

Implementasi algoritma *Naive Bayes* untuk klasifikasi penyakit gagal jantung dilakukan melalui langkah-langkah sistematis. Proses diawali dengan preprocessing data, dilanjutkan dengan pelatihan model menggunakan prinsip probabilitas Bayes, yaitu menghitung probabilitas prior dan likelihood berdasarkan fitur klinis. Model divalidasi menggunakan teknik K-Fold Cross Validation dengan berbagai nilai K untuk mengevaluasi performa dan stabilitas dalam berbagai skenario pembagian data. Pendekatan ini memungkinkan analisis yang menyeluruh dan adaptif terhadap variasi distribusi data.

2.6. Analisis dan Evaluasi

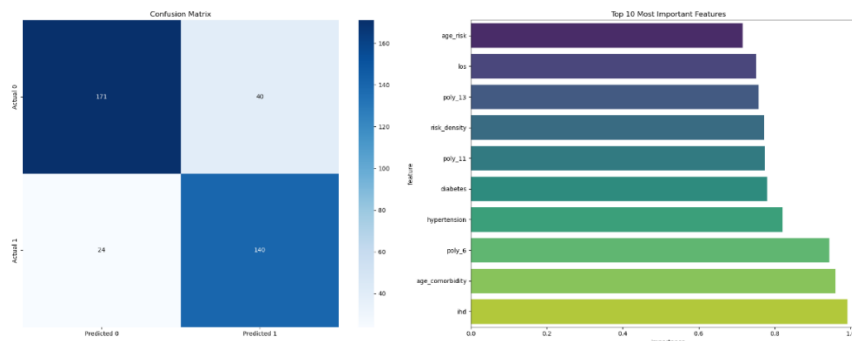
Proses analisis model dilakukan secara menyeluruh menggunakan metrik evaluasi standar: akurasi untuk mengukur ketepatan total, presisi untuk mengevaluasi keakuratan prediksi positif, recall untuk menilai sensitivitas terhadap kasus positif, dan AUC-ROC untuk mengukur kemampuan diskriminatif model. Confusion matrix memberikan gambaran detail tentang distribusi prediksi benar dan salah, sementara kurva ROC membantu mengevaluasi performa pada berbagai threshold. Konsistensi model diuji melalui perbandingan performa antar fold, memastikan stabilitas prediksi dalam skenario pembagian data yang berbeda.

3. HASIL DAN PEMBAHASAN

3.1. Hasil klasifikasi

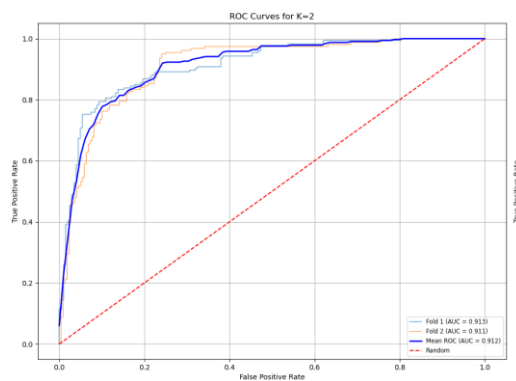
Pada proses klasifikasi menggunakan algoritma *Naive Bayes*. Pengujian dilakukan menggunakan K-fold dengan nilai k 2, 4, 5, dan 10, klasifikasi ini dilakukan untuk mengetahui hasil akurasi, presisi, dan recall untuk penyakit gagal jantung.

3.1.1. Kfold 2



Gambar 2. Confusion Matrix Kfold 2

Pada skenario K=2, model menunjukkan akurasi rata-rata 82,7% dengan recall tinggi (85,1%) dan presisi lebih rendah (77,6%), menandakan sensitivitas yang baik namun akurasi klasifikasi positif masih terbatas. Pembagian data menjadi dua fold (500 sampel per fold) memberikan data pelatihan yang besar, tetapi validasi yang minim memicu overfitting, terlihat dari 40 false positive dan 24 false negative. Performa antar fold pun kurang stabil (akurasi 82,8% dan 82,7%), mencerminkan keterbatasan validasi silang dalam menguji generalisasi model.



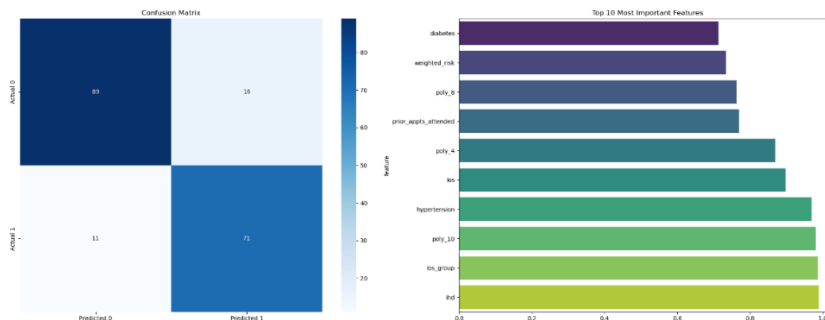
Gambar 3. ROC Curve K-fold 2

Model menunjukkan performa dasar yang kuat dengan pembagian data menjadi dua fold yang besar. Kurva kedua fold hampir identik (AUC 0.913 dan 0.911), mengindikasikan stabilitas model meski dengan validasi terbatas. Kenaikan tajam di awal kurva menunjukkan model sangat baik dalam mendeteksi kasus positif, namun pembagian data yang minimal dapat membatasi generalisasi model untuk kasus klinis yang lebih beragam.

Tabel 1 Hasil Klasifikasi Kfold 2

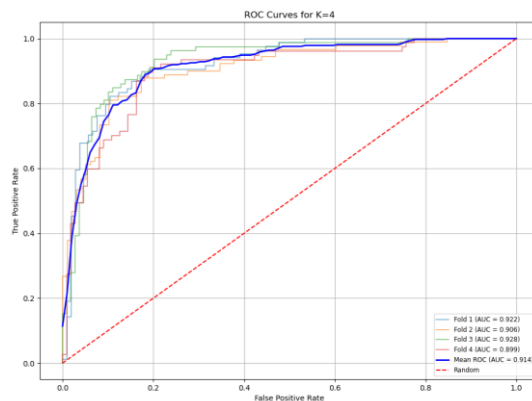
Fold	Accuracy	Precision	Recall	AUC-ROC
1	0.828	0.782	0.868	0.913
2	0.827	0.769	0.833	0.911
Rata-rata	0.827	0.776	0.851	0.912

3.1.2. Kfold 4



Gambar 4. Confusion Matrix Kfold 4

Peningkatan jumlah fold menjadi 4 (250 sampel per fold) menghasilkan pembagian data yang lebih seimbang dan validasi yang lebih robust. Model mencatat rata-rata akurasi 85,1%, presisi 81,1%, dan recall 86,1%, dengan kesalahan minimal (16 false positive, 11 false negative). Konsistensi akurasi antar fold (85,1–85,2%) menunjukkan stabilitas model yang meningkat, memungkinkan identifikasi survival dan kematian secara lebih akurat tanpa mengorbankan sensitivitas.



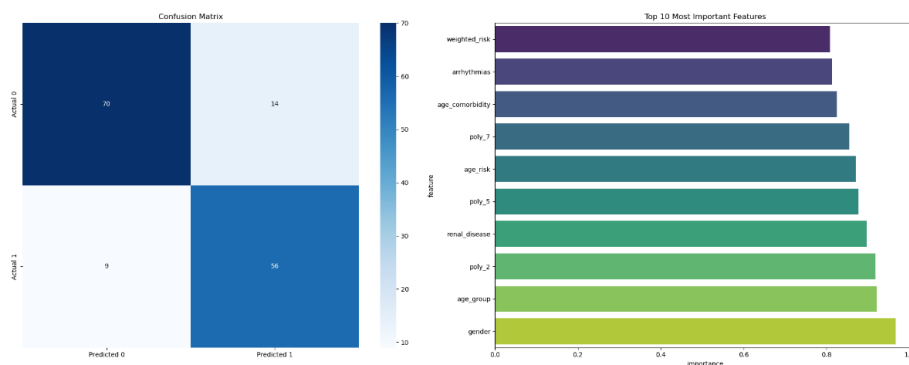
Gambar 5. ROC Curve K-fold 4

Peningkatan jumlah fold menghasilkan performa optimal dengan variasi AUC 0.899-0.928 antar fold. Model menunjukkan keseimbangan terbaik antara sensitivitas dan spesifisitas, terlihat dari kurva yang konsisten naik tajam dan stabil di area true positive rate tinggi. Pembagian data menjadi 4 fold memberikan validasi yang lebih robust sambil mempertahankan ukuran training set yang adequate, menghasilkan model yang lebih reliable untuk implementasi klinis.

Tabel 2. Hasil Klasifikasi Kfold 4

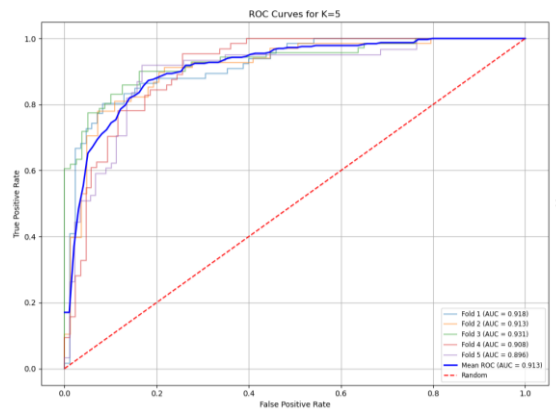
Fold	Accuracy	Precision	Recall	AUC-ROC
1	0.852	0.841	0.821	0.922
2	0.851	0.823	0.878	0.906
3	0.851	0.793	0.873	0.928
4	0.851	0.788	0.870	0.899
Rata-rata	0.851	0.811	0.861	0.914

3.1.3. Kfold 5



Gambar 6. Confusion Matrix Kfold 5

Skenario K=5 dengan 200 sampel per fold mulai menunjukkan trade-off antara frekuensi validasi dan keandalan evaluasi. Model mempertahankan performa yang stabil dengan 70 true negative dan 56 true positive, namun terjadi peningkatan false positive menjadi 17 kasus. Variabilitas performa antar fold lebih terlihat dengan rentang akurasi dari 82.0% hingga 86.8%, mengindikasikan dataset validasi yang lebih kecil mulai mempengaruhi stabilitas prediksi. Meski terjadi sedikit penurunan rata-rata akurasi menjadi 84.2%, model tetap menunjukkan keseimbangan baik antara presisi (79.6%) dan recall (85.7%), dengan nilai AUC-ROC yang stabil di 0.913.



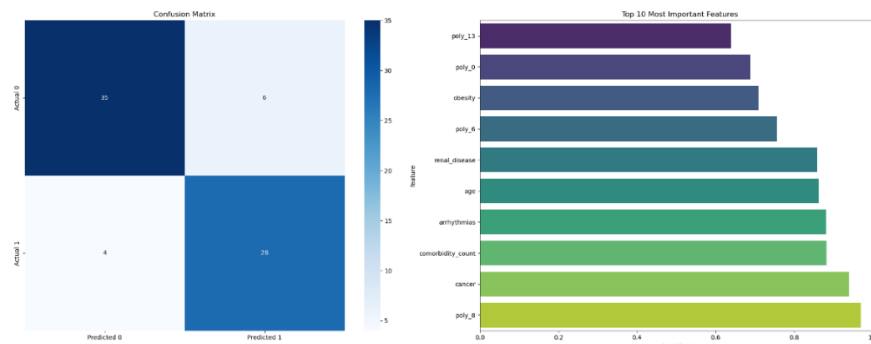
Gambar 7. ROC Curve K-fold 5

Pembagian data menjadi 5 fold mulai menunjukkan variabilitas performa (AUC 0.896-0.931). Meski rata-rata AUC tetap tinggi, fluktuasi antar fold mengindikasikan model menjadi lebih sensitif terhadap komposisi data training. Kurva ROC menunjukkan trade-off antara false positive dan true positive rate yang konsisten, namun dengan variasi lebih besar dibanding K=4.

Tabel 3. Hasil Klasifikasi Kfold 5

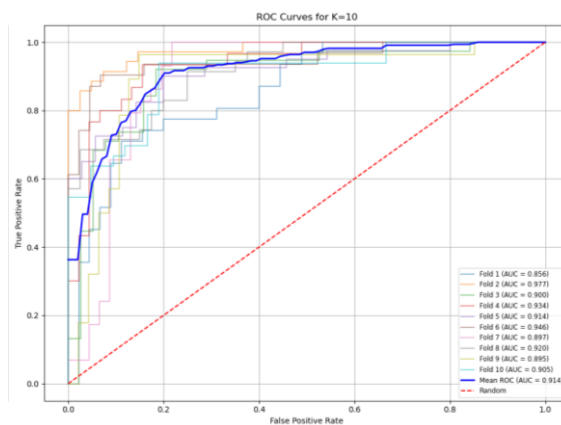
Fold	Accuracy	Precision	Recall	AUC-ROC
1	0.848	0.812	0.848	0.918
2	0.821	0.789	0.824	0.913
3	0.868	0.831	0.901	0.931
4	0.820	0.768	0.828	0.908
5	0.853	0.783	0.885	0.896
Rata-rata	0.842	0.796	0.857	0.913

3.1.4. Kfold 10



Gambar 8. Confusion Matrix Kfold 10

K=10 dengan 100 sampel per fold menunjukkan ketahanan model terhadap fragmentasi data yang tinggi. Meski ukuran subset validasi lebih kecil, model mempertahankan rata-rata akurasi 85.0% dengan presisi 80.7% dan recall 85.8%. Namun, pembagian data yang lebih granular menghasilkan variabilitas performa yang lebih tinggi antar fold, terlihat dari rentang akurasi yang lebar (78.9% hingga 90.8%) dan fluktuasi AUC-ROC (0.856 hingga 0.977). Hal ini menunjukkan bahwa meski model mampu mempertahankan performa rata-rata yang baik, reliabilitas prediksi individual menjadi lebih sensitif terhadap komposisi subset validasi. Konsistensi recall di atas 85% mengkonfirmasi model tetap handal dalam mendeteksi kasus positif meski dengan dataset yang lebih terfragmentasi.



Gambar 9. ROC Curve K-fold 10

Fragmentasi data tertinggi menghasilkan variabilitas signifikan (AUC 0.856-0.977). Meski rata-rata AUC tetap 0.914, range yang lebar menunjukkan ketergantungan tinggi terhadap komposisi fold. Dalam konteks klinis, variabilitas ini dapat mempengaruhi reliabilitas prediksi individual, meskipun performa agregat tetap baik.

Tabel 4 Hasil Klasifikasi Kfold 10

Fold	Accuracy	Precision	Recall	AUC-ROC
1	0.789	0.759	0.710	0.856
2	0.908	0.868	0.943	0.977
3	0.868	0.833	0.921	0.900
4	0.867	0.857	0.800	0.934
5	0.853	0.837	0.900	0.914
6	0.880	0.824	0.903	0.946
7	0.840	0.758	0.862	0.897
8	0.800	0.763	0.829	0.920
9	0.880	0.788	0.929	0.895
10	0.813	0.788	0.788	0.905
Rata-rata	0.850	0.807	0.858	0.914

3.2. Hasil Evaluasi dan validasi

Evaluasi *Naive Bayes* dengan *K-Fold Cross Validation* menunjukkan variasi performa berdasarkan nilai *K*:

1. *K*=2 (500 sampel/fold): akurasi 82.7%, *AUC* 0.911-0.913, stabil tapi validasi terbatas, berisiko prediksi kurang reliabel pada kasus tidak terwakili.
2. *K*=4 (250 sampel/fold): performa terbaik dengan akurasi 85.1%, presisi 81.1%, *recall* 86.1%, *AUC* 0.914, false negative rendah (11), false positive moderat (16), cocok untuk keputusan klinis dengan threshold diagnostik andal.
3. *K*=5 (200 sampel/fold): akurasi 84.2%, *AUC* 0.896-0.931, stabilitas mulai menurun, berpotensi mempengaruhi konsistensi diagnosis borderline.
4. *K*=10 (100 sampel/fold): akurasi 85.0%, *AUC* 0.856-0.977, validasi luas tapi performa individu variatif, mengurangi kepercayaan prediksi kasus tunggal.

Kesimpulan: *K*=4 adalah konfigurasi optimal, menawarkan keseimbangan akurasi, stabilitas, dan keandalan diagnostik untuk sistem pendukung keputusan klinis, dengan sensitivitas tinggi untuk meminimalkan risiko gagal deteksi. Output model dapat membantu tenaga medis dalam evaluasi risiko gagal jantung sesuai protokol klinis.

3.3. Perbandingan Hasil

Perbandingan kinerja model dengan dataset Kaggle menunjukkan peningkatan signifikan dibanding penelitian sebelumnya [13]. *Naive Bayes* dengan *K*=4 mencapai akurasi 85.1% (naik 16.1% dari 69% pada *K*=10) dan presisi 81.1% (vs 62.12%), meski *recall* sedikit menurun (86.1% vs 87.23%). *AUC-ROC* 0.914 menandakan keseimbangan metrik lebih optimal.

Peningkatan ini disebabkan oleh:

1. Pembagian data optimal (K=4) yang seimbang antara training dan validasi,
2. *Preprocessing* lebih komprehensif (*missing value* dan *feature engineering*),
3. Pemilihan *threshold* yang lebih tepat untuk diagnosis klinis.

Tabel 5. Perbandingan Hasil Akurasi Tertinggi

No	Penelitian	Algoritma	Kfold	Accuracy	Precision	Recall	AUC-ROC
1	Penelitian [13]	<i>Naive Bayes</i>	10 -fold	69%	62,12%	87,23%	-
2	Penelitian ini	<i>Naive Bayes</i>	4 -fold	85,1%	81,1%	86,1%	0.914

4. KESIMPULAN

Implementasi *Naive Bayes* dengan *K-Fold Cross Validation* menunjukkan performa optimal pada K=4 (akurasi 85.1%, presisi 81.1%, recall 86.1%, AUC-ROC 0.914). Model konsisten di semua skenario, dengan AUC stabil (0.91–0.92) dan recall selalu lebih tinggi dari presisi. Confusion matrix K=4 menunjukkan akurasi tinggi dalam klasifikasi dengan kesalahan minimal.

Model ini dapat diintegrasikan ke dalam:

1. Aplikasi web screening awal,
2. Sistem pendukung keputusan klinis,
3. Platform mobile untuk monitoring pasien.

Pengembangan lanjutan direkomendasikan dengan:

1. *Dataset multi-center* >5000 pasien,
2. Outlier handling berbasis klinis,
3. Optimasi *hyperparameter* dan seleksi fitur (LASSO),
4. *Ensemble* model dengan *weighted voting*.

Hasil menunjukkan *Naive Bayes* (K=4) efektif dan efisien untuk prediksi gagal jantung.

REFERENSI

- [1] A. Wintarti and S. Adi, “Mathunesa Tahun 2022 Komparasi Metode Support Vector Machine (SVM), K-Nearest Neighbors (KNN), dan Random Forest (RF) untuk Prediksi Penyakit Gagal Jantung,” *Jurnal Ilmiah*, 2022, doi: <https://doi.org/10.26740/mathunesa.v10n2.p258-268>.
- [2] L. B. De Azevedo, G. G. Cucato, and B. Morseth, “Editorial: Sedentary behaviour and cardiometabolic health,” *Front Cardiovasc Med*, vol. 11, Oct. 2024, doi: 10.3389/fcvm.2024.1498410.
- [3] G. Peerwani, B. Hanif, K. A. Rahim, M. Kashif, S. S. Virani, and S. Sheikh, “Presentation, management, and early outcomes of young acute coronary syndrome patients- analysis of 23,560 South Asian patients from 2012 to 2021,” *BMC Cardiovasc Disord*, vol. 24, no. 1, p. 378, Jul. 2024, doi: 10.1186/s12872-024-04036-1.
- [4] H. Li *et al.*, “Ischemic Heart Disease Mortality and Metabolic Risk Factors: A Global Analysis of Gender Disparities and Regional Variations (1980–2021),” Jan. 16, 2025. doi: 10.21203/rs.3.rs-5820382/v1.
- [5] A. O. Bilchenko *et al.*, “Acute myocardial infarction complicated by cardiogenic shock in Ukraine: multicentre registry analysis 2021–2022,” *Front Cardiovasc Med*, vol. 11, Mar. 2024, doi: 10.3389/fcvm.2024.1377969.
- [6] H. Nursita *et al.*, “Peningkatan Kualitas Hidup pada Pasien Gagal Jantung: A Narrative Review Article,” *Jurnal Berita Ilmu Keperawatan*, vol. 13, no. 1, pp. 10–21, 2020, doi: <https://dx.doi.org/10.23917/bik.v13i1.11916>.
- [7] A. Yani, Z. Azmi, and D. Suherdi, “Implementasi Data Mining Menganalisa Data Penjualan Menggunakan Algoritma K-Means Clustering,” *Jurnal Sistem Informasi Triguna Dharma (JURSI TGD)*, vol. 2, no. 2, p. 315, Mar. 2023, doi: 10.53513/jursi.v2i2.6357.
- [8] N. B. Putri and A. W. Wijayanto, “Analisis Komparasi Algoritma Klasifikasi Data Mining Dalam Klasifikasi Website Phishing,” *Komputika : Jurnal Sistem Komputer*, vol. 11, no. 1, pp. 59–66, Jan. 2022, doi: 10.34010/komputika.v11i1.4350.

- [9] D. K. Andi, Said Imam Puro, Jodi Hariyan, Ridho Rafliansyah, Muhammad Rival Aziz, and Pebro Vaulina Rajagukguk, "Analisa Data Shopping Trends Menggunakan Algoritma Klasifikasi Dengan Metode *Naive Bayes*," *Repeater : Publikasi Teknik Informatika dan Jaringan*, vol. 2, no. 3, pp. 119–134, Jul. 2024, doi: 10.62951/repeater.v2i3.118.
- [10] D. Muriyatmoko, A. Musthafa, and M. H. Wijaya, "Klasifikasi Profil Kelulusan Nilai AKPAM Dengan Metode Decision Tree C4.5," Apr. 2024.
- [11] P. Septiani, I. Pratiwi, M. Ghofar Rohman, and M. Sholihin, "Sistem Pakar Penyakit Telinga Menggunakan Metode *Naive Bayes*," 2023. doi: <https://doi.org/10.29407/gj.v7i2.19991>.
- [12] Dian, Purnawansyah, H. Darwis, and L. Nurhayati, "Klasifikasi Penyakit Bawang Merah Menggunakan *Naive Bayes* dan Convolutional Neural Network," *The Indonesian Journal of Computer Science*, vol. 12, no. 4, Aug. 2023, doi: 10.33022/ijcs.v12i4.3265.
- [13] Q. Hasanah, H. Oktavianto, and Y. D. Rahayu, "Analisis Algoritma Gaussian *Naive Bayes* Terhadap Klasifikasi Data Pasien Penderita Gagal Jantung Gaussian *Naive Bayes* Algorithm Analysis Of Data Classification Of Heart Failure Patiens," May 2022. [Online]. Available: <http://jurnal.unmuhjember.ac.id/index.php/JST>
- [14] A. Sepharni, I. E. Hendrawan, C. Rozikin, and S. Karawang, "Klasifikasi Penyakit Jantung dengan Menggunakan Algoritma C4.5," 2022. doi: <http://dx.doi.org/10.30998/string.v7i2.12012>.
- [15] Y. Pratama, A. Prayitno, D. Azrian, N. Aini, Y. Rizki, and E. Rasywir, "Klasifikasi Penyakit Gagal Jantung Menggunakan Algoritma K-Nearest Neighbor," *Bulletin of Computer Science Research*, vol. 3, no. 1, pp. 52–56, Dec. 2022, doi: 10.47065/bulletincsr.v3i1.203.
- [16] P. Tamba, "Prediksi Penyakit Gagal Jantung dengan Menggunakan Random Forest," *Jurnal Sistem Informasi dan Ilmu Komputer Prima*, vol. 5, no. 2, 2022, doi: <https://doi.org/10.34012/jurnalsisteminformasidanilmukomputer.v5i2.2445>.
- [17] A. Arifuddin, G. S. Buana, R. A. Vinarti, and A. Djunaidy, "Performance Comparison of Decision Tree and Support Vector Machine Algorithms for Heart Failure Prediction," in *Procedia Computer Science*, Elsevier B.V., 2024, pp. 628–636. doi: 10.1016/j.procs.2024.03.048.
- [18] A. Desiani, "Perbandingan Implementasi Algoritma *Naive Bayes* dan K-Nearest Neighbor Pada Klasifikasi Penyakit Hati," *SIMKOM*, vol. 7, no. 2, pp. 104–110, Jul. 2022, doi: 10.51717/simkom.v7i2.96.
- [19] I. M. A. A. D. Putra, I. M. G. Sunarya, and I. G. A. Gunadi, "Perbandingan Algoritma *Naive Bayes* Berbasis Feature Selection Gain Ratio dengan *Naive Bayes* Kovenisional dalam Prediksi Komplikasi Hipertensi," *JTIM : Jurnal Teknologi Informasi dan Multimedia*, vol. 6, no. 1, pp. 37–49, Apr. 2024, doi: 10.35746/jtim.v6i1.488.
- [20] H. Derajad Wijaya and S. Dwiasnati, "Implementasi Data Mining dengan Algoritma *Naive Bayes* pada Penjualan Obat," *JURNAL INFORMATIKA*, vol. 7, no. 1, 2020, doi: <https://doi.org/10.31294/ji.v7i1.6203>.
- [21] E. Oluwasakin *et al.*, "Minimization of high computational cost in data preprocessing and modeling using MPI4Py," *Machine Learning with Applications*, vol. 13, p. 100483, Sep. 2023, doi: 10.1016/j.mlwa.2023.100483.
- [22] S. Alam, M. S. Ayub, S. Arora, and M. A. Khan, "An investigation of the imputation techniques for *missing values* in ordinal data enhancing clustering and classification analysis validity," *Decision Analytics Journal*, vol. 9, p. 100341, Dec. 2023, doi: 10.1016/j.dajour.2023.100341.
- [23] M. R. Machado, A. Asadi, R. W. R. de Souza, and W. C. Ugulino, "Green AI in the finance industry: Exploring the impact of feature engineering on the accuracy and computational time of Machine Learning models," *Appl Soft Comput*, vol. 167, p. 112343, Dec. 2024, doi: 10.1016/j.asoc.2024.112343.
- [24] R. C R and Dr. S. C P, "Evaluating Deep Learning with different feature scaling techniques for EEG-based Music Entrainment Brain Computer Interface," *e-Prime - Advances in Electrical Engineering, Electronics and Energy*, vol. 7, p. 100448, Mar. 2024, doi: 10.1016/j.prime.2024.100448.
- [25] J. Gu and S. Lu, "An effective intrusion detection approach using SVM with *naive Bayes* feature embedding," *Comput Secur*, vol. 103, p. 102158, Apr. 2021, doi: 10.1016/j.cose.2020.102158.