



PERBANDINGAN METODE NAIVE BAYES YANG DIOPTIMALKAN DENGAN ALGORITMA GENETIKA UNTUK PRIORITAS PASIEN

Azizah Kusumah Putri¹, Meira Listyaningrum², Salwa Farhanatussaidah³,
Fajri Farid⁴, M. Syamsuddin Wisnubroto^{5*}

¹azizah.122450068@student.itera.ac.id, ²meira.122450011@student.itera.ac.id,

³salwa.122450055@student.itera.ac.id, ⁴fajri.farid@sd.itera.ac.id, ⁵syamsuddin.wisnubroto@sd.itera.ac.id
^{1,2,3,4,5}Program Studi Sains Data, Fakultas Sains, Institut Teknologi Sumatera

Abstrak

Penentuan prioritas pasien di Instalasi Gawat Darurat (IGD) merupakan proses krusial yang rentan terhadap keterbatasan sumber daya dan subjektivitas pengambilan keputusan. Penelitian ini bertujuan membandingkan kinerja model *Naïve Bayes* dengan *Naïve Bayes* yang dioptimasi menggunakan Algoritma Genetika (NB+GA) dalam klasifikasi prioritas pasien (*Red*, *Orange*, *Yellow*, *Green*). Dataset yang digunakan berasal dari Kaggle *Patient Priority Classification* dengan 17 fitur demografis dan klinis, dibagi menjadi 80% data latih dan 20% data uji. Hasil evaluasi menunjukkan adanya ketidakseimbangan data, terutama pada kelas “Yellow”. Model *Naïve Bayes* menghasilkan akurasi 64% dengan *Precision* 0,44, *Recall* 0,69, dan *F1-Score* 0,44. Setelah optimasi, Algoritma Genetika menyeleksi 8 fitur paling relevan dan meningkatkan performa model NB+GA dengan akurasi 96%, *Precision* 0,87, *Recall* 0,80, dan *F1-Score* 0,81. Hasil ini membuktikan bahwa Algoritma Genetika efektif meningkatkan kinerja *Naïve Bayes* dan berpotensi sebagai sistem pendukung keputusan triase yang objektif di IGD.

Kata kunci: Algoritma Genetika, *Naïve Bayes*, Optimasi, Prioritas Pasien, Triase.

Abstract

Determining patient priority in the Emergency Department (ED) is a crucial process that is vulnerable to resource limitations and subjectivity in decision-making. This study aims to compare the performance of the Naïve Bayes model with Naïve Bayes optimized using Genetic Algorithm (NB+GA) in patient priority classification (Red, Orange, Yellow, Green). The dataset used comes from Kaggle Patient Priority Classification with 17 demographic and clinical features, divided into 80% training data and 20% test data. The evaluation results show that there is data imbalance, especially in the “Yellow” class. The Naïve Bayes model yielded 64% accuracy with Precision 0.44, Recall 0.69, and F1-Score 0.44. After optimization, Genetic Algorithm selects the 8 most relevant features and improves the performance of the NB+GA model with an accuracy of 96%, Precision of 0.87, Recall of 0.80, and F1-Score of 0.81. These results prove that the Genetic Algorithm effectively improves the performance of Naïve Bayes and has the potential to be an objective triage decision support system in the ED.

Keywords: Genetic Algorithm, *Naïve Bayes*, Optimization, Patient Priority, Triage

1. Pendahuluan

Pengambilan keputusan mengenai perawatan pasien menjadi hal yang kompleks dan sangat penting. Salah satu aspek yang memerlukan pertimbangan matang adalah penentuan prioritas siapa yang mendapat perawatan terlebih dahulu. *Triase* menjadi langkah awal penting dalam memilah pasien agar yang membutuhkan penanganan segera dapat didahulukan, sehingga menekan angka kematian dan meningkatkan efisiensi pelayanan. Penelitian menunjukkan bahwa ketepatan proses *triase* sangat dipengaruhi oleh kemampuan perawat dalam mengambil keputusan yang tepat. Namun, tantangan utama proses ini adalah ketika kapasitas pasien melebihi kapasitas tenaga medis yang tersedia. Kondisi tersebut dapat menyebabkan beberapa kendala utama muncul seperti waktu tunggu dan perawatan lambat, beban kerja tenaga medis meningkat, bahkan dapat berdampak pada keselamatan pasien [1].

Triase memiliki peran yang krusial dalam bidang medis, namun masih banyak tenaga medis yang melakukan pengambilan keputusan secara manual. Pada penelitian, mencatat bahwa *framing bias*, *anchoring bias*, dan variabilitas dalam pelatihan dan pengalaman perawat dapat memengaruhi akurasi

keputusan. Banyak faktor yang dapat menyebabkan terjadinya kesalahan dalam penentuan prioritas, termasuk ketika kondisi IGD sibuk dan sumber daya terbatas, tekanan waktu dan kelelahan juga dapat memperburuk ketidakakuratan keputusan *triase* [2]. Karena hal-hal tersebut, diperlukan sistem pendukung pengambilan keputusan berbasis data yang mampu menstandarkan penilaian, meminimalkan bias manusia, serta memberikan rekomendasi prioritas pasien secara objektif dan cepat dengan memanfaatkan berbagai algoritma seperti *machine learning (ML)* dan *artificial intelligence (AI)*.

Salah satu solusi yang dapat digunakan untuk menjawab permasalahan di atas adalah penerapan *machine learning*. Algoritma *machine learning* yang sering digunakan dalam pengambilan keputusan adalah *Naïve Bayes* karena kesederhanaan, efisiensi dalam komputasi, dan kemampuannya dalam memberikan hasil klasifikasi yang baik [3]. Dalam studi kasus ini, *Naïve Bayes* mampu mengestimasi tingkat urgensi pasien secara cepat, sehingga dapat membantu tenaga medis dalam menentukan prioritas penanganan secara objektif. Namun, kinerja model *Naïve Bayes* sangat bergantung pada kualitas data dan pemilihan parameter probabilitas yang optimal. Jika distribusi data antar kelas tidak seimbang atau korelasi antar fitur tinggi, maka akurasi *Naïve Bayes* dapat menurun secara signifikan. Oleh karena itu, diperlukan pendekatan optimasi untuk membantu meningkatkan kemampuan model.

Salah satu pendekatan optimasi yang dapat digunakan guna mengatasi keterbatasan *Naïve Bayes* adalah Algoritma Genetika (GA). GA merupakan algoritma optimasi yang terinspirasi dari mekanisme biologis seperti seleksi alam, reproduksi, dan mutasi. Pada kasus ini, GA dapat dimanfaatkan untuk memilih fitur-fitur yang paling relevan yang akan membantu menyesuaikan bobot probabilitas sehingga dapat membantu *Naïve Bayes* menghasilkan klasifikasi yang lebih akurat dan stabil [4]. Pada penelitian yang dilakukan [5], menunjukkan bahwa integrasi kedua algoritma ini dapat menghasilkan sistem klasifikasi yang mampu mengatasi *noise* dan ketidakseimbangan data. Dengan ini, optimasi menggunakan GA menjadi pendekatan yang potensial untuk meningkatkan kinerja model *Naïve Bayes* dalam penentuan prioritas pasien di IGD.

Berbagai penelitian sebelumnya telah membuktikan efektivitas kombinasi *Naïve Bayes* dan Algoritma Genetika dalam meningkatkan akurasi klasifikasi data medis. Namun, sebagian besar studi menggunakan kombinasi algoritma tersebut untuk kasus diagnosis penyakit tertentu. Sehingga penerapannya pada konteks penentuan prioritas pasien di IGD masih jarang dibahas. Selain itu belum banyak penelitian yang membandingkan secara langsung kinerja *Naïve Bayes* konvensional dengan versi yang dioptimasi menggunakan GA.

Oleh karena itu, penelitian ini bertujuan untuk membandingkan kedua metode tersebut dalam menentukan prioritas pasien di IGD, guna memperoleh model klasifikasi yang lebih akurat dan dapat mendukung proses *triase* secara otomatis. Penelitian ini diharapkan dapat digunakan oleh perawat dan tenaga medis di IGD sebagai alat bantu dalam proses *triase*. Keputusan yang dihasilkan berhubungan dengan penentuan tingkat prioritas penanganan pasien berdasarkan kondisi klinis dan vital.

2. Metode

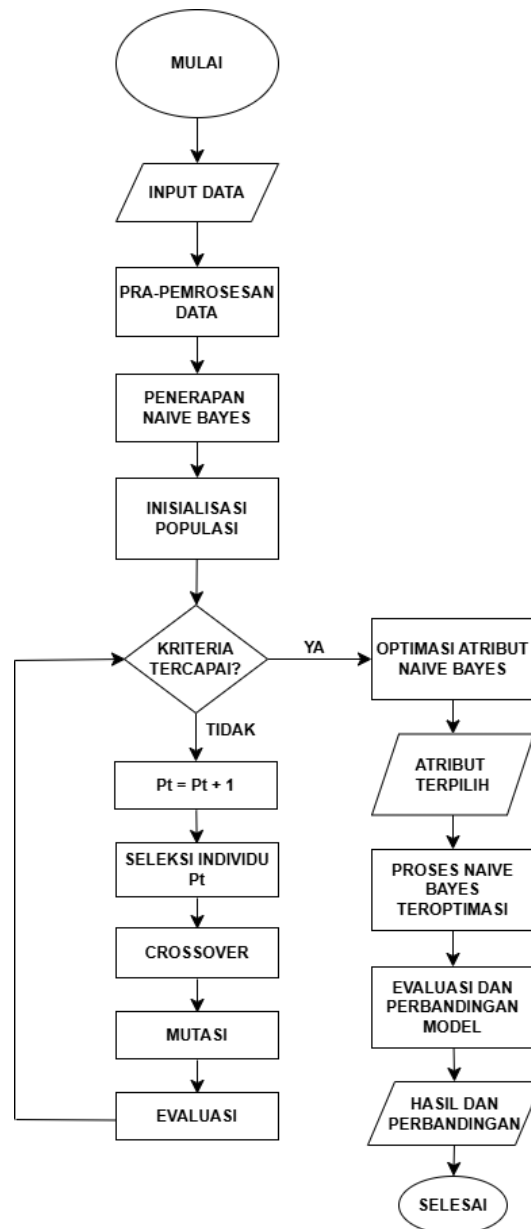
Penelitian ini menggunakan dataset publik bernama “*Patient Priority Classification*” yang dapat diakses dari platform Kaggle [6]. *Dataset* ini berisi klasifikasi prioritas pasien berdasarkan parameter klinis dan administratif, sehingga relevan untuk diterapkan dalam proses *triase* di Instalasi Gawat Darurat (IGD). *Dataset* terdiri dari 17 atribut yang merepresentasikan kondisi fisiologis, riwayat kesehatan, serta gaya hidup pasien.

Dataset penelitian ini terdiri dari berbagai variabel yang menggambarkan kondisi demografis dan klinis pasien, seperti usia, jenis kelamin, jenis nyeri dada, tekanan darah, kadar kolesterol, detak jantung maksimum, kadar glukosa, insulin, serta indeks massa tubuh (BMI). Selain itu, terdapat variabel yang berkaitan dengan riwayat kesehatan dan gaya hidup, termasuk hipertensi, penyakit jantung, status merokok, serta risiko diabetes berdasarkan keturunan. Faktor lingkungan seperti jenis tempat tinggal juga dicantumkan untuk memperkaya analisis. Variabel *triage* digunakan sebagai label atau target yang menunjukkan tingkat prioritas penanganan medis pasien berdasarkan tingkat keparahan kondisi.

Dataset ini digunakan karena menggabungkan berbagai indikator medis, seperti gaya hidup, dan demografis yang umum digunakan dalam proses penilaian kondisi pasien di IGD. Keberagaman fitur tersebut memungkinkan model *machine learning* untuk menganalisis hubungan antar variabel secara lebih komprehensif dalam menentukan tingkat urgensi pasien.

2.1. Tahapan Penelitian

Tahapan penelitian ini disusun secara sistematis supaya proses analisis dan perbandingan antara metode *Naïve Bayes* dan *Naïve Bayes* yang dioptimasi menggunakan Algoritma Genetik.



Gambar 1. Diagram Alir Penelitian

2.2. *Naïve Bayes Classifier*

Naïve Bayes merupakan salah satu metode yang bisa diterapkan dalam proses *data mining* dengan melakukan pengklasifikasian ke kelas yang dibutuhkan. *Naïve Bayes* dapat menentukan perbedaan pada data yang terdapat pada sebuah pola atau model [4]. *Bayesian Classification* terbukti memiliki akurasi dan kecepatan yang tinggi saat diaplikasikan ke dalam *database* dengan jumlah data yang besar [7]. Pada tahap pengaplikasiannya *Naïve Bayes* banyak dipakai pada berbagai bidang contohnya seperti diagnosis medis, klasifikasi teks, prediksi risiko, sampai penentuan prioritas pasien di IGD. *Naïve Bayes* juga memiliki

keunggulan yaitu mampu memproses data dalam jumlah yang besar secara efisien, dan tidak memerlukan waktu pelatihan yang lama.

Keunggulan utama *Naïve Bayes* adalah kemampuannya menangani data dengan dimensi tinggi secara efisien. Selain itu, model ini relatif mudah diimplementasikan dan tidak memerlukan proses pelatihan yang kompleks. Keterbatasan utamanya terletak pada asumsi independensi antar variabel, yang sering kali tidak sepenuhnya sesuai dengan kondisi nyata [8]. Secara umum, *Teorema Bayes* dapat dituliskan sebagai:

$$P(C|X) = \frac{P(X|C) \times P(C)}{P(X)} \quad (1)$$

Keterangan:

- $P(C | X)$: probabilitas kelas C berdasarkan data X (*posterior*)
- $P(X | C)$: probabilitas data X terjadi jika diketahui kelas C (*likelihood*)
- $P(C)$: probabilitas awal kelas C (*prior*)
- $P(X)$: probabilitas total dari data X

2.3. Algoritma Genetika

Algoritma Genetika (GA) adalah metode optimasi berbasis evolusi yang meniru mekanisme seleksi alam seperti reproduksi, mutasi, dan persilangan genetik untuk memperoleh solusi terbaik dari suatu masalah. Algoritma ini bekerja dengan cara membentuk populasi awal yang terdiri dari sejumlah solusi kandidat (kromosom). Setiap kandidat solusi dalam GA direpresentasikan dalam bentuk kromosom, yang merupakan rangkaian nilai yang melambangkan parameter-parameter yang akan dioptimalkan. Setiap kromosom mewakili satu set solusi yang mungkin, misalnya subset fitur yang akan dipakai dalam model klasifikasi. GA mengevaluasi kromosom-kromosom tersebut berdasarkan fungsi *fitness*, kemudian melalui iterasi generasi ke generasi melakukan seleksi, *crossover*, dan mutasi untuk menghasilkan keturunan yang diharapkan memiliki nilai *fitness* lebih baik [7]. Proses evolusi dalam GA dilakukan secara iteratif untuk menghasilkan populasi baru yang memiliki potensi solusi lebih baik dibandingkan generasi sebelumnya [9].

2.4. Integrasi Naïve Bayes dan Algoritma Genetika

Integrasi antara algoritma *Naïve Bayes* dan Algoritma Genetika (NB+GA) merupakan pendekatan *hybrid* yang memiliki tujuan untuk memaksimalkan kinerja klasifikasi *Naïve Bayes* dengan menjalankan optimasi pada setiap langkah seleksi fitur atau penentuan parameter terbaik. *Naïve Bayes* dikenal karena mempunyai keunggulan dalam efisiensi komputasi dan kemampuan menangani data berukuran besar, namun kelemahannya terletak pada asumsi independensi antar fitur yang tidak selalu terpenuhi pada data nyata. Oleh sebab itu, penerapan GA diharapkan dapat membantu memperbaiki kelemahan tersebut melalui metode seleksi evolusioner untuk mendapatkan kombinasi fitur paling tepat terhadap target klasifikasi.

Dalam integrasi ini, Algoritma Genetika berperan sebagai *feature selector*, yaitu memilih *subset* fitur yang optimal agar proses klasifikasi oleh *Naïve Bayes* menjadi lebih akurat dan efisien. Setiap kromosom dalam Algoritma Genetika akan menggambarkan himpunan fitur, di mana gen bernilai 1 menunjukkan fitur yang dipilih sedangkan gen bernilai 0 menunjukkan fitur yang diabaikan. Nilai *fitness* dari setiap individu dihitung berdasarkan akurasi model *Naïve Bayes* terhadap data latih, sehingga individu dengan akurasi tertinggi memiliki peluang lebih besar untuk dipertahankan atau dikombinasikan dalam generasi berikutnya [7].

2.5. Evaluasi Model dan Matrik Kinerja

Evaluasi model digunakan untuk mengetahui seberapa jauh performa algoritma *Naïve Bayes* setelah dilakukan optimasi menggunakan Algoritma Genetika dalam menentukan prioritas pasien di IGD. Pengukuran kinerja model dilakukan dengan menggunakan empat metrik utama, yaitu akurasi, presisi, *recall*, dan *F1-score* [10]. Akurasi digunakan untuk mengetahui seberapa besar proporsi prediksi yang benar terhadap keseluruhan data uji. Presisi mengukur seberapa banyak prediksi positif yang benar-benar sesuai dengan kondisi aktual. *Recall* menunjukkan kemampuan model dalam mendeteksi seluruh kasus

positif yang ada. *F1-score* digunakan sebagai ukuran keseimbangan antara presisi dan *recall*, terutama ketika distribusi data antar kelas tidak seimbang. Seluruh metrik ini dihitung berdasarkan hasil prediksi pada data uji, kemudian dibandingkan antara model *Naïve Bayes* standar dan model yang telah dioptimasi dengan Algoritma Genetika untuk menilai efektivitas proses optimasi terhadap peningkatan kinerja klasifikasi. Ketika evaluasi dilakukan secara sistematis, hasilnya tidak hanya memberikan gambaran apakah tujuan tercapai, melainkan juga menawarkan landasan untuk pengambilan keputusan, perbaikan, atau pengembangan program/sistem di masa depan [11].

3. Hasil dan Pembahasan

3.1. Statistik Deskriptif

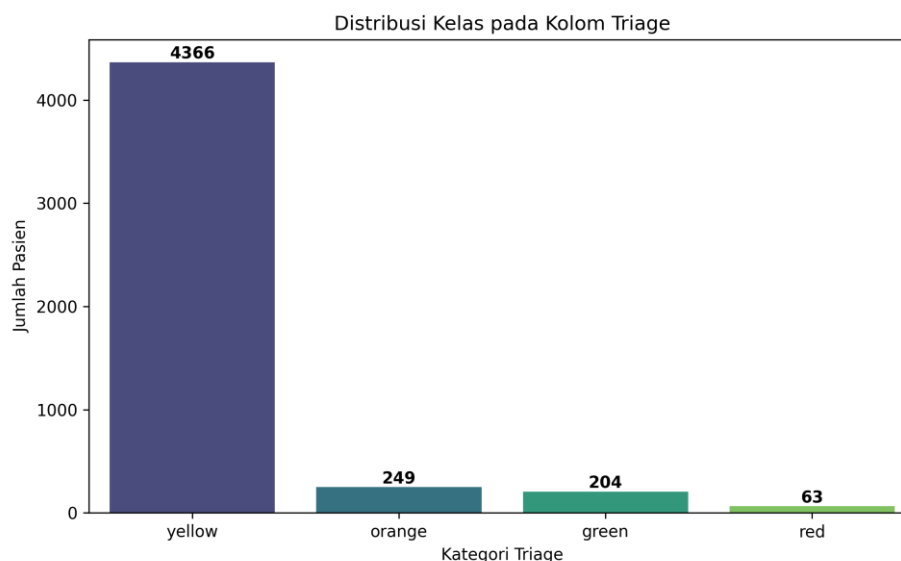
Dataset yang digunakan dalam penelitian ini terdiri dari 4882 data pasien IGD dengan menggunakan semua fitur. Nilai usia pasien berkisar antara 28 hingga 82 tahun dengan rata-rata sekitar 57 tahun. Variabel medis seperti tekanan darah memiliki rentang antara 60–164 mmHg, kolesterol berkisar antara 150–262 mg/dL, *max heart rate* berkisar antara 138–202 bpm, *plasma glucose* nilainya antara 55,12–185,31 mg/dL, *skin thickness* berkisar antara 21–98 mm, kadar insulin berkisar antara 81–160 μ IU/mL, BMI (*Body Mass Index*) berada pada rentang 10,3–49,5 kg/m², dan diabetes *pedigree* berkisar antara 0,174–0,761.

Tabel 1. Statistik Deskriptif

Variables	Mean	Median	Std	Min	Max
Age	57,964	57	11,827	28	82
Blood preassure	109,339	111	20,085	60	164
Cholesterol	178,561	177	20,085	150	262
Max heart rate	164,191	164	15,563	138	202
Plasma Glucose	95,534	91	25,255	55,12	185,31
Skin Thickness	57,392	56	22,792	21	98
Insulin	110,956	111	16,873	81	160
BMI	27,508	27	6,771	10,3	49,5
Diabetes Pedigree	0,461	0,467	0,053	0,174	0,761

Fitur dengan jenis kategorik seperti *gender*, *chest pain type*, *exercise angina*, *hypertension*, *heart disease*, *Residence type*, *smoking* status memiliki distribusi kelas yang berbeda-beda. Distribusi jenis kelamin menunjukkan bahwa pasien perempuan sedikit lebih banyak dibandingkan laki-laki, sekitar 50,88% dari total data. Distribusi fitur *chest pain type* paling banyak adalah tipe 0. Distribusi fitur *exercise angina*, *hypertension*, dan *heart disease* semuanya berada pada kelas 0. Sementara itu, mayoritas pasien berasal dari wilayah *urban*, dengan status merokok didominasi oleh kategori *never smoked*. Berikut adalah visualisasi untuk distribusi kelas pada variabel target *triage*.

Berdasarkan Gambar 2 yang menunjukkan distribusi kelas pada variabel *triage*, terlihat bahwa data sangat tidak seimbang. Sebagian besar pasien berada pada kelas *yellow* dengan 4366 pasien, sementara kelas lainnya seperti *orange*, *green*, dan *red* memiliki jumlah yang jauh lebih sedikit. Ketidakeimbangan ini menunjukkan bahwa model *Naïve Bayes* yang dilatih dengan data tersebut akan berisiko bias terhadap kelas mayoritas.



Gambar 2. Distribusi Kelas Pada Variabel Target

3.2. Hasil Evaluasi Model Naïve Bayes

Pengujian sebuah model *Naïve Bayes* melibatkan pembagian data menjadi data *training* dan data *testing* dengan perbandingan 80% data *training* dan 20% data *testing*. Data ini akan digunakan untuk evaluasi performa model. Kemudian evaluasi terhadap model *Naïve Bayes* ini digunakan sebagai pembandingan dengan model *Naïve Bayes* yang dikombinasikan dengan algoritma genetika.

Confusion matrix digunakan untuk memahami pola klasifikasi triase yang dihasilkan oleh model *Naïve Bayes*. Tabel *confusion matrix* disajikan pada Tabel 2.

Tabel 2. *Confusion Matrix* Model Naive Bayes

Akurasi = 64%		Prediction				Recall
		green	yellow	orange	red	
Actual	green	23	0	0	13	0,64
	yellow	28	564	71	215	0,64
	orange	0	0	23	24	0,49
	red	0	0	0	16	1
Precision		0,45	1	0,24	0,06	

Berdasarkan Tabel 2, *confusion matrix* menunjukkan jumlah prediksi yang benar dan salah untuk setiap kelas triase. Dari tabel tersebut, dapat diamati bahwa model cenderung melakukan kesalahan klasifikasi tertentu. Contohnya terdapat 28 pasien dari kelas *yellow* yang justru diprediksi sebagai kelas *green*, terdapat 71 pasien dari kelas *yellow* yang salah diprediksi menjadi kelas *orange*, sebanyak 13 pasien dari kelas *green* yang salah diprediksi menjadi kelas *red*, terdapat 215 pasien dari kelas *yellow* yang salah diprediksi menjadi *red*, dan sebanyak 24 pasien dari kelas *orange* yang salah diprediksi menjadi kelas *red*. Kesalahan klasifikasi ini banyak terjadi pada kelas *yellow*, yang awalnya 878 pasien, hanya 564 pasien yang diprediksi benar, sementara sisanya diklasifikasikan sebagai *green*, *red*, dan *orange*. Model kesulitan membedakan kelas *yellow* dengan kelas lain akibat adanya ketidakseimbangan data atau kemiripan fitur antar kelas.

Performa kuantitatif model *Naïve Bayes* dievaluasi menggunakan beberapa metrik evaluasi seperti nilai akurasi, presisi, *recall*, dan *F1-score*. Metrik evaluasi tersebut disajikan pada Tabel 3.

Tabel 3. Metrik Evaluasi

Metric	Value
Accuracy	0,64
Precision	0,44
Recall	0,69
F1-score	0,44

Berdasarkan Tabel 3., secara keseluruhan model *Naïve Bayes* menunjukkan akurasi sebesar 64% yang mengindikasikan kemampuan model dalam melakukan prediksi sesuai dengan label sebenarnya. Presisi cukup rendah yaitu 44% menandakan bahwa banyak prediksi model yang tidak tepat sasaran dan model sering memprediksi kelas tertentu. Sebaliknya, nilai *recall* cukup tinggi yaitu 69% menunjukkan bahwa model berhasil menangkap sebagian besar kasus aktual dari tiap kelas. Sementara nilai *F1-score* sebesar 44% menunjukkan bahwa keseimbangan antara *recall* dan *precision* masih lemah.

3.3. Proses Optimasi Algoritma Genetika

Optimasi menggunakan algoritma genetika dilakukan dengan tujuan untuk menyeleksi *subset* fitur yang paling relevan dan informatif dalam menentukan prioritas pasien. Proses ini menggunakan representasi kromosom yang setiap gennya merepresentasikan sebuah fitur dengan nilai biner 1 dan 0. Nilai 1 menandakan fitur tersebut terpilih dan nilai 0 menandakan tidak terpilih. Panjang dari kromosom adalah 17 gen, di mana masing-masing gen merepresentasikan satu fitur dalam dataset. Setelah melalui proses evolusi selama beberapa generasi, seleksi fitur menggunakan algoritma genetika mendapatkan solusi terbaik yang direpresentasikan oleh kromosom [0, 1, 1, 0, 1, 0, 0, 1, 1, 0, 0, 1, 0, 1, 1, 0, 0]. Solusi ini menunjukkan bahwa dari 17 fitur awal, hanya 8 fitur yang terpilih untuk membangun model *Naïve Bayes* yang optimal. Sementara 9 fitur lainnya dapat dieliminasi tanpa mengurangi kualitas klasifikasi. Fitur-fitur dengan indeks 2, 3, 5, 8, 9, 12, 14, dan 15 memberikan kontribusi paling signifikan dalam menentukan prioritas pasien. Fitur-fitur tersebut adalah *age*, *gender*, *blood pressure*, *exercise angina*, *plasma glucose*, *bmi*, *hypertension*, dan *heart disease*.

Fungsi objektif yang didapatkan sebesar 0,0335. Nilai fungsi objektif ini digunakan untuk mengukur kualitas solusi dan nilai yang rendah merepresentasikan tingkat *error* klasifikasi yang sangat rendah. Dalam hal ini, algoritma genetika mampu melakukan proses seleksi fitur dengan efektif, di mana dengan menggunakan hanya 47% dari total fitur awal model telah mencapai performa yang lebih baik dibandingkan menggunakan seluruh fitur.

3.4. Hasil Evaluasi Model Naïve Bayes yang Dioptimasi Algoritma Genetika

Model *Naïve Bayes* yang dioptimasi dengan algoritma genetika diuji menggunakan data *testing* dengan proporsi 20% dan data *training* dengan proporsi 80%. *Confusion matrix* yang dihasilkan dari model *Naïve Bayes* yang dioptimasi dengan algoritma genetika disajikan pada Tabel 4.

Tabel 4. *Confusion Matrix* model NB+GA

Akurasi = 96%		Prediction				Recall
		green	yellow	orange	red	
Actual	green	36	0	0	0	0,82
	yellow	7	857	14	0	0,98
	orange	0	10	37	0	0,52
	red	1	5	3	7	0,44
Precision		0,82	0,98	0,69	1	

Berdasarkan Tabel 4. *confusion matrix* dari model *Naïve Bayes* + Algoritma Genetika memperlihatkan pola klasifikasi yang lebih baik dibandingkan dengan model *Naïve Bayes* saja. Untuk kelas *green*, model mengklasifikasi semua 36 pasien dengan benar. Pada kelas *yellow* yang merupakan kelas mayoritas, model menunjukkan kinerja dengan baik dengan sedikit kesalahan, sebanyak 7 pasien

diklasifikasikan ke kelas *green* dan 14 pasien ke kelas *orange*. Untuk kelas *orange*, model berhasil mendeteksi 37 pasien dengan tepat, sementara 10 pasien diklasifikasi ke kelas *yellow*. Sementara itu, kelas *red* hanya setengah yang dapat diklasifikasi dengan tepat.

Model *Naive Bayes* yang dioptimasi dengan algoritma genetika juga meningkatkan beberapa metrik evaluasi, seperti yang disajikan pada Tabel 5.

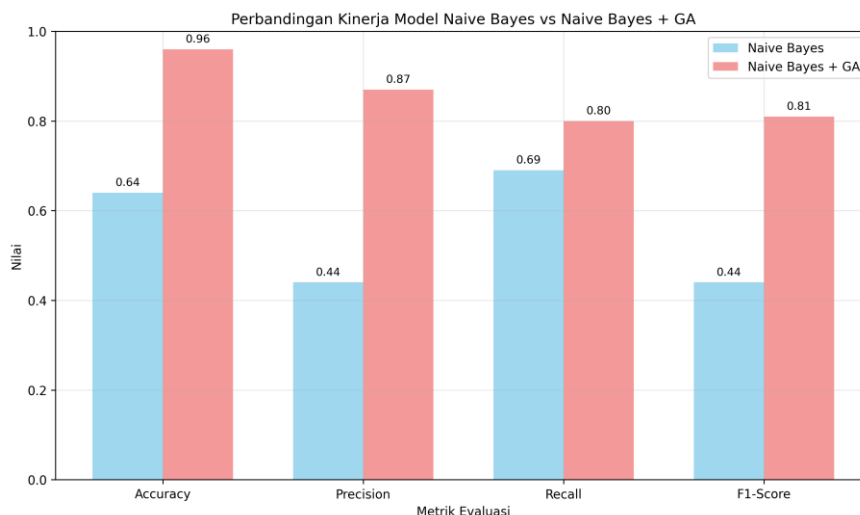
Tabel 5. Metrik Evaluasi

<i>Metric</i>	<i>Value</i>
<i>Accuracy</i>	0,64
<i>Precision</i>	0,44
<i>Recall</i>	0,69
<i>F1-score</i>	0,44

Berdasarkan metrik evaluasi pada Tabel 5., model *Naive Bayes* yang dioptimasi dengan algoritma genetika menunjukkan performa yang sangat baik dalam menentukan prioritas pasien. Nilai akurasi sebesar 0,96 menunjukkan bahwa model mampu melakukan klasifikasi dengan tingkat keberhasilan tinggi. Nilai presisi sebesar 0,87 artinya bahwa ketika model memprediksi suatu kelas maka prediksi tersebut benar sekitar 87%. Sementara itu, nilai *recall* sebesar 0,80 menunjukkan bahwa model dapat mengidentifikasi 80% dari seluruh kasus. *F1-score* menunjukkan keseimbangan antara presisi dan *recall* yang nilainya sebesar 0,81.

3.5. Analisis Perbandingan

Analisis perbandingan ini mencakup empat metrik evaluasi yaitu akurasi, presisi, *recall*, dan *F1-score*. Perbandingan ini bertujuan untuk memberikan bukti mengenai seberapa efektif optimasi algoritma genetika terhadap performa model *Naive Bayes*. Visualisasi perbandingan ini disajikan pada Gambar 3.



Gambar 3. Perbandingan Kinerja Model

Berdasarkan grafik perbandingan tersebut, dapat dilihat bahwa optimasi dengan algoritma genetika berhasil meningkatkan kinerja model. Pada metrik akurasi, model NB+GA mencapai nilai 0,96 yang lebih tinggi dibandingkan model *Naive Bayes* dasar sebesar 0,65, artinya kemampuan klasifikasi yang meningkat secara keseluruhan. Pada metrik lainnya seperti presisi, *recall*, dan *F1-score* juga mengalami peningkatan untuk model NB+GA, ini mengindikasikan bahwa model lebih teliti dalam meminimalisir *false positive*, serta keseimbangan yang lebih baik antara presisi dan *recall*. Algoritma genetika telah berhasil mengidentifikasi fitur-fitur paling relevan yang berkontribusi terhadap akurasi klasifikasi model serta mengeliminasi fitur-fitur yang kurang informatif dan tidak relevan. Model *Naive Bayes* yang dioptimasi dengan algoritma genetika terbukti lebih unggul dibandingkan dengan model *Naive Bayes* dasar.

4. Kesimpulan dan Saran

4.1. Kesimpulan

Berdasarkan hasil pengujian yang dilakukan, ditemukan bahwa model *Naïve Bayes* konvensional memperoleh akurasi sebesar 64%, dengan nilai *precision* 0.44, *recall* 0.69, dan *F1-score* 0.44. Sementara, model *Naïve Bayes* yang dioptimasi dengan Algoritma Genetika menunjukkan peningkatan signifikan dengan akurasi sebesar 96%, *precision* 0.87, *recall* 0.80, dan *F1-score* 0.81. Peningkatan tersebut menunjukkan bahwa optimasi menggunakan Algoritma Genetika mampu meningkatkan kinerja Algoritma *Naïve Bayes* melalui seleksi fitur. Hasil ini mengindikasikan bahwa kombinasi kedua algoritma tersebut dapat memberikan hasil prediksi yang lebih akurat, kuat, dan cocok digunakan dalam proses triase pasien. Secara keseluruhan, NB+GA terbukti lebih unggul dalam menentukan prioritas pasien, sehingga memiliki potensi sebagai model pendukung keputusan untuk membantu tenaga medis dalam pengambilan keputusan prioritas pasien di IGD.

4.2. Saran

Saran untuk penelitian selanjutnya, dapat dilakukan pengujian dengan jumlah data yang lebih besar dan beragam agar memperoleh hasil yang lebih general dan *robust*. Selain itu, integrasi dengan metode optimasi lain seperti *Particle Swarm Optimization (PSO)* atau *Ant Colony Optimization (ACO)* dapat dipertimbangkan untuk dibandingkan dengan hasil optimasi menggunakan Algoritma Genetika. Pengembangan sistem berbasis web atau aplikasi *real-time* juga dapat menjadi langkah lanjutan agar model *Naïve Bayes* yang dioptimasi ini dapat diterapkan langsung di lingkungan IGD sebagai sistem pendukung keputusan bagi tenaga medis dalam menentukan prioritas pasien secara cepat dan akurat.

Daftar Pustaka

- [1] M. Purnamasari Eka Putri, T. Abdur Rasyid, and Stik. Hang Tuah Pekanbaru, "Gambaran Pelaksanaan Triase di Instalasi Gawat Darurat (IGD) RSUD Raja Musa Sungai Guntung Kabupaten Indragiri Hilir," *Jurnal Keperawatan Hang Tuah (Hang Tuah Nursing Journal)*, Aug. 2022, doi: <https://www.doi.org/10.25311/jkh.Vol2.Iss2.821>.
- [2] T. N. Egoda Kapuralalage, H. F. Chan, U. Dulleck, J. A. Hughes, B. Torgler, and S. Whyte, "Clinical decision-making: Cognitive biases and heuristics in triage decisions in the emergency department," *American Journal of Emergency Medicine*, vol. 92, pp. 60–67, Jun. 2025, doi: 10.1016/j.ajem.2025.02.043.
- [3] I. Khairina, H. Malini, E. Huriani, B. Keperawatan Medikal Bedah-Keperawatan Gawat Darurat Fakultas Keperawatan Universitas Andalas ABSTRAK Kata kunci, F. Dominan Keselamatan Pasien Triase, and P. Fakultas Keperawatan, "Faktor-Faktor yang Berhubungan dengan Pengambilan Keputusan Perawat dalam Ketepatan Triase di Kota Padang," *Indonesian Journal for Health Sciences*, vol. 02, no. 01, pp. 1–6, 2018, [Online]. Available: <http://journal.umpo.ac.id/index.php/IJHS/>.
- [4] T. Dzulkarnain, D. E. Ratnawati, and B. Rahayudi, "Penggunaan Metode Naïve Bayes Classifier pada Analisis Sentimen Penilaian Masyarakat Terhadap Pelayanan Rumah Sakit di Malang," *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 11, no. 5, pp. 993–1000, Oct. 2024, doi: 10.25126/jtiik.2024117979.
- [5] M. Moustapha Mbaye, F. M H Abu Salem, and N. Tahiri, "A new machine learning workflow to create an optimal waiting list in hospitals," in *ACM International Conference Proceeding Series*, Association for Computing Machinery, May 2023, pp. 159–163. doi: 10.1145/3608298.3608328.
- [6] H. A. Aly, "Patient Priority for Clustering," Kaggle Dataset, 2023. [Online]. Available: <https://www.kaggle.com/datasets/hossamahmedaly/patient-priority-classification>.

-
- [7] M. Agustriya, M. Ula, and K. -, “Analisis Kinerja Algoritma Klasifikasi Naïve Bayes Menggunakan Genetic Algorithm dan Bagging untuk Data Publik Risiko Transaksi Kartu Kredit,” *Jurnal Sistem dan Teknologi Informasi (JustIN)*, vol. 12, no. 3, p. 584, Jul. 2024, doi: 10.26418/justin.v12i3.80136.
 - [8] Y. Arkeman, K. B. Seminar, and H. Gunawan, *Algoritma Genetika: Teori dan Aplikasinya untuk Bisnis dan Industri*. Bogor: PT Penerbit IPB Press, 2012.
 - [9] F. Sabry, *Naive Bayes Classifier: Fundamentals and Applications*. One Billion Knowledgeable, 2023.
 - [10] T. F. G. da Silva, A. G. Evsukoff, and R. M. E. M. da Costa, “A comparative study of classification algorithms for imbalanced data in sentiment analysis,” in *Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 2021, pp. 2632–2637. doi: 10.1109/SMC52423.2021.9658663.
 - [11] E. Winaryati, M. Munsarif, M. Mardiana, and S. Suwahono, *Model-Model Evaluasi, Aplikasi dan Kombinasinya*. Bantul, Yogyakarta: Penerbit KBM Indonesia, 2021.