BAB 2
LANDASAN TEORI

Bab ini membahas teori-teori yang menjadi dasar penelitian, termasuk tinjauan penelitian terdahulu, identifikasi celah penelitian, konsep malnutrisi anak berdasarkan standar WHO, dataset yang digunakan, teori mengenai algoritma *machine learning* yang digunakan, serta evaluasi model.

2.1 Penelitian Terdahulu

Penelitian terkait deteksi dini malnutrisi anak dengan pendekatan *machine learning* berkembang pesat dalam lima tahun terakhir, terutama dengan memanfaatkan data survei nasional terstandar seperti *Demographic and Health Surveys* (DHS) dan *Multiple Indicator Cluster Surveys* (MICS). Meta-analisis terbaru oleh Rao dkk. [9] yang menelaah sebelas studi *machine learning* berbasis data DHS menemukan bahwa algoritma *ensemble* berbasis *tree*, khususnya Random Forest dan gradient boosting (terutama XGBoost), secara konsisten mengungguli model statistik konvensional dalam memodelkan hubungan non-linear antara faktor antropometri, sosioekonomi, dan status gizi anak. Tinjauan sistematis yang lebih luas oleh Janssen dkk. [11] menegaskan tren ini, sekaligus menyoroti pergeseran fokus penelitian dari semata-mata mengejar performa prediksi menuju *explainable AI* yang dapat dipertanggungjawabkan secara klinis.

Rahman dkk. (2021) memanfaatkan data Bangladesh DHS 2017–2018 untuk membangun model prediksi stunting, wasting, dan underweight pada anak balita menggunakan berbagai algoritma *machine learning*. Penelitian tersebut menemukan bahwa pendekatan *ensemble* menghasilkan akurasi prediksi yang lebih baik dibandingkan *Logistic Regression*, dan mengidentifikasi pendidikan ibu, indeks kekayaan rumah tangga (*wealth index*), serta jenis sanitasi sebagai prediktor utama [12]. Pada konteks yang berbeda, Bitew dkk. (2022) membandingkan XGBoost, Random Forest, dan Logistic Regression pada data Ethiopia DHS 2016 dan menemukan bahwa XGBoost mencapai performa terbaik dengan *ROC-AUC* sekitar 0,77 untuk indikator stunting, sementara analisis *feature importance* menyoroti usia anak, *wealth index*, dan tinggi badan ibu sebagai faktor paling berpengaruh [7].

Anku dan Duah (2024) menerapkan beberapa algoritma *machine learning*

untuk memprediksi indikator *undernutrition* (*wasting*, *stunting*, dan *underweight*) pada anak balita di Ghana, serta mengidentifikasi faktor-faktor yang berkorelasi melalui pemeringkatan prediktor penting. Hasilnya menunjukkan bahwa model *ensemble*, khususnya XGBoost, dapat mencapai performa tinggi pada beberapa indikator, namun cakupannya masih terbatas pada konteks satu negara [13]. Khan dkk. (2024) memperluas pendekatan ini dengan menggabungkan XGBoost dan teknik interpretabilitas berbasis SHAP (*SHapley Additive exPlanations*) pada data Bangladesh DHS. Studi tersebut menunjukkan bahwa XGBoost dengan *tuning* yang baik mampu mencapai *ROC-AUC* yang tinggi, sementara SHAP berhasil mengungkap pendidikan ibu, usia anak, dan *wealth index* sebagai prediktor utama, baik pada level populasi maupun individu [6].

Mgomezulu dkk. (2025) mengembangkan pendekatan *ensemble* dan *deep learning* untuk deteksi dini stunting dan wasting dengan penekanan khusus pada distribusi kelas yang seimbang (*balanced class distribution*). Penelitian ini menegaskan bahwa pada data malnutrisi yang umumnya tidak seimbang, teknik penyeimbangan kelas seperti *Synthetic Minority Over-sampling Technique* (SMOTE) atau penyesuaian bobot kelas meningkatkan *recall* pada kelas minoritas, yang krusial untuk skrining kesehatan masyarakat [14]. Jemil dkk. (2026) memanfaatkan data DHS lintas negara di Afrika Timur untuk memprediksi *severe stunting* dan menemukan bahwa pemodelan lintas negara menghasilkan generalisasi yang lebih baik dibandingkan studi *single-country*, sekaligus menggunakan pendekatan interpretabilitas untuk menyoroti determinan severe stunting yang konsisten lintas konteks [15].

Sinharoy dkk. (2026) mengembangkan model prediksi stunting dengan menekankan peran variabel *WASH* (*water, sanitation, and hygiene*). Studi ini menunjukkan bahwa variabel sosioekonomi dan lingkungan memiliki kontribusi penting terhadap prediksi stunting, sehingga mendukung rasional penggabungan data sosioekonomi sebagai pelengkap data antropometri dalam pemodelan risiko malnutrisi [16]. Selain prediksi indikator status gizi, machine learning juga dimanfaatkan untuk mendukung intervensi malnutrisi akut. Studi oleh Martin dkk. (2024) pada anak penderita *acute malnutrition* di program rawat jalan menunjukkan bahwa teknik prediktif dapat membantu memperkirakan kenaikan berat badan dan mengoptimalkan strategi perawatan, walaupun fokusnya bukan pada klasifikasi indikator antropometri WHO secara penuh [17].

Ringkasan penelitian-penelitian tersebut disajikan pada Tabel 2.1. Dari penelitian yang ditinjau, dapat diidentifikasi beberapa pola penting: (1) XGBoost

dan algoritma berbasis *gradient boosting* secara konsisten mencapai performa terbaik pada data tabular DHS; (2) variabel sosioekonomi seperti pendidikan ibu dan *wealth index* secara konsisten muncul sebagai prediktor utama selain variabel antropometri; (3) tren penelitian bergerak ke arah model yang tidak hanya akurat, tetapi juga dapat diinterpretasi melalui teknik *explainable AI* seperti SHAP. Meski demikian, masih terdapat sejumlah celah penelitian yang akan dibahas pada bagian berikutnya.

Tabel 2.1. Ringkasan Penelitian Terdahulu

Peneliti (Tahun)	**Dataset**	**Algoritma**	**Hasil/Kontribusi Utama**
Rahman dkk. (2021) [12]	Bangladesh DHS 2017–2018	Logistic Regression, Random Forest, gradient boosting, ANN	Ensemble lebih unggul; pendidikan ibu & wealth index sebagai prediktor utama.
Bitew dkk. (2022) [7]	Ethiopia DHS 2016	XGBoost, Random Forest, Logistic Regression	XGBoost terbaik (ROC-AUC ∼0,77 stunting); usia anak & wealth index dominan.
Anku & Duah (2024) [13]	Ghana DHS	Beberapa ML termasuk XGBoost	XGBoost konsisten terbaik untuk 3 indikator undernutrition; terbatas single-country.
Khan dkk. (2024) [6]	Bangladesh DHS	XGBoost + SHAP	XGBoost dengan tuning baik + interpretasi SHAP; pendidikan ibu & wealth index utama.
Mgomezulu dkk. (2025) [14]	Multi-country	Ensemble + Deep Learning + balanced class	Penekanan pada penyeimbangan kelas; meningkatkan recall pada kelas minoritas.
Jemil dkk. (2026) [15]	DHS Afrika Timur (lintas negara)	Beberapa ML	Pemodelan lintas negara untuk *severe stunting*; interpretabilitas lintas konteks.
Sinharoy dkk. (2026) [16]	DHS dengan fokus WASH	ML berbasis fitur WASH	Variabel WASH (air, sanitasi, hygiene) berkontribusi penting terhadap prediksi stunting.

2.2 Research Gap

Berdasarkan tinjauan penelitian terdahulu pada Sub-bab 2.1, dapat diidentifikasi tiga celah penelitian utama yang menjadi motivasi penelitian ini.

Pertama, sebagian besar penelitian masih terbatas pada konteks satu negara, misalnya Rahman dkk. (2021) di Bangladesh [12], Bitew dkk. (2022) di Ethiopia [7], Anku dan Duah (2024) di Ghana [13], dan Khan dkk. (2024) di Bangladesh [6]. Pendekatan lintas negara mulai muncul pada studi-studi terbaru [15, 14], namun evaluasi generalisasi model lintas negara dengan tingkat sosioekonomi yang berbeda masih jarang dilakukan secara sistematis. Padahal, kemampuan model untuk berkinerja konsisten di berbagai wilayah merupakan prasyarat penting untuk penggunaan praktis pada banyak negara berpenghasilan rendah dan menengah.

Kedua, sejumlah penelitian masih menggunakan *accuracy* sebagai metrik utama, padahal data malnutrisi anak balita umumnya bersifat tidak seimbang. Prevalensi stunting global hanya sekitar 22% pada tahun 2024, sedangkan prevalensi wasting bahkan lebih rendah, yakni 6,6% [1]. Pada kondisi data yang tidak seimbang, accuracy dapat memberikan gambaran yang menyesatkan karena model yang sekadar memprediksi kelas mayoritas saja tetap dapat mencapai accuracy tinggi. Mgomezulu dkk. (2025) menegaskan pentingnya teknik penyeimbangan kelas serta penggunaan metrik yang lebih sensitif terhadap kelas minoritas, seperti *recall*, *F1-score*, dan *ROC-AUC* [14]. Pada konteks deteksi dini malnutrisi, recall menjadi sangat krusial karena biaya melewatkan anak yang sebenarnya berstatus malnutrisi (*false negative*) jauh lebih besar dibandingkan biaya melakukan intervensi pada anak yang sebenarnya tidak bermalnutrisi (*false positive*).

Ketiga, meskipun beberapa studi terbaru sudah memanfaatkan teknik *explainable AI* seperti SHAP [6, 8], studi yang secara eksplisit menyoroti kontribusi relatif fitur antropometri dibandingkan fitur sosioekonomi sebagai prediktor masih terbatas, terutama pada konteks lintas negara. Identifikasi kontribusi setiap kelompok variabel penting untuk merancang intervensi yang tepat sasaran, misalnya untuk memutuskan apakah intervensi gizi langsung pada anak lebih efektif dibandingkan intervensi pada faktor sosioekonomi keluarga, atau sebaliknya.

Berdasarkan tiga celah tersebut, penelitian ini berfokus pada deteksi dini malnutrisi anak balita dengan tiga kontribusi: (1) memanfaatkan data DHS dari 29 negara untuk menguji generalisasi model lintas konteks sosioekonomi;

(2) menggunakan kombinasi metrik evaluasi yang lebih sesuai untuk data tidak seimbang, yakni *ROC-AUC*, *balanced accuracy*, *recall*, *precision*, *F1-score*, dan *accuracy*; dan (3) menerapkan analisis interpretabilitas yang mengeksplorasi kontribusi relatif fitur antropometri dan sosioekonomi terhadap prediksi malnutrisi.

2.3 Konsep Malnutrisi Anak

Bagian ini menguraikan konsep dasar malnutrisi pada anak balita yang menjadi target penelitian, mulai dari definisi malnutrisi menurut standar internasional, indikator antropometri yang digunakan untuk mengklasifikasi status gizi, hingga faktor-faktor yang menjadi penyebab terjadinya malnutrisi.

2.3.1 Definisi Malnutrisi

Malnutrisi merupakan kondisi fisiologis yang terjadi ketika asupan nutrisi seseorang tidak memenuhi kebutuhan tubuh, atau ketika tubuh tidak mampu menyerap nutrisi dengan baik. Pada anak balita usia 0–59 bulan, malnutrisi tidak hanya merefleksikan defisit kalori, tetapi juga ketidakseimbangan zat gizi mikro yang berdampak langsung pada pertumbuhan fisik dan perkembangan kognitif anak [18]. WHO bersama UNICEF dan World Bank mengategorikan malnutrisi anak ke dalam empat kondisi utama yang masing-masing didasarkan pada perbandingan nilai antropometri terhadap kurva standar pertumbuhan WHO [1, 19]:

- Stunting: kondisi gagal tumbuh akibat kekurangan gizi kronis dan/atau infeksi berulang, ditandai dengan nilai *Height-for-Age Z-score* (HAZ) < -2 SD. Stunting umumnya berkembang perlahan dan paling banyak terdeteksi pada usia 12–23 bulan.
- Wasting: kondisi kekurangan gizi akut yang ditandai dengan berat badan sangat rendah relatif terhadap tinggi badan, atau nilai *Weight-for-Height Z-score* (WHZ) < -2 SD. Wasting berkembang lebih cepat dibandingkan stunting dan biasanya dipicu oleh kekurangan makan akut atau penyakit infeksi.
- Underweight: kondisi berat badan rendah berdasarkan umur (*Weight-for-Age Z-score*, WAZ < -2 SD). Underweight merupakan indikator komposit yang dapat mencerminkan stunting, wasting, atau gabungan keduanya.

- Overweight: kondisi kelebihan berat badan untuk tinggi badan (WHZ $> +2$ SD) yang mencerminkan ketidakseimbangan energi. Overweight juga termasuk bentuk malnutrisi, yang sering disebut *double burden of malnutrition* ketika muncul bersamaan dengan stunting atau wasting di populasi yang sama.

2.3.2 Tabel Referensi Status Gizi WHO

Penentuan status gizi anak menggunakan Z-score, yaitu skor yang membandingkan nilai pengukuran antropometri anak terhadap nilai median populasi referensi WHO pada kelompok usia dan jenis kelamin yang sama. Z-score dihitung dengan rumus:

$$Z = \frac{X - M}{S} \tag{2.1}$$

dengan X adalah nilai aktual pengukuran anak (tinggi atau berat badan), M adalah nilai median populasi referensi WHO, dan S adalah simpangan baku populasi referensi. Jika nilai Z-score mendekati 0, anak berada pada kondisi pertumbuhan setara median populasi referensi. Z-score negatif menunjukkan kondisi di bawah median, sedangkan Z-score positif menunjukkan kondisi di atas median. Anak dikategorikan mengalami malnutrisi bila nilai Z-score dari salah satu indikator berada di bawah -2 SD; sedangkan nilai di bawah -3 SD menunjukkan kondisi malnutrisi berat (*severe*) [1].

Klasifikasi status gizi anak berdasarkan rentang Z-score disajikan pada Tabel 2.2. Sebagai catatan tambahan, pengukuran antropometri di lapangan rentan terhadap kesalahan baca atau pencatatan. WHO menetapkan batas *biologically implausible value* (BIV) jika nilai Z-score berada di luar rentang ± 6 SD, sehingga data tersebut dikeluarkan dari analisis sebelum pemodelan [20].

Tabel 2.2. Klasifikasi Status Gizi Berdasarkan WHO Z-Score

Kategori	Rentang Z-Score
Normal	-2 SD $\leq Z \leq +2$ SD
Stunting (HAZ)	$Z < -2$ SD
Wasting (WHZ)	$Z < -2$ SD
Underweight (WAZ)	$Z < -2$ SD
Overweight (WHZ)	$Z > +2$ SD

Sumber: ambang klasifikasi status gizi mengikuti standar pertumbuhan anak WHO [19, 1].

2.3.3 Faktor-Faktor Penyebab Malnutrisi

Malnutrisi anak merupakan hasil interaksi kompleks antara faktor biologis anak, kondisi rumah tangga, dan lingkungan sosial-ekonomi keluarga. Berdasarkan tinjauan literatur sistematis oleh Katoch [18] dan kerangka konseptual UNICEF [19], faktor-faktor yang berkontribusi terhadap malnutrisi anak balita dapat dikelompokkan sebagai berikut:

1. Faktor Biologis: usia anak, jenis kelamin, dan berat badan lahir. Anak dengan berat badan lahir rendah berisiko lebih tinggi mengalami stunting dan underweight pada masa balita.

2. Faktor Ibu: tingkat pendidikan ibu, status nutrisi ibu, dan pengetahuan gizi. Pendidikan ibu merupakan salah satu prediktor sosioekonomi paling konsisten yang dilaporkan pada studi multi-negara [5, 21].

3. Faktor Rumah Tangga: kondisi sanitasi, akses terhadap air bersih, dan jumlah anggota keluarga. Sanitasi yang buruk meningkatkan paparan terhadap penyakit infeksi yang menghambat penyerapan zat gizi.

4. Faktor Ekonomi: pendapatan keluarga dan status kemiskinan, yang umumnya diukur melalui *wealth index* berbasis kepemilikan aset rumah tangga. Kuintil termiskin secara konsisten memiliki prevalensi stunting yang lebih tinggi dibandingkan kuintil terkaya [21].

5. Faktor Lingkungan: lokasi tempat tinggal (*urban* vs. *rural*) dan akses terhadap layanan kesehatan. Anak di daerah pedesaan umumnya memiliki risiko malnutrisi yang lebih tinggi dibandingkan anak di perkotaan.

6. Faktor Penyakit: riwayat diare, infeksi saluran pernapasan akut (ISPA), serta status imunisasi anak. Penyakit infeksi berulang merupakan penyebab langsung terjadinya wasting akut.

Kombinasi faktor-faktor di atas menjadi dasar pemilihan variabel prediktor pada penelitian ini, di mana variabel antropometri (faktor biologis anak) dipadukan dengan variabel sosioekonomi (faktor ibu, rumah tangga, ekonomi, dan lingkungan) untuk membangun model prediksi malnutrisi.

2.4 Dataset Global

Bagian ini menguraikan dataset yang menjadi sumber data utama penelitian, yaitu *Demographic and Health Surveys* (DHS). DHS dipilih karena menyediakan kombinasi data antropometri dan sosioekonomi yang terstandar di banyak negara, sehingga memungkinkan pemodelan lintas konteks geografis dan ekonomi.

2.4.1 Demographic and Health Surveys (DHS)

Demographic and Health Surveys (DHS) merupakan program survei rumah tangga berskala nasional yang dikelola oleh *Inner City Fund* (ICF) dengan dukungan dari *United States Agency for International Development* (USAID). DHS telah dilaksanakan di lebih dari 90 negara berpenghasilan rendah dan menengah sejak tahun 1984, dengan siklus survei umumnya setiap lima tahun, dan banyak digunakan untuk memantau indikator kependudukan, kesehatan ibu dan anak, serta nutrisi [22, 10].

Setiap gelombang DHS umumnya menghasilkan beberapa berkas data utama yang dapat digabungkan melalui *cluster ID* dan *household ID* [10]:

- *Household Recode* (HR): berisi karakteristik rumah tangga, termasuk akses air dan sanitasi, kepemilikan aset, dan komposisi anggota keluarga.
- *Individual Recode* (IR): berisi karakteristik wanita usia subur (15–49 tahun), termasuk tingkat pendidikan, status pekerjaan, dan riwayat kehamilan.
- *Birth Recode* (BR) dan *Kids Recode* (KR): berisi data anak balita, termasuk hasil pengukuran antropometri (berat dan tinggi badan), riwayat imunisasi, dan riwayat penyakit dalam dua minggu terakhir.

Dalam konteks pemodelan malnutrisi anak, DHS menyediakan dua kelompok variabel utama. Pertama, variabel antropometri berupa berat badan, tinggi atau panjang badan, dan usia anak yang dapat dikonversi menjadi Z-score (HAZ, WHZ, WAZ) berdasarkan standar pertumbuhan WHO. Kedua, variabel sosioekonomi seperti tingkat pendidikan ibu, indeks kekayaan rumah tangga (*wealth index* yang dihitung melalui analisis komponen utama atas kepemilikan aset), akses terhadap air minum layak, jenis sanitasi, lokasi tempat tinggal (perkotaan atau pedesaan), serta karakteristik demografis lainnya [10].

Beberapa keunggulan DHS yang menjadikannya cocok untuk penelitian ini meliputi: (1) cakupan geografis yang luas dengan lebih dari 90 negara, sehingga memungkinkan analisis lintas wilayah; (2) standarisasi metodologi sampling, kuesioner, dan pengukuran antropometri antar negara, sehingga hasil analisis dapat dibandingkan secara konsisten; (3) ketersediaan data secara terbuka untuk peneliti setelah pengajuan akses; serta (4) ukuran sampel yang cukup besar untuk mendukung pelatihan model *machine learning* dengan tingkat generalisasi yang baik. Sebagai pelengkap, terdapat program survei serupa yaitu *Multiple Indicator Cluster Surveys* (MICS) yang dikelola oleh UNICEF [23, 24], namun pada penelitian ini hanya DHS yang digunakan untuk menjaga konsistensi standardisasi data antar negara.

2.5 Teori Machine Learning

Bagian ini menguraikan teori dasar pendekatan *machine learning* yang digunakan pada penelitian ini, mulai dari paradigma *supervised learning*, algoritma XGBoost yang menjadi algoritma utama, hingga teknik penanganan data tidak seimbang yang umum digunakan dalam klasifikasi malnutrisi.

2.5.1 Supervised Learning

Penelitian ini menggunakan pendekatan *supervised learning*, yaitu jenis pembelajaran mesin di mana model dilatih menggunakan data berlabel untuk mempelajari hubungan antara fitur masukan dan keluaran target [25]. Tujuan utama *supervised learning* adalah menemukan fungsi pemetaan $f : X \rightarrow Y$ yang mampu memberikan prediksi yang baik tidak hanya pada data pelatihan, tetapi juga pada data baru yang belum pernah dilihat sebelumnya. Pada konteks penelitian ini, X merupakan kumpulan variabel antropometri dan sosioekonomi anak balita,

sedangkan Y adalah label status gizi anak (misalnya berstatus *stunted* atau tidak). Karena Y bersifat biner, masalah ini termasuk dalam kelas *binary classification*.

2.5.2 XGBoost (Extreme Gradient Boosting)

XGBoost (*Extreme Gradient Boosting*) adalah algoritma *gradient boosting* berbasis pohon keputusan (*decision tree*) [26, 27]. XGBoost menjadi salah satu algoritma paling populer untuk klasifikasi data tabular karena performanya yang tinggi, efisiensi komputasi, serta ketahanannya terhadap *overfitting* berkat fungsi regularisasi yang terintegrasi pada fungsi objektifnya.

A Konsep Tree Boosting

Boosting merupakan teknik *ensemble learning* yang menggabungkan banyak model lemah (*weak learner*), umumnya berupa pohon keputusan dangkal, menjadi satu model kuat melalui pemodelan aditif bertahap (*forward stagewise additive modeling*) [28]. Berbeda dengan pendekatan *bagging* (misalnya *Random Forest*) yang membangun banyak pohon secara paralel dan independen lalu merata-ratakan hasilnya, *boosting* membangun pohon secara berurutan (*sequential*): setiap pohon baru difokuskan untuk memperbaiki kesalahan model gabungan sebelumnya [25].

Gradient boosting [28] memformalkan gagasan tersebut sebagai penurunan gradien pada ruang fungsi (*functional gradient descent*): setiap pohon baru dilatih untuk memprediksi gradien negatif dari fungsi *loss* terhadap prediksi saat ini, yaitu residu semu (*pseudo-residual*). Dengan demikian, penambahan setiap pohon menggeser prediksi ke arah yang paling menurunkan *loss*. XGBoost merupakan implementasi *gradient boosting* yang dapat diskalakan dengan regularisasi terintegrasi dan optimasi komputasi [26].

Pada langkah ke-t, prediksi model dihitung sebagai akumulasi kontribusi dari seluruh pohon yang telah dibangun:

$$\hat{y}_i^{(t)} = \hat{y}_i^{(t-1)} + \eta\, f_t(x_i) \tag{2.2}$$

dengan f_t adalah pohon keputusan ke-t yang dilatih untuk meminimalkan kesalahan prediksi sisa, dan η adalah *learning rate* yang mengontrol seberapa besar kontribusi pohon baru terhadap model akhir. Nilai η yang kecil membuat pelatihan lebih

stabil dan tahan terhadap *overfitting*, namun memerlukan jumlah pohon yang lebih banyak.

B Fungsi Objektif dan Regularisasi

XGBoost meminimalkan fungsi objektif yang menggabungkan *loss function* dan istilah regularisasi:

$$Obj = \sum_{i=1}^{n} l(y_i, \hat{y}_i) + \sum_{k=1}^{K} \Omega(f_k) \tag{2.3}$$

dengan l adalah *loss function* yang mengukur selisih antara nilai sebenarnya y_i dan prediksi $\hat{y}_i$, sedangkan $\Omega(f_k)$ adalah fungsi regularisasi yang mengontrol kompleksitas pohon ke-k. Pada klasifikasi biner seperti pada penelitian ini, *loss function* yang digunakan adalah *binary log-loss* (atau *binary cross-entropy*):

$$l(y_i, \hat{y}_i) = -\left[y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)\right] \tag{2.4}$$

dengan $\hat{y}_i$ merupakan probabilitas prediksi yang dihasilkan melalui fungsi sigmoid:

$$\hat{y}_i = \frac{1}{1 + e^{-z_i}}, \quad z_i = \sum_{k=1}^{K} f_k(x_i) \tag{2.5}$$

Fungsi regularisasi $\Omega(f_k)$ pada XGBoost memiliki bentuk:

$$\Omega(f_k) = \gamma T_k + \tfrac{1}{2}\lambda \|w_k\|^2 + \alpha \|w_k\|_1 \tag{2.6}$$

dengan T_k adalah jumlah daun pada pohon ke-k, w_k adalah bobot daun, γ penalti per daun, λ koefisien regularisasi L_2, dan α koefisien regularisasi L_1. Adanya kombinasi regularisasi L_1 dan L_2 ini menjadi salah satu pembeda utama XGBoost dari implementasi *gradient boosting* konvensional dan membuatnya lebih tahan terhadap *overfitting*, terutama pada data berdimensi tinggi.

C Pembelajaran Aditif dan Optimasi Orde Kedua

Karena setiap pohon f_t ditambahkan secara aditif, fungsi objektif pada iterasi ke-t dapat didekati menggunakan ekspansi Taylor orde kedua di sekitar prediksi sebelumnya $\hat{y}_i^{(t-1)}$ [26]:

$$\tilde{Obj}^{(t)} \simeq \sum_{i=1}^{n} \left[g_i f_t(x_i) + \tfrac{1}{2} h_i f_t(x_i)^2 \right] + \Omega(f_t) \tag{2.7}$$

dengan $g_i = \partial_{\hat{y}^{(t-1)}} l(y_i, \hat{y}_i^{(t-1)})$ dan $h_i = \partial^2_{\hat{y}^{(t-1)}} l(y_i, \hat{y}_i^{(t-1)})$ berturut-turut adalah gradien dan Hessian dari fungsi *loss*. Inilah pembeda utama XGBoost dari *gradient boosting* klasik yang hanya menggunakan informasi orde pertama: XGBoost memanfaatkan gradien sekaligus Hessian (pendekatan mirip metode Newton), sehingga pelatihan lebih cepat dan stabil [26].

Untuk struktur pohon tertentu, bobot optimal setiap daun j serta skor kualitas struktur pohon (*structure score*) dapat diturunkan dalam bentuk tertutup:

$$w_j^* = -\frac{G_j}{H_j + \lambda} \tag{2.8}$$

$$\tilde{Obj}^* = -\frac{1}{2} \sum_{j=1}^{T} \frac{G_j^2}{H_j + \lambda} + \gamma T \tag{2.9}$$

dengan $G_j = \sum_{i \in I_j} g_i$ dan $H_j = \sum_{i \in I_j} h_i$ adalah jumlah gradien dan Hessian pada daun j, serta I_j himpunan sampel yang jatuh pada daun tersebut. Skor struktur ini mengukur seberapa baik suatu susunan pohon; semakin kecil nilainya, semakin baik strukturnya.

Karena mengevaluasi seluruh kemungkinan struktur pohon tidak praktis, XGBoost membangun pohon secara serakah (*greedy*) dengan menghitung *gain* dari setiap kandidat pembelahan (*split*):

$$\text{Gain} = \frac{1}{2} \left[\frac{G_L^2}{H_L + \lambda} + \frac{G_R^2}{H_R + \lambda} - \frac{(G_L + G_R)^2}{H_L + H_R + \lambda} \right] - \gamma \tag{2.10}$$

dengan indeks L dan R menyatakan simpul anak kiri dan kanan hasil pembelahan. Pembelahan hanya dilakukan bila *gain*-nya bernilai positif; parameter γ berperan sebagai ambang minimum sehingga pembelahan yang manfaatnya kecil akan

dipangkas (*pruning*) [26].

D Karakteristik XGBoost

Beberapa karakteristik teknis XGBoost yang relevan dengan penelitian ini meliputi:

1. Sparsity-aware split finding: XGBoost mampu menangani *missing values* secara otomatis dengan menentukan arah pembagian (*default direction*) terbaik untuk nilai hilang pada setiap node, sehingga lebih robust pada data DHS yang sering memiliki proporsi *missing* bervariasi antar variabel [26].
2. Regularisasi L_1 dan L_2 yang menekan kompleksitas pohon dan mengurangi risiko *overfitting*, terutama pada data dengan banyak fitur.
3. Pembelajaran paralel pada level pembangunan pohon, sehingga proses pelatihan menjadi cepat bahkan pada dataset gabungan lintas negara berukuran besar.
4. Kemampuan menangkap interaksi non-linear antar fitur, yang penting pada konteks malnutrisi anak di mana hubungan antara variabel sosioekonomi dan status gizi bersifat kompleks dan tidak linear.
5. Penyediaan informasi pemeringkatan fitur (*feature importance*) yang dapat dipadukan dengan teknik interpretabilitas seperti SHAP untuk menjelaskan kontribusi setiap variabel terhadap prediksi.

Berdasarkan karakteristik di atas, XGBoost dipilih sebagai algoritma tunggal pada penelitian ini. Pada implementasi, fungsi objektif yang digunakan adalah `binary:logistic` untuk klasifikasi biner status malnutrisi (*stunted*, *wasted*, atau *underweight*), dengan opsi `tree_method=hist` untuk efisiensi komputasi pada data tabular berukuran besar.

2.5.3 Penanganan Data Tidak Seimbang

Pada penelitian klasifikasi malnutrisi anak, distribusi kelas umumnya tidak seimbang: jumlah anak yang berstatus malnutrisi (kelas positif) jauh lebih sedikit dibandingkan anak normal (kelas negatif). Prevalensi global stunting pada 2024 hanya sekitar 22%, dan wasting bahkan jauh lebih rendah, yakni 6,6% [1]. Kondisi

ini menyebabkan model yang dilatih cenderung bias ke kelas mayoritas dan kurang sensitif terhadap kasus malnutrisi yang justru menjadi target deteksi.

Beberapa pendekatan umum digunakan untuk menangani data tidak seimbang [14]:

1. Penyesuaian bobot kelas (*class weighting*) dengan parameter seperti `scale_pos_weight` pada XGBoost, yang memberikan bobot lebih besar pada kelas minoritas saat menghitung *loss*.
2. Random Oversampling, yaitu menduplikasi secara acak sampel dari kelas minoritas hingga jumlahnya seimbang dengan kelas mayoritas. Teknik ini sederhana dan tidak mengubah distribusi fitur asli.
3. Synthetic Minority Over-sampling Technique (SMOTE), yang menghasilkan sampel sintetis baru dari kelas minoritas melalui interpolasi terhadap tetangga terdekatnya pada ruang fitur.

Pada penelitian ini, ketidakseimbangan kelas tidak ditangani melalui penyeimbangan ulang data (*oversampling* atau SMOTE) maupun pembobotan kelas, melainkan melalui pemilihan ambang batas keputusan yang memaksimalkan *balanced accuracy*. Dengan demikian, parameter `scale_pos_weight` dibiarkan bernilai 1 dan distribusi data latih maupun data uji tetap mencerminkan populasi aslinya. Strategi rinci penerapannya dibahas pada Bab 3.

2.6 Explainable AI (SHAP)

Explainable AI (XAI) merupakan pendekatan yang bertujuan meningkatkan transparansi dan interpretabilitas model *machine learning*, terutama pada model kompleks seperti XGBoost yang sering kali berperan sebagai *black box*. Salah satu metode XAI yang paling banyak digunakan adalah SHAP (*SHapley Additive exPlanations*) [29]. SHAP didasarkan pada konsep *Shapley Value* dalam teori permainan kooperatif, yang mengukur kontribusi setiap fitur terhadap prediksi model.

Secara matematis, nilai SHAP untuk fitur ke-i didefinisikan sebagai:

$$\phi_i = \sum_{S \subseteq F \setminus \{i\}} \frac{|S|!(|F| - |S| - 1)!}{|F|!} [f(S \cup \{i\}) - f(S)] \tag{2.11}$$

di mana F adalah himpunan seluruh fitur, dan S adalah subset fitur yang tidak mengandung i. Term $f(S \cup \{i\}) - f(S)$ menunjukkan seberapa besar tambahan informasi yang diberikan fitur *i* terhadap prediksi ketika digabungkan dengan subset S. Dengan demikian, SHAP menjelaskan kontribusi setiap fitur secara aditif terhadap prediksi akhir model.

SHAP memiliki beberapa keunggulan: (1) memberikan interpretasi lokal (per individu) maupun global (keseluruhan model), (2) konsisten secara matematis, dan (3) mampu bekerja dengan berbagai jenis model, termasuk model *ensemble* berbasis pohon seperti XGBoost yang digunakan pada penelitian ini. SHAP digunakan untuk mengidentifikasi variabel antropometri dan sosioekonomi yang paling berpengaruh dalam menentukan status gizi anak, sehingga hasil prediksi tidak hanya akurat tetapi juga dapat dipertanggungjawabkan secara klinis dan kebijakan.

2.7 Evaluasi Model

Evaluasi model merupakan tahap penting untuk menilai kinerja algoritma dalam melakukan klasifikasi. Karena data malnutrisi anak balita umumnya tidak seimbang, pemilihan metrik harus dilakukan secara hati-hati: *accuracy* saja tidak cukup karena model yang sekadar memprediksi kelas mayoritas dapat memperoleh nilai accuracy yang tinggi tanpa benar-benar mendeteksi kasus malnutrisi. Pada penelitian ini digunakan kombinasi metrik berbasis *confusion matrix*, metrik berbasis probabilitas prediksi (ROC-AUC dan PR-AUC), serta strategi pemilihan ambang batas (*threshold tuning*) untuk memaksimalkan deteksi kasus malnutrisi.

2.7.1 Confusion Matrix

Confusion matrix adalah tabel 2×2 yang merangkum kinerja model klasifikasi biner dengan membandingkan prediksi model terhadap label sebenarnya. Pada konteks deteksi malnutrisi, kelas positif adalah anak yang berstatus malnutrisi (*stunted*, *wasted*, atau *underweight*), sedangkan kelas negatif adalah anak yang berstatus normal. Struktur confusion matrix disajikan pada Tabel 2.3.

Tabel 2.3. Struktur Confusion Matrix untuk Klasifikasi Biner

	Prediksi Positif	**Prediksi Negatif**
Aktual Positif	TP (*True Positive*)	FN (*False Negative*)
Aktual Negatif	FP (*False Positive*)	TN (*True Negative*)

Sumber: struktur *confusion matrix* klasifikasi biner mengikuti konvensi umum evaluasi *machine learning* [25].

Dari keempat nilai pada confusion matrix tersebut diturunkan berbagai metrik evaluasi klasifikasi yang dijelaskan pada subseksi berikut.

2.7.2 Metrik Berbasis Confusion Matrix

Accuracy mengukur proporsi prediksi yang benar terhadap seluruh data:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{2.12}$$

Pada data tidak seimbang, *accuracy* kurang informatif karena nilai tinggi dapat dicapai dengan memprediksi kelas mayoritas saja, tanpa benar-benar mengenali kelas minoritas.

Precision (atau *Positive Predictive Value*, PPV) mengukur proporsi prediksi positif yang benar-benar berstatus positif:

$$Precision = \frac{TP}{TP + FP} \tag{2.13}$$

Recall (atau *Sensitivity*) mengukur kemampuan model dalam menemukan seluruh kasus positif:

$$Recall = \frac{TP}{TP + FN} \tag{2.14}$$

Pada konteks skrining malnutrisi, *recall* merupakan metrik yang paling kritis karena melewatkan anak yang sebenarnya malnutrisi (*false negative*) berkonsekuensi jauh lebih besar dibandingkan menyaring anak yang sebenarnya normal (*false positive*).

Specificity mengukur kemampuan model mengenali kasus negatif dengan benar:

$$Specificity = \frac{TN}{TN + FP} \tag{2.15}$$

Pelengkap dari *specificity* adalah *Negative Predictive Value* (NPV), yang mengukur proporsi prediksi negatif yang benar-benar berstatus negatif:

$$NPV = \frac{TN}{TN + FN} \tag{2.16}$$

F1-score merupakan rata-rata harmonik dari *precision* dan *recall*, sehingga memberi ukuran tunggal yang menyeimbangkan keduanya:

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{2.17}$$

Balanced Accuracy merupakan rata-rata dari *recall* dan *specificity*, dan memberikan ukuran kinerja yang lebih adil pada data tidak seimbang dibandingkan *accuracy* biasa:

$$Balanced\ Accuracy = \frac{1}{2}(Recall + Specificity) \tag{2.18}$$

2.7.3 ROC-AUC

Area Under the Receiver Operating Characteristic Curve (ROC-AUC) mengukur kemampuan diskriminasi model dalam membedakan kelas positif dan negatif pada berbagai ambang batas (*threshold*). Kurva ROC memplot *recall* (sumbu Y) terhadap *False Positive Rate* (yaitu $1 - Specificity$, sumbu X) untuk seluruh ambang batas yang mungkin. Nilai AUC berkisar dari 0 hingga 1, di mana nilai mendekati 1 menunjukkan model yang sangat baik, sementara nilai 0,5 menunjukkan kinerja yang setara dengan tebakan acak. ROC-AUC bersifat *threshold-independent*, sehingga lebih stabil dibandingkan *accuracy* sebagai indikator kinerja model [30, 31].

2.7.4 PR-AUC (Average Precision)

Precision-Recall AUC (PR-AUC), atau dikenal juga sebagai *Average Precision* (AP), mengukur luas area di bawah kurva *precision-recall* pada berbagai ambang batas. Pada konteks data yang sangat tidak seimbang seperti prediksi

wasting atau *underweight* yang prevalensinya rendah, PR-AUC lebih informatif dibandingkan ROC-AUC karena tidak terpengaruh oleh banyaknya *true negative*, sehingga lebih fokus mencerminkan kemampuan model dalam mengenali kelas minoritas [31].

2.7.5 Pemilihan Ambang Batas (Threshold Tuning)

Secara bawaan, banyak algoritma klasifikasi, termasuk XGBoost, menggunakan ambang batas (*threshold*) sebesar 0,5 untuk mengubah probabilitas prediksi menjadi keputusan biner. Pada data tidak seimbang, ambang batas tersebut sering kali tidak optimal karena model cenderung jarang memprediksi kelas minoritas, sehingga *recall* rendah. Pemilihan ambang batas yang optimal (*threshold tuning*) dilakukan dengan mencari nilai ambang batas yang memaksimalkan metrik yang paling relevan dengan tujuan penelitian, misalnya *F1-score*, *recall*, atau *balanced accuracy* (rata-rata *recall* dan *specificity*). Pada penelitian ini dipilih *balanced accuracy* agar akurasi tidak diperoleh dengan mengorbankan *recall*. Penerapan rinci *threshold tuning* pada penelitian ini, termasuk rentang nilai yang dieksplorasi dan strategi pemilihannya, dibahas pada Bab 3.