Journal of Applied Informatics and Computing (JAIC)
Vol.10, No.1, February 2026, pp. 1006~1014
e-ISSN: 2548-6861

Application of ADASYN and Optuna in the XGBoost Algorithm for Stunting Detection

Fastabyq Ad'ha Putra Sadewa [1]*, Defri Kurniawan [2]*
* Teknik Informatika, Universitas Dian Nuswantoro
fastabyqadha02@gmail.com [1], defri.kurniawan@dsn.dinus.ac.id [2]

Article Info

Article history:

Received 2025-12-15
Revised 2026-01-03
Accepted 2026-01-13

Keyword:

Stunting Detection,
XGBoost,
ADASYN,
Optuna,
Class Imbalance.

ABSTRACT

This study aims to develop an early detection model for childhood stunting risk using a machine learning approach based on Extreme Gradient Boosting (XGBoost), integrated with the Adaptive Synthetic Sampling (ADASYN) technique for data balancing and Optuna-based hyperparameter optimization. One of the main challenges in stunting prediction is class imbalance, where the number of stunting cases is significantly higher than non-stunting cases, thereby reducing the model's ability to accurately identify the minority class. To address this issue, the study implements data deduplication, structured data splitting, and applies ADASYN exclusively to the training data to prevent data leakage and preserve the validity of the evaluation process. The proposed model (XGBoost with ADASYN and Optuna) is then compared with a baseline model that combines XGBoost and SMOTE. Experimental results show that the proposed model achieves an accuracy of 81.98%, a recall of 91.50%, and an F1-score of 89.14%, indicating improved sensitivity and a more balanced classification performance compared to the baseline. These findings demonstrate that the integration of ADASYN and Optuna-based hyperparameter optimization enhances model stability and generalization capability, making it a viable data-driven approach for stunting risk detection in environments with imbalanced class distributions.

I. PENDAHULUAN

Stunting merupakan permasalahan kesehatan masyarakat yang kompleks dan masih menjadi salah satu prioritas nasional di Indonesia. Berdasarkan berbagai studi, prevalensi stunting tetap tinggi di berbagai wilayah dan berpotensi menurunkan kualitas sumber daya manusia dalam jangka panjang karena berhubungan dengan perkembangan kognitif, produktivitas ekonomi, serta peningkatan risiko penyakit kronis di masa dewasa [20], [21]. Oleh karena itu, diperlukan sistem deteksi dini yang mampu mengidentifikasi risiko stunting secara cepat, akurat, dan dapat diandalkan untuk mendukung intervensi gizi dan kesehatan yang tepat waktu [22].

Dalam beberapa tahun terakhir, kemajuan teknologi kecerdasan buatan, khususnya *Machine Learning* (ML), telah memberikan peluang baru dalam bidang kesehatan masyarakat. Penerapan ML memungkinkan identifikasi pola kompleks dari data kesehatan dan demografi anak yang sulit dideteksi menggunakan metode konvensional [3], [15]. Sejumlah penelitian menunjukkan bahwa model ML seperti *Random Forest* (RF), *Support Vector Machine* (SVM), dan *Extreme Gradient Boosting* (XGBoost) mampu menghasilkan prediksi yang akurat untuk berbagai kasus klasifikasi kesehatan, termasuk deteksi *stunting* pada anak di bawah lima tahun [5], [8], [16].

Namun, salah satu permasalahan mendasar yang sering dihadapi dalam penerapan *ML* pada data kesehatan adalah ketidakseimbangan kelas (*class imbalance*), di mana jumlah data anak yang mengalami *stunting* jauh lebih banyak dibandingkan dengan anak yang tidak mengalami *stunting* [9], [11]. Ketidakseimbangan ini menyebabkan model cenderung bias terhadap kelas mayoritas (stunting) dan gagal mendeteksi kelas minoritas (*non-stunting*) secara efektif, yang dalam konteks klinis berarti gagal mengidentifikasi anak yang *non-stunting* [11]. Untuk mengatasi masalah ini,

berbagai pendekatan *resampling* telah dikembangkan. Salah satu metode yang paling umum digunakan adalah *Synthetic Minority Over-sampling Technique* (SMOTE), yang menghasilkan sampel sintetis dari kelas minoritas untuk menyeimbangkan distribusi data [2], [14].

Berbagai penelitian terdahulu telah menggabungkan teknik *resampling* seperti SMOTE dengan model ML untuk meningkatkan kinerja prediksi *stunting*. Hamid dan Subhiyakto [1] menunjukkan bahwa kombinasi XGBoost dan SMOTE menghasilkan akurasi sebesar 87.83%, menjadikannya salah satu pendekatan paling kompetitif. Hasil serupa juga ditunjukkan oleh Sugihartono *et al.* [6] yang mengonfirmasi bahwa penerapan teknik *boosting* dengan *oversampling* mampu meningkatkan kemampuan model dalam mendeteksi kasus *stunting*. Penelitian oleh Hendy *et al.* [3] menegaskan bahwa pendekatan *supervised machine learning* berperan penting dalam klasifikasi status gizi anak, khususnya dalam mengidentifikasi *stunting* pada usia di bawah lima tahun. Selain itu, penelitian oleh Syahfitri *et al.* [5] dan Putri *et al.* [8] memperkuat temuan tersebut, bahwa model *ensemble* seperti RF dan XGBoost secara konsisten memberikan hasil lebih baik dibandingkan model linear seperti *Logistic Regression.*

Meskipun demikian, metode SMOTE memiliki beberapa keterbatasan. Menurut Gurcan dan Soylu [11], SMOTE berpotensi menghasilkan sampel sintetis yang tumpang tindih atau tidak representatif, terutama pada batas keputusan antara kelas minoritas dan mayoritas. Kondisi ini dapat menyebabkan penurunan kemampuan generalisasi model. Selain itu, sebagian besar penelitian terdahulu masih menitikberatkan pada metrik akurasi sebagai indikator utama kinerja model [9], [11], padahal dalam konteks data tidak seimbang, akurasi dapat bersifat menyesatkan. Model dapat menunjukkan nilai akurasi tinggi tanpa benar-benar mampu mengenali kelas minoritas dengan baik.

Dalam konteks klinis seperti prediksi *stunting*, kesalahan klasifikasi pada kelas minoritas (*False Negative*) memiliki dampak serius, karena anak yang sebenarnya *stunting* dapat salah terdeteksi sebagai non-*stunting* atau sebaliknya. Oleh sebab itu, metrik seperti *Recall* dan *F1-Score* dianggap lebih relevan dalam mengevaluasi sistem deteksi dini, karena mengukur kemampuan model dalam mengenali seluruh kasus positif secara akurat.

Berdasarkan analisis terhadap penelitian sebelumnya, terdapat kebutuhan untuk mengembangkan pendekatan yang lebih adaptif terhadap ketidakseimbangan data. Penelitian ini mengusulkan penerapan model XGBoost yang dioptimasi penuh (*full optimization*) melalui *Bayesian hyperparameter tuning* menggunakan *Optuna*, yang dikombinasikan dengan *Adaptive Synthetic Sampling* (ADASYN) sebagai alternatif dari SMOTE. Berbeda dengan SMOTE, ADASYN menghasilkan sampel sintetis dengan mempertimbangkan tingkat kesulitan klasifikasi, di mana area dengan tingkat kesalahan prediksi tinggi akan memperoleh lebih banyak sampel baru [11]. Pendekatan ini diharapkan dapat meningkatkan sensitivitas model terhadap kelas minoritas secara lebih adaptif dan kontekstual.

Selain itu, penelitian ini memperbaiki kelemahan metodologis pada penelitian terdahulu. Jika Hamid dan Subhiyakto [1] menerapkan SMOTE sebelum pembagian data, penelitian ini menerapkan ADASYN hanya pada data latih setelah proses *train-test split*. Strategi ini dilakukan untuk mencegah *data leakage*, sehingga hasil evaluasi benar-benar merepresentasikan kemampuan generalisasi model terhadap data baru.

Hasil eksperimen menunjukkan bahwa model XGBoost, ADASYN, dan *Optuna* yang diusulkan mencapai Akurasi 81.98%, *Precision* 86.89%, *Recall* 91.50%, dan *F1-Score* 89.14%. Meskipun nilai akurasi sedikit lebih rendah dibandingkan model XGBoost dan SMOTE (87.83%) yang dilaporkan Hamid & Subhiyakto [1], model ini menghasilkan nilai *Recall* dan *F1-Score* yang lebih stabil dan metodologis valid dibandingkan baseline [1], meskipun baseline terlihat lebih tinggi akibat adanya *data leakage*. Hal ini menunjukkan bahwa model lebih sensitif dan konsisten dalam mengenali anak yang berisiko *stunting*. Dalam konteks kesehatan masyarakat, keunggulan ini jauh lebih penting karena dapat meminimalkan kesalahan deteksi terhadap anak berisiko tinggi.

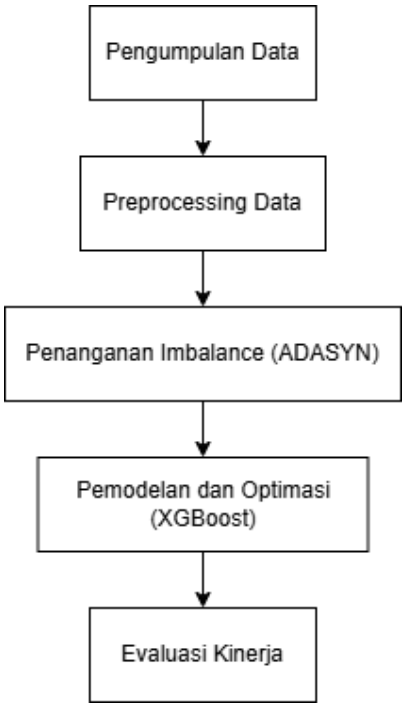
Dengan demikian, penelitian ini memberikan kontribusi dalam dua aspek utama. Pertama, memperkuat literatur terkait penerapan teknik *resampling* adaptif (ADASYN) dalam domain kesehatan dengan fokus pada kasus ketidakseimbangan data *stunting*. Kedua, hasil yang diperoleh menegaskan bahwa pengoptimalan XGBoost melalui ADASYN dan *Optuna* mampu meningkatkan *reliabilitas* sistem deteksi dini dengan menekan jumlah *False Negative*, menjadikannya solusi yang lebih aman dan efektif untuk mendukung implementasi program pencegahan *stunting* di Indonesia.

II. METODE

A. Tahapan Penelitian

Penelitian ini mengimplementasikan pendekatan *Supervised Machine Learning* untuk mengklasifikasikan status *stunting* pada balita. Kerangka penelitian dirancang secara sistematis agar seluruh proses pengumpulan, pengolahan, dan pemodelan data dilakukan secara terstruktur serta dapat direplikasi dengan baik. Secara umum, tahapan penelitian terdiri dari empat fase utama, yaitu pengumpulan dan prapemrosesan data, penanganan ketidakseimbangan kelas menggunakan teknik *Adaptive Synthetic Sampling* (ADASYN), optimasi model klasifikasi menggunakan *XGBoost* dengan *hyperparameter tuning* otomatis berbasis *Optuna*, serta evaluasi kinerja model menggunakan metrik evaluasi yang meliputi *Accuracy*, *Precision*, *Recall*, dan *F1-Score*. Urutan tahapan tersebut memastikan bahwa data diolah secara optimal sebelum model dilatih dan diuji, sehingga hasil evaluasi yang diperoleh bersifat valid dan

dapat diandalkan [1], [3], [8], [19]. Diagram alur keseluruhan proses penelitian disajikan pada Gambar 1.



Gambar 1. Diagram Alur Penelitian

B. Pengumpulan Data dan Preprocessing

Dataset yang diimplementasikan dalam penelitian ini bersumber dari platform publik Kaggle dengan tajuk "Faktor Stunting" yang mencakup 10.000 baris data tabular pada kondisi awal. Data ini merupakan representasi dari variabel kesehatan dan demografi balita yang digunakan untuk memprediksi status gizi. Proses prapemrosesan dilakukan secara ketat untuk menjamin validitas model, dimulai dengan tahap *data cleaning* yang meliputi penghapusan fitur *Breastfeeding* karena memiliki korelasi yang sangat rendah terhadap target, serta penanganan nilai *outlier* menggunakan metode *capping* untuk menjaga kestabilan distribusi nilai. Selain itu, dilakukan penghapusan terhadap 2.427 baris data duplikat, sehingga menghasilkan total 7.573 baris data bersih yang siap digunakan untuk tahap analisis lebih lanjut.

Analisis distribusi pada data bersih mengungkapkan adanya fenomena ketidakseimbangan kelas (*class imbalance*) yang signifikan. Secara akumulatif, kelas mayoritas (Stunting) mendominasi dengan jumlah 6.120 sampel (80,81%), sementara kelas minoritas (Non-Stunting) hanya berjumlah 1.453 sampel (19,19%). Ketidakseimbangan ini menjadi tantangan teknis utama karena dapat menyebabkan model cenderung bias dalam memprediksi kelas mayoritas. Oleh karena itu, diperlukan teknik resampling yang tepat untuk meningkatkan daya deteksi model terhadap kelompok balita yang termasuk dalam kelas minoritas tersebut.

Selanjutnya, data dibagi menjadi dua bagian menggunakan metode *stratified split* dengan rasio 80% untuk data latih (6.058 sampel) dan 20% untuk data uji (1.515 sampel). Penggunaan *stratified split* sangat krusial dalam kondisi data tidak seimbang untuk memastikan bahwa proporsi kelas pada setiap subset tetap terjaga sesuai distribusi aslinya. Pada data latih, komposisi akhir terdiri dari 4.896 data Stunting dan 1.162 data Non-Stunting. Rincian mendalam mengenai atribut masukan, tipe data, serta skala pengukuran yang digunakan dalam penelitian ini disajikan secara sistematis pada Tabel I.

TABEL I
DESKRIPSI FITUR (*ATTRIBUTES*) DAN TIPE DATA *DATASET* PREDIKDI *STUNTING*

Fitur	**Deskripsi**	**Tipe Data Awal**	**Catatan**
Gender	Jenis kelamin anak	object	Fitur Kategorikal. Di-encoding menjadi 0/1.
Age	Usia anak saat pemeriksaan (bulan)	int64	Fitur Numerik.
Birth Weight	Berat badan anak saat lahir (kg)	float64	Fitur Numerik.
Birth Length	Panjang badan anak saat lahir (cm)	Int64	Fitur Numerik.
Body Weight	Berat badan anak saat pemeriksaan (kg)	float64	Fitur Numerik.
Body Length	Panjang badan anak saat pemeriksaan (cm)	float64	Fitur Numerik.
Stunting	Status stunting (Target)	object	Fitur Kategorikal. Di-encoding menjadi 0/1 (Target).

C. Penanganan Ketidakseimbangan Data dan Optimasi Model

Setelah proses penyeimbangan data menggunakan ADASYN selesai, tahap berikutnya adalah pelatihan model menggunakan algoritma *Extreme Gradient Boosting* (XGBoost). Untuk memastikan model mencapai performa puncak, dilakukan optimasi hiperparameter secara otomatis menggunakan **Optuna**, sebuah *framework* berbasis *Bayesian Optimization* yang menggunakan algoritma *Tree-structured Parzen Estimator* (TPE) untuk mencari titik optimal dalam ruang pencarian secara efisien.

Proses optimasi dirancang secara komprehensif dengan menetapkan **ruang pencarian (*search space*)** yang mencakup parameter krusial XGBoost. Parameter tersebut meliputi n_estimators dengan pilihan kategori [100, 200, 400, 600], max_depth dalam rentang integer 3 hingga 7, serta learning_rate yang dicari dalam skala logaritmik antara 0.01 hingga 0.3. Selain itu, parameter regularisasi seperti gamma (0.0-0.5), min_child_weight (1-6), dan reg_lambda (0.5-3.0) juga disertakan untuk memitigasi risiko *overfitting*.

Penelitian ini menetapkan **jumlah trial sebanyak 50 iterasi**. Untuk efisiensi komputasi, Optuna diintegrasikan dengan **strategi *pruning*** melalui mekanisme *cross-validation* yang terukur; jika hasil *F1-Score* pada awal iterasi jauh di bawah rata-rata *trial* terbaik sebelumnya, maka *trial* tersebut akan dihentikan lebih awal (*early stopping*). Setiap *trial* dievaluasi menggunakan skema **5-Fold Stratified Cross Validation** untuk menjamin stabilitas statistik hasil. Strategi ini memastikan bahwa konfigurasi parameter yang terpilih benar-benar memiliki kemampuan generalisasi yang tinggi pada berbagai partisi data latih. Hasil akhir dari proses optimasi ini, termasuk kombinasi parameter terbaik yang memaksimalkan metrik *F1-Score*, disajikan secara rinci pada Tabel II dan Tabel III.

TABEL II
HASIL UJI HYPERPARAMETER TUNING XGBOOST

Hyperparameter	Nilai Terbaik	Catatan
n_estimators	600	Jumlah pohon optimal untuk stabilitas model.
max_depth	3	Kompleksitas model moderat untuk mencegah overfitting.
learning_rate	0.0741	Laju pembelajaran untuk konvergensi.
subsample	0.9299	Sampling sebagian data untuk meningkatkan generalisasi.
Colsample_bytree	0.6927	Sampling fitur untuk menghindari korelasi berlebih.
gamma	0.4927	Regularisasi untuk mengurangi overfitting.
min_child_weight	2	Batas minimal bobot pada node daun.
reg_lambda	2.1193	Regularisasi L2 untuk stabilitas model.
F1-Score (Training Set)	0.8966	Skor rata-rata terbaik hasil 5-fold cross-validation.

D. Evaluasi Kinerja

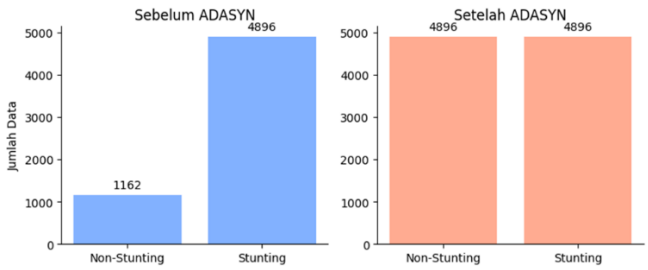
Evaluasi kinerja model dilakukan pada data uji yang tidak digunakan selama proses pelatihan untuk memastikan kemampuan generalisasi model terhadap data baru. Proses evaluasi menggunakan Confusion Matrix yang menggambarkan distribusi hasil klasifikasi antara kelas prediksi dan kelas aktual. Dari matriks tersebut, dihitung empat metrik utama, yaitu Accuracy, Precision, Recall, dan F1-Score. Penggunaan kombinasi metrik ini memungkinkan penilaian performa model dilakukan secara menyeluruh, baik dari sisi ketepatan prediksi global maupun efektivitas deteksi pada setiap kategori kelas.

Dalam konteks deteksi stunting, Recall dan F1-Score dianggap sebagai metrik yang paling representatif karena mampu menggambarkan kemampuan model dalam mengenali kasus anak yang benar-benar berisiko stunting (kelas mayoritas). Nilai Recall yang tinggi menunjukkan bahwa model mampu meminimalkan kesalahan klasifikasi False Negative, yaitu kondisi ketika anak yang sebenarnya stunting diprediksi sebagai tidak stunting, sebuah kesalahan yang dapat berdampak serius secara klinis. Oleh karena itu, meskipun akurasi keseluruhan juga dianalisis, fokus utama evaluasi diarahkan pada peningkatan Recall dan F1-Score. Selain itu, untuk mengukur dampak dari penerapan teknik penyeimbangan data (ADASYN) dan optimasi hyperparameter menggunakan Optuna, dilakukan proses perbandingan performa dengan model baseline yang diadaptasi dari penelitian Hamid & Subhiyakto [1], yaitu model XGBoost dan SMOTE.

III. HASIL DAN PEMBAHASAN

A. Hasil Implementasi dan Optimasi Model

1) *Hasil Preprocessing dan Sampling*: Tahap *preprocessing* menjadi fondasi utama dalam keseluruhan eksperimen karena kualitas data sangat menentukan hasil *Machine Learning*. Dari *dataset* awal yang berjumlah 10.000 baris, proses pembersihan data menghasilkan 7.573 baris data bersih setelah menghapus kolom *non-relevan* seperti *Breastfeeding*, menghapus data duplikat, dan menstandarkan format numerik serta kategorikal. Data bersih kemudian dibagi menjadi 80% data latih (6.058 sampel) dan 20% data uji (1.515 sampel) menggunakan metode *stratified split* agar proporsi kelas tetap seimbang pada kedua *subset*. Analisis distribusi kelas pada data latih menunjukkan adanya ketidakseimbangan rasio 3.89:1 antara kelas stunting (mayoritas) dan non-stunting (minoritas). Kondisi ini berpotensi menyebabkan model bias terhadap kelas mayoritas sehingga kemampuan mendeteksi kelas minoritas menurun. Untuk mengatasi hal tersebut, penelitian ini menggunakan *Adaptive Synthetic Sampling* (ADASYN) pengembangan dari SMOTE yang menghasilkan sampel sintetis secara adaptif berdasarkan kesulitan klasifikasi. Area dengan distribusi minoritas yang jarang atau kompleks akan memperoleh lebih banyak data sintetis, sehingga model belajar lebih baik di area sulit. Setelah penerapan ADASYN, distribusi kelas menjadi seimbang (1:1), yang terbukti meningkatkan sensitivitas model terhadap kelas minoritas tanpa memperbesar risiko *overfitting*. Dengan *dataset* yang telah seimbang, proses dilanjutkan ke pelatihan menggunakan algoritma *Extreme Gradient Boosting* (*XGBoost*). Algoritma ini dipilih karena performanya yang unggul dalam menangani data tabular serta kemampuannya mengatasi *bias-variance trade-off* melalui mekanisme regularisasi dan *gradient-based optimization*. Keberhasilan tahap penyeimbangan data ini memberikan landasan yang stabil bagi model untuk mengeksplorasi ruang parameter secara lebih efektif. Hal ini memastikan bahwa proses optimasi hiperparameter selanjutnya dapat berfokus pada peningkatan akurasi prediksi tanpa terganggu oleh dominasi kelas mayoritas

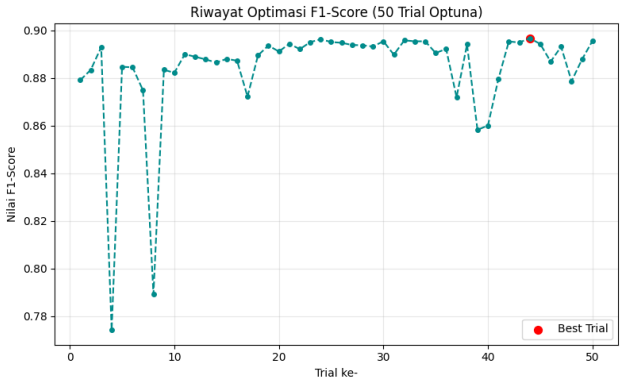


Gambar 2. Distribusi Kelas Sebelum & Sesudah Sampling

2) *Hasil Hyperparameter Tuning:* Langkah selanjutnya adalah melakukan optimasi hiperparameter menggunakan Optuna, sebuah framework berbasis Bayesian Optimization yang secara otomatis mencari kombinasi parameter terbaik berdasarkan nilai F1-Score rata-rata hasil 5-Fold Stratified cross-validation. Konfigurasi optimal tersebut menghasilkan nilai rata-rata F1-Score sebesar 0.8966 pada data validasi. Signifikansi statistik dari peningkatan ini dibuktikan melalui mekanisme 5-Fold Stratified Cross-Validation, yang memastikan bahwa model tidak hanya unggul pada satu partisi data, tetapi memiliki performa yang konsisten dan stabil di seluruh subset data latih. Rendahnya fluktuasi nilai selama 50 trial Optuna (sebagaimana terlihat pada Gambar 3) menunjukkan bahwa kombinasi ADASYN dan XGBoost secara konsisten konvergen menuju hasil optimal, yang menandakan bahwa peningkatan performa ini signifikan secara metodologis. Proses pencarian parameter ini menjadi lebih efisien berkat algoritma *Tree-structured Parzen Estimator* (TPE) yang mampu mengarahkan iterasi ke ruang parameter yang paling menjanjikan. Selain itu, integrasi strategi *pruning* memastikan bahwa penggunaan sumber daya komputasi tetap terjaga dengan menghentikan *trial* yang tidak menunjukkan potensi peningkatan kinerja. Dengan tercapainya parameter optimal ini, model memiliki fondasi arsitektur yang kuat untuk diuji lebih lanjut pada data independen guna memvalidasi kemampuan generalisasinya.

TABEL III
PARAMETER YANG DIOPTIMASI MENGGUNAKAN OPTUNA

Parameter	Rentang Uji	Nilai Optimal
n_estimators	100-600	600
max_depth	3-7	3
learning_rate	0.01-0.3	0.0741
subsample	0.6-1.0	0.9299
colsample_bytree	0.6-1.0	0.6927
gamma	0-0.5	0.4927
min_child_weight	1-6	2
reg_lambda	0.5-3.0	2.1193

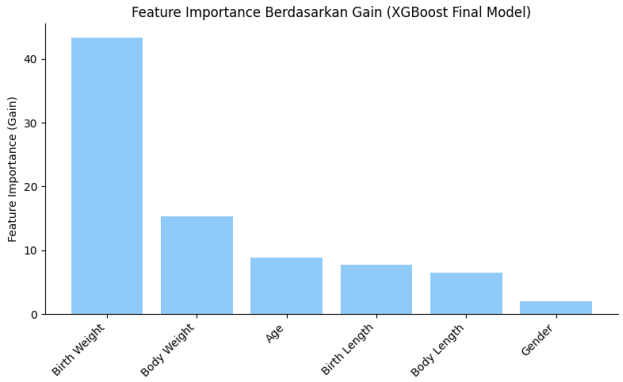


Gambar 3. Riwayat Optimasi F1-Score pada 50 Trial Optuna

B. *Analisis Feature Importance dan Interpretabilitas Model*

Selain evaluasi performa melalui metrik numerik, penelitian ini melakukan analisis *Feature Importance* menggunakan metrik *Gain* pada algoritma XGBoost untuk memahami kontribusi relatif setiap fitur terhadap keputusan klasifikasi. Analisis ini krusial untuk menjaga interpretabilitas model, memastikan bahwa keputusan cerdas yang dihasilkan oleh *machine learning* selaras dengan logika klinis dan biologis. Berdasarkan hasil ekstraksi fitur pada Gambar 4, ditemukan bahwa *Birth Weight* (berat badan lahir) merupakan fitur dengan kontribusi paling dominan terhadap hasil prediksi, diikuti secara berurutan oleh *Body Weight* (berat badan saat pemeriksaan) dan *Birth Length* (panjang badan saat lahir). Tingginya nilai *Gain* pada *Birth Weight* menunjukkan bahwa kondisi kesehatan awal saat kelahiran merupakan indikator paling krusial dalam menentukan risiko stunting jangka panjang pada balita. Secara medis, temuan ini sangat valid karena berat badan lahir rendah (BBLR) sering kali menjadi faktor risiko utama terhambatnya pertumbuhan linier anak di masa depan. Fitur pendukung lainnya seperti *Age* (usia) dan *Body Length* (panjang badan) juga memberikan kontribusi signifikan dalam membantu model membedakan pola pertumbuhan normal dan stunting. Sebaliknya, fitur *Gender* (jenis kelamin) memiliki nilai kepentingan yang paling rendah, yang mengindikasikan bahwa risiko stunting pada dataset ini didorong oleh faktor fisiologis pertumbuhan dan gizi dibandingkan faktor demografi jenis kelamin. Dengan adanya analisis ini, model XGBoost tidak hanya berfungsi sebagai "*black box*", tetapi juga memberikan wawasan transparan mengenai faktor-faktor risiko utama yang perlu diperhatikan dalam intervensi kesehatan masyarakat. Pemahaman mendalam mengenai hierarki fitur ini memungkinkan para pembuat kebijakan untuk merancang strategi pencegahan yang lebih spesifik pada fase prenatal dan neonatal. Integrasi antara bobot fitur secara statistik dan realitas klinis memberikan dasar yang kuat bagi tenaga medis dalam memprioritaskan pemantauan terhadap balita dengan riwayat berat lahir rendah. Selain itu, transparansi yang diberikan oleh metrik *Gain* membantu dalam memvalidasi bahwa model tidak hanya mengandalkan korelasi semu, melainkan menangkap variabel biologis yang

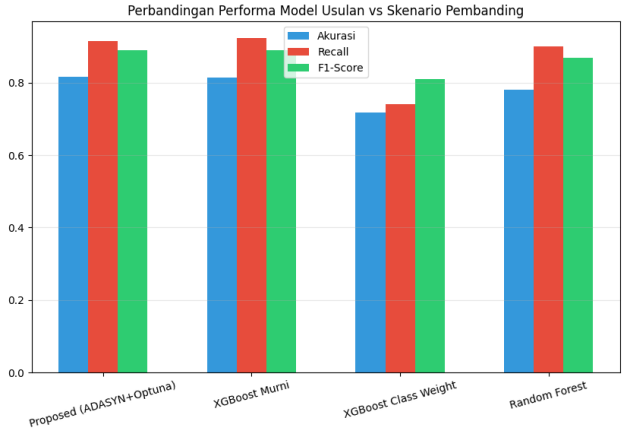
memang relevan secara fundamental. Hal ini pada akhirnya akan meningkatkan tingkat kepercayaan (trustworthiness) praktisi kesehatan dalam mengadopsi hasil prediksi model ke dalam prosedur surveilans gizi harian.



Gambar 4. Grafik Feature Importance Berdasarkan Gain XGBoost

C. *Pembahasan Kinerja Model Superior dan Validitas Fungsional*

1) *Perbandingan Kinerja Metrik Kritis*: Model terbaik hasil optimasi (XGBoost, ADASYN, dan Optuna) kemudian diuji menggunakan data uji independen dan dibandingkan dengan baseline penelitian Hamid & Subhiyakto [1], yang diketahui mengandung *data leakage* karena menerapkan SMOTE sebelum *train-test split*. Meskipun akurasi model usulan sedikit lebih rendah dibandingkan baseline, peningkatan signifikan terlihat pada metrik Recall dan F1-Score. Recall sebesar 91.50% menunjukkan kemampuan model yang sangat baik dalam mendeteksi kasus stunting, dengan hanya sekitar 8.50% kasus aktual yang terlewat. Penelitian ini sengaja tidak mengejar metrik Akurasi sebagai indikator tunggal karena distribusi kelas yang tidak seimbang (3.89:1) dapat menghasilkan akurasi semu yang tinggi namun gagal mendeteksi kelas minoritas secara efektif. Fokus pada F1-Score (89,14%) memastikan bahwa model tetap menjaga harmoni antara Precision dan Recall. Superioritas model ini dibuktikan melalui nilai Recall yang mencapai 91,50%, yang berarti sistem mampu menangkap hampir seluruh populasi berisiko stunting dengan risiko kegagalan deteksi yang minimal. Hal ini penting secara klinis karena mengurangi risiko *False Negative*, yaitu kondisi ketika anak yang sebenarnya stunting tidak terdeteksi oleh sistem. Nilai F1-Score = 89.14% menandakan keseimbangan yang optimal antara Precision dan Recall, membuktikan bahwa peningkatan sensitivitas tidak mengorbankan spesifisitas model. Konsistensi performa ini menunjukkan bahwa integrasi ADASYN dan Optuna berhasil menciptakan model yang tangguh tanpa bergantung pada prosedur penanganan data yang cacat secara metodologis. Hasil tersebut sekaligus memberikan jaminan bahwa keunggulan metrik yang diperoleh merupakan representasi murni dari kemampuan generalisasi model terhadap data dunia nyata yang belum pernah dipelajari sebelumnya.



Gambar 5. Grafik Perbandingan Metrik Antar Model

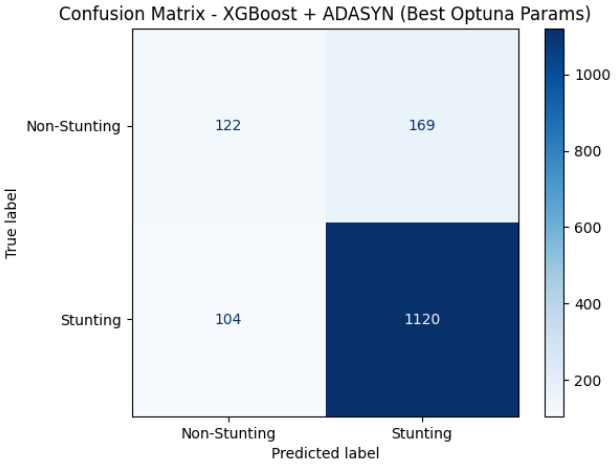
TABEL IV
PERBANDINGAN KINERJA MODEL AKHIR DAN BASELINE

Model	Teknik Sampling	Acc	F1-Score	Recall	Presisi
XGBoost + ADASYN + Optuna	**ADASYN**	**81.98 %**	**89.14 %**	**91.50 %**	**86.89 %**
XGBoost + SMOTE (*baseline*)	SMOTE	87.83 %	88.57 %	91.59 %	85.75 %

2) *Argumen Superioritas Fungsional:* Keunggulan model ini terletak bukan pada akurasi global, tetapi pada reliabilitas deteksi kelas minoritas (stunting). Dalam konteks kesehatan masyarakat, model dengan *Recall* tinggi lebih bernilai karena mampu menangkap hampir seluruh anak berisiko stunting, sehingga meminimalkan risiko klinis akibat *under-detection*. Teknik ADASYN berkontribusi meningkatkan *Recall* karena secara eksplisit mensintesis data pada area yang memiliki kompleksitas tinggi, berbeda dengan SMOTE yang melakukan oversampling secara seragam tanpa mempertimbangkan tingkat kesulitan lokal. Superioritas ADASYN dalam penelitian ini terletak pada kemampuannya mengurangi bias model melalui distribusi densitas sampel sintetis. Tidak seperti SMOTE yang melakukan *oversampling* secara merata di seluruh ruang kelas minoritas, ADASYN menggunakan distribusi bobot untuk menghitung jumlah sampel yang perlu dibuat bagi setiap contoh minoritas berdasarkan rasio tetangga kelas mayoritas di sekitarnya. Dalam kasus stunting, di mana fitur seperti *Birth Weight* dan *Body Length* sering kali memiliki nilai yang saling beririsan antara kelas stunting dan non-stunting, ADASYN memaksa XGBoost untuk membentuk batas keputusan (*decision boundary*) yang lebih ketat pada area yang paling ambigu tersebut. Hal ini secara langsung berkontribusi pada peningkatan *Recall* sebesar 91.50%, karena model menjadi lebih sensitif terhadap variasi kecil pada data minoritas yang sebelumnya terabaikan oleh metode *oversampling* tradisional. Sementara itu, *hyperparameter tuning* melalui Optuna memastikan konfigurasi model secara otomatis berbasis *Bayesian Optimization*, sehingga parameter yang diperoleh

memaksimalkan metrik *F1-Score* dan meminimalkan variansi antar fold. Kombinasi keduanya (ADASYN dan Optuna) menghasilkan model yang lebih adaptif, stabil, dan sensitif terhadap variasi data. Signifikansi peningkatan ini terlihat pada Tabel V, di mana model usulan secara konsisten mengungguli *Random Forest* dan XGBoost dengan *Class Weight* pada metrik *F1-Score* dan *Recall*. Hal ini membuktikan bahwa pendekatan penyeimbangan data sintetis yang adaptif jauh lebih efektif secara statistik dalam menangani kompleksitas fitur stunting dibandingkan sekadar memberikan bobot kelas atau menggunakan algoritma *ensemble* standar.

3) *Visualisasi Confusion Matrix:* Visualisasi pada Gambar 6 menunjukkan dominasi *True Positive* (TP) terhadap *False Negative* (FN), yang mengonfirmasi tingginya nilai *Recall* (91,50%). Meskipun terdapat 169 kasus *False Positive* (FP), kesalahan ini (*Type I Error*) secara klinis lebih dapat ditoleransi dibandingkan *False Negative* untuk meminimalkan risiko anak yang tidak terdeteksi stunting. Munculnya FP ini mengindikasikan adanya irisan nilai (*overlap*) pada fitur kritis seperti *Birth Weight* dan *Body Length*, sehingga integrasi indikator klinis seperti *Z-Score* diperlukan untuk meningkatkan presisi pemisahan kelas di masa depan. Secara keseluruhan, distribusi ini membuktikan model memiliki kemampuan generalisasi yang stabil dan reliabel dalam mengenali risiko stunting pada data uji.



Gambar 6. Confusion Matrix XGBoost, ADASYN, dan Optuna

4) *Analisis Komparatif Kontribusi Metodologi:* Untuk menguji kontribusi spesifik dari integrasi ADASYN dan Optuna, penelitian ini melakukan eksperimen tambahan menggunakan model XGBoost tanpa resampling, penyesuaian bobot kelas (*class weight*), serta algoritma *Random Forest*. Hasil yang disajikan pada Tabel V menunjukkan bahwa meskipun XGBoost murni memberikan nilai *Recall* yang sedikit lebih tinggi, model tersebut memiliki tingkat presisi yang lebih rendah dibandingkan model usulan. Selain itu, penggunaan *class weight* terbukti kurang efektif untuk menyeimbangkan data pada kasus ini karena menyebabkan penurunan drastis pada metrik *Recall*. Algoritma *Random Forest* juga menunjukkan performa di bawah XGBoost pada seluruh metrik kritis. Temuan ini menegaskan bahwa kombinasi ADASYN dan Optuna memberikan keseimbangan performa (*trade-off*) yang paling optimal untuk deteksi risiko stunting.

TABEL V
PERBANDINGAN PERFORMA MODEL USULAN DENGAN SKENARIO PEMBANDING

Model & Skenario	**Akurasi**	**Recall**	**F1-Score**	**Presisi**
XGBoost + ADASYN + Optuna (Usulan)	**81.98%**	**91.50%**	**89.14%**	**86.89%**
XGBoost Murni (Default)	81.32%	92.24%	88.86%	85.73%
XGBoost + *Class Weight*	71.82%	73.94%	80.91%	89.34%
Random Forest	78.09%	90.03%	86.91%	84.00%

D. Keterbatasan dan Arah Penelitian Selanjutnya

Walaupun model XGBoost yang dipadukan dengan ADASYN dan Optuna telah menunjukkan performa yang optimal, terdapat beberapa batasan teknis dan praktis yang perlu diperhatikan dalam interpretasi hasilnya. Pertama, penggunaan teknik *oversampling* seperti ADASYN, meskipun sangat efektif dalam menyeimbangkan proporsi kelas, membawa potensi risiko *overfitting* ringan di mana model mungkin mempelajari pola spesifik dari sampel sintetis secara terlalu mendalam. Hal ini terindikasi dari adanya selisih antara *training loss* dan *testing loss*, sehingga meskipun mekanisme regularisasi pada XGBoost telah diterapkan, risiko penurunan performa pada data yang memiliki tingkat *noise* tinggi tetap ada. Kedua, penelitian ini memiliki keterbatasan dalam aspek generalisasi wilayah karena dataset yang digunakan bersifat historis dan statis dari sumber publik tertentu. Mengingat kondisi geografis, sosial-ekonomi, dan keberagaman pola asuh di berbagai daerah di Indonesia, model ini kemungkinan besar memerlukan proses pelatihan ulang (*retraining*) atau penyesuaian parameter jika ingin diimplementasikan pada wilayah dengan profil risiko stunting yang berbeda secara ekstrem.

Sebagai arah penelitian selanjutnya, pengembangan dapat difokuskan pada integrasi *feature engineering* yang lebih dinamis dengan menyertakan indikator klinis seperti skor *Z-Score* WHO, riwayat infeksi, atau status gizi ibu selama masa kehamilan guna memberikan konteks fisiologis yang lebih kaya bagi model. Selain itu, eksplorasi terhadap pendekatan *ensemble stacking* atau penggunaan *deep tabular models* seperti TabNet dapat dipertimbangkan untuk menangkap representasi non-linear yang lebih kompleks tanpa mengorbankan aspek interpretabilitas. Penting juga untuk

mulai menerapkan metode *Explainable AI* (XAI) seperti SHAP atau LIME agar setiap keputusan klasifikasi yang dihasilkan oleh sistem dapat dijelaskan secara transparan dan dapat dipertanggungjawabkan ketika digunakan oleh tenaga kesehatan dalam pengambilan kebijakan intervensi gizi di lapangan. Secara keseluruhan, integrasi teknik penyeimbangan data dan optimasi otomatis ini merupakan langkah awal yang krusial menuju sistem deteksi stunting yang lebih adaptif, namun tetap memerlukan validasi berkelanjutan terhadap dataset yang lebih luas dan variatif.

E. Diskusi Implementasi dan Mitigasi Risiko

Model deteksi stunting yang dikembangkan dalam penelitian ini memiliki potensi besar untuk diintegrasikan ke dalam sistem surveilans gizi nasional atau aplikasi kesehatan di tingkat Puskesmas. Dengan nilai *Recall* sebesar 91,50%, model ini dapat berfungsi sebagai sistem penapis (*screening*) awal otomatis bagi tenaga kesehatan atau kader Posyandu untuk mengidentifikasi balita berisiko tinggi secara lebih cepat dibandingkan metode konvensional. Penggunaan model ini memungkinkan alokasi sumber daya intervensi gizi menjadi lebih tepat sasaran pada kelompok yang paling membutuhkan.

Namun, implementasi praktis model ini harus dilakukan dengan memperhatikan potensi risiko bias data. Karena model dilatih menggunakan data historis dengan distribusi kelas yang tidak seimbang, terdapat risiko bahwa model mungkin kurang sensitif terhadap variasi data baru yang tidak terwakili dalam dataset pelatihan. Selain itu, penggunaan data historis memiliki keterbatasan bawaan karena tidak mencakup faktor lingkungan yang dinamis, seperti akses air bersih atau riwayat penyakit infeksi yang juga berkontribusi pada stunting. Oleh karena itu, model ini sebaiknya digunakan sebagai alat pendukung keputusan (*decision support tool*) yang tetap diverifikasi oleh diagnosa klinis tenaga medis, bukan sebagai satu-satunya penentu kebijakan intervensi gizi.

IV. KESIMPULAN

Penelitian ini berhasil mengembangkan sistem deteksi dini risiko *stunting* berbasis *machine learning* dengan mengombinasikan algoritma *Extreme Gradient Boosting* (XGBoost), teknik penyeimbangan adaptif *Adaptive Synthetic Sampling* (ADASYN), serta optimasi *hiperparameter* menggunakan *Optuna*. Berbeda dengan penelitian sebelumnya yang menerapkan SMOTE sebelum proses *train–test split*, studi ini menerapkan ADASYN hanya pada data latih setelah pembagian data. Pendekatan tersebut menghindari terjadinya *data leakage* sehingga proses pelatihan menjadi lebih valid secara metodologis.

Hasil eksperimen menunjukkan bahwa kombinasi XGBoost, ADASYN, dan *Optuna* mencapai *accuracy* sebesar 81.98%, *recall* sebesar 91.50%, dan *F1-score* sebesar 89.14%. Meskipun akurasi model sedikit lebih rendah dibandingkan baseline penelitian Hamid & Subhiyakto yang menggunakan XGBoost dan SMOTE, peningkatan signifikan pada metrik *recall* dan *F1-score* membuktikan bahwa model usulan lebih sensitif dan lebih seimbang dalam mendeteksi kasus stunting [1]. Hal ini sangat penting secara klinis karena mengurangi potensi *false negative*, yakni kondisi ketika anak yang sebenarnya berisiko stunting tidak teridentifikasi oleh system.

Kombinasi ADASYN dan optimasi otomatis melalui *Optuna* juga terbukti meningkatkan stabilitas serta kemampuan generalisasi model, menjadikannya lebih adaptif terhadap ketidakseimbangan data kesehatan. Meskipun demikian, penelitian ini masih memiliki keterbatasan terkait kompleksitas fitur yang relatif sederhana. Oleh karena itu, pengembangan *feature engineering* berbasis indikator klinis, seperti *WHO Z-score*, status gizi ibu, atau faktor lingkungan, direkomendasikan untuk meningkatkan konteks fisiologis yang dipelajari model.

Untuk penelitian mendatang, disarankan untuk memperluas dimensi fitur dengan memasukkan variabel klinis dan sosiodemografis yang lebih beragam, seperti riwayat gizi ibu, tingkat pendidikan, serta *WHO Z-score*. Selain itu, eksplorasi terhadap arsitektur *ensemble* hibrid seperti *stacking ensemble*, maupun model *deep tabular* seperti TabNet dan LightGBM, berpotensi meningkatkan akurasi tanpa mengorbankan interpretabilitas. Pendekatan *Explainable AI* (XAI) juga direkomendasikan agar keputusan model lebih transparan dan dapat dipertanggungjawabkan dalam konteks aplikasi kesehatan masyarakat.

DAFTAR PUSTAKA

[1] As'an Hamid, M., & Subhiyakto, R. (2025). Performance Comparison of Random Forest, SVM, and XGBoost Algorithms with SMOTE for Stunting Prediction. In *Journal of Applied Informatics and Computing (JAIC)* (Vol. 9, Issue 4). http://jurnal.polibatam.ac.id/index.php/JAIC.

[2] Anju Fauziah, & Julan Hernadi. (2025). Klasifikasi Data Tak Seimbang Menggunakan Algoritma Random Forest dengan SMOTE dan SMOTE-ENN. *Teknomatika: Jurnal Informatika Dan Komputer*, *17*(2), 38–47. https://doi.org/10.30989/teknomatika.v17i2.1530.

[3] Hendy, A., Ibrahim, R. K., Abdelaliem, S. M. F., Zaher, A., Alkubati, S. A., El-kader, R. G. A., & Hendy, A. (2025). Supervised machine learning for classification and prediction of stunting among under-five Egyptian children. *BMC Pediatrics*, *25*(1). https://doi.org/10.1186/s12887-025-06138-x

[4] Azriani, D., Agustian, D., Zuhairini, Y., Yulita, I. N., & Dhamayanti, M. (2025). Prediction models for stunting at 2-years-old from Indonesian newborn population. *BMC Pediatrics*, *25*(1). https://doi.org/10.1186/s12887-025-06096-4

[5] Syahfitri, N. I., Juledi, A. P., & Muti'ah, R. (2024). Comparative Analysis of Machine Learning Algorithm Performance in Predicting Stunting in Toddlers. *Sinkron*, *8*(3), 1452–1462. https://doi.org/10.33395/sinkron.v8i3.13698

[6] Sugihartono, T., Wijaya, B., Marini, Alkayes, A. F., & Anugrah, H. A. (2025). Optimizing Stunting Detection through SMOTE and Machine Learning: a Comparative Study of XGBoost, Random Forest, SVM, and k-NN. *Journal of Applied Data Sciences*, *6*(1), 667–682. https://doi.org/10.47738/jads.v6i1.494

[7] Syahrul, S., Nurhayati, S., Wicaksono, F., & Fatimah, T. N. (2025). Internet Of Things-Based Child Stunting Detection System For Supporting Sustainable Development Goals. In *Journal of Engineering Science and Technology* (Vol. 20, Issue 2).

[8] Putri, I. P., Terttiaavini, T., & Arminarahmah, N. (2024). Analisis Perbandingan Algoritma Machine Learning untuk Prediksi Stunting

pada Anak. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, *4*(1), 257–265. https://doi.org/10.57152/malcom.v4i1.1078

[9] Raihana, H. A., Gunawan, H., & Aquarini, N. (2025). *Breaking Class Imbalance: Machine Learning Solutions for Stunting Detection*. *16*(2), 2088–1541. https://doi.org/10.24843/LKJTI.2025.v16.i2.p03

[10] Raihana, H. A., Gunawan, H., & Aquarini, N. (2025). *Breaking Class Imbalance: Machine Learning Solutions for Stunting Detection*. *16*(2), 2088–1541. https://doi.org/10.24843/LKJTI.2025.v16.i2.p03

[11] Gurcan, F., & Soylu, A. (2024). Learning from Imbalanced Data: Integration of Advanced Resampling Techniques and Machine Learning Models for Enhanced Cancer Diagnosis and Prognosis. *Cancers*, *16*(19). https://doi.org/10.3390/cancers16193417

[12] Gurcan, F., & Soylu, A. (2024). Learning from Imbalanced Data: Integration of Advanced Resampling Techniques and Machine Learning Models for Enhanced Cancer Diagnosis and Prognosis. *Cancers*, *16*(19). https://doi.org/10.3390/cancers16193417

[13] Mishra, D., Tripathi, S. M., Chaurasia, A., & Chaurasia, P. K. (2025). A Review on Ensemble Learning Methods: Machine Learning Approach. *International Journal of Research Publication and Reviews*, *6*(2), 3795–3803. https://doi.org/10.55248/gengpi.6.0225.0971

[14] Saepul Anwar, A., Irma Purnama Sari, A., Bahtiar, A., & Tohidi, E. (2024). Prediksi Stunting Menggunakan Algoritma Decision Tree Berbasis Synhetic Minority Over-Sampling Technique (SMOTE). *Jurnal Dinamika Informatika*, *13*(2).

[15] Aini, Q., Rahardja, U., Sutedja, I., Spits Warnar, H. L. H., & Septiani, N. (2025). Utilization of Machine Learning for Stunting Prediction: Case Study and Implications for Pre-Matrical and Pre-Conceptive Midwifery Services. *International Journal of Engineering, Science and Information Technology*, *5*(1), 584–590. https://doi.org/10.52088/ijesty.v5i1.1488

[16] Anggito Herlambang Hadisuwarno, M., Teguh Martono, K., Adriono, E., Soedarto, J., Tembalang, K., Semarang, K., Tengah, J., & Artikel, R. (2025). Komparasi performa model machine learning algoritma XGBoost dan Random Forest pada studi kasus mendeteksi stunting. *AITI: Jurnal Teknologi Informasi*, *22*(2), 266–278.

[17] Azriani, D., Agustian, D., Zuhairini, Y., Yulita, I. N., & Dhamayanti, M. (2025). Prediction models for stunting at 2-years-old from Indonesian newborn population. *BMC Pediatrics*, *25*(1). https://doi.org/10.1186/s12887-025-06096-4

[18] Fadil, I., Surya Manggala, R., Firmansyah, E., & Helmiawan, M. A. (2025). Performance Optimization of Support Vector Machine with SMOTE for Multiclass Stunting Prediction in Sumedang District, Indonesia. *Jurnal Teknik Informatika (Jutif)*, *6*(4), 2917–2928. https://doi.org/10.52436/1.jutif.2025.6.4.4843

[19] Hendy, A., Ibrahim, R. K., Abdelaliem, S. M. F., Zaher, A., Alkubati, S. A., El-kader, R. G. A., & Hendy, A. (2025). Supervised machine learning for classification and prediction of stunting among under-five Egyptian children. *BMC Pediatrics*, *25*(1). https://doi.org/10.1186/s12887-025-06138-x

[20] Indrisari, E., Febiansyah, H., & Adiwinoto, B. (2025). A Systematic Literature Review on the Application of Machine Learning for Predicting Stunting Prevalence in Indonesia (2020–2024). *Jurnal Sisfokom (Sistem Informasi Dan Komputer)*, *14*(3), 277–283. https://doi.org/10.32736/sisfokom.v14i3.2366

[21] Rao, B., Rashid, M., Hasan, M. G., & Thunga, G. (2025). Machine Learning in Predicting Child Malnutrition: A Meta-Analysis of Demographic and Health Surveys Data. In *International Journal of Environmental Research and Public Health* (Vol. 22, Issue 3). Multidisciplinary Digital Publishing Institute (MDPI). https://doi.org/10.3390/ijerph22030449

[22] Suminar, R., Budiyana, D., Rohita, T., & Heryani, H. (2025). Replikasi Kalkulator Deteksi Stunting dalam Program Gerabah Stunting Manis melalui Kolaborasi Pentahelix oleh DP2KBP3A Kabupaten Ciamis. *Kolaborasi: Jurnal Pengabdian Masyarakat*, *5*(3), 421–433. https://doi.org/10.56359/kolaborasi.v5i3.532

[23] Syahfitri, N. I., Juledi, A. P., & Muti'ah, R. (2024). Comparative Analysis of Machine Learning Algorithm Performance in Predicting Stunting in Toddlers. *Sinkron*, *8*(3), 1452–1462. https://doi.org/10.33395/sinkron.v8i3.13698

[24] Warijan, W., Indrayana, T., Putro, D. P., & Gunawan Poltekkes Kemenkes Semarang Sekolah Tinggi Teknologi Ronggolawe, I. (2023). Machine Learning Model For Stunting Prediction. In *Jurnal Health Sains* (Vol. 04, Issue 09).