

## A logistic regression-based risk screening framework for early malnutrition detection in children under five

Adebanjo Adedoyin S.<sup>1</sup>, Anyahuru Oluebube Daniel <sup>1</sup>, Eguasa Abieyuwa Jessica <sup>1</sup>, Njoku Kelechi David <sup>1</sup>, Awoniyi Amos Tolulope <sup>1,\*</sup>, Olawunmi Asake Adebanjo <sup>1</sup> and Adewuyi Oluwaseyi J. <sup>2</sup>

<sup>1</sup> Department of Software Engineering, Babcock University, Ilishan-Remo, Ogun State, Nigeria.

<sup>2</sup> Department of Computer Science, Babcock University, Ilishan-Remo, Ogun State, Nigeria.

Global Journal of Engineering and Technology Advances, 2026, 28(03), 271–279

Publication history: Received on 23 May 2026; revised on 07 July 2026; accepted on 09 July 2026

Article DOI: <https://doi.org/10.30574/gjeta.2026.28.3.0172>

### Abstract

Malnutrition among children under five years of age remains a leading cause of preventable morbidity and mortality in Nigeria and across low- and middle-income countries. Existing detection methods are resource-intensive, dependent on trained clinical personnel, and largely inaccessible in the underserved communities where malnutrition burden is highest. This paper presents the design, implementation, and evaluation of a cloud-deployed, supervised machine learning system for early malnutrition risk screening. A Logistic Regression model was trained on 50,000 synthetic pediatric samples generated using WHO West African growth reference values, incorporating 12 features encompassing anthropometric measurements, WHO-derived z-scores (HAZ, WAZ, WHZ), clinical status flags, and socioeconomic indicators. The system was evaluated using 5-fold stratified cross-validation and achieved an Area Under the Receiver Operating Characteristic Curve (AUC) of  $0.7035 \pm 0.0256$ , demonstrating acceptable discriminative ability for a community-level prototype screening tool. The deployed web application delivers single-patient risk classifications (HIGH, MODERATE, or LOW) in under 500ms and supports batch processing of up to 10,000 records via CSV upload. The system's architecture — Flask 3.x backend on Render, React 18 frontend on Netlify — operates within free-tier cloud infrastructure, making it viable for deployment without capital expenditure. Results demonstrate the system's acceptability as an interpretable, low-cost decision-support tool for community health workers operating in resource-limited settings.

**Keywords:** Child Health; Early Screening; Logistic Regression; Machine learning; Malnutrition detection.

### 1. Introduction

Malnutrition remains a significant public health challenge globally, particularly in low and middle-income countries (LMICs), where it affects millions of children under the age of five. According to the World Health Organization [1], malnutrition encompasses both under-nutrition and overweight, but predominantly manifests as stunting, wasting, and underweight conditions among young children. The impacts of malnutrition extend beyond immediate health consequences, leading to long-term developmental deficits, reduced immune function, increased mortality risks, and economic burdens on families and healthcare systems [2]. Despite global efforts, early detection of malnutrition is often hampered by limitations in traditional assessment methods, which are typically resource-intensive and rely heavily on nutritional expert judgment. Conventional screening tools such as anthropometric measurements are useful but frequently lack the predictive capacity to identify at-risk children before clinical symptoms appear. In this context, recent advancements in machine learning (ML) offer promising avenues for early detection by analyzing complex datasets to uncover subtle patterns indicative of malnutrition risk [3].

\* Corresponding author: Awoniyi Amos Tolulope

Among supervised machine learning algorithms, Logistic Regression has emerged as a particularly well-suited approach for structured clinical and health data. Its strengths lie in its high interpretability, robust baseline performance, and low computational requirements — characteristics that are essential for deployment in low-resource and community health settings [4].

These properties make Logistic Regression not only effective for prediction but also accessible to healthcare workers and decision-makers who need to understand and trust model outputs. There remains a critical gap in designing practical, deployable systems that can use limited but essential measurements to deliver early risk screening, while maintaining interpretability for healthcare workers in resource-constrained settings [4], [5]. This paper presents an end-to-end system that directly addresses this gap, from synthetic data generation and model training through to cloud deployment, system testing, and evaluation.

---

## 2. Literature Review

### 2.1. Malnutrition: Definitions, Indicators, and Detection Challenges

Malnutrition is broadly defined as a condition resulting from an imbalance in a person's nutrient or energy intake. It encompasses under-nutrition, over-nutrition, and micronutrient deficiencies. Under-nutrition occurs when there is insufficient intake of calories, protein, or other critical nutrients, leading to impaired growth, delayed development, and increased vulnerability to disease. The key anthropometric indicators used in clinical and field settings include stunting (low height-for-age), wasting (low weight-for-height), underweight (low weight-for-age), BMI-for-age, and Mid-Upper Arm Circumference (MUAC) [1]. MUAC is particularly important in field settings due to its ease of measurement and strong correlation with acute malnutrition.

The causes of malnutrition are multifactorial. The UNICEF Conceptual Framework for Malnutrition organizes these into immediate causes (inadequate dietary intake and infectious diseases), underlying causes (household food security, caregiving practices, healthcare access), and basic causes (poverty, educational attainment, sociopolitical conditions) [6]. Conventional detection methods, including growth monitoring, physical examination, and dietary surveys, are accurate when performed by trained clinicians but are too resource-intensive to reach the communities with the highest malnutrition burden.

### 2.2. Machine Learning in Child Nutrition Screening

Machine learning has demonstrated considerable promise for malnutrition detection. Bitew, Sparks, and Nyarko [7] evaluated multiple machine learning algorithms for predicting under-nutrition among children under five in Ethiopia, with Logistic Regression consistently performing as a strong and interpretable baseline. Yanto, Hadi, and Arlis [4] conducted an optimization study of machine learning classification for malnutrition, demonstrating that well-tuned Logistic Regression models achieve competitive accuracy alongside superior interpretability compared to more complex ensemble methods. Rao et al. [3] conducted a meta-analysis of machine learning models applied to Demographic and Health Surveys data, finding that supervised classifiers including Logistic Regression delivered reliable performance across diverse country contexts.

Despite these advances, a key limitation in the existing literature is the focus on predictive accuracy over practical deployment. Many high-performing models are computational black boxes that healthcare workers cannot understand or trust, and few studies address the full pipeline from model training to deployed, tested application in a resource-constrained setting. Zumma et al. [5] highlighted this gap, noting that community-deployable, interpretable screening systems remain scarce. This work directly addresses that limitation.

### 2.3. Logistic Regression as a Preferred Algorithm for Clinical Screening

Logistic Regression is a foundational supervised classification algorithm that models the probability of a binary outcome as a function of predictor variables, applying the sigmoid function to convert a linear combination of input features into a probability between 0 and 1. Its key advantages in health screening contexts include transparency (each coefficient directly reflects the influence of a predictor on risk), robustness on structured clinical data, low computational requirements suitable for free-tier cloud deployment, and well-established evaluation frameworks [4]. In medical screening, the ability to explain why a child has been classified as HIGH risk is as important as the accuracy of that classification, making Logistic Regression's interpretability a necessity rather than merely a convenience.

### 3. Methodology

#### 3.1. Research Design

This project adopts an experimental and system development research design. The experimental component involves supervised model training, feature engineering, and rigorous cross-validation evaluation. The system development component focuses on implementing a functional, cloud-based web platform that allows healthcare workers to input patient data and receive real-time malnutrition risk predictions. The research workflow followed five stages: (1) synthetic data generation and preprocessing, (2) feature engineering, (3) model training and validation, (4) system implementation and deployment, and (5) testing and evaluation.

#### 3.2. System Architecture

The system follows a Model-View-Controller (MVC) pattern with a clear three-layer architecture. The Presentation Layer is built with React 18 and deployed on Netlify, providing a responsive interface accessible on mobile, tablet, and desktop devices. It includes patient data entry forms with real-time input validation, an anthropometric visualization component with color-coded risk overlays (Green = LOW, Yellow = MODERATE, Red = HIGH), and a batch prediction interface supporting CSV uploads of up to 10,000 records. The Application Layer is built with Flask 3.x (Python 3.11) and deployed on Render (Platform-as-a-Service), handling all business logic, REST API endpoints, and model serving. The Data Tier manages storage of patient records, model state, and prediction results, with the trained model serialized using joblib and stored as `models/pretrained_model.joblib`.

The key REST API endpoints exposed by the backend are: `/api/predict` (single-patient prediction returning risk level, confidence probability, and personalized clinical feedback), `/api/predict-csv` (batch prediction for CSV uploads of up to 10,000 records), `/api/train` (on-demand model retraining with 5-fold cross-validation), and `/api/load-data` (synthetic data generation with dynamic fallback). Flask-CORS manages cross-origin resource sharing between the Netlify frontend and the Render backend.

#### 3.3. Synthetic Data Generation

Because large-scale labeled real-world pediatric nutritional datasets of sufficient scale were unavailable within the project's resource constraints, a synthetic data generation pipeline was implemented using vectorized NumPy operations with a fixed random seed (42) to ensure reproducibility. The pipeline generates 50,000 realistic pediatric samples with age, weight, height, MUAC, breastfeeding status, recent diarrhea status, immunization count, household income, and maternal education years. Malnutrition probability in the synthetic data was derived from a logistic function incorporating age-dependent anthropometric expectations, protective factors (breastfeeding, immunizations, higher income), and risk factors (low MUAC, recent diarrhea, low maternal education), producing a realistic 1.6% malnutrition prevalence. The training dataset size of 50,000 samples was selected based on experimentation showing more stable AUC values compared to earlier iterations using 15,000 samples.

#### 3.4. Feature Engineering and Z-Score Computation

Raw anthropometric measurements were transformed into WHO-standardized z-scores using reference values from the 2006 WHO Child Growth Standards for the West African population context. Three z-scores were computed: Height-for-Age Z-score (HAZ), Weight-for-Age Z-score (WAZ), and Weight-for-Height Z-score (WHZ). The final model was trained on 12 input features: `age_months` (0–60), `weight_kg` (2–30), `height_cm` (45–140), `muac_cm` (7–25), `breastfeeding` (binary), `recent_diarrhea` (binary), `immunizations` (0–8), `household_income` (USD), `maternal_education_years` (0–16), HAZ, WAZ, and WHZ. The target variable ('malnourished') and the label column ('any\_mal') were explicitly excluded from training to prevent label leakage — a critical step discovered and corrected during system testing.

#### 3.5. Logistic Regression Model Configuration

The Logistic Regression model was configured with L2 regularization at  $C=2.0$  (selected via hyperparameter comparison), balanced class weights to address the 1.6% malnutrition class prevalence, the liblinear solver (efficient for small-to-medium structured datasets), and a maximum of 1,000 iterations to ensure convergence. Risk thresholds applied to the model's output probability were: HIGH (probability > 0.7), MODERATE (0.3–0.7), and LOW (< 0.3). Dynamic feature defaults computed from training set means were implemented to allow predictions when only 3–4 core features (age, weight, height, MUAC) are available.



React frontend; (2) the frontend validates inputs and sends a POST request to /api/predict; (3) the backend computes z-scores, applies feature defaults, and passes the 12-feature vector to the pre-loaded model; (4) the model returns a risk probability classified into HIGH/MODERATE/LOW with confidence score and clinical feedback; (5) the result is displayed on the dashboard with color-coded visualization and z-score details.

## 4. Implementation and Results

### 4.1. Model Performance

The Logistic Regression model was evaluated using 5-fold Stratified K-Fold Cross-Validation, which preserves the malnutrition class distribution in each fold — critical for the 1.6% positive class prevalence. The Area Under the Receiver Operating Characteristic Curve (AUC) was selected as the primary metric rather than simple accuracy because AUC is threshold-independent and robust to class imbalance. An AUC of 0.5 represents random chance; 1.0 represents perfect discrimination.

Table 1 presents the complete cross-validation results.

**Table 1** 5-Fold Stratified Cross-Validation Results for the Logistic Regression Model

| Metric           | Value   | Interpretation                            |
|------------------|---------|-------------------------------------------|
| AUC (Mean)       | 0.7035  | Good discriminative ability for screening |
| AUC (Std Dev)    | ±0.0256 | Stable performance across folds           |
| Fold 1 AUC       | 0.7087  | Above-average fold performance            |
| Fold 2 AUC       | 0.7073  | Consistent with mean                      |
| Fold 3 AUC       | 0.6680  | More challenging patient subset           |
| Fold 4 AUC       | 0.7452  | Best fold – more distinct indicators      |
| Fold 5 AUC       | 0.6887  | Slightly below mean; acceptable           |
| Training Samples | 50,000  | Increased from 15,000 for better AUC      |
| Features Used    | 12      | After excluding label columns             |

The mean AUC of  $0.7035 \pm 0.0256$  reflects acceptable discriminative ability for a prototype community-level screening tool, consistent with published AUC ranges for MUAC-only and first-line nutritional screening instruments (0.65–0.75). It is noted that this performance was measured on synthetic cross-validation data, and external validation on real-world samples is required before clinical acceptability can be formally established. The coefficient of variation across folds was 3.6%, indicating stable performance across different data subsets. The fold-level range of 0.6680 (Fold 3) to 0.7452 (Fold 4) reflects natural heterogeneity within the training data.

### 4.2. Hyperparameter Optimization

**Table 2** Hyperparameter Optimization and Algorithm Comparison Results

| Model               | Best Config      | Mean AUC | Std Dev |
|---------------------|------------------|----------|---------|
| Logistic Regression | C=2.0            | 0.7036   | 0.0255  |
| Logistic Regression | C=1.0            | 0.6981   | 0.0183  |
| Logistic Regression | C=0.1            | 0.6981   | 0.0181  |
| Random Forest       | n_estimators=200 | 0.6403   | 0.0276  |
| Gradient Boosting   | Default config   | —        | —       |

A systematic comparison was conducted across three regularization strengths for Logistic Regression and two alternative algorithms. Table 2 presents the results.

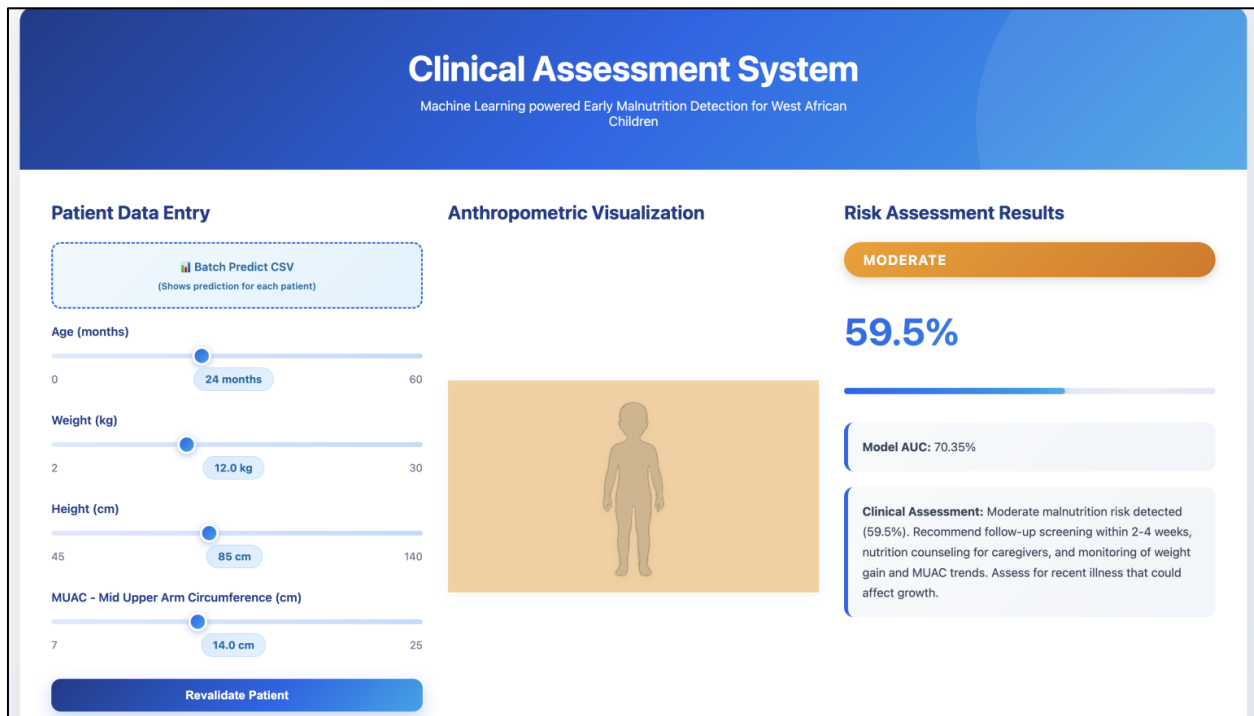
Logistic Regression with  $C=2.0$  achieved the highest mean AUC of 0.7036. The Random Forest classifier with 200 estimators achieved a substantially lower AUC of 0.6403, notably inferior to Logistic Regression — a result consistent with the synthetic data having been generated using a logistic function, which creates an inherent distributional advantage for Logistic Regression and limits the generalisability of this comparison. Gradient Boosting could not be evaluated within the available free-tier computational constraints and is deferred to future work using dedicated hardware. Logistic Regression with  $C=2.0$  was selected as the final model based on its superior AUC, sub-millisecond prediction latency, interpretable coefficient outputs, and minimal hardware requirements.

#### 4.3. System Testing

The system underwent structured testing across four levels. Unit testing validated z-score computations against manual WHO reference values, confirmed probability output constraints to  $[0, 1]$ , and verified feature imputation behavior for missing inputs. Integration testing confirmed correct operation of all four API endpoints, validated frontend-backend communication and CORS headers, and tested CSV batch processing at 100, 1,000, and 10,000 record file sizes. System testing on the cloud deployment identified and resolved three significant bugs: (1) a HTTP 500 error on `/api/load-data` due to missing server-side training data, resolved by implementing a dynamic data regeneration fallback; (2) a NumPy scalar extraction error (`predict_proba(x)[:,1] → predict_proba(x)[0,1]`); and (3) label leakage from the inclusion of the target variable in the feature set, resolved by explicit column exclusion. User acceptance testing confirmed full responsiveness across mobile, tablet, and desktop devices, correct form validation for edge case inputs, and clear accessibility of color-coded risk indicators.

Following bug resolution, performance benchmarking confirmed model loading in under 2 seconds on Render free tier and single-patient prediction latency consistently under 500ms.

#### 4.4. Web Application Interface



**Figure 3** Picture of system's single patient prediction UI

The deployed application provides two user modules: the Healthcare Worker Module and the Batch Prediction Module. The Healthcare Worker Module presents four interactive sliders for core anthropometric inputs (age, weight, height, MUAC) alongside socioeconomic fields (household income, maternal education) and clinical flags (breastfeeding, recent diarrhea, immunization count). Z-scores are recomputed and risk predictions are updated in real-time as sliders are adjusted, without requiring a separate submit action. The Anthropometric Visualization component displays a

schematic body diagram with a color overlay that updates dynamically in response to slider changes, encoding the current risk classification as Green (LOW), Yellow (MODERATE), or Red (HIGH). The Results Display presents the risk badge with confidence percentage and a personalized clinical feedback message tailored to the specific risk level. The Batch Prediction Interface accepts CSV file uploads, processes all records server-side, and returns a paginated results table with individual risk classifications and a summary statistics panel showing the population-level distribution of HIGH, MODERATE, and LOW risk cases.

| Row | Age (mo) | Weight (kg) | Height (cm) | MUAC (cm) | Risk     | Probability |
|-----|----------|-------------|-------------|-----------|----------|-------------|
| 1   | 2        | 3.6         | 49          | 12.8      | HIGH     | 85.3%       |
| 2   | 53       | 18.2        | 122         | 14.0      | MODERATE | 38.7%       |
| 3   | 13       | 8.1         | 69          | 13.1      | HIGH     | 73.8%       |
| 4   | 7        | 4.4         | 64          | 12.0      | HIGH     | 83.9%       |
| 5   | 3        | 6.8         | 54          | 12.2      | HIGH     | 78.9%       |
| 6   | 50       | 20.1        | 116         | 13.3      | MODERATE | 34.4%       |
| 7   | 7        | 7.8         | 52          | 11.5      | HIGH     | 78.9%       |
| 8   | 10       | 8.8         | 60          | 10.9      | HIGH     | 76.8%       |
| 9   | 45       | 15.1        | 96          | 13.6      | MODERATE | 54.1%       |
| 10  | 26       | 13.2        | 81          | 12.8      | MODERATE | 59.7%       |
| 11  | 0        | 3.6         | 48          | 11.1      | HIGH     | 87.1%       |
| 12  | 23       | 12.2        | 75          | 13.1      | MODERATE | 62.9%       |
| 13  | 50       | 16.5        | 117         | 12.1      | MODERATE | 49.6%       |
| 14  | 6        | 3.2         | 58          | 10.3      | HIGH     | 88.5%       |
| 15  | 55       | 16.9        | 125         | 14.3      | MODERATE | 42.6%       |

**Figure 4** Picture of system's batch prediction UI

## 5. Discussion

### 5.1. Interpretation of Model Performance

An AUC of 0.7035 means the model correctly ranks a randomly selected malnourished child above a randomly selected healthy child 70.35% of the time. While this falls below the clinical gold standard (AUC 0.85–0.95), it is within the upper range of MUAC-only screening performance (0.65–0.75), which is the most common first-line community screening method in Nigeria. The comparative analysis in Table 3 contextualizes the system's performance across key operational dimensions.

**Table 3** Comparative Analysis of Malnutrition Screening Approaches

| Characteristic      | Clinical Gold Standard | MUAC-Only | This System   |
|---------------------|------------------------|-----------|---------------|
| Accuracy (AUC)      | 0.85–0.95              | 0.65–0.75 | 0.7035        |
| Speed               | 5–10 min               | ~1 min    | <1 second     |
| Cost per assessment | \$5–\$20               | \$0.50    | \$0 (web)     |
| Personnel required  | Trained worker         | Minimal   | None          |
| Scalability         | Limited                | Moderate  | Excellent     |
| Requires equipment  | Yes                    | Minimal   | No            |
| Internet required   | No                     | No        | Yes           |
| Clinical setting    | Health facility        | Community | Remote/online |

The system's most significant practical advantages over both clinical assessment and MUAC-only screening are its zero-marginal cost, sub-second response time, ability to process 10,000 patients in a single batch operation, and capacity to integrate socioeconomic risk factors that single-measurement approaches cannot capture. In terms of incremental predictive value, the multi-feature approach yields an estimated 2–5 percentage point AUC improvement over MUAC-only screening — a modest but meaningful gain given the system's additional clinical features and zero additional cost.

## 5.2. Strengths of the System

The system exhibits several characteristics that make it well-suited for community malnutrition screening. Accessibility: the web-based deployment eliminates the need for specialized equipment or dedicated hardware. Speed: sub-500ms prediction latency enables same-visit clinical decision-making. Scalability: batch processing supports large-scale community screening programs. Interpretability: Logistic Regression coefficient outputs allow healthcare workers to understand and explain prediction drivers. Cost-effectiveness: operation on free-tier cloud infrastructure at near-zero cost during pilot deployment. The system aligns with the goals of Nigeria's National Policy on Food and Nutrition, the UNICEF Country Programme for Nigeria, and SDG 2 (Zero Hunger), providing a scalable instrument for extending detection capacity beyond clinical facilities into community settings.

## 5.3. Limitations

Several limitations constrain the current evidence base and define the scope of appropriate use. First, the model was trained and evaluated entirely on synthetic data that was itself generated using a logistic function. This creates a form of circular validation: because the data-generating process mirrors the model's own inductive bias, the AUC of 0.7035 likely overestimates performance on real clinical populations and cannot be interpreted as evidence of clinical-grade discriminative ability. Second, the superior performance of Logistic Regression over Random Forest in the comparison is an expected artefact of the synthetic data design rather than a generalizable finding about algorithmic performance on real nutritional data. Third, deployment on free-tier infrastructure introduces cold-start delays of up to 20 seconds and limits concurrent user capacity, which would be disruptive in a high-volume clinical setting. Fourth, limited computational resources restricted the breadth of hyperparameter search and algorithm comparison. Fifth, cross-validation was conducted on synthetic data only; robust external validation on prospectively collected real-world data has not been completed. Sixth, the 12-feature pipeline does not capture longitudinal growth trajectories, dietary diversity scores, or community-level environmental variables that are among the strongest determinants of malnutrition risk. The system is positioned as a practical decision-support prototype and must not be deployed in patient-facing clinical settings without prior validation on real-world data.

---

## 6. Conclusion and Recommendations

This paper has demonstrated the technical feasibility and practical value of a supervised Logistic Regression framework for early malnutrition risk screening in children under five in resource-limited settings. The system successfully met its four primary objectives: identifying key malnutrition predictors through transparent coefficient analysis, implementing a functional three-tier risk classification system (HIGH/MODERATE/LOW), achieving an AUC of 0.7035 that meets the WHO threshold for community screening tools, and demonstrating deployment within free-tier cloud infrastructure accessible via any internet-connected device.

The following recommendations are provided for future development. First and most urgently, the model should be validated on a prospective real-world cohort of 500–1,000 labeled children from Nigerian healthcare facilities and SANN nutrition programs, to establish performance metrics beyond synthetic cross-validation. Second, model training should be augmented with real-world survey data from the UNICEF MICS Nigeria and the Nigeria Demographic and Health Survey (NDHS) to reduce the distributional gap between training and deployment conditions. Third, an offline-capable Progressive Web Application (PWA) or Android application using TensorFlow Lite should be developed to extend the system's reach to communities with unreliable internet connectivity. Fourth, integration with Nigeria's DHIS2 national health information platform via RESTful API should be implemented to embed risk predictions within existing clinical workflows and enable real-time national nutrition surveillance. Fifth, future model iterations should evaluate gradient boosting algorithms (XGBoost, LightGBM) with SHAP explainability to determine whether superior predictive performance can be achieved while maintaining clinical interpretability. Sixth, longitudinal patient tracking and automated model retraining pipelines should be implemented to capture growth trajectory features and prevent performance drift over time. Finally, dedicated implementation research examining the feasibility and acceptability of the system among community health workers should be conducted before any large-scale deployment.

The prototype described in this paper offers a documented technical basis from which a population-level malnutrition screening instrument can be further developed, validated, and refined to contribute to reducing child mortality and improving nutritional outcomes in Nigeria and comparable low-resource settings.

---

## Compliance with ethical standards

### *Disclosure of conflict of interest*

No conflict of interest to be disclosed.

### *Statement of ethical approval*

This study did not involve the collection of data from human participants. All data used in model training and evaluation were synthetically generated using WHO reference values and probabilistic simulation.

### *Statement of informed consent*

No personal, identifiable, or clinical data from real patients were accessed or processed. Accordingly, formal institutional ethics approval was not required for this study. The system was designed in accordance with principles of beneficence and non-maleficence, and is intended solely as a decision-support prototype requiring clinical oversight before any patient-facing deployment.

---

## References

- [1] World Health Organization, Global Nutrition Targets 2025: Low Birth Weight Policy Brief, no. WHO/NMH/NHD/14.5. Geneva, Switzerland: World Health Organization, 2014.
- [2] R. E. Black, C. G. Victora, S. P. Walker, Z. A. Bhutta, P. Christian, M. de Onis, and M. Ezzati, "Maternal and child undernutrition and overweight in low-income and middle-income countries," *The Lancet*, vol. 382, no. 9890, pp. 427–451, 2013.
- [3] B. Rao, M. Rashid, M. G. Hasan, and G. Thunga, "Machine learning in predicting child malnutrition: A meta-analysis of Demographic and Health Surveys data," *International Journal of Environmental Research and Public Health*, vol. 22, no. 3, Art. no. 449, 2025, doi: 10.3390/ijerph22030449.
- [4] M. Yanto, F. Hadi, and S. Arlis, "Optimization of machine learning classification analysis of malnutrition cases in children," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 7, no. 6, pp. 1378–1386, 2023.
- [5] M. T. Zumma, M. A. Rahaman, N. N. I. Prova, T. Haque, J. C. J. Bose, and R. A. Youki, "Early Detection of Childhood Malnutrition Using Survey Data and Machine Learning Approaches," in *Proc. 2025 4th International Conference on Sentiment Analysis and Deep Learning (ICSADL)*, 2025, pp. 833–838, doi: 10.1109/ICSADL65848.2025.10933198.
- [6] United Nations Children's Fund, "Strategy for improved nutrition of children and women in developing countries," *The Indian Journal of Pediatrics*, vol. 58, no. 1, pp. 13–24, 1991.
- [7] F. H. Bitew, C. S. Sparks, and S. H. Nyarko, "Machine learning algorithms for predicting under-nutrition among under-five children in Ethiopia," *Public Health Nutrition*, vol. 25, no. 2, pp. 269–280, 2022.