

Sistem Deteksi Dini Anemia pada Anak Usia 0-59 Bulan Menggunakan Naïve Bayes dan Optimasi Particle Swarm Optimization

Azzikra Ramadhanti Aksan^{a,1,*}, Deviana Dyah Anggraini^{a,2}, Muhamad Fikry Maulana Ridwan^{a,3},
Herdiesel Santoso^{b,4}

^aSistem Informasi, FTTI Unjaya, Daerah Istimewa Yogyakarta, Indonesia

^bSistem Informasi, STMIK El Rahma, Daerah Istimewa Yogyakarta, Indonesia

¹azzikraaksan@gmail.com, ²devianadyaaah@gmail.com, ³fikrymaulanaaa06@gmail.com,

⁴herdiesel.santoso@stmikelrahma.ac.id

*Azzikra Ramadhanti Aksan

Diterima	Direvisi	Disetujui	Dipublikasikan
16/1/2025	8/6/2025	11/6/2025	13/6/2025

ABSTRACT

Anemia in children aged 0–59 months poses a serious health concern with long-term effects on growth and development. This study aims to develop a web-based early detection system for childhood anemia using a Naïve Bayes algorithm enhanced with Particle Swarm Optimization (PSO). The system uses secondary data from the 2018 Demographic and Health Survey in Nigeria, which includes variables such as age, nutritional status, and medical history. Although the dataset is from Nigeria, the variables are universal and relevant, making the findings applicable for similar model development in other regions. The Naïve Bayes algorithm is employed for classifying anemia levels, while PSO is applied to improve prediction accuracy by optimizing feature weights and tuning model parameters. Results show an increase in accuracy from 92.17% to 95.71% after optimization. This demonstrates PSO's effectiveness in improving model performance, especially in datasets with imbalanced class distributions. The system is implemented as a user friendly website, allowing quick and accessible anemia detection. This solution is particularly useful in regions with limited healthcare access. The findings indicate that combining Naïve Bayes with PSO can enhance predictive accuracy and support broader efforts to improve child health outcomes.

KEYWORDS

Early Detection
Anemia
Naïve Bayes
Particle Swarm
Optimization

ABSTRAK

Anemia pada anak usia 0–59 bulan merupakan masalah kesehatan yang serius dengan dampak jangka panjang terhadap pertumbuhan dan perkembangan. Penelitian ini bertujuan mengembangkan sistem deteksi dini anemia berbasis website menggunakan algoritma Naïve Bayes yang dioptimasi dengan Particle Swarm Optimization (PSO). Sistem ini memanfaatkan data sekunder dari Survei Demografi dan Kesehatan 2018 di Nigeria, yang mencakup variabel seperti usia, status gizi, dan riwayat

KATA KUNCI

Deteksi Dini
Anemia
Naïve Bayes
Particle Swarm
Optimization

kesehatan. Meskipun data berasal dari Nigeria, variabel yang digunakan bersifat universal dan relevan untuk pengembangan model serupa di Indonesia. Algoritma Naïve Bayes digunakan untuk mengklasifikasikan tingkat anemia, sementara PSO diterapkan untuk meningkatkan akurasi model melalui optimasi bobot fitur dan penyetelan parameter. Hasil evaluasi menunjukkan peningkatan akurasi dari 92,17% menjadi 95,71% setelah optimasi, menandakan efektivitas PSO dalam menangani distribusi kelas yang tidak seimbang dan meningkatkan kinerja klasifikasi. Sistem ini diimplementasikan dalam bentuk website yang mudah digunakan, memungkinkan pengguna melakukan deteksi anemia secara cepat dan praktis. Hasil penelitian menunjukkan bahwa kombinasi Naïve Bayes dan PSO mampu meningkatkan akurasi prediksi anemia secara signifikan. Sistem ini berpotensi diterapkan di wilayah dengan akses layanan kesehatan terbatas untuk mendukung peningkatan kualitas kesehatan anak-anak di Indonesia.

This is an open access article under the CC-BY-SA license.



1 PENDAHULUAN

Anemia, yang dikenal masyarakat sebagai kurang darah adalah suatu kondisi medis yang terjadi ketika kadar hemoglobin dalam darah lebih rendah dari normal, sehingga mengurangi kemampuan darah untuk membawa oksigen ke seluruh tubuh. Kondisi ini dapat menyebabkan gejala seperti kelelahan, pucat, sesak nafas, pusing, dan penurunan daya tahan tubuh, yang dapat menghambat fungsi pernapasan sel dan aktivitas tubuh secara normal [1]. Anemia umumnya terjadi ketika jumlah sel darah merah atau hemoglobin terlalu rendah untuk memenuhi kebutuhan tubuh.

Menurut data *World Health Organization* (WHO), lebih dari 496 juta orang di seluruh dunia, khususnya pada usia 15-49 tahun, mengalami anemia [2]. Di Indonesia, prevalensi anemia masih cukup tinggi, terutama di kalangan remaja putri dan ibu hamil. Hal ini dapat menyebabkan berbagai dampak negatif, seperti gangguan pertumbuhan, penurunan produktivitas, serta peningkatan risiko komplikasi kesehatan. Oleh karena itu, deteksi dini dan intervensi yang tepat sangat penting untuk menekan angka prevalensi anemia.

Deteksi dini anemia dapat dilakukan melalui berbagai metode, seperti skrining rutin yang ditargetkan berdasarkan faktor risiko, penggunaan instrumen penilaian risiko, atau tes diagnostik [3]. Dalam hal ini, teknologi *Machine Learning* (ML) menawarkan potensi besar untuk mendeteksi penyakit secara lebih akurat dan efisien. Algoritma *Machine Learning* mampu menganalisis data medis yang kompleks, seperti hasil laboratorium, riwayat kesehatan, atau faktor gaya hidup, untuk mendeteksi penyakit, termasuk anemia [4].

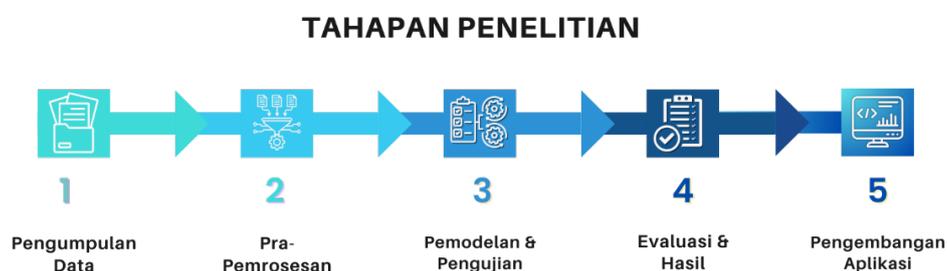
Beberapa penelitian sebelumnya telah mengkaji penggunaan *Machine Learning* untuk prediksi anemia. Penelitian Ningrum et al. (2021) menggunakan algoritma Fuzzy Tsukamoto dengan 4 tahapan untuk mendiagnosis anemia, dan mendapatkan hasil akurasi sebesar 85% [5]. Penelitian oleh Sukron (2021) yang mengoptimalkan metode Naïve Bayes dengan *Particle Swarm Optimization* (PSO) menunjukkan bahwa metode Naïve Bayes menghasilkan akurasi sebesar 86,80%, sementara setelah dioptimasi dengan PSO, akurasi meningkat menjadi 89,84%, dengan peningkatan sebesar 3,04% [6]. Selain itu, penelitian Yanti et al. (2023) menunjukkan bahwa algoritma C4.5 lebih unggul dibandingkan SVM, dengan akurasi 99,29%, presisi 98,7%, dan recall 99,69% [7]. Begitu pula dengan penelitian Sari et al. (2024) yang membandingkan algoritma C4.5 dan *K-Nearest Neighbor* (K-NN), dimana C4.5 menghasilkan akurasi 99,75%, presisi 100%, dan recall 99,5s0% [8].

Berdasarkan penelitian-penelitian tersebut, dapat disimpulkan bahwa penggunaan algoritma data mining sangat membantu tenaga kesehatan untuk mendeteksi penyakit secara lebih cepat dan akurat. Penelitian ini bertujuan untuk membangun sistem deteksi dini anemia pada anak usia 0-59 bulan dengan menggunakan algoritma Naïve Bayes yang di optimasi dengan PSO. Naïve Bayes

merupakan algoritma pembelajaran mesin yang menggunakan perhitungan probabilitas dan statistik untuk memprediksi kemungkinan berdasarkan pengalaman masa lalu yang dikembangkan oleh ilmuwan Inggris Thomas Bayes [9]. Untuk meningkatkan akurasi model, PSO digunakan untuk mengoptimasi parameter Naïve Bayes. PSO adalah algoritma optimasi yang mampu mencari nilai optimal untuk meningkatkan akurasi model [10]. Teknologi berbasis ML ini diimplementasikan dalam platform berbasis web, memberikan solusi praktis dan efektif dalam menganalisis data medis yang kompleks. Dengan adanya platform ini, deteksi anemia dapat dilakukan secara lebih luas, bahkan di daerah dengan akses terbatas ke fasilitas kesehatan, yang pada gilirannya dapat berkontribusi pada peningkatan kualitas kesehatan masyarakat, khususnya dalam menangani masalah anemia pada anak.

2 METODE PENELITIAN

Dalam penelitian ini, metodologi yang digunakan mencakup pengumpulan data, pra-pemrosesan, pemodelan, pengujian, evaluasi hasil, serta pengembangan aplikasi, yang keseluruhannya dapat dilihat dalam diagram alir pada Gambar 1 berikut.



Gambar 1. Tahap Penelitian

2.1 Pengumpulan Data

Pengumpulan data adalah usaha sistematis untuk mengumpulkan informasi atau data yang relevan dengan tujuan mengukur variabel tertentu, menjawab pertanyaan penelitian, atau menguji hipotesis yang telah dirumuskan sebelumnya [11]. Tahap ini digunakan untuk memastikan bahwa data yang diperoleh sesuai dengan kebutuhan penelitian, sehingga analisis yang dilakukan dapat memberikan hasil yang valid, reliabel, dan dapat diandalkan. Terlepas dari bidang studi atau preferensi untuk mendefinisikan data (kuantitatif, kualitatif), pengumpulan data yang akurat sangat penting untuk menjaga integritas penelitian [12]. Penelitian ini menggunakan data sekunder yang diambil dari Kaggle, sebuah platform berbasis komunitas yang menyediakan berbagai dataset dari beragam sumber. Dataset yang digunakan terdiri dari 33924 baris, dengan total 17 atribut sebelum proses pra-pemrosesan. Data ini dikumpulkan berdasarkan Survei Demografi dan Kesehatan 2018 pada anak usia 0-59 bulan di Nigeria [13]. Meskipun data berasal dari Nigeria (DHS 2018), variabel yang digunakan universal dan menjadi rujukan model yang relevan di Indonesia. Validasi lebih lanjut dengan data lokal akan menjadi fokus pengembangan sistem di masa mendatang.

Penelitian ini terdiri beberapa atribut yang datanya diambil dari situs *Kaggle Dataset*, yang dapat dilihat pada Tabel 1 berikut.

Tabel 1. Deskripsi Data

No.	Atribut	Tipe Data	Keterangan
1.	<i>Age in 5-year groups</i>	<i>Object</i>	Usia dalam kelompok 5 tahun
2.	<i>Type of place of residence</i>	<i>Object</i>	Jenis tempat tinggal
3.	<i>Highest educational level</i>	<i>Object</i>	Tingkat pendidikan tertinggi
4.	<i>Wealth index combined</i>	<i>Object</i>	Indeks kekayaan gabungan
5.	<i>Births in last five years</i>	<i>Integer</i>	Jumlah kelahiran dalam lima tahun terakhir

6.	<i>Age of respondent at 1st birth</i>	<i>Integer</i>	Usia responden saat kelahiran pertama
7.	<i>Hemoglobin level adjusted for altitude and smoking, g/dl - 1 decimal</i>	<i>Float</i>	Tingkat hemoglobin yang disesuaikan dengan ketinggian dan merokok
8.	<i>Anemia level</i>	<i>Object</i>	Tingkat anemia
9.	<i>Have mosquito bed net for sleeping (from household questionnaire)</i>	<i>Object</i>	Memiliki kelambu untuk tidur (dari kuesioner rumah tangga)
10.	<i>Smokes cigarettes</i>	<i>Object</i>	Status merokok
11.	<i>Current marital status</i>	<i>Object</i>	Status pernikahan saat ini
12.	<i>Currently residing with husband/partner</i>	<i>Object</i>	Saat ini tinggal bersama suami/pasangan
13.	<i>When child put to breast</i>	<i>Object</i>	Ketika kecil apakah menyusui?
14.	<i>Had fever in last two weeks</i>	<i>Object</i>	Pernah demam dalam dua minggu terakhir
15.	<i>Hemoglobin level adjusted for altitude, g/dl - 1 decimal</i>	<i>Float</i>	Tingkat hemoglobin yang disesuaikan dengan ketinggian
16.	<i>Anemia level.1</i>	<i>Object</i>	Anemia tingkat 1
17.	<i>Taking iron pills, sprinkles or syrup</i>	<i>Object</i>	Mengonsumsi pil besi, taburan atau sirup

2.2 Pra-Pemrosesan

Tahap pra-pemrosesan adalah fase untuk mempersiapkan data mentah untuk analisis dengan mengatasi berbagai masalah seperti ketidaklengkapan, redundansi, dan inkonsistensi. Tahap ini meningkatkan kualitas data, yang penting untuk penambahan data dan pengambilan keputusan yang efektif. Bagian berikut menguraikan komponen kunci dari pra-pemrosesan data. Tahap ini melibatkan pembersihan data, transformasi, integrasi, dan atribusi. Tahap ini memproses data mentah yang tidak sempurna untuk mempersiapkannya untuk analisis, memastikan data cocok untuk langkah-langkah selanjutnya seperti penambahan data dan analisis visual [14].

Tahap pra-pemrosesan ini dilakukan dengan menggunakan beberapa library, yaitu *pandas*, *numpy*, *scikit-learn*, *sklearn.preprocessing*, *pswarms*, dan *joblib*. Langkah pertama dilakukan pembersihan data dengan menghapus baris yang mengandung nilai NaN (*Not a Number*) atau nilai yang hilang pada kolom "Anemia level.1". Selain itu, atribut yang tidak relevan dengan proses klasifikasi juga dihapus seperti '*Current marital status*', '*Currently residing with husband/partner*', '*When child put to breast*', '*Anemia level*', dan '*Have mosquito bed net for sleeping (from household questionnaire)*'.

Selain itu, nilai yang ambigu seperti "Don't know" pada atribut '*Taking iron pills, sprinkles or syrup*' dihapus, karena nilai tersebut tidak termasuk dalam data kategorikal biner, yang seharusnya berisi nilai "Ya" atau "Tidak". Selanjutnya, dilakukan transformasi data dengan meng-encode atribut nominal menjadi atribut numerik. Atribut yang di-encode antara lain: '*Highest educational level*', '*Type of place of residence*', '*Wealth index combined*', '*Smokes cigarettes*', '*Had fever in last two weeks*', dan '*Taking iron pills, sprinkles or syrup*'.

2.3 Pemodelan dan Pengujian

Tahap pemodelan dan pengujian adalah fase kritis dalam pengembangan sistem rekayasa, yang mencakup penciptaan representasi matematika dan validasi selanjutnya dari model-model ini melalui pengujian eksperimental. Tahap ini memastikan bahwa sistem yang dirancang memenuhi kriteria kinerja yang ditentukan dan berfungsi sebagaimana dimaksudkan dalam aplikasi dunia nyata. Tahap pemodelan dan pengujian melibatkan pengembangan model, mengeksplorasi ruang parameter, dan melakukan simulasi untuk memahami interaksi dalam model [15]. Penerapan algoritma yang sesuai akan sangat mempengaruhi pemodelan dan pengujian. Algoritma yang digunakan adalah Naïve Bayes. Naïve Bayes memiliki beberapa keunggulan seperti akurat, cepat dan tinggi. Banyak peneliti telah melakukan klasifikasi pendapat menggunakan Naïve Bayes [16].

Selanjutnya dilakukan perhitungan menggunakan model GaussianNB. Namun untuk meningkatkan akurasi dan performa metode Naïve Bayes, penelitian ini menggabungkannya dengan *Particle Swarm Optimization* (PSO) [17].

Particle Swarm Optimization (PSO) adalah teknik optimasi yang sangat sederhana untuk menerapkan dan memvariasikan beberapa parameter [16]. PSO menggunakan mekanisme *swarm intelligence*, di mana di mana setiap partikel menunjukkan solusi potensial dan akan bergerak di ruang pencarian berdasarkan dua faktor utama, yaitu posisi terbaik yang pernah ditemukan oleh partikel (pbest) dan posisi terbaik dari keseluruhan partikel (gbest). Rumus berikut digunakan untuk menghitung pergerakan partikel [10]:

$$v[i] = w.v[i] + c_1.R.(pbest[i] - x[i]) + c_2.R.(gbest - x[i])$$

$$x[i] = x[i] + v[i]$$

Dimana $v[i]$ adalah kecepatan partikel pada iterasi ke- i , $x[i]$ adalah posisi partikel saat ini, w adalah bobot inersia, c_1 dan c_2 adalah konstanta pembelajaran, dan R adalah bilangan random.

Naïve Bayes adalah algoritma klasifikasi probabilistik yang didasarkan pada Teorema Bayes dan menganggap bahwa setiap fitur independen. Untuk menghitung probabilitas posterior, model ini menggunakan formula berikut [18]:

$$P(C|X) = P(X|C) \times P(C) / P(X)$$

PSO digunakan untuk mengoptimalkan subset fitur dan/atau parameter Naïve Bayes. Dalam hal ini, PSO memilih kombinasi fitur terbaik untuk memaksimalkan fungsi objektif berupa akurasi validasi.

2.4 Evaluasi dan Hasil

Tahap evaluasi dalam penambangan data sangat penting untuk menilai efektivitas dan akurasi model dan teknik yang diterapkan untuk mengekstrak wawasan yang bermakna dari data. Tahap ini melibatkan berbagai metodologi dan kerangka kerja untuk memastikan bahwa hasil yang diperoleh dapat diandalkan dan dapat diterapkan. Bagian berikut menguraikan aspek-aspek kunci dari tahap evaluasi dan hasilnya dalam penambangan data. *Particle Swarm Optimization* (PSO) adalah algoritma optimasi yang terinspirasi oleh perilaku sosial kawanan burung dan ikan [19]. Dalam konteks penelitian ini, PSO digunakan untuk mengoptimalkan parameter model Naïve Bayes, dengan tujuan meningkatkan akurasi deteksi anemia. PSO antara lain meningkatkan bobot atribut dari setiap atribut atau variabel yang digunakan, pemilihan atribut, dan seleksi fitur [8].

Dalam evaluasi model dilakukan menggunakan *confusion matrix* untuk menghitung akurasi, *precision*, *recall*, F1-score [20]. Selain itu, Area Under the Curver (AUC) digunakan untuk mengukur kemampuan model dalam membedakan antar kelas [21]. Hasil evaluasi ditampilkan dalam bentuk tabel dan visualisasi, seperti *classification report* dan *confusion matrix*, untuk menunjukkan performa klasifikasi pada masing-masing kategori tingkat anemia: tidak anemia, ringan, sedang, dan parah.

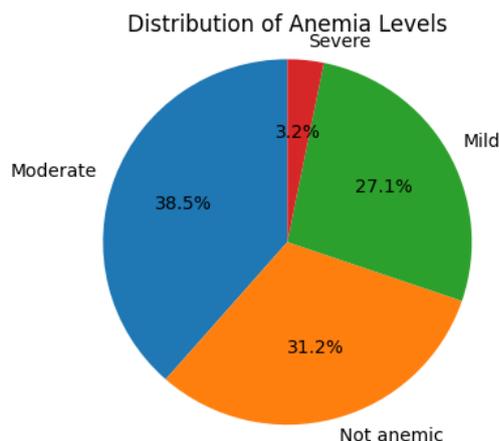
2.5 Pengembangan Aplikasi

Tahapan pengembangan aplikasi adalah serangkaian proses yang dilakukan untuk membuat aplikasi, mulai dari perencanaan hingga penerapan. Pengembangan aplikasi ini bertujuan untuk menciptakan platform berbasis web yang mendeteksi dini anemia pada anak usia 0-59 bulan.

3 HASIL DAN PEMBAHASAN

3.1 Pengumpulan Data

Gambar 2. menunjukkan visualisasi distribusi tingkat anemia pada anak usia 0-59 bulan berdasarkan dataset. Tingkat anemia dibagi menjadi empat, yaitu 31,2% *Not Anemic* (tidak anemia), 27,1% *Mild* (ringan), 38,5% *Moderate* (sedang), 3,2% *Severe* (parah). Mayoritas anak dalam data mengalami anemia tingkat sedang, yang menjadi dasar penting dalam pengembangan model deteksi. Ketidakseimbangan distribusi kelas ini menjadi tantangan tersendiri dalam proses klasifikasi, sehingga pemilihan metode dan proses optimasi model menjadi penting untuk menghindari bias terhadap kelas mayoritas.



Gambar 2. Distribusi Tingkat Anemia

3.2 Pra-Pemrosesan

Setelah dilakukan pembersihan data dan transformasi, total dataset yang digunakan yaitu 10151 dengan 11 atribut antara lain *'Age in 5-year groups'*, *'Type of place of residence'*, *'Highest educational level'*, *'Wealth index combined'*, *'Births in last five years'*, *'Age of respondent at first birth'*, *'Hemoglobin level adjusted for altitude and smoking (g/dl - 1 decimal)'*, *'Smokes cigarettes'*, *'Had fever in last two weeks'*, *'Hemoglobin level adjusted for altitude (g/dl - 1 decimal)'*, dan *'Taking iron pills, sprinkles or syrup'*. Dan terdiri dari 1 label yaitu *'Level Anemia.1'*.

Kemudian seluruh atribut dikonversi ke bentuk numerik, sehingga siap digunakan dalam proses pelatihan model klasifikasi.

3.3 Pemodelan dan Pengujian

Model yang digunakan dalam penelitian ini adalah klasifikasi menggunakan algoritma Naïve Bayes, yang diimplementasikan menggunakan bahasa pemrograman *Python* pada platform *Google Colab*. Pada tahap pemisahan data, dataset dibagi menjadi data latih 80% dan data uji 20% dengan menggunakan fungsi *train_test_split* dari pustaka *scikit-learn*. Parameter *random_seed=42* digunakan untuk memastikan pembagian data yang konsisten.

Pada tahap pemodelan, algoritma Naïve Bayes diterapkan pada data latih untuk membangun model klasifikasi anemia. Model ini kemudian dioptimasi menggunakan algoritma Particle Swarm Optimization (PSO) dengan populasi sebanyak 20 partikel. PSO merupakan metode optimasi berbasis populasi yang meniru perilaku sosial kawanan burung dalam mencari solusi terbaik. Setiap partikel merepresentasikan kandidat solusi yang bergerak di ruang pencarian parameter model, memperbarui posisi dan kecepatannya berdasarkan pengalaman terbaik partikel tersebut dan pengalaman terbaik seluruh populasi.

Dalam konteks optimasi Naïve Bayes, PSO digunakan untuk menemukan kombinasi parameter yang memaksimalkan akurasi model, seperti bobot fitur dan threshold klasifikasi. Proses ini dilakukan secara iteratif hingga konvergen pada solusi optimal. Dengan demikian, PSO membantu meningkatkan kemampuan model dalam mengenali pola data yang kompleks dan mengurangi bias terhadap kelas mayoritas.

Setelah proses optimasi, model yang telah diperbarui diuji pada data uji menggunakan metrik evaluasi seperti akurasi, presisi, dan recall. Pengujian ini memastikan bahwa model Naïve Bayes PSO yang dihasilkan mampu memberikan prediksi yang lebih akurat dan efektif dalam mendeteksi anemia dibandingkan dengan Naïve Bayes tanpa optimasi.

3.4 Evaluasi dan Hasil

Hasil evaluasi menunjukkan bahwa akurasi model Naïve Bayes mencapai 92,17%, sedangkan setelah dioptimasi dengan *Particle Swarm Optimization* (PSO), akurasi meningkat menjadi 95,71%.

Peningkatan akurasi sebesar 3,54% ini menunjukkan bahwa penerapan PSO berhasil meningkatkan performa model.

Tabel 2. Perbandingan Nilai Akurasi dan Area Under the Curve (AUC)

Nilai	Naïve Bayes	Naïve Bayes dengan PSO
Akurasi	92,17%	95,71%
AUC	0,992	0,998

Klasifikasi *report* PSO juga menunjukkan bahwa model yang dihasilkan memiliki performa yang sangat baik, dengan akurasi 96%, serta nilai precision, recall, dan F1-score yang rata-rata di atas 90% untuk masing-masing kelas. *Area Under the Curve* (AUC) yang sangat tinggi juga menunjukkan kemampuan model yang sangat baik dalam membedakan antara kelas-kelas yang ada. Namun, meskipun model menunjukkan performa yang sangat baik secara keseluruhan, masih terdapat ruang untuk perbaikan, terutama pada kelas 1, yang memiliki *precision* sedikit lebih rendah (0.87). Meskipun demikian, *recall* untuk kelas 1 tetap sangat baik dengan nilai 0.99, yang menunjukkan bahwa model sangat sensitif dalam mendeteksi kelas tersebut.

Peningkatan akurasi dan AUC setelah penerapan PSO disebabkan oleh kemampuannya dalam mengoptimalkan parameter dan bobot atribut pada model Naïve Bayes. Dengan menyesuaikan parameter tersebut secara iteratif, PSO membantu model beradaptasi lebih baik terhadap karakteristik data, termasuk distribusi kelas yang tidak seimbang. Hal ini menghasilkan klasifikasi yang lebih akurat dan sensitif, terutama dalam membedakan tingkat anemia pada data yang kompleks.

3.5 Pengembangan Aplikasi

Aplikasi ini dirancang untuk mempermudah akses deteksi anemia anak usia 0-59 bulan tanpa memerlukan alat laboratorium kompleks. Dengan mengintegrasikan algoritma Naïve Bayes yang dioptimasi menggunakan *Particle Swarm Optimization* (PSO) serta menggunakan *framework Flask*, aplikasi ini meningkatkan akurasi prediksi anemia berdasarkan data medis. Selain menyediakan antarmuka yang ramah pengguna, aplikasi ini mendukung deteksi dini dan intervensi tepat, sehingga berkontribusi pada upaya menurunkan prevalensi anemia.

Gambar 3. menunjukkan antarmuka awal aplikasi berbasis web untuk deteksi tingkat anemia. Antarmuka ini menyambut pengguna dengan pesan utama yang menekankan tujuan aplikasi dalam membantu pengguna mengetahui tingkat anemia anak dengan akurasi tinggi.

Terdapat deskripsi singkat mengenai pentingnya deteksi dini anemia pada anak, yang menjelaskan dampak anemia terhadap pertumbuhan dan perkembangan anak, serta urgensi untuk mengambil tindakan cepat guna mencegah komplikasi kesehatan yang lebih serius. Selain itu, terdapat tombol aksi dengan label "Gratis, Ayo Cek Sekarang," yang memudahkan pengguna untuk langsung memulai proses deteksi anemia melalui aplikasi ini.



Gambar 3 Tampilan Awal Aplikasi

Gambar 4. menunjukkan tampilan form deteksi anemia pada aplikasi berbasis web. Form ini dirancang untuk mengumpulkan data pengguna yang relevan sebagai langkah awal dalam proses analisis dan deteksi anemia. Data yang diminta mencakup berbagai aspek, seperti data demografis

(usia anak, jenis tempat tinggal, dan riwayat pendidikan tertinggi orang tua), kondisi sosial ekonomi (indeks pendapatan keluarga dan jumlah anak dalam lima tahun terakhir), serta data kesehatan dan riwayat medis (usia ibu saat melahirkan anak pertama, tingkat hemoglobin, kebiasaan merokok, riwayat demam dalam dua minggu terakhir, dan penggunaan obat-obatan).

Form ini bertujuan untuk mengidentifikasi faktor risiko anemia yang akan dianalisis lebih lanjut menggunakan algoritma Naïve Bayes yang dioptimasi dengan *Particle Swarm Optimization* (PSO). Setelah pengguna mengisi semua kolom yang tersedia, proses deteksi dapat dimulai dengan menekan tombol "Cek." Hasil prediksi anemia akan disajikan berdasarkan analisis data yang dimasukkan.

DETEKSI TINGKAT ANEMIA ANAK USIA 0-59 BULAN

Gambar 4. Form Deteksi Anemia

Gambar 5. menunjukkan hasil deteksi anemia dalam bentuk pop-up. Pop-up ini memberikan informasi mengenai tingkat anemia yang terdeteksi. Hasil deteksi dapat mengklasifikasikan tingkat anemia menjadi beberapa kategori, seperti tidak anemia, sedang, ringan, atau parah. Dengan demikian, pengguna dapat langsung mengetahui kondisi anemianya berdasarkan hasil yang ditampilkan.



Gambar 5. Hasil Form Deteksi Anemia

4 KESIMPULAN

Penelitian ini telah berhasil mengembangkan model klasifikasi anemia pada anak usia 0-59 bulan menggunakan algoritma Naïve Bayes yang dioptimasi dengan *Particle Swarm Optimization* (PSO). Data yang digunakan mencakup berbagai atribut terkait usia, pendidikan, status sosial ekonomi, dan kondisi medis anak serta ibu. Tahap pra-pemrosesan data melibatkan pembersihan data, penghapusan atribut yang tidak relevan, dan transformasi atribut nominal menjadi numerik. Model Naïve Bayes menunjukkan akurasi sebesar 92,17%, yang meningkat menjadi 95,71% setelah dioptimasi dengan PSO, menghasilkan peningkatan akurasi sebesar 3,54%.

Evaluasi model menunjukkan bahwa meskipun akurasi dan *Area Under Curve* (AUC) sangat tinggi, masih ada ruang untuk perbaikan, terutama pada kelas 1 yang memiliki *precision* sedikit lebih rendah meskipun *recall* tetap sangat baik. Peningkatan performa ini menggarisbawahi efektivitas PSO dalam meningkatkan akurasi prediksi anemia. Aplikasi yang dikembangkan dapat diakses dengan mudah, dirancang untuk deteksi anemia anak tanpa memerlukan alat laboratorium kompleks, dan berfungsi dengan baik di daerah dengan fasilitas terbatas. Meskipun aplikasi telah dikembangkan dan diuji secara teknis, masih ada keterbatasan dalam pengujian berbasis pengguna akhir atau di lapangan. Untuk mengetahui seberapa efektif sistem, rencana selanjutnya akan melakukan validasi oleh pengguna awam dan tenaga medis di posyandu.

5 KONTRIBUSI PENELITIAN

Kontribusi penelitian ini terhadap bidang deteksi anemia anak-anak terletak pada pengembangan model klasifikasi yang akurat menggunakan algoritma Naïve Bayes yang dioptimasi dengan PSO, yang dapat diterapkan di daerah dengan fasilitas terbatas. Penelitian ini juga memperkenalkan metode baru dalam memanfaatkan data sosial ekonomi dan medis untuk meningkatkan akurasi prediksi anemia. Dengan demikian, penelitian ini memberikan solusi praktis untuk deteksi dini anemia pada anak dan berkontribusi pada peningkatan kualitas kesehatan masyarakat.

DAFTAR PUSTAKA

- [1] R. Susanti and L. Elfianti, "Prediksi Diagnosis Anemia Melalui Metode Klasifikasi Berbasis Machine Learning," *J. Artif. Intell. Appl. (JAlA)* vol. 1, no. 1, pp. 36–43, 2024.
- [2] "Dosen FK Unpas: Kenali Anemia dan Tips Berpuasa Bagi Penderitanya." [Online]. Available: <https://www.unpas.ac.id/dosen-fk-unpas-kenali-anemia-dan-tips-berpuasa-bagi-penderitanya/>
- [3] T. A. Hindriati and N. Herawati, "Deteksi Dini dan Pencegahan Anemia Pada Ibu Hamil Trimester III di Desa Penyengat Olak Kecamatan Jambi Luar Kota," *Indones. Berdaya*, vol. 5, no. 1, pp. 141–148, 2023, doi: 10.47679/ib.2024662.
- [4] Y. Amelia, "Perbandingan Metode Machine Learning Untuk Mendeteksi Penyakit Jantung," *IDEALIS Indones. J. Inf. Syst.*, vol. 6, no. 2, pp. 220–225, 2023, doi: 10.36080/idealis.v6i2.3043.
- [5] U. A. Rheni Aprilia Ningrum, Agus Priyanto, "Implementasi Algoritma Fuzzy Tsukamoto Untuk Diagnosis Penyakit Anemia (Studi Data: Rekam Medis Pasien Ibu RSIA Bunda Arif Purwokerto)," *Infokes J. Ilm. Rekam Medis dan Inform. Kesehat.*, vol. 11, pp. 65–72, 2021, doi: <https://doi.org/10.47701/infokes.v11i2.1303>.
- [6] M. Sukron, A. Supriadi, and R. Sulton, "Optimasi Metode Naïve Bayes Menggunakan Algoritma Particle Swarm Optimization (Pso) Untuk Prediksi Penyakit Diabetes Mellitus," *COREAI J. Kecerdasan Buatan, Komputasi dan Teknol. Inf.*, vol. 2, no. 2, pp. 18–24, 2022, doi: 10.33650/coreai.v2i2.3304.
- [7] D. E. Yanti, L. Framesti, and A. Desiani, "JIP (Jurnal Informatika Polinema) PERBANDINGAN ALGORITMA C4.5 DAN SVM DALAM KLASIFIKASI PENYAKIT ANEMIA," pp. 427–434, 2022, [Online]. Available:

- <https://www.kaggle.com/datasets/biswaranjanrao/an>
- [8] N. N. K. Pada, “KOMPARASI ALGORITMA C4 . 5 DAN K- KLASIFIKASI PENYAKIT ANEMIA,” vol. 9, no. 2, pp. 110–113, 2024.
- [9] I. Arfanda, W. Ramdhan, and R. A. Yusda, “Naive Bayes Dalam Menentukan Penerima Bantuan Langsung Tunai,” *Digit. Transform. Technol.*, vol. 1, no. 1, pp. 9–16, 2021, doi: 10.47709/digitech.v1i1.1091.
- [10] W. S. Dharmawan, “I N F O R M a T I K a Dalam Prediksi Penyakit Jantung,” *J. Inform. Manaj. dan Komput.*, vol. 13, no. 2, pp. 31–41, 2021.
- [11] A. Rizky Fadilla and P. Ayu Wulandari, “Literature Review Analisis Data Kualitatif: Tahap PengumpulanData,” *Mitita J. Penelit.*, vol. 1, no. No 3, pp. 34–46, 2023.
- [12] M. E. Buchanan, “Methods of data collection,” *AORN J.*, vol. 33, no. 1, 1981, doi: 10.1016/S0001-2092(07)69400-9.
- [13] A. Rahmat, M. Syafiih, and M. Faid, “Implementasi Klasifikasi Potensi Penyakit Jantung Dengan Menggunakan Metode C4.5 Berbasis Website (Studi Kasus Kaggle.Com),” *INFOTECH J.*, vol. 9, no. 2, pp. 393–400, 2023, doi: 10.31949/infotech.v9i2.6295.
- [14] J. Cao, *Pre-processing Big Data for Business. In: E-Commerce Big Data Mining and Analytics*. Springer, Singapore, 2023. doi: https://doi.org/10.1007/978-981-99-3588-8_3.
- [15] S. W. Link, “AND PSYCHOLOGICAL MODELS,” vol. 29, no. 1, pp. 6–14, 1997.
- [16] T. D. Putra, E. Utami, and M. P. Kurniawan, “Analisis Sentimen Pemilu 2024 dengan Naive Bayes Berbasis Particle Swarm Optimization (PSO),” *Explore*, vol. 13, no. 1, pp. 1–5, 2023, doi: 10.35200/ex.v11i2.13.
- [17] Taofik Safrudin, G. Tri Pranoto, and W. Hadikristanto, “Optimasi Algoritma K- Nearest Neighbor Berbasis Particle Swarm Optimization Untuk Meningkatkan Kebutuhan Barang,” *Bull. Inf. Technol.*, vol. 4, no. 3, pp. 281–286, 2023, doi: 10.47065/bit.v4i3.724.
- [18] B. Kata and D. A. N. Frekuensi, “PENERAPAN ALGORITMA NAIVE BAYES DENGAN CHI-SQUARE UNTUK KLASIFIKASI SPAM EMAIL,” vol. 13, no. 1, 2025.
- [19] M. Ishi, J. Patil, and V. Patil, “An efficient team prediction for one day international matches using a hybrid approach of CS-PSO and machine learning algorithms,” *Array*, vol. 14, no. February, p. 100144, 2022, doi: 10.1016/j.array.2022.100144.
- [20] S. Suprihanto, I. Awaludin, M. Fadhil, and M. A. Z. Zulfikor, “Analisis Kinerja ResNet-50 dalam Klasifikasi Penyakit pada Daun Kopi Robusta,” *J. Inform.*, vol. 9, no. 2, pp. 116–122, 2022, doi: 10.31294/inf.v9i1.13049.
- [21] Y. A. Singgalen, “Analisis Performa Algoritma NBC, DT, SVM dalam Klasifikasi Data Ulasan Pengunjung Candi Borobudur Berbasis CRISP-DM,” *Build. Informatics, Technol. Sci.*, vol. 4, no. 3, 2022, doi: 10.47065/bits.v4i3.2766.