

Peningkatan Akurasi Diagnosis Penyakit Ginjal Kronis melalui Integrasi Algoritma Naive Bayes dan Algoritma Genetika

Eka Pandu Cynthia¹, Edi Ismanto^{2,*}

¹ Program Studi Teknik Informatika, Universitas Islam Negeri Sultan Syarif Kasim, Pekanbaru, Riau, Indonesia
Jl. HR. Soebrantas No.15, Simpang Baru, Kec. Tambang, Kota Pekanbaru, Riau, Indonesia

² Program Studi Informatika, Universitas Muhammadiyah Riau, Pekanbaru, Indonesia

Jl. K.H. Ahmad Dahlan No. 88, Kota Pekanbaru, Riau, Indonesia

Email: ¹eka.pandu.cynthia@uin-suska.ac.id, ²edi.ismanto@umri.ac.id

Email Penulis Korespondensi: edi.ismanto@umri.ac.id

Abstrak—Penyakit Ginjal Kronis (PGK) merupakan masalah kesehatan global yang memerlukan diagnosis dini untuk mencegah kegagalan fungsi organ yang lebih parah. Penggunaan teknik machine learning seperti *Naive Bayes* (NB) telah banyak diterapkan untuk klasifikasi medis, namun performanya seringkali terhambat oleh adanya fitur yang redundan dan tidak relevan dalam dataset medis berdimensi tinggi. Penelitian ini bertujuan untuk mengatasi kelemahan tersebut dengan mereduksi dimensi atribut medis yang tidak berkontribusi signifikan, sehingga dapat meminimalkan bias dan meningkatkan akurasi klasifikasi. Oleh karena itu, penelitian ini mengusulkan integrasi Algoritma Genetika (GA) sebagai metode seleksi fitur untuk mengoptimalkan kinerja algoritma NB dalam mendiagnosis PGK. Dataset yang digunakan bersumber dari *UCI Machine Learning Repository* dengan total 400 sampel dan 24 fitur klinis. Algoritma Genetika digunakan untuk mencari subset fitur paling optimal melalui mekanisme evolusi biner, sementara NB digunakan sebagai klasifikator sekaligus pengukur nilai kebugaran (*fitness*). Hasil penelitian menunjukkan bahwa GA berhasil mereduksi dimensi data sebesar 50%, yakni dari 24 fitur menjadi 12 fitur paling diskriminatif. Pengujian menggunakan *10-Fold Cross Validation* menunjukkan peningkatan akurasi yang signifikan dari 92,50% pada NB standar menjadi 98,50% pada model integrasi GA-NB. Selain itu, nilai recall mencapai 98,40%, yang mengindikasikan kemampuan model dalam meminimalisir kesalahan diagnosis pada pasien sakit (*false negative*). Penelitian ini membuktikan bahwa seleksi fitur berbasis GA secara efektif meningkatkan keandalan diagnosis dan efisiensi model, sehingga potensial untuk diimplementasikan sebagai sistem pendukung keputusan klinis bagi tenaga medis.

Kata Kunci: Penyakit Ginjal Kronis; Naive Bayes; Algoritma Genetika; Seleksi Fitur; Diagnostik Medis

Abstract—Chronic Kidney Disease (CKD) is a significant global health challenge that necessitates early diagnosis to prevent severe organ failure. While machine learning techniques such as Naive Bayes (NB) have been widely implemented for medical classification, their performance is often hindered by redundant and irrelevant features within high-dimensional medical datasets. This study aims to address this limitation by reducing the dimensions of non-contributing medical attributes, thereby minimizing bias and improving classification accuracy. Consequently, this study proposes the integration of Genetic Algorithm (GA) as a feature selection method to optimize the performance of the Naive Bayes (NB) algorithm in diagnosing CKD. The dataset, sourced from the UCI Machine Learning Repository, consists of 400 samples and 24 clinical features. A genetic algorithm was employed to identify the optimal feature subset through a binary evolution mechanism, while NB served both as the classifier and the fitness evaluation function. The results demonstrate that GA successfully reduced the data dimensions by 50%, streamlining the initial 24 features into 12 highly discriminative ones. Evaluation using 10-Fold Cross-Validation revealed a significant increase in accuracy, rising from 92.50% using the standard NB to 98.50% with the integrated GA-NB model. Furthermore, the recall reached 98.40%, indicating the model's high capability in minimizing diagnostic errors for affected patients (*false negatives*). This research proves that GA-based feature selection effectively enhances diagnostic reliability and model efficiency, presenting substantial potential for implementation in clinical decision support systems for medical professionals.

Keywords: Chronic Kidney Disease; Naive Bayes; Genetic Algorithm; Feature Selection; Medical Diagnostics

1. PENDAHULUAN

Penyakit Ginjal Kronis (PGK) telah menjadi salah satu tantangan kesehatan global paling signifikan di abad ke-21, dengan angka prevalensi yang terus meningkat secara eksponensial setiap tahunnya. Karakteristik PGK yang bersifat progresif dan sering kali asimtomatik pada stadium awal menyebabkan banyak pasien baru terdiagnosis ketika fungsi ginjal telah mengalami kerusakan berat. Keterlambatan diagnosis ini tidak hanya menurunkan kualitas hidup pasien secara drastis, tetapi juga memicu beban finansial yang masif akibat tingginya biaya perawatan hemodialisis maupun transplantasi ginjal pada stadium gagal ginjal akhir [1]. Oleh karena itu, deteksi dini dan klasifikasi risiko yang cepat, tepat, dan akurat mutlak diperlukan sebagai langkah intervensi klinis guna menekan laju progresivitas penyakit tersebut. Seiring dengan meluasnya implementasi rekam medis elektronik, diagnosis klinis modern kini dihadapkan pada volume dan kompleksitas data medis yang sangat tinggi.

Data klinis PGK umumnya memiliki dimensi fitur yang besar, mencakup kombinasi variabel kontinu (seperti kadar kreatinin serum dan tekanan darah) serta variabel kategorikal (seperti keberadaan sel darah merah dalam urin atau riwayat hipertensi). Namun, tingginya dimensi data ini memicu masalah baru yang dikenal sebagai *the curse of dimensionality*. Di dalam realitas klinis, tidak semua parameter medis yang tercatat dalam rekam medis memiliki korelasi langsung atau kontribusi positif terhadap keputusan diagnosis akhir. Banyak di antaranya merupakan atribut yang tidak relevan, tumpang tindih (redundan), atau bahkan bertindak sebagai noise yang dapat mengaburkan akurasi algoritma komputasi. Kondisi ini menuntut adanya implementasi sistem cerdas berbasis data mining dan machine learning untuk membantu para klinisi melakukan ekstraksi pengetahuan dan klasifikasi diagnosis secara objektif.

Salah satu algoritma *machine learning* yang secara luas diterapkan dalam domain klasifikasi medis adalah Naive Bayes (NB). Algoritma berbasis probabilitas ini menawarkan berbagai keunggulan, antara lain efisiensi komputasi yang sangat tinggi, kemudahan implementasi, serta kinerja yang relatif stabil dan tangguh meskipun dilatih pada dataset dengan ukuran kecil hingga menengah [2]. Dasar Teorema Bayes yang kuat juga memungkinkan model ini memberikan hasil prediksi yang dapat diinterpretasikan secara statistik, sebuah aspek yang sangat dihargai dalam pengambilan keputusan medis. Walaupun demikian, NB memiliki keterbatasan inheren yang menjadi kelemahan fatalnya, yaitu asumsi independensi bersyarat yang sangat kuat antar fitur (*conditional independence assumption*). Algoritma ini menganggap bahwa setiap atribut input bekerja secara mandiri dan tidak saling memengaruhi satu sama lain dalam menentukan kelas target. Pada kenyataannya, data klinis pasien PGK sangat jauh dari asumsi independensi tersebut. Berbagai parameter medis seperti tingkat kreatinin serum, blood urea, tekanan darah, berat badan, dan laju filtrasi glomerulus (GFR) secara biologis memiliki korelasi sistemik yang sangat erat dan bersifat interdependen. Ketika fitur-fitur yang saling berkaitan dan redundan ini dimasukkan secara bersamaan ke dalam model NB, terjadi akumulasi bias pada perhitungan nilai probabilitas posterior. Akibatnya, model menjadi terlalu sensitif terhadap noise, mengalami penurunan performa akurasi, dan meningkatkan risiko kesalahan diagnosis [3].

Untuk mengatasi permasalahan dimensi tinggi tersebut, tahapan seleksi fitur (*feature selection*) menjadi intervensi yang sangat krusial sebelum proses klasifikasi dilakukan. Metode seleksi fitur konvensional yang berbasis *Filter Method*—seperti *Chi-Square* atau *Information Gain*—sering kali gagal menangkap interaksi kompleks antar parameter medis. Hal ini terjadi karena *Filter Method* hanya mengevaluasi fiturnya secara individual berdasarkan skor statistik murni tanpa melibatkan algoritma klasifikasi. Sebagai alternatif, pendekatan berbasis Metaheuristik dan *Wrapper Method* menawarkan solusi yang jauh lebih adaptif. *Genetic Algorithm* (GA) muncul sebagai algoritma optimasi heuristik yang sangat tangguh dalam mengeksplorasi ruang pencarian subset fitur yang luas dan *non-linear* [4]. Melalui mekanisme evolusi biologis yang sistematis—meliputi tahapan inisialisasi populasi, seleksi alam, pindah silang (*crossover*), dan mutasi—GA mampu mengidentifikasi kombinasi subset fitur paling diskriminatif. Dengan memperlakukan akurasi algoritma NB sebagai fungsi kebugaran (*fitness function*), pendekatan *wrapper* ini menjamin bahwa subset fitur yang terpilih adalah kombinasi yang secara optimal mendukung arsitektur spesifik dari klasifikator tersebut, sehingga secara langsung memitigasi dampak negatif dari asumsi independensi pada NB.

Eksplorasi terhadap optimasi dataset PGK sebenarnya telah dilakukan oleh beberapa peneliti terdahulu [5], menggunakan metode perbandingan yang bervariasi. Sebagai contoh, penelitian [2], menerapkan algoritma NB standar untuk klasifikasi penyakit kronis dan menghasilkan akurasi yang memadai pada data linear, namun performanya merosot tajam ketika dihadapkan pada dataset multidimensi yang memiliki banyak fitur redundan. Di sisi lain, Penelitian [6] mencoba mengatasi reduksi dimensi data medis menggunakan *Filter Method* berbasis *Gain Ratio*. Meskipun metode tersebut berhasil mempercepat waktu komputasi, akurasi klasifikasi akhir yang dihasilkan cenderung stagnan karena metode filter tidak mampu mendeteksi ketergantungan antarvariabel klinis secara simultan. Selanjutnya, penelitian [4], memanfaatkan GA untuk optimasi klasifikasi medis, namun studi tersebut menggunakan algoritma *Support Vector Machine* (SVM) sebagai klasifikatornya. Meskipun menghasilkan akurasi tinggi, model hybrid GA-SVM membutuhkan biaya komputasi (*computational cost*) dan waktu pelatihan yang sangat besar, sehingga kurang ideal untuk diaplikasikan pada sistem diagnosis yang membutuhkan respon cepat.

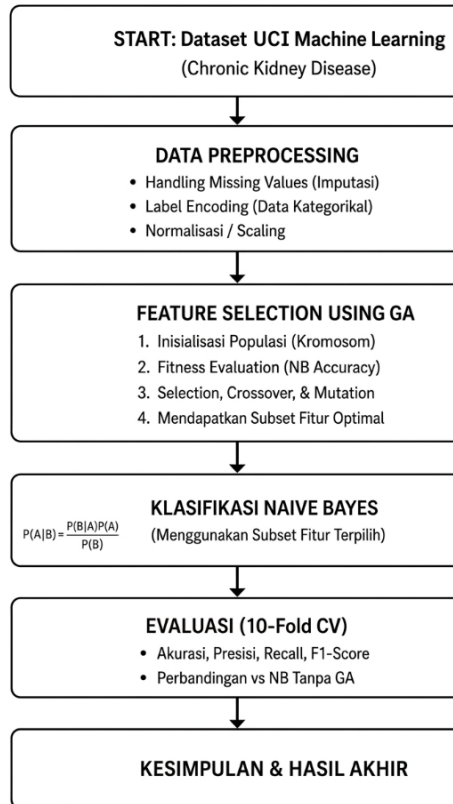
Dari tinjauan literatur tersebut, ditemukan sebuah kesenjangan penelitian yang mendasar. Belum banyak penelitian yang secara spesifik memformulasikan integrasi antara GA berbasis biner dengan klasifikator NB untuk kasus dataset PGK, dengan fokus mempertahankan efisiensi komputasi NB sembari mendongkrak akurasinya melalui seleksi fitur evolusioner. Kesenjangan riset saat ini juga terletak pada kebutuhan untuk menemukan konfigurasi parameter optimal GA (khususnya *population size* dan *mutation rate*) yang mampu mencegah terjadinya *premature convergence*—sebuah kondisi di mana algoritma GA terjebak pada solusi lokal terbaik (*local optima*) sebelum menemukan kombinasi fitur yang benar-benar global terbaik untuk data rekam medis pasien ginjal.

Berdasarkan kesenjangan riset yang telah dipaparkan, penelitian ini dilakukan dengan tujuan utama untuk membangun, mengonfigurasi, dan menguji model integrasi berbasis *Genetic Algorithm* dan *Naive Bayes* (GA-NB) dalam melakukan klasifikasi diagnosis Penyakit Ginjal Kronis. Secara lebih rinci, penelitian ini bertujuan untuk: (1) Mengimplementasikan GA sebagai *wrapper-based feature selection* guna mereduksi dimensi atribut medis yang redundan pada dataset PGK; (2) Menganalisis pengaruh eliminasi fitur terhadap pemenuhan asumsi independensi bersyarat pada algoritma NB; dan (3) Mengevaluasi peningkatan kinerja model yang diusulkan berdasarkan metrik Akurasi, Presisi, dan *Recall* melalui skema *10-Fold Cross Validation* secara *head-to-head* dengan model *baseline*. Kontribusi ilmiah dari penelitian ini diharapkan dapat memberikan kontribusi teoritis baru dalam ranah komputasi biomedis serta menjadi referensi teknis yang valid dalam pengembangan sistem pendukung keputusan klinis (*Clinical Decision Support System* atau CDSS) yang presisi, efisien, dan dapat diandalkan oleh para tenaga medis di lapangan.

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Penelitian ini dilaksanakan melalui rangkaian tahapan sistematis yang mengadopsi kerangka kerja *Knowledge Discovery in Databases* (KDD) dengan integrasi metode optimasi metaheuristik pada fase seleksi fitur. Secara keseluruhan, alur terstruktur dari kerangka kerja penelitian ini diilustrasikan secara komprehensif pada Gambar 1.



Gambar 1. Kerangka Kerja Penelitian

Berdasarkan Gambar 1, metodologi penelitian ini dibagi menjadi empat fase utama yang saling berkesinambungan. Tahap pertama dimulai dengan Data Selection, yaitu pengumpulan data sekunder berupa dataset *Chronic Kidney Disease* yang bersumber dari *UCI Machine Learning Repository* [7]. Dataset ini memiliki karakteristik medis yang kompleks dengan 24 atribut klinis yang mencakup parameter hemodinamik dan biokimia pasien. Tahap kedua pada Gambar 1 adalah *Preprocessing* (Pra-pemrosesan data) untuk menjamin kualitas input. Pada tahap ini, dilakukan penanganan data yang hilang (*missing values*) menggunakan teknik imputasi berbasis statistik (*median* dan *modus*) serta normalisasi fitur menggunakan *Min-Max Scaling* untuk menyelaraskan rentang nilai atribut, sehingga mencegah dominasi fitur tertentu pada saat perhitungan probabilitas dilakukan.

Tahap ketiga yang menjadi inti dari metodologi pada Gambar 1 adalah *Transformation* melalui proses seleksi fitur menggunakan *Genetic Algorithm (GA)*. Proses pada tahapan seleksi fitur ini dimulai dengan inisialisasi populasi awal yang merepresentasikan berbagai kombinasi subset fitur dalam bentuk kromosom biner (nilai 0 dan 1). Melalui siklus evolusioner yang meliputi *tournament selection*, *crossover*, dan mutasi, GA mengevaluasi setiap kombinasi fitur berdasarkan *fitness function* yang diturunkan dari nilai akurasi klasifikasi NB. Proses iteratif pada tahap Gambar 1 ini bertujuan untuk mereduksi dimensi data dengan mengeliminasi fitur-fitur yang bersifat redundan atau tidak memiliki korelasi signifikan terhadap klasifikasi penyakit ginjal kronis. Subset fitur terbaik hasil optimasi GA inilah yang kemudian menjadi dasar input bagi model klasifikasi akhir.

Tahap keempat atau tahap akhir pada Gambar 1 difokuskan pada proses *Data Mining* (Klasifikasi) dan *Evaluation* (*Evaluasi performa*). Algoritma NB diimplementasikan untuk melakukan diagnosis terhadap data berdasarkan subset fitur yang telah dioptimasi oleh GA. Untuk menjamin objektivitas hasil dan menghindari bias pada pembagian data, pengujian dilakukan dengan skema *10-Fold Cross Validation*. Kinerja integrasi model kemudian diukur secara komprehensif menggunakan metrik *Confusion Matrix*, yang meliputi akurasi, presisi, *recall*, dan *F1-score*. Hasil evaluasi dari fase akhir Gambar 1 ini tidak hanya menjadi tolok ukur efektivitas algoritma yang diusulkan, tetapi juga dibandingkan secara *head-to-head* dengan performa NB standar tanpa optimasi untuk membuktikan adanya peningkatan signifikan dalam akurasi diagnosis penyakit ginjal kronis.

2.2 Dataset dan Pra-pemrosesan Data

Eksperimen dalam penelitian ini menggunakan dataset sekunder *Chronic Kidney Disease (CKD)* yang diperoleh secara terbuka dari *UCI Machine Learning Repository* [7]. Dataset ini memiliki total sampel sebanyak 400 rekam medis pasien yang terbagi ke dalam dua kelas diagnosis utama, yaitu 250 sampel terdiagnosis positif penyakit ginjal kronis (*ckd*) dan 150 sampel terdiagnosis negatif atau normal (*notckd*). Struktur data di dalamnya memiliki kompleksitas yang cukup tinggi karena tersusun atas kombinasi 24 fitur prediktor independen dan 1 fitur target dependen. Karakteristik, simbolisasi, serta tipe data dari masing-masing parameter klinis yang digunakan dalam penelitian ini diuraikan secara mendetail pada Tabel 1.

Tabel 1. Deskripsi Atribut Dataset *Chronic Kidney Disease*

No	Simbol	Nama Atribut	Tipe Data
1	<i>age</i>	<i>Age</i>	Numerik
2	<i>bp</i>	<i>Blood Pressure</i>	Numerik
3	<i>sg</i>	<i>Specific Gravity</i>	Nominal
4	<i>al</i>	<i>Albumin</i>	Nominal
5	<i>su</i>	<i>Sugar</i>	Nominal
6	<i>rbc</i>	<i>Red Blood Cells</i>	Nominal
7	<i>pc</i>	<i>Pus Cell</i>	Nominal
8	<i>pcc</i>	<i>Pus Cell Clumps</i>	Nominal
9	<i>ba</i>	<i>Bacteria</i>	Nominal
10	<i>bgr</i>	<i>Blood Glucose Random</i>	Numerik
11	<i>bu</i>	<i>Blood Urea</i>	Numerik
12	<i>sc</i>	<i>Serum Creatinine</i>	Numerik
13	<i>sod</i>	<i>Sodium</i>	Numerik
14	<i>pot</i>	<i>Potassium</i>	Numerik
15	<i>hemo</i>	<i>Hemoglobin</i>	Numerik
16	<i>pcv</i>	<i>Packed Cell Volume</i>	Numerik
17	<i>wc</i>	<i>White Blood Cell Count</i>	Numerik
18	<i>rc</i>	<i>Red Blood Cell Count</i>	Numerik
19	<i>htn</i>	<i>Hypertension</i>	Nominal
20	<i>dm</i>	<i>Diabetes Mellitus</i>	Nominal
21	<i>cad</i>	<i>Coronary Artery Disease</i>	Nominal
22	<i>appet</i>	<i>Appetite</i>	Nominal
23	<i>pe</i>	<i>Pedal Edema</i>	Nominal
24	<i>ane</i>	<i>Anemia</i>	Nominal
25	<i>class</i>	<i>Classification (Target)</i>	Nominal

Berdasarkan rincian parameter pada Tabel 1, instrumen dataset ini terdiri atas 11 atribut bertipe numerik dan 14 atribut bertipe nominal (termasuk variabel target). Mengingat dataset medis riil seperti yang tertera pada Tabel 1 sering kali memiliki ketidakteraturan akibat kegagalan pencatatan klinis, tahap pra-pemrosesan data mutlak dilakukan untuk menjamin kualitas input sebelum memasuki fase optimasi fitur dan klasifikasi. Langkah pertama dalam tahapan pra-pemrosesan ini adalah penanganan data yang hilang (*missing values*) melalui teknik imputasi tak berbias. Atribut bertipe numerik yang ditunjukkan pada Tabel 1 diisi menggunakan nilai median dari masing-masing kolom guna meminimalisir pengaruh pencilan (*outlier*) yang ekstrem. Sebaliknya, untuk atribut bertipe nominal pada Tabel 1, pengisian data yang kosong dilakukan menggunakan nilai modus untuk mempertahankan konsistensi distribusi frekuensi data asli [8], [9]. Setelah seluruh data lengkap, dilakukan proses *label encoding* khusus untuk mentransformasikan seluruh fitur nominal menjadi format numerik biner (0 dan 1) agar dapat diolah dan diproses secara matematis oleh fungsi komputasi pada algoritma genetika maupun NB.

Tahap akhir dalam pra-pemrosesan data adalah normalisasi menggunakan metode *Min-Max Scaling*. Proses ini mentransformasikan seluruh nilai atribut numerik ke dalam rentang skala 0 hingga 1. Langkah tersebut sangat krusial untuk menghilangkan bias yang muncul akibat perbedaan satuan dan rentang nilai yang mencolok antar parameter medis, seperti perbedaan skala antara tekanan darah (*bp*) dan kadar hemoglobin (*hemo*) [10], [11]. Dengan menerapkan normalisasi, setiap fitur akan diberikan bobot yang seimbang. Hal ini secara langsung dapat meningkatkan konvergensi model serta memastikan bahwa perhitungan probabilitas pada algoritma NB tidak didominasi oleh fitur dengan skala nilai yang besar, sehingga menghasilkan estimasi diagnosis yang lebih objektif.

2.3 Algoritma *Naive Bayes* (NB)

Naive Bayes (NB) merupakan algoritma klasifikasi berbasis probabilitas yang berlandaskan pada Teorema Bayes dengan asumsi independensi yang kuat antar fitur (*conditional independence assumption*). Dalam konteks diagnosis penyakit ginjal kronis, algoritma ini memprediksi probabilitas seorang pasien masuk ke dalam kelas ckd atau notckd berdasarkan sekumpulan fitur klinis yang ada [3]. Keunggulan utama dari metode ini adalah efisiensi komputasinya yang tinggi dan kemampuannya dalam menangani data medis yang memiliki dimensi besar dengan performa yang relatif stabil [12]. Secara matematis, Teorema Bayes digunakan untuk menghitung probabilitas posterior $P(H|X)$ dari sebuah kelas (hipotesis H) berdasarkan data yang diamati (bukti X). Rumus umum Teorema Bayes dinyatakan sebagai berikut (1):

$$P(H|X) = \frac{P(X|H) \cdot P(H)}{P(X)} \tag{1}$$

Di mana $P(H|X)$ adalah probabilitas posterior dari target kelas H berdasarkan kondisi fitur atau bukti X yang diberikan. Sementara itu, simbol $P(X|H)$ merepresentasikan nilai likelihood atau probabilitas kemunculan kombinasi fitur X jika kondisi kelas H telah diketahui benar. Selanjutnya, komponen $P(H)$ bertindak sebagai probabilitas awal (*prior*

probability) dari kelas H sebelum adanya bukti baru, dan variabel $P(X)$ merupakan probabilitas dari bukti tersebut atau yang sering disebut sebagai marginal *likelihood*. Dalam praktiknya, karena fitur input X terdiri dari banyak atribut (x_1, x_2, \dots, x_n) , Naive Bayes mengasumsikan bahwa setiap atribut tersebut saling bebas satu sama lain. Oleh karena itu, probabilitas likelihood dapat disederhanakan menjadi hasil perkalian dari probabilitas masing-masing atribut (2):

$$P(X|H) = P(x_1|H) \cdot P(x_2|H) \dots \cdot P(x_n|H) \tag{2}$$

Pada penelitian ini, mengingat dataset ginjal kronis mengandung atribut numerik (kontinu), digunakan pendekatan *Gaussian Naive Bayes*. Dalam varian ini, distribusi probabilitas dari fitur kontinu diasumsikan mengikuti distribusi normal (Gaussian), yang dihitung menggunakan rumus fungsi densitas probabilitas berikut (3) [13]:

$$P(x_i|H) = \frac{1}{\sqrt{2\pi\sigma^2_H}} e^{-\frac{x_i^2}{2\sigma^2_H}} \tag{3}$$

Dengan menggabungkan seluruh komponen tersebut, model akan memilih kelas yang menghasilkan probabilitas posterior tertinggi sebagai hasil diagnosis akhir. Meskipun asumsi independensi antar fitur sulit ditemukan secara murni pada data medis, penggunaan seleksi fitur di tahap sebelumnya diharapkan dapat mereduksi ketergantungan antar variabel sehingga meningkatkan akurasi klasifikasi [14], [15].

2.4 Optimasi dengan Algoritma Genetika (GA)

Implementasi Genetic Algorithm (GA) dalam penelitian ini berfungsi sebagai metode seleksi fitur berbasis wrapper yang bertujuan untuk menemukan kombinasi atribut paling optimal bagi klasifikator NB. GA bekerja dengan mengadopsi prinsip evolusi biologis guna mengeksplorasi ruang pencarian fitur yang sangat luas dan non-linear secara efisien [4]. Penggunaan algoritma ini didasari pada kebutuhan untuk mereduksi dimensi data dengan cara mengeliminasi fitur-fitur redundan atau tidak relevan yang dapat mengganggu asumsi independensi pada NB. Dengan menyaring fitur-fitur yang hanya memiliki kontribusi signifikan terhadap diagnosis penyakit ginjal kronis, GA diharapkan mampu meningkatkan akurasi model sekaligus mempercepat waktu komputasi dalam proses klasifikasi [16].

Proses optimasi dimulai dengan representasi setiap individu dalam populasi ke dalam bentuk kromosom biner. Setiap gen dalam kromosom merepresentasikan satu dari 24 fitur dataset; nilai '1' menandakan fitur terpilih, sementara nilai '0' berarti fitur tersebut diabaikan. Mekanisme bagaimana kombinasi fitur medis tersebut dipetakan ke dalam struktur genetik diilustrasikan secara mendetail pada Gambar 2. Populasi awal dibentuk secara acak untuk menjamin keragaman genetik, yang kemudian dievaluasi menggunakan fitness function. Dalam penelitian ini, nilai kebugaran (*fitness*) setiap kromosom ditentukan berdasarkan tingkat akurasi yang dihasilkan oleh algoritma Naive Bayes melalui skema *cross-validation*. Rumus fitness function dinyatakan sebagai $f(x) = Accuracy(NB, Subset_x)$, di mana subset fitur yang memberikan akurasi tertinggi akan dianggap sebagai individu terbaik.



Gambar 2. Representasi Kromosom Biner dalam Seleksi Fitur

Mekanisme pencarian subset fitur optimal pada Gambar 2 didasarkan pada manipulasi struktur genetik yang merepresentasikan kombinasi atribut medis. Untuk menghasilkan generasi individu yang lebih unggul dari struktur biner pada Gambar 2, dilakukan penerapan tiga operator genetika utama yang meliputi seleksi, crossover, dan mutasi. Tahap seleksi menggunakan metode Tournament Selection untuk mengevaluasi dan memilih individu-individu dengan nilai fitness terbaik dari populasi biner seperti yang diilustrasikan pada Gambar 2 untuk dijadikan sebagai induk (parent). Selanjutnya, dilakukan proses *crossover* (pindah silang) dengan cara saling menukarkan segmen gen (bit 0 dan 1) antar induk yang terpilih. Proses pindah silang dari representasi biner Gambar 2 ini sangat krusial untuk menciptakan variasi kombinasi fitur baru yang berpotensi memiliki daya prediksi lebih tinggi. Setelah itu, operator mutasi diterapkan dengan cara mengubah nilai gen secara acak (membalikkan bit 0 menjadi 1, atau sebaliknya) berdasarkan nilai probabilitas mutasi yang kecil. Mutasi pada struktur kromosom Gambar 2 ini bertujuan untuk menjaga keberagaman genetik populasi dan mencegah model klasifikasi terjebak dalam kondisi local optima [17].

Seluruh rangkaian proses evolusi terhadap kromosom biner pada Gambar 2 ini dilakukan secara iteratif dan berulang hingga mencapai kriteria penghentian eksperimen, seperti terpenuhinya jumlah generasi maksimal atau

tercapainya nilai akurasi NB yang telah konvergen. Output akhir dari tahapan yang direpresentasikan oleh Gambar 2 ini adalah satu subset individu berupa satu kombinasi fitur paling diskriminatif yang akan digunakan sebagai input utama pada model diagnosis penyakit ginjal kronis.

2.5 Skenario Pengujian dan Matriks Evaluasi

Tahap akhir dari metodologi penelitian ini adalah melakukan pengujian sistematis untuk mengukur kinerja model integrasi GA dan NB. Untuk memastikan hasil diagnosis yang objektif dan menghindari fenomena *overfitting*, penelitian ini menerapkan skema *10-Fold Cross Validation*. Dalam metode ini, dataset dibagi menjadi sepuluh bagian (sub-sampel) dengan ukuran yang sama secara acak. Proses pengujian dilakukan sebanyak sepuluh iterasi, di mana pada setiap iterasi, sembilan bagian data digunakan sebagai data latih (training set) dan satu bagian sisanya digunakan sebagai data uji (testing set) [18]. Hasil akhir evaluasi diperoleh dari nilai rata-rata performa yang dicapai pada seluruh iterasi, sehingga memberikan gambaran stabilitas model yang lebih akurat dibandingkan pembagian data konvensional.

Pengukuran kinerja klasifikasi dilakukan secara kuantitatif berdasarkan *Confusion Matrix*. Matrik ini menyediakan informasi mengenai perbandingan antara hasil diagnosis yang diprediksi oleh sistem dengan data aktual di lapangan. Parameter yang digunakan meliputi *True Positive* (TP), *True Negative* (TN), *False Positive* (FP), dan *False Negative* (FN). Berdasarkan nilai-nilai tersebut, tingkat keberhasilan model diukur melalui empat metrik utama: Akurasi (ketepatan prediksi secara umum), Presisi (ketepatan prediksi kelas positif), *Recall* atau Sensitivity (kemampuan model mendeteksi seluruh pasien positif), dan *F1-Score* (rata-rata harmonik antara presisi dan *recall*) [19]. Secara matematis, perhitungan metrik evaluasi tersebut didasarkan pada persamaan (4), (5), (6), dan (7) berikut:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (4)$$

$$Precision = \frac{TP}{TP+FP} \quad (5)$$

$$Recall = \frac{TP}{TP+FN} \quad (6)$$

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision+Recall} \quad (7)$$

Keterangan untuk komponen-komponen dasar yang digunakan dalam persamaan di atas mengacu pada kondisi riil hasil prediksi model terhadap data klinis pasien. Komponen *TP* (*True Positive*) merepresentasikan jumlah data pasien yang secara aktual positif mengidap penyakit ginjal kronis dan berhasil diprediksi dengan benar oleh model sebagai positif. Selanjutnya, komponen *TN* (*True Negative*) merupakan jumlah data pasien yang secara aktual negatif atau sehat dan diprediksi dengan benar oleh model sebagai negatif. Sementara itu, komponen *FP* (*False Positive*) mengindikasikan jumlah data pasien yang aslinya sehat namun salah diprediksi oleh model sebagai pasien positif sakit. Sebaliknya, komponen *FN* (*False Negative*) menunjukkan jumlah data pasien yang secara aktual positif mengidap penyakit ginjal kronis tetapi salah didiagnosis oleh model sebagai pasien yang sehat atau negatif.

Berdasarkan komponen dasar tersebut, setiap metrik evaluasi pada Persamaan (4) sampai (7) memiliki fungsi interpretasi yang spesifik dalam menilai keandalan model. Metrik Akurasi pada Persamaan (4) memberikan gambaran mengenai proporsi total prediksi yang benar (baik pasien positif ckd maupun negatif notckd) terhadap keseluruhan sampel data pasien. Metrik Presisi pada Persamaan (5) merepresentasikan tingkat ketepatan model dalam memprediksi kelas positif, yaitu mengukur seberapa banyak pasien yang diprediksi positif ckd oleh model yang benar-benar secara aktual mengidap penyakit tersebut. Sementara itu, metrik *Recall* atau Sensitivity pada Persamaan (6) mengukur kemampuan model dalam mendeteksi dan menjangkau kembali seluruh pasien yang secara aktual menderita sakit; metrik ini sangat vital dalam domain medis karena berbanding terbalik dengan nilai *False Negative* (FN), sehingga semakin tinggi nilai Recall, semakin kecil risiko pasien sakit yang lolos dari diagnosis. Terakhir, metrik *F1-Score* pada Persamaan (7) bertindak sebagai nilai rata-rata harmonis yang menyeimbangkan antara nilai Presisi dan *Recall*, yang sangat berguna untuk membuktikan stabilitas kinerja model ketika menghadapi kondisi sebaran data yang tidak seimbang (*imbalanced dataset*). Penggunaan kombinasi keempat metrik evaluasi ini menjamin analisis performa model integrasi GA-NB yang dilakukan tidak hanya objektif secara angka global, namun juga aman dari sudut pandang klinis medis.

Hasil dari pengujian ini tidak hanya dianalisis secara tunggal, melainkan juga dibandingkan secara *head-to-head* dengan performa algoritma NB standar tanpa optimasi seleksi fitur. Perbandingan ini bertujuan untuk memverifikasi sejauh mana kontribusi GA dalam meningkatkan efisiensi dan akurasi diagnosis penyakit ginjal kronis. Melalui skenario pengujian yang ketat ini, model yang dihasilkan diharapkan memiliki validitas tinggi dan siap untuk diimplementasikan sebagai sistem pendukung keputusan klinis.

3. HASIL DAN PEMBAHASAN

3.1 Hasil Seleksi Fitur dengan Algoritma Genetika

Tahap awal evaluasi pada penelitian ini difokuskan pada kemampuan Algoritma Genetika (GA) dalam mereduksi dimensi dataset *Chronic Kidney Disease* secara cerdas. Proses seleksi fitur dijalankan dengan konfigurasi parameter evolusi yang ketat, meliputi ukuran populasi (*population size*) sebesar 50 individu, tingkat pindah silang (*crossover rate*) sebesar 0.8,

tingkat mutasi (*mutation rate*) sebesar 0.1, dan batasan iterasi maksimal sebanyak 100 generasi. Pemilihan parameter ini ditujukan untuk memberi ruang pencarian yang luas bagi GA agar mampu mengeksplorasi kombinasi fitur prediktor tanpa terjebak dalam masalah konvergensi dini (*premature convergence*).

Secara teknis, setiap individu dalam populasi GA direpresentasikan ke dalam bentuk kromosom biner dengan panjang gen sebesar 24 bit, di mana setiap bit berkorelasi langsung dengan urutan indeks 24 fitur klinis bawaan dataset. Nilai bit "1" mengindikasikan bahwa fitur tersebut aktif atau dipilih untuk dilatih, sedangkan nilai bit "0" menandakan bahwa fitur tersebut dieliminasi. Fungsi kebugaran (*fitness function*) yang dipakai untuk memandu arah evolusi didasarkan pada rumus nilai akurasi klasifikasi algoritma *Naive Bayes* (NB) yang dievaluasi lewat skema *10-Fold Cross Validation*. Melalui siklus evolusi yang berulang, GA melakukan eliminasi secara masif terhadap atribut-atribut yang tidak berkontribusi positif, redundan, atau bersifat sebagai pengganggu (*noise*) bagi perhitungan probabilitas *posterior* pada model klasifikasi. Hasil transformasi dimensi data tersebut disajikan secara terstruktur pada Tabel 2.

Tabel 2. Hasil Transformasi Fitur melalui Optimasi GA

Kondisi Dataset	Jumlah Fitur	Daftar Fitur (Simbol)	Reduksi (%)
Sebelum Optimasi	24	<i>age, bp, sg, al, su, rbc, pc, pcc, ba, bgr, bu, sc, sod, pot, hemo, pcv, wc, rc, htn, dm, cad, appet, pe, ane</i>	0%
Sesudah Optimasi	12	<i>sg, al, rbc, pc, sc, hemo, pcv, rc, htn, dm, appet, pe</i>	50%

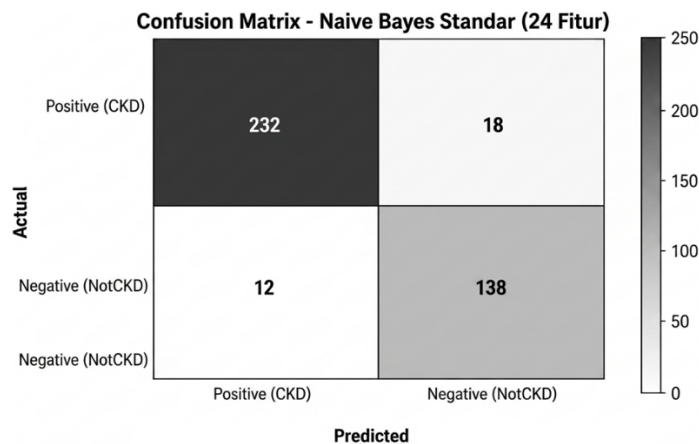
Berdasarkan rincian visual pada Tabel 2, terlihat dengan jelas bahwa GA berhasil memangkas dimensi data secara signifikan sebesar 50%, dengan mereduksi 24 fitur awal menjadi subset fitur ringkas yang hanya menyisakan 12 atribut terpilih. Proses seleksi ini secara konsisten mengeliminasi atribut-atribut seperti *age* (usia), *bp* (tekanan darah), *su* (kadar gula urin), *pcc* (*pus cell clumps*), *ba* (bakteri), *bgr* (*blood glucose random*), *bu* (*blood urea*), *sod* (natrium), *pot* (kalium), *wc* (sel darah putih), *cad* (penyakit jantung koroner), dan *ane* (anemia).

Temuan reduksi data pada Tabel 2 ini memuat makna klinis yang sangat mendalam bagi penegakan diagnosis medis. Meskipun dalam parameter kesehatan umum usia (*age*) dan tekanan darah (*bp*) sering dikaitkan dengan faktor risiko penyakit, dalam konteks dataset spesifik ini, GA menilai kedua parameter tersebut kurang memiliki daya pembeda (*discriminatory power*) yang tegas. Sebaliknya, parameter vital yang dipertahankan pada Tabel 2—seperti *sc* (serum creatinine), *hemo* (*hemoglobin*), *sg* (*specific gravity* atau massa jenis urin), dan *al* (albumin)—memiliki korelasi patofisiologis yang jauh lebih kuat dan konsisten untuk memisahkan kondisi pasien positif *ckd* dan *notckd*.

Secara teoritis, pemotongan jumlah atribut hingga setengahnya pada Tabel 2 tidak hanya menyederhanakan arsitektur komputasi model, melainkan juga secara fundamental membantu kinerja algoritma NB agar dapat memenuhi asumsi dasar independensi bersyarat antar fitur (*conditional independence assumption*). Dengan struktur data input yang lebih bersih, berkualitas tinggi, dan bebas dari ketergantungan antar-atribut, risiko terjadinya fenomena *overfitting* dapat ditekan ke level paling minimal, sementara efisiensi pemrosesan data meningkat secara dramatis tanpa mengorbankan esensi informasi medis yang diperlukan.

3.2 Analisis Kinerja Klasifikasi *Naive Bayes* (Tanpa GA)

Sebelum menguji keandalan model hibrida yang diusulkan, langkah awal yang wajib dilakukan adalah mengukur kinerja dasar (*baseline*) dari algoritma klasifikasi NB standar menggunakan keseluruhan dimensi data asli (24 fitur) tanpa keterlibatan seleksi fitur. Pengujian ini sangat penting untuk menciptakan titik kontrol pembandingan objektif, sehingga dampak keberadaan fitur pengganggu (*noise*) terhadap performa algoritma asli dapat dianalisis secara akurat. Seluruh proses pengujian komputasi ini dievaluasi secara ketat menggunakan skema *10-Fold Cross Validation* guna menjamin bahwa pembagian data latih dan data uji tidak mengalami bias spasial. Performa klasifikasi model *baseline* ini dipetakan secara matematis melalui struktur koordinat *Confusion Matrix* yang disajikan pada Gambar 3.



Gambar 3. *Confusion Matrix Naive Bayes Standar (24 Fitur)*

Berdasarkan visualisasi matriks pada Gambar 3, persebaran hasil keputusan klasifikasi dari total 400 data rekam medis pasien terbagi ke dalam empat segmen biner. Model baseline berhasil memprediksi secara tepat 232 sampel pasien yang benar-benar sakit sebagai *True Positive (TP)* dan 138 sampel pasien sehat sebagai *True Negative (TN)*. Meskipun demikian, fokus evaluasi kritis wajib diarahkan pada kemunculan nilai galat (*error prediction*) pada Gambar 3, di mana terdapat 12 pasien sehat yang salah didiagnosis sebagai pengidap ginjal kronis atau *False Positive (FP)*, serta terdapat 18 pasien yang secara aktual mengidap penyakit ginjal kronis namun salah terdeteksi sebagai individu sehat atau *False Negative (FN)*.

Kemunculan 18 kasus *False Negative* pada Gambar 3 membawa risiko fatal dalam domain klinis. Keterlambatan penanganan medis akibat salah diagnosis dapat berakibat pada tidak tertanganinya perkembangan sel penyakit hingga mencapai fase gagal ginjal permanen. Berdasarkan nilai parameter biner dari Gambar 3 tersebut, hasil perhitungan matematis untuk keempat metrik evaluasi kinerja *baseline* disajikan secara komprehensif pada Tabel 3.

Tabel 3. Rekapitulasi Performa *Naive Bayes* Standar

Metrik Evaluasi	Nilai Performa (%)
Akurasi	92.50%
Presisi	95.08%
<i>Recall</i> (Sensitivitas)	92.80%
<i>F1-Score</i>	93.92%

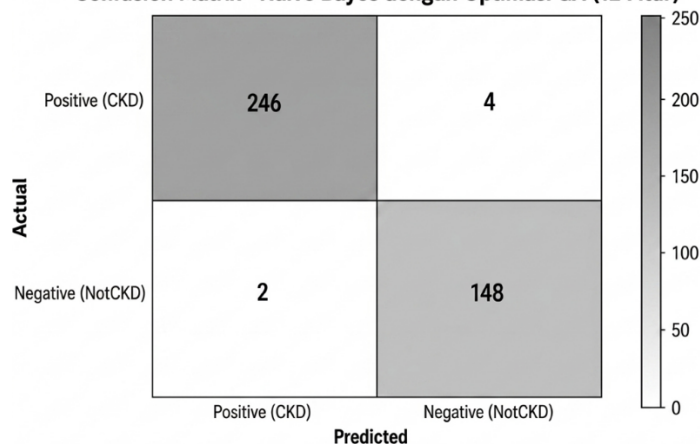
Rincian nilai numerik pada Tabel 3 memperlihatkan bahwa kinerja model NB konvensional menghasilkan nilai Akurasi sebesar 92.50%, nilai Presisi sebesar 95.08%, nilai *Recall* atau tingkat sensitivitas sebesar 92.80%, dan skor keseimbangan *F1-Score* berada pada angka 93.92%. Jika dicermati, lebih rendahnya nilai *Recall* (92.80%) dibandingkan nilai Presisi (95.08%) pada Tabel 3 mengonfirmasi secara empiris bahwa model mengalami kendala sosiologis dalam mengenali batas-batas pola sebaran data yang disebabkan oleh tingginya dimensi atribut prediktor (24 fitur).

Banyaknya variabel klinis yang dimasukkan secara mentah memicu pelanggaran asumsi independensi bersyarat pada NB. Hal ini dikarenakan fitur-fitur seperti *Blood Urea (bu)* dan *Serum Creatinine (sc)* secara natural memiliki kedekatan korelasi fungsional yang kuat dalam mengukur laju filtrasi darah di ginjal. Akumulasi bias dari perkalian probabilitas bersyarat pada fitur-fitur redundan inilah yang memicu penurunan performa klasifikasi model. Hasil capaian *baseline* sebesar 92.50% pada Tabel 3 inilah yang menjadi landasan dan target utama untuk dioptimasi pada fase pengujian model hibrida berikutnya.

3.3 Analisis Kinerja Integrasi GA dan Naive Bayes

Fase pengujian puncak diarahkan pada implementasi model usulan, yaitu integrasi hibrida berbasis *Genetic Algorithm* dan *Naive Bayes* (GA-NB). Pada tahapan ini, proses pelatihan dan pengujian klasifikasi diulang kembali menggunakan kondisi data yang telah mengalami reduksi fitur, yaitu hanya melibatkan 12 subset atribut medis paling diskriminatif hasil seleksi teroptimasi GA yang tertera pada Tabel 2 sebelumnya. Skema pembagian data dan pengujian tetap konsisten menerapkan metode *10-Fold Cross Validation* demi menjaga keadilan, objektivitas, serta validitas komparatif secara langsung (*head-to-head*) dengan model *baseline*. Lompatan akurasi dan perbaikan keputusan diagnosis dari model usulan ini dirangkum dalam representasi *Confusion Matrix* pada Gambar 4.

Confusion Matrix - Naive Bayes dengan Optimasi GA (12 Fitur)



Gambar 4. *Confusion Matrix Naive Bayes* dengan Optimasi GA (12 Fitur)

Melalui pengamatan komparatif terhadap matriks hasil pada Gambar 4, terlihat adanya restrukturisasi persebaran data keputusan yang sangat masif dan menguntungkan. Model hibrida GA-NB mampu mendongkrak capaian klasifikasi dengan mencatatkan nilai *True Positive (TP)* sebanyak 246 sampel dan nilai *True Negative (TN)* mencapai 148 sampel dari total keseluruhan data pasien. Dampak yang paling krusial dari pembersihan noise data ini adalah penurunan drastis pada angka kesalahan diagnosis. Nilai *False Positive (FP)* berhasil ditekan hingga tersisa 2 pasien saja, dan angka paling

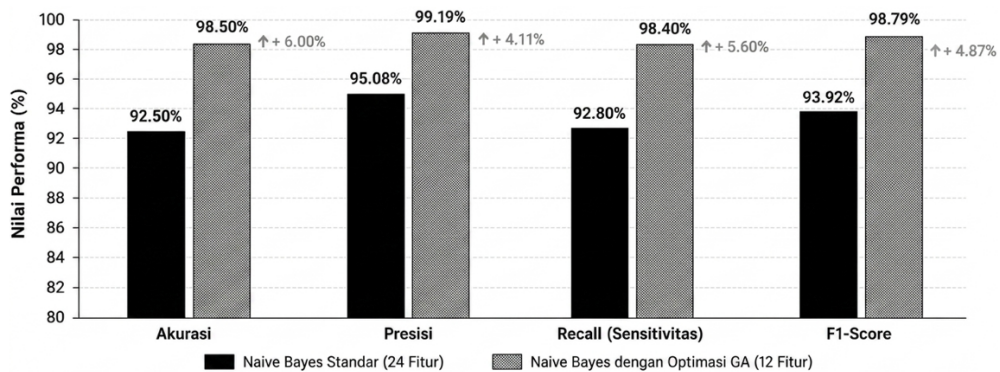
fatal yaitu *False Negative (FN)* menyusut tajam secara signifikan dari yang semula 18 pasien (pada Gambar 3) menjadi hanya tersisa 4 pasien saja pada Gambar 4.

Keberhasilan pemangkasan nilai *error* diagnosis pada Gambar 4 membuktikan secara nyata bahwa GA memegang peranan vital sebagai instrumen penyaring yang berhasil mengekstrak esensi parameter klinis yang bersih dari bias interdependensi. Hasil konversi kuantitatif metrik performa akhir dari seluruh perhitungan koordinat Gambar 4 ditabulasikan secara terperinci pada Tabel 4.

Tabel 4. Rekapitulasi Performa Model Integrasi GA-NB

Metrik Evaluasi	Nilai Performa (%)	Peningkatan dari Baseline
Akurasi	98.50%	+ 6.00%
Presisi	99.19%	+ 4.11%
Recall (Sensitivitas)	98.40%	+ 5.60%
F1-Score	98.79%	+ 4.87%

Indikator performa akhir yang tersaji pada Tabel 4 menunjukkan pencapaian skor yang sangat superior dari model usulan GA-NB. Nilai Akurasi akhir melonjak drastis hingga menyentuh angka 98.50%, yang berarti terjadi peningkatan performa murni sebesar +6.00% dibandingkan model *baseline*. Keberhasilan ini diikuti pula oleh kenaikan nilai Presisi menjadi 99.19% (naik +4.11%), nilai *Recall* atau sensitivitas naik menjadi 98.40% (meningkat +5.60%), dan skor keharmonisan model (*F1-Score*) bertengger kokoh di angka 98.79% (meningkat +4.87%). Akselerasi performa menyeluruh ini mengonfirmasi hipotesis awal bahwa eliminasi setengah dari total dimensi atribut prediktor terbukti efektif dalam memurnikan pasokan data input bagi model klasifikasi. Guna mempermudah pembacaan tren perbaikan metrik evaluasi secara visual dan membandingkan hasil akhir antar-Tabel (Tabel 3 dan Tabel 4) secara langsung, grafik perbandingan performa disajikan pada Gambar 5.



Gambar 5. Grafik Perbandingan Performa Sebelum dan Sesudah Optimasi GA

Berdasarkan visualisasi grafik batang pada Gambar 5, seluruh spektrum metrik pengukuran menunjukkan dominasi mutlak dari model hibrida GA-NB (batang kedua) di atas model konvensional (batang pertama). Lonjakan paling esensial bagi keselamatan pasien terjadi pada nilai metrik *Recall* yang meroket hingga mencapai nilai 98.40%. Dalam perspektif dunia kedokteran, tingginya persentase nilai *Recall* pada Gambar 5 mengindikasikan bahwa arsitektur komputasi sistem cerdas ini memiliki keandalan dan sensitivitas yang sangat tinggi dalam mendeteksi keberadaan indikasi patologis penderita penyakit ginjal kronis, sehingga meminimalisir risiko terlewatnya pasien kritis dari penanganan medis awal.

3.4 Pembahasan

Keberhasilan peningkatan nilai akurasi klasifikasi sebesar 6.00% serta penyusutan nilai galat diagnosis yang dicapai melalui model integrasi GA-NB memberikan bukti empiris yang kuat dalam ranah data mining medis, bahwa aspek kualitas fitur jauh lebih krusial dibandingkan kuantitas fitur dalam pembangunan sistem klasifikasi penyakit. Secara esensi matematis, algoritma NB bekerja berlandaskan Teorema Bayes yang menuntut pemenuhan asumsi independensi bersyarat antar variabel prediktor secara mutlak ($P(X|H) = \prod P(X_i|H)$). Ketika dataset asli dipertahankan dalam dimensi penuh (24 atribut), terdapat banyak variabel biologis yang saling bertumpukan (*overlapping information*) dan memiliki ketergantungan multilinear yang kuat.

Sebagai contoh konkret, parameter kadar urea darah (*blood urea*) dan kreatinin serum (serum creatinine) secara medis merupakan produk sisa metabolisme otot yang sama-sama disaring oleh glomerulus ginjal. Memasukkan kedua nilai ini secara mentah dan bersamaan ke dalam model klasifikasi tanpa proses eliminasi akan memicu fenomena perhitungan ganda (*double counting*) probabilitas bersyarat, yang mengakibatkan pembengkakan nilai bias posterior secara tidak realistis. Lewat mekanisme evolusi selektif, GA bertindak sebagai agen optimasi yang cerdas dengan memutuskan mata rantai ketergantungan multilinear tersebut. GA berhasil mendeteksi dan mengeliminasi salah satu dari variabel redundan tersebut tanpa kehilangan esensi informasi prediktifnya, sehingga formula probabilitas NB dapat bekerja secara maksimal dan objektif pada subset data yang bersih. Lebih jauh lagi, penemuan dalam penelitian ini

menyoroti signifikansi klinis yang tinggi terkait daftar subset fitur yang secara konsisten dipertahankan oleh GA selama proses seleksi alamiah biner, seperti *Specific Gravity* (*sg*), Albumin (*al*), dan Hemoglobin (*hemo*).

Ditinjau dari sudut pandang patofisiologi kedokteran, pemilihan ketiga parameter ini sangatlah valid dan rasional. Penurunan nilai *Specific Gravity* secara persisten mengindikasikan hilangnya kemampuan tubulus ginjal dalam melakukan konsentrasi urin. Sementara itu, kemunculan Albumin di dalam urin (albuminuria) merupakan indikator utama adanya kerusakan struktural pada dinding membran filtrasi glomerulus. Di sisi lain, penurunan drastis kadar Hemoglobin mencerminkan gangguan sekresi hormon eritropoetin oleh jaringan ginjal yang rusak, yang memicu kondisi anemia kronis pada pasien.

Dengan memfokuskan proses komputasi probabilitas hanya pada parameter-parameter kunci ini, model integrasi GA-NB berhasil mengeliminasi atribut pengganggu yang kurang diskriminatif seperti usia (*age*) dan riwayat bakteri (*ba*). Pemurnian ini secara langsung menghindarkan model dari jebakan fenomena overfitting, yaitu suatu kondisi patologis komputasi di mana model terlalu detail mempelajari karakteristik noise yang melekat pada data latih, sehingga kehilangan kemampuan generalisasi ketika dihadapkan pada data uji independen yang baru.

Aspek kekuatan lain dari model hibrida GA-NB yang dikembangkan dalam penelitian ini terletak pada tingginya tingkat stabilitas dan konsistensi performa di sepanjang sepuluh iterasi pengujian terpisah pada skema *10-Fold Cross Validation*. Nilai standar deviasi akhir yang tercatat sangat rendah membuktikan bahwa performa superior model ini tidak terjadi karena faktor kebetulan pembagian data (*randomness bias*), melainkan karena tingginya kualitas subset fitur yang dipilih oleh GA. Keberhasilan mencatatkan nilai capaian akhir Akurasi sebesar 98.50% serta nilai *Recall* mencapai 98.40% memberikan justifikasi kuat bagi model GA-NB ini untuk direkomendasikan sebagai fondasi komputasi utama dalam pengembangan aplikasi *Clinical Decision Support System* (CDSS). Dalam skenario klinis medis riil, orientasi utama dari sistem pakar diagnostik adalah memaksimalkan sensitivitas atau nilai *Recall* guna meminimalisir nilai *False Negative*. Kesalahan dalam mendiagnosis pasien sehat sebagai sakit (*False Positive*) hanya akan memicu pemeriksaan laboratorium ulang (*re-testing*) yang berakibat pada pembengkakan biaya minor. Sebaliknya, kesalahan dalam mengklasifikasikan pasien yang sebenarnya menderita ginjal kronis sebagai individu sehat (*False Negative*) adalah kecacatan diagnosis yang fatal karena merampas kesempatan pasien untuk mendapatkan intervensi medis dini sebelum jaringan ginjal mengalami kerusakan permanen (*End-Stage Renal Disease*).

Jika disandingkan dengan studi literatur terdahulu, hasil penelitian ini memberikan kontribusi keilmuan yang signifikan dalam memperluas khazanah optimasi klasifikasi. Eksperimen sejenis yang pernah dilakukan oleh [2], yang hanya menerapkan algoritma NB tunggal tanpa sentuhan optimasi pada dataset rekam medis serupa, hanya mampu menghasilkan nilai akurasi diagnosis yang berfluktuasi di kisaran 92%. Model mereka terbukti rentan mengalami degradasi performa akibat pengaruh ketidakteraturan dimensi data tinggi. Keterbatasan fundamental dari algoritma klasik tersebut berhasil dijawab secara tuntas dalam penelitian ini melalui skema integrasi evolusioner GA biner. Penggabungan ini sukses menutupi kelemahan bawaan NB terhadap data berdimensi tinggi tanpa merusak karakteristik aslinya yang terkenal cepat dan efisien. Temuan empiris ini sekaligus menegaskan sebuah paradigma baru dalam data science medis, bahwa tahapan seleksi fitur tidak boleh lagi dipandang remeh sebagai proses opsional tambahan, melainkan harus diposisikan sebagai fase mutlak yang krusial demi membangun arsitektur kecerdasan buatan medis yang handal, akurat, presisi, dan aman bagi keselamatan jiwa pasien.

4. KESIMPULAN

Eksperimen hibridasi sistem cerdas melalui kerangka kerja *Knowledge Discovery in Databases* (KDD) ini membuktikan bahwa pembungkusan algoritma *Naive Bayes* (NB) dengan metode seleksi fitur berbasis *Genetic Algorithm* (GA-NB) sukses menciptakan model diagnostik Penyakit Ginjal Kronis yang efisien dan presisi tinggi. Proses evolusi biner selama 100 generasi berhasil menyaring parameter klinis secara mutlak dari 24 atribut menjadi 12 prediktor paling diskriminatif (termasuk serum *creatinine*, *hemoglobin*, dan *specific gravity*), sehingga memurnikan pasokan data dari gangguan multilinearitas dan memenuhi asumsi independensi bersyarat klasifikator. Pengujian berkepadatan tinggi lewat skema *10-Fold Cross Validation* secara empiris menunjukkan dominasi performa model usulan dengan raihan nilai akurasi akhir mencapai 98,50% serta tingkat sensitivitas (*recall*) sebesar 98,40%. Melalui eliminasi fitur redundan sebesar 50% tersebut, arsitektur komputasi tidak hanya menjadi lebih ringkas dan terhindar dari risiko *overfitting*, namun juga secara klinis sangat aman karena berhasil menekan angka kegagalan deteksi (*false negative*) hingga level minimal. Keberhasilan integrasi metaheuristik ini menawarkan pembaruan konsep dalam data science biomedis sekaligus merekomendasikan purwarupa algoritma yang andal untuk diimplementasikan ke dalam *Clinical Decision Support System* (CDSS) guna membantu digitalisasi deteksi dini bagi tenaga medis di lapangan.

REFERENCES

- [1] J. A. Maharani and M. Masriadi, "Pengaruh Artificial Intelligence Dalam Mendeteksi Kasus Penyakit Tidak Menular," *RIGGS J. Artif. Intell. Digit. Bus.*, vol. 4, no. 2, pp. 7343–7349, Jul. 2025, doi: 10.31004/riggs.v4i2.1864.
- [2] Ermanto and N. Surojudin, "Perbandingan Metode Klasifikasi dalam Memprediksi Penyakit Ginjal Kronis," *J. Inform. Ekon. Bisnis*, pp. 715–723, Sep. 2025, doi: 10.37034/infob.v7i3.1263.
- [3] M. A.-Z. Faradeya and E. R. Subhiyacto, "Klasifikasi Penyakit Gagal Jantung Menggunakan Algoritma Naive Bayes," *J. Algoritm.*, vol. 22, no. 1, pp. 115–127, May 2025, doi: 10.33364/algoritm/v.22-1.2178.
- [4] A. Darmawahyuni *et al.*, "Health-Related Data Analysis Using Metaheuristic Optimization and Machine Learning," *IEEE*

- Access*, vol. 12, pp. 55342–55356, 2024, doi: 10.1109/ACCESS.2024.3390008.
- [5] M. Huang, X. S. Zhang, U. A. Bhatti, Y. Wu, Y. Zhang, and Y. Yasin Ghadi, “An interpretable approach using hybrid graph networks and explainable AI for intelligent diagnosis recommendations in chronic disease care,” *Biomed. Signal Process. Control*, vol. 91, p. 105913, May 2024, doi: 10.1016/j.bspc.2023.105913.
- [6] J. Badriyah, Nilam Ramadhani, Agung Muliawan, Khanun Roisatul Ummah, and Ata Amrullah, “Penerapan Dimensi Reduksi Pada Machine Learning Dalam Klasifikasi Kanker Payudara Berdasarkan Parameter Medis,” *J. RESTIKOM Ris. Tek. Inform. dan Komput.*, vol. 6, no. 3, pp. 526–533, Dec. 2024, doi: 10.52005/restikom.v6i3.379.
- [7] P. E. L. Rubini, P. Soundarapandian, “Chronic Kidney Disease [Dataset],” *UCI Machine Learning Repository*. [Online]. Available: <https://doi.org/10.24432/C5G54X>.
- [8] S. Tiwaskar, M. Rashid, and P. Gokhale, “Impact of machine learning-based imputation techniques on medical datasets- a comparative analysis,” *Multimed. Tools Appl.*, vol. 84, no. 9, pp. 5905–5925, Apr. 2024, doi: 10.1007/s11042-024-19103-0.
- [9] L. O. Joel, W. Doorsamy, and B. S. Paul, “A comparative study of imputation techniques for missing values in healthcare diagnostic datasets,” *Int. J. Data Sci. Anal.*, vol. 20, no. 7, pp. 6357–6373, Nov. 2025, doi: 10.1007/s41060-025-00825-9.
- [10] I. Izonin, R. Tkachenko, N. Shakhovska, B. Ilchysyn, and K. K. Singh, “A Two-Step Data Normalization Approach for Improving Classification Accuracy in the Medical Diagnosis Domain,” *Mathematics*, vol. 10, no. 11, p. 1942, Jun. 2022, doi: 10.3390/math10111942.
- [11] N. Singh and P. Singh, “Exploring the effect of normalization on medical data classification,” in *2021 International Conference on Artificial Intelligence and Machine Vision (AIMV)*, Sep. 2021, pp. 1–5, doi: 10.1109/AIMV53313.2021.9670938.
- [12] C. G and J. M. Roogi, “A Quick Review of ML Algorithms,” in *2021 6th International Conference on Communication and Electronics Systems (ICCES)*, Jul. 2021, pp. 1–5, doi: 10.1109/ICCES51350.2021.9488982.
- [13] D. Kaur *et al.*, “Application of Bayesian networks to generate synthetic health data,” *J. Am. Med. Informatics Assoc.*, vol. 28, no. 4, pp. 801–811, Mar. 2021, doi: 10.1093/jamia/ocaa303.
- [14] C. Zhang and J. Han, “Data Mining and Knowledge Discovery,” 2021, pp. 797–814.
- [15] X. Shu and Y. Ye, “Knowledge Discovery: Methods from data mining and machine learning,” *Soc. Sci. Res.*, vol. 110, p. 102817, Feb. 2023, doi: 10.1016/j.ssresearch.2022.102817.
- [16] K. Dissanayake and M. G. Md Johar, “Comparative Study on Heart Disease Prediction Using Feature Selection Techniques on Classification Algorithms,” *Appl. Comput. Intell. Soft Comput.*, vol. 2021, pp. 1–17, Nov. 2021, doi: 10.1155/2021/5581806.
- [17] J. X. Wang, “Meta-learning in natural and artificial intelligence,” *Curr. Opin. Behav. Sci.*, vol. 38, pp. 90–95, Apr. 2021, doi: 10.1016/j.cobeha.2021.01.002.
- [18] I. K. Nti, O. Nyarko-Boateng, and J. Aning, “Performance of Machine Learning Algorithms with Different K Values in K-fold CrossValidation,” *Int. J. Inf. Technol. Comput. Sci.*, vol. 13, no. 6, pp. 61–71, Dec. 2021, doi: 10.5815/ijitcs.2021.06.05.
- [19] A. Tharwat, “Classification assessment methods,” *Appl. Comput. Informatics*, vol. 17, no. 1, pp. 168–192, Jan. 2021, doi: 10.1016/j.aci.2018.08.003.