

## Penerapan Metode *Logistic Regression* dalam Sistem Prediksi Risiko Stunting Anak Berbasis Web

Rozakira Zulfa<sup>1\*</sup>, Fadli Mahyudin<sup>1</sup>, Arnawan Hasibuan<sup>2</sup>.

<sup>1</sup>Program Studi Sistem Informasi, Universitas Malikussaleh, Lhokseumawe, Indonesia

<sup>2</sup>Program Studi Teknik Elektro, Universitas Malikussaleh, Lhokseumawe, Indonesia

\*Email: rozakira.220180133@mhs.unimal.ac.id

Info Artikel	Abstrak
<p><b>Kata Kunci :</b> prediksi <i>stunting</i>, <i>logistic regression</i>, data pertumbuhan anak, <i>sistem berbasis web</i>, kesehatan anak</p> <p><b>Keywords :</b> <i>stunting prediction</i>, <i>logistic regression</i>, <i>child growth data</i>, <i>web-based system</i>, <i>child health</i></p> <p><b>Tanggal Artikel</b>                      Dikirim : 15 Juli 2025                      Direvisi : 28 Desember 2025                      Diterima : 30 Desember 2025</p>	<p><i>Stunting</i> merupakan masalah gizi kronis yang berdampak signifikan terhadap pertumbuhan dan perkembangan anak, khususnya pada periode 1000 hari pertama kehidupan. Penelitian ini mengembangkan sistem berbasis web untuk memprediksi risiko <i>stunting</i> pada anak menggunakan metode <i>Logistic Regression</i>. <i>Dataset</i> mencakup <i>fitur</i> seperti jenis kelamin, usia, berat dan panjang lahir, berat badan, panjang badan, serta status konsumsi ASI. Data diproses melalui tahapan <i>cleaning</i> dan <i>preprocessing</i> sebelum pelatihan model. <i>Logistic Regression</i> dipilih karena kemampuannya melakukan klasifikasi biner, yaitu <i>stunting</i> atau tidak. Penilaian model dilakukan dengan memanfaatkan metrik <i>accuracy</i>, <i>precision</i>, <i>recall</i>, dan AUC. Hasil eksperimen menunjukkan model mencapai <i>accuracy</i> 69,9%, <i>precision</i> 78,4%, <i>recall</i> 85,4%, dan <i>AUC</i> 0,94, sehingga dapat digunakan sebagai alat bantu untuk deteksi dini <i>stunting</i>. Sistem berbasis web ini memudahkan akses bagi tenaga kesehatan dan orang tua untuk memantau status gizi anak dan memberikan rekomendasi intervensi yang tepat. Temuan ini menunjukkan bahwa penerapan <i>Logistic Regression</i> efektif dalam mendukung pencegahan <i>stunting</i> secara lebih terarah.</p>
	<p><b>Abstract</b></p> <p><i>Stunting is a chronic nutritional problem that significantly affects children's growth and development, particularly during the first 1,000 days of life. This study develops a web-based system to predict the risk of stunting in children using the Logistic Regression method. The dataset includes features such as gender, age, birth weight and length, body weight and length, and breastfeeding status. The data were processed through cleaning and preprocessing stages before model training. Logistic Regression was chosen for its ability to perform binary classification, i.e., stunted or not stunted. The model was evaluated using accuracy, precision, recall, and AUC metrics. Experimental results show that the model achieved accuracy of 69.9%, precision of 78.4%, recall of 85.4%, and AUC of 0.94, indicating its potential as a tool for early detection of stunting. This web-based system facilitates access for healthcare workers and parents to monitor children's nutritional status and provide appropriate intervention recommendations. The findings demonstrate that Logistic Regression is effective in supporting more targeted stunting prevention efforts.</i></p>

## 1. PENDAHULUAN

*Stunting* adalah kondisi ketika seorang anak memiliki tinggi badan lebih rendah dari standar usianya, yang merupakan indikator utama dari malnutrisi kronis. Kondisi ini tidak hanya berdampak pada pertumbuhan fisik, tetapi juga berpengaruh terhadap perkembangan kognitif, prestasi pendidikan, dan risiko penyakit kronis di masa dewasa. *Stunting* masih menjadi masalah kesehatan masyarakat global, dengan prevalensi tinggi terutama di negara-negara berkembang [1].

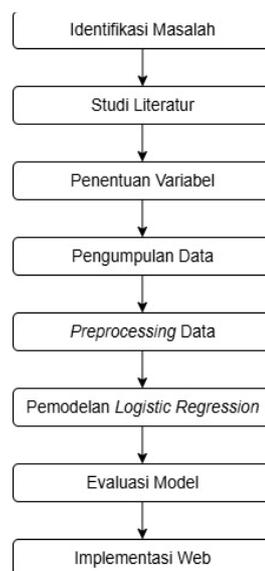
Di Indonesia, berdasarkan data Survei Status Gizi Indonesia (SSGI) yang dirilis oleh Kementerian Kesehatan, prevalensi *stunting* pada anak balita tercatat sebesar 21,5% pada tahun 2023 dan menurun menjadi 19,8% pada tahun 2024 [2]. Meskipun terjadi penurunan, angka ini menunjukkan bahwa upaya intervensi yang lebih terarah dan berkelanjutan tetap sangat diperlukan untuk mencapai target nasional. Faktor penyebab *stunting* sangat kompleks, mencakup kekurangan asupan gizi, pola makan yang tidak seimbang, sanitasi yang buruk, serta infeksi yang terjadi secara berulang [3]. Oleh karena itu, pendekatan berbasis data yang mampu mendeteksi risiko *stunting* secara dini agar intervensi gizi dapat dilakukan secara tepat waktu dan efektif.

Berbagai penelitian sebelumnya telah penerapan *machine learning* untuk memprediksi risiko *stunting* pada anak. Salah satu studi menggunakan Naive Bayes untuk mengklasifikasikan *stunting* pada balita di Semarang, menunjukkan akurasi tinggi dan potensi sebagai alat deteksi dini [4]. Pendekatan lain menggunakan *Bayesian Network* terintegrasi GIS untuk mengidentifikasi faktor risiko dan memvisualisasikan persebarannya, dengan temuan bahwa antropometri, pemberian ASI, dan pola asuh ibu berpengaruh signifikan [5]. Studi perbandingan algoritma menilai *Logistic Regression*, SVM, CNN, MLP, Naive Bayes, KNN, dan *Random Forest*. Hasil menunjukkan *Random Forest* unggul dari sisi akurasi, KNN efektif mendeteksi kasus *stunting*, sementara *Logistic Regression* stabil, efisien, mudah diinterpretasi, dan cocok untuk sistem web [6] [7]. *Logistic Regression* juga digunakan untuk skrining risiko *stunting*, dengan performa stabil dan interpretabilitas tinggi [8].

Di bidang kesehatan lain, metode *Logistic Regression* efektif untuk deteksi penyakit jantung dan diabetes, menekankan pentingnya normalisasi data [9], [10], [11]. Berdasarkan kajian ini, *machine learning*, khususnya *Logistic Regression*, berpotensi mendukung prediksi kesehatan anak. Namun, sebagian penelitian belum mengintegrasikan model ke sistem berbasis web, sehingga penelitian ini mengusulkan implementasi *Logistic Regression* dalam sistem prediksi *stunting* berbasis web untuk deteksi dini yang efisien dan mudah diakses.

Berdasarkan latar belakang tersebut, penelitian ini bertujuan untuk membangun sistem prediksi risiko *stunting* berbasis web dengan menerapkan algoritma *logistic regression*. Sistem ini dirancang menggunakan data pertumbuhan anak seperti jenis kelamin, usia, berat lahir, panjang lahir, berat badan, panjang badan, dan status konsumsi ASI sebagai variabel input. Diharapkan, sistem ini dapat menjadi alat bantu yang efektif dalam deteksi dini risiko *stunting*, serta mendukung upaya pencegahan yang lebih luas di Indonesia.

## 2. METODE PENELITIAN



Gambar 1. Tahapan Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan pemodelan *machine learning*, khususnya algoritma *Logistic Regression* karena mampu memodelkan hubungan antara variabel input dan output dalam bentuk klasifikasi biner, seperti pada kasus prediksi risiko *stunting* (berisiko/tidak berisiko). Metode regresi sebagai dasar pemodelan prediktif telah banyak diterapkan dalam berbagai bidang. Salah satu implementasinya adalah penggunaan *double log regression* untuk memprediksi konsumsi energi listrik berdasarkan data historis, yang menunjukkan tingkat akurasi yang tinggi dalam pemodelan estimasi [12]. Penelitian ini bersifat eksperimental karena melibatkan proses pelatihan model menggunakan data yang telah diproses, evaluasi performa model melalui metrik klasifikasi, serta implementasi sistem prediksi dalam bentuk aplikasi berbasis web. Tahapan penelitian dijelaskan sebagai berikut :

## 2.1 Identifikasi Masalah

*Stunting* merupakan salah satu masalah kesehatan utama di Indonesia yang berdampak jangka panjang terhadap pertumbuhan fisik dan kecerdasan anak. Oleh karena itu, penelitian ini bertujuan membangun sistem prediksi berbasis web yang dapat mendeteksi potensi *stunting* sejak dini guna mendukung proses pencegahan oleh tenaga medis dan orang tua. Klasifikasi risiko *stunting* secara dini dapat mendukung pencegahan jangka panjang terhadap dampak buruk *stunting* [9].

## 2.2 Studi Literatur

Studi literatur dilakukan terhadap berbagai penelitian terdahulu yang relevan untuk mendukung pengembangan sistem prediksi *stunting*. Beberapa penelitian sebelumnya menunjukkan bahwa *logistic regression* merupakan metode yang efektif untuk klasifikasi risiko *stunting* berdasarkan data antropometri anak [9]. Model ini dipilih karena memiliki keunggulan dalam interpretasi hasil serta efektivitas dalam memodelkan hubungan antara variabel prediktor dan keluaran biner.

Metode *logistic regression* juga telah digunakan secara luas dalam bidang kesehatan. Penelitian sebelumnya menunjukkan performa yang baik dalam klasifikasi biner pada data medis, seperti deteksi penyakit jantung [10]. Selain itu, pemilihan fitur (*feature selection*) juga menjadi aspek penting dalam membangun model prediktif yang andal. Salah satu penelitian menekankan bahwa fitur-fitur seperti status ASI, berat badan lahir, dan kondisi ibu selama kehamilan merupakan indikator yang signifikan dalam prediksi risiko *stunting* [5].

Implementasi *logistic regression* dalam klasifikasi penyakit kronis juga menunjukkan hasil yang menjanjikan. Dalam sebuah studi, tahapan normalisasi data dan pemilihan fitur yang tepat terbukti mampu meningkatkan performa model klasifikasi untuk penyakit diabetes. Studi tersebut juga menyimpulkan bahwa metrik *recall* merupakan indikator evaluasi yang paling relevan untuk kasus penyakit kronis, karena mampu meminimalkan kesalahan prediksi negatif yang berpotensi berdampak serius [11].

Perbandingan beberapa algoritma *machine learning*, seperti *Logistic Regression*, *SVM*, *CNN*, dan *MLP*, juga telah dilakukan dalam konteks deteksi status gizi anak. Hasilnya menunjukkan bahwa *Logistic Regression* unggul dalam hal interpretabilitas dan efisiensi komputasi, meskipun akurasinya lebih rendah dibandingkan dengan *SVM* dan *CNN*. Metode ini dinilai lebih sesuai untuk perangkat dengan sumber daya terbatas seperti smartphone, dan cocok untuk skenario prediksi cepat berbasis data tabular [6]. Sebagai pembanding, penelitian lain menggunakan algoritma *Random Forest* untuk prediksi *stunting* di Aceh Timur. Meskipun akurasinya hanya mencapai 58,6%, model tersebut dinilai mampu menangani multikolinearitas dan kompleksitas data, serta memberikan fleksibilitas dalam mengelola banyak variabel prediktor secara bersamaan [13].

Studi lainnya mengembangkan model prediksi diagnostik untuk mendeteksi *stunting* pada anak di Malawi dengan membandingkan tujuh metode seleksi variabel, termasuk *backward*, *stepwise*, *LASSO*, dan *random forest*. Penelitian tersebut menekankan bahwa pemilihan fitur yang tepat sangat memengaruhi akurasi model, dengan faktor-faktor seperti usia anak, berat lahir, dan indeks kekayaan rumah tangga sebagai prediktor yang paling konsisten muncul dalam semua metode [8].

## 2.3 Menentukan Variabel Penelitian

Penelitian ini menggunakan beberapa variabel independen yang terdiri dari: jenis kelamin, usia, berat lahir, panjang lahir, berat badan, panjang badan, dan status pemberian ASI. Variabel-variabel tersebut dipilih berdasarkan relevansinya terhadap risiko *stunting*, sebagaimana telah digunakan dalam penelitian sebelumnya yang melakukan klasifikasi status gizi balita dengan tingkat akurasi mencapai 85,33% [4]. Adapun variabel dependen dalam penelitian ini adalah status *stunting* anak, yang diklasifikasikan secara biner (berisiko/tidak berisiko), sesuai dengan karakteristik algoritma *logistic regression* yang digunakan dalam pemodelan.

## 2.4 Pengumpulan Data

Dataset yang digunakan dalam penelitian ini diperoleh dari platform Kaggle, yang menyediakan data terbuka mengenai kesehatan anak dan status *stunting* [14]. Data tersebut tergolong sebagai data sekunder karena dikumpulkan dan dipublikasikan oleh pihak lain, bukan secara langsung oleh peneliti.

Sebelum digunakan dalam proses pemodelan, data divalidasi melalui beberapa tahapan, yaitu pemeriksaan kelengkapan atribut, penghapusan entri duplikat, serta penanganan *missing value*. Langkah ini bertujuan untuk memastikan kualitas data agar model prediksi yang dibangun memiliki akurasi dan reliabilitas yang baik.

## 2.5 Pemodelan dan Analisis Data

### 2.5.1 Preprocessing Data

Tahapan *preprocessing* bertujuan menyiapkan data sebelum pelatihan model. Langkah pertama adalah normalisasi menggunakan *Z-Score* dengan bantuan *StandardScore* dari *Pustaka scikit-learn*. Metode ini mengubah skala fitur numerik agar memiliki nilai rata-rata 0 dan standar deviasi 1. Rumus yang Digunakan adalah :

$$X' = \frac{X - \mu}{\sigma}$$

Selain normalisasi, dilakukan pula deteksi dan penghapusan *outlier* menggunakan pendekatan statistik seperti *interquartile range (IQR)* dan *Z-score*, agar model tidak terpengaruh oleh nilai-nilai ekstrem. Pendekatan *preprocessing* seperti ini juga telah diterapkan dalam penelitian sebelumnya pada data balita untuk prediksi *stunting* menggunakan algoritma *Naive Bayes* [11]. Setelah data dibersihkan dan dinormalisasi, dataset dibagi menjadi 80% data latih (*training data*) dan 20% data uji (*testing data*). Data latih digunakan untuk melatih model, sedangkan data uji digunakan untuk mengukur performa model terhadap data yang belum dikenali sebelumnya.

### 2.5.2 Model Logistic Regression

*Logistic regression* merupakan algoritma klasifikasi biner yang digunakan untuk memprediksi probabilitas terjadinya suatu kejadian dari dua kemungkinan. Fungsi aktivasi yang digunakan dalam model ini adalah fungsi sigmoid, yang didefinisikan sebagai:

$$h_{\theta}(x) = \frac{1}{1 + e^{-\theta^T x}}$$

Nilai  $h_{\theta}(x)$  berada pada rentang 0 hingga 1, dan mempresentasikan probabilitas kelas positif. Model mempelajari parameter bobot  $\theta$  dengan meminimalkan fungsi biaya berikut :

$$J(\theta) = -\frac{1}{m} \sum_{i=1}^m [y^{(i)} \log h_{\theta}(x^{(i)}) + (1 - y^{(i)}) \log (1 - h_{\theta}(x^{(i)}))]$$

Di mana  $m$  adalah jumlah data,  $x^{(i)}$  adalah fitur pada data ke  $i$ , dan  $y^{(i)}$  adalah label aktual pada data ke-  $i$ . Pendekatan ini telah terbukti efektif dalam berbagai kasus klasifikasi biner, termasuk dalam konteks prediksi penyakit kronis berbasis data medis [11].

### 2.5.3 Evaluasi Model

Setelah model dibangun, tahap selanjutnya adalah melakukan evaluasi kinerja model menggunakan sejumlah metrik performa yang umum digunakan dalam klasifikasi biner, yaitu:

1. Akurasi (Accuracy): proporsi jumlah prediksi yang benar terhadap total keseluruhan data.
2. Presisi (Precision): proporsi prediksi positif yang benar-benar merupakan kelas positif.
3. Recall (Sensitivity): proporsi data positif yang berhasil terdeteksi dengan benar oleh model.
4. F1-score: rata-rata harmonis antara presisi dan recall, yang berguna dalam kondisi data tidak seimbang.

Evaluasi model prediksi sangat penting, terutama dalam konteks kesehatan publik, untuk meminimalkan kesalahan deteksi yang dapat berdampak besar terhadap pengambilan keputusan. Penggunaan metrik seperti *F1-score* dan *recall* menjadi krusial dalam pengembangan sistem prediksi stunting yang sensitif dan akurat, sebagaimana dijelaskan dalam penelitian sebelumnya [8].

### 3. HASIL DAN PEMBAHASAN

#### 3.1 Hasil Pengumpulan Data

Data yang digunakan dalam penelitian ini diperoleh dari dataset terbuka yang tersedia di platform Kaggle, dengan total sebanyak 10.000 entri [14]. Dataset ini terdiri dari tujuh fitur prediktor serta satu label target, yaitu status *stunting*. Sebelum digunakan dalam pemodelan, data terlebih dahulu melalui proses pembersihan, termasuk penghapusan data duplikat dan penanganan *missing values* untuk memastikan kualitas data yang optimal.

Contoh data mentah yang digunakan dalam penelitian ditampilkan pada Tabel 1, yang memuat karakteristik anak berdasarkan fitur seperti usia, berat lahir, panjang lahir, berat badan, panjang badan, status pemberian ASI, serta status stunting.

Tabel 1. Data Mentah

No	Gender	Age (week)	Birth Weight (kg)	Birth Length (cm)	Body Weight (kg)	Body Length (cm)	Breastfeeding	Stunting
1	Male	17	3	49	10	72.2	No	No
2	Female	11	2.9	49	2.9	65	Yes	Yes
3	Male	16	2.9	49	8.5	72.2	No	Yes
4	Male	31	2.8	49	6.4	63	No	Yes
5	Male	15	3.1	49	10.5	49	No	Yes
...	...	...	...	...	...	...	...	...
10000	Female	10	3	49	7.7	80	Yes	Yes

Setelah melalui tahapan *preprocessing*, data dikonversi ke dalam format numerik untuk memudahkan proses pelatihan model *machine learning*. Proses ini melibatkan normalisasi, transformasi nilai kategorikal menjadi numerik, serta pembersihan data. Contoh hasil data setelah proses *preprocessing* ditunjukkan pada Tabel 2.

Tabel 2. Data setelah tahapan *preprocessing*

No	Gender	Age (week)	Birth Weight (kg)	Birth Length (cm)	Body Weight (kg)	Body Length (cm)	Breastfeeding	Stunting
1	1	17	3	49	10	72.2	0	0
2	0	11	2.9	49	2.9	65	1	1
3	1	16	2.9	49	8.5	72.2	0	1
4	1	31	2.8	49	6.4	63	0	1
5	1	15	3.1	49	10.5	49	0	1
...	...	...	...	...	...	...	...	...
10000	0	10	3	49	7.7	80	1	1

#### 3.2 Hasil Pemodelan dan Analisis

Tahapan pemodelan dilakukan melalui empat tahap utama, yaitu: pra-pemrosesan data, pelatihan model, evaluasi performa model, dan analisis hasil.

##### 3.2.1 Pra-pemrosesan Data

Tahapan *preprocessing* bertujuan untuk menyiapkan data sebelum proses pelatihan model dilakukan. Langkah awal adalah normalisasi menggunakan metode *Z-Score* dengan bantuan fungsi *StandardScaler* dari pustaka *scikit-learn*. Metode ini

mengubah skala fitur numerik agar memiliki nilai rata-rata 0 dan standar deviasi 1. Rumus yang digunakan ditunjukkan pada Persamaan (1):

$$X' = \frac{X - \mu}{\sigma} \quad (1)$$

Selain normalisasi, dilakukan pula deteksi dan penghapusan outlier menggunakan pendekatan statistik seperti interquartile range (IQR) dan *Z-score*, agar model tidak terpengaruh oleh nilai-nilai ekstrem yang dapat menurunkan performa prediksi. Deteksi *outlier* ini membantu memastikan distribusi data lebih representatif dan stabil, sehingga model *Logistic Regression* dapat belajar dari pola yang sebenarnya. Pendekatan preprocessing seperti ini juga diterapkan dalam penelitian sebelumnya yang menggunakan metode serupa pada data balita untuk prediksi *stunting* [4], menunjukkan bahwa langkah-langkah pembersihan data merupakan tahap penting dalam pengembangan model prediksi kesehatan anak.

Setelah data dibersihkan, dinormalisasi, dan *outlier* dihapus, dataset dibagi menjadi dua bagian utama yaitu 80% data latih (Data 1–8.000) dan 20% data uji (Data 8.001–10.000). Data latih digunakan untuk membangun dan melatih model, sedangkan data uji digunakan untuk menilai performa model terhadap data yang belum dikenali sebelumnya, sehingga dapat mengukur kemampuan generalisasi model secara objektif. Pemisahan ini mengikuti praktik standar dalam *machine learning* untuk mencegah *overfitting* dan memastikan model dapat bekerja dengan baik pada data baru. Penyajian data latih dan data uji dapat dilihat pada Tabel 3 dan Tabel 4 berikut, yang menampilkan cuplikan awal, beberapa baris tengah, dan data terakhir dari dataset untuk memberikan gambaran distribusi data secara keseluruhan.

Tabel 3. Data Latih

No	Gender	Age (week)	Birth Weight (kg)	Birth Length (cm)	Body Weight (kg)	Body Length (cm)	Breastfeeding	Stunting
1	1	17	3	49	10	72.2	0	0
2	0	11	2.9	49	2.9	65	1	1
3	1	16	2.9	49	8.5	72.2	0	1
4	1	31	2.8	49	6.4	63	0	1
5	1	15	3.1	49	10.5	49	0	1
...	...	...	...	...	...	...	...	...
8000	1	9	2.8	49	6.4	76	1	1

Tabel 4. Data Uji

No	Gender	Age (week)	Birth Weight (kg)	Birth Length (cm)	Body Weight (kg)	Body Length (cm)	Breastfeeding	Stunting
8001	1	7	2.3	50	9	92.7	1	0
8002	1	13	3	49	10	49	0	1
8003	0	13	2.8	49	10	63	0	1
8004	0	10	2.8	49	8.4	49	1	1
8005	0	12	2.8	49	2.9	65	1	1
...	...	...	...	...	...	...	...	...
10000	0	10	3	49	7.7	80	1	1

Setelah data dibersihkan, nilai rata-rata (*mean*) dan standar deviasi (*standard deviation*) dari masing-masing fitur sebelum normalisasi dihitung. Perhitungan statistik deskriptif ini penting untuk memahami karakteristik dasar dataset, seperti distribusi, sebaran, dan variasi masing-masing fitur, serta memberikan gambaran awal mengenai kualitas dan konsistensi data. Informasi ini menjadi dasar utama dalam proses normalisasi menggunakan metode *Z-Score*, di mana setiap nilai fitur dikonversi menjadi skor standar dengan cara mengurangi mean dan membaginya dengan standar deviasi. Normalisasi dengan *Z-Score* memastikan semua fitur memiliki skala yang seragam, sehingga fitur dengan rentang nilai besar tidak mendominasi proses pelatihan model dibandingkan fitur dengan rentang kecil.

Pendekatan ini juga membantu mempercepat konvergensi algoritma *Logistic Regression*, meningkatkan stabilitas model, serta membuat interpretasi koefisien menjadi lebih konsisten. Selain itu, perhitungan statistik awal ini memberikan panduan untuk mendeteksi ketidakwajaran data atau bias pada dataset. Sehingga langkah *preprocessing* dapat dilakukan lebih efektif dan hasil evaluasi model menjadi lebih andal. Nilai statistik deskriptif tiap fitur ditampilkan pada Tabel 5 berikut, yang mencakup *mean* dan *standar deviasi* untuk masing-masing variabel input, memberikan gambaran menyeluruh tentang kondisi dataset sebelum normalisasi.

**Tabel 5. Rata-rata dan standar deviasi masing-masing fitur**

Fitur	Mean	Standar Deviasi
Gender	0.615	0.486595314
Age	14.406	8.109202427
Birth Weight	2.75435	0.294467447
Birth Length	49.134	0.443896384
Body Weight	7.56225	20.7284701
Body Length	68.9003	8.864038578
Breastfeeding	0.4145	0.488489736

Setelah dilakukan normalisasi, setiap nilai fitur telah ditransformasi ke dalam skala yang seragam dengan rata-rata mendekati 0 dan standar deviasi mendekati 1. Contoh hasil normalisasi ditunjukkan pada Tabel 6.

**Tabel 6. Hasil normalisasi data**

Gender	Age	Birth Weight	Birth Length	Body Weight	Body Length	Breastfeeding
-1.2638839	-0.17338	-0.1845704	-0.30187224	-0.0657187	1.252217023	-0.84853369
0.79121189	-0.66665	-2.5617432	-0.30187224	-0.2249201	-0.665644671	1.198592226

### 3.2.2 Pelatihan Model Logistic Regression

*Logistic Regression* merupakan algoritma klasifikasi biner yang digunakan untuk memprediksi probabilitas kejadian dari dua kemungkinan. Fungsi aktivasi yang digunakan adalah fungsi sigmoid, yang dinyatakan dalam Persamaan (2):

$$h\theta(x) = \frac{1}{1 + e^{-\theta^T x}} \quad (2)$$

Model mempelajari parameter bobot  $\theta$  dengan meminimalkan fungsi biaya berikut:

$$J(\theta) = -\frac{1}{m} \sum_{i=1}^m [y^{(i)} \log h_{\theta}(x^{(i)}) + (1 - y^{(i)}) \log (1 - h_{\theta}(x^{(i)}))] \quad (3)$$

Pendekatan ini telah terbukti efektif dalam berbagai kasus klasifikasi biner, termasuk dalam bidang medis [10], [15]. Salah satu penelitian menunjukkan bahwa algoritma ini mampu mendeteksi gejala stres mahasiswa dengan akurasi hingga 95% [15].

Proses pelatihan model menghasilkan dua keluaran utama yaitu nilai probabilitas dan hasil klasifikasi biner (prediksi). Prediksi dengan nilai 1 menunjukkan bahwa model mengindikasikan anak mengalami stunting, sedangkan nilai 0 menunjukkan sebaliknya. Contoh hasil pelatihan model ditunjukkan pada Tabel 7, yang memperlihatkan transformasi nilai fitur yang telah dinormalisasi menjadi probabilitas dan keputusan klasifikasi akhir.

**Tabel 7. Probabilitas dan prediksi model**

Hitung Normalisasi (z)	Probabilisas	Prediksi
-3.170837832	0.040278016	0
-2.297175065	0.091357191	0
0.03491131	0.491273059	0
-2.292098535	0.091779475	0

Untuk menilai akurasi model dalam mengklasifikasikan status stunting, dilakukan perbandingan antara label aktual dari dataset dan hasil prediksi model. Hasil evaluasi ini disajikan pada Tabel 8, yang mencakup kategori *True Positive* (TP), *False Positive* (FP), dan *False Negative* (FN), sehingga memberikan gambaran jelas tentang kinerja model dalam mendeteksi kasus *stunting*. Dari tabel tersebut, dapat dilihat bahwa mayoritas kasus *stunting* berhasil diklasifikasikan dengan benar oleh model, meskipun masih terdapat beberapa kesalahan klasifikasi yang ditunjukkan oleh nilai FP dan FN, yang menjadi pertimbangan penting dalam konteks evaluasi dan perbaikan model lebih lanjut.

**Tabel 8. Label aktual dan prediksi model**

<i>Label Aktual</i>	<i>Prediksi Model</i>	<i>Keterangan</i>
<i>Stunting</i>	<i>Stunting</i>	<i>Benar (TP)</i>
<i>Stunting</i>	<i>Stunting</i>	<i>Benar (TP)</i>
<i>Stunting</i>	<i>Stunting</i>	<i>Benar (TP)</i>
<i>Tidak Stunting</i>	<i>Stunting</i>	<i>Salah (FP)</i>
<i>Stunting</i>	<i>Stunting</i>	<i>Benar (TP)</i>
<i>Stunting</i>	<i>Tidak Stunting</i>	<i>Salah (FN)</i>
<i>Stunting</i>	<i>Stunting</i>	<i>Benar (TP)</i>

**Catatan:**

*TP* : *True Positive*,  
*FP* : *False Positive*  
*FN* : *False Negative*

Dari tabel tersebut dapat dilihat bahwa mayoritas kasus *stunting* berhasil diklasifikasikan dengan benar oleh model. Namun, masih terdapat beberapa kesalahan klasifikasi yang ditunjukkan oleh nilai FP dan FN, yang penting untuk diperhatikan dalam konteks evaluasi lebih lanjut.

**3.2.3 Evaluasi Model**

Evaluasi kinerja model dilakukan menggunakan empat metrik utama untuk klasifikasi biner, yaitu *accuracy*, *precision*, *recall*, dan *F1-score*. Rumus yang digunakan disajikan pada Persamaan (4):

$$Precision = \frac{TP}{TP+FP} \quad Recall = \frac{TP}{TP+FN} \quad F1 - Score = \frac{2 \cdot Precision \cdot Recall}{Precision+Recall} \quad (4)$$

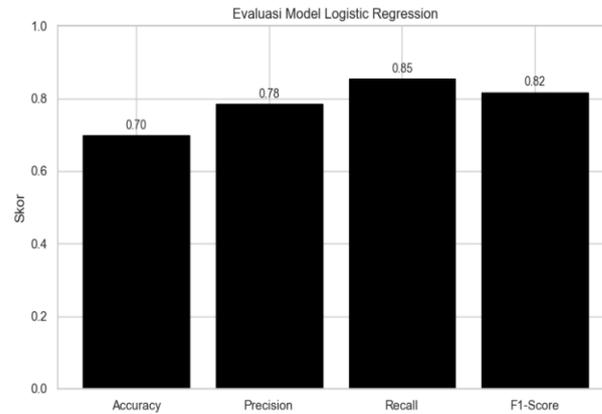
Dalam konteks prediksi risiko *stunting*, metrik evaluasi ini penting untuk mengukur sejauh mana model dapat mengenali kasus anak yang benar-benar mengalami *stunting* tanpa banyak melakukan kesalahan klasifikasi. Studi sebelumnya menunjukkan bahwa performa metrik sangat bervariasi antar algoritma. Misalnya, model *Random Forest* memiliki nilai *F1-score* tertinggi sebesar 0.922, sedangkan *K-Nearest Neighbor (KNN)* unggul dalam *recall* dengan nilai 0.967, menandakan pentingnya evaluasi holistik dalam pemilihan model [7]. Evaluasi terhadap model *Logistic Regression* pada penelitian ini menghasilkan nilai performa sebagaimana dirangkum pada Tabel 7.

**Tabel 9. Hasil evaluasi model**

<i>Metriks</i>	<i>Hasil</i>
<i>TP (True Positive)</i>	1348
<i>FP (False Positive)</i>	372
<i>TN (True Negative)</i>	50
<i>FN (False Negative)</i>	230
<i>Accuracy</i>	69.9%
<i>Precision</i>	78.4%
<i>Recall</i>	85.4 %
<i>F1-Score</i>	81.7 %

Berdasarkan hasil tersebut, model menunjukkan performa yang cukup baik. *Recall* yang tinggi (85,4%) menunjukkan bahwa model mampu mengidentifikasi sebagian besar kasus *stunting* dengan benar. Hal ini penting dalam konteks kesehatan masyarakat, karena kesalahan tipe *False Negative* (FN) dapat menghambat upaya intervensi dini. Nilai *F1-score* sebesar 81,7%

menunjukkan keseimbangan yang baik antara *precision* dan *recall*, sedangkan nilai *accuracy* sebesar 69,9% menunjukkan bahwa sekitar 70% prediksi model sesuai dengan label aktual. Secara keseluruhan, model *Logistic Regression* layak digunakan sebagai baseline dalam sistem deteksi dini stunting. Meskipun akurasi belum optimal, kemampuan model dalam menangkap kasus positif tetap menjadi nilai unggul.



**Gambar 2. Grafik Evaluasi Model Logistic Regression**

Gambar ini menampilkan empat metrik utama *Accuracy*, *Precision*, *Recall*, dan *F1-Score*. Terlihat bahwa *recall* memiliki nilai tertinggi, memperkuat peran model dalam menghindari kesalahan pendeteksian kasus stunting. *Precision* berada pada angka 78%, dan *F1-score* sebesar 82% memperlihatkan kestabilan model dalam klasifikasi dua kelas.



**Gambar 3. Grafik Confusion Matrix Model Logistic Regression**

Gambar ini memperlihatkan hasil klasifikasi model dalam bentuk confusion matrix. Model berhasil mengklasifikasikan 1348 anak stunting secara benar (*True Positive*), serta 50 anak tidak stunting secara benar (*True Negative*). Namun, terdapat 372 kasus *False Positive* dan 230 *False Negative*. Meskipun masih ada kesalahan, proporsi TP yang tinggi menjadikan model ini cukup andal dalam konteks deteksi awal, di mana mengurangi kesalahan FN merupakan prioritas utama.

### 3.3 Implementasi Aplikasi

Aplikasi prediksi risiko stunting dalam penelitian ini dikembangkan menggunakan bahasa pemrograman Python dengan framework Flask sebagai backend web. Flask dipilih karena ringan dan fleksibel dalam membangun sistem berbasis data science. Aplikasi ini juga memanfaatkan pustaka *Pandas* dan *NumPy* untuk manipulasi data, serta *scikit-learn* untuk proses pelatihan model *Logistic Regression*. Model yang telah dilatih disimpan dalam format *.pkl* menggunakan modul *pickle*, sehingga dapat digunakan kembali tanpa perlu dilakukan pelatihan ulang setiap kali aplikasi dijalankan. Implementasi sistem diawali dengan tahapan *preprocessing* data, yang mencakup konversi fitur kategorikal menjadi numerik menggunakan *LabelEncoder*, penanganan nilai kosong dengan pengisian nilai rata-rata (*mean imputation*), serta transformasi label target menjadi nilai biner (1 untuk “stunting” dan 0 untuk “tidak stunting”), sesuai dengan karakteristik *Logistic Regression*. Fitur numerik kemudian

dinormalisasi menggunakan StandardScaler agar memiliki distribusi data yang seragam. Setelah data diproses, dataset dibagi menjadi 80% untuk data pelatihan dan 20% untuk data pengujian, mengikuti praktik umum dalam pengembangan model prediktif.

Struktur aplikasi dirancang secara modular dengan tiga rute utama, yaitu halaman utama (/), halaman formulir input (/form), dan halaman hasil prediksi (/result). Setelah pengguna memasukkan data anak pada halaman input, sistem melakukan preprocessing otomatis lalu mengirimkan data tersebut ke model prediksi. Output dari model ditampilkan dalam bentuk sederhana, yakni “Stunting” atau “Tidak Stunting”. Aplikasi ini juga telah dilengkapi dengan validasi input dan umpan balik kesalahan apabila format data yang dimasukkan tidak sesuai.

Prinsip desain aplikasi ini mengacu pada praktik dalam pengembangan sistem informasi berbasis web yang efisien dan modular. Studi sebelumnya menunjukkan bahwa sistem digital berbasis web dengan arsitektur terpisah antara frontend dan backend dapat meningkatkan efisiensi pemrosesan dan interaksi pengguna [16]. Selain itu, integrasi model prediktif ke dalam sistem informasi web dinilai mampu mempercepat proses pengambilan keputusan serta menyajikan informasi secara real-time [17].

Untuk memperjelas interaksi antara pengguna dan sistem, *use case diagram* disertakan pada Gambar 4. Diagram ini menggambarkan fungsi-fungsi utama yang dapat diakses oleh pengguna, mulai dari input data hingga tampilan hasil prediksi.



Gambar 4. Use Case Diagram

*Use case* diagram di atas menjelaskan alur interaksi antara pengguna dengan sistem prediksi risiko stunting berbasis web. Dalam konteks ini, pengguna berperan sebagai aktor utama yang berinteraksi langsung dengan aplikasi. Terdapat tiga aktivitas utama yang dapat dilakukan oleh pengguna, yaitu mengakses halaman utama, mengisi formulir data anak, dan melihat hasil prediksi. Pada tahap awal, pengguna akan diarahkan ke halaman utama yang berfungsi sebagai pintu masuk menuju fitur-fitur sistem.

Selanjutnya, pengguna dapat menuju halaman formulir untuk memasukkan data anak yang meliputi usia, jenis kelamin, berat badan, panjang badan, berat lahir, panjang lahir, serta status pemberian ASI. Setelah data diinput dan dikirimkan, sistem akan memproses informasi tersebut menggunakan model *Logistic Regression* yang telah dilatih sebelumnya. Hasil prediksi kemudian ditampilkan secara real-time dalam bentuk sederhana, yaitu klasifikasi apakah anak tersebut berisiko stunting atau tidak. Alur ini dirancang agar mudah dipahami dan dioperasikan, sehingga dapat diakses oleh pengguna dari berbagai kalangan, khususnya tenaga kesehatan dan orang tua yang ingin mengetahui status pertumbuhan anak secara cepat dan efisien. Sementara itu, visualisasi tampilan aplikasi diberikan pada Gambar 5 hingga Gambar 7, yang menunjukkan alur penggunaan sistem dari halaman utama, formulir input, hingga hasil prediksi akhir.

Selain menggambarkan alur fungsional sistem, *use case* diagram ini juga menegaskan bahwa aplikasi dirancang dengan pendekatan *user-centered*, di mana setiap fitur disesuaikan dengan kebutuhan dan kemampuan pengguna. Interaksi yang sederhana tanpa proses yang berbelit-belit bertujuan untuk meminimalkan kesalahan input serta mempercepat proses prediksi. Dengan alur penggunaan yang jelas dan terstruktur, sistem ini diharapkan dapat menjadi alat bantu pendukung keputusan yang efektif dalam deteksi dini risiko *stunting*, serta mendukung upaya pencegahan melalui pemantauan pertumbuhan anak secara lebih sistematis dan berbasis data. Selain itu, *use case* diagram ini juga memberikan gambaran batasan dan tanggung jawab sistem secara jelas, di mana seluruh proses pengolahan data dan penentuan hasil prediksi sepenuhnya dilakukan oleh sistem tanpa melibatkan pengguna pada tahap perhitungan. Pengguna hanya berfokus pada penyediaan data yang dibutuhkan dan menerima hasil akhir berupa informasi risiko *stunting*. Pemisahan peran ini bertujuan untuk meningkatkan keandalan sistem serta mengurangi potensi kesalahan interpretasi oleh pengguna. Dengan demikian, aplikasi tidak hanya berfungsi sebagai media input dan output, tetapi juga sebagai sistem pendukung keputusan yang konsisten, objektif, dan dapat diandalkan dalam membantu pemantauan status gizi anak secara berkelanjutan.



Gambar 5. Halaman utama

Tampilan ini merupakan halaman pertama (beranda) yang diakses oleh pengguna saat membuka aplikasi Sistem Prediksi *Stunting* pada Anak. Halaman ini berfungsi sebagai pengantar yang memberikan gambaran umum mengenai tujuan dan fungsi utama sistem, yaitu membantu pengguna dalam memahami konsep *stunting* serta pentingnya deteksi dini terhadap risiko *stunting* pada anak. Informasi yang ditampilkan disajikan secara ringkas dan informatif agar pengguna, baik dari kalangan tenaga kesehatan maupun masyarakat umum, dapat dengan mudah memahami manfaat aplikasi sebelum melanjutkan ke tahap penggunaan fitur utama. Selain menyajikan informasi edukatif, halaman ini juga berperan sebagai navigasi awal menuju fitur inti aplikasi, yaitu proses input data untuk melakukan prediksi stunting.

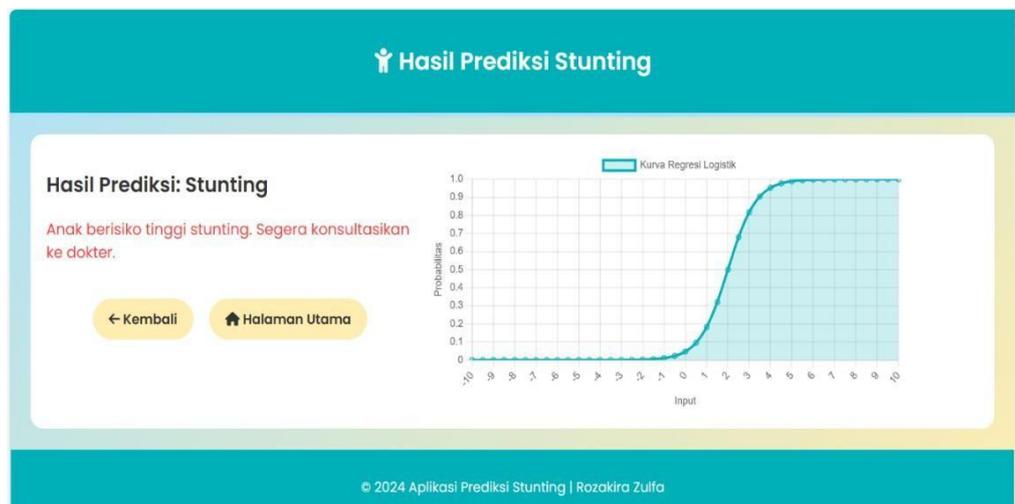
Tombol “Mulai Prediksi” disediakan secara jelas dan menonjol untuk memandu pengguna agar dapat langsung mengakses fitur prediksi tanpa kebingungan. Penempatan elemen-elemen antarmuka dirancang terstruktur dan tidak berlebihan, sehingga alur penggunaan aplikasi terasa sederhana dan ramah bagi pengguna baru. Dari segi desain, aplikasi menggunakan kombinasi warna yang lembut dan menenangkan, seperti gradasi hijau dan kuning, yang merepresentasikan kesehatan, pertumbuhan, dan kepedulian terhadap anak. Pemilihan warna ini bertujuan untuk menciptakan kesan positif, nyaman, dan tidak melelahkan mata. Selain itu, penggunaan ikon, tipografi yang jelas, serta tata letak yang rapi mendukung keterbacaan informasi dan meningkatkan pengalaman pengguna (*user experience*). Secara keseluruhan, desain halaman dibuat sederhana namun informatif agar mudah dipahami oleh semua kalangan pengguna.

Gambar 6. Halaman formulir input data

Setelah pengguna memahami gambaran umum mengenai sistem melalui halaman beranda, pengguna dapat melanjutkan ke tahap berikutnya dengan menekan tombol “Mulai Prediksi”. Tombol ini berfungsi sebagai navigasi utama yang mengarahkan pengguna ke halaman formulir input data anak. Dengan alur yang sederhana dan jelas, pengguna tidak perlu melalui banyak langkah untuk mulai menggunakan fitur utama sistem, sehingga pengalaman penggunaan menjadi lebih efisien dan ramah bagi semua kalangan. Halaman Formulir *Input Data Anak* merupakan halaman inti dalam sistem prediksi *stunting*, karena pada halaman inilah pengguna memasukkan data-data penting yang akan digunakan sebagai dasar proses prediksi. Data yang

diminta meliputi jenis kelamin anak, usia dalam bulan, berat badan saat lahir, tinggi badan saat lahir, berat badan saat ini, tinggi badan saat ini, serta riwayat pemberian ASI eksklusif. Setiap kolom dilengkapi dengan label dan *placeholder* yang jelas untuk meminimalkan kesalahan pengisian data dan membantu pengguna memahami informasi apa saja yang dibutuhkan. Dari segi desain antarmuka, halaman formulir ini dirancang dengan tampilan yang bersih dan terstruktur.

Penggunaan kombinasi warna gradasi lembut yang konsisten dengan halaman beranda memberikan kesan visual yang nyaman dan profesional. Warna hijau kebiruan pada tombol “Prediksi” dipilih untuk menonjolkan aksi utama pada halaman ini, sekaligus memberikan kesan positif dan aman yang identik dengan tema kesehatan. Tata letak *form* yang rapi dan penggunaan ikon pada setiap *input* juga membantu meningkatkan keterbacaan serta memudahkan pengguna dalam proses pengisian data. Secara keseluruhan, halaman formulir *input* data berperan sebagai jembatan antara informasi awal pada halaman beranda dan hasil prediksi yang akan ditampilkan oleh sistem. Dengan desain yang sederhana, warna yang menenangkan, serta alur penggunaan yang jelas, halaman ini diharapkan dapat mendukung pengguna dalam memasukkan data secara akurat sehingga hasil prediksi stunting yang dihasilkan sistem menjadi lebih optimal dan dapat dijadikan bahan pertimbangan awal dalam pemantauan pertumbuhan anak.



Gambar 7. Halaman hasil prediksi *stunting*

Halaman hasil prediksi merupakan halaman lanjutan yang ditampilkan setelah pengguna mengisi dan mengirimkan data anak melalui halaman *stunting* formulir input data. Halaman ini berfungsi untuk menyajikan hasil analisis sistem terhadap data yang telah dimasukkan, sehingga pengguna dapat mengetahui status risiko *stunting* pada anak secara langsung. Informasi disajikan secara ringkas dan jelas agar mudah dipahami oleh pengguna dari berbagai latar belakang, baik tenaga kesehatan maupun masyarakat umum. Pada halaman ini ditampilkan hasil prediksi dalam bentuk teks yang menunjukkan status anak, misalnya berisiko *stunting* atau tidak, disertai dengan pesan rekomendasi yang bersifat informatif. Jika anak terindikasi memiliki risiko *stunting*, sistem memberikan imbauan kepada pengguna untuk segera melakukan konsultasi dengan tenaga medis sebagai langkah pencegahan dini. Selain itu, halaman ini juga dilengkapi dengan visualisasi grafik kurva regresi logistik yang menggambarkan probabilitas hasil prediksi berdasarkan *input* data, sehingga pengguna dapat memperoleh gambaran analitis dari proses prediksi yang dilakukan oleh sistem.

Dari segi desain, halaman hasil prediksi menggunakan kombinasi warna yang konsisten dengan halaman sebelumnya, yaitu nuansa biru dan hijau yang memberikan kesan tenang, profesional, serta berkaitan dengan tema kesehatan. Tata letak dibuat seimbang antara teks hasil prediksi dan grafik, sehingga informasi utama mudah ditangkap tanpa membuat pengguna merasa bingung. Tombol navigasi seperti Kembali dan Halaman Utama juga disediakan untuk memudahkan pengguna dalam mengulang proses input data atau kembali ke halaman beranda, sehingga alur penggunaan aplikasi tetap sederhana dan intuitif. Dengan adanya halaman ini, pengguna tidak hanya memperoleh hasil prediksi, tetapi juga mendapatkan gambaran visual dan arahan awal untuk pengambilan keputusan yang lebih tepat terkait kesehatan anak.

### 3.4 Perbandingan Hasil dengan Penelitian Terdahulu

Hasil penelitian ini menunjukkan bahwa model *Logistic Regression* mampu memprediksi *stunting* dengan performa yang baik, khususnya pada metrik *recall* sebesar 85,4% dan *F1-score* sebesar 81,7%. Nilai *recall* yang tinggi sangat penting dalam konteks deteksi dini *stunting*, karena berfokus pada kemampuan model dalam mengenali sebanyak mungkin kasus positif (anak

yang mengalami *stunting*), sehingga intervensi dapat segera dilakukan. Hasil ini sejalan dengan penelitian yang dilakukan oleh Pangaribuan et al. [10] yang menggunakan metode *Logistic Regression* dalam mendeteksi penyakit jantung. Penelitian tersebut menunjukkan bahwa *Logistic Regression* mampu menghasilkan tingkat sensitivitas (*recall*) sebesar 88,54%, yang berarti model tersebut berhasil mengidentifikasi sebagian besar kasus positif dengan tepat. Keberhasilan *Logistic Regression* dalam konteks tersebut menunjukkan bahwa metode ini memiliki stabilitas performa yang baik untuk klasifikasi biner di bidang kesehatan.

Selain performa yang cukup tinggi, *Logistic Regression* juga dikenal karena kemudahannya dalam interpretasi. Seperti yang juga ditunjukkan oleh Pangaribuan et al. [10], metode ini mempermudah dalam memahami pengaruh setiap variabel terhadap kemungkinan hasil prediksi. Keunggulan ini menjadikan *Logistic Regression* sangat cocok digunakan dalam pengambilan keputusan berbasis data, termasuk untuk mendukung deteksi awal *stunting* secara efisien dan transparan. Dengan demikian, temuan dalam penelitian ini memperkuat bukti bahwa *Logistic Regression* adalah metode yang relevan dan dapat diandalkan dalam sistem prediksi kesehatan, serta dapat menjadi baseline yang baik untuk pengembangan model yang lebih kompleks ke depannya.

#### 4. KESIMPULAN

Penelitian ini berhasil membangun sistem prediksi *stunting* anak berbasis web menggunakan algoritma *Logistic Regression*. Berdasarkan hasil evaluasi model, diperoleh nilai akurasi sebesar 69,9%, *precision* 78,4%, *recall* 85,4%, dan *F1-score* 81,7%. Nilai *recall* yang tinggi menandakan bahwa model mampu mendeteksi sebagian besar kasus anak yang mengalami *stunting*, yang sangat penting dalam konteks deteksi dini dan pencegahan. Proses pengembangan aplikasi dimulai dari pra-pemrosesan data seperti normalisasi dan penghapusan outlier, pelatihan model dengan data latih, serta integrasi model ke dalam web berbasis *Python Flask*.

Aplikasi yang dihasilkan memungkinkan pengguna untuk melakukan prediksi secara *real-time* dengan antarmuka yang sederhana dan mudah digunakan. Hasil prediksi dan visualisasi evaluasi menunjukkan bahwa *Logistic Regression* cukup layak digunakan sebagai *baseline* dalam sistem prediksi awal *stunting*. Keunggulannya dalam interpretabilitas dan efisiensi membuat metode ini cocok digunakan untuk pengambilan keputusan awal, yang dapat dikembangkan lebih lanjut dengan metode yang lebih kompleks untuk peningkatan akurasi.

#### DAFTAR PUSTAKA

- [1] W. H. Organization, "Levels and Trends in Child Malnutrition." Accessed: Jul. 12, 2025. [Online]. Available: <https://www.who.int/publications/i/item/9789240025257>
- [2] K. K. R. Indonesia, "Survei Status Gizi Indonesia (SSGI) 2024."
- [3] UNICEF, "Improving Child Nutrition: The Achievable Imperative for Global Progress." Accessed: Jul. 12, 2025. [Online]. Available: [https://www.unicef.org/publications/index\\_68661.html](https://www.unicef.org/publications/index_68661.html)
- [4] W. C. Wahyudin, F. M. Hana, and A. Prihandono, "Prediksi Stunting Pada Balita Di Rumah Sakit Kota Semarang Menggunakan Naive Bayes," *J. Ilmu Komput. dan Matematika*, vol. 2019, pp. 32–36, 2023.
- [5] R. W. Febriana and E. Setyati, "Sistem Prediksi Risiko Stunting Menggunakan Bayesian Network Berbasis GIS," *Eksplor. Teknol. Enterp. dan Sist. Inf.*, vol. 1, no. 1, pp. 9–19, 2022, doi: 10.59039/ekstensi.v1i1.2.
- [6] F. T. Sabillillah, C. A. Sari, R. B. Abiyi, and A. Susanto, "Comparison Of Machine Learning Algorithms On Stunting Detection For 'Centing' Mobile Application To Prevent Stunting," *Sinkron*, vol. 8, no. 4, pp. 2360–2368, 2024, doi: 10.33395/sinkron.v8i4.13967.
- [7] I. P. Putri, T. Terttiaavini, and N. Arminarahmah, "Analisis Perbandingan Algoritma Machine Learning untuk Prediksi Stunting pada Anak," *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 4, no. 1, pp. 257–265, 2024, doi: 10.57152/malcom.v4i1.1078.
- [8] J. Mkungudza, H. S. Twabi, and S. O. M. Manda, "Development of a diagnostic predictive model for determining child stunting in Malawi: a comparative analysis of variable selection approaches," *BMC Med. Res. Methodol.*, vol. 24, no. 1, pp. 1–12, 2024, doi: 10.1186/s12874-024-02283-6.
- [9] C. Raras, A. Widiawati, L. Nurazizah, and I. R. Yunita, "Implementasi Algoritma Logistic Regression pada Pembuatan Website Sederhana untuk Prediksi Penyakit Jantung," *Infotekmesin*, vol. 15, no. 01, pp. 117–122, 2024, doi: 10.35970/infotekmesin.v15i1.2048.
- [10] J. Junifer Pangaribuan, H. Tanjaya, and K. Kenichi, "Mendeteksi Penyakit Jantung Menggunakan Machine Learning Dengan Algoritma Logistic Regression," *J. Inf. Syst. Dev.*, vol. 06, no. 02, pp. 1–10, 2021.
- [11] I. W. T. Wahyudi and I. G. A. G. A. Kadyanan, "Implementasi Logistic Regression dalam Sistem Diagnosa Penyakit Diabetes dengan KNN," *JELIKU (Jurnal Elektron. Ilmu Komput. Udayana)*, vol. 11, no. 4, p. 743, 2023, doi:

- 10.24843/jlk.2023.v11.i04.p12.
- [12] A. Hasibuan, W. V. Siregar, M. Isa, E. Warman, R. Finata, and M. Mursalin, "The Use of Regression Method on Simple E for Estimating Electrical Energy Consumption," *HighTech Innov. J.*, vol. 3, no. 3, pp. 306–318, 2022, doi: 10.28991/HIJ-SP2022-03-06.
- [13] R. M. Permana, "Machine Learning Approaches To Child Stunting Prediction," *Klover Multidiscip. J. Eng.*, vol. 7, no. 2, pp. 1–8, 2023, [Online]. Available: <http://kloverjournals.org/index.php/Engineering/article/view/719>
- [14] L. Valentina, "Gizi Anak." [Online]. Available: <https://www.kaggle.com/datasets/laurenvalentina/gizi-anak>
- [15] A. A. A. Hakim, R. A. Saputra, and S. Stiswaty, "Deteksi Dini Gejala Stres pada Mahasiswa Berdasarkan Faktor-Faktor Penyebabnya Menggunakan Metode Logistic Regression," 2024. doi: 10.20961/ijai.v8i2.83441.
- [16] P. R. Fhonna and A. Marzuki, "Sistem Informasi Absensi Pegawai Pada Biro Kominfo Kantor Bupati Kabupaten Aceh Utara Berbasis Web," *J. Ilmu Komput. dan Sist. Inf.*, vol. 3, no. 3, pp. 333–340, 2021.
- [17] R. P. Fhonna, A. T. Hidayat, and M. Abdullah, "Pemanfaatan Aplikasi Bank Soal Untuk Meningkatkan Kualitas Proses Belajar Mengajar Pada SMP Negeri 2 Nisam Antara," *J. Pengabd. Kpd. Masy. Nusant.*, vol. 4, no. 4, pp. 3680–3686, 2023.