

Optimasi Hyperparameter Grid Search untuk Multi-Algoritma Klasifikasi Tingkat Hidrasi pada Daily Water Intake

Hyperparameter Grid Search Optimization for Multi-Algorithm Hydration Level Classification on Daily Water Intake

Julfiana Rizkiah Futri^{1,*}, Heni Sulistiani²

¹ Fakultas Teknik dan Ilmu Komputer, Sistem Informasi, Universitas Teknokrat Indonesia, Lampung, Indonesia

² Fakultas Teknik dan Ilmu Komputer, Magister Ilmu Komputer, Universitas Teknokrat Indonesia, Lampung, Indonesia

Email: ^{1,*}julfiana_rizkiah_futri@teknokrat.ac.id, ²henisulistiani@teknokrat.ac.id

Email Penulis Korespondensi: julfiana_rizkiah_futri@teknokrat.ac.id

Received: 02/07/2026, Revised: 18/07/2026, Accepted: 01/09/2026, Published: 08/09/2026

Abstrak—Air merupakan elemen vital bagi kelangsungan fungsi fisiologis manusia, mulai dari proses metabolisme hingga performa kognitif dan fisik. Kurangnya konsumsi air harian dapat memicu berbagai gangguan kesehatan yang menurunkan kualitas hidup individu secara signifikan. *Machine learning* menyediakan pendekatan berbasis data untuk mengklasifikasikan tingkat hidrasi secara lebih objektif berdasarkan atribut individu dan lingkungan. Penelitian ini mengevaluasi dan membandingkan kinerja tiga algoritma klasifikasi *Random Forest*, *Naive Bayes*, dan *K-Nearest Neighbor* dalam memprediksi status hidrasi (*Good* atau *Poor*) menggunakan dataset publik *Daily Water Intake* yang terdiri dari 30.000 sampel. Ketidakseimbangan kelas pada dataset (79,7% *Good* dan 20,3% *Poor*) diatasi melalui SMOTE, sementara *Stratified K-Fold Cross Validation* dengan K=10 digunakan untuk evaluasi model. Optimasi *hyperparameter* dilakukan menggunakan *Grid Search*, dan performa model dinilai melalui metrik akurasi, presisi, *recall*, *F1-Score*, dan AUC-ROC. Hasil penelitian membuktikan bahwa *Random Forest* mencapai akurasi tertinggi sebesar 99,50% dan AUC-ROC 0,9999 setelah optimasi berkat kemampuan ensemblenya menangkap hubungan non-linear antarfitur, unggul dibandingkan *Naive Bayes* yang dibatasi oleh asumsi independensi fitur. KNN menunjukkan peningkatan performa paling signifikan pascaoptimasi dengan kenaikan akurasi 0,62% (dari 97,62% menjadi 98,24%), sementara *Naive Bayes* tidak mengalami perubahan dengan akurasi tetap 83,53% karena parameter optimalnya identik dengan nilai *default*. Temuan ini memberikan panduan berbasis bukti untuk pemilihan algoritma dalam sistem klasifikasi tingkat hidrasi.

Kata Kunci: Random Forest; Naive Bayes; K-Nearest Neighbor; Grid Search; Tingkat Hidrasi; SMOTE; Cross Validation

Abstract—Water is a vital element for sustaining human physiological functions, ranging from metabolic processes to cognitive and physical performance. Insufficient daily water intake can trigger various health problems that significantly reduce individual well being. Machine learning provides a data driven approach to classify hydration levels more objectively based on individual and environmental attributes. This study evaluates and compares the performance of three classification algorithms Random Forest, Naive Bayes, and K-Nearest Neighbor in predicting hydration status (Good or Poor) using the Daily Water Intake public dataset comprising 30,000 samples. Class imbalance in the dataset (79.7% Good, 20.3% Poor) was addressed through SMOTE, while Stratified K-Fold Cross Validation with K=10 was employed for model evaluation. Hyperparameter optimization was performed using Grid Search, and model performance was assessed through accuracy, precision, recall, F1-Score, and AUC-ROC metrics. Results demonstrate that Random Forest achieved the highest accuracy of 99.50% and AUC-ROC of 0.9999 after optimization thanks to its ensemble mechanism, which captures non-linear relationships among features, outperforming Naive Bayes, which is constrained by its feature-independence assumption. KNN showed the most notable improvement post optimization with an accuracy gain of 0.62% (from 97.62% to 98.24%), while Naive Bayes remained unchanged at 83.53% as its optimal parameter matched the default value. These findings offer evidence based guidance for algorithm selection in hydration classification systems.

Keywords: Random Forest; Naive Bayes; K-Nearest Neighbor; Grid Search; Hydration Level; SMOTE; Cross Validation

1. PENDAHULUAN

Hidrasi yang memadai merupakan salah satu pilar utama dalam menjaga kesehatan tubuh manusia. Air berperan dalam hampir seluruh proses biologis, termasuk regulasi suhu tubuh, transportasi nutrisi, serta mendukung fungsi kognitif dan fisik yang optimal. Ketika asupan cairan tidak mencukupi, berbagai gangguan kesehatan dapat muncul, mulai dari penurunan konsentrasi dan kelelahan hingga gangguan fungsi organ yang lebih serius [1]. Oleh sebab itu, kemampuan untuk memantau dan mengklasifikasikan tingkat hidrasi seseorang secara akurat menjadi aspek yang semakin relevan dalam konteks kesehatan berbasis data.

Pendekatan konvensional dalam penilaian tingkat hidrasi umumnya masih bersifat subjektif dan tidak mempertimbangkan keragaman karakteristik individu, seperti usia, berat badan, intensitas aktivitas fisik, maupun kondisi lingkungan sekitar. Ketidakakuratan pendekatan ini berpotensi menghasilkan rekomendasi hidrasi yang kurang tepat sasaran [2]. Kemajuan di bidang *machine learning* membuka peluang untuk mengembangkan model prediktif yang mampu mengklasifikasikan tingkat hidrasi secara lebih objektif, terukur, dan berbasis data nyata.

Kajian terhadap literatur klasifikasi data kesehatan berbasis *machine learning* di Indonesia memperlihatkan bahwa kombinasi *Random Forest*, *Naive Bayes*, dan KNN telah banyak digunakan untuk mengolah data fisiologis dan perilaku individu menjadi kategori status kesehatan tertentu. Salah satu penelitian membandingkan *Naive Bayes* dan KNN untuk klasifikasi data kesehatan dan menunjukkan bahwa performa kedua algoritma sangat dipengaruhi oleh karakteristik fitur yang digunakan [3]. Penelitian lain yang menerapkan KNN pada data kesehatan turut menegaskan

bahwa akurasi model sangat bergantung pada pemilihan parameter dan proporsi kelas dalam data latih [4]. Studi serupa yang membandingkan KNN dengan algoritma lain pada data kesehatan juga menunjukkan bahwa strategi praproses memberikan pengaruh besar terhadap keandalan hasil klasifikasi [5]. Temuan-temuan ini relevan dengan konteks klasifikasi tingkat hidrasi pada penelitian ini, mengingat data asupan air harian juga bersifat individual, dipengaruhi oleh banyak atribut fisiologis dan lingkungan, serta rentan terhadap ketidakseimbangan proporsi kelas, sehingga pemilihan teknik penanganan ketidakseimbangan kelas dan strategi validasi menjadi faktor penentu keandalan model klasifikasi hidrasi yang dibangun. Pola umum yang konsisten ditemukan pada studi-studi tersebut adalah bahwa performa ketiga algoritma sangat dipengaruhi oleh karakteristik data, proporsi kelas, dan strategi praproses yang diterapkan.

Berdasarkan uraian tersebut terdapat kesamaan hasil penelitian yaitu performa Naive Bayes dan KNN pada data kesehatan sangat bergantung pada karakteristik fitur, proporsi kelas, dan strategi praproses yang diterapkan, sehingga temuan ketiganya saling menguatkan meskipun diterapkan pada domain penyakit yang berbeda (sindrom metabolik, gagal ginjal, dan penyakit jantung). Namun, celah yang konsisten muncul pada ketiga studi tersebut adalah belum diterapkannya penanganan ketidakseimbangan kelas secara eksplisit maupun optimasi *hyperparameter* secara sistematis, sehingga performa yang dilaporkan berpotensi belum mencerminkan kemampuan optimal masing-masing algoritma. Celah inilah yang menjadi salah satu dasar penelitian ini.

Tiga algoritma *machine learning* yang banyak digunakan dalam permasalahan klasifikasi bidang kesehatan adalah *Random Forest*, *Naive Bayes*, dan *K-Nearest Neighbor* (KNN) [6]. *Random Forest* merupakan algoritma *ensemble* yang membangun sejumlah besar pohon keputusan dan menggabungkan hasilnya melalui mekanisme *majority voting*, sehingga menghasilkan model yang stabil dan tahan terhadap *overfitting* [7]. *Naive Bayes* adalah algoritma berbasis probabilitas yang memanfaatkan Teorema Bayes dengan asumsi bahwa setiap fitur bersifat independen satu sama lain, ringan secara komputasi namun rentan mengalami penurunan performa ketika fitur-fiturnya saling berkorelasi [8]. Sementara itu, KNN mengklasifikasikan data baru berdasarkan mayoritas kelas dari sejumlah K tetangga terdekatnya dalam ruang fitur, sederhana dan intuitif namun sangat sensitif terhadap pemilihan nilai K , metrik jarak, dan skala data [9].

Dataset Daily Water Intake yang digunakan dalam penelitian ini memiliki distribusi kelas yang tidak seimbang, dengan proporsi 79,7% untuk kelas *Good* dan hanya 20,3% untuk kelas *Poor*. Ketidakseimbangan kelas seperti ini merupakan tantangan umum dalam klasifikasi data kesehatan karena model cenderung bias terhadap kelas mayoritas dan mengabaikan kelas minoritas yang justru sering kali memiliki signifikansi klinis lebih tinggi [10]. Untuk mengatasi kondisi tersebut, teknik SMOTE (*Synthetic Minority Oversampling Technique*) diterapkan dengan cara membangkitkan sampel sintesis pada kelas minoritas hingga distribusi kelas menjadi lebih proporsional sebelum proses pelatihan model dimulai [11].

Penelitian sebelumnya telah menerapkan ketiga algoritma tersebut, yaitu *Naive Bayes*, KNN, dan *Random Forest*, untuk mengklasifikasikan tingkat hidrasi menggunakan dataset publik *Daily Water Intake* [12]. Hasil penelitian mereka menunjukkan bahwa *Random Forest* unggul dengan akurasi 99,17%, diikuti KNN sebesar 97,07%, dan *Naive Bayes* sebesar 87,65%. Meskipun demikian, penelitian tersebut memiliki sejumlah keterbatasan metodologis yang perlu disempurnakan. Pertama, tidak ada penanganan terhadap ketidakseimbangan kelas pada dataset, yang kemungkinan besar memengaruhi performa *Naive Bayes* secara signifikan, terlihat dari selisih besar antara akurasi (87,65%) dan F1-Score (66,93%). Kedua, evaluasi hanya menggunakan metode *Holdout* dengan pembagian 80:20 yang rentan terhadap bias hasil. Ketiga, tidak ada upaya optimasi *hyperparameter* pada ketiga algoritma, padahal pemilihan parameter yang tepat berpotensi meningkatkan performa model secara signifikan. Penelitian tersebut secara eksplisit mengakui keterbatasan ini dan menyarankan penelitian selanjutnya untuk menerapkan teknik optimasi parameter.

Permasalahan serupa terkait ketidakseimbangan kelas dan minimnya optimasi parameter juga banyak ditemukan pada penelitian klasifikasi data kesehatan lain di Indonesia. Salah satu penelitian menunjukkan bahwa akurasi model klasifikasi data kesehatan sangat bergantung pada distribusi data latih yang digunakan [13]. Penelitian lain mengungkapkan adanya kesenjangan performa model akibat tidak adanya penanganan kelas minoritas secara eksplisit pada tahap praproses [14]. Beberapa studi yang menerapkan teknik penyeimbangan data juga menunjukkan kontribusi signifikan terhadap peningkatan *recall* pada kelas minoritas dibandingkan model tanpa penyeimbangan [15]. Temuan-temuan tersebut memperkuat urgensi penerapan SMOTE serta strategi validasi yang lebih *robust* pada penelitian klasifikasi tingkat hidrasi yang dilakukan saat ini, mengingat dataset *Daily Water Intake* juga memiliki karakteristik ketidakseimbangan kelas yang serupa.

Mengacu pada saran tersebut, penelitian ini hadir sebagai kelanjutan dari penelitian sebelumnya [12] dengan menerapkan optimasi *hyperparameter* menggunakan Grid Search sebagai kontribusi utama. Algoritma yang digunakan tetap sama, yaitu *Random Forest*, *Naive Bayes*, dan KNN, agar hasil dapat dibandingkan secara langsung. Dataset yang digunakan memiliki karakteristik distribusi kelas yang tidak seimbang, di mana kelas *Good* mendominasi sebesar 79,7% sementara kelas *Poor* hanya sebesar 20,3%. Kondisi ketidakseimbangan ini berpotensi menyebabkan model lebih condong memprediksi kelas mayoritas sehingga performa pada kelas minoritas menjadi tidak optimal. Oleh karena itu, teknik SMOTE diterapkan untuk menyeimbangkan distribusi kelas sebelum proses pelatihan model dilakukan. Selain itu, *Stratified K-Fold Cross Validation* ($K=10$) digunakan untuk memperoleh evaluasi yang lebih *robust* dan objektif. Penelitian ini bertujuan untuk menganalisis sejauh mana optimasi *Grid Search* mampu

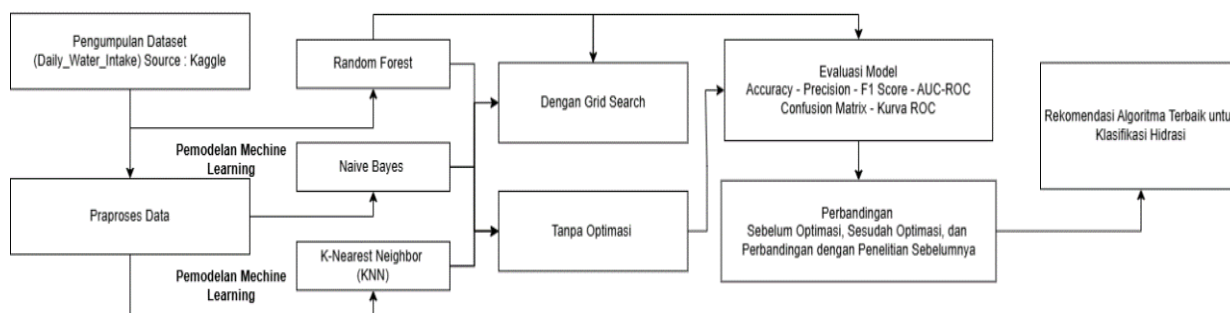
meningkatkan performa masing-masing algoritma dan mem-berikan rekomendasi algoritma terbaik untuk sistem klasifikasi tingkat hidrasi berbasis machine learning.

Secara keseluruhan, kontribusi penelitian ini terletak pada penggabungan tiga elemen metodologis sekaligus, yaitu penanganan ketidakseimbangan kelas melalui SMOTE, evaluasi yang robust melalui *Stratified K-Fold Cross Validation*, dan optimasi *hyperparameter* melalui *Grid Search*, yang belum diterapkan secara bersamaan pada penelitian klasifikasi tingkat hidrasi sebelumnya [12]. Penelitian lain yang menerapkan optimasi pada algoritma probabilistik turut menjadi rujukan tambahan dalam memahami keterbatasan ruang *hyperparameter* pada *Naive Bayes* [16]. Penelitian klasifikasi data kesehatan lain yang menggunakan *Naive Bayes* juga menunjukkan konsistensi performa algoritma tersebut ketika parameter optimalnya identik dengan nilai default [17]. Dengan demikian, penelitian ini diharapkan dapat memberikan kontribusi metodologis sekaligus praktis bagi pengembangan sistem klasifikasi tingkat hidrasi yang lebih akurat dan andal.

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Penelitian ini dilaksanakan secara sistematis melalui beberapa tahapan utama yang saling berkaitan, meliputi pengumpulan data, praproses, penanganan ketidakseimbangan kelas, pemodelan dengan tiga algoritma *machine learning*, optimasi *hyperparameter* menggunakan *Grid Search*, serta evaluasi performa model menggunakan metrik akurasi, presisi, *recall*, *F1-Score*, dan AUC-ROC. Seluruh proses komputasi dijalankan menggunakan bahasa pemrograman *Python* dengan memanfaatkan *library scikit-learn* dan *imbalanced-learn* pada platform *Google Colaboratory*. Gambar 1 menyajikan tahapan penelitian secara keseluruhan.



Gambar 1. Tahapan Penelitian

Gambar 1 menunjukkan bahwa proses penelitian berjalan secara linear dan berurutan, dimulai dari pengumpulan dataset, dilanjutkan dengan praproses data (pembersihan duplikat, encoding, normalisasi, dan penyeimbangan kelas melalui SMOTE), pemodelan menggunakan tiga algoritma klasifikasi, optimasi *hyperparameter* melalui *Grid Search*, hingga tahap akhir berupa evaluasi performa model menggunakan metrik akurasi, presisi, *recall*, *F1-Score*, dan AUC-ROC. Alur bertahap ini memastikan bahwa setiap komponen metodologi diterapkan secara konsisten pada ketiga algoritma sebelum hasilnya dibandingkan.

2.2 Dataset Penelitian

Dataset yang digunakan dalam penelitian ini adalah *Daily Water Intake*, sebuah dataset publik yang diperoleh dari platform Kaggle [22]. Dataset ini memuat 30.000 catatan data individu terkait konsumsi air harian, karakteristik fisik, dan kondisi lingkungan. Dataset ini mencakup tujuh atribut utama, yaitu usia (*Age*), jenis kelamin (*Gender*), berat badan (*Weight*), tingkat aktivitas fisik (*Physical Activity Level*), kondisi cuaca (*Weather*), jumlah asupan air harian (*Daily Water Intake*), dan tingkat hidrasi (*Hydration Level*) sebagai variabel target. Variabel target terdiri dari dua kategori, yaitu *Good* yang mewakili kondisi hidrasi yang baik dan *Poor* yang menandakan kondisi dehidrasi. Distribusi awal dataset menunjukkan ketidakseimbangan kelas yang cukup signifikan, dengan proporsi 79,7% untuk kelas *Good* dan 20,3% untuk kelas *Poor*.

Pemilihan dataset publik berskala besar dengan 30.000 catatan ini memberikan keuntungan tersendiri dibandingkan dataset kesehatan berskala kecil yang umum digunakan pada penelitian klasifikasi sejenis di Indonesia, karena jumlah sampel yang besar memungkinkan model mempelajari variasi pola data secara lebih menyeluruh sekaligus mengurangi risiko *overfitting* pada tahap pelatihan. Ketujuh atribut yang tersedia, baik yang bersifat fisiologis seperti usia dan berat badan, perilaku seperti tingkat aktivitas fisik, maupun lingkungan seperti kondisi cuaca, merepresentasikan faktor-faktor yang secara teoretis memengaruhi kebutuhan cairan harian seseorang. Karakteristik multidimensi pada atribut tersebut menjadi dasar pertimbangan dalam memilih ketiga algoritma yang diuji, mengingat masing-masing algoritma memiliki kemampuan berbeda dalam menangkap interaksi antar fitur, mulai dari pendekatan berbasis pohon pada Random Forest, probabilistik pada Naive Bayes, hingga berbasis kedekatan jarak pada KNN.

Tabel 1. Dataset *Daily_Water_Intake*

Age	Gender	Weight (kg)	Daily Water Intake (liters)	Physical Activity Level	Weather	Hydration Level
56	Male	96	04.23	Moderate	Hot	Good
60	Male	105	0,190972	High	Normal	Good
36	Male	68	02.39	Moderate	Cold	Good
19	Female	74	03.13	Moderate	Hot	Good
38	Male	77	02.11	Low	Normal	Poor
39	Female	93	0,139583	Low	Normal	Good
59	Male	104	0,186806	Moderate	Hot	Good
64	Female	106	0,179167	Moderate	Hot	Good
69	Female	108	0,146528	Low	Normal	Poor

Tabel 1 menyajikan contoh sembilan baris data yang merepresentasikan variasi karakteristik individu dan lingkungan yang tercakup dalam dataset, mulai dari kombinasi usia dan berat badan yang beragam hingga variasi kondisi cuaca dan tingkat aktivitas fisik yang memengaruhi nilai asupan air harian aktual serta label tingkat hidrasi akhir (*Good* atau *Poor*).

2.3 Praproses Data

Tahap praproses data dilakukan untuk memastikan kualitas dan konsistensi data sebelum memasuki proses pemodelan. Terdapat empat langkah yang dilaksanakan secara berurutan.

- Identifikasi dan penghapusan data duplikat, di mana ditemukan 338 baris data yang identik dan selanjutnya dihapus untuk menghindari bias dalam proses pelatihan model.
- Encoding atribut kategorikal. Atribut *Gender*, *Physical Activity Level*, *Weather*, dan *Hydration Level* diubah menjadi representasi numerik menggunakan Label *Encoder* agar dapat diproses oleh algoritma *machine learning*.
- Normalisasi data, di mana atribut numerik dinormalisasi menggunakan *Min-Max Scaler* untuk menyamakan skala antar fitur dan mencegah dominasi fitur tertentu dalam proses pembelajaran.
- Penanganan ketidakseimbangan kelas. Mengingat distribusi kelas awal yang tidak seimbang (79,7% *Good* vs 20,3% *Poor*), teknik SMOTE diterapkan untuk membangkitkan sampel sintetis pada kelas minoritas sehingga distribusi kelas menjadi seimbang sebelum model dilatih.

Keempat tahapan praproses tersebut dilakukan secara berurutan agar kualitas data yang masuk ke tahap pemodelan benar-benar terjaga. Penghapusan duplikasi mencegah model mempelajari pola yang berulang secara artifisial, encoding kategorikal memastikan seluruh atribut dapat diproses secara numerik oleh ketiga algoritma, normalisasi menjaga agar fitur dengan rentang nilai besar seperti berat badan tidak mendominasi perhitungan jarak pada KNN, sementara SMOTE memastikan model tidak bias terhadap kelas mayoritas. Penelitian klasifikasi kesehatan lain di Indonesia yang menerapkan SMOTE sebelum pelatihan turut menegaskan bahwa langkah ini memberikan kontribusi nyata terhadap peningkatan sensitivitas model dalam mengenali kelas minoritas pada data yang tidak seimbang [18]. Dengan demikian, rangkaian praproses dalam penelitian ini dirancang agar konsisten dengan praktik terbaik yang telah terbukti efektif pada studi klasifikasi data kesehatan, sehingga dapat diandalkan pula untuk kasus klasifikasi tingkat hidrasi.

2.4 Random Forest

Random Forest adalah algoritma ensemble yang membangun *T* pohon keputusan secara paralel menggunakan teknik bootstrap sampling dan pemilihan fitur secara acak [7]. Prediksi akhir ditentukan melalui mekanisme *majority voting* sebagaimana dirumuskan pada Persamaan (1):

$$y = \arg \max_k \sum_{t=1}^T I(ht(x) = k) \quad (1)$$

Keterangan: *T* = jumlah pohon (*T*=100, hasil *Grid Search*), *ht(x)* = prediksi pohon ke-*t*, *k* {*Good*, *Poor*} = label kelas, *I(·)* adalah fungsi indikator yang bernilai 1 apabila prediksi pohon ke-*t* sama dengan kelas *k*, dan bernilai 0 jika sebaliknya. Mekanisme ensemble ini menghasilkan model yang stabil dan tahan terhadap *overfitting* [7]. Mekanisme *bootstrap sampling* pada *Random Forest* memungkinkan setiap pohon dilatih menggunakan subset data yang berbeda secara acak dengan pengembalian, sehingga setiap pohon memiliki sudut pandang yang sedikit berbeda terhadap pola data secara keseluruhan. Kombinasi antara keragaman pohon akibat bootstrap sampling dan pemilihan fitur acak pada setiap pemisahan node menjadikan korelasi antar pohon relatif rendah, sehingga kesalahan prediksi pada satu pohon cenderung tidak terjadi secara sistematis pada pohon-pohon lainnya. Karakteristik inilah yang menjelaskan mengapa *Random Forest* secara konsisten unggul dalam berbagai studi klasifikasi data kesehatan di Indonesia, baik pada kasus dengan jumlah sampel besar maupun kecil, karena mekanisme ensemble mampu meredam noise dan variansi yang muncul akibat keterbatasan data pelatihan tunggal.

2.5 Naive Bayes

Naive Bayes adalah algoritma klasifikasi probabilistik yang didasarkan pada Teorema Bayes dengan asumsi independensi antar fitur [5]. Probabilitas posterior setiap kelas dihitung menggunakan Persamaan (2):

$$P(C|X) = \frac{P(C)\pi P(x_i|C)}{P(X)} \quad (2)$$

Keterangan: $C \{Good, Poor\}$ = kelas target, $X = (x_1, \dots, x_6)$ = vektor 6 fitur, $P(C)$ = probabilitas prior kelas, $P(x_i|C)$ = *likelihood* fitur ke- i , $P(X)$ = probabilitas evidence. Parameter *var_smoothing* = $1e-9$ diterapkan untuk mencegah probabilitas nol pada fitur tertentu [5].

2.6 K-Nearest Neighbor (KNN)

KNN mengklasifikasikan data baru berdasarkan mayoritas kelas dari $K=3$ tetangga terdekatnya dalam ruang fitur [6]. Kedekatan antar data dihitung menggunakan jarak *Euclidean* pada Persamaan (3), dengan pembobotan berbasis jarak (*weights=distance*) pada Persamaan (4):

$$d(x, x_i) = \sqrt{\sum_{j=1}^n (x_j - x_{ij})^2} \quad (3)$$

Keterangan: $n = 6$ (jumlah fitur), $K = 3$ (hasil *Grid Search*), w_i = bobot tetangga ke- i berdasarkan jarak. Skema pembobotan *distance* memungkinkan tetangga yang lebih dekat memberikan pengaruh lebih besar dalam penentuan kelas akhir, sehingga meningkatkan presisi klasifikasi dibandingkan skema seragam [6]. Penerapan normalisasi *Min-Max Scaler* pada tahap pra-proses menjadi faktor penting yang mendukung kinerja KNN, mengingat algoritma ini sangat bergantung pada perhitungan jarak *Euclidean* antar titik data. Tanpa normalisasi, atribut dengan rentang nilai besar seperti berat badan berpotensi mendominasi perhitungan jarak dan mengabaikan kontribusi atribut lain yang secara klinis tidak kalah penting, seperti tingkat aktivitas fisik. Dengan menyamakan skala seluruh fitur ke rentang yang seragam, kontribusi setiap atribut terhadap penentuan tetangga terdekat menjadi lebih proporsional, sehingga keputusan klasifikasi yang dihasilkan KNN benar-benar mencerminkan kemiripan karakteristik individu secara menyeluruh, bukan hanya didominasi oleh atribut dengan satuan numerik yang lebih besar.

2.7 Optimasi Hyperparameter dengan Grid Search

Grid Search merupakan metode pencarian hyperparameter yang bekerja dengan cara mengevaluasi secara menyeluruh setiap kombinasi nilai parameter yang telah ditentukan dalam sebuah grid pencarian [19]. Pemilihan rentang nilai parameter didasarkan pada nilai default sebagai titik awal, rekomendasi literatur, dan karakteristik dataset. Untuk parameter berskala kecil seperti *var_smoothing*, pengujian dilakukan dalam skala logaritmik. Untuk *n_neighbors*, dipilih nilai ganjil guna menghindari hasil seri dalam proses voting klasifikasi biner. Pada setiap kombinasi, model dievaluasi menggunakan *Stratified K-Fold Cross Validation*, dan kombinasi yang menghasilkan skor *F1-weighted* tertinggi dipilih sebagai konfigurasi optimal. Tabel 1 menyajikan parameter beserta rentang nilai yang diuji.

Tabel 2. OPTIMASI HYPERPARAMETER

Algoritma	Hyperparameter	Rentang Nilai yang Diuji
Random Forest	<i>n_estimators</i>	50, 100, 200
	<i>max_depth</i>	None, 10, 20
	<i>min_samples_split</i>	2, 5, 10
Naive Bayes	<i>var_smoothing</i>	$1e-9$, $1e-8$, $1e-7$, $1e-6$, $1e-5$
KNN	<i>n_neighbors</i>	3, 5, 7, 9, 11
	<i>metric</i>	euclidean, manhattan
	<i>weights</i>	uniform, distance

Rentang nilai pada Tabel 2 dirancang cukup luas untuk menangkap kemungkinan konfigurasi optimal tiap algoritma: *Random Forest* diuji pada 27 kombinasi ($3 \times 3 \times 3$) dari jumlah pohon, kedalaman, dan minimum sampel split; Naive Bayes diuji pada lima titik *var_smoothing* berskala logaritmik mengingat algoritma ini hanya memiliki satu *hyperparameter* utama; sementara KNN diuji pada 20 kombinasi ($5 \times 2 \times 2$) dari jumlah tetangga, metrik jarak, dan skema pembobotan. Cakupan rentang yang proporsional terhadap kompleksitas masing-masing algoritma ini memastikan proses *Grid Search* dapat mengeksplorasi ruang parameter secara memadai sebelum konfigurasi terbaik ditentukan melalui *Stratified K-Fold Cross Validation*.

2.8 Evaluasi Model

Evaluasi performa model dilakukan menggunakan *Stratified K-Fold Cross Validation* dengan $K=10$. Pendekatan ini memastikan bahwa proporsi kelas pada setiap fold tetap merepresentasikan distribusi kelas keseluruhan, sehingga hasil evaluasi tidak bias akibat pembagian data yang tidak proporsional. Berbeda dengan metode *Holdout* yang hanya melakukan satu kali pembagian data, *Stratified K-Fold* mengevaluasi model sebanyak 10 kali dengan subset data yang berbeda, kemudian merata-ratakan hasilnya untuk memperoleh estimasi performa yang lebih andal.

Metrik evaluasi yang digunakan mencakup lima ukuran kinerja: (1) Akurasi mengukur proporsi prediksi yang benar dari keseluruhan data uji; (2) Presisi menggambarkan tingkat ketepatan prediksi kelas positif; (3) *Recall* mengukur kemampuan model dalam mendeteksi seluruh sampel kelas positif; (4) *F1-Score* menyeimbangkan presisi dan *recall* dalam satu nilai harmonik; dan (5) AUC-ROC mengukur kemampuan model membedakan kedua kelas di berbagai ambang keputusan, di mana metrik ini sangat relevan untuk data tidak seimbang karena tidak terpengaruh

oleh distribusi proporsi kelas. Performa model dievaluasi dalam dua kondisi: sebelum optimasi sebagai baseline dan sesudah optimasi *Grid Search*.

Penggunaan *Stratified K-Fold Cross Validation* dengan $K=10$ pada penelitian ini sejalan dengan praktik validasi pada berbagai studi klasifikasi kesehatan yang menekankan pentingnya pembagian data yang merepresentasikan proporsi kelas secara konsisten pada setiap iterasi pengujian. Pendekatan ini berbeda dari metode Holdout sederhana yang umum digunakan pada penelitian klasifikasi kesehatan terdahulu, yang cenderung menghasilkan estimasi performa kurang stabil ketika ukuran data uji relatif kecil atau distribusi kelasnya tidak proporsional. Dengan merata-ratakan hasil dari sepuluh kali pengujian pada subset data yang berbeda, kerangka evaluasi pada penelitian ini diharapkan mampu memberikan gambaran performa model yang lebih representatif dan dapat dipertanggungjawabkan secara metodologis sebelum dibandingkan dengan hasil penelitian sebelumnya.

3. HASIL DAN PEMBAHASAN

Dataset *Daily Water Intake* yang digunakan pada penelitian ini terdiri dari 30.000 data awal yang menyusut menjadi 29.662 baris setelah penghapusan 338 data duplikat, dengan distribusi kelas asli 79,7% *Good* dan 20,3% *Poor* (Tabel 1). Setelah penerapan SMOTE pada tahap praproses, jumlah data efektif yang digunakan dalam proses pelatihan dan evaluasi model bertambah menjadi 47.460 sampel dengan distribusi kelas yang seimbang. Ringkasan karakteristik dataset ini menjadi acuan penting dalam menginterpretasikan seluruh hasil evaluasi model. Bagian ini memaparkan hasil pengujian dan analisis perbandingan kinerja tiga algoritma klasifikasi machine learning dalam memprediksi tingkat hidrasi manusia. Pengujian dilakukan dalam dua kondisi, yaitu tanpa optimasi sebagai *baseline* dan setelah penerapan optimasi hyperparameter menggunakan *Grid Search*, guna mengukur sejauh mana optimasi mampu meningkatkan performa masing-masing algoritma. Hasil yang diperoleh selanjutnya dianalisis secara komparatif antar algoritma dan dibandingkan langsung dengan penelitian sebelumnya [12] untuk mengukur kontribusi penelitian ini secara kuantitatif.

Pembahasan pada bagian ini disusun secara sistematis, dimulai dari pemaparan parameter optimal hasil *Grid Search*, dilanjutkan dengan perbandingan performa sebelum dan sesudah optimasi, pembahasan mendalam untuk masing-masing algoritma, visualisasi *confusion matrix* dan kurva ROC, hingga perbandingan menyeluruh dengan hasil penelitian sebelumnya [12]. Alur pembahasan ini dirancang agar pembaca dapat memahami secara bertahap bagaimana setiap komponen metodologi, mulai dari penanganan ketidakseimbangan kelas hingga optimasi *hyperparameter*, saling berkontribusi terhadap peningkatan kualitas model klasifikasi tingkat hidrasi yang dihasilkan.

3.1 Parameter Optimal Hasil Grid Search

Proses *Grid Search* berhasil mengidentifikasi kombinasi *hyperparameter* terbaik untuk setiap algoritma dengan mengevaluasi seluruh kombinasi parameter secara sistematis menggunakan *Stratified K-Fold Cross Validation*. Tabel 3 menyajikan parameter optimal yang ditemukan beserta interpretasinya. *Random Forest* menghasilkan konfigurasi terbaik dengan $n_estimators=100$, $max_depth=20$, dan $min_samples_split=2$. Kedalaman pohon sebesar 20 level sudah cukup untuk menangkap pola data secara menyeluruh tanpa risiko *overfitting*, sementara 100 pohon terbukti menghasilkan prediksi yang stabil melalui mekanisme *majority voting*. *Naive Bayes* memperoleh nilai optimal pada $var_smoothing=1e-9$ yang identik dengan nilai *default scikit-learn*. Kondisi ini mencerminkan keterbatasan ruang pencarian parameter *Naive Bayes* yang hanya memiliki satu *hyperparameter*, sehingga *Grid Search* tidak menemukan konfigurasi yang lebih unggul dari pengaturan bawaannya. KNN menunjukkan konfigurasi terbaik pada $n_neighbors=3$, $metric=euclidean$, dan $weights=distance$. Penggunaan pembobotan berbasis jarak memungkinkan tetangga yang lebih dekat mendapat pengaruh lebih besar dalam proses klasifikasi, yang terbukti meningkatkan presisi prediksi secara signifikan dibandingkan skema pembobotan seragam.

Tabel 3. Parameter Optimal Hasil Grid Search

Algoritma	Hyperparameter	Nilai Optimal	Keterangan
Random Forest	$n_estimators$	100	Kedalaman sedang sudah cukup optimal
	max_depth	20	
	$min_samples_split$	2	
Naive Bayes	$var_smoothing$	1e-9	Identik dengan nilai default
KNN	$n_neighbors$	3	Pembobotan jarak meningkatkan diskriminasi kelas
	$metric$	euclidean	
	$weights$	distance	

Secara umum, Tabel 3 menunjukkan bahwa hanya KNN dan Random Forest yang memperoleh manfaat nyata dari proses *Grid Search*, ditandai dengan parameter optimal yang berbeda dari nilai default ($max_depth=20$ dan $n_estimators=100$ untuk Random Forest; $n_neighbors=3$ dan $weights=distance$ untuk KNN), sementara Naive Bayes tidak menunjukkan perubahan karena parameter optimalnya identik dengan default. Pola ini mengindikasikan bahwa algoritma dengan ruang hyperparameter yang lebih kaya (Random Forest dan KNN) memiliki peluang lebih besar untuk ditingkatkan melalui penyetelan parameter dibandingkan algoritma probabilistik sederhana seperti Naive Bayes.

3.2 Perbandingan Performa Sebelum dan Sesudah Optimasi

Perbandingan performa ketiga algoritma pada lima metrik evaluasi dalam dua kondisi, yaitu tanpa optimasi sebagai *baseline* dan setelah penerapan *Grid Search*, yang keduanya dievaluasi menggunakan *Stratified K-Fold Cross Validation (K=10)*. Berikut Tabel 4 menyajikan perbandingan performa sebelum dan sesudah optimasi. Seluruh hasil yang ditampilkan, *Random Forest* secara konsisten berada di posisi teratas, disusul KNN, dan *Naive Bayes* di urutan terakhir. Perubahan yang terjadi pada *Random Forest* dan *Naive Bayes* setelah optimasi praktis tidak tampak secara visual, sejalan dengan hasil numerik yang menunjukkan minimnya pengaruh optimasi terhadap kedua algoritma tersebut. KNN menjadi satu-satunya algoritma yang menunjukkan perbedaan hasil yang cukup jelas. Pada grafik AUC-ROC, *Random Forest* meraih nilai 0,99 yang hampir sempurna, sementara KNN mengalami penurunan tipis sebesar -0,0021 sebagai *trade-off* wajar dari optimasi yang memprioritaskan *F1-Score*. *Naive Bayes* mempertahankan nilai AUC-ROC 0,92 yang tidak berubah.

Tabel 4. Perbandingan Performa Sebelum Dan Sesudah Optimasi

Algoritma	Metrik	Tanpa Optimasi	Dengan Optimasi	Selisih	Keterangan
Random Forest	Accuracy	0.9949	0.9950	+0.0001	Peningkatan minimal
	Precision	0.9950	0.9950	+0.0	Tidak berubah
	Recall	0.9949	0.9950	+0.0001	Peningkatan minimal
	F1-Score	0.9949	0.9950	+0.0001	Peningkatan minimal
	AUC-ROC	0.9999	0.9999	0.0	Tidak berubah
Naive Bayes	Accuracy	0.8353	0.8353	0.0	Tidak berubah
	Precision	0.8400	0.8400	0.0	Tidak berubah
	Recall	0.8353	0.8353	0.0	Tidak berubah
	F1-Score	0.8347	0.8347	0.0	Tidak berubah
	AUC-ROC	0.9268	0.9268	0.0	Tidak berubah
KNN	Accuracy	0.9762	0.9824	+0.0062	Peningkatan signifikan
	Precision	0.9770	0.9828	+0.0058	Peningkatan signifikan
	Recall	0.9762	0.9824	+0.0062	Peningkatan signifikan
	F1-Score	0.9762	0.9824	+0.0062	Peningkatan signifikan
	AUC-ROC	0.9946	0.9925	-0.0021	Sedikit menurun

Secara ringkas, Tabel 4 menegaskan bahwa optimasi *Grid Search* memberikan dampak yang sangat tidak merata antar algoritma: KNN memperoleh peningkatan pada seluruh metrik (accuracy, precision, recall, F1-Score naik +0,58% hingga +0,62%), *Random Forest* hanya memperoleh peningkatan marginal pada tiga metrik (+0,0001), sementara *Naive Bayes* tidak mengalami perubahan sama sekali pada kelima metrik (selisih 0,0). Pola ini konsisten dengan ukuran ruang hyperparameter masing-masing algoritma, di mana algoritma dengan lebih banyak parameter yang dapat disetel (KNN dan *Random Forest*) memiliki peluang lebih besar untuk menemukan konfigurasi yang lebih baik dibandingkan *Naive Bayes* yang hanya memiliki satu hyperparameter.

3.3 Random Forest

Sebagai algoritma ensemble yang menggabungkan banyak pohon keputusan, *Random Forest* terbukti menjadi model terbaik dalam penelitian ini dengan akurasi mencapai 99,50% dan AUC-ROC 0,9999 setelah optimasi. Peningkatan dari *Grid Search* sangat kecil, hanya +0,0001 untuk akurasi, recall, dan F1-Score, sementara presisi dan AUC-ROC tidak berubah sama sekali. Tingginya nilai AUC-ROC yang mendekati sempurna dapat dijelaskan oleh tiga faktor utama. Pertama, mekanisme ensemble yang menggabungkan 100 pohon keputusan diduga mampu dalam memisahkan kelas *Good* dan *Poor*. Kedua, setiap pohon dibangun dari subset fitur yang dipilih secara acak kemungkinan disebabkan oleh seluruh atribut secara bersamaan. Ketiga, penerapan SMOTE sebelum pelatihan memastikan model mendapat paparan seimbang terhadap kedua kelas sehingga kurva ROC terbentuk secara proporsional. Keunggulan ini sejalan dengan penelitian sebelumnya [12] yang juga menempatkan *Random Forest* sebagai algoritma terbaik, namun penelitian ini berhasil menghasilkan akurasi yang lebih tinggi (99,50% vs 99,17%) berkat penerapan SMOTE dan evaluasi *Stratified K-Fold* yang lebih ketat.

Temuan ini sejalan dengan berbagai penelitian klasifikasi data kesehatan berbasis *Random Forest* di Indonesia yang juga menempatkan algoritma ensemble ini sebagai model dengan performa tertinggi dibandingkan algoritma pembanding lainnya [20]. Karakteristik *Random Forest* yang mampu menggabungkan banyak pohon keputusan independen membuatnya relatif tahan terhadap variasi data uji, sehingga hasil prediksi cenderung stabil meskipun konfigurasi parameter diubah dalam rentang yang wajar. Konsistensi performa *Random Forest* pada penelitian ini, baik sebelum maupun sesudah optimasi, mengindikasikan bahwa parameter default *scikit-learn* untuk algoritma tersebut sudah cukup representatif terhadap kompleksitas dataset Daily Water Intake, sehingga *Grid Search* hanya memberikan penyempurnaan marginal dibandingkan kontribusinya pada algoritma KNN.

Secara struktural, keunggulan *Random Forest* dibandingkan *Naive Bayes* pada dataset ini dapat dijelaskan oleh perbedaan mendasar cara kedua algoritma menangani hubungan antar fitur: mekanisme ensemble *Random Forest* mampu menangkap interaksi non-linear antara usia, berat badan, aktivitas fisik, dan cuaca secara otomatis melalui

pemisahan node berjenjang, sementara Naive Bayes mengasumsikan independensi antar fitur yang pada kenyataannya saling berkorelasi pada data fisiologis, sehingga batas keputusan yang dihasilkan Naive Bayes menjadi kurang presisi dibandingkan Random Forest.

3.4 Naive Bayes

Naive Bayes merupakan satu-satunya algoritma yang tidak menunjukkan perubahan performa setelah penerapan *Grid Search*, dengan akurasi tetap di angka 83,53% dan F1-Score 83,47%. Kondisi ini terjadi karena parameter terbaik yang ditemukan, yaitu $var_smoothing=1e-9$, identik dengan nilai default *scikit-learn*, sehingga tidak ada perbedaan model yang dihasilkan. Performa yang lebih rendah dibandingkan dua algoritma lainnya mencerminkan keterbatasan asumsi independensi antar fitur yang menjadi landasan Naive Bayes. Pada dataset Daily Water Intake, terdapat keterkaitan inheren antara atribut seperti tingkat aktivitas fisik, berat badan, dan jumlah asupan air harian, sehingga pendekatan probabilistik sederhana *Naive Bayes* kurang mampu menangkap kompleksitas pola data secara menyeluruh. Meskipun demikian, F1-Score *Naive Bayes* dalam penelitian ini (83,47%) jauh lebih tinggi dari hasil penelitian sebelumnya sebesar 66,93% [12], membuktikan bahwa penerapan SMOTE memberikan dampak positif yang signifikan dalam mengenali kelas minoritas *Poor*.

Proses *Grid Search* untuk Naive Bayes menguji lima nilai $var_smoothing$ pada skala logaritmik ($1e-9$ hingga $1e-5$), namun seluruh kombinasi tersebut menghasilkan skor *F1-weighted* yang identik dengan pengaturan *default*, sehingga tidak ditemukan konfigurasi yang lebih unggul. Implikasinya terhadap tujuan penelitian ini adalah bahwa strategi peningkatan performa pada *Naive Bayes* tidak dapat mengandalkan optimasi *hyperparameter* semata, melainkan lebih bergantung pada perbaikan kualitas data melalui SMOTE dan strategi praproses lain, mengingat ruang pencarian parameter algoritma ini secara inheren terbatas dibandingkan *Random Forest* maupun KNN.

Keterbatasan asumsi independensi fitur pada *Naive Bayes* juga teramati pada beberapa studi klasifikasi data kesehatan di Indonesia. Salah satu penelitian yang membandingkan *Naive Bayes* dengan KNN menunjukkan bahwa *Naive Bayes* cenderung memberikan akurasi lebih rendah ketika fitur-fitur dalam dataset saling berkorelasi [3]. Penelitian lain yang menerapkan teknik optimasi pada *Naive Bayes* turut menegaskan keterbatasan ruang *hyperparameter* algoritma ini dibandingkan algoritma berbasis pohon keputusan maupun ensemble [16]. Pada konteks penelitian ini, atribut seperti aktivitas fisik, berat badan, dan asupan air harian secara alami saling memengaruhi status hidrasi seseorang, sehingga model probabilistik sederhana kesulitan memisahkan kontribusi masing-masing fitur secara independen. Meskipun demikian, peningkatan signifikan pada F1-Score dibandingkan penelitian sebelumnya [12] tetap menunjukkan bahwa kombinasi SMOTE dengan evaluasi yang lebih ketat mampu memperbaiki keandalan *Naive Bayes* meskipun tanpa optimasi *hyperparameter* yang berarti.

3.5 K-Nearest Neighbor (KNN)

KNN mencatatkan peningkatan paling signifikan pascaoptimasi dengan kenaikan akurasi sebesar +0,62% dari 97,62% menjadi 98,24%, serta peningkatan serupa pada presisi (+0,58%), *recall* (+0,62%), dan *F1-Score* (+0,62%). Hasil ini mempertegas bahwa KNN sangat responsif terhadap pemilihan *hyperparameter* yang tepat. Konfigurasi terbaik adalah $n_neighbors=3$, $metric=euclidean$, dan $weights=distance$, di mana pembobotan berbasis jarak memungkinkan tetangga yang lebih dekat memiliki pengaruh lebih dominan dalam proses *voting*. Satu-satunya penurunan terjadi pada AUC-ROC sebesar -0,0021, yang merupakan *trade-off* wajar karena optimasi dilakukan berdasarkan skor *F1-weighted* sebagai target utama. Nilai AUC-ROC KNN setelah optimasi (0,9925) tetap berada dalam kategori sangat baik.

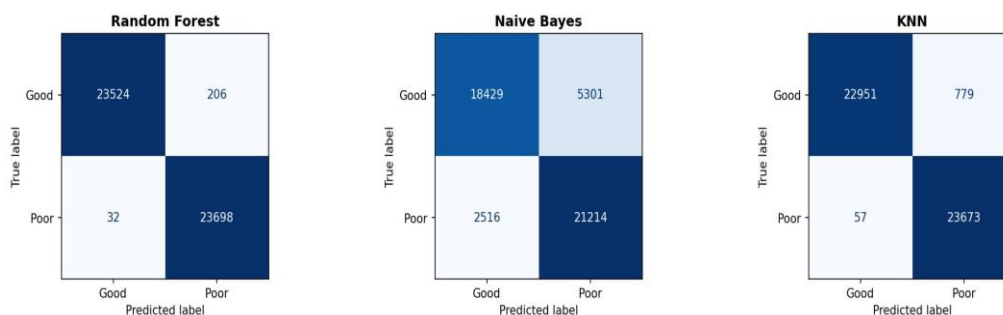
Penurunan AUC-ROC pada KNN dapat dijelaskan dari karakteristik probabilitas prediksi yang dihasilkan skema pembobotan *distance*. Ketika tetangga terdekat diberi bobot dominan, probabilitas keluaran model cenderung terpolarisasi mendekati 0 atau 1 untuk sebagian besar sampel, alih-alih tersebar secara bertahap di berbagai ambang keputusan. Kondisi ini meningkatkan ketepatan klasifikasi pada ambang baku (0,5) sebagaimana tercermin dari kenaikan akurasi dan *F1-Score*, namun sedikit mengurangi kemampuan model membedakan kedua kelas secara konsisten di seluruh rentang ambang keputusan yang menjadi dasar perhitungan AUC-ROC, sehingga penurunan tipis sebesar -0,0021 merupakan konsekuensi wajar dari perubahan karakteristik probabilitas tersebut, bukan indikasi penurunan kualitas model secara keseluruhan. Temuan ini konsisten dengan hasil penelitian [21], yang menunjukkan bahwa peningkatan performa klasifikasi pada varian KNN berbobot tidak selalu diiringi oleh kualitas kalibrasi probabilitas yang lebih baik dibandingkan KNN standar. Dalam penelitian tersebut, skema pembobotan pada KNN cenderung menghasilkan estimasi probabilitas kelas yang kurang terkalibrasi ditandai dengan nilai *Brier score* dan *Expected Calibration Error* (ECE) yang lebih tinggi meskipun akurasi klasifikasinya meningkat. Hal ini memperkuat argumen bahwa penurunan AUC-ROC pada penelitian ini bukan disebabkan oleh menurunnya kualitas diskriminasi model, melainkan oleh perubahan karakteristik sebaran probabilitas keluaran akibat skema pembobotan *distance*, yang cenderung menghasilkan skor probabilitas yang lebih terpolarisasi mendekati 0 atau 1.

Sensitivitas KNN terhadap pemilihan *hyperparameter* sejalan dengan temuan pada berbagai penelitian klasifikasi data kesehatan di Indonesia. Salah satu penelitian yang membandingkan KNN dengan algoritma lain menunjukkan bahwa pemilihan nilai K dan metrik jarak berpengaruh besar terhadap akurasi akhir model [5]. Penelitian lain yang menerapkan KNN pada data kesehatan turut menegaskan pentingnya penyesuaian parameter terhadap karakteristik data yang digunakan [13]. Nilai $n_neighbors$ yang relatif kecil ($K=3$) pada penelitian ini sesuai dengan ukuran dataset yang besar setelah resampling SMOTE, di mana tetangga terdekat yang lebih sedikit namun

relevan secara lokal mampu memberikan keputusan klasifikasi yang lebih presisi dibandingkan jumlah tetangga yang lebih besar namun berpotensi mencakup data dari kelas yang berbeda.

3.6 Confusion Matrix

Confusion matrix dari ketiga model setelah penerapan optimasi *Grid Search*, yang memberikan gambaran detail mengenai distribusi prediksi benar dan salah untuk masing-masing kelas. Berikut adalah gambar hasil ketiga model yang disajikan dalam bentuk *Confusion Matrix* setelah optimasi pada Gambar 2.



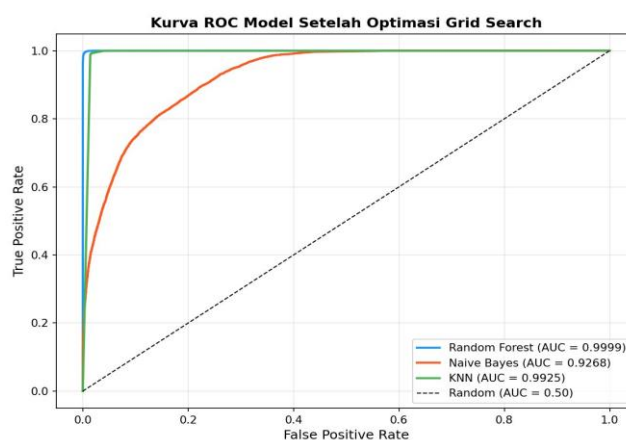
Gambar 2. Confusion Matrix Ketiga Model Setelah Optimasi Grid Search

Berdasarkan Gambar 2, Random Forest mencatat kesalahan klasifikasi paling sedikit dari total 47.460 data hasil resampling SMOTE, dengan hanya 206 *false positive* dan 32 *false negative*. KNN menunjukkan performa yang juga baik dengan 779 *false positive* dan hanya 57 *false negative*, di mana angka *false negative* yang sangat rendah mengindikasikan KNN sangat handal mendeteksi kasus *Poor*. *Naive Bayes* mencatat kesalahan terbesar dengan 5.301 sampel *Good* yang salah diklasifikasikan sebagai *Poor*, mencerminkan kesulitan model dalam menetapkan batas keputusan yang tepat antara kedua kelas.

3.7 Kurva ROC dan AUC

Kurva ROC (*Receiver Operating Characteristic*) menggambarkan hubungan antara *True Positive Rate* (TPR) dan *False Positive Rate* (FPR) di berbagai nilai ambang keputusan, sementara AUC merupakan representasi numerik luas area di bawah kurva tersebut. Semakin tinggi nilai AUC dan semakin mendekati pojok kiri atas grafik, semakin baik kemampuan model membedakan kedua kelas. Perlu dicatat bahwa penelitian sebelumnya [12] tidak melaporkan AUC-ROC, sehingga metrik ini menjadi kontribusi evaluasi tambahan yang membedakan penelitian ini secara metodologis.

Gambar 3 memperlihatkan bahwa kurva Random Forest dan KNN naik secara hampir vertikal mendekati pojok kiri atas grafik, mengindikasikan *True Positive Rate* yang sangat tinggi dengan *False Positive Rate* yang sangat rendah. *Random Forest* mencapai AUC tertinggi sebesar 0,99, KNN menyusul dengan 0,9925, sementara *Naive Bayes* memperoleh 0,92 dengan lekukan kurva yang lebih landai. Berdasarkan standar interpretasi AUC, nilai di atas 0,90 dikategorikan sebagai *excellent classifier*. Dengan demikian, ketiga algoritma dalam penelitian ini masuk dalam kategori tersebut meskipun dengan tingkat keunggulan yang berbeda.



Gambar 3. Kurva ROC Ketiga Model Setelah Optimasi Grid Search

3.8 Perbandingan dengan Penelitian Sebelumnya

Perbandingan menyeluruh antara hasil penelitian ini dan penelitian sebelumnya [12] guna menilai sejauh mana kontribusi penelitian ini, hasil yang diperoleh pada seluruh metrik evaluasi dibandingkan secara langsung. Pada algoritma Random Forest, penelitian sebelumnya [12] melaporkan *accuracy* 0,9917, *precision* 0,9891, *recall* 0,9695,

dan *F1-Score* 0,9792; penelitian ini memperoleh *accuracy* 0,9949, *precision* 0,9950, *recall* 0,9949, dan *F1-Score* 0,9949 tanpa optimasi, yang meningkat menjadi *accuracy* 0,9950, *precision* 0,9950, *recall* 0,9950, dan *F1-Score* 0,9950 setelah optimasi (AUC-ROC 0,9999, tidak dilaporkan pada penelitian sebelumnya), sehingga selisih akhir terhadap penelitian sebelumnya adalah +0,0033 (*accuracy*), +0,0059 (*precision*), +0,0255 (*recall*), dan +0,0158 (*F1-Score*). Pada algoritma *Naive Bayes*, penelitian sebelumnya melaporkan *accuracy* 0,8765, *precision* 0,7289, *recall* 0,6188, dan *F1-Score* 0,6693; penelitian ini menghasilkan nilai yang identik baik tanpa maupun dengan optimasi, yaitu *accuracy* 0,8353, *precision* 0,8400, *recall* 0,8353, dan *F1-Score* 0,8347 (AUC-ROC 0,9268), sehingga selisihnya adalah -0,0412 (*accuracy*, satu-satunya penurunan pada seluruh perbandingan), +0,1111 (*precision*), +0,2165 (*recall*), dan +0,1654 (*F1-Score*). Pada algoritma KNN, penelitian sebelumnya melaporkan *accuracy* 0,9707, *precision* 0,9375, *recall* 0,9158, dan *F1-Score* 0,9265; penelitian ini memperoleh *accuracy* 0,9762, *precision* 0,9770, *recall* 0,9762, dan *F1-Score* 0,9762 tanpa optimasi, meningkat menjadi *accuracy* 0,9824, *precision* 0,9828, *recall* 0,9824, dan *F1-Score* 0,9824 setelah optimasi (AUC-ROC turun tipis dari 0,9946 menjadi 0,9925), sehingga selisih akhir terhadap penelitian sebelumnya adalah +0,0117 (*accuracy*), +0,0453 (*precision*), +0,0666 (*recall*), dan +0,0559 (*F1-Score*). Tanda (+) pada seluruh nilai selisih menunjukkan peningkatan, sedangkan tanda (-) menunjukkan penurunan dibandingkan penelitian sebelumnya [12]; AUC-ROC tidak dilaporkan pada penelitian sebelumnya sehingga ditandai N/A pada seluruh perbandingan.

Beberapa temuan penting dapat diidentifikasi dari perbandingan tersebut [12]. *Random Forest* dan KNN menghasilkan performa yang lebih unggul pada seluruh metrik dibandingkan hasil pada [12]. Peningkatan paling mencolok terjadi pada *recall* dan *F1-Score*, mengindikasikan bahwa penerapan SMOTE secara signifikan memperbaiki kemampuan model mengenali kelas minoritas Poor yang sebelumnya terlewatkan. Satu-satunya nilai yang lebih rendah dari penelitian sebelumnya [12] adalah akurasi *Naive Bayes* (83,53% vs 87,65%). Namun kondisi ini justru mencerminkan evaluasi yang lebih ketat melalui *Stratified K-Fold* dibandingkan *Holdout* yang berpotensi menghasilkan estimasi terlalu optimistik. Terlebih lagi, *F1-Score Naive Bayes* dalam penelitian ini (83,47%) jauh lebih tinggi dari hasil pada penelitian sebelumnya [12] sebesar 66,93%, membuktikan kontribusi positif SMOTE dalam mengatasi ketidakseimbangan kelas.

Secara umum, perbandingan menyeluruh di atas menegaskan bahwa kerangka metodologis yang diterapkan pada penelitian ini, yakni kombinasi SMOTE, *Stratified K-Fold Cross Validation*, dan *Grid Search*, memberikan kontribusi nyata terhadap perbaikan kualitas evaluasi dibandingkan pendekatan *Holdout* pada penelitian sebelumnya [12]. Pola ini juga konsisten dengan temuan pada studi klasifikasi kesehatan lain di Indonesia, di mana penerapan teknik penyeimbangan kelas secara umum berkontribusi terhadap peningkatan *recall* dan *F1-Score* pada kelas minoritas tanpa mengorbankan akurasi keseluruhan model [15]. Dengan demikian, hasil penelitian ini tidak hanya memperkuat posisi *Random Forest* sebagai algoritma unggulan, tetapi juga memberikan bukti empiris mengenai pentingnya kerangka evaluasi yang ketat dalam menghasilkan kesimpulan klasifikasi kesehatan yang dapat dipercaya.

4. KESIMPULAN

Penelitian ini telah menganalisis pengaruh optimasi *hyperparameter* menggunakan *Grid Search* terhadap kinerja *Random Forest*, *Naive Bayes*, dan *K-Nearest Neighbor* dalam mengklasifikasikan tingkat hidrasi manusia berdasarkan dataset *Daily Water Intake*, dengan tiga temuan utama: KNN paling diuntungkan oleh optimasi (akurasi naik 0,62% menjadi 98,24% pada konfigurasi *n_neighbors=3*, *metric=euclidean*, *weights=distance*), *Random Forest* tetap menjadi algoritma terbaik secara keseluruhan (akurasi 99,50%, AUC-ROC 0,9999) meski peningkatan dari optimasi sangat minimal karena parameter *default* sudah mendekati optimal, dan *Naive Bayes* tidak mengalami perubahan performa karena parameter optimalnya identik dengan nilai *default*, mencerminkan keterbatasan ruang *hyperparameter* algoritma tersebut. Dibandingkan penelitian terdahulu [12], penelitian ini unggul pada mayoritas metrik untuk *Random Forest* dan KNN serta *F1-Score Naive Bayes* yang jauh lebih baik (83,47% vs 66,93%), membuktikan kontribusi positif SMOTE dan *Stratified K-Fold* terhadap kualitas dan objektivitas evaluasi. Penelitian ini memiliki keterbatasan berupa penggunaan satu sumber dataset publik, optimasi yang hanya menggunakan *Grid Search* tanpa pembandingan metode lain seperti *Random Search* atau *Bayesian Optimization*, serta cakupan hanya tiga algoritma klasik tanpa menyertakan ensemble modern seperti XGBoost atau LightGBM, sehingga penelitian selanjutnya disarankan untuk mengeksplorasi hal-hal tersebut sekaligus mengkaji teknik interpretability seperti SHAP atau LIME. Secara praktis, *Random Forest* direkomendasikan sebagai algoritma utama bagi pengembangan sistem pemantauan hidrasi digital yang memprioritaskan akurasi maksimal, sementara KNN hasil optimasi tetap menjadi alternatif kompetitif ketika sumber daya komputasi terbatas dan interpretabilitas menjadi pertimbangan utama.

5. PERNYATAAN PENULIS

5.1 Kontribusi Penulis

Konseptualisasi: J.R.F., dan H.S.;

Metodologi: J.R.F., dan H.S.;

Perangkat Lunak: J.R.F.;

Validasi: J.R.F., dan H.S.;

Analisis Formal: J.R.F., dan H.S.;

Investigasi: J.R.F., dan H.S.;

Sumber Daya: J.R.F.;

Kurasi Data: J.R.F.;

Penulisan Draf Awal: J.R.F., dan H.S.;

Penulisan, Peninjauan, dan Penyuntingan: J.R.F., dan H.S.;

Visualisasi: J.R.F.;

Seluruh penulis telah membaca dan menyetujui versi manuskrip yang telah diterbitkan.

5.2 Ketersediaan Data

Data yang disajikan dalam penelitian ini tersedia berdasarkan permintaan kepada penulis korespondensi.

5.3 Pendanaan

Penelitian ini tidak menerima pendanaan eksternal.

5.4 Pernyataan Dewan Peninjau Institusional

Tidak Berlaku (Not applicable).

5.5 Pernyataan Persetujuan Berdasarkan Informasi

Tidak Berlaku (Not applicable).

5.6 Pernyataan Konflik Kepentingan

Para penulis menyatakan bahwa mereka tidak memiliki kepentingan keuangan yang bersaing atau hubungan pribadi yang diketahui yang dapat dianggap memengaruhi pekerjaan yang dilaporkan dalam artikel ini.

5.7 Penggunaan AI Generatif dan Teknologi Berbantuan AI dalam Proses Penulisan

Para penulis menggunakan AI Generatif untuk meningkatkan kejelasan penulisan dalam artikel ini. Para penulis telah meninjau dan menyunting konten yang dihasilkan dengan bantuan AI serta bertanggung jawab penuh atas versi akhir publikasi.

REFERENCES

- [1] A. Dolci, T. Vanhaecke, J. Qiu, R. Ceccato, R. Arboretti, and L. Salmaso, "Personalized prediction of optimal water intake in adult population by blended use of machine learning and clinical data," *Sci. Rep.*, vol. 12, no. 1, Dec. 2022, doi: 10.1038/s41598-022-21869-y.
- [2] Y. Zeng and F. Cheng, "Medical and Health Data Classification Method Based on Machine Learning," *J. Healthc. Eng.*, vol. 2021, 2021, doi: 10.1155/2021/2722854.
- [3] F. Sholekhah, A. D. Putri, R. Rahmaddeni, and L. Efrizoni, "Perbandingan Algoritma Naïve Bayes dan K-Nearest Neighbors untuk Klasifikasi Metabolik Sindrom," *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 4, no. 2, pp. 507–514, Feb. 2024, doi: 10.57152/malcom.v4i2.1249.
- [4] P. A. Rakhma Devi, "Klasifikasi Penyakit Gagal Ginjal Kronis Dengan Metode Knn (Studi Kasus RS Di Kab Gresik)," *JIPi (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)*, vol. 9, no. 3, pp. 1739–1748, Sep. 2024, doi: 10.29100/jipi.v9i3.6226.
- [5] Sahar, "Analisis Perbandingan Metode K-Nearest Neighbor dan Naïve Bayes Classifier pada Data Set Penyakit Jantung," *Indonesian Journal of Data and Science (IJODAS)*, vol. 1, no. 3, pp. 79–86, 2020.
- [6] V. Wulandari, W. J. Sari, Z. Alfian, L. Legito, and T. Arifianto, "Implementasi Algoritma Naïve Bayes Classifier dan K-Nearest Neighbor untuk Klasifikasi Penyakit Ginjal Kronik," *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 4, no. 2, pp. 710–718, Apr. 2024, doi: 10.57152/malcom.v4i2.1229.
- [7] A. F. Fadhlullah and T. Widiyaningtyas, "Comparative Analysis of Decision Tree and Random Forest Algorithms for Diabetes Prediction," *JTAM (Jurnal Teori dan Aplikasi Matematika)*, vol. 8, no. 4, p. 1121, Oct. 2024, doi: 10.31764/jtam.v8i4.24388.
- [8] Rohana Tatang, Novita Yulia Hilda, and Nurlaelasari Euis, "Komparasi Algoritma Machine Learning Dalam Memprediksi Kapasitas Produksi Potensial Air Bersih Di Indonesia," *Jurnal Teknologi Terpadu*, vol. 11, no. 1, pp. 44–52, Jul. 2025, doi: 10.54914/jtt.v11i1.1732.
- [9] G. L. Pritalia, "Analisis Komparatif Algoritme Machine Learning pada Klasifikasi Kualitas Air Layak Minum," 2022.
- [10] H. Hairani, T. Widiyaningtyas, and D. Dwi Prasetya, "Addressing Class Imbalance of Health Data: A Systematic Literature Review on Modified Synthetic Minority Oversampling Technique (SMOTE) Strategies," *JOIV: International Journal on Informatics Visualization*, vol. 8, no. 3, pp. 1310–1318, Sep. 2024, doi: 10.62527/joiv.8.3.2283.
- [11] A. Syukron, E. Saputro, and P. Widodo, "Penerapan Metode Smote Untuk Mengatasi Ketidakseimbangan Kelas Pada Prediksi Gagal Jantung," 2023. [Online]. Available: <https://doi.org/10.25047/jtit.v10i1.312>
- [12] A. Arisman, N. Nurhayati, and T. Tanti, "Analisis Klasifikasi Tingkat Hidrasi Manusia Menggunakan Algoritma Machine Learning Berbasis Dataset Asupan Air Harian," *Jurnal Minfo Polgan*, vol. 14, no. 2, pp. 3533–3541, Jan. 2026, doi: 10.33395/jmp.v14i2.15829.



- [13] A. M. Argina, “Penerapan Metode Klasifikasi K-Nearest Neighbor pada Dataset Penderita Penyakit Diabetes,” *Indonesian Journal of Data and Science*, vol. 1, no. 2, pp. 29–33, Jul. 2020, doi: 10.33096/ijodas.v1i2.11.
- [14] N. Marito Putry and B. Nurina Sari, “Komparasi Algoritma Knn Dan Naïve Bayes Untuk Klasifikasi Diagnosis Penyakit Diabetes Melitus,” *Jurnal Sains dan Manajemen*, vol. 10, no. 1, 2022.
- [15] M. D. Purbolaksono, M. I. Tantowi, A. I. Hidayat, and Adiwijaya, “Perbandingan Support Vector Machine dan Modified Balanced Random Forest dalam Deteksi Pasien Penyakit Diabetes,” *Jurnal RESTI*, vol. 5, no. 2, pp. 393–399, Apr. 2021, doi: 10.29207/resti.v5i2.3008.
- [16] A. Ridwan, “Penerapan Algoritma Naïve Bayes untuk Klasifikasi Penyakit Diabetes Mellitus,” *Jurnal Sistem Komputer dan Kecerdasan Buatan (SISKOM-KB)*, vol. 4, no. 1, pp. 15–21, 2020, doi: 10.47970/siskom-kb.v4i1.169.
- [17] A. Byna and M. Basit, “Penerapan Metode Adaboost Untuk Mengoptimasi Prediksi Penyakit Stroke Dengan Algoritma Naïve Bayes,” *Jurnal Sisfokom (Sistem Informasi dan Komputer)*, vol. 9, no. 3, pp. 407–411, Nov. 2020, doi: 10.32736/sisfokom.v9i3.1023.
- [18] R. Irfannandhy, L. B. Handoko, and N. Ariyanto, “Analisis Performa Model Random Forest dan CatBoost dengan Teknik SMOTE dalam Prediksi Risiko Diabetes,” *Edumatic: Jurnal Pendidikan Informatika*, vol. 8, no. 2, pp. 714–723, Dec. 2024, doi: 10.29408/edumatic.v8i2.27990.
- [19] W. Nugraha and A. Sasongko, “Hyperparameter Tuning pada Algoritma Klasifikasi dengan Grid Search,” *SISTEMASI: Jurnal Sistem Informasi*, vol. 11, no. 2, pp. 391–401, 2022, doi: 10.32520/stmsi.v11i2.1750.
- [20] C. Z. V. Junus, T. Tarno, and P. Kartikasari, “Klasifikasi Menggunakan Metode Support Vector Machine Dan Random Forest Untuk Deteksi Awal Risiko Diabetes Melitus,” *Jurnal Gaussian*, vol. 11, no. 3, pp. 386–396, Jan. 2023, doi: 10.14710/j.gauss.11.3.386-396.
- [21] B. Pfeifer and M. Kreuzthaler, “Calibrated kNN classification via second-layer neighborhood analysis,” *Adv. Data Anal. Classif.*, vol. 20, no. 1, pp. 145–165, Mar. 2026, doi: 10.1007/s11634-025-00654-5.
- [22] S. Shinde, “Daily Water Intake & Hydration Patterns Dataset,” Kaggle, 2025. [Online]. Available: <https://www.kaggle.com/datasets/sonalshinde123/daily-water-intake-and-hydration-patterns-dataset>. [Accessed: Jan. 29, 2026].